

УНИВЕРЗИТЕТ У БЕОГРАДУ  
ФАКУЛТЕТ ОРГАНИЗАЦИОНИХ НАУКА

Никола Д. Војтек

**СИСТЕМ ЗА ПРЕДВИЂАЊЕ НЕДОЛАСКА  
ПУТНИКА НА ЛЕТ ЗАСНОВАН НА  
ТЕХНИКАМА РАЧУНАРСКЕ  
ИНТЕЛИГЕНЦИЈЕ**

докторска дисертација

Београд, 2020.

UNIVERSITY OF BELGRADE  
FACULTY OF ORGANIZATIONAL SCIENCES

Nikola D. Vojtek

**NO SHOW PASSENGERS PREDICTION  
SYSTEM BASED ON COMPUTATIONAL  
INTELLIGENCE TECHNIQUES**

Doctorial Dissertation

Belgrade, 2020.

Ментор:

---

др Братислав Петровић, редовни професор  
Факултет организационих наука, Универзитет у Београду

Чланови комисије:

---

Др Ивана Мијатовић, редовни професор  
Факултет организационих наука, Универзитет у Београду

---

Др Павле Милошевић, доцент  
Факултет организационих наука, Универзитет у Београду

---

Др Ивана Драговић, доцент  
Факултет организационих наука, Универзитет у Београду

---

Др Наташа Бојковић, ванредни професор  
Саобраћајни факултет, Универзитет у Београду

Датум одбране: \_\_\_\_\_

## *Изјаве захвалности - Позив на авантуру*

*Пресудан тренутак за започињање овог дела дисертације на овај начин представља моје упознавање са књигом Џозефа Кемпбела, Херој са хиљаду лица (енг. Joseph Campbell, The Hero With a Thousand Faces). Сматрам да је путовање кроз које сам прошао током докторских студија једно свеобухватно искуство и не би било могуће без људи који су били ту уз мене и пружили велику подршку. Могу слободно рећи да је ово путовање утицало на промену мог размишљања, свести, начина схватања одређених ствари, али и помогло да спознам сопствене могућности. Поред оствареног доприноса у научној заједници и стицања звања, такође сматрам да је искуство један неизмерно битан позитиван елемент који произилази из докторских студија.*

*Како Кемпбел истиче, херој је особа која је успела да достигне свој пун потенцијал, односно која је успела да се избори са својим личним и ограничењима из окружења. Аутор описује пут хероја као универзални шаблон једне авантуре и наглашава унутрашњу трансформацију особе која одлучи да прихвати позив ка авантури. Поглавља у Кемпбеловој књизи представљају кораке пута хероја, који се може поделити у више етапа, а које сам препознао у својој авантури.*

*Етапа 1: обичан живот. Пре него што сам почео и да размишљам о уписивању докторских студија, приводио сам крају обавезе на мастер студијама. Тада сам размишљао у потпуно другом правцу, тј. о даљим корацима у погледу посла и стицања искуства јер сам увелико радио.*

*Етапа 2: позив ка авантури. Први наговештај да почнем и да размишљам о докторским студијама је уследио након консултација у вези са завршним мастер радом са проф. др. Иваном Мијатовић на катедри Управљање квалитетом. То је уствари био први позив ка авантури према путу који Кемпбел наводи, где сам увидео да ако би прихватио тај позив, морао би да изађем из своје зоне комфора и утиснем се у нешто непознато. До тада, некако су основне и мастер студије биле нешто што је за мене било сасвим нормално да завршим, испит по испит, нешто на шта сам се навикао и током 5 година постало је део мог окружења и свакодневнице. Међутим, упуштање у истраживачке воде и*

академски приступ одређеним проблемима на том нивоу су мени донекле били нови и скоро непознати.

Етапа 3: одбијање позива. Након повратка са консултација, размишљао сам неколико дана о позиву за авантуру који ми је уручен. Схватао сам да ће то изискивати више новца, да нисам сигуран у којој области бих усмерио студије, а понајвише сам био несигуран по питању начина на који бих могао да остварим научни допринос у одабраној области. Узевши све у обзир, решио сам да не размишљам више о томе, па чак и почео да смишљам разлоге због којих би оправдао своје одустајање.

Етапа 4: сусрет са учитељем. Након неког времена, уследиле су нове консултације са проф. Др Иваном Мијатовић. Више смо тада причали о докторским студијама и начинима како започети, проћи кроз и завршити то путовање.

Етапа 5: прелазак првог прага и напуштање свакодневнице. После додатног размишљања одлучио сам да прихватим позив, односно да се пријавим на конкурс за докторске студије. Одлучио сам се за смер управљачки системи јер сам био свестан да морам да изађем из зоне комфора и након уписа, упознао новог ментора проф. др Братислава Петровића.

Етапа 6: учење правила новог света. Било ми је потребно доста времена како би разумео нове појмове у области коју сам одабрао за спровођење истраживања и која је донекле мени до тада била такоређи страна. Помоћ у учењу сам добио од Павла Милошевића који је тада био асистент на катедри за управљање системима, у виду различитих литература, књига, радова и осталих белешки а које су се односиле на испите.

Етапа 7: усвајање нових приступа и идеја. Током прве две године студија, било је неколико неуспелих покушаја да се дефинише тема истраживања, на чему би се и базирала докторска дисертација. Након што сам научио правила новог света, односно нове појмове у области управљачких система, трагао сам за проблемом из свог познатог окружења који бих могао да решим новим приступом који сам научио. Након неколико покушаја који нису дали очекиване резултате, био сам на корак одустајања. Међутим, у том тренутку сам добио прилику за запослењем у авио компанији и отвориле су ми се нове прилике из авио

индустрије које сам могао да искористим како бих формулисао проблем за истраживање.

Етапа 8: сусрет са великом препреком. Проблем у авио индустрији који се односи на предвиђање путника који се неће појавити на лету је представљао област погодну за примену техника рачунарске интелигенције. Спровео сам истраживање, применио одређене технике и знања, и добио задовољавајуће резултате. Такође, спровео сам анализу других предлога за решавање овог проблема који су били доступни кроз разну литературу. Уз помоћ професора и асистента, све наведено сам искористио и формулисао кроз рад који је било неопходно да објавим у неком од доступних међународних часописа из моје области. Овде сам се сусрео са до сада највећом препреком у току докторских студија. Часописа је било много, већина са превисоким критеријумима за прихватање рада, а било је потребно и дуго чекати на први одговор од уредника и рецензата.

Етапа 9: награда за достизање циља. Коментаре и сугестије које сам добио као повратни одговор из неколико часописа сам искористио за унапређење квалитета самог истраживања. После одређеног периода, као награда за претходно уложен труд и упорност, 07.04.2020. године сам добио одговор из часописа о прихватању рада, са чиме је главна препрека ка достизању циља премошћена.

Етапа 10: започињање повратка назад. Након добијања повратног позитивног одговора из часописа, крећем у реализацију преосталих активности, како би путовање привео крају успешно и вратио се назад у своје познато окружење. Сада имам толико нових идеја које бих желео да реализујем након завршетка докторских студија.

Етапа 11: суочавање са последњим тестом. Пре повратка назад, нова препрека, наизглед борба са временом али у ствари борба са самим собом, да ли ја ово могу да приведем крају? Да ли имам снаге да завршим писање тезе? да допуњујем, прерављам и да направим овај последњи корак?

Етапа 12: повратак кући са новим знањем. Након авантуре докторских студија, вратио сам се у своје првобитно наизглед познато окружење, богатији за нова знања и искуства које сам стекао. Међутим, за мене сада то окружење

*представља нови свет јер сам се током пута променио и схватам да више никад ништа неће бити исто.*

*Људима око себе, који су ме пратили у овој авантури, неизмерно сам захвалан.*

*Прво бих желео да се захвалим свом ментору, проф. др Братиславу Петровићу који ми је пружио шансу и имао разумевање док сам учио о новим појмовима у области техника рачунарске интелигенције. Велику захвалност дугујем и проф. др Ивани Мијатовић која ме је, поред многобројних савета и подршке током основних и мастер студија, подстакла да размишљам у правцу докторских студија. Током ове авантуре, остварио сам дугогодишњу сарадњу са доцентом др Павлом Милошевићем са Катедре за управљање системима, коме се овим путем захваљујем на труду и раду који је уложио и помоћи са истраживањем. Захвалио бих се и доценту др Ивани Драговић и проф. др Наташи Бојковић које су помогле да се истраживање сагледа из другог угла и унапреди.*

*Велико хвала дугујем и пријатељу и колеги Бојану Смуђи, на неисцрпној снази да заједно са мном учи и открива нове концепте и појмове у авио индустрији.*

*Не постоје речи које бих могао да употребим и захвалим се својој породици. Родитељима Весни и Душану, који су ме поред многих ствари научили да никад не одустајем и да увек будем позитиван. Брату Стефану који је увек истицао важност ове моје авантуре, и супрузи Милицы која је често знала да говори „не одустај, тако си близу“, а некад је то изговарала и на самом почетку. На крају, надам се да ће мојој деци ова авантура служити као позитиван пример и да ће их подстаћи да истражују, уче нове ствари и достигну свој пун потенцијал.*

# **Систем за предвиђање недоласка путника на лет заснован на техникама рачунарске интелигенције**

**Резиме:** Тема овог рада представља предлог система за предвиђање путника који се неће појавити на лету („no-show“), који се заснива на техникама рачунарске интелигенције. Предвиђање броја „no-show“ путника представља специфичан и уско формулисан проблем који је већ дужи низ година веома актуелан у авио индустрији како са теоријског, тако и са практичног аспекта. На основу очекиваног броја „no show“ путника, као и других фактора, авио компаније доносе одлуку о додатном броју места који ће бити доступан кроз резервациони систем. На овај начин, авио компаније могу остварити додатан профит, поготову када се ради о летовима који су попуњени у потпуности.

Предложени систем за предвиђање „no-show“ путника се састоји од две компоненте. Прва компонента се односи на избор најпрецизнијег модела за предвиђање, а друга на примену и валидацију система. Модел за предвиђање се састоји из алгорита који се заснива на техници закључивања на основу случаја и интерполативне Булове алгебре. Даље, модел комбинује предлог који је генерисан од стране алгорита и предлог који препоручује експерт. На овај начин предложени систем обједињује и узима у обзир објективну и субјективну димензију приликом предвиђања. Сличност између летова се израчунава коришћењем традиционалних мера (Еуклидска и Менхетн) и ИБА мере сличности. Такође, ИБА приступ употпуњује традиционални алгоритам технике закључивања на основу случаја кроз омогућавање логичке агрегације вредности, односно моделовањем постојећих нелинеарних зависности између података.

Примена предложеног система је представљена коришћењем података о лету на релацији Београд - Амстердам, за период од годину дана. Добијени резултати показују да је неопходно укључити препоруку експерта у процес предвиђања, као и да сам алгоритам није довољан да би се добили довољно прецизни резултати. Такође, добијени резултати указују да су модели који су засновани на ИБА приступу и који комбинују резултате алгорита и препоруку експерта, прецизнији од модела који користе традиционалне мере за израчунавање сличности. Сходно томе, потврђено је да логички приступ моделовању сличности представља перспективан правац примене у оквиру



технике закључивања на основу случаја. Са практичне стране, предложено решење је једноставно за разумевање у погледу функционисања, и може се доста једноставно имплементирати и прилагодити специфичностима и операцијама авио компаније.

**Кључне речи:** закључивање на основу случаја, интерполативна Булова алгебра, предвиђање „no-show“ путника, авио индустрија

**Научна област:** Техничке науке

**Ужа научна област:** Управљање системима

УДК број:

# **No-show passengers prediction system based on computational intelligence techniques**

**Abstract:** In this doctoral dissertation no-show passengers prediction system based on computational intelligence techniques is proposed. Predicting no-show passengers represents a specific and concisely formulated problem that actively persists for a longer period of time in the airline industry from both theoretical and practical perspective. Based on the expected number of no-show passengers, as well as some other factors, airlines are making decisions about how many additional seats will be allowed for overbooking through reservation system. This way, airlines could make additional profit, especially when it comes to the high demanding flights that are fully booked.

Proposed prediction system for no-show passengers consists of two major components. First component considers selecting the best performing prediction model from the available model pool, and the second component is related to the model validation and application. Prediction model is based on the algorithm that combines case based reasoning technique and interpolative Boolean algebra (IBA) approach. Furthermore, model combines prediction recommendation generated by algorithm and recommendation provided from the domain expert. This way, the proposed system considers and takes into account both objective and subjective dimension. Similarity between flights is determined using traditional metrics (Euclidean and Manhattan) and IBA similarity measure. Also, IBA approach is enhancing the conventional CBR algorithm by enabling logical aggregation of values, i.e. capturing existing nonlinear dependencies in the data.

The usage of the proposed system is illustrated in the numerical example regarding a single leg flight on the Belgrade-Amsterdam route and covers a one-year period. The obtained results show the necessity to include expert knowledge in prediction process, i.e. the CBR algorithm used alone is insufficient to produce results that are accurate enough. Furthermore, the results are indicating that the IBA-based models that combine the results of the CBR algorithm and expert recommendations perform better than distance-based models. Therefore, it is confirmed that the logic-based approach of similarity modelling is the prospective direction within the CBR algorithm. From a practical side, proposed solution is easy for understanding from the

functional aspect, and could be easily implemented and adjusted according to airline operations.

**Keywords:** case based reasoning, interpolative Boolean algebra, no-show passengers forecasting, airline industry.

**Scientific field:** Technical sciences

**Specific scientific field:** System control

UDK number:

# Садржај

1. УВОД.....	1
1.1. Проблем, предмет и циљ истраживања.....	1
1.2. Полазне хипотезе.....	4
1.3. Структура рада.....	5
2. ОСНОВЕ РАЧУНАРСКЕ ИНТЕЛИГЕНЦИЈЕ.....	8
2.1. Основне карактеристике техника рачунарске интелигенције.....	9
2.2. Фази логика.....	10
2.3. Неуронске мреже.....	12
3. ПРЕГЛЕД ПОСТОЈЕЋИХ РЕШЕЊА ЗА УПРАВЉАЊЕ РЕСУРСИМА У АВИО ИНДУСТРИЈИ.....	17
3.1. Решења заснована на статистичким техникама.....	19
3.1.1. Мерење перформанси авио компаније.....	19
3.1.2. Оптимизација инвентара (седишта).....	20
3.1.3. Предвиђање броја путника.....	21
3.1.4. Одређивање „overbooking“ лимита.....	22
3.2. Решења заснована на техникама рачунарске интелигенције.....	23
3.2.1. Мерење перформанси авио компаније.....	24
3.2.2. Оптимизација инвентара (седишта).....	26
3.2.3. Предвиђање броја путника.....	26
3.2.4. Одређивање „overbooking“ лимита.....	27
3.2.5. Предвиђање броја „no-show“ путника.....	28
4. ИНТЕРПОЛАТИВНА БУЛОВА АЛГЕБРА.....	30
4.1. Основни појмови.....	32
4.1.1. Пример 1 – поређење попуњености авиона.....	33
4.1.2. Пример 2 – поређење резервација путника.....	35
4.2. ИБА структура.....	35
4.3. Симболички и вредносни ниво.....	39
4.3.1. Симболички ниво.....	39
4.3.2. Вредносни ниво.....	41
4.4. Правци примене.....	46

4.4.1. Логичка агрегација .....	46
4.4.2. ИБА мера сличности/различитости .....	47
5. ЗАКЉУЧИВАЊЕ НА ОСНОВУ СЛУЧАЈА.....	51
5.1. Структура и основни елементи .....	52
5.2. Процес примене .....	53
5.3. Поузданост решења.....	54
6. ПРИМЕНА ТЕХНИКЕ ЗАКЉУЧИВАЊА НА ОСНОВУ СЛУЧАЈА: СЛУЧАЈ НИСКОБУЏЕТНЕ АВИО КОМПАНИЈЕ.....	56
6.1. Опис проблема.....	57
6.2. Формулација модела на бази технике закључивања на основу случаја.....	59
6.3. Резултати и разматрање коришћеног модела .....	61
7. СИСТЕМ ЗА ПРЕДВИЂАЊЕ „NO-SHOW“ ПУТНИКА У АВИО ИНДУСТРИЈИ .....	63
7.1. Опис проблема.....	64
7.2. Економски аспект проблема.....	65
7.3. Предлог система за предвиђање „no-show“ путника .....	68
7.3.1. Компонента 1 - Избор модела .....	69
7.3.2. Компонента 2 - Примена система .....	70
7.3.3. Компонента 3 - Ажурирање базе .....	70
7.3.4. Процес избора модела за предвиђање „no-show“ путника.....	71
8. ПРИМЕНА ПРЕДЛОЖЕНОГ СИСТЕМА.....	79
8.1. Анализа иницијалног скупа података.....	79
8.1.1. Графички приказ података – појединачни атрибути .....	81
8.1.2. Графички приказ података – повезаност атрибута са зависном варијаблом .....	87
8.2. Избор модела - претпроцесирање података.....	92
8.3. Избор модела - алгоритам за одређивање сличности .....	96
8.4. Тест значајности одабраних модела .....	107
8.4.1. Фридманов тест .....	107
8.4.2. Вилкоксонев тест.....	108
8.5. Примена система и валидација .....	110
8.6. Поређење са моделом базираним на неуронским мрежама .....	111
9. ИМПЛЕМЕНТАЦИЈА РЕШЕЊА У ДОМЕНУ АВИО КОМПАНИЈЕ.....	113

9.1. Основне карактеристике предложеног решења .....	113
9.2. Домен/сектор примене .....	117
9.3. Смернице за процес имплементације .....	119
10. ЗАКЉУЧАК .....	121
10.1. Осврт на постављене хипотезе и остварене доприносе.....	124
10.2. Могућности и правци даљег истраживања .....	127
11. ЛИТЕРАТУРА .....	129
12. ПРИЛОЗИ.....	143
12.1. Списак слика .....	143
12.2. Списак Табела .....	145
12.3. Списак скраћеница .....	145
Биографија аутора .....	147
ИЗЈАВА О АУТОРСТВУ .....	150
ИЗЈАВА О ИСТОВЕТНОСТИ ШТАМПАНЕ И ЕЛЕКТРОНСКЕ ВЕРЗИЈЕ ДОКТОРСКОГ РАДА.....	151
ИЗЈАВА О КОРИШЋЕЊУ .....	152

*„If you want it, you can fly, you just have to trust you a lot.“*

Steven Paul Jobs

## 1. УВОД

### 1.1. Проблем, предмет и циљ истраживања

Предвиђање броја путника који се неће појавити на лету („no-show“) представља специфичан и уско формулисан проблем који је већ дужи низ година веома актуелан у авио индустрији како са теоријског, тако и са практичног аспекта. На основу очекиваног броја „no-show“ путника, као и других фактора, авио компаније доносе одлуку о додатном броју места који ће бити доступан кроз систем за резервацију. На овај начин, авио компаније остварују додатан профит, поготову када се ради о летовима који су попуњени у потпуности.

Гледано са теоријског, односно научног аспекта, овај проблем је препознат као изузетно битан јер се доводи у везу са „overbooking“ лимитом постојећег инвентара, а који, како је истакнуто у (Somboon & Amaruchkul, 2017), сам по себи доприноси повећању профита. Авио компаније користе различите стратегије и приступе којима одређују овај ниво, односно број седишта које је могуће резервисати преко максималног капацитета авиона. Такође, може се рећи и да се авио компаније на неки начин штите од путника који потенцијално могу одустати од саме услуге у последњем тренутку (на многим примерима се може видети да је ово доста честа ситуација). За разлику од других производа и услуга које потенцијални купци могу вратити након куповине, односно отказати резервацију, у авио индустрији је ситуација специфична. Када авион полети, седишта на том лету се не могу поново продати. Сходно томе, када авио компаније не би могле да се заштите од губитка прихода који настаје кроз примену „overbooking“ приступа, биле би у неку руку присиљене да реализују лет са празним седиштима која су могла бити попуњена. Један од ефикаснијих начина за одређивање „overbooking“ лимита, представља предвиђање броја путника који се неће појавити на лету (Fildes, Nikolopoulos, Crone & Syntetos, 2008; Kunnumkal, Talluri & Topaloglu, 2012). Узевши ово у обзир, евидентно је да постоји значајнији ниво интересовања

са научног аспекта, за проналажење решења и дефинисање система који ће дати што боље резулате предвиђања.

У литератури је истакнуто да је за предвиђање броја путника који се неће појавити на лету неопходно анализирати и обрадити велики број података. Често су базе података којима располажу авио компаније сачињене од непотпуних, непрецизних, као и од података који су међусобно нелинеарно повезани. Како је истакнуто од стране бројних аутора (Min & Joo, 2016; Pineda, Liou, Hsu & Chuang, 2018) развој техника рачунарске интелигенције је у значајној мери олакшао и убрзао авио компанијама обраду података. Може се рећи да су технике рачунарске интелигенције у овом конкретном проблему у предности у односу на класичне технике, јер пружају могућност моделовања нелинеарних веза и толеришу непотпуне и непрецизне податке (Lim, Balas & Do, 2013).

Предвиђање „no-show“ путника се базира на искуствима из прошлости, што отвара простор за примену технике закључивања на основу случаја. Као што је дато у (Leake, 1996) ова техника користи базе знања и симулира људско закључивање у ситуацијама у којима је могуће користити искуство из прошлости. На тај начин се нови проблем (случај) може решити проналажењем сличних случаја из прошлости који се налазе у бази знања и применом њихових решења. Процес је објашњен у (Aamodt & Plaza, 1994) и састоји се из 4 корака који се у литератури називају „4R“ (енг. Retrieve, Reuse, Revise, Retain) (Aamodt & Plaza, 1994). Приликом примене ове технике веома је битно одабрати одговарајућу меру сличности, као и оператор који ће се користити за агрегацију добијених вредности. Управо се овим отвара простор за увођење интерполативне Булове алгебре.

Интерполативна Булова алгебра (ИБА) је предложена од стране Драгана Радојевића (Radojević, 2000; 2008a) и представља конзистентну  $[0,1]$  -реализацију Булове алгебре јер чува све законе на којима почива Булова алгебра. Састоји се из симболичког и вредносног нивоа који су детаљно описани у (Radojević, 2008a). На симболичком нивоу, атрибути се посматрају независно од њихових вредности, и пратећи принцип структурне функционалности, рачуна се структура сложеног логичког израза (Radojević, 2008d). На вредносном нивоу, атрибутима се додељују вредности, бира се одговарајући оператор у зависности од природе атрибута и/или



њихове статистичке међузависности, и рачуна сама вредност израза остајући у Буловом оквиру. Два тренутно најуспешнија правца примене ИБА се односе на логичку агрегацију и меру сличности/различитости (Radojević, 1999; 2008c; Poledica, Milošević, Dragović, Petrović & Radojević, 2015; Milošević et al., 2018). Ово су управо две области у оквиру технике закључивања на основу случаја које изискују одређена унапређења услед специфичности проблема предвиђања „no-show“ путника.

Како би предвиђање „no-show“ путника било делотворно и са практичног аспекта саме авио компаније, потребно је фокусирати се на летове који су потпуно или готово потпуно попуњени. На овај начин се повећава могућност остваривања додатног профита. У оваквим ситуацијама често може помоћи знање експерта у области/домену као што је комерцијала у оквиру које подпада управљање седиштима, класама и капацитетима авиона. Додатну вредност експерт може донети јер нпр. има искуства са конкретним летом и зна колико се путника није појавило у прошлости, или је нпр. упознат са специфичним екстерним факторима који су се изненада појавили (терористички напади, елементарне непогоде, штрајкови на аеродрому...). Узевши у обзир целокупну ситуацију и све додатне факторе који услед изненадног настанка нису моделовани и укључени у алгоритам, експерт може дати своју процену очекиваног броја „no-show“ путника. Тачније, експертска процена може допринети процени добијеној од стране алгоритма, те тако комбинован систем може понудити прецизније крајње решење.

**Проблем** истраживања докторског рада је могућност израде система за предвиђање „no-show“ путника у авио индустрији који би комбиновао резултате добијене алгоритмом и процену од стране експерта. Алгоритам би се базирао на техници закључивања на основу случаја и користио интерполативну Булову алгебру (Radojević, 2008a) за логичку агрегацију вредности променљивих и за израчунавање сличности између атрибута (Radojević, 1999; 2008c).

**Предмет** докторског рада је примена интерполативне Булове алгебре у комбинацији са техником закључивања на основу случаја како би се формулисао систем за предвиђање „no-show“ путника у авио индустрији. У оквиру рада биће разматрано више комбинација по питању избора логичког оператора и коришћења

неколико различитих мера сличности. Такође, биће тестирана решења са и без укључивања експертског знања.

**Циљ** докторског рада је да се предложи нови систем за предвиђање „no-show“ путника у авио индустрији заснован на комбинацији техника рачунарске интелигенције (закључивање на основу случаја и интерполативна Булова алгебра). Предложени систем ће бити унапређен узимањем у обзир експертског знања у оквиру процеса предвиђања које ће бити агрегирано са предвиђањем генерисаним од стране алгоритма. Такође, циљ је и да се предложени систем тестира на практичном примеру авио компаније, како би се донео утврдила његова ефикасност, односно ваљаност.

Посматрано са практичног аспекта, предложени систем за предвиђање „no-show“ путника може у значајној мери помоћи и унапредити одређене процесе у оквиру авио компаније. Своју примену може наћи у секторима који за циљ има предвиђање броја путника који се неће појавити на лету и одређивање „overbooking“ лимита. Најчешће су ти сектори или домени у авио компанијама комерцијала и комерцијални послови, управљање приходима, управљање инвентаром и управљање резервационим системом. Информације о томе које системе и технике користе авио компаније у ове сврхе нису јавно доступне или је до њих веома тешко доћи, али се са сигурношћу може рећи да систем базиран на техникама рачунарске интелигенције може донети додатну вредност предвиђањем „no-show“ путника.

## 1.2. Полазне хипотезе

У докторском раду полази се од следећих хипотеза:

Основна хипотеза:

- Предвиђања „no-show“ путника у авио индустрији се може унапредити коришћењем техника рачунарске интелигенције у комбинацији са експертским знањем.

Помоћне хипотезе:

- Закључивање на основу случаја се може унапредити увођењем интерполативне Булове алгебре;
- Логичка агрегација и мера сличности/различитости засноване на ИБА се могу користити у предвиђање броја путника који се неће појавити на лету;
- Могуће је развити систем за предвиђање „no-show“ путника који комбинује технике рачунарске интелигенције и експертско знање;
- Укључивање експертског знања може да побољша резултате предвиђања предложеног система.

### 1.3. Структура рада

Докторска дисертација се састоји од 10 поглавља и дела рада у коме је наведена коришћена литература.

У првом поглављу рада описани су проблем, предмет и циљеви истраживања овог рада. Такође, дефинисане су основна и помоћне хипотезе које се односе на истраживање. Затим је дат опис структуре рада по поглављима.

У другом поглављу су објашњени основни концепти рачунарске интелигенције, а фази логика и неуронске мреже су посебно објашњене, са акцентом на могућности примене.

У трећем поглављу је дат преглед постојећих решења за предвиђање броја „no-show“ путника у авио индустрији, као и за оптимизацију капацитета авиона, максимално искоришћење инвентара, предвиђање броја путника и за предвиђање „overbooking“ лимита.

У четвртом поглављу су наведени основни теоријски појмови и дефиниције које се односе на интерполативну Булову алгебру (ИБА). Објашњени су основни појмови коришћењем два примера резервације седишта и посматрања попуњености лета у авио индустрији, дата је ИБА структура, представљени су симболички и вредносни ниво, као и два правца примене – логичка агрегација и мера сличности/различитости.

У петом поглављу су дати основни теоријски појмови и дефиниције које се односе на технику закључивања на основу случаја. Прво је дефинисана структура

саме технике, заједно са основним елементима. Након тога, објашњен је процес примене технике кроз четири корака. У оквиру последњег потпоглавља, наведени су фактори који могу утицати на поузданост добијеног решења.

У оквиру шестог поглавља је представљена примена технике закључивања на основу случаја на примеру ниско буџетне авио компаније и на једноставнијем проблему предвиђања укупног броја путника на лету.

У седмом поглављу је описан проблем непојављивања путника на лету и указано је на економски аспект кроз три различита случаја. У последњем потпоглављу, дат је предлог решења које се односи на предвиђање броја „no-show“ путника.

У осмом поглављу је представљена примена предложеног система на реалном примеру у оквиру авио индустрије. Избор модела, тестирање и валидација су спроведени коришћењем података о лету на релацији Београд – Амстердам. На крају, у последњем потпоглављу, претходно добијени резултати су упоређени са резултатима предвиђања неуронским мрежама.

У деветом поглављу је дат практичан осврт на применљивост решења у оквиру авио компаније. Прво су истакнуте су основне карактеристике предложеног решења. Затим је идентификован домен, односно сектор примене и на крају су дате смернице за сам процес имплементације.

У последњем поглављу дат је закључак, са посебним освртом на постављене хипотезе, остварене доприносе у оквиру дисертације и могуће правце даљег истраживања.

На крају рада наведена је литература која је коришћена приликом израде дисертације.

Од основних метода научног истраживања у раду су коришћене следеће:

- Анализа прикупљене литературе са циљем генерисања прегледа постојећих решења која се користе за предвиђање и управљање ресурсима у авио индустрији;
- Критичка анализа технике закључивања на основу случаја, као и интерполативне Булове алгебре са посебном пажњом на два правца примене - логичка агрегација и мера сличности/различитости;

- Развој новог система за предвиђање „no-show“ путника комбиновањем технике закључивања на основу случаја и интерполативне Булове алгебре, као и укључивањем експертске димензије;
- Експериментална провера предложеног система и спровођење тестирања над расположивим подацима из праксе. Подаци који су коришћени за експерименталну проверу нису јавно доступни;

„Onaj ko voli praksu bez teorije je poput moreplovca koji se ukrca na brod bez kormila i kompasa, te nikad ne zna na koju obalu će pristati.“

Леонардо Да Винчи

## 2. ОСНОВЕ РАЧУНАРСКЕ ИНТЕЛИГЕНЦИЈЕ

У оквиру другог поглавља су објашњени основни концепти рачунарске интелигенције, што у ствари представља сам увод у наставак и даље разматрање примене техника рачунарске интелигенције у поглављу 3. Након идентификовања области и датих дефиниција, у потпоглављу 2.1. су дате основне карактеристике техника рачунарске интелигенције. Потпоглавље 2.2. дефинише основне појмове и елементе фази логике а потпоглавље 2.3. се бави објашњавањем концепта неуронских мрежа са акцентом на могућности примене.

Рачунарска интелигенција представља област рачунарских наука која се фокусира на изучавање карактеристика који су блиски начину на који људи размишљају и доносе закључке. Као што је истакнуто од стране Назмула и Хојата (Nazmul & Hojjat, 2013) рачунарска интелигенција се састоји од скупа рачунарских методологија и приступа који се користе у решавању сложених реалних проблема код којих употреба традиционалних математичких модела не даје очекиване резултате. Као главне разлоге изостанка очекиваних резултата, поменути аутори у књизи наводе да процеси могу бити превише сложени, да постоје ситуације у оквиру којих проблем и процеси излазе из основног математичког оквира, као и да сам процес може бити стохастичке природе, односно да елементе и понашање у оквиру процеса није могуће са сигурношћу предвидети.

Сам назив рачунарска интелигенција (енг. *computational intelligence*) је први пут употребљен 1990. године од стране међународног Института инжењера електротехнике и електронике (енг. *IEEE - Institute of Electrical and Electronics Engineers*), тачније од Савета за Неуронске Мреже (енг. *Neural Networks Council*) као члана поменуте организације. Почев од тог периода па све до данас, системи базирани на принципима рачунарске интелигенције су нашли примену у скоро свим индустријама (Chou, Liu, Huang, Yih & Han, 2011; Shaheen, Rashed & Cheng,

2010; Cavalcante, Brasileiro, Souza, Nobrega & Oliveira, 2016). Када је у питању обрада и процесирање велике количине података, рачунарска интелигенција помаже у изазовима и проблемима који се често јављају у овим ситуацијама. Као што је истакнуто у једном од скоријих истраживања (Reddy & Mehta, 2019), ови проблеми се често односе на неконзистентност у подацима, непостојање одређених података, као и на недостатак времена и прецизно формулисаног модела за обраду великих количина података. Као што се може видети и према једном од такође релативно новијем истраживању (Zhou, Wang, Guo & Zhang, 2019), и даље постоји велико интересовање за развојем система који користе концепте и методе рачунарске интелигенције и комбинују их са другим актуелним технолошким концептима.

Како истиче Енглберт (*Andries P. Engelbrecht*) у књизи „*Computational Intelligence An Introduction*“ (2007, 3-5), рачунарска интелигенција је област која се бави изучавањем механизма који користе интелигентно понашање како би се прилагодили ситуацији и окружењу. Како се даље у књизи појашњава, ти механизми укључују парадигме вештачке интелигенције (енг. *Artificial Inteligence*) које показују способност да уче и прилагођавају се новим ситуацијама, да уопштавају, сажимају, откривају нове елементе и факторе, и да повезују информације.

## 2.1. Основне карактеристике техника рачунарске интелигенције

За разлику од конвенционалних, технике рачунарске интелигенције имају циљ да нађу/предложе приближно решење, како за проблеме који су јасно дефинисани, тако и за оне који су неодређени. Због тога, неретко се може наћи у литератури да се технике рачунарске интелигенције, као што су неуронске мреже или фази логика, називају и меко рачунарство (енг. *soft computing*). Као што је истакнуто од стране Задеа (Zadeh, 1994), конвенционалне методе не толеришу непрецизност и делимичну истинитост, док се основни принцип метода рачунарске интелигенције односи на коришћење способности за толерисање непрецизности и могућност моделовања делимичне истинитости, како би се дошло до решења проблема које је робустно, знатно лакше за имплементацију и

изискује мање трошкова. Узимајући све у обзир, може се рећи да технике рачунарске интелигенције нуде много могућности за моделовање, дефинисање и унапређење процеса, као што је и истакнуто у (Saridakis & Dentsoras, 2008). Пример интеграције техника рачунарске интелигенције са другим техникама је дат у оквиру (Zeleznikow & Nolan, 2001).

Како би један систем био у могућности да обради информације на исти или сличан начин као људски ум, мора поседовати следеће три карактеристике: отвореност, робустност (употребљивост) и могућност обраде података у реалном времену (Zadeh, 1994). Отвореност представља способност прилагођавања новонасталој ситуацији и променама у окружењу. Робустност се односи на стабилност и на претходно напоменуто способност обраде непотпуних и непрецизних података. Последња карактеристика указује на потребу реаговања на одређени проблем у оквиру разумног временског оквира (у зависности од ситуације).

Три најчешће коришћене технике, односно концепта у оквиру рачунарске интелигенције су фази логика, неуронске мреже и алгоритми еволутивног рачунарства. С обзиром да је у овом раду фокус на примени елемената фази логике и неуронских мрежа, у одређеним аспектима и у оквиру одређених корака истраживања, алгоритми еволутивног рачунарства неће бити додатно објашњени. Као што је истакнуто у (Walczak, 2001), резултати који произилазе из употребе неуронских мрежа за решавање одређених проблема се базирају на историјским подацима о конкретном процесу. С друге стране, фази логика главну примену проналази кроз могућност обраде описних променљивих употребом фази бројева (Jamshidi, 2003) и фази скупова (Lalla, Facchinetti & Mastroleo, 2008). Детаљније објашњење ових техника је дато у наредна два потпоглавља. Такође, у циљу израде свеобухватног и комплексног система, често долази до комбинације ове две технике (нпр. (Kar, Das & Ghosh, 2014; Chang, Kuo, Wu & Tzeng, 2015).

## 2.2. Фази логика

Већина проблема у економским и организационим наукама често нема јасно дефинисану структуру, и може се рећи да то представља један од главних



разлога и мотиватора за константно развијање модела заснованих на фази логици (Zadeh, 1965), који се користе како би се премостили поменути проблеми (Skala, 1978). Као што је појашњено од стране Енглберта у поменутој књизи „*Computational Intelligence An Introduction*“ (2007, 10-11), традиционална теорија скупова захтева да елемент или буде елемент скупа или не, а слично је и са бинарном логиком, где вредности одређеног атрибута могу бити или 0 или 1. Међутим, људски мозак у већини ситуација није искључив, односно не може једнозначно разграничити да ли одређени елемент припада или не припада одређеном скупу. Другим речима, људско посматрање и закључивање углавном укључује и одређену дозу несигурности.

Фази скупови и фази логика дозвољавају управо закључивање са одређеном дозом апроксимације (Ramot, Friedman, Langholz & Kandel, 2003). На основу фази логике и теорије фази скупова, дефинисани су фази системи који се користе за доношење одлуке. Фази логика представља генерализацију класичне Булове логике (Zadeh, 1972), а основни допринос се огледа кроз могућност обраде лингвистичких променљивих (Sugeno & Tanaka, 1991).

У оквиру фази скупа елемент припада скупу у одређеном степену, а сама фази логика омогућава закључивање на основу припадности елемента одређеном скупу. У поређењу са класичним скупом, фази скуп представља његову генерализацију, а степен припадности скупа одређеног елемента фази скупу се најчешће дефинише као вредност између интервала  $[0,1]$ , а сама вредност се испољава у датом интервалу кроз функцију припадности (енг. *membership function*)  $\mu_A(x) = A(x)$ . Предмет интересовања фази логике су сви елементи интервала  $[0,1]$ , односно и крајњи и међуслучајеви. Према томе, елемент  $x$  у потпуности припада фази скупу  $A$ , када је  $\mu_A(x) = 1$ , док за супротни крајњи случај  $\mu_A(x) = 0$ ,  $x$  у потпуности не припада фази скупу  $A$ .

Кроз коришћење одговарајућих метода, фази логика може бити веома делотворан алат као саставни део процеса за доношење одлука у ситуацијама где није могуће јасно утврдити све факторе (Grabisich, 1996; Narukawa & Torra, 2007; Zhu & Xu, 2014). Веома битна област примене фази логике се односи на одређивање ефикасности и способности процеса да обавља задате функције (Lozano, 2014; Kaaya & Kahraman, 2011; Wu & Liao, 2014). Иако је сам концепт

фази логике дефинисан знатно раније, управо због константне потврде резултата који су били постигнути употребном фази логике, мноштво истраживања је спроведено са циљем унапређења концепта. Један од таквих примера се може наћи у раду (Yazdanbakhsh & Dick, 2018).

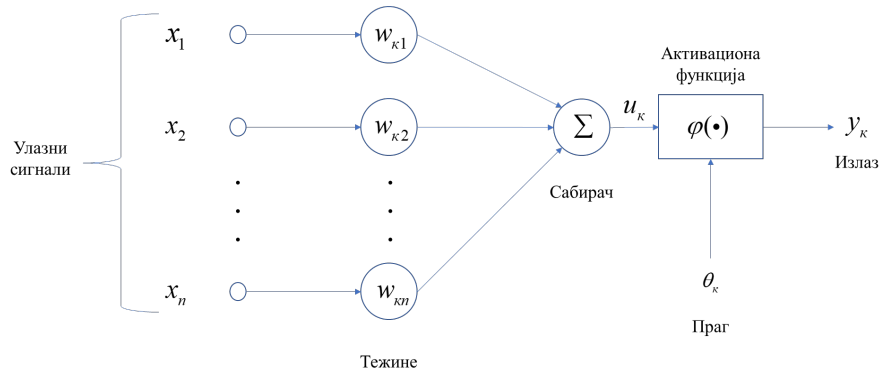
Такође, веома је битно истаћи у овом делу да се у оквиру фази логике не поштује закон искључења трећег и контрадикције. Међутим, као што је и истакнуто у (Hájek, 1998), то се није показало као велико ограничење када је у питању примена, а многи сматрају и као корисну техничку претпоставку. Посматрано са теоријског становишта, овај проблем је врло значајан и предмет је многих истраживања, а једно од решења које се односи на ИБА концепт је представљено и у (Radojević, 2000).

### 2.3. Неуронске мреже

Основна мотивација за истраживање у правцу моделовања биолошких неуронских система кроз алгоритамски приступ се огледа кроз покушај моделовања људског мозга (Engelbrecht, 2007). Узевши у обзир комплексност и нелинеарност, људски мозак је способан да обавља више паралелних активности, препознаје шаблоне, да опажа, учи, памти и уопштава информације. Процењује се да постоји између 10-500 билиона неурона и 60 трилиона спојева између њих у људском мозгу. Неурони су груписани у приближно 1000 главних модула, а сваки од модула има око 500 неуронских мрежа. Узевши ово у обзир, веома актуелно питање, које поставља и Енглберт у поменутој књизи, је да ли је могуће стварно моделовати целокупан људски мозак? Свакако је развој неуронских мрежа један значајан корак ка томе.

Неуронске мреже представљају технику рачунарске интелигенције која се базира на сложеном алгоритму (Hagan, Demuth, Beale & DeJesus, 2014), а чија је примена изразито популарна и распрострањена у свету, поготову када се ради о решавању сложених проблема (Cichy & Kaiser, 2019). По питању сложености алгоритма, постоје различите врсте неуронских мрежа, а узимајући у обзир повезаност неурона, броја слојева, као и могућности учења помоћу поратне спреге. У ситуацији када постоје више слојева, неурони у оквиру једног слоја

могу бити повезани потпуно или делимично са неуронима наредног слоја. Учење са повратном спрегом се односи на постојање веза у оквиру неуронске мреже, које омогућавају слање повратне информације и које се могу користити за учење. Општи модел неурона је дат на следећој Слици 1.



**Слика 1.** Општи модел неурона

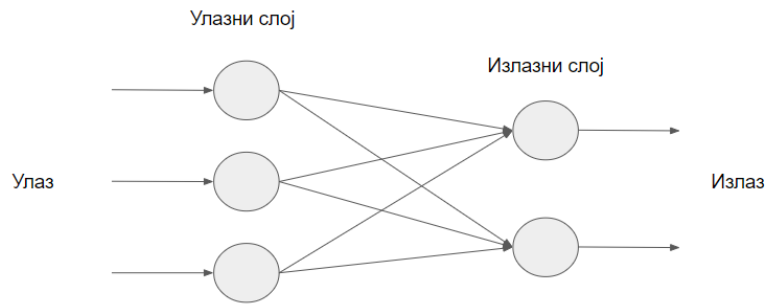
Један неурон  $k$  се описује коришћењем следеће две једначине:

$$u_k = \sum_{j=1}^n w_{kj} x_j \quad (2.1)$$

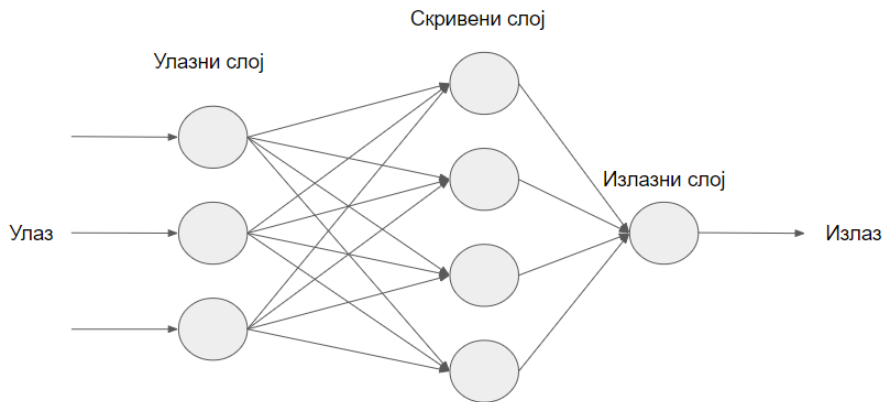
$$y_k = \varphi(u_k - \theta_k)$$

У оквиру ових једначина, а како је и приказано на претходној Слици 1,  $x_j$  где  $j=1, \dots, n$  представља улазни сигнал, а тежина за неурон  $k$  је  $w_{kj}$  где  $j=1, \dots, n$ . Линеарна комбинација свих улаза је означена као  $u_k$ , праг је  $\theta_k$ , активациона функција је  $\varphi(\cdot)$ , а излазни сигнал неурона је означен као  $y_k$ .

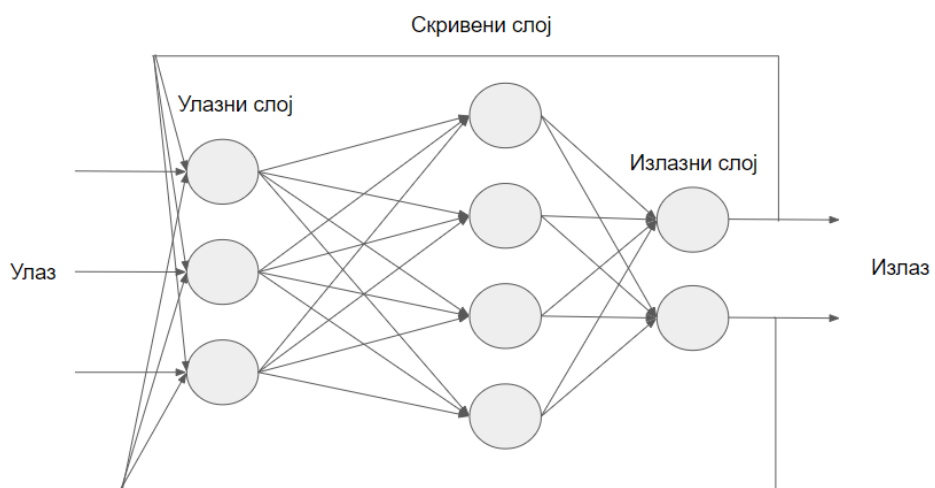
У зависности од тока прослеђивања информација и фунционишања, неуронске мреже се могу поделити на мреже са простирањем сигнала унапред (енг. *feedforward neural network*) и мреже са простирањем сигнала уназад, односно од излазних слојева ка улазним (енг. *feedback neural networks*). Неки од често коришћених модела неуронских мрежа су: (1) једнослојна неуронска мрежа са прослеђивањем сигнала унапред, (2) неуронска мрежа са једним скривеним слојем са прослеђивањем унапред и (3) рекурентна неуронска мрежа са скривеним слојем. Ради лакшег појашњења и истицања разлика у сложености алгорита, ова три типична пример неуронских мрежа су дата на следећим Сликама, а илустроване у (Engelbrecht, 2007).



**Слика 2.** Једнослојна неуронска мрежа са прослеђивањем унапред



**Слика 3.** Неуронска мрежа са једним скривеним слојем са прослеђивањем унапред



**Слика 4.** Рекурентна неуронска мрежа са скривеним слојем

У поређењу са конвенционалним технологијама, неуронске мреже нуде мноштво могућности системима за доношење одлука, за решавање проблема и примену савремених технологија у различитим областима (Hill & Remus, 1994; Aiken & Bsat, 1999). Један од могућности примене неуронских мрежа у ситуацији у којој не постоји јасно дефинисан модел је дат у (Dutta, Shekhar & Wong, 1994). У оквиру истраживања спроведеног у ранијем периоду, Бансал, Кауфман и Веиц (Bansal, Kauffman & Weitz, 1994) истичу да су применом неуронских мрежа добијени поузданији резултати, а само истраживање су спровели на скупу нелинеарних и непрецизних података. Нешто касније, друга група аутора (Zhang, Patuwo & Hu, 2001) је такође дошла до сличних закључних разматрања. Резултати постигнути у оквиру неколико каснијих истраживања такође потврђују да су неуронске мреже поузданије (Kim, Yang & Kim, 2008; Marques, Lacerda, Camargo & Teixeira, 2014; Schmidhuber, 2015) и погодније за примену у односу на конвенционалне моделе (Joseph, Larrain & Turner, 2014), када се ради о базама података које чине непрецизни и нелинеарни подаци.

Неуронске мреже могу симултано да обрађују различите типове нумеричких података (Calderon & Cheh, 2002), као и да укључе у обраду сезонске варијације у подацима, трендове и кретања (Zhang & Qi, 2005). Смит и Гупта (Smith & Gupta, 2000) истичу као основну разлику између неуронских мрежа и конвенционалних метода, не узимање претпоставки у погледу карактеристика података и расподеле. Једно од истраживања из ранијег периода, у оквиру кога је дато поређење неуронских мрежа и конвенционалних метода, је дато у (Ansuji, Camargo & Radharaman, 1996). На основу добијених резултата, аутори истичу да су неуронске мреже оствариле боље перформансе предвиђања (мању грешке у предвиђању очекиваних параметара). Слично истраживање је спроведено исте године од стране (Agrawal & Schorling, 1996), у оквиру кога су аутори упоређивали резултате неуронских мрежа са три различите конвенционалне технике, и дошли до истих закључака.

Услед карактеристика који су описани и истакнути у овом поглављу, евидентно је да су неуронске мреже нашле примену у многим доменима и индустријама, за решавање различитих врста проблема и као помоћ системима за

доношење одлука. Неке од примера примене неуронских мрежа се могу наћи у (Thomassey & Happiette, 2007; Sun, Choi, Au & Yu, 2008; Wong & Guo, 2010; Mansur & Kuncoro 2012; Moosmayer, Chong, Liu & Schuppar, 2013).

*„It’s foolish to think that anything we do is the so-called right way. It’s often simply one choice of many.“*

David Reimer<sup>1</sup>

### 3. ПРЕГЛЕД ПОСТОЈЕЋИХ РЕШЕЊА ЗА УПРАВЉАЊЕ РЕСУРСИМА У АВИО ИНДУСТРИЈИ

У овом поглављу је дат преглед постојећих решења за предвиђање броја „no-show“ путника у авио индустрији, као и за оптимизацију капацитета авиона, максимално искоришћење инвентара, предвиђање броја путника и за предвиђање „overbooking“ лимита. Посматрана су и анализирана решења заснована на техникама рачунарске интелигенције, као и решења заснована на статистичким техникама. Решења која укључују људски фактор су посебно назначена.

Обрада и анализа великих количина података је данас много лакша применом модерних компјутерских технологија, што се може видети у авио индустрији. Авио компаније данас свакодневно анализирају велике количине података много брже него у ранијим периодима када је ова индустрија била у развоју, па тако и могу доносити одлуке у реалном времену (Min & Joo, 2016; Pineda, Liou, Hsu & Chuang, 2018). Постоји мноштво техника које се базирају на математици и статистици, а које се могу применити за решавање сложенијих проблема и у свакодневном раду. Као што је дато у (Lim, Balas & Do, 2013), узимајући у обзир природу проблема и друге факторе који могу утицати на начин решавања, разликују се статистичке и технике рачунарске интелигенције. Статистичке, односно конвенционалне технике функционишу на принципу прецизности и извесности. Могу изискивати значајно више времена за анализу и обраду података, а као предуслов, већина мора имати и испуњене одређене претпоставке везане за улазне податке (нпр. расподелу). За разлику од статистичких, технике рачунарске интелигенције могу толерисати непрецизност података и присутност фактора неизвесности. Такође, изискују значајно мање времена за анализу и обраду података, и могу дати боља и прихватљивија решења

---

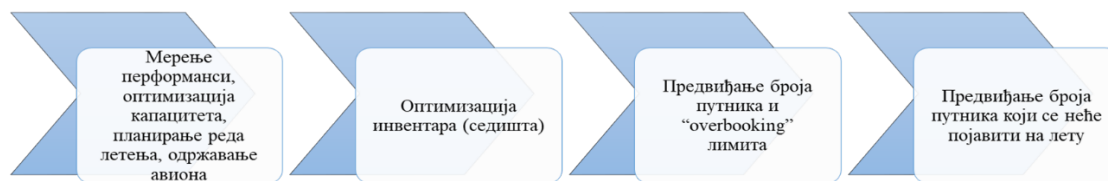
<sup>1</sup> Из књиге: Count like an Egyptian – A hands-on introduction to ancient mathematics

у ситуацијама када не постоје прецизни и потпуни подаци. Детаљнија анализа предложених решења на бази техника рачунарске интелигенције у оквиру различитих индустрија и пословних домена је дата у (Vojtek, 2017).

У данашње време, авио индустрија се усмерава ка коришћењу компјутерских технологија у готово свим доменима где потреба постоји. Бројна решења у виду система, модела и процеса су развијена и користе се у областима оптимизације капацитета (инвентара) авиона, предвиђање потражње за летовима, предвиђање броја путника који се неће појавити на лету, распоређивање посаде на различитим рутама и типовима авиона, итд. Најчешће ова решења комбинују већ поменути статистичке и технике рачунарске интелигенције, како би се постигла ефикасност и ефективност у одређеним областима и процесима, а све у циљу максимизације профита. Детаљан преглед литературе и решења заснована на статистичким и техникама рачунарске интелигенције која се користе у авио индустрији је дат у оквиру (Vojtek, Poledica & Petrovic, 2018).

Од поменутих домена, предвиђање броја путника који се неће појавити на лету (у даљем тексту „no-show“ путника) је област која је предмет истраживања овог рада. Ова област такође изискује обраду великог броја података, па због тога постоји потреба за применом компјутерских техника. У пракси, авио компаније често не располажу са комплетним подацима које ће користити за предвиђање „no-show“ путника. Узимајући то у обзир, ово поглавље има за циљ да представи одређена постојећа решења која су предложена и користе се за предивђање „no-show“ путника од стране авио компанија. Пошто је одабрана област од интереса за ово истраживање уско специјализована и односи се на конкретан проблем који даље није могуће декомпоновати, прво су анализирана свеобухватнија и сложенија решења. Као што је приказано на Слици 4, полазећи од решења која се односе на управљање свих ресурса, преко решења која се односе на управљање седиштима, препознате су и идентификоване уско специјализоване области које се односе на „overbooking“ концепт и предвиђање „no-show“ путника. Неопходно је истаћи да главни циљ свих ових решења представља унапређење пословања и повећање профита авио компаније, кроз управљање ресурсима.





**Слика 1.** Обухваћене области примене статистичких и техника рачунарске интелигенције у авио индустрији

Хронолошки посматрано, решења која су развијена за потребе авио индустрије су се у ранијим периодима више базирала на статистичким техникама. Касније, почевши од 80-тих и 90-тих година прошлог века, дефинисане су и све више почеле да се користе процесивне предности техника рачунарске интелигенције. Сходно томе, анализа релевантне литературе у овом поглављу прати хронологију развоја и примене компјутерских техника у авио индустрији, у области унапређења пословања и начина искоришћења капацитета. Анализа је подељена у два потпоглавља. Потпоглавље 3.1 садржи преглед постојећих решења која се заснивају на статистичким техникама, док су у потпоглављу 3.2 представљена решења која се заснивају на техникама рачунарске интелигенције. Ради боље прегледности, анализирана решења су груписана према предходно поменути областима проблема на које се односе.

### 3.1. Решења заснована на статистичким техникама

Као што је истакнуто у уводном делу овог поглавља, статистичке технике изискују знатно више времена и прецизно формулисан аналитички модел за процесуирање података. Решења заснована на статистичким техникама могу дати одличне резултате за решавање проблема код којих су наведени услови испуњени.

#### 3.1.1. Мерење перформанси авио компаније

Почевши од перформанси авио компанија на највишем нивоу, Ветхерфорд и Белобаба (Weatherford & Belobaba, 2002) истичу важност мерења финансијских резултата. Аутори предлажу коришћење симулационе анализе, а сам модел за доношење стратешких одлука дефинишу кроз хеуристичка правила. Истраживање у истој области је спроведено и годину дана касније од стране Бојда и Билегана

(Boyd & Vilegan, 2003) који у закључним разматрањима наводе као најважнији фактор за постизање профита механизме за максимално искоришћење инвентара (седишта у авиону). Ово је такође потврђено и у (Chen, Günther & Johnson, 2003), а аутори посебно истичу да су стратегије за оптимизацију седишта кључне јер омогућују поделу и груписање седишта у авиону на више и ниже ценовне тарифе. Као крајњи циљ, долази до максимизације профита авио компаније.

### 3.1.2. Оптимизација инвентара (седишта)

Значајно интересовање авио компанија за имплементацију и коришћење решења којима је могуће максимално искористити ограничен број седишта у авионима је било у периоду између 70-тих и 80-тих година прошлог века. Како бројни аутори истичу, примена предложених оптимизационих модела је довела до значајног повећања прихода авио компанија у том периоду, па је тиме и интересовање ка проналажењу што ефикаснијих модела било веће. Група аутора (Yu, Chang & Chen, 2016) кроз истраживање које је спроведено много касније, истиче да авио компаније и даље морају наставити да се фокусирају на што боље искоришћење инвентара којим располажу (у зависности од типа авиона, спецификација и броја седишта). Ово је веома битно за нискобуџетне (енг. *low-cost*) авио компаније, које свој комплетан бизнис модел граде на принципу максималног искоришћења инвентара. Један од првих пионира у овој области је свакако Белобаба, који кроз истраживање објављено 1989 (Belobaba, 1989) предлаже EMSR (енг. *expected marginal seat revenue*) модел за управљање седиштима и максимизацију капацитета авиона. Модел је тестиран и пуштен у примену кад додатак тадашњем аутоматизованом систему резервација у Вестерн Аирлајнс авио компанији (Western Airlines). Белобаба предлаже и даје препоруку да се овај модел користи на летовима са изразито високом потражњом. Пошто такви летови генеришу и највише профита, улагање додатног напора у виду додатног планирања капацитета и инвестирање ресурса је оправдана. Временом се показало да Белобабин EMSR модел даје очекиване резултате, па је унапређен неколико година касније од стране Робинсона (Robinson, 1995) кроз увођење оптималних стратегија. Предложене стратегије се активирају у тренутку када је потребно донети одлуку да ли откључати вишу ценовну тарифу, остати на истој

или спустити на нижу тарифу. На овај начин омогућено је авио компанијама да оставе одређени број слободних седишта потенцијалним путницима који би желели да резервишу седиште са вишом ценовном тарифом, која носи са собом низ других погодности а самим тим и већу цену.

Томпсон је спровео једно од првих истраживања (Thompson, 1961) које је обухватило проблем резервисања седишта са више различитих ценовних тарифа. Нешто касније, Ротстеин се бави анализом овог истог проблема (Rothstein, 1984), али фокусом на технике операционих истраживања које су авио компаније тада примењивале у домену резервације седишта са више различитих ценовних категорија. Као што је раније истакнуто, максимално искоришћење инвентара доводи до повећања профита, па је тако овај проблем истакнут као један од приоритетнијих за решавање у оквиру комерцијалног сектора. Група аутора 90-тих година спроводи истраживање (Smith, Leimkuhler & Darrow, 1992) у оквиру кога пореди перформансе и резултате постигнуте применом више различитих модела који су се користили за оптимизацију седишта и одређивање ценовних тарифа. Почетком 21-ог века, друга група аутора (Van Ryzin & McGill, 2000) сугерише нешто једноставније решење које комбинује адаптивне алгоритме и стохастичку теорију апроксимације. Слично, у литератури се могу наћи и предложена решења новијег датума која комбинују статичке и динамичке стратегије (Lan, Gao, Ball & Karaesmen, 2008), која се фокусирају на технике стохастичног програмирања (Ahmed & Pojari, 2008; Gu & Zhu, 2016). Одређени модели моделују и ризик и снижења у ценама седишта (нпр. погледати: Fan & Wang, 2013) и предлажу проширење постојећих модела на паралелне трансферне летове (Huang, Ge, Zhang & Xu, 2013).

### 3.1.3. Предвиђање броја путника

Паралелно са развојем модела за оптимизацију, повећавала се потражња за летовима. Сходно томе, авио компаније су почеле више да се фокусирају на примену решења којим ће што боље моћи да предвиде број путника на летовима и да одреде „overbooking“ лимите. Дефиниција „overbooking“ концепта је дата у уводном делу овог рада, а односи се на омогућавање већег броја резервација у односу на расположиви капацитет авиона, и то на летовима који су већ попуњени.

Предвиђање броја путника на летовима је веома битно за авио компанију (Grammig, Hujer & Scheidler, 2005), а значајни финансијски резултати се могу остварити прецизнијим моделима (Riddington, 1987). Када се узму у обзир статистичке технике које се користе за предвиђање боја путника, постојећи модели најчешће се базирају на временским серијама. Ридингтон је спровео једно од можда и првих истраживања на ову тему (Riddington, 1987), у оквиру кога је истакнута значајност прецизности самих решења за предвиђање. Ридингтон предлаже решење које се заснива на комбинацији ARIMA (енг. *autoregressive integrated moving average*) модела, кеш менаџмент процесом и сигурносним залихама, како би се остварила што већа прецизност. Пример унапређења постојећег решења увођењем новог аспекта се може видети у (Benitez, Paredes, Lodewijks & Nabais, 2013). Ова група аутора предлаже увођење временски променљивих коефицијената у постојећи модел који се заснива на временским серијама. Овако унапређено решење је тестирано на 5857 различитих дестинација, што је обухватило и више различитих авио компанија. Аутори истичу да су остварени бољи резултати у поређењу са постојећим много једноставнијим моделом без временски променљивих коефицијената. Друга група аутора (Tsui, Balli, Gilbey & Gow, 2014) тестира два такође унапређена ARIMA модела, SARIMA (Box-Jenkins seasonal ARIMA) и ARIMAX, а предложена унапређења имају за циљ бољу идентификацију и моделовање тренда. За потребе валидације, аутори користе податке са аеродрома Хонг Конг о путницима у различитим деловима сезоне. Аутори закључују да су оба нова модела остварила прецизније резултате, а за оцену перформанси модела коришћени су показатељи MAPE (енг. *Mean absolute percentage error*) и RMSE (енг. *Root mean square error*). На сличан начин, Филдер, Веи и Исмаил (Filder, Wei & Ismail, 2011) анализирају економетријске моделе и оцењују перформансе кроз више различитих мера грешака, односно показатеља.

#### 3.1.4. Одређивање „overbooking“ лимита

Примена „overbooking“ лимита на постојећи инвентар доприноси повећању профита авио компанија (Somboon & Amaruchkul, 2017). Анализа понашања

путника и начина на који спроводе избор седишта и ценовних тарифа је спроведена у не тако давно објављеној анализи 2015. године (погледати: Zeng & Li, 2015). Аутори у свом раду предлажу управо коришћење „overbooking“ стратегије које се могу активирати у различитим ситуацијама, у зависности од попуњености авиона. За одређивање „overbooking“ лимита, Сомбон и Амарчукул (Somboon & Amaruchkul, 2017) предлажу примену модела који се базира на вероватноћи појављивања путника на лету. Након спроведеног тестирања, аутори потврђују да долази до повећања профита у ситуацији када су примењени „overbooking“ лимити. Још један модел који такође користи вероватноћу појављивања путника на лету је предложен у оквиру (Amaruchkul & Sae-Lim, 2011). Решење је тестирано коришћењем две расподеле појављивања путника на лету – нормалну, бета, као и детерминистички модел. Модел који је користио бета расподелу је дао најбоље резултате. Упоредо са поменутиим моделима који користе вероватноћу појављивања путника на лету, неопходно је представити и моделе који се фокусирају на губитке које „overbooking“ приступ може проузроковати. Ови губици настају када на једном лету се појави више путника (са потврђеном резервацијом и купљеном картом) од капацитета са којим авион располаже. Узимајући ово у обзир, груп аутора (Lan, Ball, Karaesmen, Zhang & Liu, 2015) предлаже модел који акценат ставља управо на ове ситуације, и генерише решење које може произвести најмање губитке уколико се најнеповољнији сценарији остваре.

Предвиђање путника који се неће појавити на лету, представља један од ефикаснијих начина на основу кога авио компаније могу одредити „overbooking“ лимите. Ово је препознато од стране више аутора (Lawrence, Hong & Cherrier, 2003; Fildes, Nikolopoulos, Crone & Syntetos, 2008; Kunnumkal, Talluri & Topaloglu, 2012), а представља и ужу област која ће се даље обрађивати у овом раду.

### 3.2. Решења заснована на техникама рачунарске интелигенције

Авио индустрију данас, као и многе друге индустрије, карактеришу сложене и учестале промене које долазе и наметнуте су од стране окружења. Како би опстале на тржишту, одржале конкурентност и максимизовале профит, авио

компаније су приморане да често доносе одлуке које се темеље на непрецизним и непотпуним подацима, базиране и на симулацијама и очекивањима. Управо ово представља прилику за искоришћење пуног потенцијала техника рачунарске интелигенције, како би се премостио проблем непрецизних и непотпуних података.

Фази логика (енг. *fuzzy logic*) и неуронске мреже (енг. *neural networks*) представљају две најпопуларније и најчешће коришћене технике рачунарске интелигенције. Када је у питању фази логика, потенцијал примене лингвистичких варијабли је најчешће искоришћен код мерења квалитета услуге авио компаније. Неуронске мреже су до сада највише коришћене за предвиђање потражње за летовима, односно броја путника који се може очекивати на лету. Како би се повећала прецизност резултата, најчешће су коришћене у комбинацији са другим статистичким и техникама рачунарске интелигенције. Поред две поменуте технике, кроз ово истраживање су обухваћена решења базирана на генетским алгоритмима (енг. *genetic algorithms*). Техника закључивања на основу случаја је од посебног интереса за целокупан рад, те су потенцијалне прилике за њену употребу у специфичним доменима пословања авио компанија посебно наглашене.

Као и код претходног потпоглавља, решења заснована на техникама рачунарске интелигенције а која се односе на авио индустрију су груписана према области на коју се односе.

### 3.2.1. Мерење перформанси авио компаније

Користећи пун потенцијал компјутерских технологија, група аутора (Pineda, Liou, Hsu & Chuang, 2018) предлаже модел заснован на техникама рачунарске интелигенције који за циљ има идентификацију и рангирање авио компанија на основу њихових финансијских и оперативних перформанси. Предложени хибридни систем комбинује DRSA, DEMATEL, DANP и VIKOR методе. Демирел и Девеци (Demirel & Devenci, 2017) предлажу употребу генетског алгоритма који за циљ има проналажење оптималне комбинације посаде која би изискивала минималне трошкове, узела у обзир све летове и испуњавала све

регулаторне критеријуме. Слично, за решење проблема оптимизације посаде, друга група аутора (Koticha, Sanghani & Gambhava, 2004) такође предлаже решење засновано на генетском алгоритму. Тестирање предложеног модела је спроведено на 28 проблема из праксе и како се у закључку истиче, обећавајући резултати су остварени. Атли и Кахраман (Atli & Kahraman, 2012) су применили фази алгоритам за проналажење критичног пута при планирању одржавања авиона. Као финални резултат, успели су да применом фази модела смање време које је потребно за планирање одржавања авиона. Друга група аутора (Chiu, Chiu & Hsu, 2004), предлаже комбинацију генетског алгоритма са техником закључивања на основу случаја за употребу у оквиру процеса одржавања авиона. Циљ предложеног система је да се одаберу слични случајеви из базе знања и да се на основу генетског алгоритма израчунају динамичке тежине. Нешто касније, аутори (Ahmed & Roojari, 2008) су показали да се стохастичко програмирање може применити у процесу планирања авио компаније. Преиспитивањем коришћења оптимizacionих модела у главним процесима за планирање реда летења, аутори истичу да је битно да решење буде робусно и предлажу коришћење стохастичког програмирања како би се моделовала неизвесност приликом планирања реда летења. Овај проблем је такође анализиран од стране друге групе аутора (Deveci, Demirel & Ahmetoğlu, 2017) који су користили фази TOPSIS вишекритеријумски метод за доношење одлука приликом избора нових дестинација.

Поред мерења перформанси, неопходно је истаћи и област оцене квалитета услуге које авио компаније пружају. Применом фази логике у овом домену, предложена су решења која могу дати изванредне резултате. За мерење квалитета услуге коју авио компаније пружају путницима, група аутора (Chou, Liu, Huang, Yih & Han, 2011) предлаже SERVQUAL метод са тежинским факторима. Како су људске оцене често неодређене, аутори истичу да је много боље користити лингвистички приступ у описивању очекиване вредности. Предложен метод је примењен у Тајванској авио компанији, оцене о квалитету услуге су прикупљене од путника, анализирани и предлози за побољшања су формулисани. Још један хибридни систем је развијен од стране (Li, Yu, Pei, Zhao & Tian, 2017) који се састоји из аналитички хијерархијског процеса и двоструког фази лингвистичког

метода. Квалитет услуге у току лета је мерен код три авио превозника из Кине. Коришћењем комбинованог фази приступа за доношење одлука, (Perçin, 2017) предлаже метод за оцену квалитета услуге авио компаније у Турској.

### 3.2.2. Оптимизација инвентара (седишта)

У циљу повећања перформанси система за управљање седиштима, (Chang, Hsieh, Yeh & Liu, 2006) предлажу систем за расподелу седишта на бази технике закључивања на основу случаја. Користећи податке авио компанија које располажу великим капацитетима, тестирање је спроведено и утврђено да предложени систем даје боље резултате од основног метода који се тренутно користио (на бази ко први дође, први буде услужен, енг. *First Come First Served*).

### 3.2.3. Предвиђање броја путника

Различите технике су коришћене при изради решења у виду система и модела за предвиђање броја путника на летовима. Упоредни приказ више модела се може наћи у (Cheng & Mengting, 2018), где је посебан акценат стављен на предиктивну моћ неуронских мрежа, фази и осталих модела заснованих на техникама рачунарске интелигенције. Од свих, можда је и најизраженија примена неуронских мрежа вишеструких структура са различитим алгоритмима учења. Једна од ранијих примена неуронских мрежа за предвиђање у авио индустрији је предложена од стране Нама и Шафера (Nam & Schaefer, 1995). Аутори су препознали неопходност што прецизнијег предвиђања броја путника на лету како би авио превозник могао донети одлуку о расподели седишта и класа. Управо због потребе за што прецизнијим предвиђањима, аутори пореде резултате добијене употребом неуронских мрежа са резултатима које су произвеле статистичке технике. У закључку наводе да је решење засновано на неуронским мрежама дало прецизније резултате и предлажу укључивање истог у процес предвиђања. Група аутора (Weatherford, Gentry & Wilamowski, 2003) такође предлаже неуронске мреже за предвиђање броја путника на лету, на основу поређења резултата са



неколико различитих статистичких техника. У њиховом истраживању, чак је и неуронска мрежа са основном структуром дала боље резултате. Слично поређење неуронских мрежа са статистичким техникама (тачније са Box-Jenkins ARIMA моделом) се може наћи у (Chikr-El-Mezouar & Mohamed Hassan Gabr, 2011). Аутори су за тестирање користили две базе података и у оба случаја неуронске мреже су дале боље резултате.

#### 3.2.4. Одређивање „overbooking“ лимита

Још један веома битан и популаран концепт, који авио компаније користе за повећање прихода по лету, представља „overbooking“ (Zeng & Li, 2015). Проблем одређивања „overbooking“ лимита је у фокусу већ десетак година, а авио компаније константно траже нова решења која би што прецизније могла да одреде ове лимите. Бенефити од овог концепта су изражени на летовима који су попуњени у потпуности, а може се представити као разлика између изгубљених прихода услед не појављивања путника или касног отказивања резервације (Nambisan, 2003). Томсон спроводи једно од првих истраживања које се може наћи у литератури на ову тему, у оквиру кога препознаје „overbooking“ као један од начина за повећање профита са постојећим капацитетима (погледати: Thompson, 1961). Нешто касније Ротстеин (Rothstein, 1984) у свом истраживању које је у претходним поглављима овог рада већ разматрано, наводи два главна разлога због којих авио компаније морају увести „overbooking“ концепт у своје редовно пословање. То су пре свега тенденција путника да отказују резервацију 3 дана или чак 12 сати пре самог лета и једноставно непојављивање одређених путника на лету. Посматрано из угла авио компаније, разлог за отказивање или не појављивање није битан, јер свакако авион ће полетети и отићи до дестинације са једним или више празних седишта, а та иста седишта су могла бити попуњена другим потенцијалним путницима који су приликом резервације одбијени јер је авион у том тренутку био попуњен. У зависности од специфичности ситуације на лету за који се примењује, приступ „overbooking“ може донети са собом како позитивне, тако и негативне резултате по авио компанију (Suzuki, 2006). Негативни аспекти се могу јавити у ситуацији када лет има више путника са

резервацијама него капацитет авиона, а потенцијални губитак и трошак по авио компанију може настати уколико другим путницима компаније мора обезбедити други лет, или чак платити пуну резервацију код конкурентске авио компаније.

### 3.2.5. Предвиђање броја „no-show“ путника

Као што је и наведено у претходном потпоглављу, информације о броју путника који се неће појавити на лету или ће отказати лет у неком тренутку, могу бити веома значајне за авио компаније, јер се на основу њих могу одредити „overbooking“ лимити. Број путника који се неће појавити на лету када су у питању распродати летови може износити приближно 15%, како је указано у једном од ранијих истраживања (погледати: Smith, Leimkuhler & Darrow, 1992). Ово указује да ће значајан број седишта остати празан на конкретном лету, те да авио компаније морају настојати да продају бар одређени део тих седишта другим потенцијалним путницима. Сходно томе, јасно је да авио компаније теже да користе што прецизнија решења за предвишање „no-show“ путника.

Фаравеј и Чатфилд су 1998. године (Faraway & Chatfield, 1998) формулисали једно од решења које се заснива на веома популарној техници која спада у групу техника рачунарских интелигенција. Наиме, аутори користе неуронске мреже у свом моделу за предвиђање „no-show“ путника. Добијени резултати су још тада указивали да неуронске мреже, употребљене у овом домену, се могу показати као веома ефикасан алат. У оквиру даљих разматрања, Фаравеј и Чатфилд наводе да би предложено решење било још прецизније уколико би се укључило експертско знање у овој области, односно препорука експерта. Приликом истраживања доступне литературе, није пронађено пуно радова који се односе на примену техника рачунарске интелигенције у домену предвиђања „no-show“ путника, па чак ни у домену одређивања „overbooking“ лимита. Такође, доста је ограничен број радова који предлажу употребу техника рачунарске интелигенције за предвиђање броја путника, односно потражње за самим летовима. Закључивање на основу случаја, као једна од техника рачунарске интелигенције, може бити веома погодна за употребу у поменутим доменима, јер симулира људско закључивање за решавање специфичних проблема. Што се тиче авио индустрије, а како је и истакнуто у ранијим деловима овог рада, техника

закључивања на основу случаја је до сада предлагана више пута за коришћење у оквиру домена расподеле седишта (Chang, Hsieh, Yeh & Liu, 2006), али и у оквиру процеса одржавања авиона (Magaldi, 1994; Richard, 1997; Chiu, Chiu & Hsu, 2004).

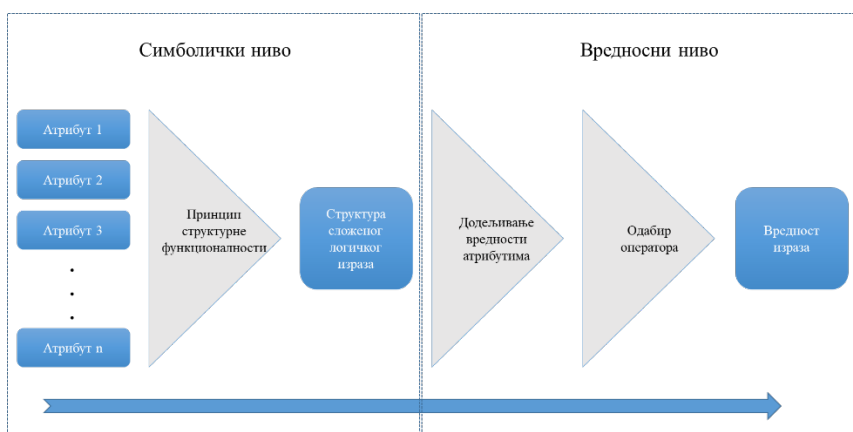
*„Истински интелигентна особа поздравља нове идеје, јер оне допуњавају синергију других акумулираних идеја..“*

Роберт Киосаки

## 4. ИНТЕРПОЛАТИВНА БУЛОВА АЛГЕБРА

Интерполативна Булова алгебра (ИБА) представља конзистентну  $[0,1]$  - реализацију Булове алгебре јер чува све законе на којима почива Булова алгебра. Предложена је од стране Драгана Радојевића (Radojević, 2008a) као одговор на проблеме неважења свих Булових закона у конвенционалној фази и већини вишеверносних логика. У оквиру ИБА реализације, сви Булови закони су задовољени у  $[0,1]$  оквиру, укључујући и законе искључење трећег и контрадикције.

Интерполативна Булова алгебра се састоји из два нивоа – симболичког и вредносног (Radojević, 2008a). На симболичком нивоу, атрибути се посматрају независно од њихових вредности, и пратећи принцип структурне функционалности, рачуна се структура сложеног логичког израза (Radojević, 2008d). На вредносном нивоу, атрибутима се додељују вредности, бира се одговарајући оператор у зависности од природе атрибута и/или њихове статистичке међузависности, и рачуна сама вредност израза остајући у Буловом оквиру. Следећа Слика илуструје поступак, полазећи од појединачних атрибута па све до вредности логичког израза.



Слика 5. ИБА – симболички и вредносни ниво

Два тренутно најуспешнија правца примене ИБА се односне на логичку агрегацију и меру сличности/различитости (Radojević, 1999; 2008c; Poledica, Milošević, Dragović, Petrović & Radojević, 2015; Milošević et al., 2018). Основна математичка својства, као и могућност и погодност примене ИБА мере сличности у различитим областима се могу наћи у (Radojević, 2006; Poledica, Milošević, Dragović, Petrović & Radojević, 2015; Milošević et al., 2018).

Сама теорија ИБА је до сада имала значајан утицај на теорију и праксу, а представља и основ неколико докторских дисертација у различитим областима. У оквиру (Milošević, 2017), проучавана је могућност и оправданост увођења интерполативне Булове алгебре у теорију интуиционистичких фази скупова. Логички приступ моделовању сличности који се заснован на интерполативној Буловој алгебри је предложен и уведен у (Poledica, 2016). Други пример коришћења ИБА теорије се односи на истраживање које је имало за циљ дефинисање новог система закључивања који изражава знање експерата на лакши, прикладнији и природнији начин (Dragović, 2016). Једна од две најчешће коришћене технике рачунарске интелигенције која је и објашњена у поглављу 2. овог рада, конзистентна фази логика, је коришћена за формализацију и објашњење концепта у случају градације (Dobrić, 2016). Фази логика је коришћена и код дефинисања методологије система раног упозорења које карактерише вербална интерпретабилност (Kovačević, 2014). Такође, ИБА је са успехом примењивана у различитим доменима пословања, као што су алгоритамско трговање (Rakićević, Simeunović, Petrović & Milić, 2018), електронска набавка (Bobar, 2015) и развој информационих система (Turajlić, 2014).

Ово поглавље је организовано на следећи начин. У потпоглављу 4.1. су изложени основни појмови ИБА, а однос принципа истинитостне и структурне функционалности је објашњен кроз два примера у авио индустрији. Након тога, ИБА структура је представљена у потпоглављу 4.2, а симболички и вредносни ниво су објашњени у потпоглављу 4.3. У потпоглављу 4.4 су дефинисана два правца примене ИБА – логичка агрегација и ИБА мера сличности, односно различитости.

## 4.1. Основни појмови

Интерполативна Булова алгебра (ИБА) представља конзистентну  $[0,1]$  - реализацију Булове алгебре јер чува све законе на којима почива Булова алгебра (Radojević, 2008a). Ради бољег објашњења ИБА концепта, неопходно је осврнути се на дефиницију Булове алгебре.

**Дефиниција 1 (Brown, 2012).** Булова алгебра је уређена петорка  $\langle B, \wedge, \vee, 1, 0 \rangle$ , где је  $B$  скуп елемената,  $\wedge$  и  $\vee$  бинарне операције на  $B$ , а 0 и 1 крајњи/супротни елементи  $B$ , таква да за свако  $x, y, z \in B$  важи:

(т1)  $x \vee y = y \vee x$ ,  $x \wedge y = y \wedge x$  (комутативност);

(т2)  $x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$ ,  $x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$  (дистрибутивност);

(т3)  $x \vee 0 = x$ ,  $x \wedge 1 = x$  (идентичност);

(т4)  $x \vee \neg x = 1$ ,  $x \wedge \neg x = 0$  (закон искључења трећег и контрадикције). ■

Пратећи нотацију и означавање које се користи у радовима који дефинишу Интерполативну Булову алгебру, у даљим деловима овог рада, за скуп атрибута ће се уместо ознаке  $B$  користити  $\Omega$ . Сходно томе, елементи скупа  $\Omega = \{a_1, \dots, a_n\}$  представљају атрибуте, односно посматране особине, које ће бити објашњене на примерима који следе у оквиру овог потпоглавља. Оно што је битно истаћи, у оквиру класичне Булове алгебре, вредносна реализација ових атрибута подразумева коришћење две вредности  $a_i^v \in \{0, 1\}$ ,  $i = 1, \dots, n$ . С друге стране, кад се посматра ИБА оквир, ови атрибути имају  $[0,1]$ -вредносну реализацију  $a_i^v \in [0, 1]$ ,  $i = 1, \dots, n$ . ИБА структура је детаљније објашњена у потпоглављу 4.2.

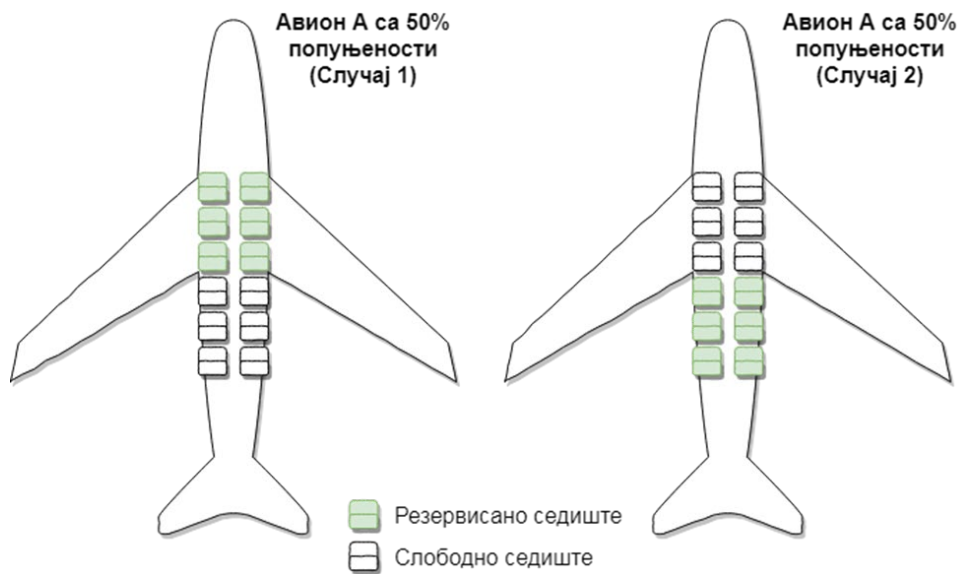
Интерполативна Булова алгебра се састоји из два нивоа – симболичког и вредносног (Radojević, 2008a). Ради лакшег разумевања нивоа, неопходно је прво истаћи да је ИБА заснована на принципу структурне функционалности (енг. *Principle of structural functionality*). Као што је објашњено у (Radojević, 2008d), то подразумева да структура било ког елемента ИБА може бити директно одређена

на основу структуре њених компонената. Сходно томе, на симболичком нивоу се атрибути (компоненте) посматрају кроз структуру, независно од њихове вредности. На вредносном нивоу се атрибутима додељују вредности, бира се одговарајући оператор (у зависности од природе атрибута и/или њихове статистичке међузависности) и рачуна вредност израза остајући у Буловом оквиру.

Посебан случај принципа истинитосне функционалности представља принцип структурне функционалности, који налаже да се ИБА трансформације обављају на симболичком нивоу, пре увођења вредности. На овај начин је обезбеђено да се негација третира, прво на структурном нивоу, а не да се негираној променљивој прво додели вредност. За потребе овог рада, принцип структурне функционалности је објашњен кроз два једноставна примера у авио индустрију – (1) поређење попуњености авиона А и (2) поређење резервација путника.

#### 4.1.1. Пример 1 – поређење попуњености авиона

У оквиру овог примера посматрајмо авион А са капацитетом од 100 места (седишта). Два случаја се разматрају, у оквиру којих је авион оба пута био 50% попуњен, међутим у питању су била различита седишта у путничкој кабини. Неједнакост авиона А у ова два случаја попуњености је приказано на следећој слици.



**Слика 6.** Неједнакост Авиона А у два случаја попуњености

Према Слици 6. иако је у оба случаја авион А 50% попуњен, и има исту тачку поласка и одредиште, исти датум и време полетања и слетања, као и исти број празних седишта, разликује се на структурном нивоу. Наиме, празна седишта у оба случаја (компоненте / атрибути) нису исте класе, нису повезана са истим додатним услугама и не садрже исте тарифне карактеристике. За потребе овог примера, сматраћемо да су првих шест седишта на претходној слици означена као седишта бизнис класе (уз све погодности које у оквиру резервације припадају). Сходно томе, како би се одредиле вредности празних седишта, односно вредност авиона А у оба случаја, прво је потребно посматрати њихове компоненте (седишта) на структурном нивоу. Ако случај 1 означимо као  $a = 0.5$ , онда случај 2 можемо означити као  $-a = 0.5$ . Ако се пореде вредности, једна половина (50%) је једнака једној половини, али променљива  $a$  се не може сматрати једнаком негираној вредности исте променљиве  $-a$ . Према наведеном примеру се може закључити да се принцип истините функционалности фокусира на структуру посматране променљиве а не на њену саму вредност. Другим речима, ИБА је погодна за решавање проблема у оквиру којих су наизглед вредности једнаке, али се структуре, природе и порекло посматраних променљивих разликују.



#### 4.1.2. Пример 2 – поређење резервација путника

У другом примеру се пореде две резервације, два различита путника који седе један поред другог у авиону. Може се закључити да оба путника путују истом класом (или бизнис или економска), путују до исте дестинације и имају исте погодности у авиону. Иако су наизглед ове две резервације исте, разликује се на структурном нивоу. Атрибути од интереса који се овде могу посматрати су канал куповине карте (резервације), цена и начин плаћања, могућност пријаве на лет и погодности за чланове одређених програма (енг. *Loyalty memberships*).

Путник  $X$  је своју карту купио преко веб сајта авио компаније и платио кредитном картицом, док је путник  $Y$  своју карту купио преко овлашћене агенције и платио је у готовини. Погоднију, односно „бољу“ цену је добио путник  $Y$ , пошто овлашћена агенција баш за ту дестинацију добија повољније тарифе од авио компаније (честа је пракса да авио компаније одређене дестинације у одређеним временским периодима промовишу и самим тим одобравају ниже цене овлашћеним агенцијама како би их мотивисале ка продаји резервација на тим дестинацијама). С друге стране, пошто је купио преко веб сајта, путник  $X$  је у могућности да се пријави на лет преко интернета 48 сати пре лета (енг. *online check-in*), док путнику  $Y$ , карту може „чекирати“ искључиво агенција. На крају, путник  $X$  може приликом куповине карте унети на веб сајту број чланске картице (енг. *loyalty membership card*) и на основу дестинације бити награђен са одређеним бројем миља на свом рачуну који касније може искористити за покриће дела цене приликом куповине следеће карте. Путник  $Y$  нема могућност добијања награде у виду миља јер обавља куповину преко овлашћене агенције по повољнијој цени.

#### 4.2. ИБА структура

У оквиру ИБА структуре елементи скупа  $\Omega = \{a_1, \dots, a_n\}$  представљају атрибуте. Сваки од атрибута  $a_i, i = 1, \dots, n$  представља по једну посматрану особину. У наведеном првом примеру из претходног потпоглавља, особине седишта би биле: класа ( $a_1$ ), број повезаних додатних услуга ( $a_2$ ), рефундациона

опција ( $a_3$ ), итд. У описаном другом примеру, особине резервације би биле: канал куповине карте ( $a_1$ ), цена и начин плаћања ( $a_2$ ), могућност пријаве на лет ( $a_3$ ), погодности за чланове одређених програма ( $a_4$ ), итд. Ови атрибути представљају елементе логичких функција у оквиру ИБА. Број логичких функција које се могу формирати над  $n$  атрибутима (елементима) је  $2^{2^n}$ . Па тако, уколико се авионско седиште (место) код првог наведеног примера описује помоћу два атрибута, могуће је формирати укупно 16 логичких функција. Уколико се број атрибута повећа само за један, односно, уколико се резервација из другог примера састоји од нпр. 3 атрибута, могуће је формирати укупно 256 логичких функција. Сходно томе, може се констатовати да сложеност логичког израза у ИБА расте експоненцијално са повећавањем броја атрибута.

За израчунавање структуре, ИБА користи атомске елементе Булове алгебре (Radojević, 2008a). Они представљају најједноставније елементе Булове алгебре.

**Дефиниција 2 (Radojević, 2008a).** Атомски елемент  $\alpha(S)(a_1, \dots, a_n)$ ,  $S \in P(\Omega)$  је најједноставнији елемент Булове алгебре  $B(\Omega)$  у смислу да у себи не садржи ниједан други елемент, осим тривијалне константе 0:

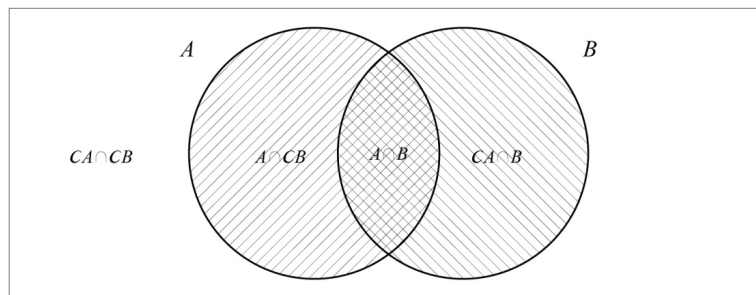
$$\alpha(S)(a_1, \dots, a_n) = \bigwedge_{a_i \in S} a_i \bigwedge_{a_i \in \Omega \setminus S} \neg a_i.$$

Сходно дефиницији, атомски елементи Булове алгебре су логичке функције које не садрже ниједан други елемент Булове алгебре осим елемента 0 (нула). Унија свих атома представља потпун скуп, док је пресек било која два атома у суштини празан скуп. Код најједноставнијих елемената, односно атома, постоји  $2^n$  најједноставнијих логичких функција од  $n$  елемената.

**Дефиниција 3 (Radojević, 2008c).** Структурна функција  $\sigma_\varphi$  анализираниог елемента Булове алгебре  $\varphi(a_1, \dots, a_n) \in B(\Omega)$  је скуп функција  $\sigma_\varphi : P(\Omega) \rightarrow \{0, 1\}$  такав да за  $S \in P(\Omega)$  важи:

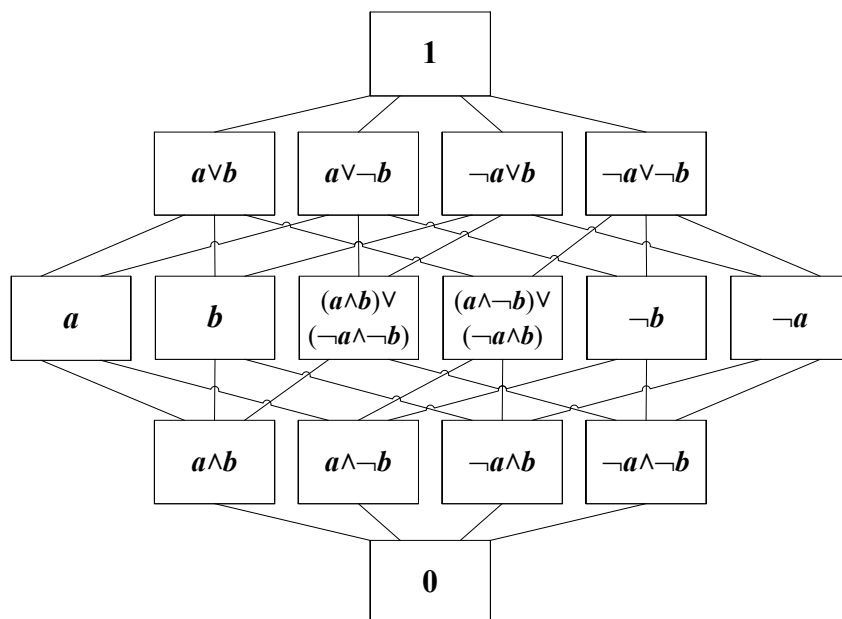
$$\sigma_\varphi(S) = \begin{cases} 1, & \alpha(S)(a_1, \dots, a_n) \subseteq \varphi(a_1, \dots, a_n) \\ 0, & \alpha(S)(a_1, \dots, a_n) \not\subseteq \varphi(a_1, \dots, a_n) \end{cases} \quad \blacksquare$$

У циљу што једноставнијег појашњења ИБА структуре, потребно је приказати структуру елемената Булове алгебре. Посматрано кроз стандардне логичке операције дисјункције ( $a \vee b$ ) и конјункције ( $a \wedge b$ ) атрибута, њихова конјункција је једнака универзуму, док је дисјунктивност приказана Веновим дијаграмом кроз скуповну нотацију на Слици 7. Комплемент скупа је означен као  $C$ .



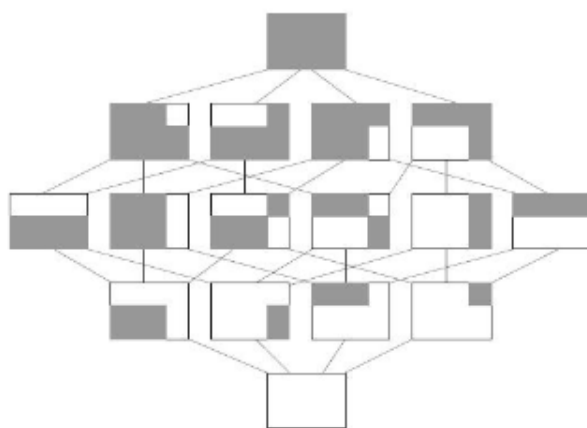
**Слика 7.** Венов дијаграм дисјунктивности атома Булове алгебре над атрибутима  $a$  и  $b$

У случају 2 атрибута  $a$  и  $b$ , могу се формирати 4 атома у оквиру Булове алгебре:  $a \wedge b, \neg a \wedge b, a \wedge \neg b, \neg a \wedge \neg b$ . Ови атоми су приказани кроз Хасеов дијаграм на следећој Слици.









**Слика 8.** Хасеов дијаграм за Булову алгебру над атрибутима  $a$  и  $b$  (Radojević, 2008c)

Атомски елементи се налазе у претпоследњем реду, а линије представљају релације укључености. У оквиру Хасеовог дијаграма ознака  $\neg$  представља ознаку за негацију. Као што је претходно наведено, пресек било која два елемента дају празан скуп (0), док унија било која два елемента представља потпун скуп (1). Праћењем релација на дијаграму, можемо рећи да логичка функција  $\varphi(a,b) = \neg a \vee \neg b$  укључује атоме  $a \wedge \neg b, \neg a \wedge b, \neg a, \neg b$ . Управо ова релација укључености атома у логичку функцију представља основ структурног нивоа ИБА. Аналогно Хасеовом дијаграму, ИБА структура је приказана у (Radojević, 2005) и дата на следећој Слици.



Слика 9. Приказ структуре ИБА елемената са атрибутима  $a$  и  $b$

На претходној Слици се јасно види укљученост једноставнијих елемената (атома) у сложеније елементе. Потпуно бео правоугаоник представља празан скуп (0) а четири правоугаоника изнад њега представљају 4 атома:  $a \wedge b$ ,  $\neg a \wedge b$ ,  $a \wedge \neg b$ ,  $\neg a \wedge \neg b$ . Па тако је елемент  $a$  представљен као , елемент  $b$  као , а њихов пресек ( $a \wedge b$ ) као . С друге стране, елемент  $a$  представљен као  се састоји од два атомска елемента  $a \wedge b$  и  $a \wedge \neg b$  који су представљени као  и . Очигледно је да  $a$  представља збир површина ова два елемента.

### 4.3. Симболички и вредносни ниво

Имајући у виду ИБА симболички и вредносни ниво (Radojević, 2008a), прво се одређује структура целе логичке функције на симболичком нивоу (на основу структура елемената те функције), а затим прелази на вредносни ниво ИБА.

#### 4.3.1. Симболички ниво

Сходно принципу структурне функционалности (Radojević, 2008d), сва правила Булове алгебре су вредносно неутрална и посматрају се независно од реализације на симболичком нивоу. Према овом принципу, структура било ког

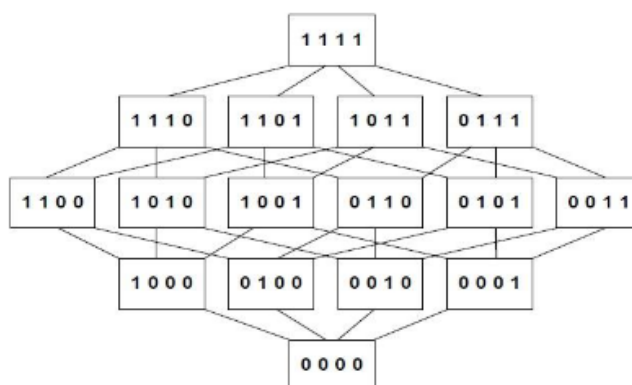
ИБА елемента се може израчунати на основу структуре његових компонената. Овај принцип је формализован кроз процедуру којом се логичке функције трансформишу у генерализоване Булове полиноме (ГБП). Процедура је дефинисана у (Radojević, 2008a) и аутоматизована у (Milošević, Petrović, Radojević & Kovačević, 2014). Као што је наведено у (Radojević, 2006), ИБА се технички заснива на генерализованим Буловим полиномима и свака Булова функција се може трансформисати у одговарајући ГБП.

Структурну функцију је могуће приказати помоћу структурног вектора (вектора структурних функција). Димензије структурног вектора су једнаке броју атома, а за елементе има вредности 1 и 0, што се односи на укљученост атома у посматрану структуру.

**Дефиниција 4 (Radojević, 2008b).** Структурни вектор  $\vec{\sigma}_\varphi$  анализираниог елемента Булове алгебре  $\varphi \in B(\Omega)$  је  $1 \times N$  димензиони вектор вредности структурних функција за све атоме:

$$\vec{\sigma}_\varphi = [\sigma_\varphi(S) | S \in P(\Omega)]. \quad \blacksquare$$

Структурни вектори свих елемената над 2 атрибута  $a$  и  $b$  су представљени помоћу Хасеовог дијаграма на следећој Слици.



**Слика 10.** Структурни вектор свих елемената над атрибутима  $a$  и  $b$  (Radojević, 2008c)

Аналогно дијаграмима са претходних Слика 8 и 9, елементи вектора одговарају структурној функцији атома и то редом:

$$a \wedge b \sim (1000), a \wedge \neg b \sim (0100), \neg a \wedge b \sim (0010), \neg a \wedge \neg b \sim (0001)$$

односно

$$a \vee b \sim (1110), a \vee \neg b \sim (1101), \neg a \vee b \sim (1011), \neg a \vee \neg b \sim (0111).$$

На основу претходне дефиниције 4 структурног вектора  $\overline{\sigma_\varphi}$ , може се истаћи да се његовим коришћењем сваки елемент Булове алгебре  $\varphi(a_1, \dots, a_n)$  може представити као производ његове структуре и вектора колоне атомских елемената.

$$\varphi(a_1, \dots, a_n) = \overline{\sigma_\varphi} \cdot \begin{bmatrix} \alpha_1(S)(a_1, \dots, a_n) \\ \vdots \\ \alpha_{2^n}(S)(a_1, \dots, a_n) \end{bmatrix} \quad (4.1)$$

Према принципу структурне функционалности, структура сложене логичке функције се може одредити на основу структуре њених компонената. Како се све логичке операције могу свести на операције конјункције, дисјункције и негације, правила за структурну трансформацију логичких функција су:

$$\sigma_{\varphi \wedge \psi}(S) = \sigma_\varphi(S) \wedge \sigma_\psi(S) \quad (4.3)$$

$$\sigma_{\varphi \vee \psi}(S) = \sigma_\varphi(S) \vee \sigma_\psi(S) \quad (4.4)$$

$$\sigma_{\neg\varphi}(S) = \neg\sigma_\varphi(S) \quad (4.5)$$

Ова правила трансформације логичких функција представљају кључне делове принципа структурне функционалности и ИБА. У следећем потпоглављу 4.4 ће бити коришћена за трансформацију генерализованог Буловог полинома.

#### 4.3.2. Вредносни ниво

Вредности примарних атрибута се уводе на вредносном нивоу ИБА, где је сваки елемент Булове алгебре вредносно реализован преко генерализованог

Буловог полинома (Radojević, 2008a). ГБП може садржати три различита оператора - класични плус (+), класични минус (-) и генерализовани производ ( $\otimes$ ). За генерализовани производ важи да је дистрибутиван у односу на плус и минус, као и да је комутативан, асоцијативан и идемпотентан на структурном нивоу. Елементи ГБП су вредности примарних атрибута  $a_i^v \in [0,1]$ ,  $i=1,\dots,n$ . Сходно томе, може се констатовати да генерализовани производ има значајан утицај на саму реализацију ГБП-а.

**Дефиниција 5 (Radojević, 2008a).** Генерализовани производ је бинарни оператор  $\otimes$  на јединичном интервалу  $[0,1]$  који је поткласа  $t$ -норме и задовољава услов ненегативности. Генерализовани производ је функција  $\otimes: [0,1]^2 \rightarrow [0,1]$  таква да за све реализације  $a_i^v \in [0,1]$  примарних атрибута  $a_i \in \Omega$ ,  $i=1,\dots,n$  важи:

$$(T1) \ a_i^v \otimes a_j^v = a_j^v \otimes a_i^v \text{ (комутативност);}$$

$$(T2) \ a_i^v \otimes (a_j^v \otimes a_k^v) = (a_i^v \otimes a_j^v) \otimes a_k^v \text{ (асоцијативност);}$$

$$(T3) \ a_i^v \otimes a_j^v \leq a_i^v \otimes a_k^v \text{ када је } a_j^v \leq a_k^v \text{ (монотоност);}$$

$$(T4) \ a_i^v \otimes 1 = a_i^v \text{ (неутрални елемент);}$$

$$(T5) \ \sum_{K \in P(\Omega \setminus S)} (-1)^{|K|} \otimes_{a_i \in S \cup K} a_i^v \geq 0 \text{ (ненегативност).} \quad \blacksquare$$

Сходно последњем услову ненегативности (T5), обезбеђује се да вредности атома  $\alpha^{\otimes}(a_1^v, \dots, a_n^v)$  припадају јединичном интервалу  $[0,1]$ , те утиче да генерализовани производ (ГП) представља подскуп скупа  $t$ -норми. Према томе, у случају два атрибута  $\Omega = \{a, b\}$ , ГП може бити било која  $t$ -норма која даје резултат већи од Лукашиевичеве  $t$ -норме, а мањи од минимума (Milosevic et al., 2018).

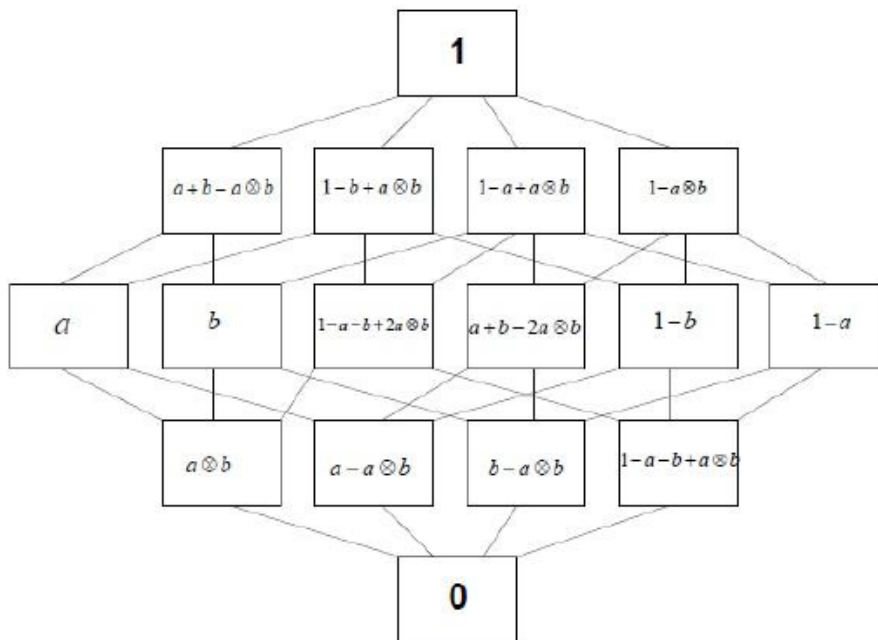
$$\max(x + y - 1, 0) \leq x \otimes y \leq \min(x, y) \quad (4.6)$$



На вредносном нивоу сваки елемент Булове алгебре  $\varphi(a_1, \dots, a_n)$  се може представити одговарајућим ГБП (Radojević, 2006). Па тако се релација  $x \Leftrightarrow p = (x \Rightarrow p) \wedge (p \Rightarrow x)$  (логичка релација еквиваленције) може трансформисати у ГБП на следећи начин:

$$\begin{aligned}
 (x \Leftrightarrow p)^\otimes &= ((x \Rightarrow p) \wedge (p \Rightarrow x))^\otimes = \\
 &= (x \Rightarrow p)^\otimes \otimes (p \Rightarrow x)^\otimes = \\
 &= (1 - x + x \otimes p) \otimes (1 - p + x \otimes p) = \\
 &= 1 - p + x \otimes p - x + x \otimes p - x \otimes x \otimes p + \\
 &+ x \otimes p - x \otimes p \otimes p + x \otimes p \otimes x \otimes p = \\
 &= 1 - p + x \otimes p - x + x \otimes p - x \otimes p + x \otimes p - x \otimes p + x \otimes p = \\
 &= 1 - p - x + 2 \cdot x \otimes p
 \end{aligned}
 \tag{4.7}$$

Пратећи ова иста правила, могуће је израчунати и остале елементе Булове алгебре над 2 атрибута  $\Omega = \{a, b\}$ . Њихови одговарајући ГБП су приказани Хасеовим дијаграмом на слици 10.



Слика 11. ГБП свих елемената Булове алгебре са два атрибута  $a$  и  $b$

Као што је већ назначено, ГБП може садржати три различита оператора – класични плус (+), класични минус (-) и генерализовани производ ( $\otimes$ ), чији избор зависи од природе атрибута и/или њихове статистичке међузависности.

*Случај 1. – стандардни минимум*

За изразито међузависне и корелисане атрибуте који су истог или сличног порекла, између њих се користи класични минимум, тј.  $x \otimes p = \min(x, p)$ . На примеру авио индустрије и резервација на лету, ово се односи на ситуацију два атрибута, од којих би један представљао број продатих резервација преко веб сајта, а други број продатих резервација преко агенције.

*Случај 2. – Лукашијевичев оператор*

Уколико су атрибути исте или сличне природе али нису међузависни, потребно је користити Лукашијевичев оператор (Lukasiewicz оператор), тј.  $x \otimes p = \max(x + p - 1, 0)$ . На примеру авио индустрије и резервација на лету, ово се односи на ситуацију два атрибута, од којих би један представљао број непродатих карата, а други цену карте.

*Случај 3. – стандардни производ*

За ситуације у којима су посматрани атрибути различите природе и статистички независни у потпуности, користи се стандардни производ, тј.  $x \otimes p = x \cdot p$ . На примеру авио индустрије и седишта на лету, ово се односи на ситуацију два атрибута, од којих би један представљао датум или време лета (или неки други временски показатељ), а други укупан број резервација у бизнис класи.

*Случај 4. – више од два атрибута*

Претходно наведени примери се односе на два атрибута, али се често у ситуацији могу наћи више од два атрибута, од који неки могу бити исте а неки различите природе. У тој ситуацији, потребно је прво груписати атрибуте исте

природе, одредити њихове вредности, а затим те вредности агрегисати као атрибуте различите природе. Узмимо пример следећих атрибута:

- (1) број продатих резервација преко веб сајта
- (2) број продатих резервација преко агенције
- (3) број резервација у бизнис класи
- (4) датум лета
- (5) сезонски показатељ

Најпре је потребно груписати атрибуте на следећи начин:

- Прва група

- (1) број продатих резервација преко веб сајта -  $a_1$
- (2) број продатих резервација преко агенције -  $a_2$
- (3) број резервација у бизнис класи -  $a_3$

- Друга група

- (4) датум лета -  $b_1$
- (5) сезонски показатељ -  $b_2$

Затим је потребно одредити вредности атрибута по групама. Како су у обе групе атрибути исте природе и међусобно зависни, користи се класични минус, тј.

$x \otimes p = \min(x, p)$ . Сходно томе имамо:

- За прву групу

$$a_1 \otimes a_2 \otimes a_3 = \min(a_1, a_2, a_3)$$

- За другу групу

$$b_1 \otimes b_2 = \min(b_1, b_2)$$

На крају, добијене вредности се агрегишу по принципу атрибута различите природе, тј.  $x \otimes p = x \cdot p$ . Целокупан приказ ове две групе атрибута је дат у наставку:

$$a_1 \otimes a_2 \otimes a_3 \otimes b_1 \otimes b_2 = \min(a_1, a_2, a_3) \cdot \min(b_1, b_2)$$

Више примера и пратећих појашњења о редоследу агрегације је дато у (Milošević, 2017). Такође, додатна појашњења у вези саме реализације ГП се могу наћи у (Dragović, Turajlić, Radojević & Petrović, 2014; Milošević et al., 2018).

#### 4.4. Правци примене

Два тренутно најуспешнија правца примене ИБА се односе на логичку агрегацију и меру сличности/различитости (Radojević, 1999; 2008c; Poledica, Milošević, Dragović, Petrović & Radojević, 2015; Milošević et al., 2018).

##### 4.4.1. Логичка агрегација

Генерално посматрано, агрегација представља поступак или процедуру за добијање збирног податка који репрезентативно представља посматрани објекат. Једну реализацију ове процедуре засновану на ИБА је предложио Радојевић (1999, 2008c) у виду логичке агрегације која узима у обзир логичку повезаност објеката.

Логичку агрегацију карактерише конзистентност и транспарентност, и састоји се из два корака:

- (1) нормализација вредности атрибута и
- (2) агрегација нормализованих вредности у једну репрезентативну вредност коришћењем логичке/псеудо-логичке функције.

Сама процедура почиње са дефинисањем агрегационог модела у виду логичког израза (Radojević, 2008c). Даље се вредности атрибута нормализују у

оквиру  $[0,1]$  интервала коришћењем одабране функције за нормализацију  $\|\cdot\|: \Omega \rightarrow [0,1]$ . Након тога, вредности се агрегишу у репрезентативну вредност коришћењем логичке/псеудо-логичке функције као оператора.

Логичка агрегација зависи од мере која је одабрана за агрегацију, односно логичке/псеудо-логичке функције која се трансформише у ГБП, као и од оператора ГП. По питању мере агрегације, логички оператор за агрегацију  $LA: [0,1]^n \rightarrow [0,1]$  у општем случају је псеудо-логичка функција, односно линеарна конвексна комбинација ГБП, која се представља у следећем облику:

$$LA_{i=1}^m(x_i) = \sum_{j=1}^m w_j \cdot \varphi_j^{\otimes}(x_1, \dots, x_m) \quad (4.8)$$

где  $\sum_{j=1}^m w_j = 1$ ,  $w_j \geq 1$ ,  $j = 1, \dots, m$  и  $\varphi_j^{\otimes}(x_1, \dots, x_m)$  су ГБП.

Уопштено говорећи, логичка агрегација није монотона функција и пружа многе могућности за моделовање различитих интеракција између атрибута (Dragović et al., 2014). У ситуацијама код којих је нарушена монотоност, логичка агрегација не задовољава све услове за формалну дефиницију оператора агрегације, али је нашла пут за различите примене. Одступањем од обавезног услова монотоности, процедура логичке агрегације нуди више опција за моделовање.

#### 4.4.2. ИБА мера сличности/различитости

Поред логичке агрегације, други најчешћи правац примене ИБА се односи на меру сличности и различитости (Rakićević et al., 2018). Мерење сличности и различитости између објеката представља основу многих алгоритама за селекцију, класификацију, предвиђање и сл. Неки од тих алгоритама, а који се често примењују, су kNN (енг. *k-Nearest Neighbour*), техника закључивања на основу случаја, TOPSIS, као и класификација на основу прототипа. У оквиру ИБА, сличност се посматра са аспекта логике, те се ИБА мера сличности и различитости користи за мерење сличности између атрибута/објеката. Као што је

проучавано у (Radojević, 2010; Poledica et al., 2015; Rakićević et al., 2018), релација ИБА еквиваленције представља меру сличности, а ИБА ексклузивна дисјункција задовољава услове да буде мера различитости. Релација ИБА еквиваленције је формулисана по узору на класичну дефиницију двоструке импликације:

$$x \Leftrightarrow p = (x \Rightarrow p) \wedge (p \Rightarrow x) \quad (4.9)$$

За примену логичке функције у оквиру ИБА, први корак је трансформација логичке релације у ГБП. Трансформациона процедура релације еквиваленције у ГБП је урађена исто као што је представљено у формули 4.7. и следећи ГБП је добијен:

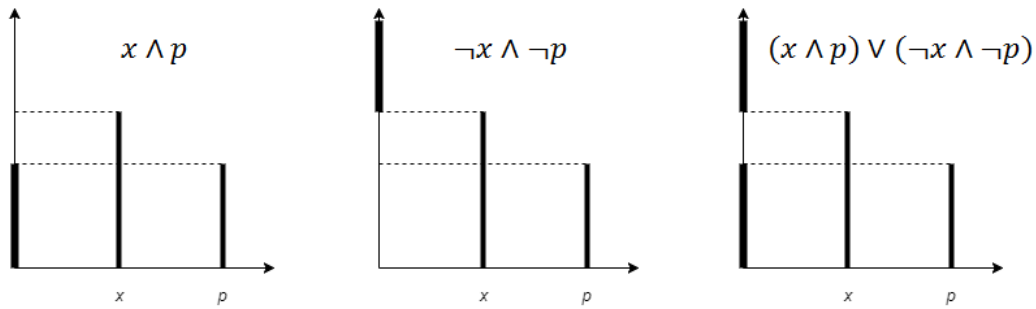
$$S_{IBA}(x, p) = (x \Leftrightarrow p)^{\otimes} = ((x \Rightarrow p) \wedge (p \Rightarrow x))^{\otimes} = 1 - p - x + 2 \cdot x \otimes p \quad (4.9)$$

Пошто је поређење једино могуће са истим атрибутима, на вредносном нивоу ГП оператор за ИБА меру сличности је класичан минимум. Могућности примене ИБА мера сличности у различитим доменама су приказани у (Poledica et al., 2015; Milošević et al., 2018).

Приликом одређивања сличности између два објекта, ИБА мера сличности сматра да су слични уколико оба или поседују или не поседују одређене атрибуте (Poledica et al., 2015). Сходно томе, меру сличности је могуће и представити на следећи начин:

$$S_{IBA}(x, p) = ((x \Rightarrow p) \wedge (p \Rightarrow x))^{\otimes} = ((x \wedge p) \vee (\neg x \wedge \neg p))^{\otimes} \quad (4.10)$$

Како се класичан минимум користи као ГП оператор, уобичајни графички приказ интерпретације ИБА мере сличности је дат на Слици 12.



Слика 12. ИБА мера сличности

Логички приступ при мерењу сличности базиран на ИБА мери сличности се састоји од два корака (Milosevic et al., 2018):

1. Припрема података (нормализација атрибута, утврђивање природе података, односно утврђивање логичких и статистичких зависности које постоје између посматраних атрибута),
2. Моделовање сличности на основу ИБА (подразумева примену једног од три модела: поређење објеката на нивоу атрибута, поређење на нивоу објекта и генерални модел).

#### Поређење објекта на нивоу атрибута

Када се пореде два објекта (нпр.  $X$  и  $Y$ ) на нивоу атрибута, а описани су кроз  $m$  атрибута, сличности се може моделовати као логичка агрегација ИБА сличности по атрибутима.

$$s_{IBA}^A(X, Y) = LA_{i=1}^m(s_{IBA}(x_i, y_i)) \quad (4.11)$$

#### Поређење на нивоу објекта

Када се пореде два објекта (нпр.  $X$  и  $Y$ ) на нивоу објекта, а описани су кроз  $m$  атрибута, најпре се ти објекти представе логичком агрегацијом атрибута, а затим се оне упоређују коришћењем ИБА мере сличности.

$$s_{IBA}^O(X, Y) = s_{IBA} \left( \underset{i=1}{\overset{m}{LA}}(x_i), \underset{i=1}{\overset{m}{LA}}(y_i) \right) \quad (4.12)$$

### Генерализовани модел

Генерализовани модел представља комбинацију претходна два описана модела. Користи се за одређивање сличности између два објекта (нпр.  $X$  и  $Y$ ) који су описани кроз  $m$  атрибута, али само у ситуацијама у којима се утврди да постоји потреба за поређење атрибута и за поређење самих објеката.

$$s_{IBA}^{gen}(X, Y) = \sum_{j=1}^m w_j \cdot s_j^O(X, Y) \quad (4.13)$$

Увођењем различитих логичких агрегација сличности / различитости (Poledica, Milošević, Dragović, Radojević & Petrović, 2013), ове мере су нашле примену за решавање различитих проблема попут кластеровања (Rakićević, Nešić & Radojević, 2013), класификације (Dobrić, Kovačević, Petrović, Radojević & Milošević, 2015) и одређивање консензуса (Poledica et al., 2013).



*„For a great player, however, the math is the easy part of the game. The hard part is reading the faces of the other players.“*

Michael Lewis<sup>2</sup>

## 5. ЗАКЉУЧИВАЊЕ НА ОСНОВУ СЛУЧАЈА

Закључивање на основу случаја (енг. *case based reasoning*) представља технику која користи базе знања и симулира људско закључивање у ситуацијама у којима је могуће користити искуство из прошлости (Leake, 1996). Ова техника је развијена у касним 80-тим годинама у Сједињеним Америчким Државама (Kolodner, 1993). Према Вотсону и Мариру (Watson & Marir, 1994), закључивање на основу случаја је у том периоду привукло пажњу истраживачке заједнице, али пошто је мало људи имало искуства у примени саме технике приликом решавања специфичних проблема, сама техника је тек касније почела да се примењује у различитим индустријама. Данас се примењује засебно (погледати нпр. Liu, Hu & Li, 2012) или у комбинацији са другим статистичким и техникама рачунарске интелигенције (погледати нпр. Faia, Pinto, Abrishambaf, Fernandes, Vale & Corchado, 2017).

На основу ове технике, нови проблем се може решити проналажењем сличних ситуација и проблема из прошлости и применом њихових решења. У зависности од врсте проблема, техника закључивања на основу случаја се може поделити на – (1) технику намењену класификацији случаја и (2) технику намењену изради предлога решења (Kolodner, 1993). У првом случају, основни циљ је класификовати проблем и донети одлуку да ли се нови проблем или ситуација могу посматрати као претходни проблеми. Одлука се доноси на основу сличности и разлика између новог и „старог“ проблема. У другом случају, основни циљ примене технике је израда предлога решења које се може применити за решавање новог проблема. Предлог решења се базира на сличним ситуацијама из прошлости. Као што је Колоднер (Kolodner, 1993) навео, ова подела може бити корисна са теоријског аспекта, али пошто у реалним ситуацијама проблеми често

---

<sup>2</sup> Из књиге: *Poker lažova* (енг. *Liar's Poker*)

имају карактеристике оба приступа, модели који се заснивају на техници закључивања на основу случаја користе комбинацију оба приступа.

Ово поглавље је организовано на следећи начин. Потпоглавље 5.1. објашњава структуру и основне елементе технике закључивања на основу случаја. У потпоглављу 5.2. се описује процес примене технике, а у потпоглављу 5.3. наводе и објашњавају четири групе фактора која могу утицати на поузданост добијеног решења.

### 5.1. Структура и основни елементи

Закључивање на основу случаја се састоји из две групе елемената – независне променљиве или атрибути (енг. *Features*) и зависна променљива (енг. *Label*). Базу знања чине решени проблеми (случајеви) који се састоје из независних и зависне променљиве. Структура базе је приказана следећом генеричком матрицом:

$$\begin{array}{ccccccc}
 x_{1,1} & x_{2,1} & x_{3,1} & \cdots & x_{i,1} & \Rightarrow & y_1 \\
 x_{1,2} & x_{2,2} & x_{3,2} & \cdots & x_{i,2} & \Rightarrow & y_2 \\
 \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
 x_{1,j} & x_{2,j} & x_{3,j} & \cdots & x_{i,j} & \Rightarrow & y_j
 \end{array} \tag{3.1}$$

Независне променљиве, односно „features“ се означавају са  $x_{i,j}, i = 1, \dots, n,$   $j = 1, \dots, m,$  где  $m$  представља број решених проблема а  $n$  број независних променљивих. Зависна променљива је означена са  $y_j, j = 1, \dots, m,$  где  $m$  представља број решених проблема (случајева) у бази знања.

## 5.2. Процес примене

Закључивање на основу случаја је процес који се састоји од 4 корака, који се у литератури још називају и 4 “R” (енг. *Retrieve, Reuse, Revise, Retain*) (Aamodt & Plaza, 1994):

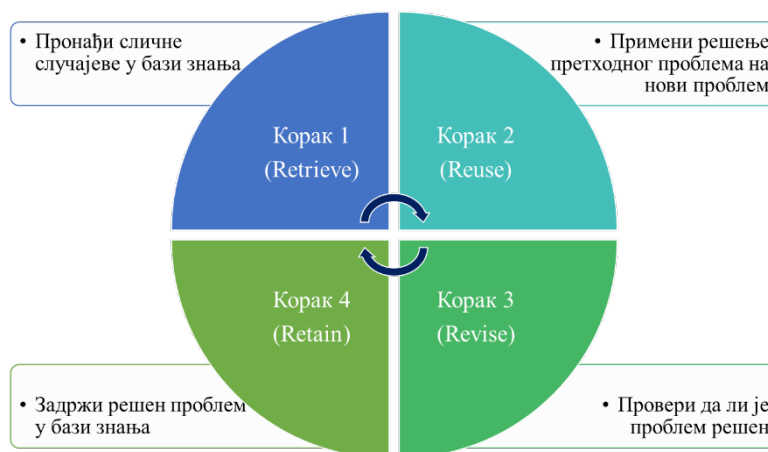
Корак 1: Пронађи сличне случајеве у бази знања

Корак 2: Примени решење претходног проблема на нови проблем

Корак 3: Провери да ли је проблем решен

Корак 4: Задржи решен проблем у бази знања

Као што је наведено, процес почиње проналажењем случаја који највише одговарају проблему, што представља први корак. У наредном кораку, начин на који је решен претходни проблем се користи за решавање новог проблема. У трећем кораку се проверава да ли је предложено решење решило проблем и да ли је вредно чувања у бази. Након тога, у четвртном кораку се у бази знања задржавају делови који могу бити употребљени за решавање будућих проблема. Процес закључивања на основу случаја је приказан на следећој Слици:



Слика 13. Процес закључивања на основу случаја

Као што је објашњено од стране (Aamodt & Plaza, 1994), случај представља проблем који може бити или сачуван у бази знања (познати случај) или може чекати да се реши (нови случај). Проблем који је сачуван у бази знања као

познати случај представља ситуацију која се десила у прошлости и за коју постоји решење које се може поново употребити за решавање будућих проблема. Проблем који чека да се реши је нови случај који је потребно решити анализом атрибута познатих случаја. Поставља се питање како одабрати познате случајеве из базе знања који су исти или слични са новим случајем и имају исте или сличне атрибуте. Као што је објашњено у (Kim & Shin, 2000), овај избор зависи од мере сличности која ће се користити за израчунавање разлика и удаљености парова атрибута. Крајњи циљ је вратити одређени број познатих случајева који су довољно слични новом случају (Aamodt, 1993). Сходно томе, постоји много мера сличности које се могу применити, и њихов избор зависи од саме природе проблема.

### 5.3. Поузданост решења

Постоје четири групе фактора према (Xu, Wang, Ma & Lin, 2010) које могу утицати на поузданост решења које се предлаже техником закључивања на основу случаја. То су:

1. Избор атрибута
2. Избор познатих случајева
3. Алгоритам
4. Људски фактори

Како се сваки случај може састојати из одређеног броја атрибута, избор оних који ће се користити у формулисању модела је веома битан. Потребно је одабрати оне атрибуте који најбоље описују проблем и у корелацији су са решењем проблема. Даље, с обзиром да познати случајеви чине базу знања, може се рећи и да поузданост модела зависи од квалитета те базе знања. Овде се пре свега мисли на које све случајеве, укључујући њихове атрибуте и решења је потребно посматрати, а који се могу изоставити из базе знања. Према већ поменутој групи аутора (Xu et al., 2010), два основна фактора која се везују за алгоритам су избор мера сличности и стратегије за предлог решења. По питању мера сличности, могуће је користити нпр. Менхетн растојање, Еуклидско растојање, као и многе друге мере сличности. Како се ова техника примењује за

израду модела за доношење одлука, не може функционисати независно од доносиоца одлука и људских фактора. Приказ ових група фактора и утицај који остварују је приказан на следећој Слици 14.



**Слика 14.** Групе фактора и њихов утицај на технику закључивања на основу случаја

Већ је истакнуто да израчунавање сличности између познатих и новог случаја представља веома битан фактор од кога у великој мери зависи поузданост решења. Сходно томе Деза и Деза (Deza & Deza, 2009) дају свеобухватан преглед мера које се могу применити за израчунавање раздаљине (удаљености) између два објеката. Многе од предложених мера се могу успешно применити у оквиру алгоритма технике закључивања на основу случаја (нпр. погледати (Chang et al., 2006, Chiu, Chiu & Hsu, 2004)).

*„If I had asked people what they wanted, they would have said faster horses.“*

Henry Ford

## 6. ПРИМЕНА ТЕХНИКЕ ЗАКЉУЧИВАЊА НА ОСНОВУ СЛУЧАЈА: СЛУЧАЈ НИСКОБУЏЕТНЕ АВИО КОМПАНИЈЕ

У оквиру овог поглавља, а пре преласка на разматрање проблема непојављивања путника на лету у оквиру знатно сложенијих околности, описана је примена технике закључивања на основу случаја на примеру нискобуџетне авио компаније и на једноставнијем проблему предвиђања укупног броја путника на лету. Циљ овог поглавља је да укаже, кроз једноставнији пример нискобуџетне компаније, који су основни кораци и на који начин је могуће применити технику закључивања на основу случаја за предвиђање броја путника на лету. Приликом саме примене технике, сумирани су и изазови и сагледана су ограничења при примени, класичног алгоритма. Такође, идеја овог примера је да се сагледа потенцијални простор за унапређење класичног алгоритма технике закључивања на основу случаја увођењем ИБА.

У првом потпоглављу је дат опис проблема и дефинисани су атрибути који су доступни и ограничени операцијама и специфичношћу ниско буџетне авио компаније, а које је могуће користити за предвиђање броја путника на лету. У другом потпоглављу је дата формулација модела за предвиђање на бази технике закључивања на основу случаја. У последњем трећем потпоглављу, приказани су резултати који су добијени коришћењем претходно формулисаног модела, указано је на неопходност проширења модела услед постојања сложенијих проблема и дата су разматрања на основу који су формулисана наредна поглавља овог рада.

## 6.1. Опис проблема

Узевши у обзир потенцијал за остваривање прихода и разне могућности за примену техника рачунарских интелигација и техника које прате последња светска кретања, евидентно је да авио индустрија већ дуже време привлачи пажњу многих стручњака из поменутих области и шире. Увидом у извештај међународне организације IATA (IATA, 2018) долази се до података који указују да постоји преко 5000 регистрованих авио компанија у свету и у фискалној 2018-ој години је процењено да је превезено нешто више од 4.3 билиона путника, што је омогућило авио компанијама оперативни профит од нешто више од 55.8 билиона долара. С друге стране, услед бројних екстерних фактора, долази до константних промена у цени горива (Lim & Hong, 2014), као и повећања осталих оперативних трошкова (Zuidberg, 2014). Ово су само једни од основних показатеља који расту из године у годину, и указују да авио компаније морају планирати оптимално коришћење капацитета (седишта), као и услугу коју пружају а односи се на сам лет. По питању дела планирања услуге, веома је битно истаћи и део који се односи на унапређење услуге, на чијим активностима авио компаније морају радити како би биле конкурентне, а приликом којих је неопходно узети у обзир поред глобалних показатеља и повратне информације које авио компаније могу добити од путника.

За планирање оптималног коришћења капацитета, авио компаније користе мноштво алата којима спроводе предвиђање броја путника које могу очекивати на лету. У зависности од пословног модела који је присутан код конкретне авио компаније, сложености података и окружења, алати за предвиђање се прилагођавају и имплементирају у оквиру постојећих операција и процеса авио компаније. Као што је евидентно и према поглављу 3. у оквиру кога је дат преглед литературе, мноштво алата за предвиђање се базира на техникама рачунарске интелигенције.

Пре преласка на разматрање знатно сложенијег проблема предвиђања непојављивања путника на лету, прво је било неопходно сагледати и реализовати технику закључивања на основу случаја једноставнијем проблему. Два циља која су постигнута овим приступом се односе на:

1. указивање потенцијалне сложености у подацима који се могу јавити код проблема већих размера, као што је предвиђање недоласка путника на лету, и
2. препознавање и потврђивање елемената у корацима реализације технике закључивања на основу случаја које је могуће унапредити и који нуде могућност примене сложенијих и свеобухватнијих концепата.

Као што је истакнуто на почетку, у овом потпоглављу је дат опис проблема и дефинисани су атрибути који су доступни и ограничени операцијама и специфичношћу нискобуџетне авио компаније, а које је могуће користити за предвиђање броја путника на лету. За потребе овог поглавља, користиле се подаци о лету нискобуџетне авио компаније на релацији Цирих – Хавана. Предвиђање путника на лету ће бити реализовано кроз историјске податке који су прикупљени за период март – септембар 2017. године. У овом временском периоду, укупно је прикупљено 57 записа о лету, што је довољно за презентацију начина примене технике закључивања на основу случаја за предвиђање. Променљиве које ће се користити за потребе формулације модела су:

- Атрибути
  - Месец
  - Викенд показатељ (указује да ли је лет био током викенда или радног дана)
  - Процент резервација на лету у оквиру пословне (бизнис) класе
  - Процент резервација на лету у оквиру економске класе
  - Процент резервација на лету које су реализоване преко веб сајта авио компаније (енг. *Web channel*)
  - Процент резервација на лету које су реализоване преко агенција као треће стране (енг. *Agency channel*)
  - Процент попуњености лета - искоришћеност капацитета (број путника који се укрцао на лет) (енг. *Flight flown rate*)
- Зависна променљива
  - Укупан број резервација на лету



Атрибути који се користе за предвиђање, а који су наведени се могу груписати у четири категорије. Прва категорија се односи на временске показатеље, као што су датум лета, време и показатељ да ли је лет био реализован преко викенда или радног дана. Друга група се односи на показатеље који разликују путника по врсти кабине/седишта које је резервисано (економска и бизнис класа). С обзиром да је истакнуто да ће се основна примена технике закључивања на основу случаја реализовати на основном проблему, подаци који ће бити коришћени садрже само резервације економских класа. Трећа категорија представља продајне канале, а последња категорија се односи на искоришћеност капацитета кроз проценат путника који се укрцао на лет (искоришћеност капацитета). Овај атрибут представља однос између броја путника који су се појавили на лету, укрцао на авион и одлетели и укупног броја резервација за тај конкретан лет.

## 6.2. Формулација модела на бази технике закључивања на основу случаја

Примена технике закључивања на основу случаја за предвиђање броја путника на лету се састоји из следећих корака:

1. Припрема базе података
2. Скалирање података
3. Израчунавање удаљености
  - a. Израчунавање удаљености између атрибута
  - b. Израчунавање удаљености између тестних случајева
4. Избор сличних случајева
5. Израчунавање грешке мерења као показатеља перформанси модела

**Припрема базе података.** Пре иницирања процеса технике закључивања на основу случаја, а који је описан у оквиру трећег поглавља овог рада, неопходно је припремити податке који постоје у бази. Ово подразумева предузимање одређених корективних активности за податке који недостају, као и предузимање

активности по питању идентификације и третирања екстремних вредности (енг. *outliers*). Додатно, за потребе предвиђања и остварења што бољих перформанси, неопходно је да база садржи податке о конкретном лету за најмање једну сезону која се може посматрати као целокупан период.

**Скалирање података.** Атрибути који се користе у оквиру модела за предвиђање су различитих редова величина, па је неопходно да се спроведе одређено скалирање самих података коришћењем неког од постојећих приступа. За потребе овог поглавља, коришћена је метода стандардизације података, као једна од метода која се веома често користи у оквиру сличних анализа и истраживања.

**Израчунавање удаљености.** Специфичности технике закључивања на основу случаја налажу да се приликом тестирања сваки тестни случај пореди са осталим случајевима у бази. Како би се одабрали најсличнији случајеви из базе, са случајем који се тестира, удаљеност између два случаја се израчунава кроз два корака. Прво се рачуна појединачна удаљеност/одстојање између атрибута тестног и случаја из базе. Након тога, одређује се укупна удаљеност између два посматрана случаја, а сам овај процес се понавља за сваки тестни случај. У погледу мера које се користе за израчунавање удаљености, а за потребе овог рада, коришћене су две мере - еуклидско и менхетн растојање.

**Избор сличних случајева.** На основу претходно утврђених граница за избор сличних случајева, израчунава се просек вредности зависне променљиве свих прихваћених случајева који су раније решени и који се налазе у бази. Ово значи да су сви решени случајеви са израчунатим мерама сличности које су једнаке или мање од утврђених граница, одабрани и њихове зависне променљиве се третирају даље у оквиру процеса предвиђања.

**Израчунавање грешке мерења као показатеља перформанси модела.** За потребе израчунавања перформанси модела за предвиђање, користи се средња апсолутна процентуална грешка (енг. *Mean absolute percentage error*). Овај тип грешке је детаљније објашњен у оквиру потпоглавља 7.3.4. (корак 9).

### 6.3. Резултати и разматрање коришћеног модела

Пратећи претходно описане кораке, спроведена је техника закључивања на основу случајева, урађена је валидација. С обзиром да је за потребе овог примера коришћен мали узорак, коришћена је унакрсна валидација са једним издвојеним елементом, с обзиром да овај вид валидације у оваквим ситуацијама даје нарeпрезентативније резултате. Након тога, израчунате су грешке мерења као показатеља перформанси модела и приказане у следећој Табели 1.

**Табела 1.** Показатељи перформанси модела за предвиђање броја резервација на лету

Модел	MAPE
Модел за предвиђање базиран на Еуклидској мери за израчунавање растојања	12.082%
Модел за предвиђање базиран на Менхетн мери за израчунавање растојања	11.823%

Модел који је користио Менхетн меру за израчунавање растојања између случајева је остварио боље перформансе предвиђања према подацима из Табеле 1 (MAPE од 11.823%). Међутим, и други модел, у оквиру кога је коришћена Еуклидска мера за израчунавање растојања, је такође остварио задовољавајуће перформансе предвиђања, али са мало слабијим MAPE показатељем који је 12.082%. Узимајући у обзир природу проблема, нагласак на једноставности у коришћеним подацима, чињеницу да техника закључивања на основу случаја није модификована и да не укључује додатне елементе, може се констатовати да су оба модела остварила перформансе са задовољавајућим грешкама у предвиђању. Поред тога, евидентно је и да постоји доста простора за потенцијална унапређења. Такође, евидентно је да није могуће остварити поређење овог модела са сличним моделима који су доступни и описани кроз литературу, а неки од кључних разлога су:

- везе између атрибута нису детаљније проучене,
- није одређен значај атрибута,
- за одређивање сличности/различитости коришћене су мере за просторно распоређене атрибуте

- није претпостављена, па самим тим ни моделована никаква интеракција између атрибута.

Први циљ је био указивање потенцијалне сложености у подацима који се могу јавити код проблема већих размера, као што је предвиђање недоласка путника на лету. На једноставном примеру је показано које све варијације у атрибутима могу постојати, као што су нпр. присутност више од једне класе седишта, узимање у обзир додатних временских показатеља (нпр. показатељ који одређује да ли је у току сезона или је лет реализован ван сезоне), постојање више канала преко којих путник може резервисати место и купити карту, као и коришћење знатно сложенијих показатеља капацитета попуњености авиона. Како ови атрибути представљају основне елементе и одреднице у оквиру једног записа о лету, очекивано је да се исти или већина њих (у зависности и од расположивости података), користе и за предвиђање путника који се неће појавити на лету.

Други циљ се односио на препознавање и потврђивање елемената у корацима реализације технике закључивања на основу случаја, које је могуће унапредити и који нуде могућност примене сложенијих и свеобухватнијих концепата. Након проласка кроз основне кораке примене технике закључивања на основу случаја, евидентно је да постоји простор за коришћење сложенијих приступа приликом скалирања података, одређивања сличности између случаја/атрибута (коришћење других мера за израчунавање удаљености атрибута), агрегације података, начина дефинисања граница прихватања, као и коришћење релевантнијих мера грешака који ће служити као показатељи перформанси модела у погледу предвиђања.

*„If you are working on something exciting that you really care about, you don't have to be pushed. The vision pulls you.“*

Steven Paul Jobs

## 7. СИСТЕМ ЗА ПРЕДВИЂАЊЕ „NO-SHOW“ ПУТНИКА У АВИО ИНДУСТРИЈИ

У оквиру овог поглавља, проблем непојављивања путника на лету је описан и дат је предлог решења које се односи на предвиђање броја ових путника (Vojtek, Petrović & Milošević, Forthcoming 2021). За потребе предвиђања неопходно је користити одређене атрибуте који описују лет, за који се спроводи предвиђање. Узимајући у обзир врсту проблема, економски аспект и доступност података, предложени систем за предвиђање је конципиран кроз комбинацију две технике – закључивање на основу случаја и ИБА. У ранијем поглављу 4. су истакнути домени и области из теорије и праксе у оквиру којих је ИБА концепт и одређени елементи разматрани, предложени и примењени. У погледу предложеног комбиновања ИБА са техником закључивања на основу случаја, слична примена је истакнута и анализирана у (Poledica, Vogicevic-Arsic & Petrovic, 2010) на проблему тржишта некретнина и дала је обећавајуће резултате.

У првом потпоглављу је дат опис проблема и истакнуте су групе атрибута који се најчешће користе за предвиђање броја путника који се неће појавити на лету. У другом потпоглављу је дат општи пример овог проблема са нагласком на економски аспект, како би се истакла важност самог проблема, односно зашто је неопходно да авио компаније што прецизније покушају да предвиде број пуника који се неће појавити на лету. У последњем трећем потпоглављу, систем за предвиђање броја путника који се неће појавити на лету је предложен и његови елементи су детаљно описани.

## 7.1. Опис проблема

Предвиђање путника који се неће појавити на лету се као процес најчешће заснива на историјским подацима о посматраном лету. Атрибути који се најчешће користе за предвиђање се могу груписати у три категорије:

- Време
- Искоришћеност капацитета, односно инвентара (седиште)
- Резервације

У погледу временских атрибута, најчешће се користи време полетања (дан, месец или чак тачно време) и показатељ који се односи на викенд и радни дан. Овај последњи показатељ је веома битан јер може постојати значајна разлика у броју путника који се неће појавити на летовима током викенда и радних дана. Такође, препоручује се да база података која се користи у себи садржи информације о летовима барем за једну целу календарску годину. На овај начин могуће је обухватити сезонске трендове и укључити их у процес предвиђања, јер се проценат “no-show“ путника током и ван сезоне може разликовати у великој мери. Учесталост лета (да ли се ради о дневном лету, једном недељно или месечно) утиче на укупан број записа о лету који ће се наћи у бази података, што такође може значајно утицати на резултате предвиђања.

Показатељ који се односи на искоришћеност капацитета (седишта) лета на одређени дан је проценат попуњености. Овај показатељ представља однос између броја резервација и укупног броја седишта који је доступан у авиону. Процент попуњености указује да је нпр. 90% седишта резервисано.

Трећа група показатеља се односи на резервационе канале, односно на глобалне дистрибутивне системе (енг. *GDS - global distribution systems*). Показатељи могу указивати на број резервација направљених кроз најзаступљенији домаћи ГДС или кроз најзаступљенији интернационални ГДС.

## 7.2. Економски аспект проблема

Како би се указало на значај представљеног проблема, пример стандардног Ербас А320 (енг. AIRBUS A320) са капацитетом од 174 седишта је представљен са економског аспекта на релацији Београд-Амстердам. У питању је недељни лет. Посматра се ситуација када је 97% седишта у авиону резервисано три дана пре самог лета, односно 169 седишта је резервисано и само 5 је остало слободно. На авио компанији је да донесе одлуку да ли ће дозволити „overbooking“ (који је објашњен у једном од претходних потпоглавља овог рада - погледати потпоглавље 2.1.) и за колико седишта. Ова одлука се доноси у већини случајева на основу потенцијалног прихода који се може остварити и на основу броја путника који се нису појавили на конкретном лету у прошлости. У овом примеру, авио компанија је донела одлуку да дозволи „overbooking“ од 10 седишта. Два дана пре самог лета, сва седишта су распродата и само су остала слободна 10 „overbooking“ седишта. За мерење прихода по путнику користи се показатељ приход за једно седиште по миљи пута - RASM (енг. *revenue per available seat mile*). Овај показатељ се користи као стандард на нивоу целе авио индустрије и израчунава се за сваку авио компанију посебно. На овом примеру ће се користити RASM авио компаније Луфтханса (Lufthansa), који је према компанијском годишњем извештају за 2016-ту годину износио 7.8 еуро центи (euro cents)<sup>3</sup>. У циљу сагледавања економског аспекта описаног проблема, три случаја ће бити посматрана, и то:

(1) Одређени проценат „overbooking“ седишта продат и више путника се није појавило на лету

(2) Одређени проценат „overbooking“ седишта продат и мање путника се није појавило на лету

(3) „overbooking“ седишта нису продата и неколико путника се није појавило на лету

---

<sup>3</sup> Луфтханса (Lufthansa) Годишњи извештај 2016. <https://investor-relations.lufthansagroup.com/en/finanzberichte/annual-report/2016.html> Последњи пут прегледано 02. Фебруар 2018. године

Случај 1. „overbooking“ мањи од „no-show“ путника

У овом случају 3 додатна „overbooking“ седишта су продата пре посматраног недељног лета на релацији Београд-Амстердам, што представља 30% од укупно дозвољеног броја „overbooking“ седишта. Како је укупна удаљеност од Београда до Амстердама 903.83 миље<sup>4</sup>, додатни приход који авио компанија остварује износи:

$$3 \text{ путника} * 903.83 * 0.078 \text{ €} = 211.49 \text{ €}$$

Напомена: удаљеност је изражена у миљама јер је то најчешће коришћена мерна јединица која се користи у авио индустрији, а и RASM показатељ се односи на приход по миљи пређеног пута.

Како је у питању недељни лет, потенцијални додатни приход на годишњем нивоу износи 10,997.48 € (52 \* 211.49 €). Поред продата 3 додатна „overbooking“ седишта, у овом случају се на самом лету није појавило 5 путника, односно било је 5 путника са „no-show“ статусом у систему након што је авион полетео. То значи да је авио компанија била ускраћена за додатни профит за 2 седишта. Такође, овај случај није проузроковао додатне трошкове по авио компанију.

Случај 2. „overbooking“ већи од „no-show“ путника

У другом случају поново су 3 додатна „overbooking“ седишта продата пре посматраног недељног лета на релацији Београд-Амстердам, што представља 30% од укупно дозвољеног броја „overbooking“ седишта. Пошто је у питању иста удаљеност, као што је и израчунато за случај 1, додатни приход који авио компанија остварује износи 211.49 €.

---

<sup>4</sup> Калкулатор миља за летове. [https://www.worldatlas.com/travelaids/flight\\_distance.htm](https://www.worldatlas.com/travelaids/flight_distance.htm)  
Последњи пут прегледано 23. Април 2018. године



Поред продата 3 додатна „overbooking“ седишта, у овом случају се на самом лету није појавило 2 путника, односно било је 2 путника са „no-show“ статусом у систему након што је авион полетео. То значи да је авио компанија на том лету имала једног путника више од максималног капацитета авиона, те је морала да путника преусмери на други лет на истој релацији али који је у власништву друге авио компаније. Како би се решила настала ситуација, и како би се смањила шанса да дође до незадовољства путника, пракса је да у том случају авио компанија путнику обезбеди први следећи лет вишом класом, односно са већом тарифном категоријом. Углавном, путник добија седиште у бизнис класи са одређеним додатним погодностима. Трошкови који у овој ситуацији настају се углавном решавају између две авио компаније, али могу, поређења ради, износити и 800 €, па чак и 1,200 €. Када се узме у обзир додатни приход који је авио компанија остварила продајом „overbooking“ седишта и одузму ови трошкови, јасно је да је авио компанија у губитку.

### Случај 3. Без продатих „overbooking“ седишта

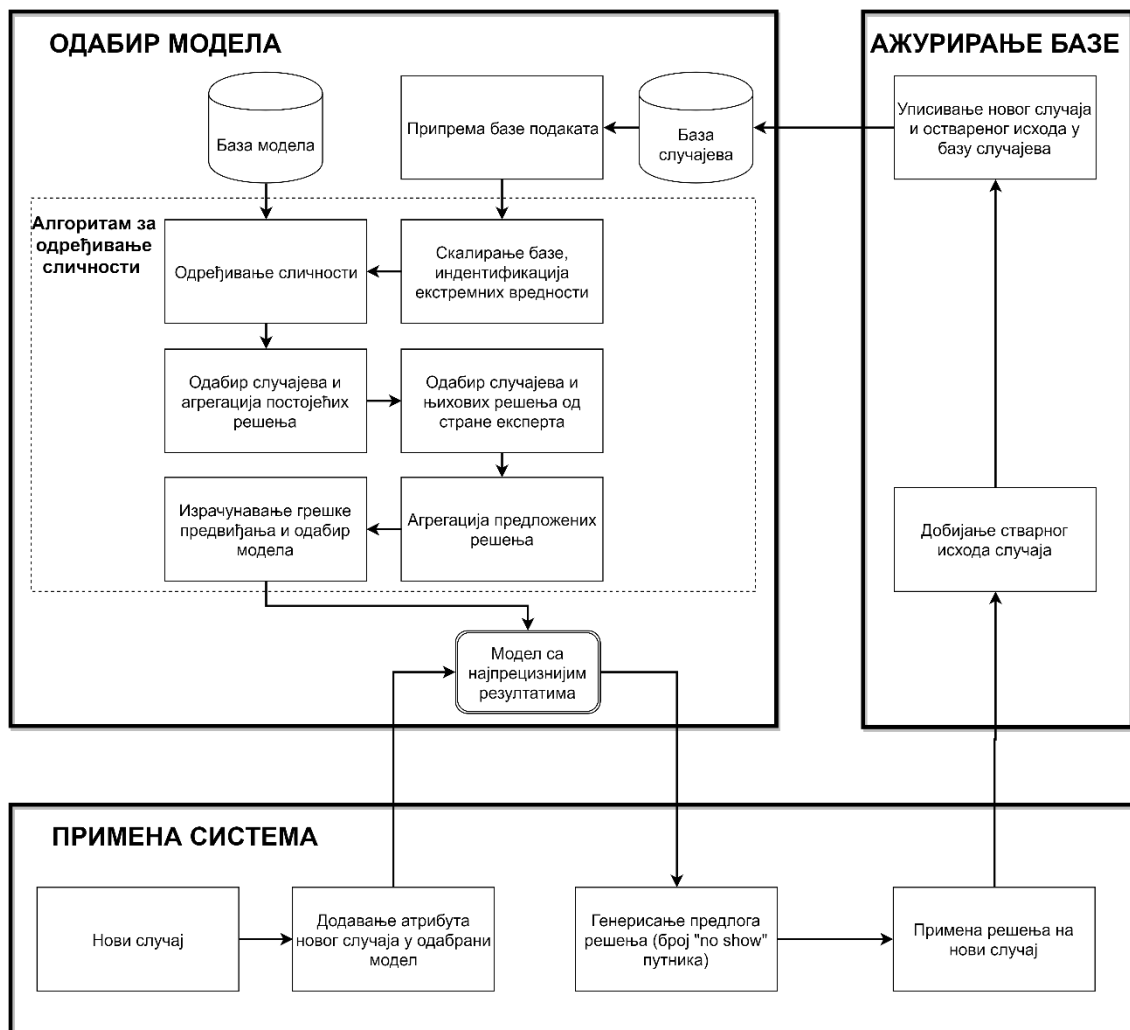
У последњем случају није дошло до продаје додатних „overbooking“ седишта, те компанија није остварила никакав додатни приход. Сходно томе, не долази до изражаја ни податак колико се путника није појавило на лету. Једино што је битно је да колико год путника је било „no-show“, то представља изгубљен додатни профит за авио компанију.

Сва три претходно описана случаја представљају свакодневне ситуације са којима се авио компаније сусрећу на својим летовима са високом потражњом. Најчешће су ови летови минимум два пута недељно, што значи да очекивани годишњи додатни профит може бити 20,000.00 € по лету. Оно што је евидентно је да грешке приликом предвиђања могу скупо коштати авио компанију, па није лако одредити „overbooking“ лимит на одређеном лету. Узевши све ово у обзир, проблем предвиђања путника који се неће појавити на лету је веома битан за авио компаније које имају више летова са високом потражњом код којих су сва или већина седишта резервисана. Што већа прецизност у предвиђању „no-show“ путника помаже авио компанији приликом одређивања „overbooking“ лимита, и

смањује шансу за остваривање случаја 2 у коме је више продатих "overbooking" седишта него "no-show" путника.

### 7.3. Предлог система за предвиђање „no-show“ путника

Проблем предвиђања „no-show“ путника је описан у потпоглављу 7.1. а на његову значајност са економског аспекта је указано кроз три различита случаја у потпоглављу 7.2. Узевши све у обзир, у овом потпоглављу предложен је систем за предвиђање „no-show“ путника (Vojtek, Petrović & Milošević, Forthcoming 2021) који се састоји из три главне компоненте - (1) Избор модела, (2) Примена система и (3) Ажурирање базе знања. Систем за предвиђање „no-show“ путника је приказан на следећој слици.



Слика 15. Структура система за предвиђање „no-show“ путника

Као што је приказано на претходној слици, предложени систем може функционисати динамички. Другим речима, компонента примена система се активира сваки пут када се нови случај заједно са решењем унесе у базу знања.

### 7.3.1. Компонента 1 - Избор модела

Избор модела је прва компонента којом се од више модела утврђује који модел за предвиђање даје најбоље резултате за конкретну базу знања расположиву у том тренутку. У суштини, варијације модела се односе на могуће алгоритме технике закључивања на основу случаја. Па тако нпр. један модел може користити Менхетн растојање за одређивање сличности између два случаја, а други Еуклидско растојање. Такође, први модел може имати рестриктивнију

границу за прихватање сличних случајева, док други модел мање рестриктивну. О овоме ће бити више речи у последњем делу овог поглавља, у коме је описан процес примене система за предвиђање „no-show“ путника. Узевши све у обзир, ова прва компонента се састоји из два елемента - базе случајева и алгоритма за одређивање сличности између случајева.

Алгоритам за одређивање сличности између случајева се састоји из следећих корака:

1. Скалирање података
2. Одређивање сличности атрибута
3. Агрегација вредности атрибута на нивоу случаја
4. Агрегација вредности одабраних случаја

Ови кораци се реализују различито за сваки модел из групе доступних модела. Сви модели се прво оцењују према резултатима предвиђања коришћењем одабраних мера грешке, као што су средња апсолутна процентуална грешка и средњи апсолутни трошак жаљења (енг. *Mean absolute cost regret*). Модел који даје најпрецизније, односно резултате са најмањом грешком се узима за решавање новог проблема и користи се у следећој компоненти.

### 7.3.2. Компонента 2 - Примена система

Примена система за предвиђање „no-show“ путника у ствари представља симулацију процеса за доношење одлука. У оквиру ове компоненте, модел који је одабран из претходне компоненте даје предлог у виду броја „no-show“ путника за нови случај. Након тога, сам предлог је потребно да се анализира од стране одговорне особе која доноси одлуку да ли ће се дозволити „overbooking“ према предложеном броју „no-show“ путника, или ће се чак тај број кориговати.

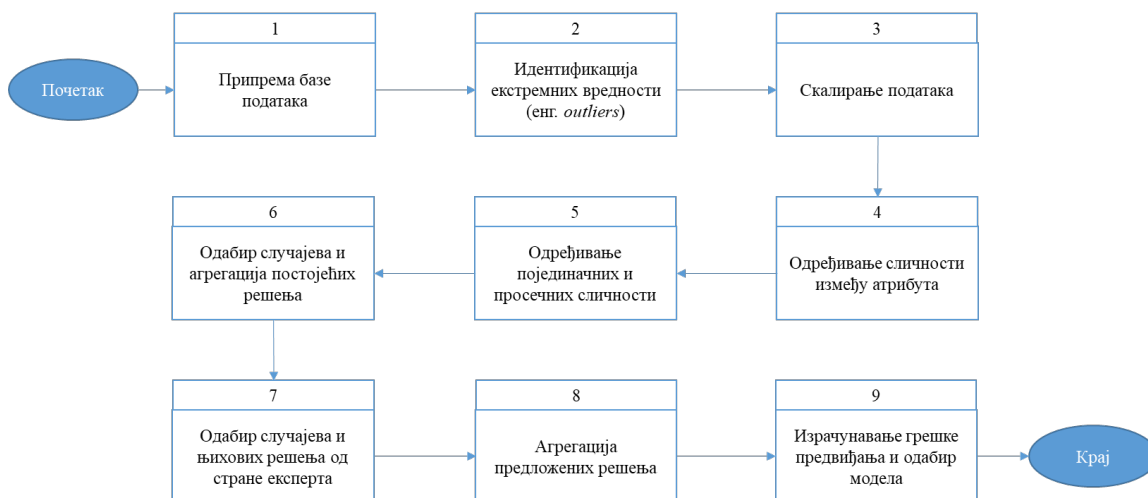
### 7.3.3. Компонента 3 - Ажурирање базе

База са познатим случајевима, укључујући атрибуте и решења, се ажурира након што се за нови случај, који је био анализиран у оквиру претходне компоненте 2, добије решење. На примеру „no-show“ путника у авио индустрији,

решење се односи на стварни број путника који се није појавио на лету који је посматран као нови случај у оквиру компоненте 2. Поред решења, у базу се уносе и атрибути лета.

### 7.3.4. Процес избора модела за предвиђање „no-show“ путника

Процес избора модела за предвиђање „no-show“ путника је дат на следећој Слици.



Слика 16. Процес избора модела за предвиђање „no-show“ путника

#### 1. Припрема базе података

За потребе предложеног система, база података се мора креирати на нивоу лета. Неопходно је да база садржи податке о лету за најмање годину дана. На овај начин се осигурава да се у бази обухвате подаци из два временска оквира, односно у току и ван сезоне. Један запис представља податке о једном лету за одређени дан (нпр. Београд - Амстердам, петак, 25. мај 2018). Припрема и уређивање саме базе се састоји из следећа два корака:

- Корак 1. Записе у оквиру којих је број „no-show“ путника једнак или мањи од нуле је потребно избацити из базе.
- Корак 2. Записе у оквиру којих је остварен низак проценат попуњености лета (нпр. испод 90%) је такође потребно избацити из базе.

Основна сврха предложеног система је предвиђање путника који се неће појавити на лету, па сходно томе, неопходно је узети у обзир само случајеве (записе) код којих је забележен одређени број „no-show“ путника. Случајеви у оквиру којих је број „no-show“ путника мањи од нуле, у ствари представљају ситуације у којима се појавило више путника него што је капацитет авиона (један од разлога може бити и тај да је број „no-show“ путника погрешно процењен и да је у систему дозвољен већи „overbooking“ од очекиваног). Укључивање ових случајева би знатно утицало на резултате система, а могле би се десити и ситуације где би предлог „no-show“ путника био негативан. У погледу процента попуњености, сама вредност од 90% је сугерисана од стране аутора у овом раду и посматраће се као хиперпараметар који је могуће подесити пре самог покретања система, а на основу специфичних потреба авио компаније и других фактора. Могуће је наћи и оптималну вредност која се може разликовати од случаја до случаја, али то би захтевало детаљније анализе и прикупљање додатних података. У оквиру техника рачунарске интелигенције, хиперпараметар представља параметер чија се вредност подешава пре покретања процеса учења.

## 2. Идентификација екстремних вредности (енг. *outliers*)

За идентификацију и третирање необичних и екстремних вредности користи се позната трансформациона процедура која је добила назив по Чарлсу Винзору (енг. *Winsorizing*, Charles P. Winsor, 1895–1951). Након што се екстремне вредности идентификују, њихове вредности се замењују са најближим неекстремним вредностима. Овај приступ се често сматра прикладнији у односу на стандардно искључивање ових вредности, јер би то значило искључивање многих случајева из базе, а самим тим би и компонента база знања била слабија.

## 3. Скалирање података

Пошто су атрибути који се користе у оквиру алгоритма за одређивање сличности случајева различитих вредности и користе различите скале, скалирање података је потребно спровести пре иницирања самог алгоритма. Веома је тешко предложити приступ за скалирање, чак и у ситуацијама у којима су многи фактори познати и под контролом. Сходно томе, у циљу проналажења алгоритма

који ће дати најбоље резултате, односно проналажење најбоље комбинације приступа и техника, користиће се нормализација (*min-max*) и конвенционална стандардизација за скалирање података. Предложени систем је могуће унапредити додатно, коришћењем и других нормализационих функција у оквиру овог корака, што може бити предлог за неке наредне верзије система.

#### 4. Одређивање сличности између атрибута

Овај корак практично представља почетак тестирања, у оквиру кога се сви случајеви из базе користе, а сваки случај се тестира у односу на друге случајеве. За израчунавање сличности између атрибута, са  $x$  је означен „no-show“ познатих случаја а са  $p$  „no-show“ случаја који се тестира. Другим речима, случај који се тестира у ствари представља симулацију новог случаја. За одређивање сличности између случајева, следеће мере се користе:

- a. Еуклидско растојање (енг. *Euclidean distance - ED*)

$$D_{ED}(X, P) = \sqrt{\sum_{i=1}^n (x_{i,j} - p_i)^2} \quad (5.1)$$

- b. Менхетн растојање (енг. *Manhattan distance - MD*)

$$D_{MD}(X, P) = \sum_{i=1}^n |x_{i,j} - p_i| \quad (5.2)$$

- c. ИБА мера сличности (енг. *IBA similarity measure - IBA*)

$$S_{IBA}(x, p) = 1 - p - x + 2 \cdot x \otimes p = 1 - p - x + 2 \cdot \min(x, p) \quad (5.3)$$

Прве две функције представљају стандардне мере за израчунавање удаљености између два посматрана објекта. Трећа која се базира на интерполативној Буловој алгебри, одређује сличност између објеката са становишта логике. Као што је и у поглављу 4. истакнуто, ГП оператор за ИБА меру сличности је класичан минимум, јер је поређење на вредносном нивоу једино и могуће са истим атрибутима. Неопходно је истаћи да сличност у ствари предстаља негацију различитости/удаљености.

ИБА функција за израчунавање сличности је дата у наставку:

$$D_{IBA}(x, p) = 1 - S_{IBA}(x, p) = p + x - 2 \cdot p \otimes x \quad (5.4)$$

Претходно наведене функције се могу користити појединачно, у зависности од саме ситуације и природе проблема. Примена ИБА мере сличности у оквиру алгоритма за закључивање на основу случаја, пружа нову перспективу у оквир самог закључивања (Poledica, Bogicevic-Arsic & Petrovic, 2010). Овај приступ се може посматрати као алтернатива много стандарднијем приступу у оквиру технике закључивања на основу случаја који се базира на коришћењу конвенционалних мера за израчунавање удаљености/сличности између два објекта.

##### 5. Одређивање појединачних и агрегисаних сличности

У оквиру овог корака, прво је потребно израчунати сличност између тестног случаја и свих осталих случајева. Тако се први случај посматра као тестни у односу на све друге, након тога се други случај посматра као тестни у односу на друге и тако до последњег случаја. Појединачне сличности се добијају агрегацијом сличности између атрибута, чије су вредности добијене у оквиру претходног корака. Ова функција агрегације може бити више традиционална или се може заснивати на логици. На пример, у случају ИБА, може се користити логичка агрегација:

$$P = S_{IBA}(X, P) = LA(S_{IBA}(x, p)) \quad (5.5)$$

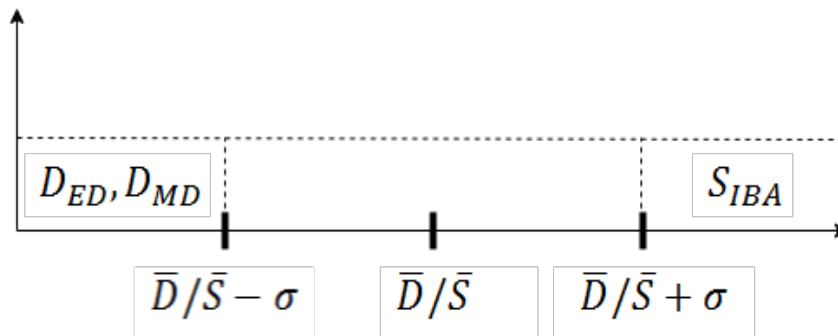
Након што су појединачне сличности одређене, потребно је израчунати просечну сличност и стандардно одступање, односно стандардну девијацију. Ово је неопходно како би се формирала граница прихватања случаја и користиће се у следећем кораку.

##### 6. Избор случајева и агрегација постојећих решења

На основу границе прихватања, која се формира на основу просечне сличности/удаљености  $(\bar{D} / \bar{S})$  и стандардне девијације израчунатих



сличности/удаљености ( $\sigma_{D/S}$ ) из претходног корака, прихватљиви случајеви су одабрани. Финалне вредности сличности су означене са  $D_{ED}^j, D_{MD}^j, S_{IBA}^j$  где је  $j = 1, \dots, m$  а  $m$  представља број постојећих случаја. У зависности од тога да ли је израчунат ниво сличност или удаљеност између постојећих случајева, два нивоа/границе прихватања су посматрана, као што је приказано на следећој слици 14.



Слика 17. Ниво/граница прихватања постојећих случајева

Када се користе мере за израчунавање удаљености, два објекта који имају исту вредност ће имати удаљеност једнаку нули (0). Сходно томе, граница прихватања за прве две функције ( $D_{ED}$  и  $D_{MD}$ ) је дефинисана као  $D_{ED}, D_{MD} \leq \bar{D} - \sigma$ . На основу овога, сви постојећи случајеви код којих је удаљеност мања или једнака са средњом вредношћу удаљености умањеној за стандардну девијацију су прихваћени и даље су процесуирани кроз алгоритам. У ситуацији када се користи ИБА мера сличности, два иста објекта имају вредност јединици (1). Граница прихватања је у овом случају  $S_{IBA} \geq \bar{S} + \sigma$ . Након што су случајеви одабрани из базе, потребно је урадити агрегацију њихових решења (процента “no-show“ путника). За агрегацију ће се користити два приступа. Први се односи на ИБА логички приступ реализован уз помоћ *min* оператора, а други на израчунавање просечне вредности.

#### 7. Избор случајева и њихових решења од стране експерта

Као што је и раније било наведено, поред предлога решења који ће бити генерисан од стране описаног алгоритма, експертско знање је такође укључено у процес предвиђања. У оквиру овог корака, експерт бира случајеве из базе

познатих случајева и њихових решења који су по мишљењу експерта слични случају који се посматра. Неопходно је да минимални број случајева, који је потребно да експерт одабере, буде једнак броју случајева који су одабрани од стране алгоритма у претходном кораку. На овај начин могуће је применити у наредном кораку различите функције за агрегацију података.

#### 8. Агрегација предложених решења

Након што су предлози решења добијени од алгоритма и експерта, иста се агрегишу коришћењем одабраног оператора. За разлику од једноставног рачунања просечне вредности, предложено је коришћење ИБА логичке агрегације. С обзиром на природу проблема (предложена решења су исте/сличне природе), логичка агрегација је спроведена коришћењем  $\min$  оператора. Како би се моделовале могуће нелинеарне зависности између података, могуће је применити у овом кораку и много комплекснију функцију логичке агрегације. Овај елемент може представљати надоградњу у некој наредној итерацији предложеног система за предвиђање.

#### 9. Израчунавање грешке предвиђања и избор модела

Коришћењем различитих мера за одређивање удаљености/сличности између атрибута, укључивањем и искључивањем препорука експерта, као и употребом различитих оператора агрегације, више различитих модела је формулисано. У последњем кораку избора модела, прецизност предвиђања сваког модела се утврђује. За оцену перформанси предвиђања, различити критеријуми се могу користити, као што су грешке предвиђања, брзина генерисања предлога решења, могућност интерпретације резултата и други. Од многих, може се рећи да су грешке предвиђања најбитније. Сходно томе, две мере у виду грешака предвиђања ће бити коришћене за оцену перформанси модела - средња апсолутна процентуална грешка (енг. *Mean absolute percentage error*) и - средњи апсолутни трошак жаљења (енг. *Mean absolute cost regret*).

Како је главни циљ модела да генерише проценат “no-show“ путника, процентуална грешка је веома битна са аспекта авио компаније. Сходно томе,

акценат је на средњој апсолутној процентуалној грешци, која је и као показатељ највише коришћена за тумачење добијених резултата у наредном поглављу.

Поред овога, у погледу трошкова је веома битно раздвојити и третирати различито ситуације код којих је потцењено и прецењено предвиђање. Потцењено предвиђање се односи на случај када је предвиђен број “no-show“ путника мањи од стварног броја путника који се нису појавили. Прецењено предвиђање се односи на случај када је предвиђен број “no-show“ путника већи од стварног броја путника који се нису појавили. У погледу трошкова, потцењено предвиђање може мање коштати авио компанију од ситуације у којој је дошло до прецењивања очекиваног броја “no-show“ путника. У циљу појашњења, дат је следећи пример. Авио компанија предвиђа да се 10 путника неће појавити на лету и сходно томе омогућава „overbooking“ за додатних 10 седишта, а како се ради о лету високе потражње, свих 10 седишта бивају продата. На самом лету се не појаве 8 путника, па сходно томе, авио компанија има 2 путника више од максималног капацитета авиона, за које ће морати да плаћа додатни трошак како би им обезбедила путовање. Најчешће у оваквим ситуацијама, авио компанија резервише скупља седишта код других авио компанија, или мора примити путника на неком другом свом лету и платити путнику додатне компензације. Ово изискује много веће трошкове од трошкова који могу настати приликом потцењивања броја „no-show“ путника, где настаје само трошак посматран кроз неискоришћене прилике. На основу свега овога, средњи апсолутни трошак жаљења је уведен како би смањιο у одређеној мери значај модела који генеришу више прецењених предвиђања. Ова грешка је представљена следећом формулом:

$$MACR = \frac{1}{n} \sum_{i=1}^n cr_i \quad (5.6)$$

$$cr_i = \begin{cases} |x_i - p_i| \cdot w_1, & \text{for } x_i - p_i \geq 0 \\ |x_i - p_i| \cdot w_2, & \text{for } x_i - p_i < 0 \end{cases}$$

Тежински коефицијенти  $w_1$  и  $w_2$  представљају однос између потцењеног и прецењеног предвиђања, и на свакој авио компанији је да одреди ове параметре према свом моделу пословања и стратегији. Након израчунавања грешака

предвиђања, модели се рангирају и модел са најбољим резултатима предвиђања је одабран и користи се у оквиру наредне компоненте примене система.

*“People don't buy what you do; they buy why you do it. And what you do simply proves what you believe”*

Simon Sinek

## 8. ПРИМЕНА ПРЕДЛОЖЕНОГ СИСТЕМА

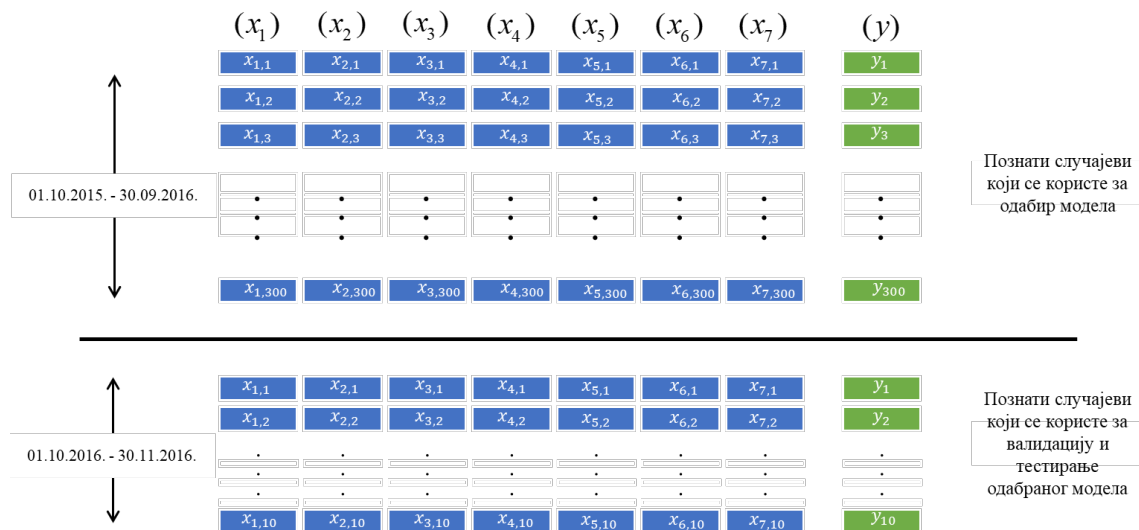
У овом поглављу је представљена примена предложеног система на реалном примеру у оквиру авио индустрије. Предложени систем је тестиран коришћењем података о лету на релацији Београд - Амстердам за период од 1.10.2015. до 30.11.2016. године. Пре саме валидације рада и резултата предложеног система, у потпоглављу 8.1 је дат дескриптивни и графички приказ података који су се користити за тестирање предложеног решења. Након тога, у потпоглављу 8.2 је урађено претпроцесирање података, а у потпоглављу 8.3 је спроведен избор модела уз помоћ претходно описаног алгоритма за одређивање сличности. За добијене резултати, у виду предвиђачких перформанси модела, урађен је тест значајности у потпоглављу 8.4. Након тога, одабрани модели су валидирани на додатном мањем скупу података у потпоглављу 8.5. На крају, у потпоглављу 8.6. добијени резултати су упоређени са резултатима предвиђања неуронским мрежама.

### 8.1. Анализа иницијалног скупа података

Информације о резервацијама путника су прикупљене од стране авио компаније, односно извучене из резервационог система за лет Београд - Амстердам за период од 1.10.2015. до 30.11.2016. године. У оквиру самог резервационог система авио компаније, један лет је описан помоћу низа карактеристика. У конкретном случају, у питању је дневни лет за који су анализирани следећи подаци:

- Атрибути
  - Сезонски показатељ ( $x_1$ )
  - Месец полетања ( $x_2$ )
  - Викенд показатељ ( $x_3$ )
  - Процент попуњености лета ( $x_4$ )
  - Број резервација у оквиру домаћег ГДС-а ( $x_5$ )
  - Број резервација у оквиру међународног ГДС-а ( $x_6$ )
  - Број резервација у свим осталим ГДС-овима ( $x_7$ )
- Зависна променљива
  - Процент „no-show“ путника ( $y$ )

Прва три атрибута ( $x_1, x_2, x_3$ ) се односе на временску димензију посматраног лета. Четврти атрибут ( $x_4$ ) указује на капацитет авиона и показује попуњеност расположивог броја седишта. Последња три атрибута ( $x_5, x_6, x_7$ ) представљају три различита канала продаје, односно различите системе кроз које су направљене резервације. Подела података према временској димензији је приказана на следећој Слици 18.



**Слика 18.** Подела података према временској димензији коришћених за примену предложеног система

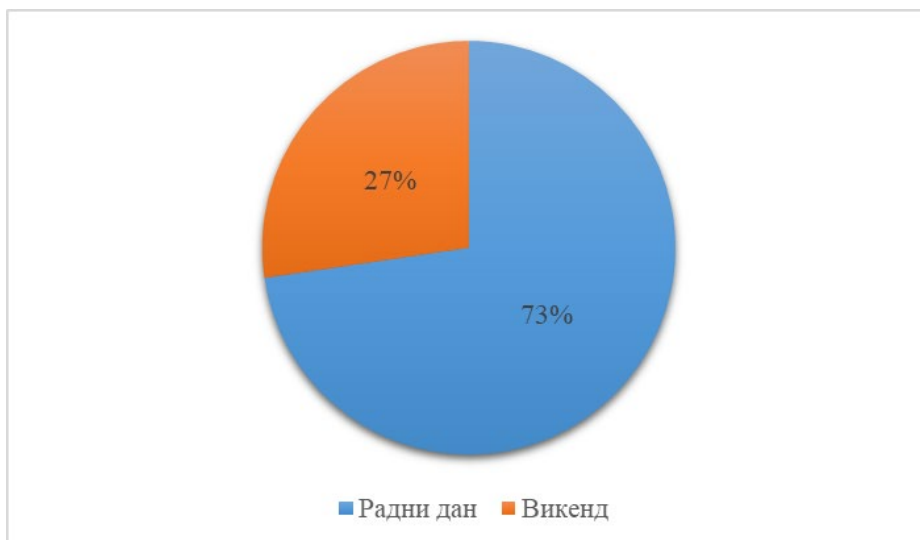
Пошто се ради о дневном лету, и узевши у обзир неколико отказивања самог лета у току посматране године, 300 записа (који представљају познате случајеве) је доступно почевши од 1.10.2015. до 30.9.2016. године. Ови подаци у ствари представљају базу знања и биће коришћени за избор одговарајућег модела, односно комбинације различитих алгоритама и мера. Такође, додатних 10 записа у периоду од 1.10.2016. до 30.11.2016. ће бити коришћена за валидацију и тестирање одабраног модела.

### 8.1.1. Графички приказ података – појединачни атрибути

На следећим Сликама је дат основни приказ података који су коришћени за избор модела, у оквиру којих је сваки атрибут посебно представљен.

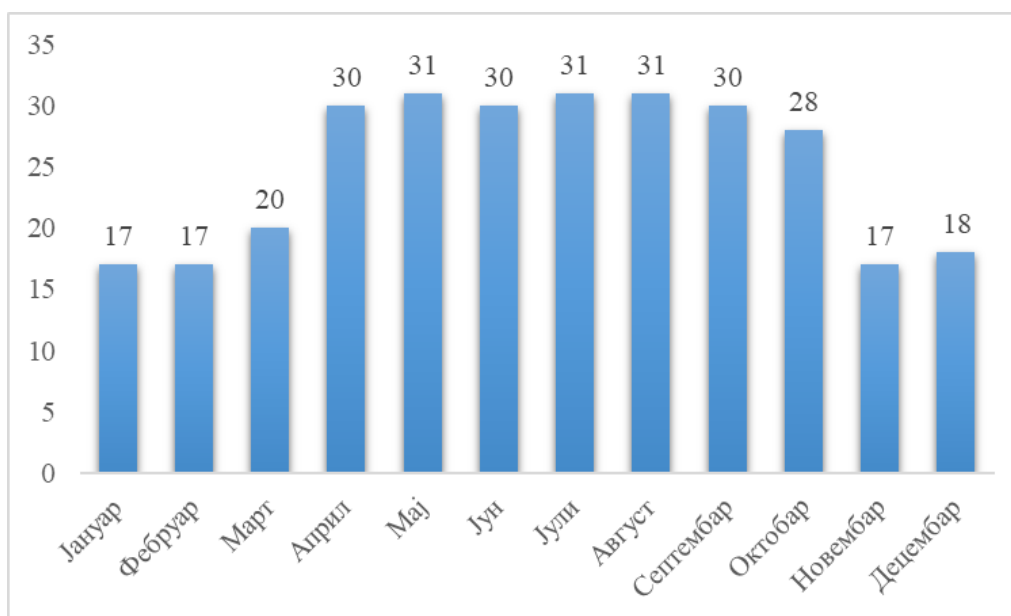


**Слика 19.** Број летова у току и ван сезоне



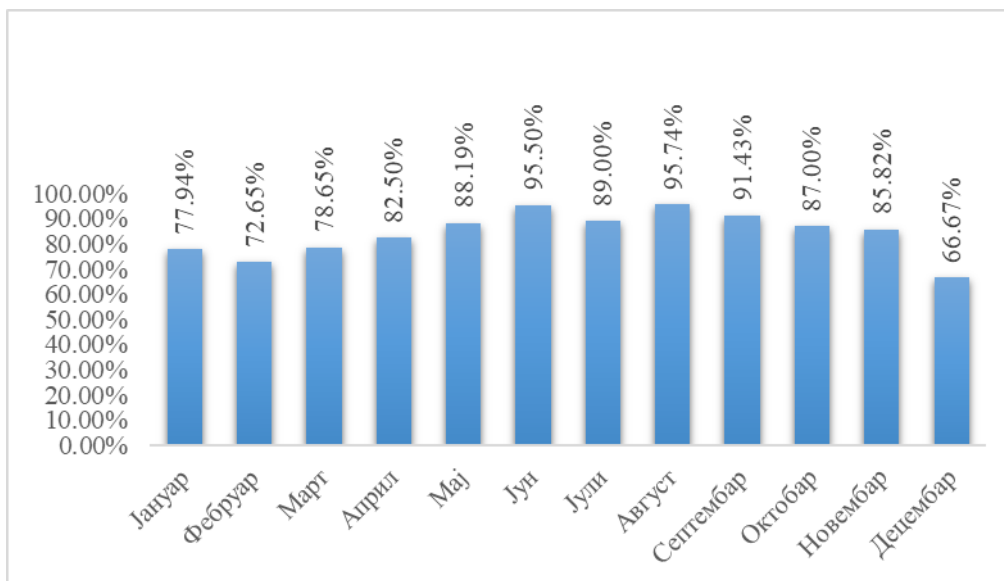
**Слика 20.** Број летова викендом и радним данима

Слика 19 показује да је конкретан лет на релацији Београд – Амстердам, посматрано на годишњем нивоу, претежно био у сезони (71%) а отприлике једна трећина записа се односи на лет у ван сезонском периоду (29%). Слично, према Слици 20 лет је претежно био током радних дана (73%) а нешто мање од једне трећине записа (27%) указује на ситуације када је овај лет био викендом.



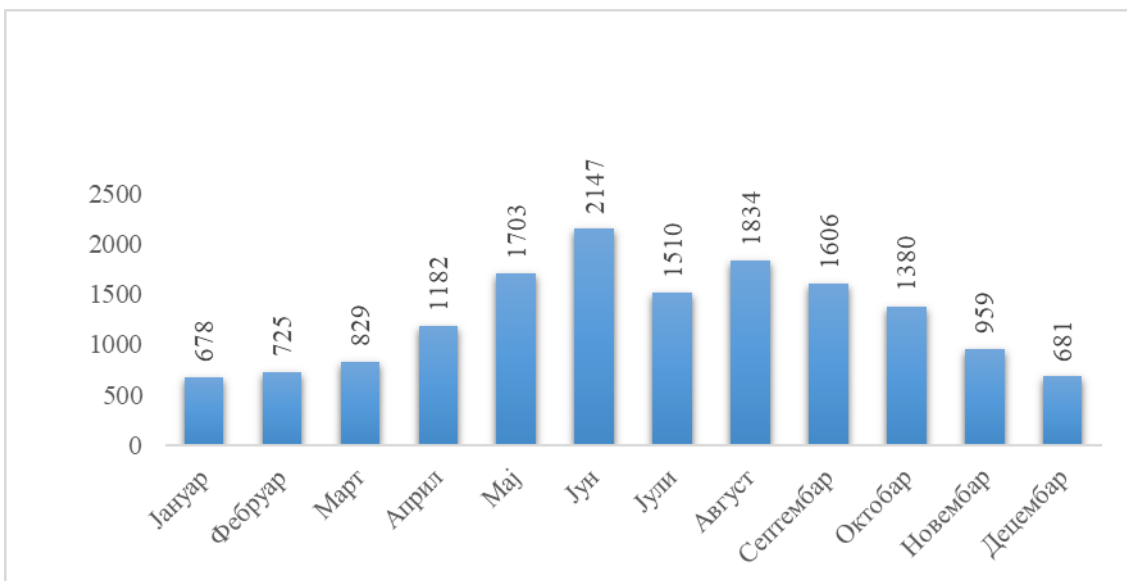
**Слика 21.** Број летова по месецима



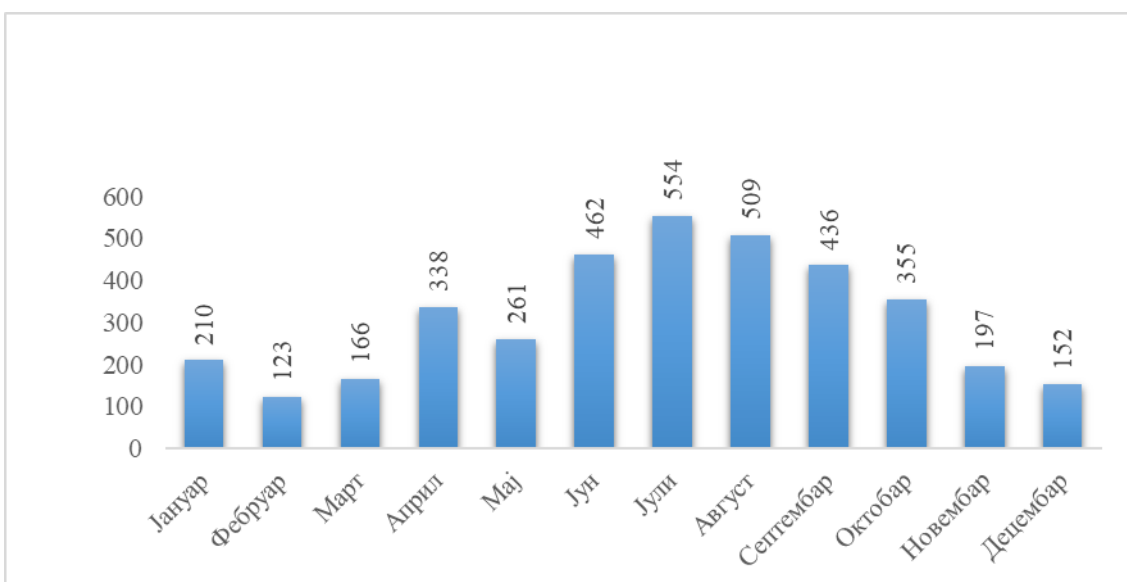


**Слика 22.** Просечан проценат попуњености лета по месецима

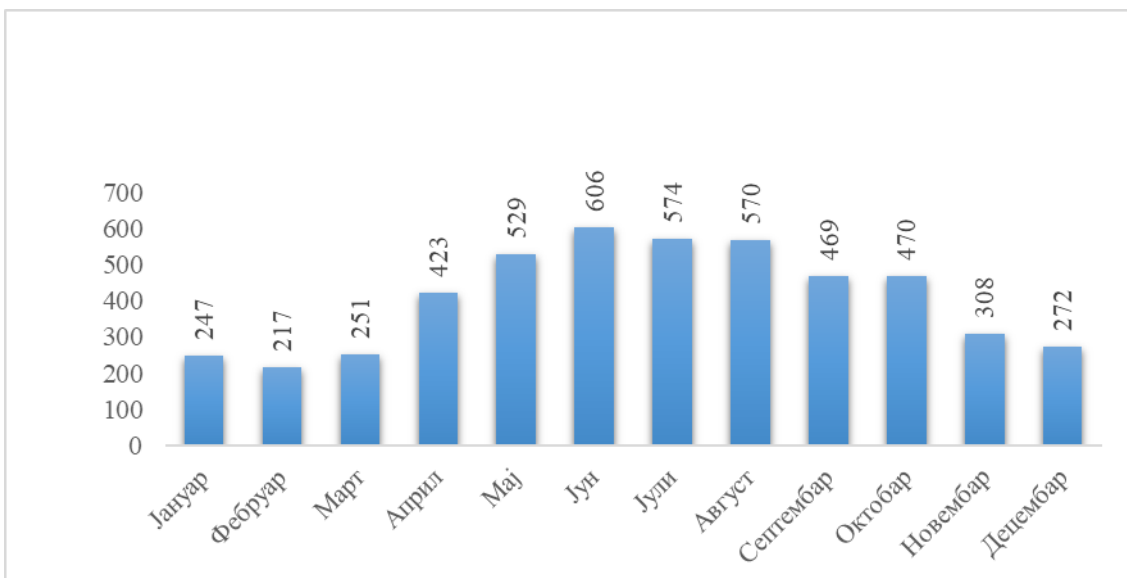
Већ је истакнуто да је већина летова била у сезони за посматран временски период од почетка октобра 2015. до краја септембра 2016. Сходно томе, на Слици 21 се може видети да је лет на релацији Београд – Амстердам од априла па све до септембра (што и представља сезонски период) био сваког дана. Међутим, ово не значи да је на сваком од тих летова било путника у статусу „no-show“, па ће се у наредном кораку претпроцесирања података овај број летова сигурно редуковати у одређеној мери. Посматраће се само летови код којих је било „no-show“ путника. За иницијалну анализу, веома је битно истаћи фреквентност посматраног лета у сезонском периоду, што представља један од показатеља да се ради о лету са високом потражњом. На Слици 22 се може видети просечан проценат попуњености лета по месецима. У јуну (95.5%) и августу (95.74%) су забележени највиши проценти попуњености посматраног лета, што представља такође и сезонски период.



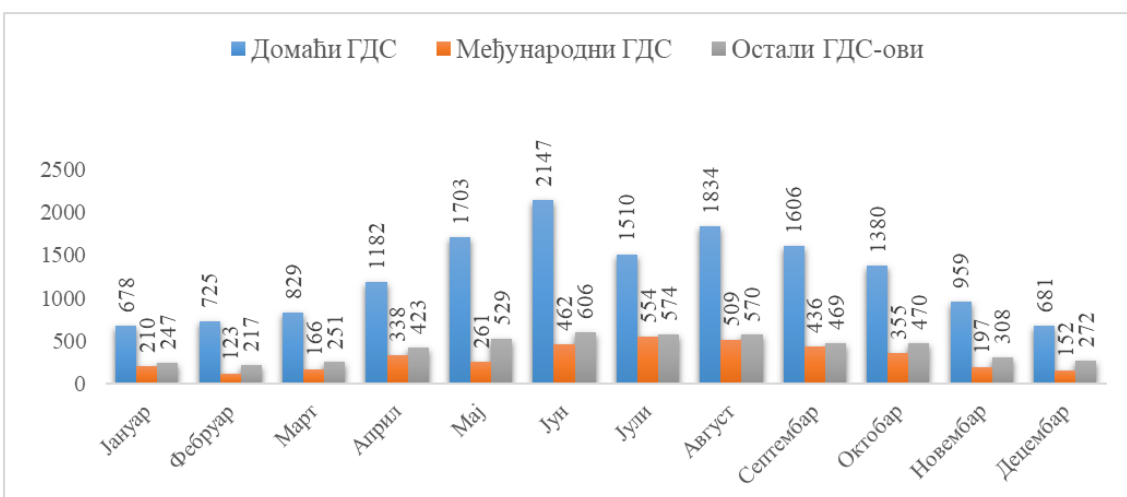
**Слика 23.** Број резервација у оквиру домаћег ГДС-а по месецима



**Слика 24.** Број резервација у оквиру међународног ГДС-а по месецима



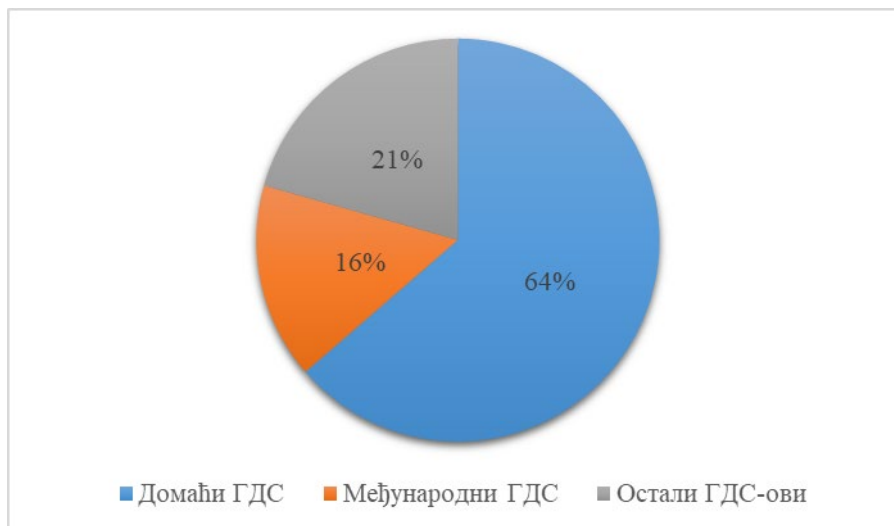
Слика 25. Број резервација у свим осталим ГДС-овима по месецима



Слика 26. Упоредни преглед резервација у ГДС-овима по месецима

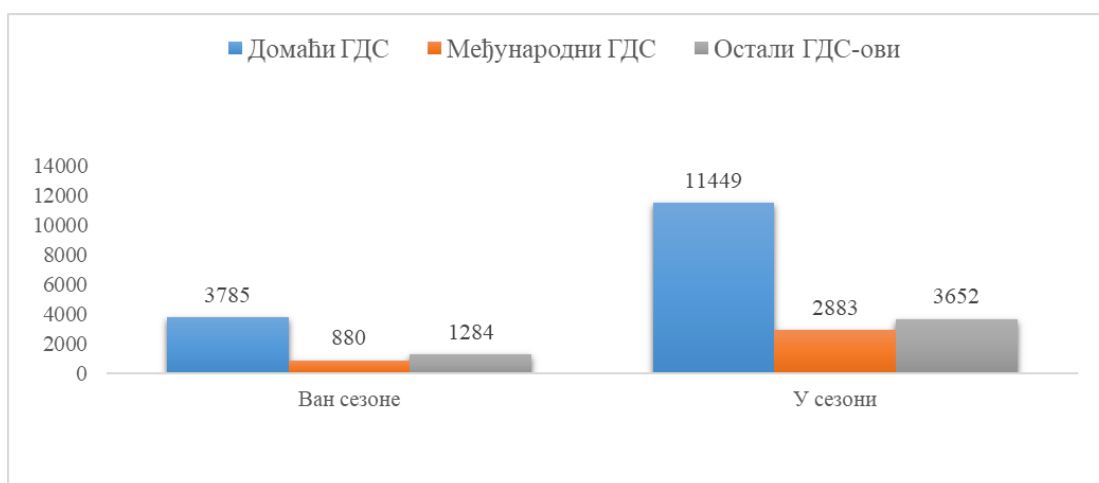
На Сликама 23, 24, 25 и 26 су приказани појединачни и упоредни преглед резервација по месецима у односу на различите ГДС-ове (домаћи, међународни и остали). Забележено је значајно више резервација у сваком месецу које су направљене кроз домаћи ГДС за посматрани лет. Додатно, у јулу је забележен скоро подједнак број резервација направљених кроз међународни ГДС и кроз остале ГДС-ове. На једноставном примеру, ако авио компанија ХУ користи као главни (домаћи) ГДС Амадеус систем, ово би значило да је већина резервација реализована кроз овај систем, без обзира да ли је то било преко веб сајта авио компаније, агенција које исто користе Амадеус ГДС или другог интернет извора.

Остале резервације су спроведене преко агенција или осталих интернет извора (нпр. *Kayak* или *Kiwi*) које претежно користе друге ГДС системе (нпр. *Galileo*, *Sabre*).

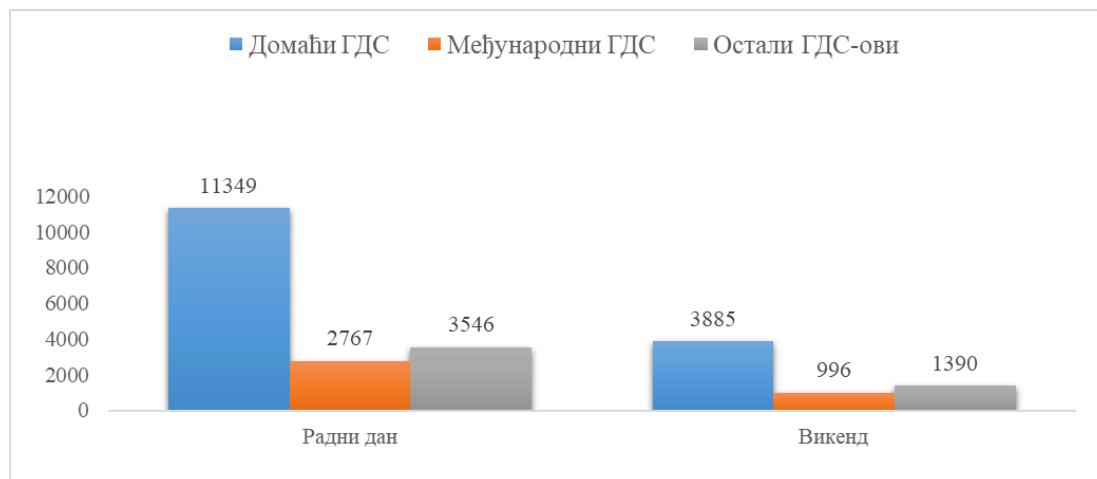


Слика 27. Укупан број резервација по ГДС-овима

Слика 27 приказује преглед укупног броја резервација по ГДС-овима. Као што се може видети, удео резервација које су направљене у оквиру домаћег ГДС-а је значајно већи (64%) у односу на удео резервација направљених у осталим ГДС-овима (21%) и међународном ГДС-у (16%).



Слика 28. Преглед резервација у ГДС-овима у односу на сезонски показатељ

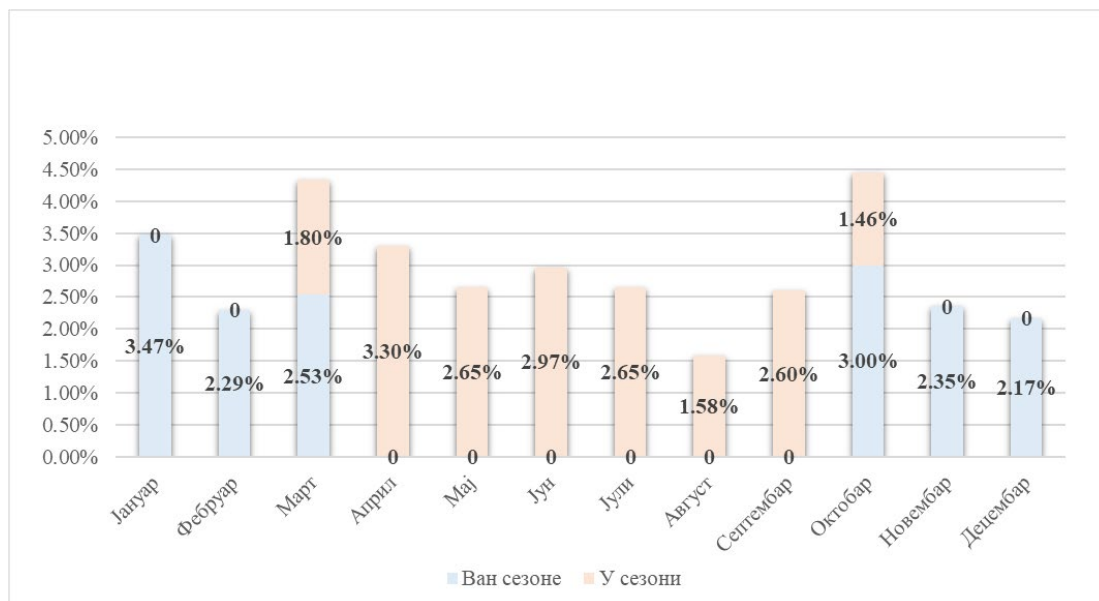


**Слика 29.** Преглед резервација у ГДС-овима у односу на викенд и радне дане

У погледу сезонског показатеља, а према Слици 28, значајно је више резервација у свим посматраним ГДС-овима у оквиру сезоне него ван сезоне. Па тако је у оквиру домаћег ГДС-а у сезони направљено 11449 резервација а ван сезоне само 3785. Ово представља отприлике 75% резервација у сезони наспрам 25% ван сезоне, а сличан однос је присутан и када је у питању међународни ГДС (76% - 24%) и остали ГДС-ови (74% - 26%). На основу Слике 29, може се закључити да је значајно више резервација у оквиру посмартаних ГДС-ова реализовано током радних дана, а проценти су слични као и код сезонског показатеља када се гледа однос. У оквиру домаћег ГДС-а током радних дана је направљено 11349 резервација а током викенда 3885, што представља однос 74% - 26% у корист резервација током радних дана. Слично, овај однос код међународног ГДС-а је 74% - 26% а код осталих ГДС-ова је 72% - 28%. Када се узме у обзир величина авио компаније, може се рећи да је овај однос присутан код већине националних и авио компанија средњих величина (у просеку годишње превезу између 2 и 20 милиона путника).

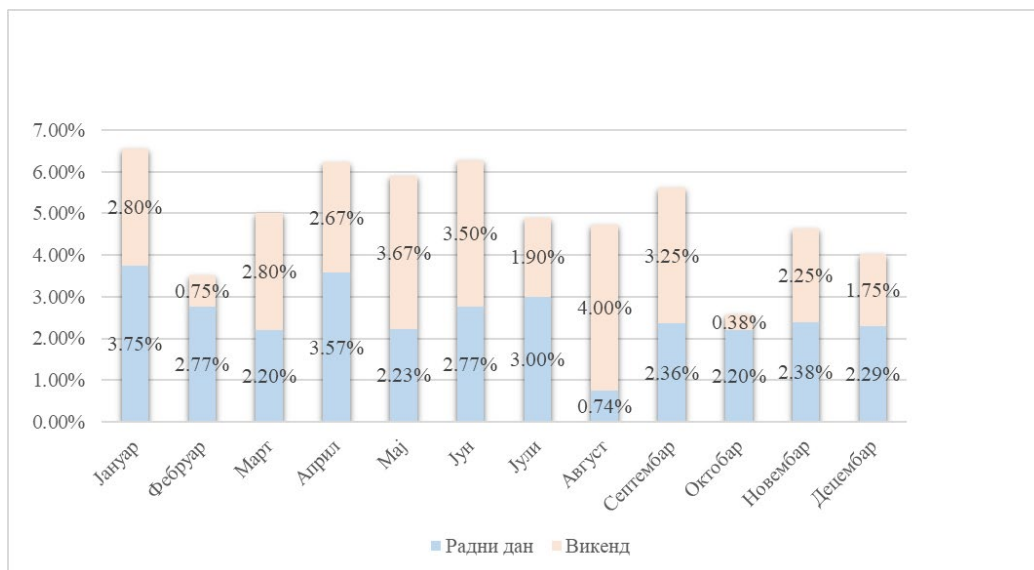
### 8.1.2. Графички приказ података – повезаност атрибута са зависном варијаблом

На сликама које следе је дат приказ повезаности посматраних атрибута са зависном варијаблом (процент „no-show“ путника) из базе знања.



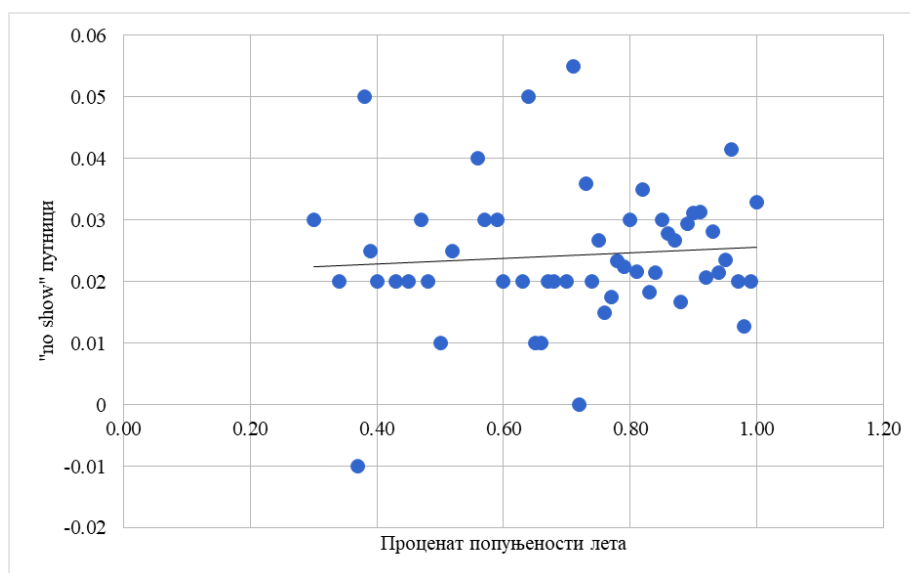
**Слика 30.** Процент „no-show“ путника у току и ван сезоне

Слика 30 показује разлику између броја „no-show“ путника у току и ван сезоне. У марту (4.33%) и октобру (4.46%) је просечно посматрано било највише путника који се нису појавили на својим летовима, а ово су уједно и месеци у оквиру којих авио компаније мењају своје редове летења. Уопштено гледајући и узимајући у обзир навике људи који путују, промене у терминима летова и учесталости могу се довести у везу са непојављивањем одређеног процента путника на летовима. У другој половини марта за већину авио компанија почиње сезона и завршава се у другој половини октобра. Најмањи број путника који се нису појавили на лету је забележен у августу, што је и очекивано с обзиром да је то сезона одмора и већину резервација путници реализују у ту сврху. Просечан проценат „no-show“ путника је сличан у току (2.47%) и ван сезоне (2.58%).



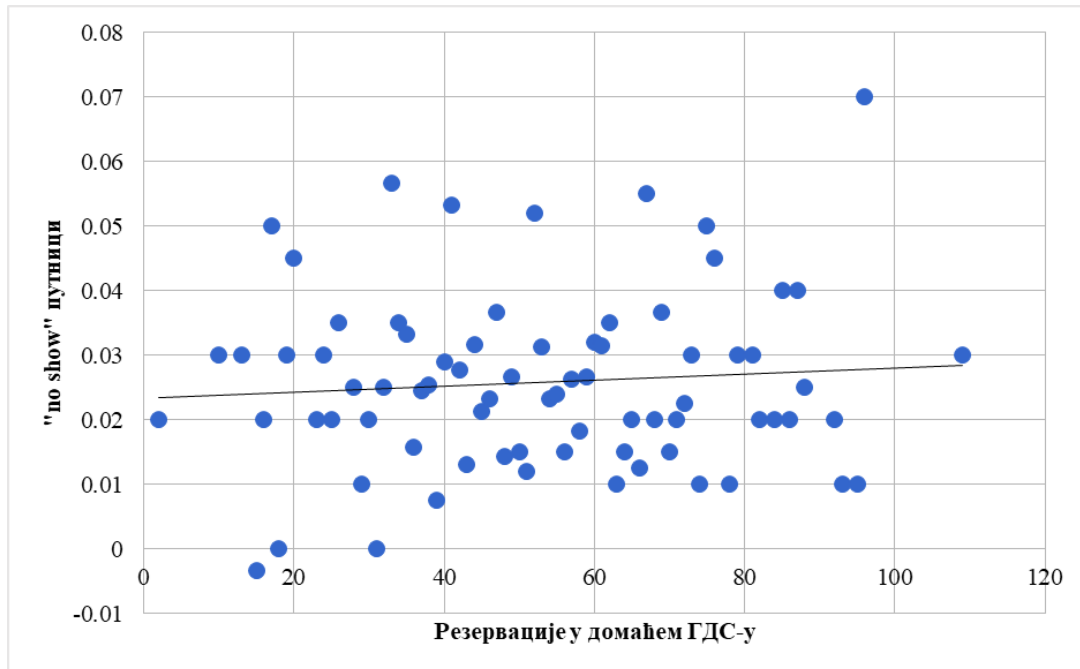
**Слика 31.** Процент „no-show“ путника викендом и радним данима

На Слици 31 се може видети разлика у броју „no-show“ путника током викенда и радних дана. Значајно је мање „no-show“ путника викендима у фебруару (0.75%) и октобру (0.38%), а значајно више забележено је у августу (4.00%). Просечан број путника који се неће појавити на лету је значајно већи током радних дана (2.29%) у односу на викенде (1.75%).



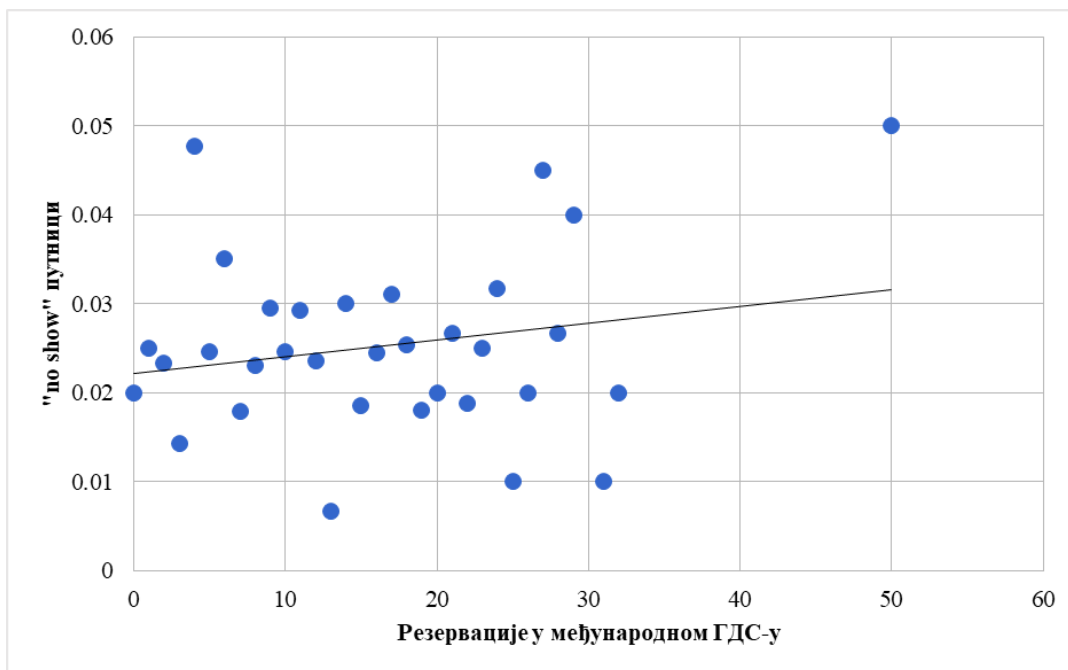
**Слика 32.** Приказ зависности „no-show“ путника и процента попуњености лета

Дијаграм растурања (енг. *scatter plot*) који је приказан на Слици 32 указује на слабу линеарну зависност између „no-show“ путника и процента попуњености лета.

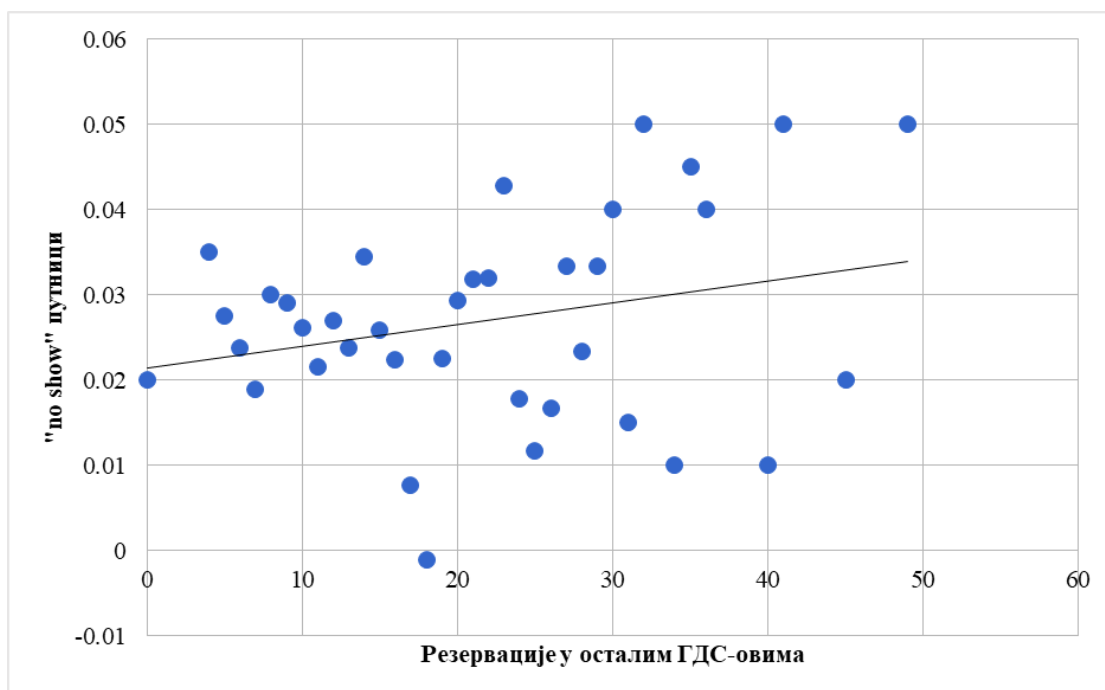


**Слика 33.** Приказ зависности „no-show“ путника и резервација у оквиру домаћег ГДС-а





**Слика 34.** Приказ зависности „no-show“ путника и резервација у оквиру међународног ГДС-а



**Слика 35.** Приказ зависности „no-show“ путника и резервација у оквиру осталих ГДС-ова

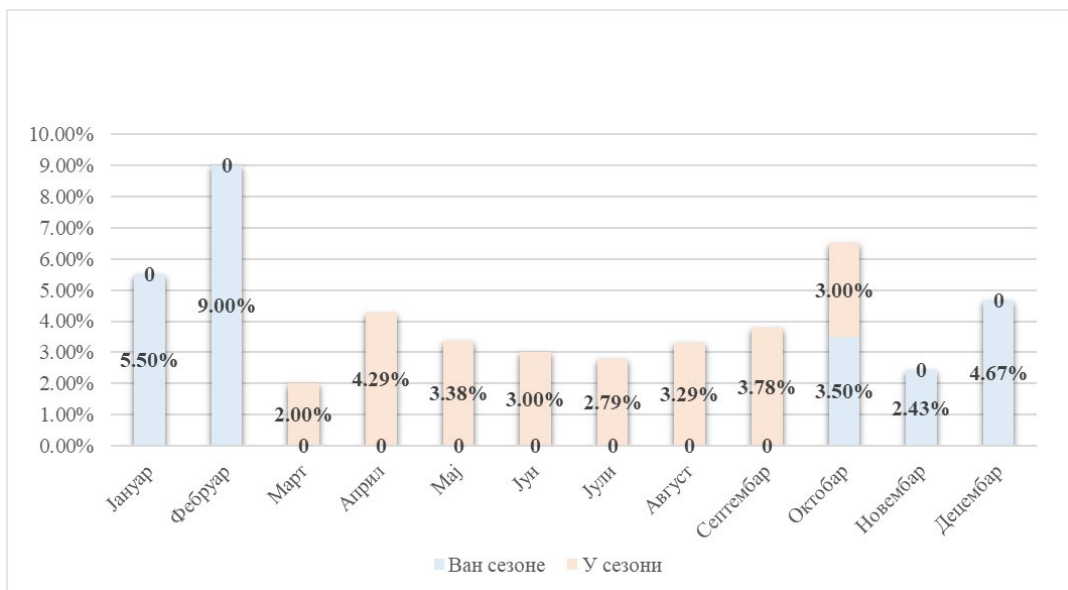
Као што се може видети, дијаграми растурања који су приказани на Сликама 33, 34 и 35 указују на слабу повезаност између „no-show“ путника и броја резервација у оквиру домаћег, међународног и осталих ГДС-ова.

## 8.2. Избор модела - претпроцесирање података

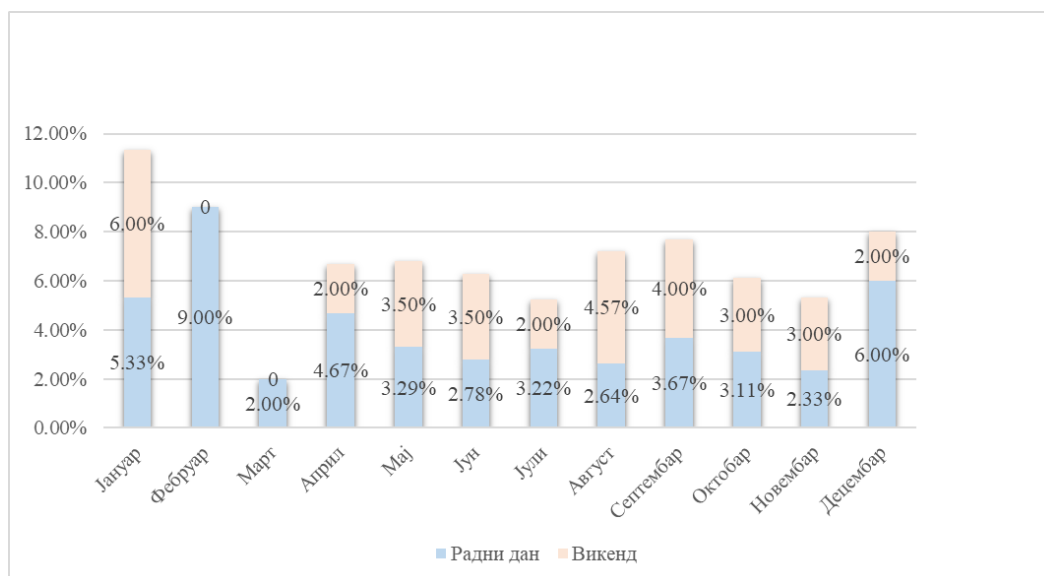
Подаци о иницијалних 300 познатих случајева садрже у себи одређени ниво случајних грешака приликом евидентирања и извлачења из резервационог система, као и празних поља. Такође, неки записи о летовима не могу бити коришћени јер или не постоји на тим летовима путника који се нису појавили или имају превише мали проценат попуњености лета. Релативно низак проценат попуњености указује да није постојала висока потражња за тим одређеним летовима и авио компанија није имала интереса да омогући „overbooking“ на њима. Узевши све ово у обзир, неопходно је спровести одређене активности припреме базе података.

### *Корак 1. Припрема базе података*

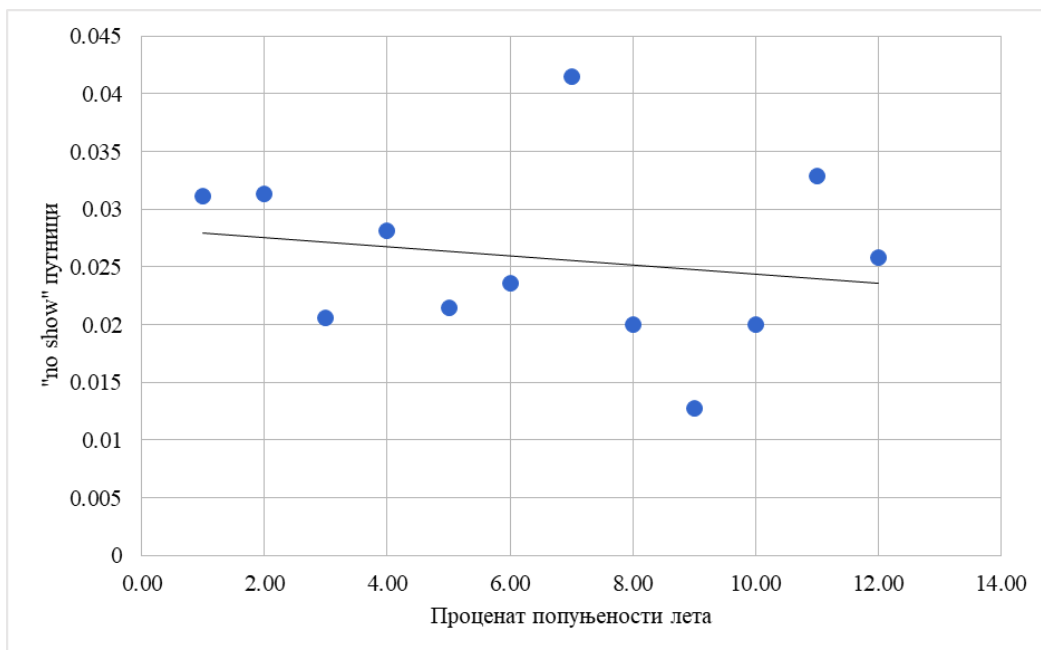
Узевши у обзир дате приказе и оно на шта указују несређени подаци, неопходно је спровести одређене активности како би се ти исти подаци прилагодили и могли користити за избор модела за предвиђање путника који се неће појавити на лету. Наиме, потребно је размотрити има ли потребе за уклањањем записа (односно летова) у оквиру којих је број „no-show“ путника једнак или мањи од нуле. Такође, потребно је уклонити записе у оквиру којих је остварен низак проценат попуњености лета, односно мањи од 90%, јер такви летови нису од интереса са „overboking“ аспекта. Другим речима, веома је мала вероватноћа да код летова који нису довољно попуњени (нпр. попуњеност од 80%), дође до резервације свих 20% седишта у наредних неколико дана пре лета. Ова вредност је сугерисана од стране аутора (Vojtek, Petrović & Milošević, Forthcoming 2021) и посматраће се као параметар који је могуће подесити сваки пут пре пуштања система у примену, као што је и објашњено у потпоглављу 5.3.4. Након уклањања ових записа, на следећим сликама су приказани сређени подаци у оквиру базе података који ће бити коришћени за избор модела.



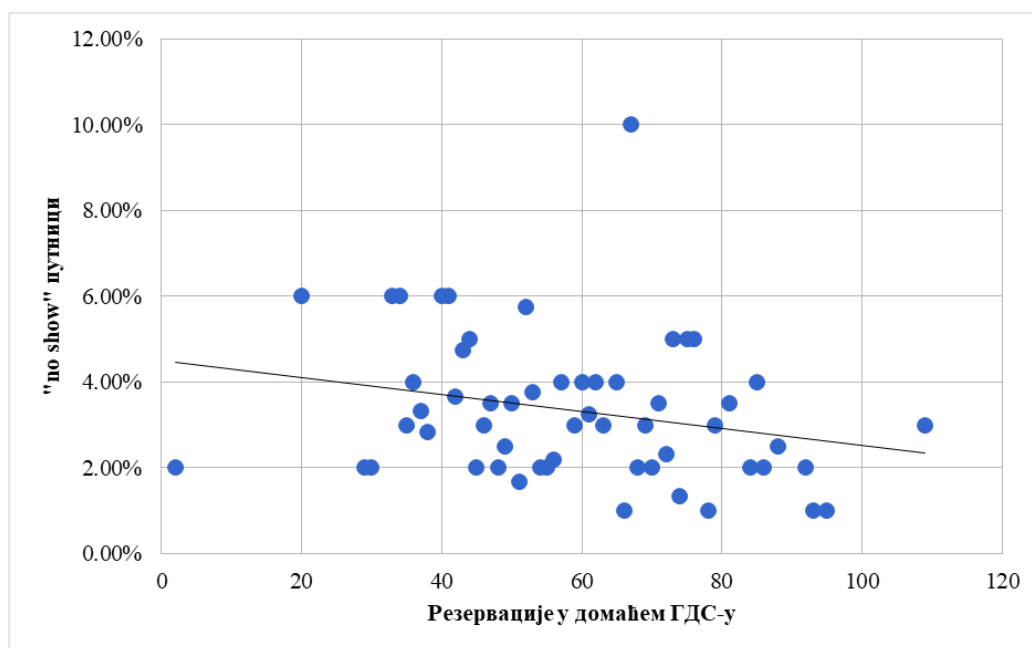
Слика 36. Процент „no-show“ путника у току и ван сезоне



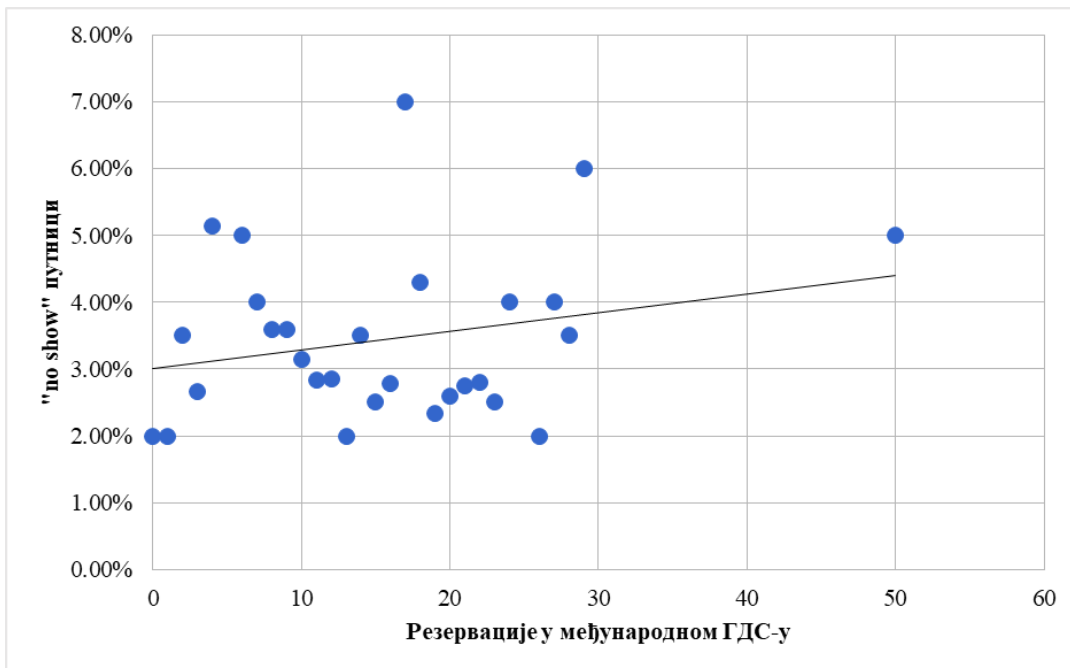
Слика 37. Процент „no-show“ путника викендом и радним данима



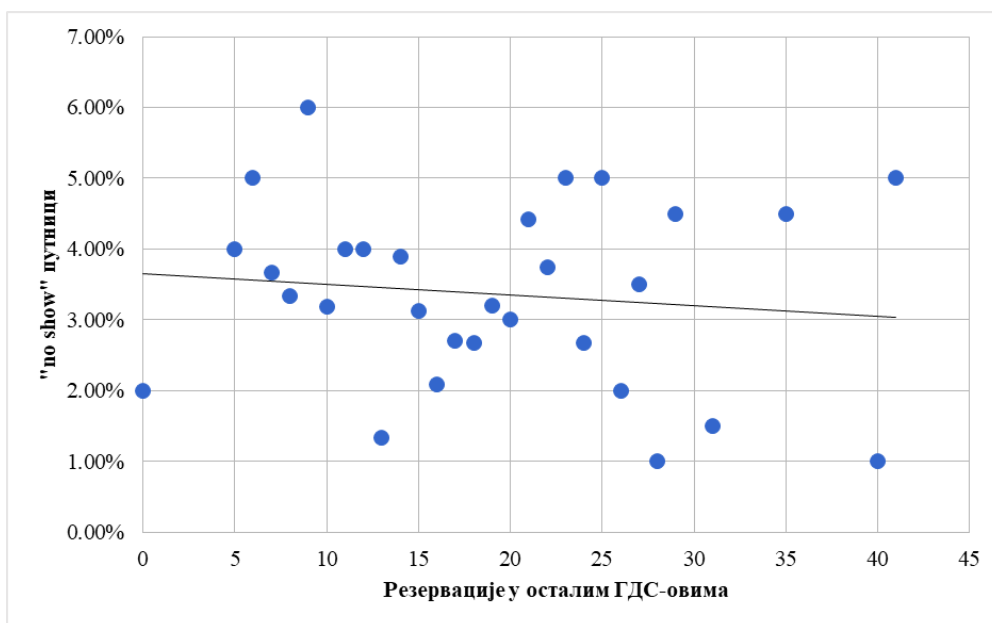
**Слика 38.** Приказ зависности „no-show“ путника и процента попуњености лета



**Слика 39.** Приказ зависности „no-show“ путника и резервација у оквиру домаћег ГДС-а



**Слика 40.** Приказ зависности „no-show“ путника и резервација у оквиру међународног ГДС-а



**Слика 41.** Приказ зависности „no-show“ путника и резервација у оквиру осталих ГДС-ова

Према Слици 36, сређени подаци сада показују значајну разлику у процентима „no-show“ путника у току (3.25%) и ван сезоне (4.06%), а посебно се истиче месец фебруар где је проценат „no-show“ путника значајно виши (9%). С друге стране, као што је приказано на Слици 37, разлика између непојављивања

путника на лету радним данима (3.29%) и викендом (3.53%) је сада смањена, где је мало више “no-show“ путника током викенда. Дијаграми растурања на Сликама 38, 39, 40 и 41 сада показују јачу повезаност (позитивну и негативну) између “no-show“ путника и процента попуњености лета и резервација у оквиру ГДС-ова. Након сређивања базе података, 123 записа, односно познатих случајева је коришћено за избор модела, што је представљено у следећем потпоглављу. У циљу испитивања и потврђивања значајности модела који дају најбоље резултате, два статистичка теста значајности су спроведена у оквиру потпоглавља 8.4. Модели са најбољим резултатима су примењени на додатних 10 познатих случајева у циљу валидације целокупног система у оквиру потпоглавља 8.5.

### 8.3. Избор модела - алгоритам за одређивање сличности

Сам процес избора модела је почео првим кораком у оквиру претходног потпоглавља где су спроведене активности припреме података. Пречишћена база података се састоји од 123 позната случајева који ће бити коришћени у оквиру даљих корака процеса. Применом различитих техника скалирања података, различитих мера за израчунавање удаљености и сличности, као и употребом различитих агрегатора, дефинисана су укупно 18 различитих модела од којих је потребно одабрати онај који даље најбоље резултате предвиђања. Модели су дати у оквиру следеће Табеле 2.

**Табела 2.** Листа потенцијалних модела

Модел #	Процедура за скалирање података	Мера сличности	Граница прихватања (ниво атрибута)	Агрегација (ниво случаја)	Агрегација (виши ниво)	Експертска препорука
1	Стандардизација	Еуклидска	$D_{ED} \leq \bar{D} - \sigma$	Просек	-	Искључена
2	Стандардизација	Менхетн	$D_{MD} \leq \bar{D} - \sigma$	Просек	-	Искључена
3	Стандардизација (+ Нормализација за ИБА)	ИБА	$S_{IBA} \geq \bar{S} + \sigma$	Просек	-	Искључена
4	Стандардизација	Еуклидска	$D_{ED} \leq \bar{D} - \sigma$	Просек	Просек	Укључена
5	Стандардизација	Менхетн	$D_{MD} \leq \bar{D} - \sigma$	Просек	Просек	Укључена
6	Стандардизација (+ Нормализација за ИБА)	ИБА	$S_{IBA} \geq \bar{S} + \sigma$	Просек	Просек	Укључена
7	Стандардизација	Еуклидска	$D_{ED} \leq \bar{D} - \sigma$	Просек	ИБА ( <i>min</i> )	Укључена
8	Стандардизација	Менхетн	$D_{MD} \leq \bar{D} - \sigma$	Просек	ИБА ( <i>min</i> )	Укључена
9	Стандардизација (+ Нормализација за ИБА)	ИБА	$S_{IBA} \geq \bar{S} + \sigma$	Просек	ИБА ( <i>min</i> )	Укључена
10	Нормализација	Еуклидска	$D_{ED} \leq \bar{D} - \sigma$	Просек	-	Искључена
11	Нормализација	Менхетн	$D_{MD} \leq \bar{D} - \sigma$	Просек	-	Искључена
12	Нормализација	ИБА	$S_{IBA} \geq \bar{S} + \sigma$	Просек	-	Искључена
13	Нормализација	Еуклидска	$D_{ED} \leq \bar{D} - \sigma$	Просек	Просек	Укључена
14	Нормализација	Менхетн	$D_{MD} \leq \bar{D} - \sigma$	Просек	Просек	Укључена
15	Нормализација	ИБА	$S_{IBA} \geq \bar{S} + \sigma$	Просек	Просек	Укључена
16	Нормализација	Еуклидска	$D_{ED} \leq \bar{D} - \sigma$	Просек	ИБА ( <i>min</i> )	Укључена
17	Нормализација	Менхетн	$D_{MD} \leq \bar{D} - \sigma$	Просек	ИБА ( <i>min</i> )	Укључена
18	Нормализација	ИБА	$S_{IBA} \geq \bar{S} + \sigma$	Просек	ИБА ( <i>min</i> )	Укључена

Посматрано у глобалу, наведени модели укључују два приступа - први који се односи на израчунавање удаљености и стандардну агрегацију и други логички приступ који се базира на примени ИБА концепта. За скалирање података, две процедуре су коришћене у оквиру модела – Стандардизација и Нормализација. Такође, у Табели су јасно означени модели код којих је експертска препорука укључена и код којих није узета у обзир (искључена из модела).

### *Корак 2. Идентификација екстремних вредности*

Као што је описано у поглављу 7, утицај екстремних вредности је смањен применом познате трансформационе процедуре у оквиру које се идентификоване вредности замењују најближим неекстремним вредностима (енг. *Winsorizing*).

Узевши у обзир природу података, утицај екстремних вредности је могуће смањити код три атрибута – *Број резервација у оквиру домаћег ГДС-а* ( $x_5$ ), *Број*

резервација у оквиру међународног ГДС-а ( $x_6$ ) и Број резервација у свим осталим ГДС-овима ( $x_7$ ). Како би се прво идентификовале екстремне вредности, неопходно је груписати податке и израчунати број њиховог појављивања, односно фреквенцију. У оквиру следећих Табела 3, 4 и 5 су дати груписани подаци о броју резервација у ГДС-овима, као и појединачна и кумулативна фреквенција.

**Табела 3.** Број резервација у оквиру домаћег ГДС-а

Опсег	Број резервација	Кумулативно %
10	1	0.71%
20	0	0.71%
30	3	2.84%
40	17	14.89%
50	29	35.46%
60	41	64.54%
70	20	78.72%
80	17	90.78%
90	9	97.16%
100	3	99.29%
110	1	100.00%
Укупно		100.00%

**Табела 4.** Број резервација у оквиру међународног ГДС-а

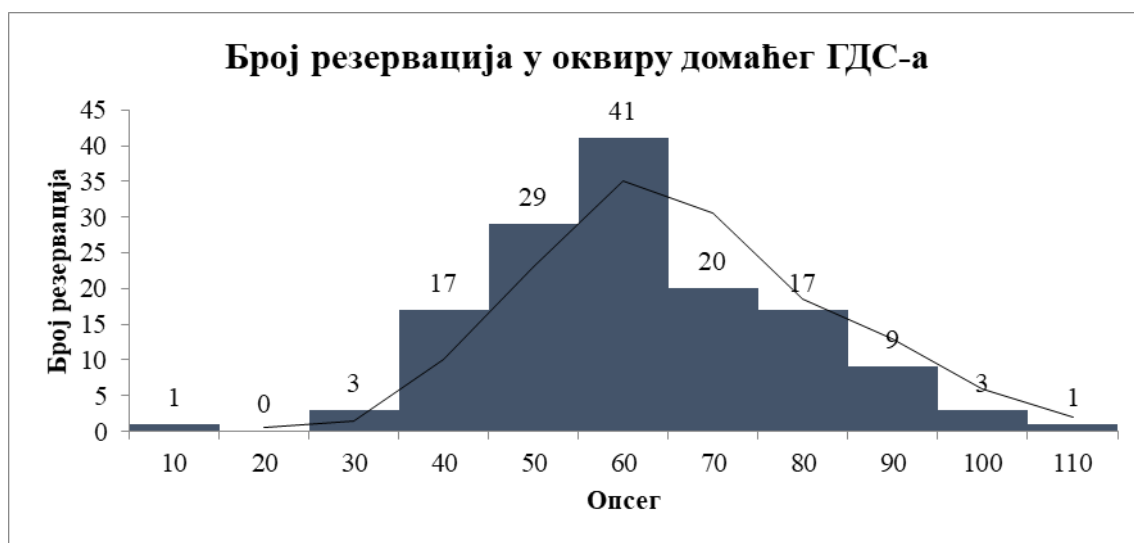
Опсег	Број резервација	Кумулативно %
10	49	34.75%
20	68	82.98%
30	23	99.29%
40	0	99.29%
50	1	100.00%
60	0	100.00%
70	0	100.00%
80	0	100.00%
90	0	100.00%
100	0	100.00%
110	0	100.00%
Укупно		100.00%



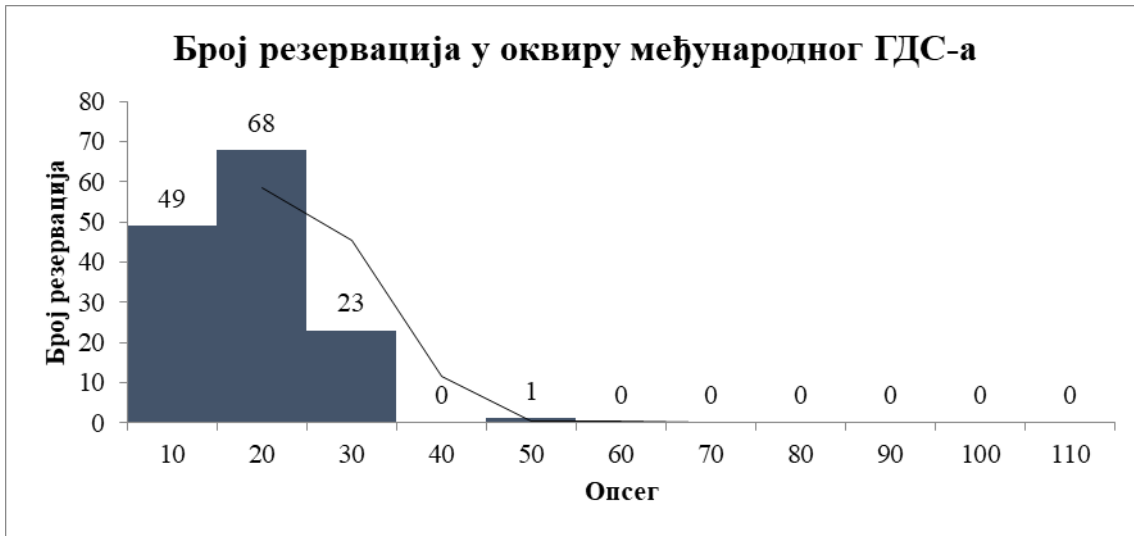
**Табела 5.** Број резервација у оквиру осталих ГДС-ова

Опсег	Број резервација	Кумулативно %
10	24	17.02%
20	69	65.96%
30	42	95.74%
40	5	99.29%
50	1	100.00%
60	0	100.00%
70	0	100.00%
80	0	100.00%
90	0	100.00%
100	0	100.00%
110	0	100.00%
Укупно		100.00%

Подаци из Табела су приказани и на следећим хистограмима, где су екстремне вредности видљивије.



**Слика 42.** Број резервација у оквиру домаћег ГДС-а



**Слика 43.** Број резервација у оквиру међународног ГДС-а



**Слика 44.** Број резервација у оквиру осталих ГДС-ова

Као што се може видети и на претходним сликама 42, 43 и 44, присуство екстремних вредности је евидентно. Како ове вредности могу утицати на резултате добијене коришћењем предложеног алгорита, примењена је трансформациона процедура која је предложена у потпоглављу 7.3.4. Идентификоване екстремне вредности се замењују са најближим неекстремним вредностима (енг. *Winsorizing*). Овако сређени подаци су прво приказани у оквиру следећих Табела 6, 7 и 8.

**Табела 6.** Број резервација у домаћем ГДС-у - Без екстремних вредности

Опсег	Број резервација	Кумулативно %
10	0	0.00%
20	0	0.00%
30	0	0.00%
40	21	14.89%
50	29	35.46%
60	41	64.54%
70	20	78.72%
80	17	90.78%
90	13	100.00%
100	0	100.00%
110	0	100.00%
Укупно		100.00%

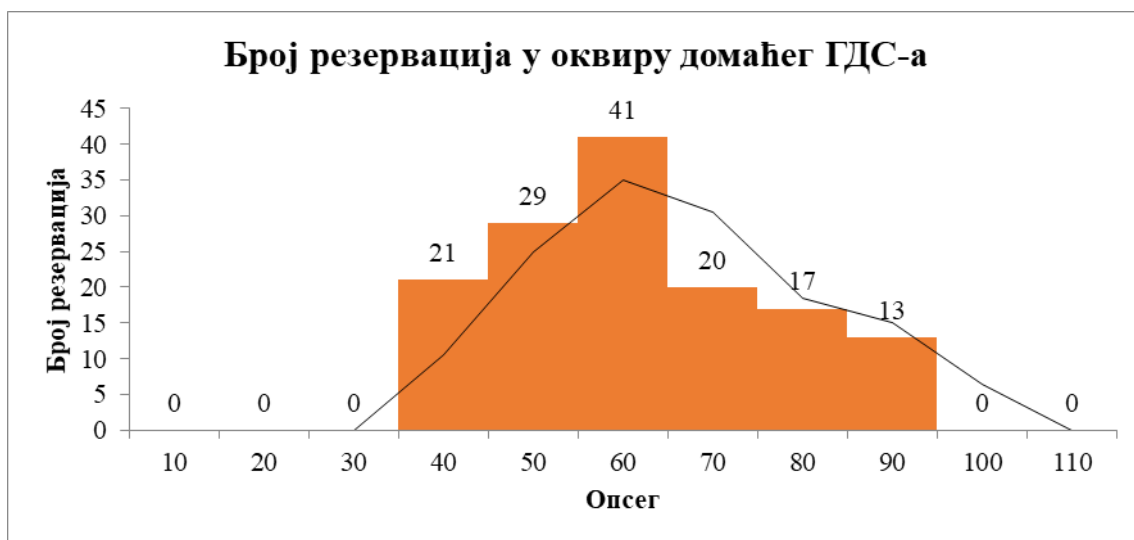
**Табела 7.** Број резервација у међународном ГДС-у - Без екстремних вредности

Опсег	Број резервација	Кумулативно %
10	49	34.75%
20	68	82.98%
30	24	100.00%
40	0	100.00%
50	0	100.00%
60	0	100.00%
70	0	100.00%
80	0	100.00%
90	0	100.00%
100	0	100.00%
110	0	100.00%
Укупно		100.00%

**Табела 8.** Број резервација у осталим ГДС-овима - Без екстремних вредности

Опсег	Број резервација	Кумулативно %
10	24	17.02%
20	69	65.96%
30	48	100.00%
40	0	100.00%
50	0	100.00%
60	0	100.00%
70	0	100.00%
80	0	100.00%
90	0	100.00%
100	0	100.00%
110	0	100.00%
Укупно		100.00%

Сређени подаци су приказани и на следећим сликама кроз хистограме, где се може видети да је утицај екстремних вредности смањен након примене трансформационе процедуре.



Слика 45. Број резервација у домаћем ГДС-у - Без екстремних вредности



Слика 46. Број резервација у међународном ГДС-у - Без екстремних вредности



Слика 47. Број резервација у осталим ГДС-овима - Без екстремних вредности

Након што је утицај екстремних вредности смањен, прелази се на следећи корак у оквиру процеса избора модела за предвиђање „no-show“ путника.

#### *Корак 3. Скалирање података*

Атрибути су скалирани коришћењем два приступа - нормализације (*min - max*) и конвенционалне стандардизације. Додатно, комбинација алгорита која укључује ИБА приступ за израчунавање сличности и стандардизацију за скалирање, захтева још један додатан корак нормализације података. Ово је предуслов за коришћење ИБА приступа и примену ИБА мере сличности.

#### *Корак 4. Одређивање сличности између атрибута*

Након што су подаци сређени, утицај екстремних вредности смањен и подаци скалирани, следећи корак представља почетак тестирања, у оквиру кога се сви случајеви из базе користе, а сваки случај се тестира у односу на друге случајеве. Прво се одређује сличност између атрибута првог случаја и атрибута свих осталих случаја у бази, затим се поступак понавља за други случај и све до последњег случаја који се налази у бази и одређивање сличности његових атрибута са атрибутима осталих случајева. У оквиру предложеног система, три мере су биле коришћене за одређивање сличности (Еуклидско и Менхетн растојање, као и ИБА мера сличности) као део технике закључивања на основу случаја.

#### *Korak 5. Oдређивање појединачних и просечних сличности*

На основу одређених сличности и израчунатих удаљености између атрибута, израчунате су вредности на нивоу случаја. Ове вредности представљају појединачне сличности свих случајева са првим случајем, затим свих осталих случајева са другим случајем и тако све док се и сличности свих случаја са последњим случајем у бази не одреди. Сходно томе, добијене су вредности које представљају појединачне сличности између првог случаја са осталим случајевима. На основу овога, израчунате су просечна вредност (сличност) и стандардно одступање које ће се користити у наредном кораку за дефинисање границе прихватања. Истим поступком, коришћене су добијене појединачне сличности између другог случаја и осталих случајева и израчунати поменуте мере. Поступак је поновљен до последњег случаја у бази.

#### *Korak 6. Избор случајева и агрегација постојећих решења*

Границе прихватања су претходно дефинисане у оквиру потпоглавља 5.3. Приликом коришћења мера удаљености, два објекта који имају исту вредност ће имати удаљеност једнаку нули (0). Код ИБА мере сличности, два иста објекта ће имати вредност један (1). На основу овога, случајеви из базе су одабрани и урађена је агрегација њихових решења, односно процента “no-show“ путника. За агрегацију су коришћена два приступа. Први се односи на ИБА логички приступ који је реализован коришћењем *min* оператора у смислу логичке конјункције, а други на израчунавање просечне вредности.

#### *Korak 7. Избор случајева и њихових решења од стране експерта*

За сваки посматрани и тестирани случај, експерт је одабрао, као и претходно алгоритам, исти број осталих постојећих случаја из базе који су најсличнији. Сам поступак избора се састоји из неколико корака. Прво, експерт анализира тестни случај, односно карактеристике лета за који је потребно проценити број путника који се неће појавити. Овде експерт акценат ставља на временске показатеље и број резервација у различитим ГДС-овима, као и процентуалну попуњеност лета. Затим, експерт пролази кроз постојеће случајеве, односно летове из прошлости који су се десили у истим или сличном временским

периодима. Након тога, експерт анализира број резервација и њихову расподелу у оквиру ГДС-ова. Такође, експерт узима у обзир и проценат попуњености лета и на основу свега овога бира најсличније случајеве из базе. На крају, израчунава се просечна вредност процента „no-show“ путника код одабраних случајева, што представља предлог решења од стране експерта.

#### *Корак 8. Агрегација предложених решења*

Просечне вредности за добијање предложеног решења су коришћене без и у комбинацији са предлогом експерта. У оквиру модела где је експертски предлог укључен, за добијање предлога решења коришћен је додатно и логички ИБА приступ кроз употребу *min* оператора.

#### *Корак 9. Израчунавање грешке предвиђања и избор модела*

Перформансе предвиђања свих модела су утврђене коришћењем две мере у виду грешака - средња апсолутна процентуална грешка и средњи апсолутни трошак жаљења. Прва мера је одабрана јер се очекује да ће дати реалну и јасну слику грешке предвиђања у процентима која ће моћи и лако да се протумачи и сагледа. Као што је раније назначено у оквиру описа целокупног процеса, друга мера је уведена како би се раздвојиле и посматрале различито ситуације код којих је потцењено и прецењено предвиђање, поготову са аспекта трошкова који настају за авио компанију. За потцењена предвиђања коришћен је тежински коефицијент  $w_1 = 0.4$ , а за прецењена  $w_2 = 0.6$ . Перформансе модела су дате у оквиру следеће Табеле 9.

**Табела 9.** Оцена перформанси модела

Модел #	Скалирање података	Мера сличности	Граница прихватања (ниво атрибута)	Агрег. (ниво случаја)	Агрег. (виши ниво)	Експертска препорука	MAPE	MACR
1	Станд.	Еуклидска	$D_{ED} \leq \bar{D} - \sigma$	Просек	NA	Искључена	0.407	0.0036
2	Станд.	Менхетн	$D_{MD} \leq \bar{D} - \sigma$	Просек	NA	Искључена	0.437	0.0044
3	Станд. (+ Норм. за ИБА)	ИБА	$S_{IBA} \geq \bar{S} + \sigma$	Просек	NA	Искључена	0.611	0.0056
4	Станд.	Еуклидска	$D_{ED} \leq \bar{D} - \sigma$	Просек	Просек	Укључена	0.288	0.0029
5	Станд.	Менхетн	$D_{MD} \leq \bar{D} - \sigma$	Просек	Просек	Укључена	0.304	0.0030
6	Станд. (+ Норм. за ИБА)	ИБА	$S_{IBA} \geq \bar{S} + \sigma$	Просек	Просек	Укључена	0.392	0.0037
7	Станд.	Еуклидска	$D_{ED} \leq \bar{D} - \sigma$	Просек	ИБА ( <i>min</i> )	Укључена	0.275	0.0031
8	Станд.	Менхетн	$D_{MD} \leq \bar{D} - \sigma$	Просек	ИБА ( <i>min</i> )	Укључена	0.289	0.0033
<b>9</b>	<b>Станд. (+ Норм. за ИБА)</b>	<b>ИБА</b>	<b><math>S_{IBA} \geq \bar{S} + \sigma</math></b>	<b>Просек</b>	<b>ИБА (<i>min</i>)</b>	<b>Укључена</b>	<b>0.251</b>	<b>0.0028</b>
10	Норм.	Еуклидска	$D_{ED} \leq \bar{D} - \sigma$	Просек	NA	Искључена	0.431	0.0043
11	Норм.	Менхетн	$D_{MD} \leq \bar{D} - \sigma$	Просек	NA	Искључена	0.455	0.0045
12	Норм.	ИБА	$S_{IBA} \geq \bar{S} + \sigma$	Просек	NA	Искључена	0.616	0.0056
13	Норм.	Еуклидска	$D_{ED} \leq \bar{D} - \sigma$	Просек	Просек	Укључена	0.307	0.0031
14	Норм.	Менхетн	$D_{MD} \leq \bar{D} - \sigma$	Просек	Просек	Укључена	0.321	0.0032
15	Норм.	ИБА	$S_{IBA} \geq \bar{S} + \sigma$	Просек	Просек	Укључена	0.394	0.0037
<b>16</b>	<b>Норм.</b>	<b>Еуклидска</b>	<b><math>D_{ED} \leq \bar{D} - \sigma</math></b>	<b>Просек</b>	<b>ИБА (<i>min</i>)</b>	<b>Укључена</b>	<b>0.271</b>	<b>0.0031</b>
17	Норм.	Менхетн	$D_{MD} \leq \bar{D} - \sigma$	Просек	ИБА ( <i>min</i> )	Укључена	0.285	0.0033
<b>18</b>	<b>Норм.</b>	<b>ИБА</b>	<b><math>S_{IBA} \geq \bar{S} + \sigma</math></b>	<b>Просек</b>	<b>ИБА (<i>min</i>)</b>	<b>Укључена</b>	<b>0.250</b>	<b>0.0028</b>

Према првој мери МАРЕ, модел #18 је остварио најбоље предвиђачке перформансе, а састоји се из *min-max* нормализације, ИБА мере сличности, *min* оператора за агрегацију и укључује препоруку експерта. Други по реду са најбољим перформансама је модел #9 који се разликује од #18 само у процедури која је коришћена за скалирање података, где је коришћена процедура стандардизације. Модел #16 је трећи најбољи по питању предвиђачких перформанси. Он се састоји из *min-max* нормализације, еуклидске удаљености, *min* оператора за агрегацију и укључује препоруку експерта. На основу ове прве мере (МАРЕ) може се закључити да постоје мале разлике у грешкама предвиђања између прва три најбоља модела која су истакнута.

Када се погледа MACR показатељ, евидентно је да су први и други модел са најбољим предвиђачким перформансама исти као и у случају претходне мере



(модел #18 и #9). Што се тиче трећег најбољег модела, MACR мера која истиче трошковну димензију, препознаје модел #4 као модел са бољим предвиђачким перформансама у поређењу са моделом #16. Другим речима, модел #4 даје мању грешку у смислу трошка жаљења.

#### 8.4. Тест значајности одабраних модела

Пре преласка на примену одабраног модела, примену и валидацију целокупног предложеног система, два статистичка теста су урађена како би се потврдила значајност модела са најбољим предвиђачким перформансама. Као што се наводи у (Derrac et al., 2011), статистички тестови су коришћени у последње време као додатак области рачунарске интелигенције како би се побољшао процес оцењивања перформанси модела и нових метода. Сходно томе, у оквиру овог рада је прво коришћен Фридманов тест (енг. *Friedman rang test*) како би се испитале разлике између средњих вредности ранга модела. Након тога, коришћен је Вилкоксон тест (енг. *Wilcoxon two-sided rank sum test*) у циљу испитивања који модели, тачније њихове перформансе, се међусобно разликују.

##### 8.4.1. Фридманов тест

У наредним Табелама 10 и 11 су прво приказани резултати Фридмановог теста.

**Табела 10.** Фридманов тест - резултати теста

N	Chi-Square	f	Asymp. Sig.
50	130.524	7	.000

**Табела 11.** Фридманов тест - средње вредности ранга

Модел	Средња вредност ранга	Модел	Средња вредност ранга	Модел	Средња вредност ранга
m1	11.07	m7	8.50	m13	7.40
m2	11.53	m8	9.82	m14	8.12
m3	13.16	m9	7.46	m15	9.30
m4	6.89	m10	11.14	m16	8.58
m5	7.16	m11	11.98	m17	9.33
m6	9.16	m12	13.18	m18	7.22

Нулта хипотеза за Фридманов тест гласи: *Не постоји разлика између средњих вредности рангова модела.* Као што је показано у Табели 10, вредност теста  $p$  је мања од 0.05, што представља и дефинисан ниво значајности у овом случају. Сходно томе, нулта хипотеза се одбацује и закључује се да се најмање два модела значајно разликују једно од другог, што представља алтернативну хипотезу. На основу Табеле 11, може се закључити да модели са највећим рангом су #4, #5 и #18.

#### 8.4.2. Вилкоксонов тест

У следећој Табели 12 су прво приказани резултати Вилкоксоновог теста, укључујући и одговарајуће  $p$  вредности теста за сваки пар модела појединачно.

**Табела 12.** Резултати Вилкоксоновог теста

	m1	m2	m3	m4	m5	m6	m7	m8	m9	m10	m11	m12	m13	m14	m15	m16	m17	m18
m1	1.000	0.255	0.002	0.000	0.000	0.566	0.001	0.083	0.000	0.912	0.131	0.001	0.000	0.001	0.546	0.003	0.044	0.000
m2		1.000	0.007	0.000	0.000	0.372	0.001	0.004	0.000	0.647	0.332	0.007	0.000	0.000	0.367	0.001	0.015	0.000
m3			1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.011	0.303	0.000	0.000	0.000	0.000	0.000	0.000
m4				1.000	0.082	0.003	0.619	0.270	0.449	0.000	0.000	0.000	0.585	0.026	0.002	0.870	0.378	0.430
m5					1.000	0.008	0.664	0.258	0.391	0.000	0.000	0.000	0.919	0.069	0.008	0.909	0.340	0.379
m6						1.000	0.075	0.149	0.009	0.222	0.102	0.000	0.013	0.049	0.303	0.049	0.117	0.009
m7							1.000	0.037	0.144	0.006	0.001	0.000	0.981	0.798	0.075	0.751	0.111	0.136
m8								1.000	0.008	0.088	0.016	0.000	0.472	0.566	0.149	0.005	0.509	0.009
m9									1.000	0.000	0.000	0.000	0.229	0.119	0.010	0.097	0.014	0.283
m10										1.000	0.117	0.001	0.000	0.000	0.226	0.001	0.036	0.000
m11											1.000	0.007	0.000	0.000	0.106	0.000	0.003	0.000
m12												1.000	0.000	0.000	0.000	0.000	0.000	0.000
m13													1.000	0.100	0.008	0.927	0.626	0.218
m14														1.000	0.037	0.606	0.725	0.112
m15															1.000	0.048	0.119	0.009
m16																1.000	0.040	0.097
m17																	1.000	0.015
m18																		1.000

Нулта хипотеза за Вилкоксонев тест гласи: *Не постоји статистички значајна разлика између редоследа два модела која се упоређују.*

Код поређења парова модела код којих је  $p$  вредност мања од 0.05, што представља и дефинисан ниво значајности у овом случају, нулта хипотеза се одбацује. Ово значи да се та два модела статистички значајно разликују, што у ствари представља алтернативну хипотезу. Резултати Вилкоксоневог теста у виду  $p$  вредности су дати у Табели 12. Прво је неопходно сагледати поређење модела са најбољим предвиђачким перформансама #18 са осталим моделима. На основу овог теста евидентно је да не постоји статистички значајна разлика у поређењу са:

- Моделима који не користе ИБА а укључују препоруку експерта (#4, #5, #7, #13 #14);
- Моделима који користе ИБА и укључују препоруку експерта (#9 и #16).

Слична је ситуација и са другим најбољим моделом #9, у погледу предвиђачких перформанси. Не постоји статистички значајна разлика у поређењу са:

- Моделима који не користе ИБА а укључују препоруку експерта (#4, #5, #7, #13 #14);
- Моделима који користе ИБА и укључују препоруку експерта (#16 и #18).

Из резултата ових тестова проистичу два закључка. Први, да постоји статистички значајна разлика између модела који укључују ИБА и препоруку експерта и свих осталих модела који не користе ИБА а укључују препоруку експерта. Други, да не постоји статистички значајна разлика између прва два најбоља модела (#18 и #16) по питању перформанси предвиђања, а оба модела укључују ИБА и препоруку експерта.

## 8.5. Примена система и валидација

Додатних 10 случајева (забележени летови у периоду октобар – новембар 2016. године) су коришћена за валидацију перформанси предвиђања четири најбоља модела из потпоглавља 8.3. За ове случајеве су такође поновљени претходно описани кораци за добијање предлога решења. Резултати валидације су дати у оквиру следеће Табеле 13.

**Табела 13.** Валидација одабраних модела

Модел #	Скалирање података	Мера сличности	Граница прихватања (ниво атрибута)	Агрег. (виши ниво)	Експертска препорука	МАРЕ	МАСР
18	Норм.	ИБА	$S_{IBA} \geq \bar{S} + \sigma$	ИБА ( <i>min</i> )	Укључена	0.2312	0.0040
9	Станд. (+ Норм. за ИБА)	ИБА	$S_{IBA} \geq \bar{S} + \sigma$	ИБА ( <i>min</i> )	Укључена	0.2386	0.0037
16	Норм.	Еуклидска	$D_{ED} \leq \bar{D} - \sigma$	ИБА ( <i>min</i> )	Укључена	0.2768	0.0047
4	Станд.	Еуклидска	$D_{ED} \leq \bar{D} - \sigma$	Просек	Укључена	0.3680	0.0024

Модел #18 је генерисао најбоље резултате предвиђања у погледу мере грешке МАРЕ која указује на најмању грешку предвиђања од 23.1%. Ова грешка је чак и мања од грешке која је добијена за овај модел у оквиру компоненте избор модела. Међутим, када се погледа други показатељ МАСР, модел #18 је остварио значајно лошије резултате, односно знатно већу грешку када су трошкови у питању. Приликом избора модела та грешка је износила 0.0028 а приликом валидације износи 0.0040. Упркос овим разликама, остварене грешке код оба показатеља у оквиру фазе валидације, и даље сврставају модел #18 у прва три модела по питању предвиђачких перформанси. Сличан случај је и са моделом #9, за који МАРЕ показатељ износи 23.9% грешке у предвиђању, а иако је и МАСР показатељ повећан, овај модел је и даље други најбољи. Модел #16 је остварио мало слабије резултате у оквиру фазе валидације, где је забележено повећање у МАРЕ показатељу са 27.1% на 27.6%, као и за МАСР показатељ са 0.0031 на 0.0047. Модел #4 је одабран и укључен у фазу валидације јер је у оквиру корака оцене перформанси модела остварио најнижи МАСР показатељ, односно најнижу грешку у предвиђању када су трошкови у питању. Ово се такође поновило и у оквиру валидације модела, међутим, како се МАРЕ показатељ значајно повећао

(са 28.8% на 36.8%), модел #4 ће и даље остати као четврти најбоље рангиран модел.

## 8.6. Поређење са моделом базираним на неуронским мрежама

Узимајући у обзир да су подаци који су коришћени у овом експерименту специфични и нису коришћени за потребе других истраживања, није оправдано упоређивати добијене резултате са резултатима који су остварили модели и решења која су дата у оквиру прегледа литературе. Како би се омогућило прикладно поређење, спроведено је још једно предвиђање броја путника који се неће појавити на лету, у оквиру експеримента мањег обима, као дела целокупног истраживања. Предвиђање, уз коришћење истог скупа података, је спроведено са неуронским мрежама. Као што је и назначено кроз преглед литературе (Faraway & Chatfield, 1998; Weatherford, Gentry & Wilamowski, 2003), решења базирана на овој техници су се показала веома успешна за примену и решавање проблема предвиђања недоласка путника на лет у авио индустрији.

За потребе овог експеримента, коришћене су неуронске мреже различитих структура са алгоритмом са пропацијом грешке уназад (енг. *backpropagation algorithm*), где је сам број неурона у оквиру скривеног слоја варирао (10, 20 и 40 неурона). Експеримент са свакој од три мреже је поновљен 10 пута, узимајући у обзир стохастичку природу неуронских мрежа. Након добијања резултата, израчунати су МАРЕ показатељи и упоређени су са МАРЕ показатељима који су израчунати за одабране моделе (из претходног потпоглавља) у оквиру предложеног система за предвиђање недоласка путника на лет. Како су првобитни резултати неуронских мрежа над скупом одабраних показатеља били знатно лошији, експеримент је потврдио закључке из претходних поглавља, да је неопходно укључити експертско знање.

Сходно томе, експеримент је био поновљен уз укључивање и обједињавање експертског знања заједно са резултатима које су генерисале неуронске мреже. Израчунате су просечне вредности резултата предвиђања који су добијени неуронским мрежама и пратећи кораке у претходно предложеном систему у оквиру поглавља 7., те вредности су обједињене са предложеним

вредностима од стране експерта. Резултати валидације података су дати у оквиру наредне Табеле 14 и упоређени са моделом #18 који је генерисао најбоље резултате предвиђања у погледу мере грешке MAPE у оквиру предложеног система за предвиђање.

**Табела 14.** Грешке предвиђања модела #18 и модела базирани на неуронским мрежама

Модел	Експертска препорука	MAPE
Модел #18 са ИБА мером сличности	Укључена	0.2312
Неуронске мреже, 10 неурона	Укључена	0.2486
Неуронске мреже, 20 неурона	Укључена	0.2248
Неуронске мреже, 40 неурона	Укључена	0.3548

У погледу добијених резултата и грешака предвиђања, евидентно је да је модел #18 са ИБА мером сличности и укљученим експертским знањем остварио знатно боље резултате од неуронске мреже са 40 неурона и укљученим експертским знањем. Грешке предвиђања друге две неуронске мреже са једноставнијим структурама (10 и 20 неурона) су доста сличне као и грешке предвиђања модела #18. Међутим, моделима који су коришћени за поређење у овом потпоглављу, а који су засновани на неуронским мрежама, недостаје интерпретација и нису толико једноставне за разумевање у погледу функционисања. Другим речима, неуронске мреже су пример система црне кутије, а њихов начин рада и покушаји интерпретације свакако нису блиски менаџменту у области авио индустрије.

У данашњем времену, интерпретација примене рачунарских технологија на реалним примерима и проблемима се сматра као изузетно значајно својство код алгоритама који се користе за предвиђање. Сходно томе, можемо закључити да предложени систем за предвиђање путника који се неће појавити на лету представља боље решење за крајњег корисника, јер користи више различитих модела који су једноставни за разумевање, а који остварују перформансе (у погледу грешке предвиђања) које су сличне и боље од перформанси које остварују сложене технике рачунарске интелигенције. Такође, модели који су саставни део предложеног система омогућавају дубље разумевање области недоласка путника на лет и могу дати резултате који даље могу да се користе како би авио компанија дефинисала одређене корективне активности.

*„My favorite things in life don't cost any money. It's really clear that the most precious resource we all have is time.“*

Steven Paul Jobs

## 9. ИМПЛЕМЕНТАЦИЈА РЕШЕЊА У ДОМЕНУ АВИО КОМПАНИЈЕ

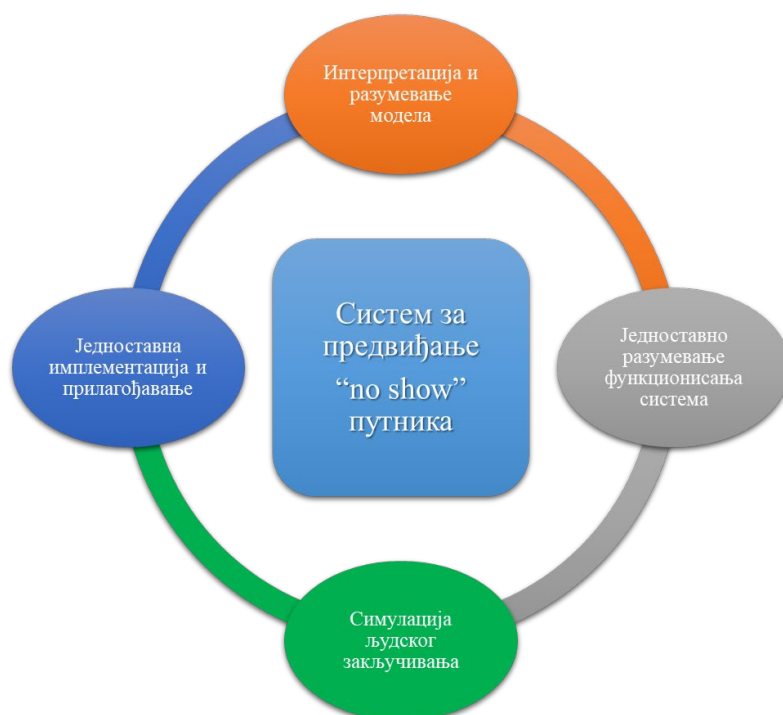
Предлог система за предвиђање путника који се неће појавити на лету је подржан са аспекта прегледа литературе и постојећих решења у истом и сличном домену. Решење се заснива на теоријским основама и комбинацији технике закључивања на основу случаја и интерполативне Булове алгебре. Након што је приказана примена система и спроведена валидација на реалном примеру, у оквиру овог поглавља је дат практичан осврт на применљивост решења у оквиру авио компаније. У потпоглављу 9.1. се полази од основних карактеристика предложеног решења, да би се у оквиру потпоглавља 9.2. идентификовао домен, односно сектор примене у авио компанији. На крају, у потпоглављу 9.3. су дате смернице за сам процес имплементације.

### 9.1. Основне карактеристике предложеног решења

Предвиђање путника који се неће појавити на лету представља веома значајан аспект за авио компанију јер помаже приликом одређивања броја „overbooking“ седишта за конкретан лет. На овај начин, авио компанија свесно омогућава продају већег броја седишта у односу на сам капацитет авиона, и то на специфичним летовима, који су у већини случаја са високом потражњом. Примена „overbooking“ принципа, дефинитивно може донети бенефите авио компанији. У зависности од различитих сценарија, који су описани у ранијим поглављима овог рада, авио компанија може остварити додатни профит који на годишњем нивоу представља значајан удео у целокупном профиту. С друге стране, потребно је сагледати и трошковну страну, која се односи на саму имплементацију решења од стране авио компаније. Посматрано чак и кроз краћи

временски период од нпр. годину дана, може се очекивати да додатни профит премашује трошкове имплементације.

Присуство информационих технологија у авио индустрији је евидентно и из године у годину се развијају нова решења заснована на модерним технологијама. Сходно томе, поред сагледавања додатног профита као главног бенефита који се добија имплементацијом предложеног решења, потребно је осврнути се и на остале карактеристике које предложени систем нуди (Слика 48).



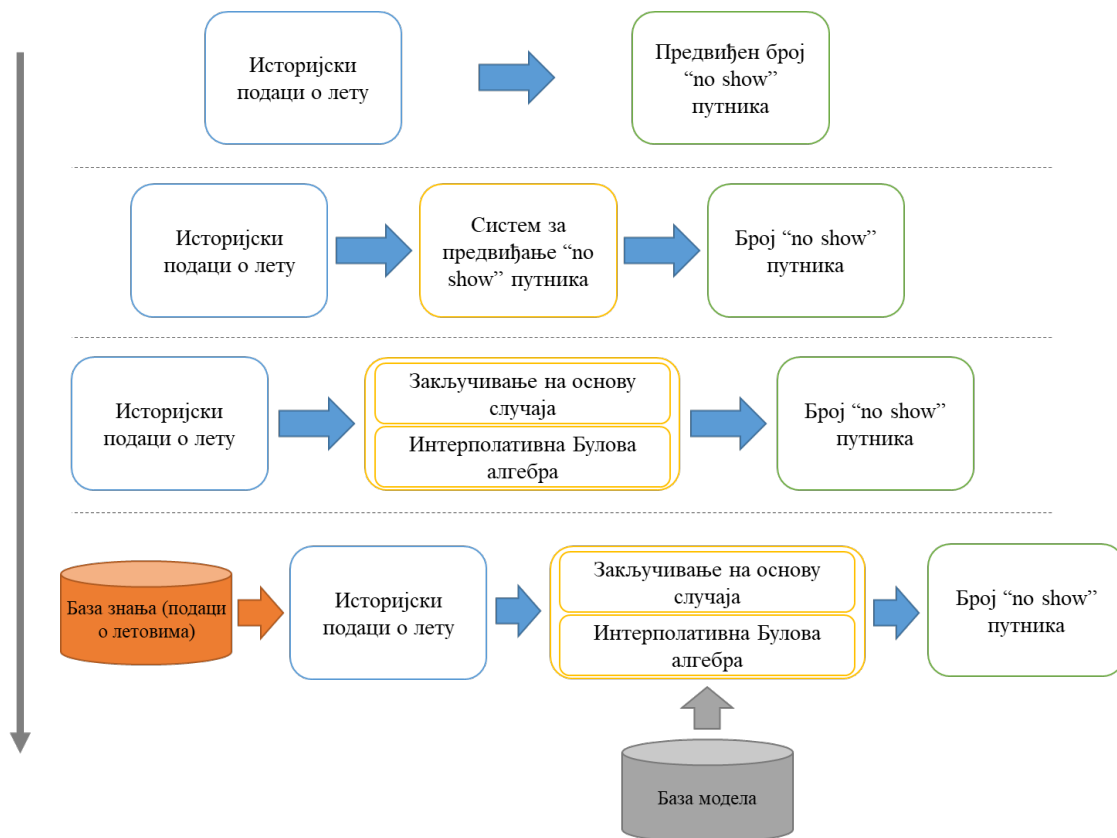
**Слика 48.** Основне карактеристике предложеног решења

Увођењем логичког аспекта кроз коришћење ИБА мере сличности и агрегације, постигнута је лакша интерпретација и разумевање модела са корисничког аспекта. Као што је и раније истакнуто, приликом поређења два лета, неопходно је сагледати подједнако атрибуте које оба лета поседују као и карактеристике које оба лета не поседују. Ово је омогућено увођењем ИБА мере сличности. Сам циљ постизања боље интерпретације и разумевања модела је лакше разумевање корисника у погледу корака који су довели до избора модела и предлога решења, односно очекиваног броја „no-show“ путника. Сваки појединачни корак технике закључивања на основу случаја се може једноставно



испратити. Резултати који су добијени у оквиру сваког корака се користе даље као улазни подаци за следећи корак, а могу бити разматрани и сагледани са различитих страна.

Предложени систем је једноставан за разумевање у погледу функционисања. Посматрано са највишег нивоа, потребно је на основу одређених историјских података о лету предвидети број „no-show“ путника. Декомпоновањем се даље историјски подаци о лету третирају као улазни елементи, сам систем који трансформише улазне податке се базира на комбинацији технике закључивања на основу случаја и интерполативне Булове алгебре, а излазни подаци представљају процене броја „no-show“ путника. Да би систем функционисао, неопходна су два предуслова - база података о летовима и база модела. База података у ствари представља записе о лету по данима и неопходно је да се узме у обзир минимални период од претходних годину дана како би се у моделовање укључио и сезонски тренд. База модела садржи потенцијалне и могуће комбинације алгоритма за предвиђање „no-show“ путника укључујући различито скалирање података, употребу различитих мера за израчунавање сличности, као и узимања у обзир препоруке експерта. Шематски приказ предложеног система за предвиђање „no-show“ путника са функционалног аспекта је представљен на следећој Слици 49.



**Слика 49.** Шематски приказ предложеног система за предвиђање са функционалног аспекта

Симулацијом људског закључивања на основу случаја из прошлости, може се рећи да предложени систем функционише на сличан начин као и људска перцепција. Ако се погледа алгоритам на коме се базира предложени систем, евидентно је да се предлог броја „no-show“ путника генерише на основу логичког комбиновања и процесирање историјских података. Коришћење података из прошлости представља једну од главних карактеристика технике закључивања на основу случаја, а сам процес закључивања се базира на посматрању свих атрибута који представљају један случај. Уколико је нови случај довољно сличан са одређеним случајем или са више случајева који су се десили у прошлости, њихова решења се разматрају као потенцијална за решавање новог случаја. Начин на који се одређује сличност између два или више случајева се заснива на логичком приступу.

На крају, са аспекта саме компаније, предложено решење се може доста једноствано имплементирати и прилагодити специфичностима и операцијама авио компаније. Један од најбитнијих корака имплементације се односи на

прикупљање података о летовима из резервационог система авио компаније. Уколико се предложени систем посматра као посебна апликација, неопходно је повезати ту апликацију са резервационим системом како би се могла формирати база података о летовима из прошлости. У зависности од специфичности операција авио компаније, могуће је узети друге атрибуте у обзир, па чак и изоставити неке од оних који су расположиви. Такође, у зависности од начина на који је организован рад, може се омогућити да више експерата допринесе предлогу решења кроз сугестије које се базирају на искуству у овом домену. Смернице за процес имплементације су дате у последњем потпоглављу 9.3.

## 9.2. Домен/сектор примене

Структура, начин функционисања и организација активности у оквиру авио компанији зависе од више фактора као што су величина летачке флоте, број дестинација и територијална покривеност (домаћи, регионални, међународни), специфичности канала продаје, па чак и укупан број запослених. Специфичност једне авио компаније се огледа у међузависности операција. Па тако пилот неће моћи да полети уколико авион претходно није припремљен и преконтролисан у погледу исправности, а супервизори неће спровести инспекцију авиона уколико претходно на авиону није урађен сервис. Аеродромско особље неће укрцати путнике у тај исти авион (иако је све претходно урађено) уколико не постоји или није у функцији резервациони систем са подацима о пријављеним путницима. Из истог резервационог система, кабинско особље мора добити информације о специфичности услуге коју је потребно пружити појединачном путнику у погледу седишта и додатних погодности на лету (приоритетно укрцавање, седиште са више простора, додатан оброк на лету и итд.).

За одржавање авиона се користе посебни програми који прате исправност делова, планирају њихово редовно одржавање и замену, а све је повезано у већини случаја са програмима који планирају и ажурирају распоред летења по дестинацијама и авионима. Ови програми су даље повезани са резервационим системом који садржи све податке о путницима, услугама које је путник купио, као и све што је потребно како би се кроз систем пријавио путник, издала карта

при доласку путника на аеродром и сам путник укрцао на лет. На основу свега овога, може се закључити да су информациони системи и њихова повезаност у веома битан фактор за функционисање авио компаније и обављање редовних операција.

Када се сагледа шта је све потребно како би авио компанија пружила крајњу услугу путнику у виду превоза, могу се у погледу организационе структуре издвојити неколико целина, сектора, односно дивизија. Базична организациона структура авио компаније се у већини случајева састоји из следећих операција и групе послова:

- Аеродромске операције
- Летачке операције
- Послови одржавања флоте
- Продаја и маркетинг
- Комерцијала и управљање резервацијама
- Информационе технологије и инфраструктура
- Финансије
- Административне, кадровске и правне операције
- Подуговарачке операције (најчешће контакт центар и рачуноводство)

Подела и груписање операција у секторе и дивизије је специфична и разликује се, као и у случају било које друге индустрије, од компаније до компаније.

Предложени систем за предвиђање броја путника који се неће појавити на лету може бити коришћен у оквиру комерцијалног или резервационог сектора авио компаније. Препоруке у виду броја “no-show“ путника се могу користити од стране особа који су задужени за ажурирање резервационог система, подешавање инвентара и „overbooking“ лимита. Пратећи тренутно присутне трендове у авио компанијама, те позиције могу бити менаџер или супервизор инвентара (енг. *Inventory manager* или *Inventory supervisor*), па чак и менаџер или супервизор сектора за комерцијална питања (енг. *Commercial manager* или *Commercial supervisor*). Наравно, могуће је целокупни процес аутоматизовати и омогућити да

се број седишта на лету у оквиру резервационог система аутоматски ажурира на основу резултата предвиђања, али је свакако неопходно да постоји људски надзор, односно особа која ће пратити дозвољене „overbooking“ лимите.

### 9.3. Смернице за процес имплементације

Како би предложени систем функционисао на очекивани начин и давао резултате, предуслов је да постоји апликација, односно програм који је повезан и комуницира са резервационим системом авио компаније. Од резервационог система добија неопходне податке о летовима, инвентару и броју доступних седишта. На основу добијених историјских и података о новом лету, сврха програма је да праћењем претходно описаног алгоритма предлаже корекцију инвентара у виду „overbooking“ лимита. Поред овога, неопходно је да експерт има опцију да унесе своју препоруку у виду очекиваног броја путника који се неће појавити на том лету у оквиру саме апликације, како би се генерисао јединствени предлог. Целокупан процес је описан на следећој Слици 50.



**Слика 50.** Повезаност предложеног система са резервационим системом

Сходно томе, како би предложени систем и сам програм могао да се употребљава, неопходно је спровести одређене кораке имплементације. Један од могућих процеса имплементације би се састојао из следећих корака:

1. Иницијално истраживање и анализа летова и броја “no-show“ путника
2. Избор летова са високом потражњом
3. Дефинисање могућности и начина за повезивање предложеног система са резервационим системом
4. Мапирање података који су потребни за функционисање предложеног система
5. Дефинисање одговорности и овлашћења експерта, односно особе која ће и пратити функционисање предложеног система и кориговати инвентар/број седишта (уносити „overbooking“ лимите)
6. Дефинисање система извештавања
7. Иницијално повезивање са резервационим системом и тестирање програма
8. Тренинг запослених (експерта) за употребу програма
9. Пуштање програма у рад
10. Праћење резултата (у прва 3 месеца) и корекције уколико је потребно
11. Редовно одржавање

На крају, сам процес имплементације је специфичан и може се разликовати у зависности од операција авио компаније, али и многих других фактора као што су начин попуњавања седишта у систему, врсте продајних канала, а и сама доступност информација о резервацијама по појединачном каналу. Када се ради о авио компанији која на годишњем нивоу превезе 2, 3 или више милиона путника, са преко 400 дестинација и неколико хиљаде летова, чување података о појединачним резервацијама или летовима може изискивати значајна улагања у инфраструктуру и системе. Из овог разлога, неретко је случај да се авио компаније опредељују на чување само одређених података, а може се рећи и да су на неки начин и лимитирани у томе. Ово може утицати на предложени систем за предвиђање путника који се неће појавити на лету кроз доступност атрибута који се користе у алгоритму за одређивање сличности.

*„As tennis players, we work and we sacrifice many things. To lose, that's not a happy thing - I mean sure, I was disappointed. You have to come back strong. But to win the last point in a grand slam tournament, that's the most beautiful and most satisfying feeling you can get as a tennis player. It's worth it.“*

Bjorn Borg

## 10. ЗАКЉУЧАК

Авио компаније послују у веома турбулентном окружењу, а кретања одређених фактора који утичу на пословање је веома тешко или скоро немогуће предвидети. Интензивним развојем информационих технологија је омогућено прикупљање великих количина података у веома кратком времену. Једно од главних питања је на који начин компаније у авио, и другим индустријама, могу искористити ове информације и донети одлуке којима ће се унапредити пословање, односно остварио профит. Прикупљање података је захтеван део процеса, у већини случајева изискује и додатне напоре, улагања ресурса и времена, али свакако није толико интензиван и сложен, као и остале активности које следе након тога. Када је реч о наредном кораку, односно обради података, авио компаније користе широк спектар решења која се темеље на примени савремених технологија. Од начина на који се подаци процесирају, зависи доношење одлуке, као и даљи кораци који се предузимају. Узевши у обзир количину података који су свакодневно доступни, изазови са којим се сусрећу авио компаније се односе на избор података које је потребно користити, као и на премошћавање препрека у виду недостатака података. Систем који је предложен у овом раду, представља резултат истраживања које је настојало да обухвати активне проблеме и изазове са којима се сусрећу компаније у авио индустрији. Пажљивом селекцијом података, као и применом техника које се заснивају на рачунарској интелигенцији, предложен је систем који може да одговори на актуелне изазове у авио индустрији, а уз одређене додатне напоре, може се генерализовати и прилагодити и применити на решавање истих проблема и у оквиру осталих видова транспорта. Основни принципи техника рачунарске интелигенције, који се односе на способност толерисања непрецизности и

могућност моделовања делимичне истинитости, су управо омогућили да се применом система који је базиран на овим техникама, предложи решење проблема које је робушно, знатно лакше за имплементацију и изискује мање трошкова.

Тема овог рада представља предлог система за предвиђање путника који се неће појавити на лету („no-show“), који се заснива на техникама рачунарске интелигенције. Предвиђање броја „no-show“ путника представља специфичан и уско формулисан проблем који је већ дужи низ година веома актуелан у авио индустрији како са теоријског, тако и са практичног аспекта. На основу очекиваног броја „no-show“ путника, као и других фактора, авио компаније доносе одлуку о додатном броју места који ће бити доступан кроз резервациони систем. На овај начин, авио компаније могу остварити додатан профит, поготову када се ради о летовима који су попуњени у потпуности.

У раду су најпре објашњени основни концепти рачунарске интелигенције, а након тога је дат преглед постојећих решења за предвиђање броја „no-show“ путника у авио индустрији, као и за оптимизацију капацитета авиона, максимално искоришћење инвентара, предвиђање броја путника и за предвиђање „overbooking“ лимита. Посматрана су и анализирана решења заснована на техникама рачунарске интелигенције, као и решења заснована на статистичким техникама, а решења која укључују људски фактор су посебно назначена. Затим су наведени основни теоријски појмови и дефиниције које се односе на интерполативну Булову алгебру, која представља основу за Буловски конзистентну фази логику. Након тога, дати су основни теоријски појмови и дефиниције који се односе на технику закључивања на основу случаја. Пре преласка на разматрање проблема непојављивања путника на лету у оквиру знатно сложенијих околности, прво је описана примена технике закључивања на основу случаја на примеру ниско буџетне авио компаније и на једноставнијем проблему предвиђања укупног броја путника на лету. Након тога, а комбинацијом претходно дефинисана два приступа, дат је предлог система за предвиђање броја „no-show“ путника и спроведена је валидација на примеру лета Београд – Амстердам. Предложени систем не може функционисати без историјских података о лету које је могуће извући из резервационог система авио компаније.



Сходно томе, у раду је дат и практичан осврт на имплементацију и применљивост решења у авио компанији.

Предложени систем пружа авио компанији свеобухватни начин за предвиђање “no-show“ путника. Систем се састоји из две компоненте: (1) избор најпрецизнијег модела и (2) примена и валидација. Модел за предвиђање се састоји из алгоритма који се заснива на техници закључивања на основу случаја и интерполативне Булове алгебре. Модел комбинује предлог који се генерише од стране алгоритма и предлог који препоручује експерт. На овај начин предложено је решење које обједињује објективну и субјективну димензију. Сличност између летова се израчунава коришћењем традиционалних мера различитости (Еуклидска и Менхетн), као и ИБА логичком мером сличности. Такође, ИБА приступ употпуњује традиционални алгоритам технике закључивања на основу случаја кроз омогућавање логичке агрегације вредности, односно моделовањем постојећих нелинеарних зависности између података.

Примена предложеног система је приказана коришћењем података о лету на релацији Београд-Амстердам. База података се састоји од записа о конкретном лету по дану за период од годину дана. Добијени резултати показују да је неопходно укључити препоруку експерта у процес предвиђања, односно само алгоритам није довољан да би се добили довољно прецизни резултати. Такође, добијени резултати указују да су модели који су засновани на ИБА приступу и који комбинују резултате алгоритма и препоруку експерта, дали прецизније резултате од модела који користе традиционалне мере за израчунавање сличности. Сходно томе, потврђено је да логички приступ моделовању сличности представља перспективан правац примене у оквиру технике закључивања на основу случаја.

Допринос овог рада се може сагледати подједнако кроз теоријску и практичну димензију. Кроз увођење логичких зависности и ИБА мере сличности као део процеса закључивања на основу случаја, добијени су бољи резултати предвиђања и постигнуто је боље тумачење и разумевање модела. Са практичне стране, предложено решење је једноставно за разумевање у погледу функционисања, и може се доста једноставно имплементирати и прилагодити специфичностима и операцијама авио компаније. Крајњи циљ са практичне стране

је повећање профита кроз боље искоришћење инвентара, односно седишта на лету. Предложени систем за предвиђање „no-show“ путника може наћи своје место у оквиру комерцијалног или резервационог сектора авио компаније. Препоруке у виду броја „no-show“ путника се користе за ажурирање резервационог система, односно за подешавање инвентара и „overbooking“ лимита на посматраном лету.

### 10.1. Осврт на постављене хипотезе и остварене доприносе

У овом делу рада дат је осврт на постављене хипотезе и остварене научне доприносе. Са гледишта постављених **помоћних хипотеза** може се закључити следеће:

- *Закључивање на основу случаја се може унапредити увођењем интерполативне Булове алгебре;*

Детаљним прегледом литературе је прво указано и утврђена је могућност унапређења процедуре технике закључивања на основу случаја применом ИБА. Затим су, кроз анализу начина функционисања технике закључивања на основу случаја, идентификована два корака у оквиру којих је могуће применити тренутно два најуспешнија правца примене ИБА. Приликом одређивање сличности између атрибута, поред традиционалних мера, коришћена је и ИБА мера сличности. У кораку избор случајева и агрегације постојећих решења, поред израчунавања просечних вредности, коришћен је и ИБА логички приступ који је реализован кроз *min* оператор. Увођење интерполативне Булове алгебре у оквиру процеса закључивања на основу случаја је приказано у 7. поглављу овог рада. Након што је дефинисан комбиновани алгоритам, у оквиру поглавља 8. је спроведена валидација и тестирање модела. Поређењем резултата свих модела указано је да се техника закључивања на основу случаја може унапредити увођењем интерполативне Булове алгебре. Модели чији се алгоритам заснива на комбинацији ИБА приступа и технике закључивања на основу случаја су дали боље резултате у погледу предвиђања “no-show“ путника.

- *Логичка агрегација и мера сличности/различитости засноване на ИБА се могу користити у предвиђање броја путника који се неће појавити на лету;*

Потреба за логичким приступом и употребом ИБА мере сличности је идентификована у потпоглављу 4.1. кроз два примера у авио индустрији. Први се односи на поређење попуњености два авиона, а други на поређење резервација путника. Специфичност проблема изискује потребу за посматрање структуре посматране променљиве, а не посматрање искључиво њене вредности. Ово се управо постиже кроз ИБА принцип структурне функционалности који фокус ставља на структуру посматране променљиве. Као што је истакнуто, ИБА је погодна за решавање проблема у оквиру којих су наизглед вредности једнаке, али се структуре посматраних променљивих разликују. Проблем предвиђања путника који се неће појавити на лету је описан у потпоглављу 5.1. где је такође указано на потребу да се посматрају и упоређују два лета код којих је забележен одређени проценат “no-show“ путника са структурног нивоа. Код оба лета проценат “no-show“ путника може бити исти, али се летови разликују по осталим атрибутима. Другим речима, вредности код два лета могу бити иста, али се по структури разликују. Сходно томе, потврђено је да се логичка агрегација и мера сличности/различитости засноване на ИБА могу користити у решавању проблема предвиђања броја “no-show“ путника.

- *Могуће је развити систем за предвиђање „no-show“ путника који комбинује технике рачунарске интелигенције и експертско знање;*

Систем који комбинује технике рачунарске интелигенције и експертско знање је приказан у петом делу овог рада. Предложени систем за предвиђање „no-show“ путника узима у разматрање предлог генерисан од стране алгоритма и предлог који се добија од стране експерта на бази знања и искуства у поменутој области. У оквиру процеса избора модела, у кораку 7. Избор случајева и њихових решења од стране експерта, је описано на који начин је експертско знање укључено у процес предвиђања. Како би се у наредном кораку примениле различите функције за агрегацију података, неопходно је да експерт одабере најмање једнак број случајева колико је одабрано и од стране алгоритма.

- *Укључивање експертског знања може да побољша резултате предвиђања предложеног система.*

Прегледом литературе у поглављу 3., је уочен одређени број радова који инсистира и указује на потребу укључивања експертског знања у оквиру процеса предвиђања у авио индустрији. У поглављу 8. овог рада је урађена валидација и тестирање расположивих модела за предвиђање “no-show“ путника. Применом различитих техника скалирања података, различитих мера за израчунавање удаљености и сличности, као и употребом различитих агрегатора, дефинисана су укупно 18 различитих модела. У оквиру дефинисаних модела, посматрају се две групе - модели који укључују препоруку експерта и модели који не укључују препоруку експерта, односно предвиђање заснивају само на алгоритму. За оцену перформанси коришћене су две мере у виду грешака предвиђања - средња апсолутна процентуална грешка и средњи апсолутни трошак жаљења. Поређењем резултата свих модела истакнуто је да се увођењем експертског знања могу побољшати резултати предвиђања предложеног система. Модели који комбинују резултате алгоритма и препоруку експерта дају боље резултате предвиђања “no-show“ путника.

Узимајући у обзир резултате остварене применом предложеног система, као и резултате предвиђања различитих модела, указано је на неопходност укључивања препоруке експерта у процес предвиђања, односно сам алгоритам није довољан да би се добили довољно прецизни резултати. Узимајући ово у обзир, као и претходно истакнуте аргументе који потврђују помоћне хипотезе може се закључити да је потврђена и **основна хипотеза**:

- *Предвиђања „no-show“ путника у авио индустрији се може унапредити коришћењем техника рачунарске интелигенције у комбинацији са експертским знањем.*

Међу оствареним научним доприносима издвајају се:

- Преглед и анализа постојећих решења за предвиђање у авио индустрији;
- Предлог новог система за предвиђање „no-show“ путника у авио индустрији;
- Примена предложеног система на примеру из праксе користећи реалне податке авио компаније надмашује постојећа традиционална решења;
- Идентификација домена/сектора/области у авио компанији у којој се предложени систем може применити;
- Друштвени допринос се огледа у лакој могућности имплементације предложеног система у различитим авио компанијама, као и могућности прилагођавања система потребама истих;

## 10.2. Могућности и правци даљег истраживања

У циљу искоришћења пуног потенцијала решења које је предложено, општа препорука је да се за даље истраживање користи, уколико је могуће због доступности, база која садржи податке о летовима за више година. Узевши то у обзир, као и закључке из претходног потпоглавља, посебно су се издвојили следећи могући правци за будуће истраживање и потенцијална унапређења модела:

### *1. Примена модела на лету са једним или више опција преседања.*

У раду је тестирана примена предложеног решења на директном лету Београд - Амстердам, односно лету без преседања. Сходно томе, један од праваца за даље истраживање би се односио на проблем предвиђања “no-show“ путника код летова са преседањем. Ови летови су специфични јер у таквим ситуацијама постоје путници који могу да се не појаве на почетку путовања али и на дестинацији која се налази између почетне и крајње дестинације. За детаљније сагледавање овог проблема могуће је посматрати следећи пример руте: Београд (почетна дестинација) -> Праг (преседање) -> Амстердам (крајња дестинација). Одређени путници се укрцавају на лет у Београду, а постоје и неки путници који си купили карту од Прага до Амстердама, па се они укрцавају на лет у Прагу. Као

што се може видети на основу овога, могуће је имати путнике у “no-show“ статусу приликом укрцавања у Београду и у Прагу.

*2. Укључивање екстерних атрибута који могу утицати на непојављивање путника на лету.*

Поред поменутих интерних атрибута, као други правац за даље истраживање и могуће унапређење иницијалног модела би се односио на укључивање екстерних атрибута-фактора као што су временски услови, елементарне непогоде, посебни јавни масовни догађаји, празници, као и гужве на аеродромима. Све ово наведено може утицати на непојављивања путника на лету. Наравно, пре укључивања додатних атрибута, неопходно је спровести анализу и испитати међузависност атрибута и зависне променљиве, односно броја “no-show“ путника.

*3. Примена модела на нискобуџетну авио компанију.*

Још један правац за даље истраживање би се односио на проширење саме примене и тестирања модела. У раду је коришћен пример лета редовне/традиционалне авио компаније Београд - Амстердам (енг. *legacy airline*). Препорука је да се размотри прилагођавање модела и примена предложеног решења на пример нискобуџетне авио компаније. Као полазна основа за разматрање овог предлога правца даље примене може бити и поглавље 6., у оквиру кога је приказана примена технике закључивања на основу случаја на примеру ниско буџетне авио компаније и на једноставнијем проблему предвиђања укупног броја путника на лету. Присутност нискобуџетних компанија је из дана у дан све већа, а све је и већи број традиционалних авио компанија које се окрећу ка овом моделу. Сходно томе, уколико се истраживање прошири у овом правцу, евидентно је да предложено решење за предвиђање „no-show“ путника може бити од интереса већем броју авио компанија.

## 11. ЛИТЕРАТУРА

- Aamodt, A. (1993). Explanation-driven retrieval, reuse, and learning of cases, In EWCBR-93: First European Workshop on Case-Based Reasoning. University of Kaiserslautern SEKI Report SR-93-12 (SFB 314) (Kaiserslautern, Germany, pp. 279-284.
- Aamodt, A. & Plaza, E. (1994). Case-based reasoning: foundational issues, methodological variations and system approaches. *AI Communications*, 7(1), 39-59.
- Agrawal, D., & Schorling, C. (1996). Market Share Forecasting: An empirical comparison of ANN and Multinomial Logit Model. *Journal of Retailing*, (72), pp. 383–407.
- Ahmed, A., H. & Poojari, C., A. (2008). An Overview of the Issues in the Airline Industry and the Role of Optimization Models and Algorithms. *The Journal of the Opr Research Society*, 59(3), 267-277.
- Aiken, M., & Bsat, M. (1999). Forecasting market trends with Neural Networks. *Information Systems Management*, pp. 42-48.
- Amaruchkul, K. & Sae-Lim, P. (2011). Airline overbooking models with misspecification. *Journal of Air Transport Management*, 17, 143-147.
- Ansuji, A.P., Camargo, M.E., Radharamanan, R., & Petty, D.G., (1996). Sales forecasting using time series and neural networks. *Computers & Industrial Engineering*, 31(1–2), pp. 421–424.
- Atli, O. & Kahraman, C. (2012). Aircraft Maintenance Planning Using Fuzzy Critical Path Analysis. *International Journal of Computational Intelligence Systems*, 5(3), 553-567.
- Bansal, A., Kauffman, R.J., & Weitz R.R. (1994). Comparing the Modelling Performance of Regression and Neural Networks as Data Quality Varies: A Business Value Approach. *Journal of Management Information System*, 10(1), pp. 11-32.
- Belobaba, P., P. (1989). Application of a Probabilistic Decision Model to Airline Seat Inventory Control. *Operations Research*, 37(2), 183-197.
- Benítez, R.B.C., Paredes, R.B.C., Lodewijks, G. & Nabais, L.J. (2013). Damp trend Grey Model forecasting method for airline industry. *Expert Systems with Applications*, 40, 4915–4921.

- Bobar, V. (2015). Possibilities to improvement of business decision making system in e-procurement: doctoral dissertation. Faculty of organizational sciences. University of Belgrade, Belgrade.
- Boyd, E., A. & Bilegan, I., C. (2003). Revenue Management and ECommerce. *Management Science, Special Issue on E-Business and Management Science*, 49(10), 1363-1386.
- Brown, F. M. (2012). *Boolean Reasoning: the Logic of Boolean Equations* (2nd ed.). Dover Publications.
- Calderon, T.G., & Cheh, J.J. (2002). A roadmap for future neural networks research in auditing and riskassessment. *International Journal of Accounting Information Systems*, 3, pp. 203–236.
- Cavalcante, R.C., Brasileiro, R.C., Souza, L.F., Nobrega, J.P., & Oliveira, A.L.I. (2016). Computational Intelligence and Financial Markets: A Survey and Future Directions, *Expert Systems with Applications*, 55, pp. 194-211. Doi: <https://doi.org/10.1016/j.eswa.2016.02.006>.
- Chang, P.C., Hsieh, J.C., Yeh, C.H. & Liu, C.H. (2006) A Case-Based Seat Allocation System for Airline Revenue Management. In: Huang D.S., Li K., Irwin G.W. (eds) *Intelligent Computing. ICIC 2006. Lecture Notes in Computer Science*, vol 4113. Springer, Berlin, Heidelberg.
- Chang, B., Kuo, C., Wu, C-H., & Tzeng G-H. (2015). Using Fuzzy Analytic Network Process to assess the risks in enterprise resource planning system implementation. *Applied Soft Computing*, 28, pp. 196-207. Doi: <https://doi.org/10.1016/j.asoc.2014.11.025>
- Chen, C.P.V., Günther, D. & Johnson, L.E. (2003). Solving for an Optimal Airline Yield Management Policy via Statistical Learning. *J of the Royal Statistical Society. Series C (Applied Statistics)*, 52(1), 19-30.
- Cheng, L. & Mengting, X. (2018). A Review of Research on Airline Passenger Volume Forecasting. 4th International Conference on Machinery, Materials and Computer (MACMC 2017). Published in *Advances in Engineering Research*, 150, 3112-3119.
- Chikr-El-Mezouar, Z. & Mohamed Hassan Gabr, M. (2011). Iterated neural network models for time series analysis. *International Journal of Statistics*, 69(2), 129-149. doi: <https://doi.org/10.1007/BF03263553>



- Chiu, C., Chiu, N.-H. & Hsu, C.-I. (2004). Intelligent aircraft maintenance support system using genetic algorithms and case-based reasoning. *Int. J. Adv. Manuf. Technol.*, 24, 440–446. doi: 10.1007/s00170-003-1707-x
- Chou, C-C., Liu, L-J., Huang, S-F., Yih, J-M. & Han, T-C. (2011). An evaluation of airline service quality using the fuzzy weighted SERVQUAL method. *Applied Soft Computing*, 11(2), 2117-2128. doi: <https://doi.org/10.1016/j.asoc.2010.07.010>
- Cichy, R.M., & Kaiser, D. (2019). Deep Neural Networks as Scientific Models, *Trends in Cognitive Sciences*, 23(4), pp. 305-317. Doi: <https://doi.org/10.1016/j.tics.2019.01.009>.
- Demirel, Ç. N. & Deveci, M. (2017). Novel search space updating heuristics-based genetic algorithm for optimizing medium-scale airline crew pairing problems. *Int.J. of Comp Intel. Sys.*, 10, 1082–1101.
- Derrac, J., García, S., Molina, D. & Herrera, F. (2011). A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms. *Swarm and Evolutionary Computation*, 1, 3-18. Doi: 10.1016/j.swevo.2011.02.002
- Deveci, M., Demirel, N. C. & Ahmetoğlu, E. (2017). Airline new route selection based on interval type-2 fuzzy MCDM: A case study of new route between Turkey - North American region destinations. *Journal of Air Transport Management*, 59, 83-99. doi: <https://doi.org/10.1016/j.jairtraman.2016.11.013>
- Deza, M. M., & Deza, E. (2009). *Encyclopedia of distances*. Springer, Dordrecht.
- Dobrić, V. (2016). *Concepts modelling using consistent fuzzy logic: doctoral dissertation*. Faculty of organizataional sciences. Univeristy of Belgrade, Belgrade.
- Dobrić, V., Kovačević, D., Petrović, B., Radojević, D., & Milošević, P. (2015). Formalization of human categorization process using Interpolative Boolean algebra. *Mathematical Problems in Engineering*, 2015, Article ID 620797.
- Dragović, I. (2016). *A consistent neuro-fuzzy inference system: doctoral dissertation*. Faculty of organizataional sciences. Univeristy of Belgrade, Belgrade.
- Dragović, I., Turajlić, N., Radojević, D. & Petrović, B. (2014). Combining Boolean consistent fuzzy logic and AHP illustrated on the web service selection problem. *Int. J. Comp. Int. Syst.* 7 (supp. 1): 84–93. doi:10.1080/18756891.2014.853935

- Dutta, S., Shekhar, S., & Wong, W.Y. (1994). Decision support in non-conservative domains: Generalization with neural networks. *Decision Support Systems*, 11, pp. 527-544.
- Engelbrecht, A. P. (2007). *Computational intelligence: an introduction*. 2nd ed. John Wiley & Sons Ltd.
- Faia, R., Pinto, T., Abrishambaf, O., Fernandes, F., Vale, Z. & Corchado, J. M. (2017). Case based reasoning with expert system and swarm intelligence to determine energy reduction in buildings energy management, *Energy and Buildings*, 155, pp. 269-281. Doi: <https://doi.org/10.1016/j.enbuild.2017.09.020>
- Fan, W. & Wang, J. (2013). A model for airline seat control considering risk and discount. *Proceedings of the 2nd Int. Conf. on Computer Science and Electronics Engineering (ICCSEE 2013)*, 0114-0117.
- Faraway, J. & Chatfield, C. (1998). Time Series Forecasting with Neural Networks: A Comparative Study Using the Airline Data. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 47(2), 231-250.
- Filder, R., Wei, Y. & Ismail, S. (2011). Evaluating the forecasting performance of econometric models of air passenger traffic flows using multiple error measures. *Int.J.of Forecasting*, 27, 902–922.
- Fildes, R., Nikolopoulos, K., Crone, S. F. & Syntetos, A. A. (2008). Forecasting and Operational Research: A Review. *The Journal of the Operational Research Society*, 59(9), 1150-1172.
- Grabisch, M. (1996). The application of fuzzy integrals in multicriteria decision making. *European Journal of Operational Research*, 89, pp. 445 - 456.
- Grammig, J., Hujer, R. & Scheidler, M. (2005). Discrete Choice Modelling in Airline Network Management. *Journal of Applied Econometrics*, 20(4), 467-486.
- Gu, Y. & Zhu, J. (2016). Study on Seat Capacity Allocation in Airline Alliance. *Advances in Social Science, Education and Humanities Research*, 63, pp. 358-364.
- Hagan, M., Demuth, H., Beale, M., & DeJesus O. (2014). *Neural Network Design* (2nd ed.). Martin Hagan.
- Hájek, P. (1998). *Metamathematics of Fuzzy Logic*. Trends in Logic 4. Dordrecht: Springer.

- Hill, T., & Remus, W. (1994). Neural network models for intelligent support of managerial decision making. *Decision Support Systems*, 11, pp. 449-459.
- Huang, Y., Ge, Y., Zhang, X. & Xu, Y. (2013). Overbooking for parallel flights with transference. *Int. J. Production Economics*, 144, pp. 582–589.
- IATA (2018). 2017 Marked by Strong Passenger Demand, Record Load Factor. Последњи пут прегледано 18.05. 2019. године на: <https://www.iata.org/pressroom/pr/Pages/2018-02-01-01.aspx>
- Jamshidi, M., (2003). Tools for Intelligent Control: Fuzzy Controllers, Neural Networks and Genetic Algorithms. *Philosophical Transactions: Mathematical, Physical and Engineering Sciences. Information, Knowledge and Technology*, 361(1809), pp. 1781-1808.
- Joseph, A., Larrain, M., & Turner, C. (2014). The Treasury Bill Rate, the Great Recession, and Neural Networks estimates of Real Business Sales. *Procedia Computer Science*, 36, pp. 227-233. Doi: 10.1016/j.procs.2014.09.084
- Kar, S., Das, S., & Ghosh, P.K., (2014). Applications of neuro fuzzy systems: A brief review and future out-line. *Applied Soft Computing*, 15, pp. 234-259. Doi: <http://dx.doi.org/10.1016/j.asoc.2013.10.014>
- Kaya, I., & Kahraman, C. (2011). Process capability analyses based on fuzzy measurements and fuzzy control charts. *Expert Systems with Applications*, 38(4), pp. 3172 - 3184. Doi: 10.1016/j.eswa.2010.09.004
- Kim, C.N., Yang, K.H., & Kim, J. (2008). Human decision-making behavior and modelling effects. *Decision Support Systems*, 45, pp. 517-527. Doi: 10.1016/j.dss.2007.06.011
- Kim, S. H. & Shin, S. W. (2000). Identifying the impact of decision variables for nonlinear classification tasks. *Expert Systems with Applications*, 18(3), pp. 201–214.
- Kolodner, J. (1993). *Case-based Reasoning*. Morgan-Kaufman.
- Kotecha K., Sanghani G. & Gambhava N. (2004) Genetic Algorithm for Airline Crew Scheduling Problem Using Cost-Based Uniform Crossover. In: Manandhar S., Austin J., Desai U., Oyanagi Y., Talukder A.K. (eds) *Applied Computing*. AACC 2004. *Lecture Notes in Computer Science*, vol 3285. Springer, Berlin, Heidelberg. doi: [https://doi.org/10.1007/978-3-540-30176-9\\_11](https://doi.org/10.1007/978-3-540-30176-9_11)

- Kovačević, D. (2014). Fuzzy causal early warning systems: doctoral dissertation. Faculty of organizational sciences. University of Belgrade, Belgrade.
- Kunnumkal, S., Talluri, K. & Topaloglu, H. (2012). A randomized linear programming method for network revenue management with product-specific no-shows. *Transportation Science*, 46(1), 90–108.
- Lalla, M., Facchinetti, G., & Mastroleo, G., (2008). Vagueness evaluation of the crisp output in a fuzzy in-ference system. *Fuzzy Sets and Systems*, 159, pp. 3297-3312. Doi: 10.1016/j.fss.2008.03.002
- Lan, Y., Ball, O. M., Karaesmen, Z. I., Zhang, X. J. & Liu, X. G. (2015). Analysis of seat allocation and overbooking decisions with hybrid information. *European J. of Operational Research*, 240, 493–504.
- Lan, Y., Gao, H., Ball, O. M. & Karaesmen, I. (2008). Revenue Management with Limited Demand Information. *Management Science*, 54(9), 1594-1609.
- Lawrence, R. D., Hong, S. J., & Cherrier, J. (2003). Passenger-based predictive modeling of airline no-show rates. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 397-406.
- Leake, D. B. *Case-based Reasoning: Experiences, Lessons and Future Directions*, (MIT Press, Cambridge, MA, 1996).
- Lim, C.P., Balas, V.E. & Do, Q. (2013). Special issue recent advances in soft computing: Theories and applications. *Journal of Intelligent & Fuzzy Systems*, 24, 415-416. doi:10.3233/IFS-2012-0562
- Lim, S.H. & Hong, Y. (2014). Fuel hedging and airline operating costs. *Journal of Air Transport Management*, 36, pp. 33-40. Doi: <https://doi.org/10.1016/j.jairtraman.2013.12.009>
- Li, W., Yu, S., Pei, H., Zhao, C. & Tian, B. (2017). A hybrid approach based on fuzzy AHP and 2-tuple fuzzy linguistic method for evaluation in-flight service quality. *Journal of Air Transport Management*, 60, 49-64. doi: <https://doi.org/10.1016/j.jairtraman.2017.01.006>
- Liu, W., Hu, G. & Li, J. (2012). Emergency resources demand prediction using case-based reasoning, *Safety Science*, 50(3), pp. 530-534. Doi: <https://doi.org/10.1016/j.ssci.2011.11.007>

- Lozano, S. (2014). Process efficiency of two-stage systems with fuzzy data. *Fuzzy Sets and Systems*, 243, pp. 36 - 49. Doi: <http://dx.doi.org/10.1016/j.fss.2013.05.012>
- Magaldi, R. V. (1994). Case-based reasoning: prospects for applications. *IEE Colloquium*, pp. 6/1–6/9.
- Mansur, A., & Kuncoro, T. (2012). Product Inventory Predictions at Small Medium Enterprise Using Mar-ket Basket Analysis Approach-Neural Networks. *Procedia Economics and Finance. International Conference on Small and Medium Enterprises Development with a Theme “Innovation and Sustainability inSME Development” (ICSMED)*, 4, pp. 312-320. Doi: 10.1016/S2212-5671(12)00346-2
- Marques, A., Lacerda, D.P., Camargo, L.F.R., & Teixeira, R.. (2014). Exploring the relationship between marketing and operations: Neural network analysis of marketing decision impacts on delivery performance. *International Journal of Production Economics*, 153, pp. 178-190. Doi: <http://dx.doi.org/10.1016/j.ijpe.2014.02.020>
- Milošević, P. (2017). IFS-IBA approach: Interpolative Boolean algebra in intuitionistic fuzzy set theory. Faculty of organizational sciences. Univeristy of Belgrade, Belgrade.
- Milošević, P., Petrović, B., Radojević, D., & Kovačević, D. (2014). A software tool for uncertainty modeling using Interpolative Boolean algebra. *Knowledge-Based Systems*, 62, 1-10.
- Milošević, P., Poledica, A., Rakićević, A., Dobrić, V., Petrović, B., & Radojević, D. (2018). IBA-based framework for modeling similarity. *International Journal of Computational Intelligence Systems*. 11(1), 206 - 218. doi:10.2991/ijcis.11.1.16
- Min, H. & Joo, S-J. (2016). A comparative performance analysis of airline strategic alliances using data envelopment analysis. *Journal of Air Transport Management*, 52, 99-110. doi: <https://doi.org/10.1016/j.jairtraman.2015.12.003>
- Moosmayer, D.C., Chong, A.Y-L., Liu, M.J., & Schuppar, B. (2013). A neural network approach to pre-dicting price negotiation outcomes in business-to-business contexts. *Expert Systems with applications*, 40, pp. 3028-3035. Doi: <http://dx.doi.org/10.1016/j.eswa.2012.12.018>
- Nam, K. & Schaefer, T. (1995). Forecasting international airline passenger traffic using neural networks. *Logistics and Transportation Review*, 31(3), :239.

- Nambisan, S.S. (2003). A decision-support tool for airline yield management using genetic algorithms. *Computer-Aided Civil and Infrastructure Engineering*, 18, pp. 214–223.
- Narukawa, Y., & Torra, V. (2007). Fuzzy measures and integrals in evaluation of strategies. *Information Sciences*, 177, pp. 4686 - 4695. Doi: 10.1016/j.ins.2007.05.010
- Nazmul, S., & Hojjat, A. (2013). *Computational Intelligence: Synergies of Fuzzy Logic, Neural Networks and Evolutionary Computing*. John Wiley & Sons. ISBN 978-1-118-53481-6.
- Perçin, S. (2017). Evaluating airline service quality using a combined fuzzy decision-making approach. *Journal of Air Transport Management*. In press. doi: <https://doi.org/10.1016/j.jairtraman.2017.07.004>
- Pineda, P.J.G., Liou, J.J.H., Hsu, C-C. & Chuang, Y-C. (2018). An integrated MCDM model for improving airline operational and financial performance. *Journal of Air Transport Management*, 68, 103-117.
- Poledica, A. (2016). A logic-based approach to similarity modeling: doctoral dissertation. Faculty of organizational sciences. University of Belgrade, Belgrade.
- Poledica, A., Bogicevic-Arsic, V. & Petrovic, B. (2010). Logical aggregation as similarity measure in case-based reasoning. In *Computational Intelligence: Foundations and Applications - Proceedings of the 9th International FLINS Conference*, pp. 585-590.
- Poledica, A., Milošević, P., Dragović, I., Radojević, D. & Petrović, B. (2013). A consensus model in group decision making based on interpolative Boolean algebra. In G. Pasi, J. Montero & D. Ciucci (Eds.), *Proceedings of the 8th conference of the European Society for Fuzzy Logic and Technology (EUSFLAT-13)* (pp. 648-654). Amsterdam: Atlantis Press.
- Poledica, A., Milošević, P., Dragović, I., Petrović, B. & Radojević, D. (2015). Modeling consensus using logic-based similarity measures. *Soft Computing*, 19(11), 3209–3219. doi: <https://doi.org/10.1007/s00500-014-1476-5>
- Radojević, D. (1999). Logical interpretation of discrete choquet integral defined by general measure. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 7(6), 577-588.

- Radojević, D. (2000). (0, 1)-valued logic: A natural generalization of Boolean logic. *Yugoslav Journal of Operations Research*, 10(2), 185-216.
- Radojević, D. (2005). Interpolative relations and interpolative preference structures. *Yugoslav Journal of Operations Research*, 15(2).
- Radojević, D. (2006). Boolean Frame is Adequate for Treatment of Gradation or Fuzziness Equally as for Two-Valued or Classical Case. In *Proceedings of the 4th Serbian-Hungarian Joint Symposium on Intelligent Systems–SISY 2006*, pp. 43-57.
- Radojević, D. (2006). Interpolative Realization of Boolean Algebra, The 8th Symposium on Neural Network Application in Electrical Engineering, Conference Publications: 201-206. doi: 10.1109/NEUREL.2006.341214
- Radojević, D. (2008a). Interpolative realization of Boolean algebra as a consistent frame for gradation and/or fuzziness. In M. Nikraves, J. Kacprzyk & L. Zadeh (Eds.), *Forging New Frontiers: Fuzzy Pioneers II* (295–317). Berlin: Springer. doi: [https://doi.org/10.1007/978-3-540-73185-6\\_13](https://doi.org/10.1007/978-3-540-73185-6_13)
- Radojević, D. (2008b). Fuzzy Set Theory in Boolean Frame, *International Journal of Computers Communications and Control* 3(S): 121-131.
- Radojević, D. (2008c). Logical aggregation based on interpolative Boolean algebra, *Mathware Soft Comp.* 15 (1), 125-141.
- Radojević, D. (2008d). Real sets as consistent Boolean generalization of classical sets. In L. Zadeh, D. Tufis, F. Filip & I. Dzitac (Eds.), *From Natural Language to Soft Computing: New Paradigms in Artificial Intelligence* (150–171). Bucharest: Editing House of Romanian Academy.
- Radojević, D. (2010). Generalizovane (realno-vrednosne) relacije poretka i ekvivalencije. In B. Forca, M. Kovač, O. Čabarkapa & D. Petrović (Eds.), *XXXVII Simpozijum o operacionim istraživanjima, SYM-OP-IS 2010* (pp. 451-454). Beograd: Medija centar Odbrana.
- Rakićević, A., Nešić, I. & Radojević, D. (2013). A novel approach to hierarchical clustering based on logical measure of dissimilarity. In N. Mladenović, G. Savić, M. Kuzmanović, D. Makajić-Nikolić & M. Stanojević (Eds.), *Proceedings of the 11th Balkan Conference on Operational Research* (pp. 147-155). Smederevo: New Press.

- Rakićević, A., Simeunović, V., Petrović, B., & Milić, S. (2018). An automated system for stock market trading based on logical clustering. *Technical Gazette - Tehnički vjesnik*, 25(4), pp. 970 - 978.
- Ramot, D., Friedman, M., Langholz, G., & Kandel, A. (2003). Complex fuzzy logic, in *IEEE Transactions on Fuzzy Systems*, 11(4), pp. 450 - 461.
- Reddy D.V.S., & Mehta R.V.K. (2019) Study on Computational Intelligence Approaches and Big Data Analytics in Smart Transportation System. In: *Soft Computing and Medical Bioinformatics*. SpringerBriefs in Applied Sciences and Technology. Springer, Singapore.
- Richard, H. (1997). Improving the quality of technical data for development case based reasoning diagnostic software for aircraft maintenance. In: *Data engineering proceedings, 13th international conference on IEEE*, 584.
- Riddington, G., L. (1987). Forecast Accuracy: The Financial Benefits to a Small Airline. *The Journal of the Operational Research Society*, 38(6), 479-485.
- Robinson, L., W. (1995). Optimal and Approximate Control Policies for Airline Booking with Sequential Nonmonotonic Fare Classes. *Operations Research*, 43(2), 252-263.
- Rothstein, M. (1984). OR and the Airline Overbooking Problem. *Operations Research*, 33(2), 237-248.
- Saridakis, K.M., & Dentsoras, A.J., (2008). Soft computing in engineering design – A review. *Advanced Engineering Informatics*, 22, pp. 202-221. Doi: 10.1016/j.aei.2007.10.001
- Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, 61, pp. 85-117. Doi: <http://dx.doi.org/10.1016/j.neunet.2014.09.003>
- Shaheen, H.I., Rashed, G.I., & Cheng, S.J. (2010) Application and comparison of computational intelligence techniques for optimal location and parameter setting of UPFC. *Engineering Applications of Artificial Intelligence*, 23 (2), pp. 203-216. doi: <https://doi.org/10.1016/j.engappai.2009.12.002>
- Skala, H.J. (1978). On many-valued logics, fuzzy sets, fuzzy logics and their applications. *Fuzzy Sets and Systems*, 1, pp. 129-149.



- Smith, K.A., & Gupta, J.N.D. (2000). Neural networks in business: techniques and applications for the operations researcher. *Computers & Operations Research*, 27, pp. 1023-1044.
- Smith, C., Leimkuhler, J.F., & Darrow, R.M. (1992). Yield management at American Airlines. *Interfaces*, 22, 8–31.
- Somboon, M. & Amaruchkul, K. (2017). Applied Two-Class Overbooking Model in Thailand's Passenger Airline Data. *The Asian Journal of Shipping and Logistics*, 33(4), 189-198.
- Sugeno, M., & Tanaka, K. (1991). Successive identification of a fuzzy model and its applications to prediction of a complex system. *Fuzzy Sets and Systems*, 42, pp. 315-334.
- Sun, Z-L., Choi, T-M., Au, K-F., & Yu, Y. (2008). Sales forecasting using extreme learning machine with applications in fashion retailing. *Decision Support Systems*, 46, pp. 411-419. Doi: 10.1016/j.dss.2008.07.009
- Suzuki, Y. (2006). The net benefit of airline overbooking. *Transportation Research Part E*, 42, pp. 1–19.
- Thompson, H.R. (1961). Statistical problems in airline reservation control. *Operations Research Quarterly*, 12, 167–185.
- Thomassey, S., & Happiette, M. (2007). A neural clustering and classification system for sales forecasting of new apparel items. *Applied Soft Computing*, 7, pp. 1177-1187.
- Tsui, K.H.W., Balli, O.H., Gilbey, A. & Gow, H. (2014). Forecasting of Hong Kong airport's passenger throughput. *Tourism Management*, 42, 62-76.
- Turajlić, N. (2014). Novel models and methods for the selection and composition of web services based on non-functional requirements: doctoral dissertation. Faculty of organizational sciences. University of Belgrade, Belgrade.
- Van Ryzin, G. & McGill, J. (2000). Revenue Management Without Forecasting of Optimization: An Adaptive Algorithm for Determining Airline Seat Protection Levels. *Management Science*, 46(6), 760-775.
- Vojtek, N. (2017). Managing business processes using soft computing techniques – a literature review, *Management: Journal Of Sustainable Business And Management*

- Solutions In Emerging Economies, 23(1), 63-71. doi: 10.7595/management.fon.2017.0023
- Vojtek, N., Petrović, B., & Milošević, P. (2021). Decision Support System for Predicting the Number of No-Show Passengers in Airline Industry, *Technical Gazette - Tehnicki Vjesnik*, 28(1), in press. Doi: 10.17559/TV-20191215144655
- Vojtek, N., Poledica, A. & Petrović, B. (2018). Statistical and soft computing techniques in airline industry – a literature review, in: *Proceedings of the XVI International Symposium SymOrg 2018: Doing Business in the Digital Age: Challenges, Approaches and Solutions*, Faculty of Organizational Sciences, Belgrade.
- Vojtek, N. & Smudja, B. (2019). Improving the passenger feedback process in airline industry. *International Journal for Traffic and Transport Engineering*, 9(2), 255 – 269. Doi: [http://dx.doi.org/10.7708/ijtte.2019.9\(2\).10](http://dx.doi.org/10.7708/ijtte.2019.9(2).10)
- Walczak S., (2001). Neural networks as a tool for developing and validating business heuristics. *ExpertSystems with Applications*, 21(1), pp. 31–36.
- Watson, I. & Marir, F. (1994). Case-based reasoning: A review. *The Knowledge Engineering Review*, 9(4), 327-354. Doi: 10.1017/S0269888900007098
- Weatherford, L., R. & Belobaba, P., P. (2002). Revenue Impacts of Fare Input and Demand Forecast Accuracy in Airline Yield Management. *The Journal of the Operational Research Society*, 53(8), 811-821.
- Weatherford, L., Gentry, T. & Wilamowski, B. (2003). Neural network forecasting for airlines: A comparative analysis. *Journal of Revenue and Pricing Management*, 1(4), 319-331. doi: <https://doi.org/10.1057/palgrave.rpm.5170036>
- Wong, W.K., & Guo, Z.X. (2010). A hybrid intelligent model for medium-term sales forecasting in fashionretail supply chains using extreme learning machines and harmony search algorithm. *International Journal of Production Economics*, 128, pp. 614-624. Doi: 10.1016/j.ijpe.2010.07.008
- Wu, C-W., & Liao, M-Y. (2014). Fuzzy nonlinear programming approach for evaluating and ranking process yields with imprecise data. *Fuzzy Sets and Systems*, 246, pp. 142 - 155. Doi: 10.1016/j.fss.2013.10.014

- Xu, X., Wang, K., Ma, W. & Lin, J. (2010). Improving the Reliability of Case-Based Reasoning Systems. *International Journal of Computational Intelligence Systems*, 3(3), pp. 256-265.
- Yazdanbakhsh, O., & Dick, S. (2018). A systematic review of complex fuzzy sets and logic. *Fuzzy Sets and Systems*, 338, pp. 1-22, Doi: <https://doi.org/10.1016/j.fss.2017.01.010>.
- Yu, M-M., Chang, Y-C. & Chen, L-H. (2016). Measurement of airlines' capacity utilization and cost gap: Evidence from low-cost carriers. *Journal of Air Transport Management*, 53, 186-198.
- Zadeh L.A., (1965). Fuzzy Sets. *Information and Control*, (8), pp. 338-353.
- Zadeh, L.A. (1972). A fuzzy-set-theoretic interpretation of linguistic hedges. *Journal of Cybernetics*, 2(3), pp. 4-34.
- Zadeh L.A., (1994). Fuzzy Logic, Neural Networks, and Soft Computing. *Communications of the ACM*, 37(3).
- Zeleznikow, J., & Nolan, J.R., (2001). Using soft computing to build real world intelligent decision support systems in uncertain domains. *Decision Support Systems*, (31), 263–285.
- Zeng, X. & Li, Y. (2015). Research on Passenger Seat Choice Behavior in Airline Revenue Management. 3rd International Conference on Management, Education, Information and Control (MEICI 2015). Atlantis Press. 1407-1413.
- Zhang, G.P., Patuwo, B.E., & Hu, M.Y. (2001). A simulation study of artificial neural networks for nonli-near time-series forecasting. *Computers & Operations Research*, 28, pp. 381-396.
- Zhang, G.P., & Qi, M. (2005). Neural network forecasting for seasonal and trend time series. *EuropeanJournal of Operational Research*, 160(2), pp. 501-514. Doi: [10.1016/j.ejor.2003.08.037](https://doi.org/10.1016/j.ejor.2003.08.037)
- Zhou, Z., Wang, B., Guo, Y., & Zhang, Y. (2019). Blockchain and Computational Intelligence Inspired Incentive-Compatible Demand Response in Internet of Electric Vehicles, in *IEEE Transactions on Emerging Topics in Computational Intelligence*, 3(3), pp. 205-216. Doi: [10.1109/TETCI.2018.2880693](https://doi.org/10.1109/TETCI.2018.2880693)

Zhu, B., & Xu, Z. (2014). A fuzzy linear programming method for group decision making with additive reciprocal fuzzy preference relations. *Fuzzy Sets and Systems*, 246, pp. 19-33.

Zuidberg, J. (2014). Identifying airline cost economies: An econometric analysis of the factors affecting aircraft operating costs. *Journal of Air Transport Management*, 40, pp. 86-95. Doi: <https://doi.org/10.1016/j.jairtraman.2014.06.007>

## 12. ПРИЛОЗИ

### 12.1. Списак слика

Слика 1. Општи модел неурона .....	13
Слика 2. Једнослојна неуронска мрежа са прослеђивањем унапред.....	14
Слика 3. Неуронска мрежа са једним скривеним слојем са прослеђивањем унапред .....	14
Слика 4. Рекурентна неуронска мрежа са скривеним слојем.....	14
Слика 5. ИБА – симболички и вредносни ниво.....	30
Слика 6. Неједнакост Авиона А у два случаја попуњености .....	34
Слика 7. Венов дијаграм дисјунктности атома Булове алгебре над атрибутима $a$ и $b$ .....	37
Слика 8. Хасеов дијаграм за Булову алгебру над атрибутима $a$ и $b$ (Radojević, 2008c).....	38
Слика 9. Приказ структуре ИБА елемената са атрибутима $a$ и $b$ .....	39
Слика 10. Структурни вектор свих елемената над атрибутима $a$ и $b$ (Radojević, 2008c).....	40
Слика 11. ГБП свих елемената Булове алгебре са два атрибута $a$ и $b$ .....	43
Слика 12. ИБА мера сличности.....	49
Слика 13. Процес закључивања на основу случаја .....	53
Слика 14. Групе фактора и њихов утицај на технику закључивања на основу случаја.....	55
Слика 15. Структура система за предвиђање „no-show“ путника .....	69
Слика 16. Процес избора модела за предвиђање „no-show“ путника .....	71
Слика 17. Ниво/граница прихватања постојећих случајева .....	75
Слика 18. Подела података према временској димензији коришћених за примену предложеног система .....	80
Слика 19. Број летова у току и ван сезоне .....	81
Слика 20. Број летова викендом и радним данима .....	82
Слика 21. Број летова по месецима .....	82

Слика 22. Просечан проценат попуњености лета по месецима .....	83
Слика 23. Број резервација у оквиру домаћег ГДС-а по месецима .....	84
Слика 24. Број резервација у оквиру међународног ГДС-а по месецима .....	84
Слика 25. Број резервација у свим осталим ГДС-овима по месецима .....	85
Слика 26. Упоредни преглед резервација у ГДС-овима по месецима .....	85
Слика 27. Укупан број резервација по ГДС-овима .....	86
Слика 28. Преглед резервација у ГДС-овима у односу на сезонски показатељ.....	86
Слика 29. Преглед резервација у ГДС-овима у односу на викенд и радне дане....	87
Слика 30. Процент „no-show“ путника у току и ван сезоне .....	88
Слика 31. Процент „no-show“ путника викендом и радним данима .....	89
Слика 32. Приказ зависности „no-show“ путника и процента попуњености лета .	89
Слика 33. Приказ зависности „no-show“ путника и резервација у оквиру домаћег ГДС-а .....	90
Слика 34. Приказ зависности „no-show“ путника и резервација у оквиру међународног ГДС-а .....	91
Слика 35. Приказ зависности „no-show“ путника и резервација у оквиру осталих ГДС-ова .....	91
Слика 36. Процент „no-show“ путника у току и ван сезоне .....	93
Слика 37. Процент „no-show“ путника викендом и радним данима .....	93
Слика 38. Приказ зависности „no-show“ путника и процента попуњености лета .	94
Слика 39. Приказ зависности „no-show“ путника и резервација у оквиру домаћег ГДС-а .....	94
Слика 40. Приказ зависности „no-show“ путника и резервација у оквиру међународног ГДС-а .....	95
Слика 41. Приказ зависности „no-show“ путника и резервација у оквиру осталих ГДС-ова .....	95
Слика 42. Број резервација у оквиру домаћег ГДС-а.....	99
Слика 43. Број резервација у оквиру међународног ГДС-а .....	100
Слика 44. Број резервација у оквиру осталих ГДС-ова .....	100
Слика 45. Број резервација у домаћем ГДС-у - Без екстремних вредности .....	102
Слика 46. Број резервација у међународном ГДС-у - Без екстремних вредности	102
Слика 47. Број резервација у осталим ГДС-овима - Без екстремних вредности..	103

Слика 48. Основне карактеристике предложеног решења.....	114
Слика 49. Шематски приказ предложеног система за предвиђање са функционалног аспекта.....	116
Слика 50. Повезаност предложеног система са резервационим системом.....	119

## 12.2. Списак Табела

Табела 1. Показатељи перформанси модела за предвиђање броја резервација на лету.....	61
Табела 2. Листа потенцијалних модела.....	97
Табела 3. Број резервација у оквиру домаћег ГДС-а .....	98
Табела 4. Број резервација у оквиру међународног ГДС-а .....	98
Табела 5. Број резервација у оквиру осталих ГДС-ова.....	99
Табела 6. Број резервација у домаћем ГДС-у - Без екстремних вредности.....	101
Табела 7. Број резервација у међународном ГДС-у - Без екстремних вредности	101
Табела 8. Број резервација у осталим ГДС-овима - Без екстремних вредности ..	101
Табела 9. Оцена перформанси модела.....	106
Табела 10. Фридманов тест - резултати теста.....	107
Табела 11. Фридманов тест - средње вредности ранга .....	108
Табела 12. Резултати Вилкоксеновог теста .....	108
Табела 13. Валидација одабраних модела.....	110
Табела 14. Грешке предвиђања модела #18 и модела базирани на неуронским мрежама.....	112

## 12.3. Списак скраћеница

ГБП	генерализовани Буловог полином
ГДС	глобални дистрибуциони систем
ГП	генерализовани производ
ИБА	интерполативна Булова алгебра
ЛА	логичка агрегација

MAPE средња апсолутна процентуална грешка (енг. *Mean absolute percentage error*)

MACR средњи апсолутни трошак жаљења (енг. *Mean absolute cost regret*).



## Биографија аутора

Никола Војтек рођен је 14. августа 1986. године у Вршцу. Основну школу „Олга Петров Радишић“ је завршио 2001. године и исте године уписује Економску средњу школу у Панчеву са истуреним одељењем у Гимназији у Вршцу коју завршава 2005. године. Основну и средњу школу завршава као носилац Вукове дипломе уз бројне стипендије на нивоу општине Вршац. Током основне и средње школе учествовао је на бројним општинским и републичким такмичењима из математике и рачуноводства.

Факултет организационих наука у Београду уписује 2008. године и опредељује се за смер Управљање квалитетом у оквиру модула Менаџмент и организација. У току основних студија био је активни члан студентске организације. Основне студије завршава у року са просечном оценом 8,96 и дипломира на тему „Увођење концепта Balanced Scorecard у издавачко предузеће“ са оценом 10.

Никола Војтек 2012. године уписује мастер академске студије на Факултету организационих наука, студијски програм Управљање квалитетом, студијска група Инжењеринг квалитета. Завршни мастер рад на тему „Примена Balanced Scorecard-а у стандардизованим системима менаџмента“ је одбранио са оценом 10.

Школске 2013/2014. године уписао је докторске студије на Факултету организационих наука, студијски програм Управљање системима. Научно-истраживачке области којима се Никола Војтек бави су: теорија система, фази логика, неуронске мреже, закључивање на основу случаја, методе предвиђања, итд. Током досадашњег рада је написао/учествовао у писању радова који су објављени у часописима или презентовани на конференцијама међународног и националног значаја.

### Радно искуство

Никола Војтек је пројект менаџер тренутно запослен у компанији Нитес. У претходних 7 година је био ангажован на различитим мултинационалним пројектима у оквиру авио и финансијске индустрије, као и ланца снабдевања.

Тренутно је ангажован на пројектима у оквиру енергетске индустрије. Фокус његовог истраживања је на развоју метода и решења за предвиђање у авио индустрији.

#### Списак објављених радова

Никола Војтек је, у сарадњи са другим ауторима, објавио следеће научне радове у часописима, као и у зборницима са домаћих и међународних конференција.

#### *Радови објављени у научним часописима међународног значаја (M23)*

- **Vojtek, N., Petrović, B., & Milošević, P. (2021).** Decision Support System for Predicting the Number of No-Show Passengers in Airline Industry, Technical Gazette - Tehnicki Vjesnik, in press. DOI: 10.17559/TV-20191215144655 (M23, IF2018: 0.644)

#### *Радови објављени у научним часописима међународног значаја (M24)*

- **Vojtek, N. (2017).** Managing business processes using soft computing techniques – a literature review. Management: Journal of Sustainable Business and Management Solutions in Emerging Economies, 23(1), 63-71. doi:10.7595/management.fon.2017.0023

#### *Радови објављени у научним часописима међународног значаја (M51)*

- **Vojtek, N. & Smudja, B. (2019).** Improving the passenger feedback process in Airline industry. International Journal for Traffic and Transport Engineering, 9 (2), 255-269. Doi: 10.7708/ijtte.2019.9(2).10

#### *Радови саопитени на скупу међународног значаја штампани у целини (M33)*

- **Vojtek, N., Poledica, A., & Petrović, B. (2018).** Statistical and soft computing techniques in airline industry – a literature review, in: Proceedings of the XVI International Symposium SymOrg 2018: Doing Business in the Digital Age:

Challenges, Approaches and Solutions, Faculty of Organizational Sciences, Belgrade, *in press*.

- Smuđa, B., & **Vojtek, N.** (2017). Protecting inventory capacity in airline industry, in: Proceedings of the III international scientific-business conference Limen 2017: Leadership & Management: Integrated Politics of Research and Innovations, Belgrade, Serbia, ISBN 978-86-80194-09-7.
- **Vojtek, N.**, & Smuđa, B. (2016). Implementing IT system: consultant vs do it yourself approach, in: Proceedings of the II international scientific-business conference Limen 2016: Leadership & Management: Integrated Politics of Research and Innovations, Belgrade, Serbia, ISBN 978-86-80194-05-9.
- **Vojtek, N.** (2015). Dealing with linguistic variables in Business Process Management using fuzzy logic, in: Proceedings of the 12th International Conference: Standardization, Prototypes and Quality: A means of Balkan Countries' Collaboration, ISBN: 978-605-83983-0-6.
- Vučinić, J., & **Vojtek, N.** (2014). Manager's Perception of the Role of Organizations for Standardization, in: Proceedings of the 11th International Conference "Standardization, Prototypes and Quality: A Means of Balkan Countries' Collaboration", University of Belgrade, Faculty of Organizational Sciences, ISBN: 978-86-7680-299-9.
- **Vojtek, N.** (2014). The evaluating of the fulfilment of management system standard requirements based on fuzzy logic, in: Proceedings of the 11th International Conference "Standardization, Prototypes and Quality: A Means of Balkan Countries' Collaboration", University of Belgrade, Faculty of Organizational Sciences, ISBN: 978-86-7680-299-9.
- **Vojtek, N.**, & Pribičević, I. (2013). Key process management in the integrated management system, in: Proceedings of the 4th International DQM conference on life cycle engineering and management (ICDQM-2013), The DQM Research Center, ISBN: 978-86-86355-14-0.
- **Vojtek, N.**, Gajić, M., & Vitošević, M. (2012). Comparative review and implementation of the standards ISO 14001 and ISO 50001, in: Levi Jakšić, M., Barjaktarović Rakočević, S. (Eds.), Proceedings of the XIII International Symposium SymOrg 2012: Innovative Management and Business Performance, Faculty of Organizational Sciences, Belgrade, ISBN: 978-86-7680-254-8.

*Радови саопштени на скупу међународног значаја штампани у изводу (M34)*

- **Vojtek, N.**, (2015). Business Processes Results Prediction Using Neural Networks, in: Proceedings of the 1st EWG-DSS International Conference on Decision Support System Technology (ICDSSST 2015), Belgrade, ISBN:978-86-7680-313-2.

## ИЗЈАВА О АУТОРСТВУ

Потписан: Никола Војтек

Број индекса: 5014/2013

Изјављујем

да је докторска дисертација под називом:

СИСТЕМ ЗА ПРЕДВИЂАЊЕ НЕДОЛАСКА ПУТНИКА НА ЛЕТ  
ЗАСНОВАН НА ТЕХНИКАМА РАЧУНАРСКЕ ИНТЕЛИГЕНЦИЈЕ

- резултат сопственог истраживачког рада;
- да предложена докторска теза у целини ни у деловима није била предложена за добијање било које дипломе према студијским програмима других факултета;
- да су резултати коректно наведени и
- да нисам кршио ауторска права и користио интелектуалну својину других лица.

Потпис докторанда

У Београду, \_\_\_\_\_

\_\_\_\_\_

## ИЗЈАВА О ИСТОВЕТНОСТИ ШТАМПАНЕ И ЕЛЕКТРОНСКЕ ВЕРЗИЈЕ ДОКТОРСКОГ РАДА

Име и презиме аутора: Никола Војтек

Број индекса: 5014/2013

Докторски студијски програм: Информациони системи и квантитативни менаџмент (изборно подручје Управљање системима)

Наслов докторске дисертације: Систем за предвиђање недоласка путника на лет заснован на техникама рачунарске интелигенције

Ментор: др Братислав Петровић, редовни професор

Изјављујем да је штампана верзија моје докторске дисертације истоветна електронској верзији коју сам предао за објављивање на порталу Дигиталног репозиторијума Универзитета у Београду.

Дозвољавам да се објаве моји лични подаци везани за добијање академског звања доктора наука, као што су име и презиме, година и место рођења и датум одбране рада.

Ови лични подаци могу се објавити на мрежним страницама дигиталне библиотеке, у електронском каталогу и у публикацијама Универзитета у Београду.

Потпис докторанда

У Београду, \_\_\_\_\_

\_\_\_\_\_

## ИЗЈАВА О КОРИШЋЕЊУ

Овлашћујем Универзитетску библиотеку „Светозар Марковић“ да у Дигитални репозиторијум Универзитета у Београду унесе моју докторску дисертацију под насловом:

### СИСТЕМ ЗА ПРЕДВИЂАЊЕ НЕДОЛАСКА ПУТНИКА НА ЛЕТ ЗАСНОВАН НА ТЕХНИКАМА РАЧУНАРСКЕ ИНТЕЛИГЕНЦИЈЕ

која је моје ауторско дело.

Дисертацију са свим прилозима предао сам у електронском формату погодном за трајно архивирање.

Моју докторску дисертацију похрањену у Дигиталном репозиторијуму Универзитета у Београду и доступну у отвореном приступу могу да користе сви који поштују одредбе садржане у одабраном типу лиценце Креативне заједнице (Creative Commons) за коју сам се одлучио.

1. Ауторство (CC BY);
2. Ауторство – некомерцијално (CC BY-NC);
3. Ауторство – некомерцијално – без прерада (CC BY-NC-ND);
4. Ауторство – некомерцијално – делити под истим условима (CC BY-NC-SA);
5. Ауторство – без прерада (CC BY-ND);
6. Ауторство – делити под истим условима.

(Молимо да заокружите само једну од шест понуђених лиценци.  
Кратак опис лиценци је саставни део ове изјаве).

Потпис докторанда

У Београду, \_\_\_\_\_

\_\_\_\_\_

1. Ауторство. Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце, чак и у комерцијалне сврхе. Ово је најслободнија од свих лиценци.
2. Ауторство – некомерцијално. Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца не дозвољава комерцијалну употребу дела.
3. Ауторство – некомерцијално – без прерада. Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, без промена, преобликовања или употребе дела у свом делу, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца не дозвољава комерцијалну употребу дела. У односу на све остале лиценце, овом лиценцом се ограничава највећи обим права коришћења дела.
4. Ауторство – некомерцијално – делити под истим условима. Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце и ако се прерада дистрибуира под истом или сличном лиценцом. Ова лиценца не дозвољава комерцијалну употребу дела и прерада.
5. Ауторство – без прерада. Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, без промена, преобликовања или употребе дела у свом делу, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца дозвољава комерцијалну употребу дела.
6. Ауторство – делити под истим условима. Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце и ако се прерада дистрибуира под истом или сличном лиценцом. Ова лиценца дозвољава комерцијалну употребу дела и прерада. Слична је софтверским лиценцама, односно лиценцама отвореног кода.