

# Predviđanje ponašanja pješaka u blizini pješačkih prijelaza namijenjena optimalnom upravljanju brzinom autonomnog vozila

---

**Bićanić, Ana**

**Master's thesis / Diplomski rad**

**2023**

*Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj:* **University of Zagreb, Faculty of Mechanical Engineering and Naval Architecture / Sveučilište u Zagrebu, Fakultet strojarstva i brodogradnje**

*Permanent link / Trajna poveznica:* <https://urn.nsk.hr/urn:nbn:hr:235:267578>

*Rights / Prava:* [Attribution-NonCommercial-ShareAlike 4.0 International/Imenovanje-Nekomercijalno-Dijeli pod istim uvjetima 4.0 međunarodna](#)

*Download date / Datum preuzimanja:* **2025-02-20**

*Repository / Repozitorij:*

[Repository of Faculty of Mechanical Engineering and Naval Architecture University of Zagreb](#)



SVEUČILIŠTE U ZAGREBU  
FAKULTET STROJARSTVA I BRODOGRADNJE

# DIPLOMSKI RAD

**Ana Bićanić**

Zagreb, 2023.

SVEUČILIŠTE U ZAGREBU  
FAKULTET STROJARSTVA I BRODOGRADNJE

# DIPLOMSKI RAD

Mentor:

Doc. dr. sc. Branimir Škugor, mag. ing.

Student:

Ana Bićanić

Zagreb, 2023.

Izjavljujem da sam ovaj rad izradila samostalno koristeći znanja stečena tijekom studija i navedenu literaturu.

Zahvaljujem profesoru i mentoru, doc. dr. sc. Branimiru Škugoru, na utrošenom vremenu, pruženoj pomoći, na svim savjetima i smjernicama pri izradi ovoga rada.

Zahvaljujem svojoj obitelji, najviše svojim roditeljima, koji su mi omogućili ovaj studij i neprekidno me podržavali i ohrabrivali tijekom cijelog studija.

Također zahvaljujem dečku Josipu, svojim prijateljima i kolegama na pruženoj podršci i pomoći tijekom studiranja.

Ana Bićanić



SVEUČILIŠTE U ZAGREBU  
FAKULTET STROJARSTVA I BRODOGRADNJE



Središnje povjerenstvo za završne i diplomske ispite  
Povjerenstvo za diplomske ispite studija strojarstva za smjerove:  
Proizvodno inženjerstvo, inženjerstvo materijala, industrijsko inženjerstvo i menadžment,  
mehatronika i robotika, autonomni sustavi i računalna inteligencija

Sveučilište u Zagrebu Fakultet strojarstva i brodogradnje	
Datum: 11. 11. 2023.	Prilog
Klasa: 602 - 04 / 23 - 6 / 1	
Ur.broj: 15 - 23 - 466	

## DIPLOMSKI ZADATAK

Student: **Ana Bičanić**

JMBAG: 0035215772

Naslov rada na hrvatskom jeziku: **Predviđanje ponašanja pješaka u blizini pješačkih prijelaza namijenjena optimalnom upravljanju brzinom autonomnog vozila**

Naslov rada na engleskom jeziku: **Prediction of pedestrian crossing behaviour at vicinity of crosswalks aimed for optimal speed control of autonomous vehicle**

Opis zadatka:

Iako su brojni zadaci autonomne vožnje niže razine u dobroj mjeri riješeni, autonomna vožnja koja iziskuje neki vid interakcije s drugim sudionicima u prometu kao što su primjerice pješaci i biciklisti još uvijek predstavlja značajan izazov. Zbog kompleksnosti zadatka, predikcija ponašanja ostalih sudionika u prometu u sklopu upravljačke strategije autonomnog vozila od neophodne je važnosti za postizanje sigurne i udobne vožnje. U sklopu diplomskog rada stoga je potrebno:

- provesti pregled literature na temu predikcije ponašanja pješaka u blizini pješačkih prijelaza bez semafora
- proučiti postojeće simulacijsko okruženje, te provesti usporednu analizu pripadnih modela pješaka
- postaviti računalno efikasan predikcijski model ponašanja pješaka koji bi bio prikladan za adaptaciju u stvarnom vremenu (primjerice temeljen na logističkoj regresiji)
- predložiti metodu adaptacije parametara predikcijskog modela koja bi iziskivala minimalnu količinu potrebnih podataka
- demonstrirati funkcionalnost adaptacije predloženog modela u navedenom simulacijskom okruženju za slučaj vođenja autonomnog vozila u otvorenoj petlji
- razmotriti integraciju predloženog predikcijskog modela s postojećom optimalnom upravljačkom strategijom autonomnog vozila.

U radu je potrebno navesti korištenu literaturu i eventualno dobivenu pomoć.

Zadatak zadan:

Datum predaje rada:

Predviđeni datumi obrane:

28. rujna 2023.

30. studenoga 2023.

4. – 8. prosinca 2023.

Zadatak zadao:

Predsjednik Povjerenstva:

Doc. dr. sc. Branimir Škugor

Prof. dr. sc. Ivica Garašić

## SADRŽAJ

SADRŽAJ .....	I
POPIS SLIKA .....	II
POPIS TABLICA.....	IV
POPIS OZNAKA .....	V
POPIS KRATICA .....	VII
SAŽETAK.....	VIII
SUMMARY .....	IX
1. UVOD.....	1
2. PREGLED METODA PREDVIĐANJA PONAŠANJA PJEŠAKA NA PJEŠAČKOM PRIJELAZU BEZ SEMAFORA .....	5
2.1. Metoda temeljena na klasifikacijskoj metodi <i>Random Forest</i> [12] .....	5
2.2. Metoda temeljena na procjeni relevantnosti značajki [17] .....	9
2.2.1. Odabrane značajke .....	10
2.2.2. Evaluacija.....	11
2.2.2.1. Grupna eliminacija.....	11
2.2.2.2. Pojedinačna eliminacija .....	13
2.2.2.3. Evaluacija najboljeg skupa značajki .....	13
3. POSTOJEĆE SIMULACIJSKO OKRUŽENJE .....	16
3.1. Temeljne karakteristike simulacijskog okruženja.....	16
3.2. Generiranje simulacijskih podataka namijenjenih treniranju i testiranju predikcijskog modela.....	18
3.3. Početna analiza simulacijskih podataka .....	19
3.4. Različiti simulacijski modeli pješaka.....	21
4. RAČUNALNO EFIKASAN PREDIKCIJSKI MODEL PONAŠANJA PJEŠAKA TEMELJEN NA LOGISTIČKOJ REGRESIJI .....	24
4.1. Općenito o logističkoj regresiji .....	24
4.1.1. Logistička (sigmoidna) funkcija .....	25
4.1.2. Funkcija troška.....	26
4.1.3. Gradijentni spust .....	27
4.2. Predikcijski model temeljen na logističkoj regresiji .....	27
4.2.1. Analiza rezultata .....	28
4.2.2. Filtracija podataka.....	46
4.2.2.1. Analiza filtriranih podataka .....	47
5. INTEGRACIJA PREDIKCIJSKOG MODELA S POSTOJEĆOM UPRAVLJAČKOM STRATEGIJOM.....	55
6. ZAKLJUČAK.....	58
LITERATURA.....	59

**POPIS SLIKA**

Slika 1 Ilustracija komunikacije autonomnog vozila s centralnim sustavom na oblaku [8] .....	4
Slika 2 Segmentirano raskrižje gdje plava boja predstavlja nogostup, a crvena cestu [12].....	6
Slika 3 Rezultati klasifikacije za putanje koje se križaju [12] .....	8
Slika 4 Rezultati klasifikacije za putanje koje se ne križaju [12].....	8
Slika 5 Pješački prijelaz s trajektorijama [17].....	10
Slika 6 Točnost klasifikacije spram broja pojedinačnih značajki [17].....	13
Slika 7 Tipične trajektorije pješaka, teško razdvojive [17].....	14
Slika 8 Tipične trajektorije pješaka, lako razdvojive [17] .....	15
Slika 9 Tipične trajektorije pješaka koji prelaze preko pješačkog prijelaza [17].....	15
Slika 10 Prikaz simulacijskog okvira za jedno vozilo/jedan pješak [21].....	16
Slika 11 Grafički prikaz modela za umjerenog pješaka .....	20
Slika 12 Udio vjerojatnosti realiziranog ishoda .....	20
Slika 13 Grafički prikaz modela (lijevo), udio vjerojatnosti realiziranog ishoda za jako perturbiranog pješaka (desno) .....	21
Slika 14 Grafički prikaz modela (lijevo), udio vjerojatnosti realiziranog ishoda za konzervativnog pješaka (desno) .....	22
Slika 15 Grafički prikaz modela (lijevo), udio vjerojatnosti realiziranog ishoda za agresivnog pješaka (desno) .....	22
Slika 16 Ilustracija klasifikacije pomoću logističke regresije [25] .....	24
Slika 17 Sigmoidna funkcija [26].....	26
Slika 18 Usporedba točnosti nakon 1000 prolazaka na skupu za treniranje .....	29
Slika 19 Usporedba točnosti nakon 5000 (lijevo) i 10000 (desno) prolazaka na skupu za treniranje.....	30
Slika 20 Usporedba točnosti nakon 1000 prolazaka na skupu za testiranje.....	30
Slika 21 Usporedba točnosti nakon 5000 (lijevo) i 10000 (desno) prolazaka na skupu za testiranje .....	30
Slika 22 Usporedba funkcije troška nakon 1000 prolazaka na skupu za treniranje .....	31
Slika 23 Usporedba funkcije troška nakon 5000 (lijevo) i 10000 prolazaka (desno) na skupu za treniranje .....	31
Slika 24 Usporedba funkcije cilja nakon 1000 prolazaka na skupu za testiranje .....	32
Slika 25 Usporedba funkcije troška nakon 5000 (lijevo) i 10000 prolazaka (desno) na skupu za testiranje.....	32
Slika 26 Usporedba $\alpha + \beta 1$ parametara nakon 1000 prolazaka.....	33
Slika 27 Usporedba $\alpha + \beta 1$ parametara nakon 5000 (lijevo) i 10000 (desno) prolazaka .....	33
Slika 28 Usporedba $\beta 2$ parametara nakon 1000 prolazaka .....	33
Slika 29 Usporedba $\beta 2$ parametara nakon 5000 (lijevo) i 10000 (desno) prolazaka .....	34
Slika 30 Usporedba $\beta 3$ parametara nakon 1000 prolazaka .....	34
Slika 31 Usporedba $\beta 3$ parametara nakon 5000 (lijevo) i 10000 (desno) prolazaka .....	34
Slika 32 Usporedba točnosti nakon 1000 prolazaka na skupu za treniranje .....	35
Slika 33 Usporedba točnosti nakon 1000 prolazaka na skupu za testiranje.....	36
Slika 34 Usporedba funkcije troška nakon 1000 prolazaka na skupu za treniranje .....	36
Slika 35 Usporedba funkcije cilja nakon 1000 prolazaka na skupu za testiranje .....	37
Slika 36 Usporedba $\alpha + \beta 1$ parametara nakon 1000 prolazaka.....	37
Slika 37 Usporedba $\alpha + \beta 1$ parametara nakon 5000 (lijevo) i 10000 (desno) prolazaka .....	38
Slika 38 Usporedba $\beta 2$ parametara nakon 1000 prolazaka .....	38
Slika 39 Usporedba $\beta 3$ parametara nakon 1000 prolazaka .....	38
Slika 40 Usporedba točnosti nakon 1000 prolazaka na skupu za treniranje .....	39
Slika 41 Usporedba točnosti nakon 1000 prolazaka na skupu za testiranje.....	40

Slika 42 Usporedba funkcije troška nakon 1000 prolazaka na skupu za treniranje .....	40
Slika 43 Usporedba funkcije cilja nakon 1000 prolazaka na skupu za testiranje .....	41
Slika 44 Usporedba $\alpha + \beta 1$ parametara nakon 1000 (lijevo) i 5000 (desno) prolazaka .....	41
Slika 45 Usporedba $\beta 2$ parametara nakon 1000 (lijevo) i 5000 (desno) prolazaka .....	42
Slika 46 Usporedba $\beta 3$ parametara nakon 1000 (lijevo) i 5000 (desno) prolazaka .....	42
Slika 47 Usporedba točnosti nakon 1000 prolazaka na skupu za treniranje .....	43
Slika 48 Usporedba točnosti nakon 1000 prolazaka na skupu za testiranje .....	43
Slika 49 Usporedba funkcije troška nakon 1000 prolazaka na skupu za treniranje .....	44
Slika 50 Usporedba funkcije cilja nakon 1000 (lijevo) i 10000 (desno) prolazaka na skupu za testiranje .....	44
Slika 51 Usporedba $\alpha + \beta 1$ parametara nakon 1000 (lijevo) i 10000 (desno) prolazaka .....	45
Slika 52 Usporedba $\beta 2$ parametara nakon 1000 (lijevo) i 10000 (desno) prolazaka .....	45
Slika 53 Usporedba $\beta 3$ parametara nakon 1000 (lijevo) i 10000 (desno) prolazaka .....	45
Slika 54 Točnost treniranje (lijevo) u odnosu na testiranje (desno) .....	47
Slika 55 Funkcija troška treniranje (lijevo) u odnosu na testiranje (desno) .....	47
Slika 56 Broj filtriranih točaka nakon 1. iteracije .....	48
Slika 57 Broj filtriranih točaka nakon 5. iteracije .....	49
Slika 58 Broj filtriranih točaka nakon 10. iteracije .....	49
Slika 59 Broj filtriranih točaka nakon 15. iteracije .....	50
Slika 60 Broj filtriranih točaka nakon 20. iteracije .....	50
Slika 61 Točnost treniranje (lijevo) u odnosu na testiranje (desno) .....	51
Slika 62 Funkcija troška treniranje (lijevo) u odnosu na testiranje (desno) .....	51
Slika 63 Broj filtriranih točaka nakon 1. iteracije .....	52
Slika 64 Broj filtriranih točaka nakon 5. iteracije .....	52
Slika 65 Broj filtriranih točaka nakon 10. iteracije .....	53
Slika 66 Broj filtriranih točaka nakon 15. iteracije .....	53
Slika 67 Broj filtriranih točaka nakon 20. iteracije .....	54
Slika 68 Organizacija strategije sigurne kontrole brzine [21] .....	55
Slika 69 Ilustracija potencijalnih scenarija preko vremenskog horizonta u ovisnosti o pješakovoju odluci [28] .....	57



---

**POPIS TABLICA**

Tablica 1 Rezultat eliminacije prve skupine s odgovarajućom ukupnom klasifikacijskom točnošću i lažno pozitivnom stopom [17] .....	12
Tablica 2 Rezultat grupne eliminacije temeljene na vremenskom odmaku s odgovarajućom ukupnom klasifikacijskom točnošću i lažno pozitivnom stopom [17].....	12
Tablica 3 Ocjena najboljeg skupa značajki: ukupan broj pozitivnih i negativnih trajektorija za diskretizirane mjere udaljenosti <i>dtcross</i> i <i>dtcurb</i> [17].....	14
Tablica 4 Parametri simulacijskog modela pješaka. ....	18
Tablica 5 Parametri ostalih modela.....	21
Tablica 6 Usporedba točnosti .....	23
Tablica 7 Tablični pregled analize grafova .....	46

## POPIS OZNAKA

Oznaka	Jedinica	Opis
$a$	-	međuvarijabla točnosti
$a_v$	$m/s^2$	ubrzanje vozila
$\alpha$	-	stopa učenja
$\alpha, \beta_{1,2,3}$	-	parametri modela pješaka
$c$	-	klasa
$d_{cross}$	-	minimalna udaljenost pješaka do pješačkog prijelaza
$d_{curb}$	-	minimalna ortogonalna udaljenost pješaka do najbližeg ruba
$D_p$	-	binarna varijabla odluke pješaka o prelasku preko pješačkog prijelaza
$d_{ped,car}$	m	udaljenost između automobila i pješaka
$e$	-	Baza prirodnog logaritma
$h_\theta$	-	hipoteza modela
$\hat{t}_{p,pd}$	-	vremenski korak u kojem pješak donosi odluku
$\hat{t}_{v,en}$	-	vrijeme ulaska vozila na pješački prijelaz
$J$	-	ciljna funkcija
$J_{Act1(2,3)}$	-	ciljna funkcija pojedinih upravljačkih akcija
$k$	-	diskretni vremenski korak
$l_{cross}$	m	duljina zone pješačkog prijelaza
$l_p$	m	širina zone pješačkog prijelaza
$l_v$	m	duljina pješačkog prijelaza i vozila
$l_{veh}$	m	duljina vozila
$m$	-	broj parametara
$p_{actual,i}$	-	vjerojatnost realiziranog ishoda
$p_{cross}$	-	vjerojatnost prelaska pješaka
$p_i$	-	vjerojatnost prelaska za trenutnu iteraciju
$p_{yield}$	-	vjerojatnost ne prelaska pješaka
$s_{car}$	m	prijeđena udaljenost automobila
$s_p$	m	položaj pješaka
$s_{p,0}$	m	inicijalni položaj pješaka
$s_{ped}$	m	prijeđena udaljenost pješaka
$s_v$	m	položaj vozila
$s_{v,0}$	m	inicijalni položaj vozila
$s_x$	-	status agenta
$t$	-	oznaka jednog stabla u šumi (kod Random Forest metode klasifikacije)
$\Delta T$	s	vrijeme diskretizacije
$\theta$	-	parametri predikcijskog modela pješaka

---

$U_{ped}$	-	međuvarijabla u logističkoj regresiji
$v$	-	skup ulaznih značajki
$v_{car}$	m/s	brzina automobila
$v_p$	m/s	brzina pješaka
$v_{p,0}$	m/s	inicijalna brzina pješaka
$v_{ped}$	m/s	brzina pješaka
$v_{rel}$	m/s	relativna brzina
$v_v$	m/s	brzina automobila
$v_{v,pr}$	m/s	preferirana brzina vozila
$w_a$	-	težinski faktor u ciljnim funkcijama upravljačke strategije
$X$	-	ulazni parametri
$x(y)_{car}$	m	prijeđena udaljenost automobila u 2D koordinatnom sustavu
$x(y)_{ped}$	m	prijeđena udaljenost pješaka u 2D koordinatnom sustavu
$y$	-	realizirani ishod interakcije vozila i pješaka, klasa
$\frac{\partial}{\partial \theta_j}$	-	derivacija po parametru predikcijskog modela pješaka

**POPIS KRATICA**

---

<b>Oznaka</b>	<b>Opis</b>
AV	Autonomno vozilo / Autonomous vehicle
CV	Cross validation
DARPA	Defense Advanced Research Projects Agency
MPC	Model predictive control
OTA	Over the air
RF	Random forest
RFE	Recursive feature elimination
SVM	Support vector machines
TP	True positive
TN	True negative

**SAŽETAK**

Brojni zadaci niže razine autonomne vožnje mogu se smatrati u dobroj mjeri riješenima. Međutim, kada se radi o autonomnoj vožnji koja iziskuje neki oblik interakcije s drugim sudionicima u prometu, kao što su biciklisti i pješaci, još uvijek postoje značajni izazovi koje treba riješiti. S obzirom na kompleksnost zadatka, predviđanje ponašanja ostalih sudionika u prometu u sklopu upravljačke strategije autonomnog vozila od neophodne je važnosti za postizanje sigurne i udobne vožnje. Na tom tragu u diplomskom radu se razmatra računalno učinkovit predikcijski model ponašanja pješaka u blizini pješačkih prijelaza bez semafora, koji je prikladan za adaptaciju u stvarnom vremenu. U radu je predložena metoda stohastičkog filtriranja podataka snimljenih tijekom svake interakcije vozila i pješaka, čija je glavna svrha propuštanje samo najinformativnijih podataka za potrebe adaptacije modela, te se na taj način ostvaruje minimiziranje prometa podataka ako bi se adaptacija modela provodila u oblaku. Demonstracija funkcionalnosti predloženog modela te njegova adaptacija provedena je u postojećem simulacijskom okruženju za slučaj jednog vozila i jednog pješaka, pri čemu je autonomno vozilo vođeno u otvorenoj petlji.

**Ključne riječi:** autonomna vožnja, pješak, pješački prijelaz, predikcijski model, logistička regresija, simulacija.

---

**SUMMARY**

A number of tasks of lower level of autonomous driving can be considered to a good extent solved. However, when it comes to autonomous driving that requires some form of interaction with other road users, such as cyclists and pedestrians, there are still significant challenges to be addressed. Given the complexity of the task, predicting the behavior of other road users as a part of autonomous vehicle's management strategy is essential for achieving safe and comfortable driving. On this track, the thesis considers a computationally efficient predictive model of pedestrian behavior near pedestrian crossings without traffic lights, which is suitable for real-time adaptation. The paper proposes a method of stochastic filtering of data recorded during each vehicle-pedestrian interaction, the main purpose of which is to pass only the most informative data for the purpose of model adaptation, thus minimizing data traffic if the model adaptation is carried out in the cloud. The demonstration of the functionality of the proposed model and its adaptation was carried out in the existing simulation environment for the case of one vehicle and one pedestrian, where an autonomous vehicle was driven in an open loop.

**Keywords:** autonomous driving, pedestrian, crosswalk, prediction model, logistic regression, simulation.

## 1. UVOD

Autonomna vozila (AV) su računalno upravljana vozila koja se na temelju raspoznavanja okoline mogu kretati samostalno, uz ograničeni utjecaj ili bez utjecaja vozača [1]. Ultimativna razina autonomne vožnje bi značila da čovjek nije dužan preuzeti kontrolu nad vozilom ni u kojem trenutku, niti je njegovo prisustvo u vozilu uopće potrebno. Tijekom posljednjih nekoliko godina, zbog načina i dinamike života i poslovanja, interes za autonomna vozila značajno je porastao. Razvoj AV predstavlja iznimno intrigantan niz tehnoloških inovacija i razvoja, a povijest ovih vozila ima svoje korijene još u proteklim desetljećima. Koncept AV može se pratiti unazad do 1920-ih godina kada je Francis Houdina, inženjer elektrotehnike, demonstrirao jedan od najranijih prototipova [2]. Bio je to radio-upravljani automobil, koji je mogao pokrenuti motor, mijenjati brzine i potrubiti. Ovaj automobil nudio je pogled u budućnost AV, ali nije bio u potpunosti autonoman. Zatim je na Svjetskom sajmu 1939. godine, General Motors predstavio prvi model samovozećeg automobila [2]. Bio je to električni automobil koji se kretao putem radio-kontroliranih elektromagnetskih polja, pokretan magnetiziranim metalnim šiljcima ugrađenim u kolnik. Ova inovacija postala je stvarnost 1958. godine, kada je vozilo opremljeno sensorima koji su mogli detektirati struju kroz žicu ugrađenu u cestu. Pomoću te struje moglo se upravljati volanom, omogućujući vozilu autonomno kretanje lijevo ili desno. Sljedeća važna prekretnica bila je 1961. godine kada je James Adams stvorio Stanford Cart, opremljen kamerama i programiran za autonomno otkrivanje i praćenje linije na tlu. Ovo je bila prva upotreba kamera u AV, koja je vitalni element današnjih AV. 1977. godine ta je ideja poboljšana sustavom kamera koji je podatke prenosio na računalo za obradu slika ceste što je dovelo do testiranja prvog samovozećeg putničkog vozila na svijetu koje je moglo postići brzinu do 20 milja na sat. Devedesetih godina neuronske mreže za obradu slike počele su se integrirati i na upravljačke sustave vozila. U ranim 2000-ima, industrija samovozećih automobila bila je u punom zamahu jer je tih godina agencija Ministarstva obrane SAD-a, DARPA (*eng. - Defense Advanced Research Projects Agency*), sponzorirala niz izazova kako bi se ubrzao razvoj takvih automobila [2].

Napredak kakav poznajemo danas u području AV ostvaren je zahvaljujući kombiniranom unaprjeđenju računalne moći ugradbenih računala, umjetne inteligencije, strojnog učenja i senzorske tehnologije. Današnja se AV oslanjaju na senzore, izvršne elemente, složene algoritme, sustave strojnog učenja i moćne procesore za izvršavanje softvera. Stvaraju i

održavaju kartu svoje okoline na temelju niza senzora koji se nalaze u različitim dijelovima vozila. Radarski senzori prate položaj vozila u blizini. Video kamere otkrivaju semafore, čitaju prometne znakove, prate druga vozila i traže pješake. Lidar senzori, za otkrivanje svjetla i određivanje dometa, odbijaju pulseve svjetlosti od okoline automobila kako bi izmjerili udaljenosti, otkrili rubove ceste i identificirali oznake na trakama. Ultrazvučni senzori u kotačima detektiraju rubnjake i druga vozila prilikom parkiranja. Sofisticirani softver zatim obrađuje sav taj senzorski unos, iscrtava putanju i šalje upute aktuatorima automobila koji kontroliraju ubrzanje, kočenje i skretanje vozila. Upravo zbog kompleksnosti takvih sustava, postoje i mnogi izazovi koji sežu od zakonodavnih i ekoloških do etičkih i tehnoloških. Zakonodavni problemi su zakonski propisi, koji se mogu razlikovati od države do države. U Sjedinjenim Američkim Državama su neke države predložile „porez na milju“ za AV kako bi spriječile porast „zombi automobila“ koji se voze bez putnika. Zakonodavci su također napisali zakone u kojima se predlaže da sva AV moraju biti vozila s nultom emisijom i moraju imati ugrađenu tipku za paniku. Nadalje, etička pitanja se odnose na odgovornost za nesreće uzrokovane AV gdje se može raspravljati o čimbenicima koji su prouzrokovali nesreću. Proizvođači autonomnih vozila bi trebali snositi odgovornost za sigurnost svojih vozila jer su odgovorni za projektiranje, proizvodnju i programiranje sustava AV. Međutim, također postoji određena razina odgovornosti za vozače kada su prisutni u vozilu, zato što postoje upute o sigurnom korištenju vozila.

AV su programirana da budu iznimno precizna u svojim radnjama i reakcijama. Čvrsto kodirana pravila, algoritmi za izbjegavanje prepreka, prediktivno modeliranje i prepoznavanje objekata pomažu upravljačkom softveru da slijedi prometna pravila i prevladava prepreke. Podešenje cjelokupnog upravljačkog sustava predstavlja značajan izazov s obzirom na neizmerno velik broj različitih situacija koje se mogu sresti u stvarnoj vožnji. Osim inicijalnih podešenja kroz simulacije, jedan od načina za iterativno poboljšanje autonomne vožnje koji primjerice koristi tvrtka Tesla na svojim vozilima je takozvani način sjene (eng. *Shadow mode*), u kojem automobilom zapravo upravlja čovjek ili čovjek s pomoću autopilota [3]. To se tipično radi za testiranje nove verzije softvera autopilota koji prima podatke od senzora, ali ni na koji način ne preuzima kontrolu nad automobilom. Umjesto toga, donosi odluke o tome kako voziti na temelju senzora, a te se odluke mogu usporediti s odlukama vozača ili starijom verzijom autopilota. U slučajevima kada se odluke novog softvera razlikuju od odluka vozača ili starijeg autopilota, incidenti se označavaju za daljnju analizu te se ispituje koliko je važna ta razlika. Podskup tih incidenata zatim se daje ljudskim recenzentima da ispituju i utvrde je li novi softver



napravio pogrešku. Ako se otkrije pogreška, ona se označava za ispravak i proces testiranja se nastavlja. Zatim, kada se svi ti podaci obrade, spremaju se i prosljeđuju svim vozilima pomoću OTA (*eng. Over the air*) sustava koji pruža mogućnost ažuriranja softvera i firmvera u vozilima na daljinu, bez potrebe za fizičkim pristupom vozilu [4]. Ovo je ključna značajka u razvoju i održavanju autonomnih vozila, budući da proizvođačima omogućuje kontinuirano poboljšanje softvera vozila temeljem velike količine stvarnih podataka, popravljjanje grešaka i uvođenje novih značajki [5].

U kontekstu interakcije AV s ostalim sudionicima u prometu potrebno bi bilo predviđati njihovo ponašanje u neposrednoj budućnosti u svrhu ostvarenja sigurne i udobne vožnje [6]. Ovaj rad se stoga bavi predviđanjem ponašanja pješaka u blizini pješačkih prijelaza bez semafora, s naglaskom na odabir računalno efikasnog predikcijskog modela i minimalnog broja podataka za potrebe učenja tog modela. Problem predviđanja ponašanja pješaka formuliran je kao problem klasifikacije s dva moguća ishoda: i) pješak ide preko pješačkog prijelaza prije vozila, i ii) pješak ide preko pješačkog prijelaza nakon vozila. Za ovaj problem odabran je model zasnovan na binarnoj logističkoj regresiji, koji je jednostavne strukture i minimalnog broja parametara i stoga udovoljava zahtjevu na računalnu efikasnost. Slično kao i kod navedenog centralnog testiranja i ažuriranja autopilota Tesla sustava preko OTA sustava, zamislivo bi bilo da se podaci snimljeni tijekom individualnih interakcija vozila s pješacima također šalju u centralni sustav na oblaku (*eng. cloud*), gdje bi se predikcijski model ažurirao temeljem velike količine podataka prikupljenih s mnogo vozila (v. npr. ilustraciju na Slika 1). Ažurirani model bi se potom slao natrag na sva pojedinačna vozila, koja bi ga koristila za potrebe autonomne vožnje. U svrhu smanjenja prometa potencijalno velike količine podataka zamišljeno je da se na oblak šalju samo relevantni informativni podaci. Stoga je u ovom radu predložena metoda stohastičkog filtriranja podataka, koja na temelju lokalnog predikcijskog modela (na samom vozilu) određuje koliko su neki podaci relevantni. Ako su podaci prema trenutnom modelu očekivani onda je postavljeno da vjerojatnost njihovog propuštanja/slanja prema oblaku bude manja i obrnuto. U ovom radu su kao ulaz u model razmotreni samo kinematski podaci kao što su pozicije i brzine vozila i pješaka. U slučaju mogućnosti interpretacije suptilnih znakova i neverbalne komunikacije između vozila i pješaka koji bi potencijalno mogli poboljšati točnost predikcije [7], broj podataka bi znatno narastao i u tom slučaju bi prednosti metode filtriranja podataka još više došle do izražaja.

Rad je organiziran u 6 poglavlja. U drugom poglavlju nalazi se pregled dviju metoda predviđanja ponašanja pješaka na pješačkom prijelazu bez semafora. U trećem poglavlju

opisano je postojeće simulacijsko okruženje koje je bilo temelj za razvoj predikcijskog modela, koji je opisan u četvrtom poglavlju. Na kraju je opisana integracija razmatranog predikcijskog modela s postojećom upravljačkom strategijom, u petom poglavlju. Zaključne napomene dane su u zadnjem poglavlju.



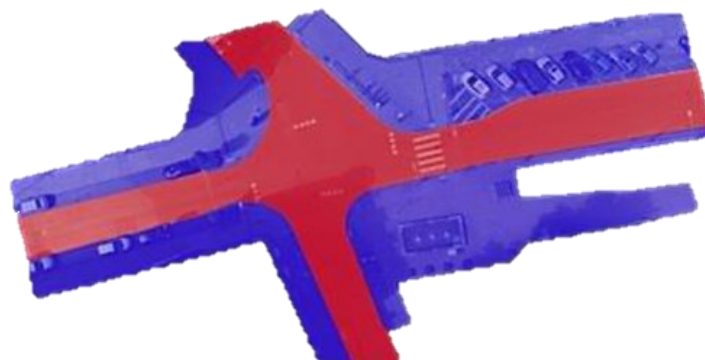
**Slika 1 Ilustracija komunikacije autonomnog vozila s centralnim sustavom na oblaku [8]**

## 2. PREGLED METODA PREDVIĐANJA PONAŠANJA PJEŠAKA NA PJEŠAČKOM PRIJELAZU BEZ SEMAFORA

Na putu prema potpuno automatiziranoj vožnji u urbanim okruženjima, računalni sustavi se moraju nositi sa sve većim brojem zadataka koje trenutačno obavljaju samo ljudi. Jedan od najkompleksnijih i najvažnijih zadataka je razmišljati o drugim sudionicima u prometu. Pješaci zbog brojnih faktora koji utječu na njihove odluke posebno predstavljaju veliki izazov. Ti faktori uključuju demografiju pješaka [9], dinamiku prometa [10], okolišne uvjete, uvjete na cesti [11] i još mnogo toga. Iako se kreću znatno sporije od drugih sudionika u prometu, poput automobila, obično su puno okretniji. Pješaci mogu promijeniti smjer kretanja i brzinu unutar djelića sekunde. U urbanom okruženju to može značiti da će pješak koji je dugo hodao paralelno s cestom iznenada prijeći cestu. Takvo okruženje je vrlo složeno i postavlja mnoge situacije koje je teško predvidjeti. Iskorištavanje strukture okoline ključno je za precizno predviđanje u takvim situacijama. Postoje mnogi strukturalni elementi, poput ulica, nogostupa, križanja ili pješačkih prijelaza, koji pružaju važne smjernice za kretanje svih sudionika u prometu. Osim toga, postoje i pravila, kako zakonska tako i društvena, koja definiraju kako se svi sudionici u prometu moraju kretati ili barem trebaju kretati u datom okruženju. To uključuje, primjerice, pravila o prednosti prolaza na križanjima ili pješačkim prijelazima. Sudionici u prometu obično djeluju racionalno, prate pravila i nastoje izbjeći nesreće, ali često inzistiraju na svojem pravu prednosti (npr. na pješačkom prijelazu). Iako pješaci ponekad krše prometna pravila, na primjer, trčeći na crveno svjetlo, i dalje nastoje izbjeći nesreće. To obično rezultira ponašanjem koje minimizira rizik, kao što je prelazak ceste samo ako je prazna ili ako je najbliži automobil još daleko. U nastavku ovoga poglavlja bit će prikazane metode za proučavanje i predikciju ponašanja pješaka i ispitivanja kako su navedeni faktori međusobno povezani.

### 2.1. Metoda temeljena na klasifikacijskoj metodi *Random Forest* [12]

Ključni fokus ove studije je predviđanje namjere prelaska pješaka, s naglaskom na upotrebu modela strojnog učenja, do tri sekunde prije stvarnog prelaska, koristeći skup podataka snimljen iz perspektive drona. Oslanja se na javno dostupan skup podataka [13] za obuku i testiranje predloženog modela. U ovoj studiji odabrano je raskrižje u njemačkom gradu Aachenu, prikazano na Slika 2, jer sadrži najveći broj pješaka, zbog svog položaja u blizini centra grada.



**Slika 2 Segmentirano raskrižje gdje plava boja predstavlja nogostup, a crvena cestu [12]**

Problem predviđanja namjere prelaska pješaka je formuliran kao zadatak binarne klasifikacije gdje je cilj odrediti hoće li određeni pješak prijeći ulicu ili neće. Predloženi model može predvidjeti namjeru prelaska pješaka koji hodaju na različitim lokacijama oko nesignaliziranog raskrižja (npr. pješački prijelaz, nogostup ili ispred rubnika) s obzirom na popis odabranih značajki i izvedenih značajki koje pruža javno dostupan skup podataka i podataka pretprocesiranja.

U ovom slučaju, za klasifikaciju je korištena metoda *Random Forest* (RF). RF je fleksibilan algoritam strojnog učenja jednostavan za korištenje, a koji se najčešće koristi za zadatke klasifikacije, ali često i za zadatke regresije [14]. Temelji se na konceptu ansambla, što znači da kombinira predikcije više pojedinačnih modela kako bi se postigla bolja ukupna prediktivna snaga. Osnovna ideja je konstruirati više stabala odlučivanja tijekom treninga i kombinirati njihove rezultate prilikom donošenja konačne odluke. Spomenuto stablo odlučivanja je također model strojnog učenja koji se koristi za donošenje odluka na temelju niza pravila. Stabla odlučivanja klasificiraju primjere organizirajući ih kroz hijerarhijsku strukturu stabla, od korijena prema listovima ili završnim čvorovima. Svaki list ili završni čvor u stablu daje klasifikaciju određenog primjera. Proces izgradnje stabla je rekurzivan: svaki čvor u stablu predstavlja test za određeni atribut, a grane koje se spuštaju od čvora predstavljaju različite mogućnosti odgovora na taj test. Ovaj rekurzivni proces se ponavlja za svako podstablo koje se ukorijeni u novom čvoru, omogućujući tako organizaciju i klasifikaciju primjera kroz višestruke testove i odluke [15]. Obzirom na skup ulaznih značajki za predikciju  $\mathbf{v}$ , obučeni klasifikator RF će ispisati klasu  $c$  kojoj pripadaju i vjerojatnost  $p_t$  da će biti u toj klasi. Predviđene vjerojatnosti klasa vektora ulaznih značajki izračunavaju se kao srednje predviđene vjerojatnosti klasa stabala u šumi, označene s  $T$ . Vjerojatnost klase jednog stabla je udio uzoraka iste klase u listu. U konačnici, kada dobijemo ulazne značajke, svako stablo u ansamblu

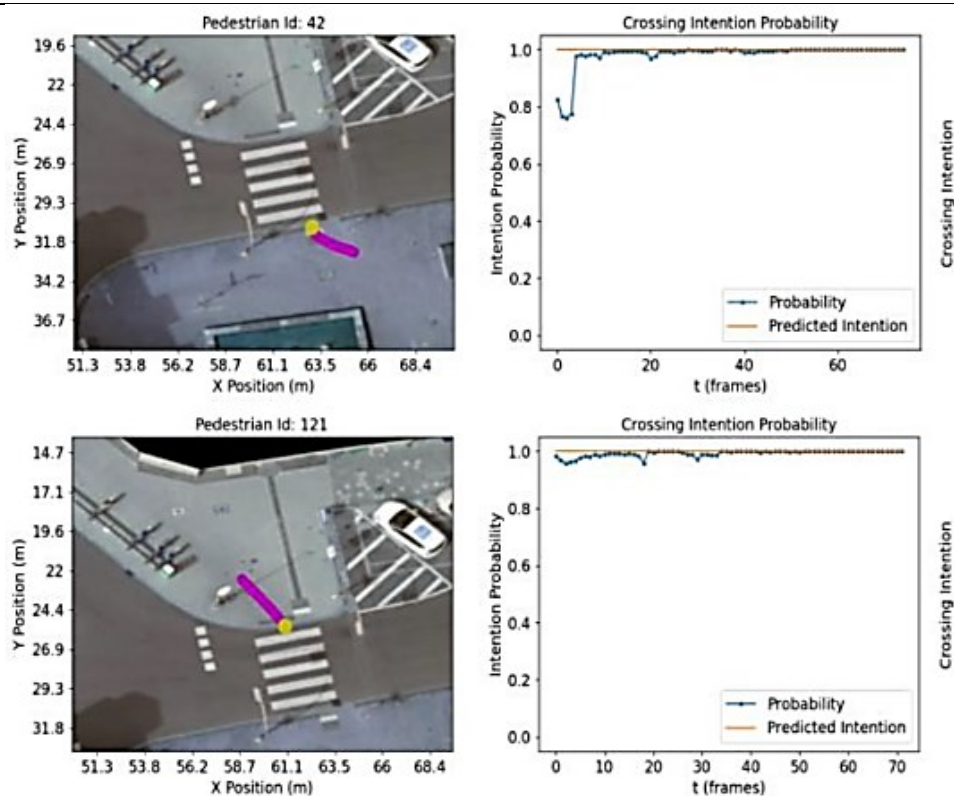
daje svoju vlastitu predikciju, a predviđene vjerojatnosti klasa se usrednjavaju kako bismo dobili konačnu predikciju ansambla, što je izračunato pomoću sljedeće formule:

$$p(c|\mathbf{v}) = \frac{1}{T} \sum_t^T p_t(c|\mathbf{v}), \quad (1)$$

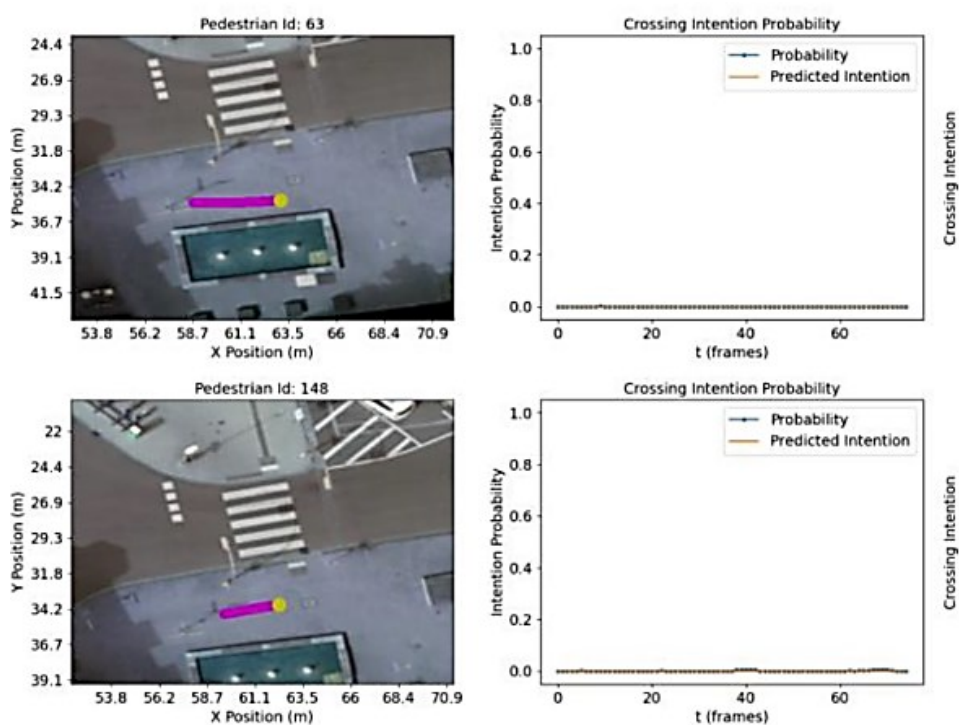
gdje je:  $p(c|v)$  konačna predikcija ansambla,  $T$  ukupan broj stabala u RF,  $p_t(c|v)$  predikcija jednog stabla,  $c$  je klasa, a  $\mathbf{v}$  predstavlja skup ulaznih značajki.

Još jedan dio predikcije je i proces predobrade podataka koji je uključivao sljedeće korake. Prvo, ručno je provedena osnovna semantička segmentacija okoline kako bi se slika podijelila na dva glavna segmenta: nogostup i cestu kao što je označeno bojama na Slika 2. Ova segmentacija omogućila je prepoznavanje trenutka kada pješak počinje prelaziti cestu. Zatim je uvedena značajka nazvana "*distanceToRoad*". Ova značajka uključivala je određivanje vektora koji se proteže od trenutne pozicije pješaka prema cesti u smjeru kretanja pješaka. Duljina tog vektora predstavljala je udaljenost do ceste. Kada bi se detektirala radnja prijelaza temeljem semantičke segmentacije i značajke "*distanceToRoad*," označio se svaki prethodni vremenski korak (do tri sekunde unatrag) kao "C" (što označava ponašanje prijelaza). Svi ostali vremenski koraci smatrali su se "NC" (bez prelaska), osim kada bi se utvrdilo da su pješaci već bili na cesti, u tom slučaju bi se ti koraci također označili kao "C." Ove oznake su pružile osnovu za označavanje namjere prijelaza pješaka. Za parametriranje modela korištena je Python biblioteka za strojno učenje *Scikit-learn* [16].

Skup podataka je temeljito promiješan i podijeljen na dvije skupine - skup za treniranje (70%) i skup za testiranje (30%) radi prilagodbe modela i evaluacije njegove učinkovitosti. Posebna pažnja posvećena je tome da testni skup ne sadrži duplikate trajektorija koje se već nalaze u skupu za obuku. Kako bi se dodatno provjerila izvedba modela, eksperiment je uključivao deset putanja koje se križaju i deset putanja koje se ne križaju, a koje model prije nije vidio. Model je uspješno predviđao namjeru pješaka u ovim scenarijima s točnošću od 97%. Na Slika 3 i Slika 4 prikazani su rezultati za različite scenarije prelaska pješaka na raskrižju. Putanje pješaka prikazane su ružičastom bojom. Točka putanje prikazana žutom bojom označava posljednji uzorak putanje u nizu. Model je uspješno predviđao namjeru pješaka u različitim situacijama, uključujući i one na manje korištenim površinama raskrižja.



Slika 3 Rezultati klasifikacije za putanje koje se križaju [12]



Slika 4 Rezultati klasifikacije za putanje koje se ne križaju [12]

Opisani rad predstavio RF klasifikator za predviđanje namjere prijelaza pješaka preko pješačkog prijelaza na prometnom raskrižju u središtu grada. Koristeći značajke pješaka kao što su položaj, brzina i smjer, koje pruža javno dostupan skup podataka, te izvedene značajke,

model je mogao predvidjeti namjeru prelaska unutar vremenskog okvira od tri sekunde i točnosti od 0,98.

Za kraj je analiziran odgovor modela na situacije u kojima pješak pokazuje neodlučnost prije prelaska ceste. Model je ispitivan na primjerima pješaka koji su značajno usporili prije prelaska. Ovi slučajevi su istaknuli važnost proučavanja interakcije pješaka s drugim sudionicima u prometu te su ukazali na mogućnost poboljšanja modela dodavanjem dodatnih kontekstualnih informacija kao što su poza tijela ili smjer pogleda pješaka. U budućnosti bi se tu mogle dodati spomenute kontekstualne informacije, izradom skupa podataka snimljenih s fiksnog senzora i izdvajanjem tih značajki. Također su istraženi scenariji u kojima namjera prelaska nije bila jasna, npr. usporavanje pješaka prije prelaska, otkrivajući potencijalnu važnost proučavanja interakcije pješaka s drugim agentima u sceni.

## **2.2. Metoda temeljena na procjeni relevantnosti značajki [17]**

Sljedeći rad koristi bazu podataka stvarnih trajektorija pješaka kako bi se naučio model koji može predviđati hoće li pješak prijeći cestu. Prvo se uvodi veliki skup mogućih značajki koje bi mogle biti prikladne za opisivanje ponašanja. Nakon toga provodi se određivanje relevantnosti kako bi se identificirale one značajke koje su potrebne za postizanje najbolje moguće opće izvedbe. Pružaju se eksperimentalni rezultati na podacima prikupljenim na pješačkom prijelazu u gradu na jugu Njemačke.

I u ovome radu riječ je o problemu klasifikacije te se u tu svrhu koristila nelinearna metoda potpornih vektora (eng. SVM- *support vector machines*) [18]. Cilj SVM-a je pronaći najbolju moguću liniju ili granicu-hiperravninu odluke koja razdvaja podatkovne točke različitih klasa podataka te se kao takav pokazao idealnim za ovaj zadatak jer je odluka o tome prelazi li pješak ili ne binarna. Rad obuhvaća dva važna istraživačka pitanja: a) koje informacije su potrebne za shvatiti namjere pješaka, i b) u kojoj mjeri namjere pješaka utječu na druge sudionike u prometu. Da bi se dobio odgovor na ova pitanja, definirao se veliki skup mogućih značajki za algoritam učenja. Te značajke se mogu podijeliti u dva osnovna tipa: značajke koje opisuju kretanje pješaka u lokalnom koordinatnom okviru i značajke koje karakteriziraju interakciju između pješaka i automobila. Koristeći ove značajke, naučeno je nekoliko različitih modela i ocijenjeno prema mjeri kvalitete te se koristio iterativni algoritam eliminacije značajki za određivanje najrelevantnijih značajki. Podaci koji se koriste u ovom radu snimljeni su na pješačkom prijelazu prikazanom na Slika 5, koja prikazuje i trajektorije za pješake koji prelaze cestu (crno), i za one koji to ne čine (bijelo).



**Slika 5 Pješački prijelaz s trajektorijama [17]**

Dva skupa podataka koji su korišteni za treniranje i evaluaciju potpuno su različita. Skup za obuku sadržavao je samo trajektorije snimljene tijekom jednog dana, dok je testni skup sadržavao trajektorije snimljene drugih dana. Postojalo je nekoliko mogućih kriterija za završetak iteracija treniranja. Razmatrajući dva različita skupa podataka, iteracija bi bila završena ako bi bila postignuta određena točnost klasifikacije. Gledajući samo skup za treniranje, korištena je greška unakrsne validacije (eng. CV- *cross-validation*). Cilj unakrsne validacije je testirati sposobnost modela da predvidi nove podatke koji nisu korišteni u njegovoj procjeni, kako bi se označili problemi kao što su prekomjerno prilagođavanje ili pristranost odabira i kako bi se dobio uvid u to kako će se model generalizirati na neovisni skup podataka. Ova metoda se koristila kako bi se zadovoljio kriterij za završetak, tj. ako je ta greška bila ispod unaprijed definiranog praga.

### **2.2.1. Odabrane značajke**

Od nekoliko mogućih fizičkih vrijednosti koje bi mogle opisati kretanje pješaka, jedna od najvažnijih je brzina pješaka  $v_{ped}$ . Korištenjem pružene karte zajedno s izmjerenom pozicijom pješaka, mogu se izračunati sljedeće relativne mjere udaljenosti:  $dt_{curb}$  kao minimalna ortogonalna udaljenost do najbližeg ruba, gdje je  $dt_{curb} < 0$  ako je pješak na cesti, a  $dt_{curb} \geq 0$  ako nije; te  $dt_{cross}$  kao minimalna udaljenost do pješačkog prijelaza, gdje je  $dt_{cross} = 0$  ako je pješak na pješačkom prijelazu, a  $dt_{cross} \geq 0$  ako nije. Dodatno, koristi se i prijeđena udaljenost između prethodnog i trenutnog vremenskog koraka  $s_{ped}$ . Dodane su i značajke za automobile kako bi se istražio njihov utjecaj na ponašanje pješaka. Automobili se identificiraju prema udaljenosti od pješaka, eliminirajući one koji su već prošli pješački prijelaz, uz primjenu jednostavne funkcije troška za odabir automobila koji su blizu pješaka i brzo će dosegnuti pješački prijelaz. Slično kao i za predložene značajke za pješake, koristi se brzina automobila  $v_{car}$ , ali i prijeđena udaljenost automobila  $s_{car}$ . Koristeći informacije o karti, izračunava se



udaljenost do pješačkog prijelaza  $d_{tcrosscar}$  kao relativna mjera udaljenosti, gdje je  $d_{tcrosscar} = 0$  ako je automobil na pješačkom prijelazu, a  $d_{tcrosscar} \geq 0$  inače. Razmatrajući odnos između pješaka i automobila, u prostor značajki mogu se uključiti i udaljenost između njih:

$$d_{ped,car}(t) = \begin{vmatrix} x_{ped} - x_{car} \\ y_{ped} - y_{car} \end{vmatrix}, \quad (2)$$

i njihova relativna brzina:

$$v_{rel} = v_{car} - v_{ped}, \quad (3)$$

gdje su:  $x(y)_{ped}$  prijeđena udaljenost pješaka u 2D koordinatnom sustavu,  $x(y)_{car}$  prijeđena udaljenost automobila u 2D koordinatnom sustavu,  $v_{rel}$  relativna brzina.

Za svaki vremenski korak, povijest svih značajki iz posljednja četiri vremenska koraka uključena su u prostor značajki. Na primjer: umjesto samo  $s_{ped}(t)$ , koriste se vrijednosti:  $s_{ped}(t)$ ,  $s_{ped}(t - 1)$ ,  $s_{ped}(t - 2)$ ,  $s_{ped}(t - 3)$  i  $s_{ped}(t - 4)$ . Ukupno postoje pet vrijednosti za svaku značajku što vodi na ukupno 75 značajki.

Osim definiranja mogućih značajki, glavni cilj je identifikacija najvažnijih značajki dobivena rekurzivnom eliminacijom značajki (eng. RFE- *Recursive feature elimination*) [19]. Algoritam se sastoji od iterativnih koraka, uključujući obuku SVM-a, izračun kriterija rangiranja za sve značajke temeljen na sirovim težinama SVM-a te uklanjanje značajki s najmanjim kriterijem rangiranja. Analizira se važnost svake značajke, a postupak odabira provodi se dok se sve značajke ne eliminiraju. Rezultat ovog algoritma je rangiranje svih značajki.

### 2.2.2. Evaluacija

Za evaluaciju se koristi prethodno opisani RFE za obuku i testiranje više SVM modela. Prvo su predstavljeni rezultate grupne i pojedinačne eliminacije. Zatim su detaljnije analizirani rezultati klasifikacije za određeni skup značajki. Tijekom klasifikacije donosi se odluka o tome hoće li pješak prijeći ili neće prijeći. Dalje u tekstu će se na to odnositi kao pozitivan (za pješaka koji prelazi) ili negativan (za pješaka koji ne prelazi) rezultat.

#### 2.2.2.1. Grupna eliminacija

Korištena su dva pristupa grupne eliminacije. Prvi pristup odvaja podatke prema njihovom stvarnom fizičkom mjerenju, generirajući 15 grupa, svaka s pet članova, koje pripadaju izvornim pojedinačnim značajkama. Rezultirajuće rangiranje, prikazano u Tablica 1,

ukazuje da su najbolje rangirane grupe značajki za fizičku separaciju :  $v_{ped,x}$ ,  $d_{tcurb}$  i  $d_{tcross}$ , postižući ukupnu točnost klasifikacije od 91,67%.

**Tablica 1 Rezultat eliminacije prve skupine s odgovarajućom ukupnom klasifikacijskom točnošću i lažno pozitivnom stopom [17]**

Number of Groups	Physical Measure Ranking	Overall Classification Accuracy	True Positive Rate
1	$v_{ped,x}$	0.6518	0.0514
2	$d_{tcross}$	0.8777	0.8401
<b>3</b>	<b><math>d_{tcurb}</math></b>	<b>0.9167</b>	<b>0.9120</b>
4	$ v_{ped} $	0.8981	0.8831
5	$v_{ped,y}$	0.8733	0.8472
6	$s_{ped}$	0.8502	0.8289
7	$v_{rel,x}$	0.8420	0.8437
8	$ v_{rel} $	0.8276	0.8514
9	$d_{ped,car}$	0.8239	0.8507
10	$v_{car,x}$	0.8048	0.8510
11	$ v_{car} $	0.7917	0.8500
12	$v_{rel,y}$	0.7855	0.8493
13	$v_{car,y}$	0.7768	0.8437
14	$d_{tcross_{car}}$	0.7778	0.8429
15	$s_{car}$	0.7679	0.8418

Drugi pristup grupne eliminacije razdvaja podatke prema vremenskim koracima, dobivajući pet grupa s 15 članova. Tablica 1 i Tablica 2 prikazuju rezultate ove grupe eliminacije. Ukupna točnost klasifikacije za svaki redak u obje tablice pokazuje izvedbu modela koji sadrži određene grupe, pri čemu redci iznad utječu na rezultate. Iz Tablica 1 proizlazi da su značajke povezane s automobilima na kraju rangiranja i pružaju inferiornije rezultate klasifikacije. Najbolji rezultat postiže se kombinacijom vremenskih koraka  $t$  i  $t-4$  s ukupnom točnošću klasifikacije od 83,31%.

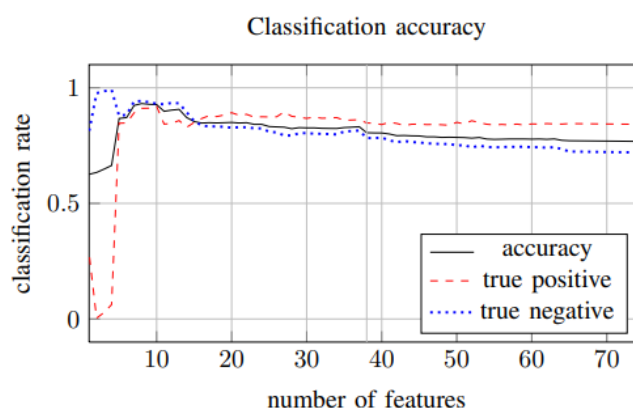
**Tablica 2 Rezultat grupne eliminacije temeljene na vremenskom odmaku s odgovarajućom ukupnom klasifikacijskom točnošću i lažno pozitivnom stopom [17]**

Number of Groups	Time Step Ranking	Overall Classification Accuracy	True Positive Rate
1	$t$	0.6845	0.6768
<b>2</b>	<b><math>t - 4</math></b>	<b>0.8331</b>	<b>0.8232</b>
3	$t - 1$	0.8078	0.8451
4	$t - 3$	0.7867	0.8563
5	$t - 2$	0.7679	0.8418

Usporedba rezultata dviju grupa eliminacije sugerira da prvi pristup, fizička separacija, daje bolje rezultate, a uklanjanje određenih vremenskih koraka može poboljšati rezultate.

### 2.2.2.2. Pojedinačna eliminacija

U ovom odjeljku primijenjen je standardni algoritam RFE za neovisno eliminiranje svih značajki, a evaluirane su standardne metrike binarne klasifikacije, uključujući točnost, stvarni postotak pozitivnih i stvarni postotak negativnih. Evaluacija se temeljila na vremenskim koracima, fokusirajući se na procjenu trajektorija pješaka u dijelovima umjesto na cjelokupnim putanjama. Rezultati su prikazani na Slika 6.



**Slika 6 Točnost klasifikacije spram broja pojedinačnih značajki [17]**

Budući da se morao odabrati određeni skup značajki, sve stope klasifikacije moraju se procijeniti zajedno. Da bi se postigla najbolja ukupna izvedba, mora se pronaći skup značajki koji maksimizira sve krivulje. Posebno je visoka stopa stvarnih pozitivnih od velike važnosti jer pogrešna klasifikacija pješaka koji za stvarno prelazi cestu može predstavljati visoki sigurnosni rizik. Pregledom svih triju krivulja zajedno može se pronaći optimum za točno 10 korištenih značajki. Vrlo važna stopa stvarnih pozitivnih doseže globalni maksimum u tom trenutku. Iako postoje kombinacije značajki koje pružaju sličnu ili malo bolju točnost i stopu stvarnih negativnih, odgovarajući pad stope stvarnih pozitivnih zabranjuje upotrebu tih skupova značajki. Za veći broj korištenih značajki, može se primijetiti da, iako veliki skupovi značajki i dalje pružaju dobre rezultate, točnost se postupno smanjuje. Ovaj pad ponovno pripada značajkama povezanim s automobilima. Slika 6 prikazuje mali pad u točnosti negdje između 35 i 40 korištenih značajki. To je otprilike točka na kojoj se prve značajke povezane s automobilima dodaju u prostor značajki. Nasuprot tome, 10 značajki koje pružaju najbolji rezultat ponovno su dio grupa  $v_{ped,x}$ ,  $dtcurb$  i  $dtcross$ .

### 2.2.2.3. Evaluacija najboljeg skupa značajki

U posljednjem dijelu istraživanja identificiran je optimalan skup značajki koji pruža najbolje rezultate klasifikacije. Dosadašnja evaluacija fokusirala se isključivo na vremenske

korake, tretirajući podatke u svakoj trajektoriji kao neovisna promatranja. Sada se prelazi na evaluaciju temeljenu na trajektorijama, u kojoj se uvodi nova mjera za ocjenjivanje oznaka tragova. Konkretno, nova mjera označava trajektoriju kao točnu ili netočnu ovisno o omjeru vremenskih koraka koji su ispravno klasificirani:

$$\text{oznaka} = \begin{cases} \text{točno,} & \text{ako } \frac{\text{broj ispravnih vremenskih koraka}}{\text{broj svih vremenskih koraka}} > y, \\ \text{netočno,} & \text{inače.} \end{cases} y \in \{0,1\}. \quad (4)$$

Tablica 3 prikazuje broj trajektorija koji se svrstavaju u različite diskretne intervale udaljenosti, te informacije o broju stvarnih pozitivnih i stvarnih negativnih trajektorija za vrijednosti  $y = 0.5$  i  $y = 0.8$ .

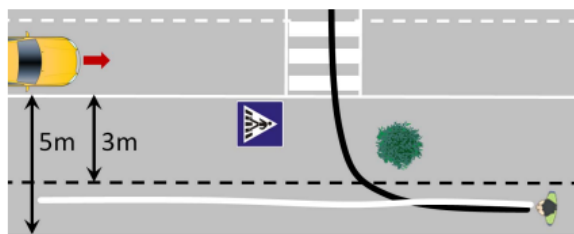
**Tablica 3 Ocjena najboljeg skupa značajki: ukupan broj pozitivnih i negativnih trajektorija za diskretizirane mjere udaljenosti  $d_{t\text{cross}}$  i  $d_{t\text{curb}}$  [17]**

Distance $x$ in $m$	$x = d_{t\text{cross}}$					
	all	positive (cross)		negative (not cross)		
		$y = 0.5$	$y = 0.8$	all	true $y = 0.5$   $y = 0.8$	
$\leq 1$	36	36	36	52	52	52
1 – 2	26	26	25	39	39	38
2 – 3	20	19	16	28	28	25
3 – 4	28	23	18	38	31	26
4 – 5	15	12	8	41	37	33
$> 5$	18	12	11	30	27	26

Distance $x$ in $m$	$x = d_{t\text{curb}}$					
	all	positive (cross)		negative (not cross)		
		$y = 0.5$	$y = 0.8$	all	true $y = 0.5$   $y = 0.8$	
$\leq 1$	38	38	38	55	55	55
1 – 2	28	27	26	39	39	38
2 – 3	26	22	21	33	29	28
3 – 4	28	22	17	50	45	40
4 – 5	10	7	7	17	15	15

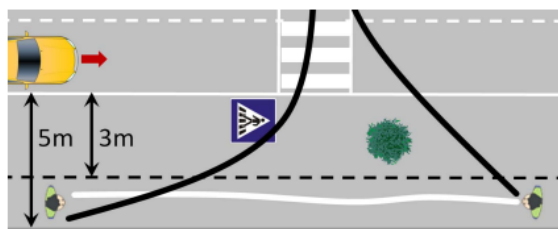
Slika 7 prikazuje dvije tipične trajektorije pješaka. Jedna prikazuje pješaka koji prolazi pješački prijelaz s konstantnom vrijednošću  $d_{t\text{curb}}$ , dok druga pokazuje pješaka koji dugo vremena ide paralelno s pješačkim prijelazom prije nego što prijeđe cestu. Ove dvije trajektorije, koje su početno teške za razdvajanje, glavni su uzrok inferiornih stopa klasifikacije, posebno stvarnih pozitivnih, u intervalu  $3m < d_{t\text{curb}} \leq 5m$ .



**Slika 7 Tipične trajektorije pješaka, teško razdvojive [17]**

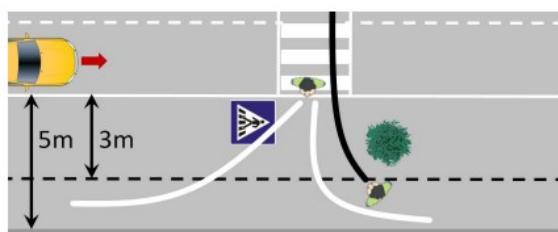
Nasuprot tome, Slika 8 prikazuje one prelazne trajektorije koje se lako mogu razdvojiti od prolazećeg pješaka jer se mogu identificirati, primjerice, zbog postojane smanjujuće

vrijednosti  $d_{curb}$ . Većina prelaznih trajektorija pokazuje ovakvo ponašanje. To rezultira snažnom osnovom stvarnih pozitivnih, čak i za velike udaljenosti.



**Slika 8 Tipične trajektorije pješaka, lako razdvojive [17]**

Na Slika 9 prikazane su uobičajene trajektorije pješaka blizu pješačkog prijelaza. Važno je napomenuti, da u ovom kontekstu, pješaci koji ne prelaze često predstavljaju one koji su već prešli ulicu u prošlosti i sada se udaljavaju od pješačkog prijelaza. Razdvajanje ovih trajektorija je lako jer se vrijednosti  $d_{cross}$  i  $d_{curb}$  obično identične i ili se smanjuju (za pješake koji prelaze) ili se povećavaju (za pješake koji ne prelaze) tijekom vremena. U konačnici, zbunjenost između različitih pokreta gotovo je nemoguća. Stoga, točnost klasifikacije uvijek je vrlo visoka za male vrijednosti  $d_{cross}$ .



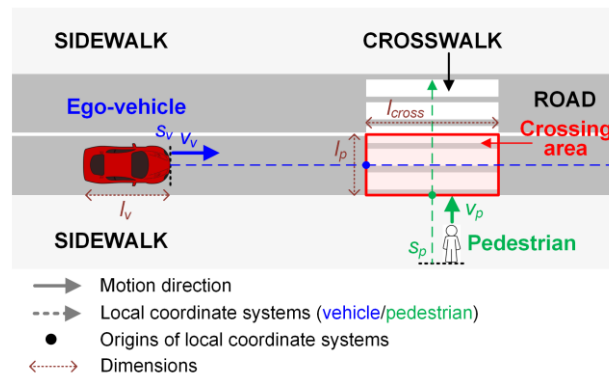
**Slika 9 Tipične trajektorije pješaka koji prelaze preko pješačkog prijelaza [17]**

Rezultati ovoga rada pokazuju da mali podskup ovih značajki, koje ovise samo o trajektorijama pješaka i lokalnoj karti, daje najbolje rezultate. Sve značajke koje ovise o prisutnosti drugih sudionika u prometu manjeg su značaja. Čini se da ne pružaju dodatne informacije za ovaj specifičan zadatak. Za automatizirano vozilo, prikazani rezultati prvi su korak prema složenom sustavu koji može predvidjeti kretanje pješaka u različitim situacijama.

### 3. POSTOJEĆE SIMULACIJSKO OKRUŽENJE

#### 3.1. Temeljne karakteristike simulacijskog okruženja

Simulacijski okvir, koji je poslužio kao temelj za razvoj modela predviđanja ponašanja pješaka korištenjem metode logističke regresije, uključuje pojednostavljenu interakciju jednog pješaka i jednog automobila na pješačkom prijelazu bez semafora (v. Slika 10). U ovome scenariju, ponašanje pješaka i njegove odluke počivaju na modelu prethodno korištenom u sklopu kompleksnijeg modela zasnovanog na teoriji igara [20]. Automobil se kreće putanjom referentne brzine propisanom na početku simulacije i to kroz dva propisana upravljačka parametra. Prvi parametar je ciljana brzina, a drugi je ciljano ubrzanje prema ciljanoj brzini. Interakcija je maksimalno pojednostavljena u smislu da u nju ulazi samo jedno vozilo i samo jedan pješak te se pješak giba konstantnom brzinom, dok ne dosegne rub pješačkog prijelaza gdje donosi konačnu odluku, nastaviti s prijelazom ili stati i pričekati vozilo da prođe. Područje potencijalnog sudara između automobila i pješaka jest zajednička zona na pješačkom prijelazu koja se može smatrati dijeljenim resursom oko kojeg se natječu i pješak i automobil. Vozilo je namješteno tako da prilazi pješačkom prijelazu isključivo uzdužno u svojoj voznoj traci, dok pješak prilazi okomito u odnosu na cestu.



Slika 10 Prikaz simulacijskog okvira za jedno vozilo/jedan pješak [21]

Varijable stanja automobila i pješaka uključuju njihove stvarne pozicije, definirane relativno u odnosu na područje pješačkog prijelaza, i preko njihove stvarne brzine. Gibanje vozila simulirano je na temelju sljedećeg modela u prostoru stanja zadanog u vremenski-diskretnom obliku:

$$\begin{bmatrix} s_v(k+1) \\ v_v(k+1) \end{bmatrix} = \begin{bmatrix} 1 & \Delta T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} s_v(k) \\ v_v(k) \end{bmatrix} + \begin{bmatrix} 0.5\Delta T^2 \\ \Delta T \end{bmatrix} a_v(k), \quad (5)$$

gdje su:  $s_v$  i  $v_v$  varijable položaja i brzine vozila,  $a_v$  varijabla za upravljanje ubrzanjem vozila,  $\Delta T$  je vrijeme diskretizacije (postavljeno na  $\Delta T = 0.1$  s) i  $k$  je diskretni vremenski korak. U

slučaju pješaka, položaj  $s_p$  odabran je kao varijabla stanja, dok je brzina  $v_p$  upravljajčka varijabla. Pripadna jednadžba stanja, također u vremenski-diskretnom obliku, glasi:

$$s_p(k+1) = s_p(k) + v_p(k)\Delta T. \quad (6)$$

Položaji vozila i pješaka,  $s_v$  i  $s_p$ , dani su s obzirom na njihove lokalne koordinatne sustave kako je prikazano na Slika 10. Dakle, negativna vrijednost položaja označava da pripadni agent još nije stigao do pješačkog prijelaza.

Status ( $S_x$ ) svakog agenta, gdje indeks  $x = v$  označava vozilo, te  $x = p$  pješaka definira se kao:

$$S_x(k) = \begin{cases} -2, & \text{ako } S_x(k) \ll 0 \text{ m,} \\ -1, & \text{inače ako } S_x(k) \leq 0 \text{ m,} \\ 0, & \text{inače ako } 0 < S_x(k) < l_x, \\ 1, & \text{inače ako } S_x(k) \geq l_x, \end{cases} \quad (7)$$

što označava da se odgovarajući agent još nije pojavio u blizini pješačkog prijelaza  $S_x = -2$ , ili se nalazi ispred pješačkog prijelaza  $S_x = -1$ , unutar područja pješačkog  $S_x = 0$ , ili je već prošao pješački  $S_x = 1$ . Parametar  $l_x$  označava širinu pješačkog prijelaza iz perspektive pješaka ( $l_p = 2.5$  m) i duljinu pješačkog prijelaza iz perspektive vozila ( $l_v = 9$  m, što uključuje i duljinu pješačkog prijelaza  $l_{cross}$  i duljinu vozila  $l_{veh}$ ,  $l_v = l_{veh} + l_{cross}$ , v. Slika 10). Definirano je da se sudar smatra događajem kada su oba agenta prisutna unutar područja prijelaza u isto vrijeme, odnosno barem u jednom vremenskom koraku  $k$ :

$$s_p(k) = 0 \wedge s_v(k) = 0, \exists k. \quad (8)$$

Ova definicija je oprezna i naglašava sigurnost, uzimajući u obzir mogućnost da se i vozilo i pješak nalaze unutar područja prijelaza, ali do sudara ne dolazi zbog ograničenih fizičkih dimenzija agenata. Na primjer, može se dogoditi da pješak upravo izlazi iz područja prijelaza dok vozilo tek ulazi, čime se izbjegavaju stvarni kontakti između agenata.

Kao što je već spomenuto, kretanje pješaka je pojednostavljeno pretpostavkom da se pješak kreće prema pješačkom prijelazu konstantnom brzinom, koja je fiksirana na  $v_{p,0} = 1$  m/s, i donosi odluku tek kada dođe do ruba pješačkog prijelaza (tj. do ishodišta lokalnog koordinatnog sustava pješaka na Slika 10). Ova točka odluke otkriva se na temelju sljedećeg uvjeta:

$$s_p(k) \leq 0 \wedge \hat{s}_p(k+1) > 0, \quad (9)$$

gdje je  $\hat{s}_p(k+1)$  predviđena pozicija pješaka u sljedećem ( $k+1$ ) koraku izračunatom prema (6). Ako je status vozila u točki donošenja odluke  $s_v = -1$ , odnosno, nalazi se ispred pješačkog

prijelaza prema izrazu iz (7) odluka o prijelazu pješaka ili ne prijelazu predviđa se prema sljedećem modelu očekivanja (preuzeto iz [22]):

$$U = \alpha + \beta_1 v_p + \beta_2 v_v + \beta_3 |s_v|, \quad (10a)$$

$$p_{cross} = \frac{e^U}{1 + e^U}, \quad (10b)$$

koji kao ulaze uzima brzinu pješaka  $v_p$ , brzinu vozila  $v_v$ , udaljenost vozila od pješačkog prijelaza  $|s_v|$ , te na izlazu daje vjerojatnost  $p_{cross}$  da se pješak odlučuje za odluku prijelaza  $D_p = 1$ , a gdje je vjerojatnost suprotnog događaja, odluke odustajanja od prelaska  $D_p = 0$ , jednostavno definirana kao  $p_{yield} = 1 - p_{cross}$ . Konačna odluka pješaka donosi se generiranjem slučajnog broja  $r$  iz jedinične uniformne distribucije vjerojatnosti  $U(0,1)$  i njezinom usporedbom s dobivenom vjerojatnošću prelaska  $p_{cross}$  kao:

$$D_p = \begin{cases} 0, & \text{ako je } r > p_{cross}, \\ 1, & \text{ako je } r \leq p_{cross}, \end{cases} \quad (11)$$

gdje je u slučaju  $D_p(k) = 0$ , brzina pješaka  $v_p(k)$  u (6) postavljena na 0, a u slučaju  $D_p(k) = 1$ , je namještena na  $v_{p,0}(k)$ , odnosno na 1. Kada pješak odluči odustati od prelaska na rubu pješačkog, pješak čeka da automobil prijeđe pješački prijelaz. Mehanizam donošenja odluke (11) se deaktivira, a odluka pješaka postavlja se na odustajanje  $D_p(k) = 0$  ako je vozilo na pješačkom prijelazu ( $S_v(k) = 0$ ) ili na prelazak  $D_p(k) = 1$  ako je vozilo već prošlo pješački prijelaz ( $S_v(k) = 1$ ).

### 3.2. Generiranje simulacijskih podataka namijenjenih treniranju i testiranju predikcijskog modela

Podaci su generirani provođenjem 1000 simulacija interakcije vozila i pješaka na temelju simulacijskog okruženja opisanog u prethodnom potpoglavlju. Izvodi se s vremenskom diskretizacijom (korakom)  $\Delta T$  od 0,1 s. Svaka se simulacija izvodi za različitu slučajno generiranu kombinaciju početnih položaja vozila ( $s_{v,0}$ ) i početne brzine vozila ( $v_{v,0}$ ), ali uvijek za isti početni položaj pješaka koji je postavljen na  $s_{p,0} = -4$  m. Parametri simulacijskog modela pješaka (10) preuzeti su iz [21], gdje je pješak s tim parametrima klasificiran kao pješak umjerenog ponašanja (v. vrijednost pripadnih parametara u Tablica 4).

Tablica 4 Parametri simulacijskog modela pješaka.

TIP PJEŠAKA	$\alpha$	$\beta_1$	$\beta_2$	$\beta_3$
umjereni	-12.3448	16.2870	-1.6019	0.6628



Iz svake od 1000 simulacija, ekstrahirane su vrijednosti brzine vozila  $v_v$ , i položaja vozila  $s_v$  u trenutku kad je pješak dosegao rub pješačkog prijelaza. Vrijednosti  $v_v$  kreću se od 10 m/s do 5 m/s, a njihov položaj  $s_v$  od +10 m do -40 m, gdje vrijednosti u plusu znače da je vozilo već prošlo pješački prijelaz u trenutku odluke pješaka, a negativne vrijednosti znače da se tek približava pješačkom prijelazu. Isto tako, dobiven je i realizirani ishod  $y$ , koji može biti 0 ako je pješak odlučio stati, ili 1 ako je odlučio prijeći preko pješačkog prijelaza prije vozila. U slučaju ranije spomenute kolizije, uzeta je vrijednost  $y = 1$ , jer je to svakako značilo da je pješak (isto kao i vozilo) stupio na pješački prijelaz tj. odlučio je prijeći prije vozila. Vrijednosti koje su bitne za daljnju analizu su brzine i položaj vozila točno u trenutku odluke pješaka, te hoće li zakoračiti na pješački prijelaz ili neće. Ti podaci spremljeni su u tekstualnu datoteku te su učitani u programski kod. Kod modela napisan je u programskom jeziku Matlab. Na isti način dobiveno je i 1000 podataka za skup za testiranje.

### 3.3. Početna analiza simulacijskih podataka

U okruženju za pred-analizu podataka, ispitano je koliko dobro model predviđa vjerojatnosti realiziranog ishoda, a kao temelj su uzeti idealni parametri umjerenog simulacijskog modela pješaka (navedeni u Tablica 4).

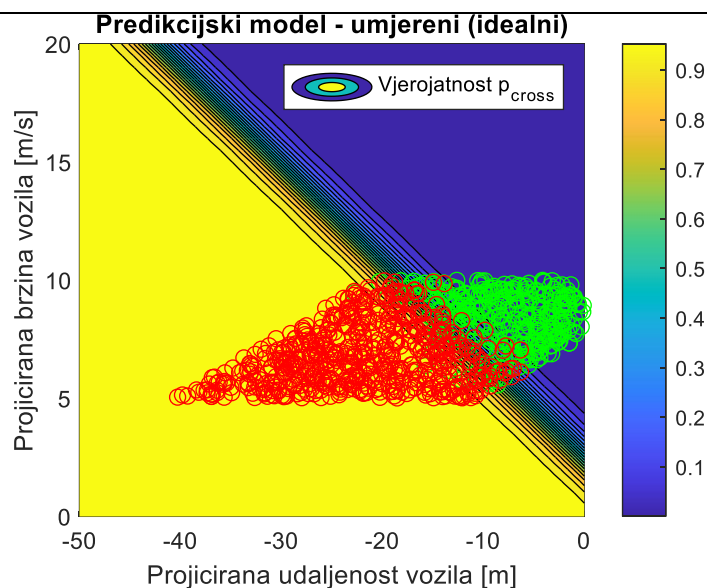
Formula kojom model računa vjerojatnost prelaska pješaka opisana je u (10). Zatim slijedi računanje već spomenute vjerojatnosti realiziranog ishoda označene sa  $p_{actual,i}$ , a koja se računa formulama:

$$p_{actual,i} = 1 - p_i, \quad \text{za } y(i) = 0, \quad (12)$$

odnosno,

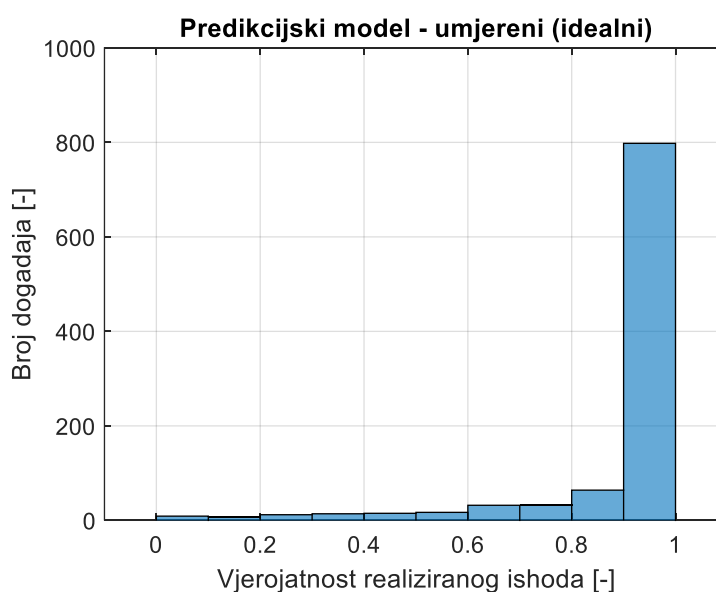
$$p_{actual,i} = p_i, \quad \text{za } y(i) = 1. \quad (13)$$

Veći iznos ove vjerojatnosti ocrta bolju točnost modela (u idealnom slučaju poprima vrijednost 1). U nastavku je dana vizualna analiza modela za slučaj različitih vrijednosti pripadnih parametara, spram simulacijom generiranih točaka.



**Slika 11 Grafički prikaz modela za umjerenog pješaka**

Na slici 11 su podaci za svaku od 1000 simulacija prikazani kružićima, dakle brzinom i pozicijom vozila u trenutku kad pješak dođe na rub pješačkog prijelaza, pri čemu zeleni kružići predstavljaju vrijednost realiziranog ishoda  $y = 0$ , odnosno da je pješak stao, te crveni prikazuju vrijednost realiziranog ishoda  $y = 1$  odnosno da je pješak odlučio prijeći preko pješačkog prijelaza. Plavo područje na grafu označava vjerojatnost koju daje razmatrani model  $p_i$  jednaku 0, a žuti dio  $p_i$  jednaku 1. Analizu ovoga grafa najjednostavnije je provesti tako da gledamo da je model točan za poklapanje crvenih točaka u žutom području te zelenih točaka u plavom području. Ove točke koje se ne poklapaju kako je navedeno, detaljnije su analizirane histogramom na Slika 12.



**Slika 12 Udio vjerojatnosti realiziranog ishoda**

Iz ovoga histograma se vidi da postoji jako mali udio (broj) vjerojatnosti realiziranog ishoda koji su manji od 1, odnosno, da je vjerojatnost da će se realizirani ishod  $y$  stvarno dogoditi, jako velika. Može se zaključiti da ovaj model dobro predviđa kada su inicijalne vrijednosti parametra podešene na umjerenog pješaka i kada se učitavaju podaci koji su generirani na temelju istih tih vrijednosti. Ovo je i očekivano obzirom da su korištenjem ovog modela i pripadnih parametara analizirane točke i generirane.

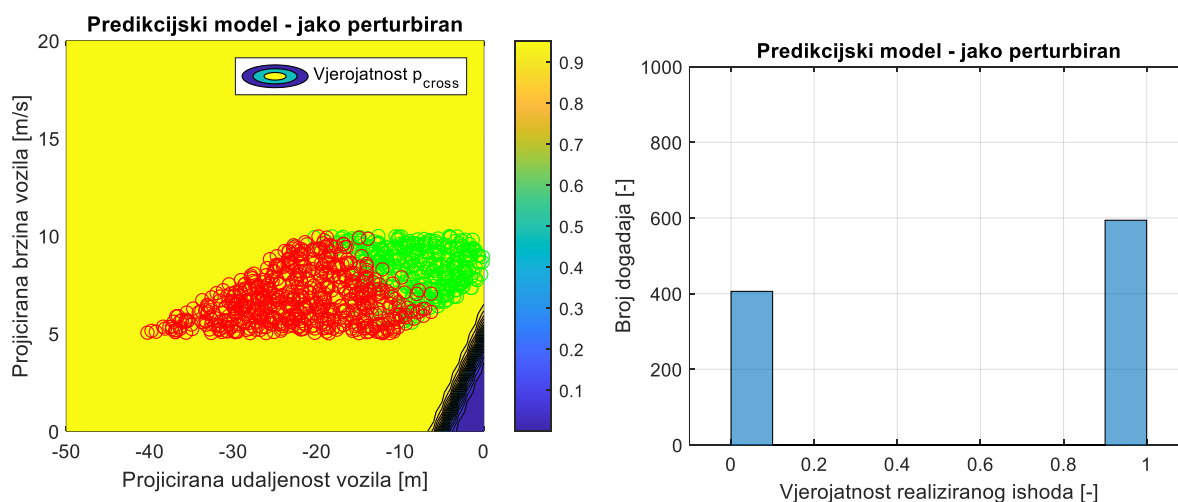
### 3.4. Različiti simulacijski modeli pješaka

Osim umjerenog tipa pješaka, za potrebe različitih testiranja odabrano je još nekoliko skupova parametara simulacijskog modela pješaka (dani u Tablica 5).

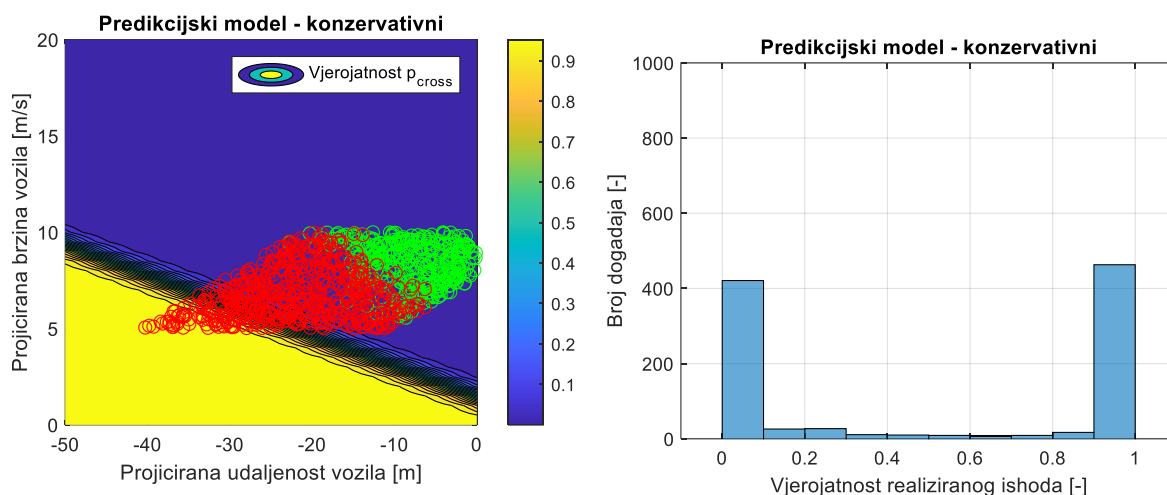
Tablica 5 Parametri ostalih modela

TIP PJEŠAKA	$\alpha$	$\beta_1$	$\beta_2$	$\beta_3$
jako perturbirani	-5	-5	2	2
konzervativni	-13.292	17.915	-3.135	0.495
agresivni	-0.9362	9.7593	-1.0759	0.2439

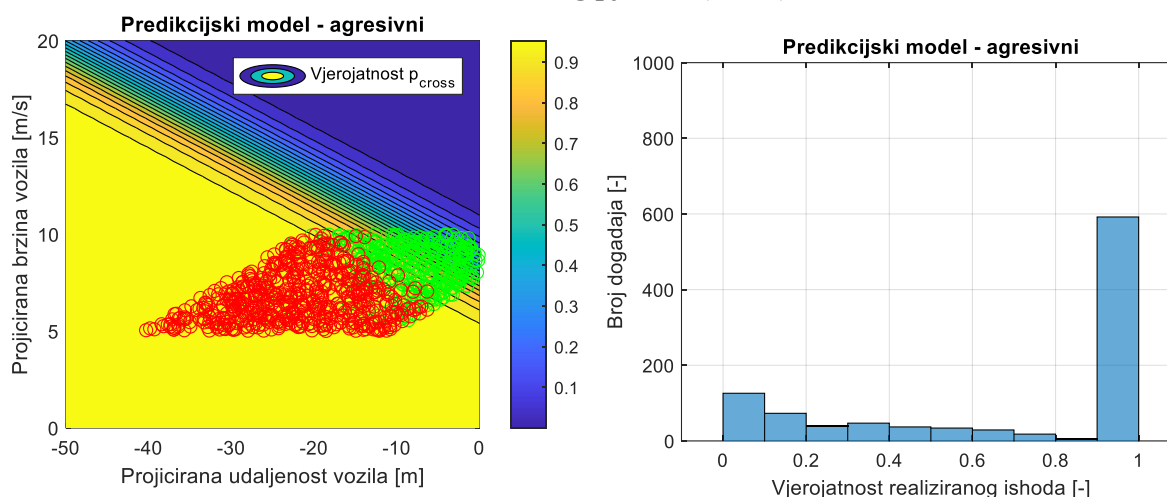
Jako perturbirani pješak je podešen tako da ima značajno različite parametre od ostalih modela, dok su parametri konzervativnog modela pješaka preuzeti iz [22], a parametri agresivnog modela su dobiveni perturbiranjem parametara umjerenog modela pješaka. On služi kao neka vrsta testiranja modela, u smislu da se može promatrati koliko dobro takvi početni parametri mogu odgovoriti na postavljeni model. Način na koji svaki od pješaka predviđa vjerojatnosti realiziranog ishoda najbolje je usporediti grafički slikama u nastavku.



Slika 13 Grafički prikaz modela (lijevo), udio vjerojatnosti realiziranog ishoda za jako perturbiranog pješaka (desno)



**Slika 14 Grafički prikaz modela (lijevo), udio vjerojatnosti realiziranog ishoda za konzervativnog pješaka (desno)**



**Slika 15 Grafički prikaz modela (lijevo), udio vjerojatnosti realiziranog ishoda za agresivnog pješaka (desno)**

Analiza se provodi isto kao i za umjerenog pješaka, te se iz grafova može zaključiti da su parametri jako perturbiranog pješaka loši u smislu da će za sve vrijednosti brzina i udaljenosti vozila izračunati vjerojatnost  $p_i$  jednaku 1 što bi značilo da će pješak uvijek prijeći, a to nije istina. Pješak prelazi samo u slučaju  $y = 1$ , označeno crvenim kružićima na grafu, a ne i za sve ove zelene kružice koji su smješteni u žuto područje. To se može vidjeti i iz histograma, gdje je za 40% slučajeva model izračunao vjerojatnost jednaku 0 da će se ti slučajevi dogoditi.

Sljedeći pješak je konzervativni. Za njega se vidi da ima jako suženo žuto područje, što znači da će u većini slučajeva davati vjerojatnost prelaska  $p_i$  jednaku 0. Znači da je vjerojatnost kriva za sve ove crvene točke smještene u plavome području. Histogram mu se ne razlikuje značajno od histograma jako perturbiranog pješaka.

Zadnji po redu je agresivni pješak. On za širi raspon pozicija i brzina vozila daje prilično veliku vjerojatnost prelaska, ali manje nego jako perturbirani pješak. Može se vidjeti da on čak dosta dobro predviđa vjerojatnosti prelaska  $p_i$ , za one vrijednosti koje su uistinu 1. Slabije smješta vjerojatnosti 0, odnosno, one se nalaze u ovom graničnom području između žutog i plavog dijela grafa. Histogram pokazuje dosta dobre rezultate, sličnije onima za umjerenog pješaka. Zaključak je taj da pomoću podataka generiranih na temelju umjerenog pješaka, agresivni pješak može najbolje predvidjeti pomoću svojih parametara, a konzervativni i jako perturbirani daju jako slične odzive. Još jedan od pokazatelja koliko dobro model predviđa je i vrijednost točnosti. Ona se računa formulama:

$$a = a + (p_i \geq 0.5 \wedge y(i) = 1) \vee (p_i < 0.5 \wedge y(i) = 0), \quad (14)$$

gdje  $a$  označava varijablu koja se povećava za 1 za svaku točku podataka kada model dobro predviđa odluku pješaka prema vjerojatnosti  $p_i$  i stvarnoj odluci  $y$ . Nakon što su sve točke podataka obrađene, računa se ukupna točnost modela prema formuli:

$$\text{točnost} = \frac{TP + TN}{\text{br\_podataka}}, \quad (15)$$

gdje  $TP$  predstavlja broj točnih pozitivnih predviđanja (pješak je prešao cestu i model je točno predvidio), a  $TN$  predstavlja broj točnih negativnih predviđanja (pješak nije prešao cestu i model je točno predvidio). U Tablica 6 prikazane su uspoređene vrijednosti točnosti za sve tipove pješaka, uključujući i umjerenog.

**Tablica 6 Usporedba točnosti**

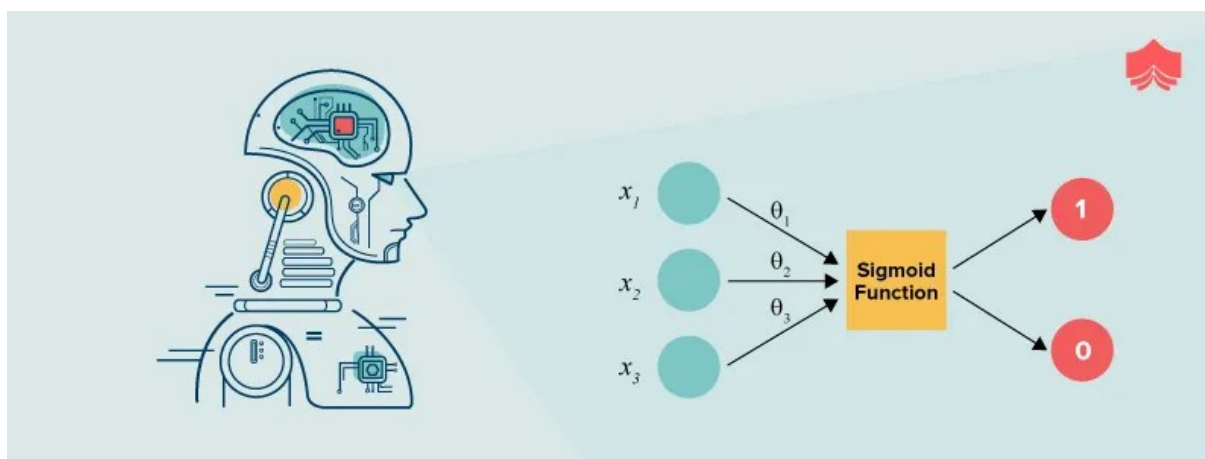
TIP PJEŠAKA	Jako perturbirani	Konzervativni	Umjereni	Agresivni
TOČNOST	59%	50.5%	94.3%	67.8%

Vrijednosti točnosti su očekivane, kada se usporede s prethodnom analizom grafova.

## 4. RAČUNALNO EFIKASAN PREDIKCIJSKI MODEL PONAŠANJA PJEŠAKA TEMELJEN NA LOGISTIČKOJ REGRESIJI

### 4.1. Općenito o logističkoj regresiji

Definicija logističke regresije kaže da je to tehnika u kojoj se nepoznate vrijednosti diskretne varijable predviđaju na temelju poznatih vrijednosti jedne ili više kontinuiranih i/ili diskretnih varijabli [23]. Još jedna interpretacija da je to odgovarajuća regresijska analiza za provođenje kada je zavisna varijabla dihotomna (binarna). Kao i sve regresijske analize, logistička regresija je prediktivna analiza. Logistička regresija koristi se za opisivanje podataka i za objašnjenje odnosa između jedne ovisne binarne varijable i jedne ili više nominalnih, ordinalnih, intervalnih ili omjernih neovisnih varijabli [24]. Ukratko, njezin cilj je predviđanje vjerojatnosti da će neki ulazni primjerak pripadati određenoj klasi (0 ili 1). Primjer klasifikacije ilustriran je Slika 16.



Slika 16 Ilustracija klasifikacije pomoću logističke regresije [25]

Sada kada je objašnjeno što je to logistička regresija, treba spomenuti i da postoje različite vrste logističke regresije. Prva je binarna logistička regresija koja se koristi za predviđanje vjerojatnosti binarnog ishoda, kao što je da ili ne, točno ili netočno ili 0 ili 1. Na primjer, može se koristiti za predviđanje ima li pacijent bolest ili ne, hoće li se kredit vratiti ili ne i sl. Sljedeća vrsta je multinomna logistička regresija koja se koristi za predviđanje vjerojatnosti jednog od tri ili više mogućih ishoda, kao što je vrsta proizvoda koju će kupac kupiti, ocjena koju će kupac dati proizvodu itd. Posljednja je ordinalna logistička regresija i koristi se za predviđanje vjerojatnosti ishoda koji spada u unaprijed određeni redoslijed, kao što je razina zadovoljstva kupaca, ozbiljnost bolesti itd. U ovome radu koristi se prva spomenuta vrsta, binarna logistička regresija. Ona je implementirana sljedećim koracima:

1. Priprema podataka: podaci su u formatu tako da svaki redak predstavlja jedno opažanje, a svaki stupac drugu varijablu. Ciljana varijabla, ona koju želimo predvidjeti, je zapisana u binarnom formatu (0/1).
2. Uvježbavanje modela: model se uči tako da mu se „pokazuju“ podaci za uvježbavanje. To uključuje pronalaženje vrijednosti parametara modela koji minimiziraju pogrešku u podacima o obuci.
3. Ocjenjivanje modela: na temelju zadržanih testnih podataka kako bi se procijenila njegova izvedba na nevidljivim podacima.
4. Korištenje modela za predviđanje: Nakon što je model obučen i procijenjen, može se koristiti za predviđanje ishoda na novim podacima.

#### 4.1.1. *Logistička (sigmoidna) funkcija*

Pretpostavka na kojoj se temelji logistička regresija glasi:

$$0 \leq h_{\theta}(x) \leq 1, \quad (16)$$

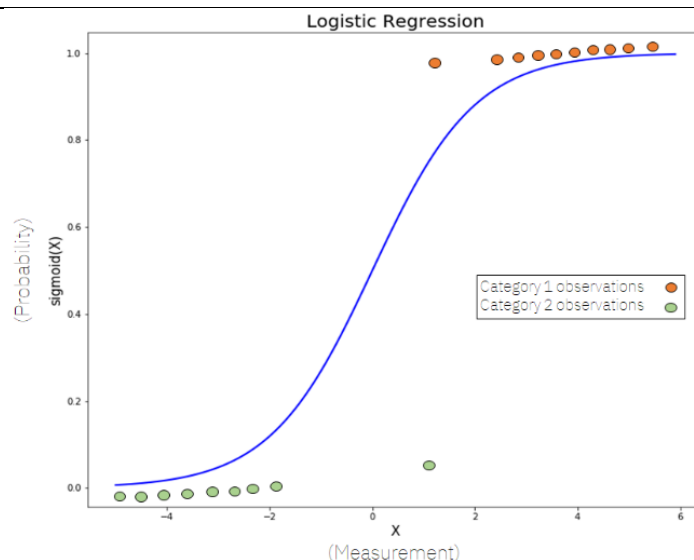
gdje je:

$$h_{\theta}(x) = g(\theta^T x) = \frac{1}{1 + e^{-\theta^T x}}. \quad (17)$$

Ako  $\theta^T x$  zamijenimo s  $z$ , dobivamo izraz:

$$h_{\theta}(x) = \frac{1}{1 + e^{-z}}, \quad (18)$$

gdje je  $h_{\theta}(x)$  hipoteza modela, a  $\theta$  su parametri modela koji se žele naučiti. To su oni isti parametri spomenuti u poglavlju 3. Formula (18) prikazuje konačnu sigmoidnu funkciju koja će se koristiti u ovome radu, a primjer njezinog grafičkog prikazan je na Slika 17. Sigmoidna funkcija tvori graf u obliku slova S, što znači da kako  $x$  raste vjerojatnost se asimptotski približava 1, a kako se  $x$  približava negativnoj beskonačnosti, vjerojatnost postaje 0. Model postavlja prag koji odlučuje koji se raspon vjerojatnosti preslikava na koju binarnu varijablu. Pretpostavka je da postoje 2 moguća ishoda, istinito i lažno, i prag se tipično postavlja na 0,5. Vjerojatnost manja od 0,5 bila bi preslikana u ishod lažan, a vjerojatnost veća ili jednaka 0,5 bila bi preslikana u ishod istinit.



Slika 17 Sigmoidna funkcija [26]

#### 4.1.2. Funkcija troška

Nakon prikaza sigmoidalne funkcije, slijedi definiranje funkcije troška. Funkcija troška predstavlja cilj optimizacije, tj. stvara se funkcija troška i minimizira se tako da se može razviti točan model s minimalnom pogreškom. Ona mjeri razliku između stvarnih izlaza (tj. stvarnih klasa) i predviđenih izlaza modela te se koristi kako bi se ažurirali parametri modela tijekom postupka treniranja. Za logističku regresiju, funkcija troška prikazana je u obliku (preuzeto iz [27]):

$$\text{Cost}(h_{\theta}(x), y) = \begin{cases} -\log(h_{\theta}(x)) & \text{ako je } y = 1, \\ -\log(1 - h_{\theta}(x)) & \text{ako je } y = 0. \end{cases} \quad (19)$$

Funkcije iz (19) mogu biti spojene u jednu funkciju:

$$J(\theta) = -\frac{1}{m} \sum [y^{(i)} \log(h_{\theta}(x^{(i)})) + (1 - y^{(i)}) \log(1 - h_{\theta}(x^{(i)}))], \quad (20)$$

gdje su:

$J(\theta)$  funkcija troška koju treba minimizirati,

$m$  broj primjera u skupu podataka za treniranje,

$\theta$  su parametri modela koje želimo naučiti,

$x^{(i)}$  su ulazni primjeri,

$h_{\theta}(x^{(i)})$  je hipoteza modela tj. predviđena vjerojatnost da  $x^{(i)}$  pripada klasi 1,

$y^{(i)}$  je stvarna klasa primjera  $x^{(i)}$ .



### 4.1.3. *Gradijentni spust*

Sljedeće pitanje je kako pronaći iznose parametara modela koji minimiziraju funkciju troška, a to se postiže pomoću metode gradijentnog spusta. Taj gradijentni spust se mora provesti na svakom parametru:

$$\theta_j = \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta), \quad (21)$$

$$\frac{\partial}{\partial \theta_j} J(\theta) = \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)}) x_j^{(i)} \quad (22)$$

gdje  $\alpha$  predstavlja stopu učenja. Nakon što je opisana metoda logističke regresije, slijedi predstavljanje načina na koji je ona primijenjena na predstavljeni problem predviđanja ponašanja pješaka na nesignaliziranom pješačkom prijelazu.

## 4.2. *Predikcijski model temeljen na logističkoj regresiji*

Niže su navedene jednadžbe preuzete iz [21] prema kojima se vidi da je odabrana ista struktura predikcijskog modela kao i ona simulacijskog modela dane jednadžbom (10). Parametri predikcijskog modela označeni su simbolima „ $\hat{\cdot}$ “ kako bi se naglasilo da se mogu razlikovati od stvarnih/idealnih parametara simulacijskog modela. Važno je napomenuti da predikcijski model ima samo četiri parametra, što ga čini računalno učinkovitim.

$$U = \hat{\alpha} + \hat{\beta}_1 v_p + \hat{\beta}_2 v_v + \hat{\beta}_3 |s_v|, \quad (23)$$

$$\hat{p}_{cross} = \frac{e^U}{1 + e^U}, \quad (24)$$

U poglavlju 3 predstavljena je pred-analiza podataka, koja sada služi za stvaranje predikcijskog modela za slučaj vođenja autonomnog vozila u otvorenoj petlji. Postupak kreće vrlo slično onome iz pred-analize, učitavanjem istih onih podataka provođenjem 1000 simulacija interakcije umjerenog pješaka i vozila na temelju prethodno opisanog simulacijskog modela. Međutim, osim toga skupa, unose se i podaci koji će biti korišteni za testiranje modela. Oni su generirani na isti način, samo je to skup od 1000 skroz novih podataka koje model nije „vidio“ u fazi treniranja, što je bitno kako bi testiranje bilo nepristrano.

Osim učitavanja podataka, inicijalna konfiguracija obuhvaća inicijalizaciju maksimalnog broja podataka za koji se vrši treniranje, grupa podataka predstavlja broj podataka koji će se koristiti u svakoj iteraciji treniranja, a postavljena je na iterativno dodavanje 50 po 50 podataka. Zatim se namješta već spomenuta stopa učenja  $\alpha$  iz jednadžbe (21), a uključena je u izračun gradijentnog spusta. Sljedeće što se namješta je broj prolazaka gradijentnog spusta,

odnosno koliko puta će model proći kroz funkciju troška s ciljem njezine minimizacije. Kasnije će biti prikazano koliko rezultati treniranja, a posljedično testiranja, ovise o podešavanju tog broja prolazaka. Slijedi postavljanje praga odlučivanja, koji se koristi za klasifikaciju izlaza modela u binarnu odluku (0 ili 1), što je već spomenuto u potpoglavlju 4.1.1. Iduća stavka je jedna od najvažnijih, a to je postavljanje inicijalnih parametara modela. Ovdje se za primjer treniranja mogu uzeti bilo koja od četiri tipa prethodno definiranih pješaka (v. Tablica 4 i Tablica 5), odnosno, njihovi parametri nazvani „beta\_param\_ped...“, a kao idealni parametri postavljeni su parametri umjerenog pješaka i njemu se teži tijekom treniranja i testiranja modela.

Nakon inicijalizacije, pokreće se *while* petlja koja je glavna petlja za iterativno treniranje modela, a koja se izvršava dok je broj grupiranih podataka manji ili jednak ukupnom broju dostupnih podataka, u ovome slučaju 1000. Unutar te petlje izvršava se evaluacija idealnog modela, pomoću već predstavljene funkcije troška (20) i pomoću izračuna predviđene vrijednosti vjerojatnosti idealnog modela na treniranju, prema izrazu:

$$\text{vjerojatnost} = \text{sigmoid}(\mathbf{X} \cdot \boldsymbol{\theta}), \quad (25)$$

gdje je  $\mathbf{X}$  matrica ulaznih značajki (brzina i položaj vozila te brzina pješaka koja je =1), a  $\boldsymbol{\theta}$  je vektor parametara odabranog modela, koji je za slučaj evaluacije postavljen na idealnog (umjerenog) pješaka. *Sigmoid* je zatim funkcija koja pretvara linearnu kombinaciju ulaznih značajki i parametara modela u vrijednosti između 0 i 1 (v. Slika 17), što se interpretira kao vjerojatnost pozitivnog ishoda tj. vjerojatnosti da će pješak odabrati prijelaz preko pješačkog prijelaza  $p_i$ . Još jedan način evaluacije idealnog modela je preko izračuna točnosti korištenjem izraza (15). Zatim se ista stvar radi i na skupu podataka za testiranje, učitavanjem testnih podataka. Jedina razlika je u tome što se testni podaci učitavaju svi odjednom, a ne nadodavanjem njih 50 po 50 kao što je to slučaj s podacima za treniranje.

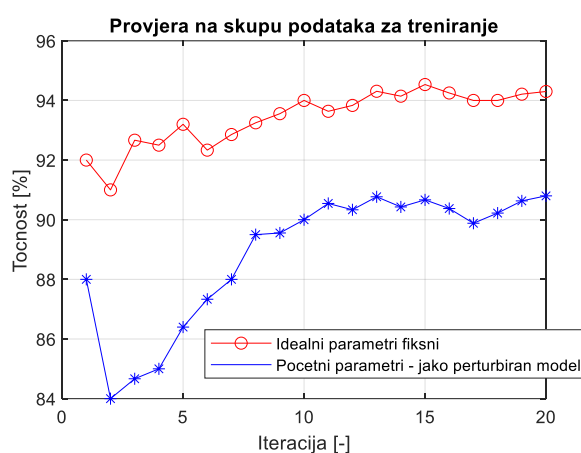
Nakon toga slijedi treniranje modela na trenutnoj grupi podataka, pomoću gradijentnog spusta definiranih izrazima (21) i (22). Pomoću njega se dobiju novi parametri modela i nova funkcija troška. Zatim se vrednuje trenirani model, na isti način kao što se evaluira i idealni model, samo umjesto idealnih parametara, u izraz ulazi trenirani model parametara koji su proizašli iz prethodnog koraka. Ti Nakon izvršavanja *while* petlje, slijedi vizualizacija rezultata pomoću nekoliko grafova. Upravo će preko tih grafova biti provedena analiza rezultata.

#### 4.2.1. Analiza rezultata

Iterativna procedura treniranja modela objašnjena je u potpoglavlju 4.2. Početna inicijalizacija postavljena je na 1000 prolazaka gradijentnog spusta kroz funkciju troška, na

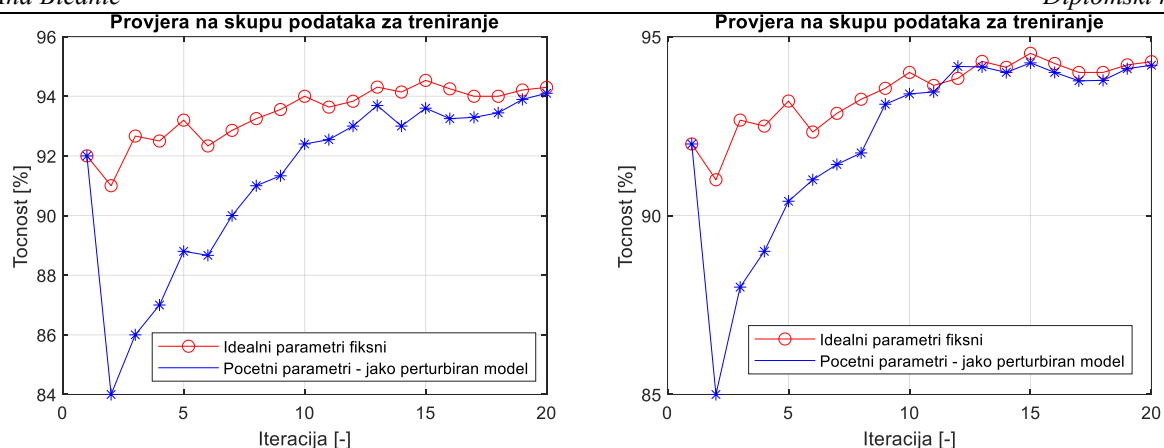
stopu učenja  $\alpha = 0.005$  i na jako perturbiranog tipa pješaka (v. Tablica 5). Nakon toga će se početna inicijalizacija postaviti na 5000 pa na 10000 prolazaka algoritma treniranja, kako bi se dobiveni rezultati mogli bolje objasniti i usporediti.

Analiza započinje usporedbom vrijednosti točnosti i funkcije troška između idealnog modela (s idealnim parametrima) i modela kada su parametri jako perturbirani i to oboje na skupu podataka za treniranje i za testiranje. Također se uspoređuju i vrijednosti parametara. Međutim, obzirom da je brzina pješaka uvijek fiksirana na 1 m/s, pripadni parametar  $\beta_1$  u ovim analizama nije relevantan, te se stoga u analizi  $\alpha$  i  $\beta_1$  razmatraju kombinirano kao  $\alpha + \beta_1$ . Prva slika (v. Slika 18) prikazuje usporedbu točnosti nakon inicijalno postavljenih 1000 prolazaka, na skupu za treniranje.



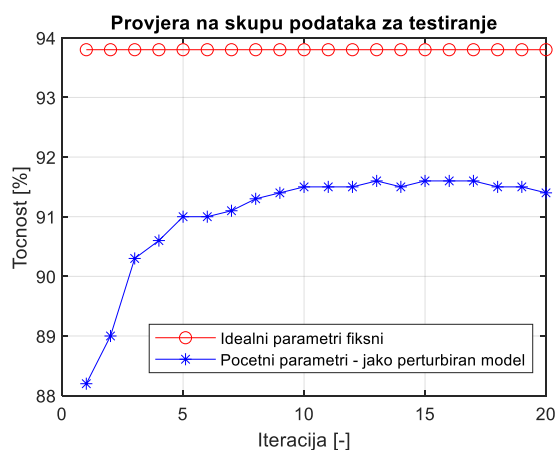
**Slika 18 Usporedba točnosti nakon 1000 prolazaka na skupu za treniranje**

Zanimljiva karakteristika koja se vidi već iz prvoga grafa je ta da za prvu iteraciju, odnosno za skup od 50 ulaznih podataka za treniranje, model temeljen na jako perturbiranom pješaku daje dosta veliku točnost (88%), a zatim se naglo smanji već u idućoj iteraciji. Uzrok tome je da tih 50 podataka nije reprezentativno i da model može imati dobru performansu točnosti na relativno malom skupu podataka, a da pritom ne generalizira dobro. Sada slijedi prikaz točnosti za 5000 i 10000 prolazaka. Zatim slijedi ista usporedba, ali za slučaj kada je broj prolazaka postavljen na 5000 (Slika 19, lijevo) i na 10000 (Slika 19, desno).

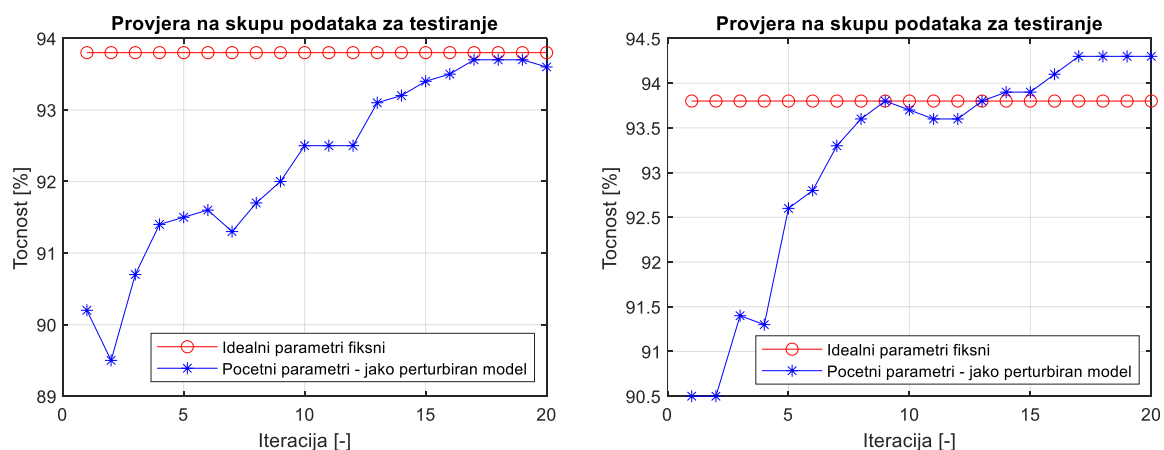


**Slika 19 Usporedba točnosti nakon 5000 (lijevo) i 10000 (desno) prolazaka na skupu za treniranje**

Iz grafova se može iščitati da se za 5000 prolazaka model tek na kraju približi idealnoj točnosti, a za 10000 već nakon 12 iteracija (600 podataka) dobro prati idealnu točnost skupa za treniranje. Slijedi prikaz usporedbe točnosti, ali sada na skupu za testiranje, za 1000 (Slika 20), 5000 (Slika 21, lijevo) i 10000 (Slika 21, desno) prolazaka.

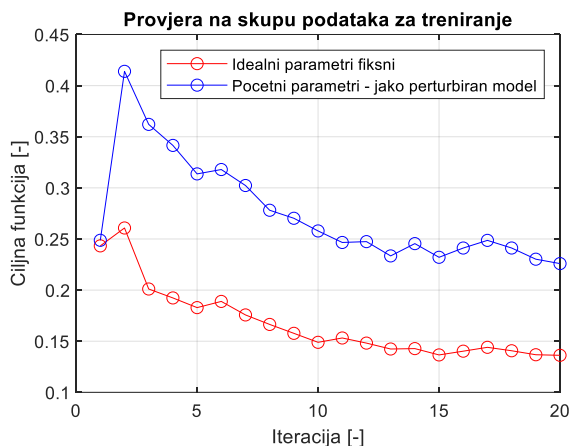


**Slika 20 Usporedba točnosti nakon 1000 prolazaka na skupu za testiranje**

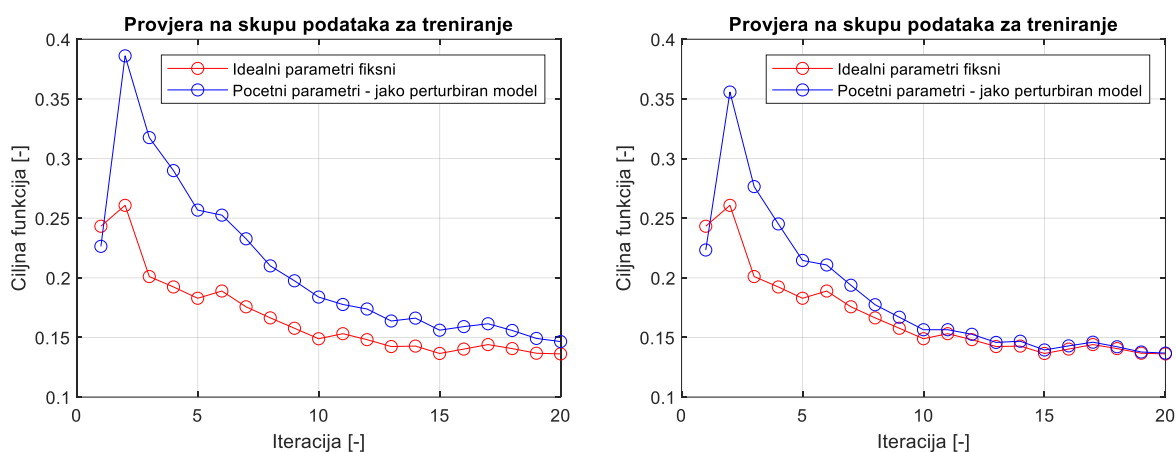


**Slika 21 Usporedba točnosti nakon 5000 (lijevo) i 10000 (desno) prolazaka na skupu za testiranje**

Zanimljivo je primijetiti da se za 5000 prolazaka model približava idealnoj točnosti skupa za testiranje sličnom brzinom kao i na skupu za treniranje, ali i da za 10000 prolazaka model može postići točnost i veću od one idealne. Idealnu točnost dostiže već nakon devet iteracija. Slijedi analiza funkcije troška na skupu za treniranje, prikazana na slikama Slika 22 i Slika 23.

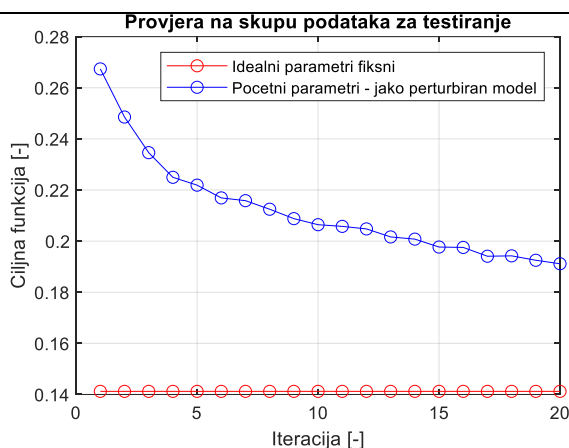


Slika 22 Usporedba funkcije troška nakon 1000 prolazaka na skupu za treniranje

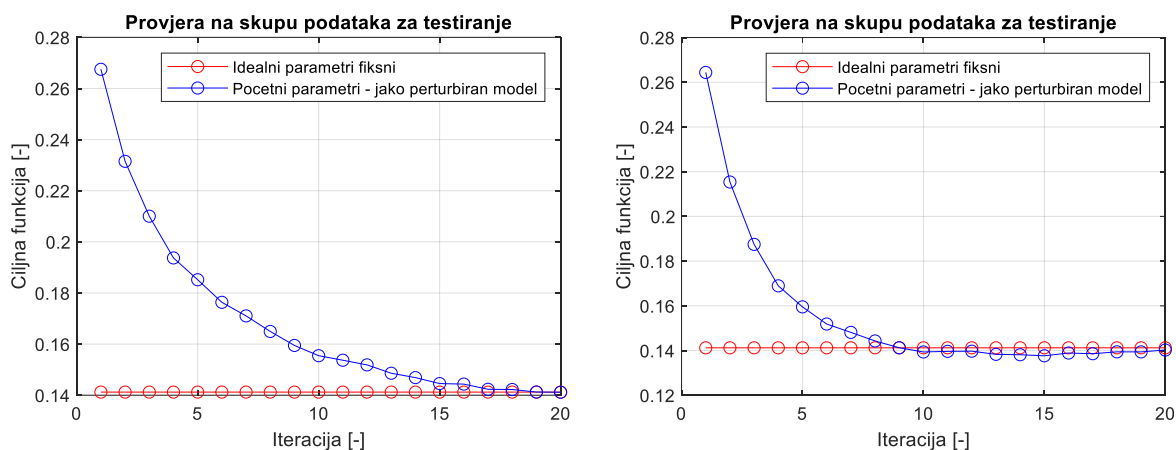


Slika 23 Usporedba funkcije troška nakon 5000 (lijevo) i 10000 (desno) prolazaka na skupu za treniranje

Iz grafa na Slika 22, vidi se da se model tijekom 1000 prolazaka uopće ne približi idealnoj vrijednosti. Na grafovima na Slika 23 se može iščitati da se za 5000 prolazaka model tek na kraju približi idealnoj funkciji cilja, a za 10000 već nakon 12 iteracija (600 podataka) dobro prati idealnu funkciju cilja, isto kao što je to bio slučaj s točnosti. Slijedi prikaz usporedbe funkcije troška, ali sada na skupu za testiranje, za 1000 (Slika 24), 5000 (Slika 25, lijevo) i 10000 (Slika 25, desno) prolazaka.

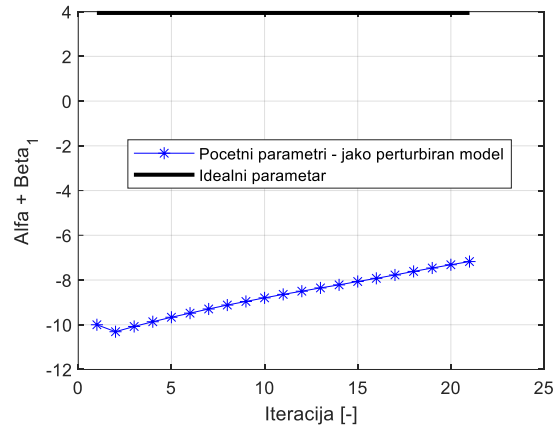


Slika 24 Usporedba funkcije cilja nakon 1000 prolazaka na skupu za testiranje

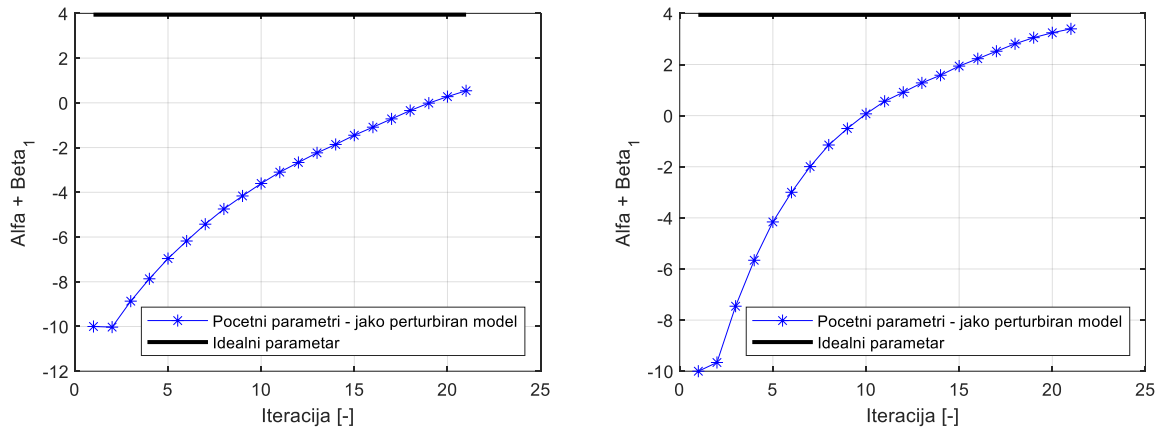


Slika 25 Usporedba funkcije troška nakon 5000 (lijevo) i 10000 prolazaka (desno) na skupu za testiranje

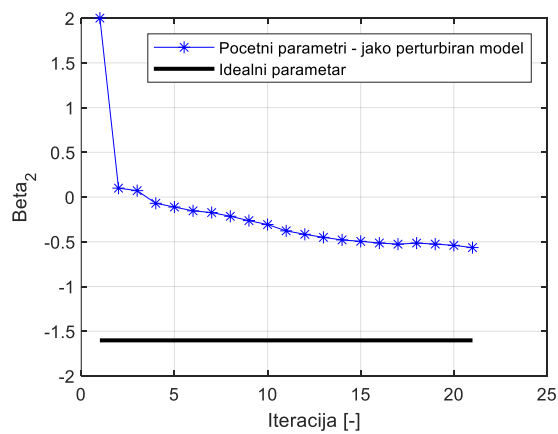
Zanimljivo je primijetiti da se za 5000 prolazaka model približava idealnoj funkciji cilja skupa za testiranje sličnom brzinom kao i na skupu za treniranje, ali i da za 10000 prolazaka model već nakon 9 iteracija dostiže ciljanu funkciju troška. Isto kao i kod točnosti, za 1000 prolazaka se model uopće ne približi idealnoj vrijednosti. Sličan slučaj je bio i sa točnosti u usporedbi s podacima za testiranje. Na posljertku, slijedi analiza grafova koji predstavljaju usporedbu parametara sa onima koji su postavljeni na idealne vrijednosti. Na slikama Slika 26 i Slika 27 prikazani su parametri  $\alpha + \beta_1$ , usporedno za 1000, 5000 i 10000 prolazaka. Slijede ih slike Slika 28 i Slika 29 koje prikazuju istu stvar, samo za parametar  $\beta_2$ , te slike Slika 30 i Slika 31 za  $\beta_3$  parametar.



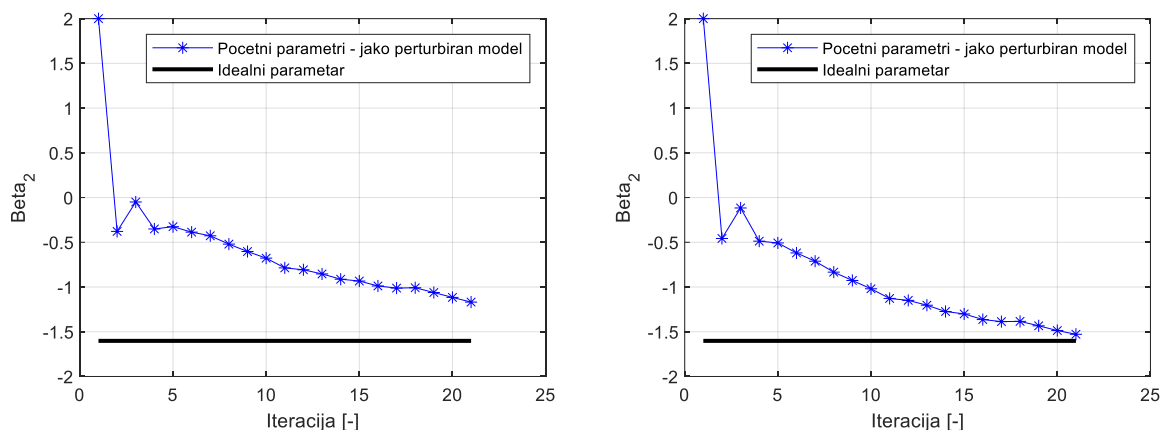
Slika 26 Usporedba  $\alpha + \beta_1$  parametara nakon 1000 prolazaka



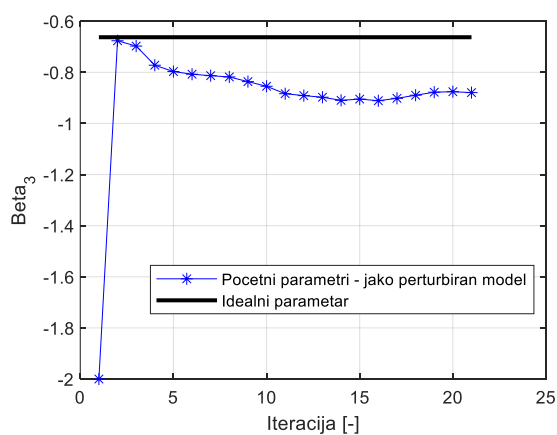
Slika 27 Usporedba  $\alpha + \beta_1$  parametara nakon 5000 (lijevo) i 10000 (desno) prolazaka



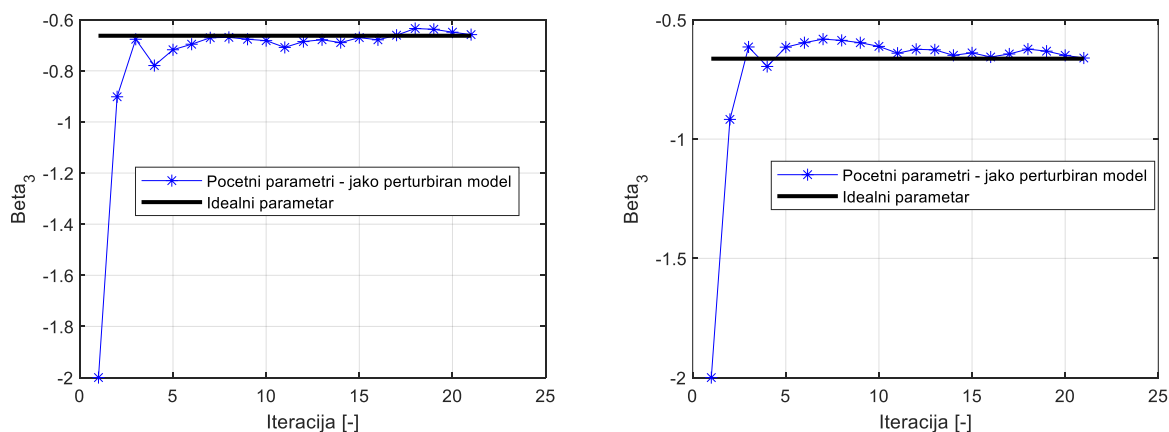
Slika 28 Usporedba  $\beta_2$  parametara nakon 1000 prolazaka



Slika 29 Usporedba  $\beta_2$  parametara nakon 5000 (lijevo) i 10000 (desno) prolazaka



Slika 30 Usporedba  $\beta_3$  parametara nakon 1000 prolazaka



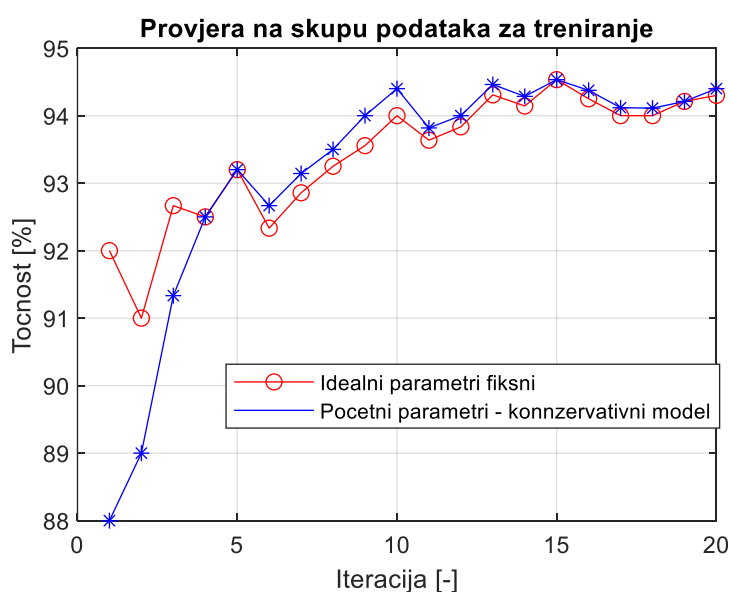
Slika 31 Usporedba  $\beta_3$  parametara nakon 5000 (lijevo) i 10000 (desno) prolazaka

Prikazani grafovi mogu se analizirati skupno, jer svi prikazuju da se parametri jako perturbiranog pješaka tek nakon 10000 prolazaka približe idealnoj vrijednosti, iako ni za 5000 prolazaka nisu daleko od nje.  $\beta_3$  parametar dosegne idealnu vrijednost i nakon 5000 prolazaka



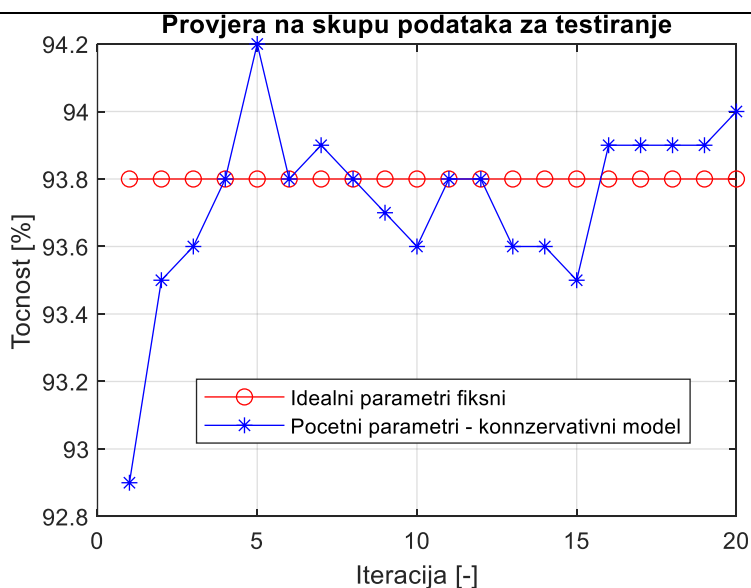
i to već nakon treće iteracije. Samo je za 1000 prolazaka evidentno da to nije dovoljno da bi se parametri približili idealnima.

Ovo je bila analiza za slučaj jako perturbiranog pješaka, sada slijede preostala tri tipa na isti način. Početna inicijalizacija sada je ponovo postavljena na 1000 prolazaka gradijentnog spusta kroz funkciju troška, na stopu učenja  $\alpha = 0.005$  i na konzervativnog tipa pješaka. Analiza započinje usporedbom vrijednosti točnosti i funkcije troška između idealnog modela (sa idealnim parametrima) i to oboje na skupu podataka za treniranje i za testiranje. Prva slika (v. Slika 32) prikazuje usporedbu točnosti nakon inicijalno postavljenih 1000 prolazaka, na skupu za treniranje.



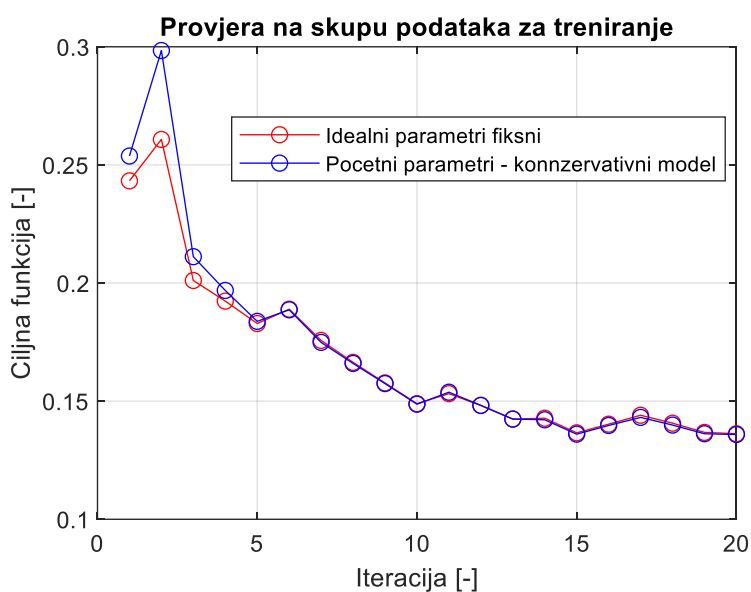
**Slika 32 Usporedba točnosti nakon 1000 prolazaka na skupu za treniranje**

Iz ovoga grafa se može vidjeti da model temeljen na konzervativnom pješaku daje dosta slične vrijednosti točnosti u odnosu na idealne vrijednosti već nakon 5 iteracija (250 podataka), te dalje prati tu točnost sve do zadnje iteracije. Nema smisla ponavljati simulaciju za 5000 prolazaka, kada se sa njih 1000 dobivaju zadovoljavajući rezultati. Slijedi prikaz usporedbe točnosti, ali sada na skupu za testiranje, isto za 1000 prolazaka (Slika 33).



**Slika 33 Usporedba točnosti nakon 1000 prolazaka na skupu za testiranje**

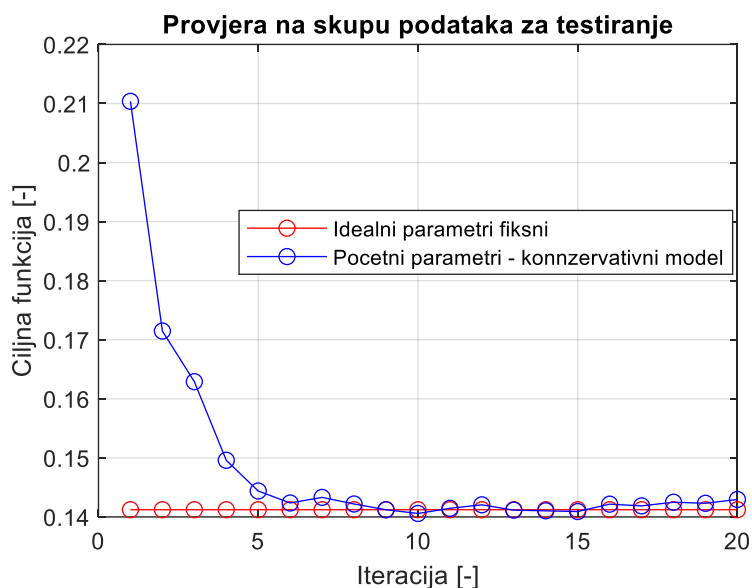
Što se tiče točnosti skupa za testiranje u usporedbi s idealnom vrijednosti, može se vidjeti da ona nakon četiri iteracije dosegne tu idealnu vrijednost, ali nakon toga stalno varira oko nje i to u vrijednostima  $\pm 0.4\%$ . Slijedi analiza funkcije troška na skupu za treniranje, prikazana na Slika 34.



**Slika 34 Usporedba funkcije troška nakon 1000 prolazaka na skupu za treniranje**

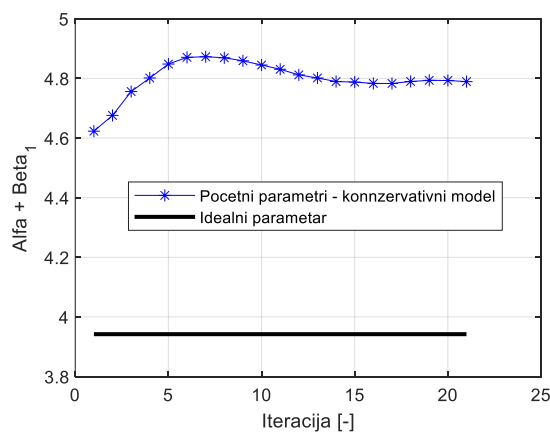
Isto kao i u slučaju točnosti, iz ovoga grafa se može vidjeti da model temeljen na konzervativnom pješaku daje dosta slične vrijednosti funkcije troška u odnosu na idealne vrijednosti već nakon pet iteracija (250 podataka), te dalje prati tu funkciju troška sve do zadnje iteracije. Nema smisla ponavljati simulaciju za 5000 prolazaka, kada se sa njih 1000 dobivaju

zadovoljavajući rezultati. Isto se događa i na skupu za testiranje, samo nešto kasnije, oko 10.-te iteracije, što se vidi na grafu ispod (Slika 35).

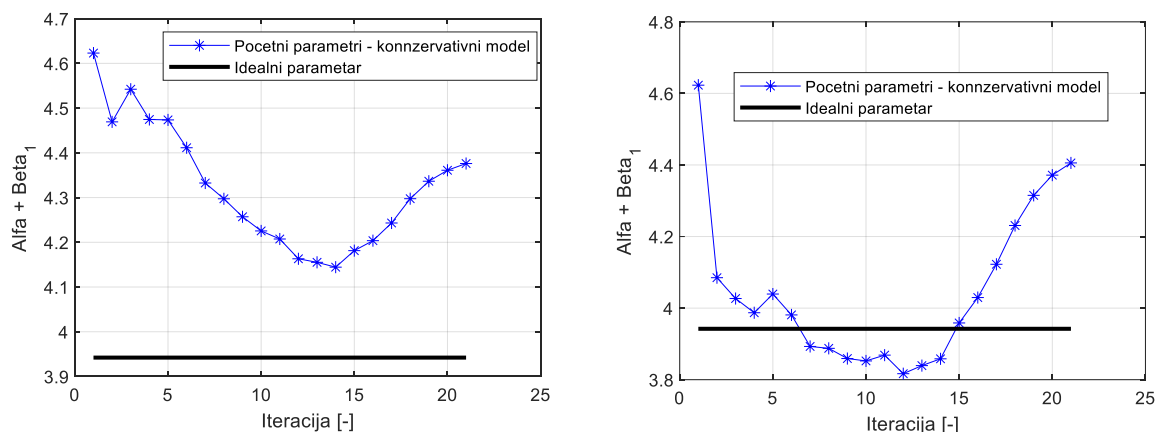


**Slika 35 Usporedba funkcije cilja nakon 1000 prolazaka na skupu za testiranje**

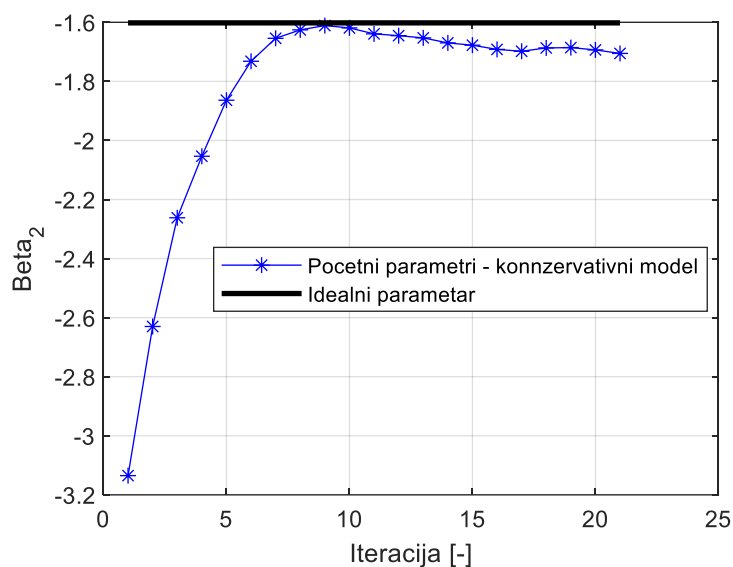
Naposljetku opet slijedi analiza grafova koji predstavljaju usporedbu parametara sa onima koji su postavljeni na idealne vrijednosti. Na slikama Slika 36 i Slika 37 prikazani su parametri  $\alpha + \beta_1$ , usporedno za 1000, 5000 i 10000 prolazaka. Slijede ih slike Slika 38 koje prikazuju parametar  $\beta_2$  samo za 1000 prolazaka i Slika 39 koja prikazuje istu stvar, samo za parametar  $\beta_3$ .



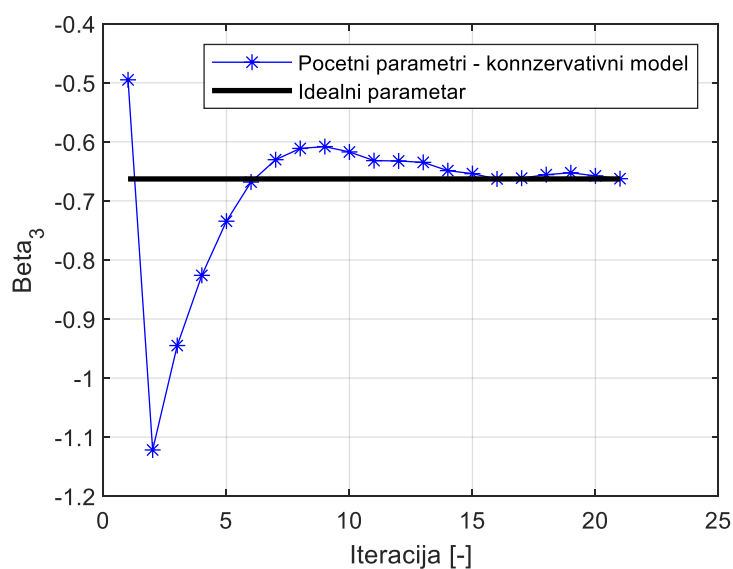
**Slika 36 Usporedba  $\alpha + \beta_1$  parametara nakon 1000 prolazaka**



Slika 37 Usporedba  $\alpha + \beta_1$  parametara nakon 5000 (lijevo) i 10000 (desno) prolazaka



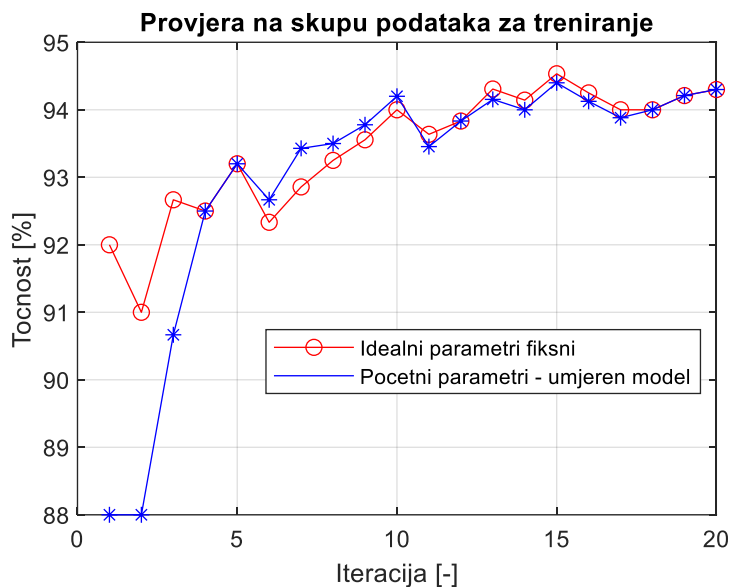
Slika 38 Usporedba  $\beta_2$  parametara nakon 1000 prolazaka



Slika 39 Usporedba  $\beta_3$  parametara nakon 1000 prolazaka

Za grafove na slikama Slika 38 i Slika 39 može se primijetiti da već nakon 1000 prolazaka dolaze do idealne vrijednosti, ali za parametre  $\alpha + \beta 1$  prikazani su i rezultati nakon 5000 i 10000. Iz tih grafova se vidi da tek kada se broj prolazaka podesi na 10000, parametri uspiju doći do idealne vrijednosti, ali vrlo brzo se i odmaknu od nje.

Sljedeća analiza je za parametre umjerenog pješaka. Prva slika (v. Slika 40) prikazuje usporedbu točnosti nakon inicijalno postavljenih 1000 prolazaka, na skupu za treniranje.



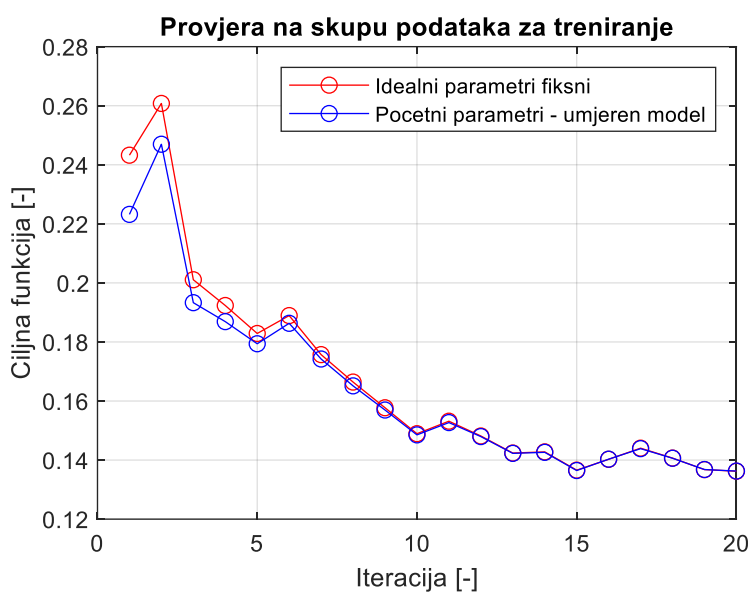
**Slika 40 Usporedba točnosti nakon 1000 prolazaka na skupu za treniranje**

Iz ovoga grafa se može vidjeti da model temeljen na umjerenom pješaku, očekivano, daje dosta slične vrijednosti točnosti u odnosu na idealne vrijednosti već nakon četiri iteracije (200 podataka), te dalje prati tu točnost sve do zadnje iteracije. Nema smisla ponavljati simulaciju za 5000 prolazaka, kada se sa njih 1000 dobivaju zadovoljavajući rezultati. Slijedi prikaz usporedbe točnosti, ali sada na skupu za testiranje, isto za 1000 prolazaka (Slika 41).



**Slika 41 Usporedba točnosti nakon 1000 prolazaka na skupu za testiranje**

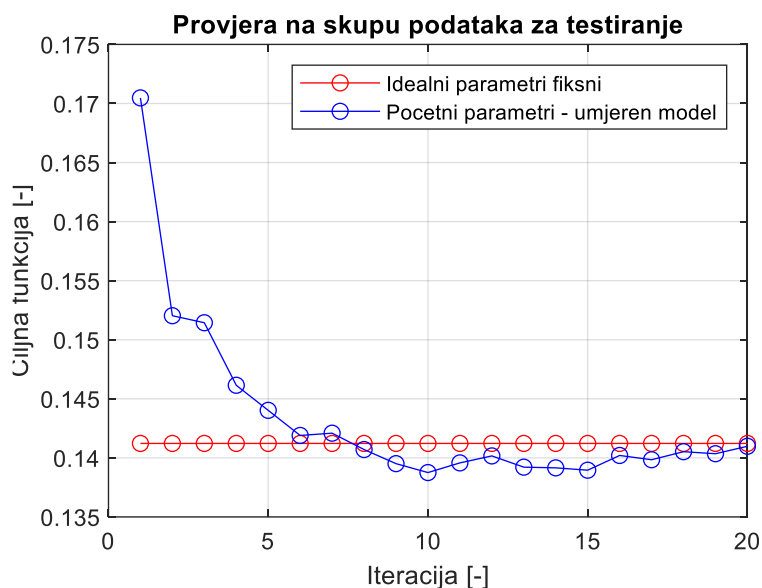
Što se tiče točnosti skupa za testiranje u usporedbi s idealnom vrijednosti, može se vidjeti da ona već nakon 2 iteracije premaši tu idealnu vrijednost te nakon toga stalno varira oko nje i to u jako malom rasponu ( $\pm 0.5\%$ ). Slijedi analiza funkcije troška na skupu za treniranje, prikazana na Slika 42.



**Slika 42 Usporedba funkcije troška nakon 1000 prolazaka na skupu za treniranje**

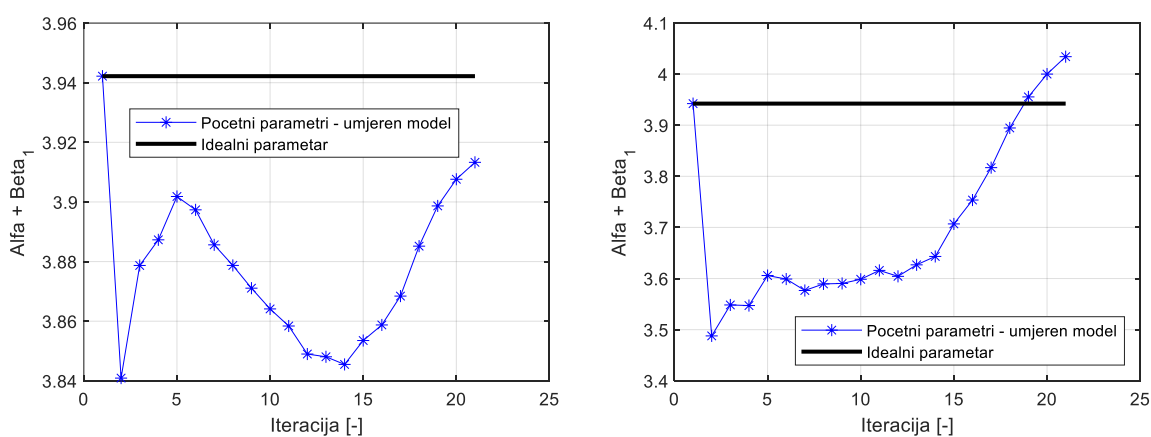
Isto kao i u slučaju točnosti, iz ovoga grafa se može vidjeti da model temeljen na umjerenom pješaku daje dosta slične vrijednosti funkcije troška u odnosu na idealne vrijednosti, nakon otprilike pet iteracija (250 podataka), te nakon desete iteracije prati tu funkciju troška poprimajući identičnu vrijednost. Nema smisla ponavljati simulaciju za 5000 prolazaka, kada se sa njih 1000 dobivaju zadovoljavajući rezultati. Slično se događa i na skupu za testiranje,

gdje se nakon 7 iteracija postigne idealna vrijednost i kasnije se dobivaju vrijednosti jako bliske toj idealnoj, što se vidi na grafu ispod, Slika 43).

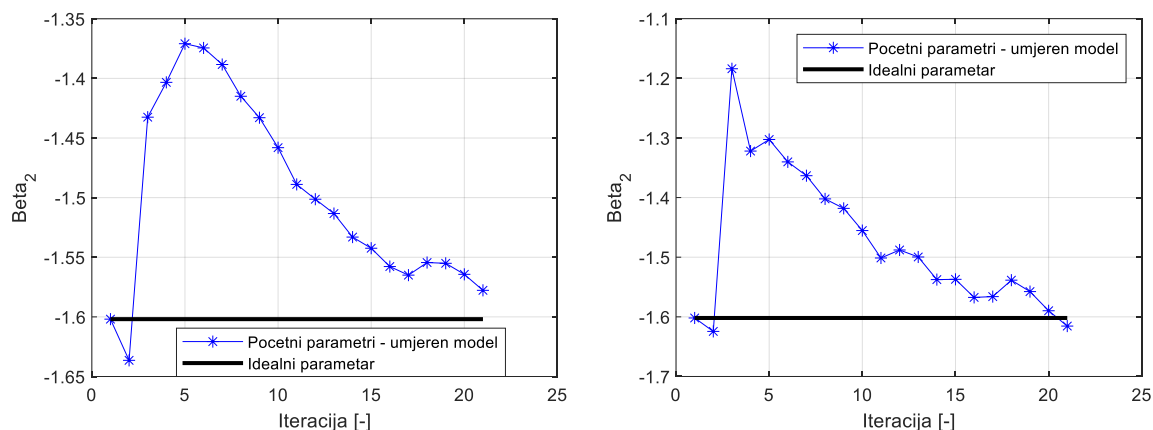


**Slika 43 Usporedba funkcije cilja nakon 1000 prolazaka na skupu za testiranje**

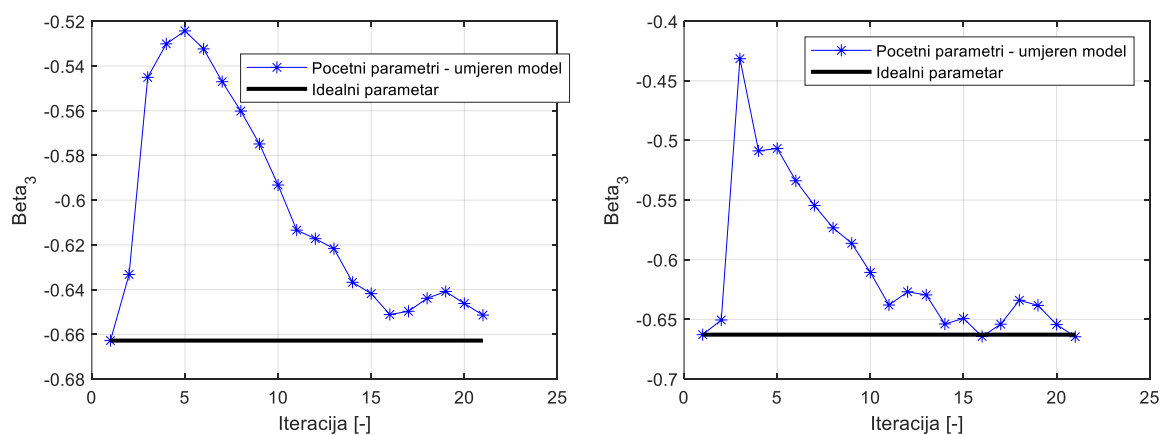
Naposlijetku opet slijedi analiza grafova koji predstavljaju usporedbu parametara sa onima koji su postavljeni na idealne vrijednosti. Ni jedan parametar nije uspio doći do idealne vrijednosti nakon 1000 ponavljanja, ali nakon 5000 svi dolaze do idealne. Na Slika 44 prikazani su parametri  $\alpha + \beta_1$ , usporedno za 1000 i 5000 prolazaka. Slijedi ju Slika 45 koja prikazuje istu stvar samo za parametar  $\beta_2$ , i Slika 46 za parametar  $\beta_3$ .



**Slika 44 Usporedba  $\alpha + \beta_1$  parametara nakon 1000 (lijevo) i 5000 (desno) prolazaka**



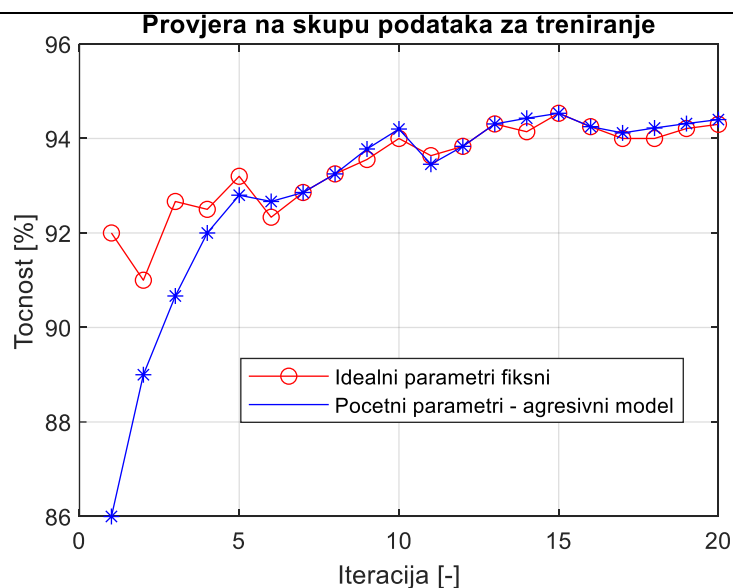
**Slika 45 Usporedba  $\beta_2$  parametara nakon 1000 (lijevo) i 5000 (desno) prolazaka**



**Slika 46 Usporedba  $\beta_3$  parametara nakon 1000 (lijevo) i 5000 (desno) prolazaka**

Još preostaje analizirati podatke kada se kao tip pješaka uzme agresivni. Početna inicijalizacija sada je ponovo postavljena na 1000 prolazaka gradijentnog spusta kroz funkciju troška, na stopu učenja  $\alpha = 0.005$  i na agresivnog tipa pješaka. Slika 47 prikazuje usporedbu točnosti nakon inicijalno postavljenih 1000 prolazaka, na skupu za treniranje





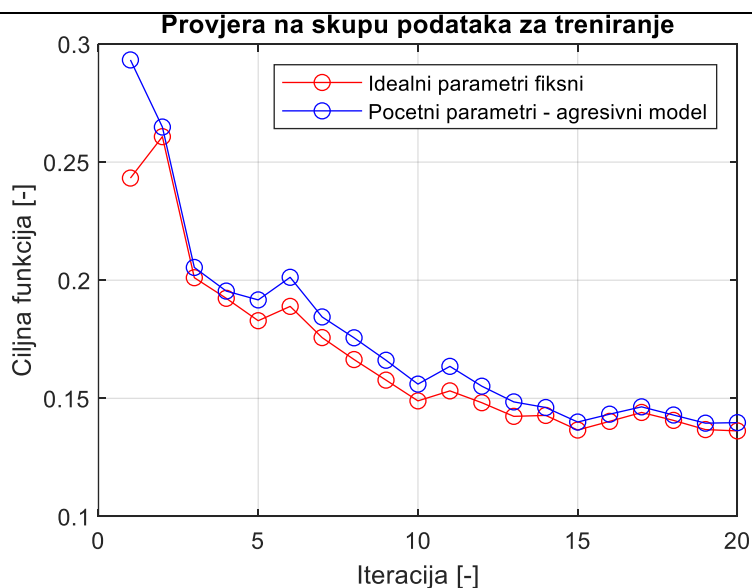
**Slika 47 Usporedba točnosti nakon 1000 prolazaka na skupu za treniranje**

Iz grafa se može vidjeti da model temeljen na agresivnom pješaku, daje dosta slične vrijednosti točnosti u odnosu na idealne vrijednosti već nakon sedme iteracije (350 podataka), te dalje prati tu točnost sve do zadnje iteracije. Slijedi prikaz usporedbe točnosti, ali sada na skupu za testiranje, isto za 1000 prolazaka (Slika 48).



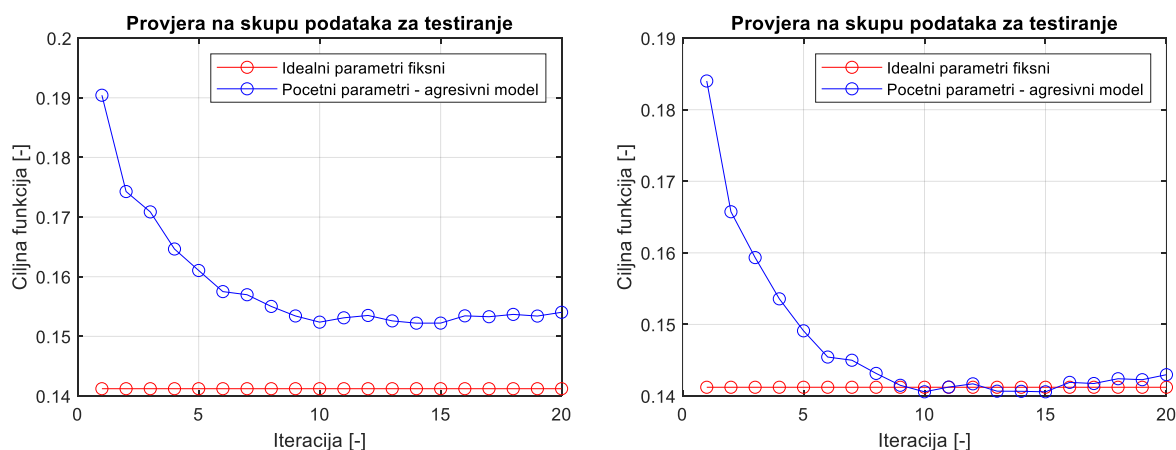
**Slika 48 Usporedba točnosti nakon 1000 prolazaka na skupu za testiranje**

Što se tiče točnosti skupa za testiranje u usporedbi s idealnom vrijednosti, može se vidjeti da ona nakon 7 iteracije dosegne idealnu vrijednost te nakon toga stalno varira oko nje i to u jako malom rasponu ( $\pm 0.6\%$ ). Slijedi analiza funkcije troška na skupu za treniranje, prikazana na Slika 49.



**Slika 49 Usporedba funkcije troška nakon 1000 prolazaka na skupu za treniranje**

Isto kao i u slučaju točnosti, iz ovoga grafa se može vidjeti da model temeljen na agresivnom pješaku daje dosta slične vrijednosti funkcije troška u odnosu na idealne vrijednosti, i to već nakon 2. iteracije te ju dalje prati u jako malom rasponu pogreške. Međutim, na skupu za testiranje nije isti slučaj pa se za usporedbu prikazuje i slučaj za 10000 ponavljanja. Iz grafova se vidi da se nakon 1000 prolazaka (Slika 50, lijevo) vrijednost funkcije troška jako približi idealnoj već nakon 10 iteracija. Pokrenuta je i simulacija za 5000 prolazaka, međutim, ni u jednom trenutku nije dosegnuta idealna funkcija pa su prikazani samo rezultati za 10000 prolazaka (Slika 50, desno).

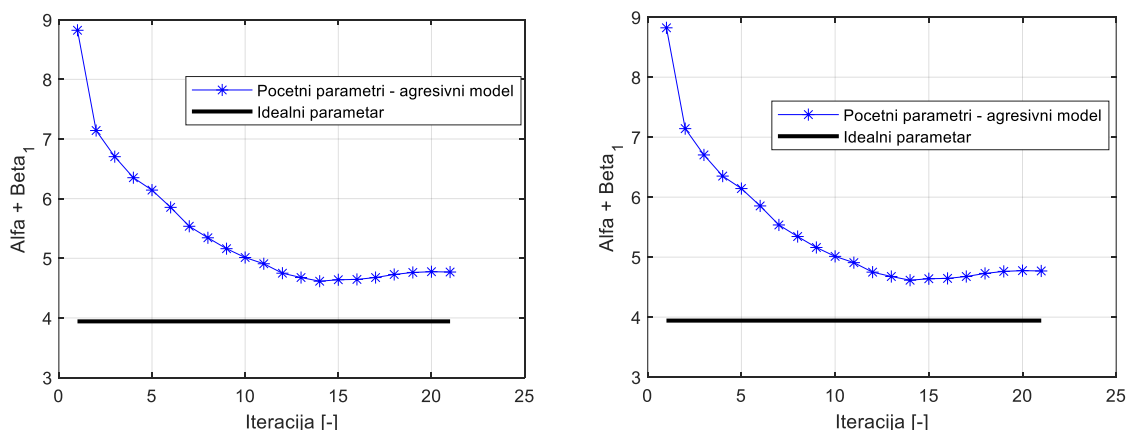


**Slika 50 Usporedba funkcije cilja nakon 1000 (lijevo) i 10000 (desno) prolazaka na skupu za testiranje**

Na poslijetku ponovo slijedi analiza grafova koji predstavljaju usporedbu parametara sa onima koji su postavljeni na idealne vrijednosti. Na Slika 51 prikazani su parametri  $\alpha + \beta_1$ , usporedno

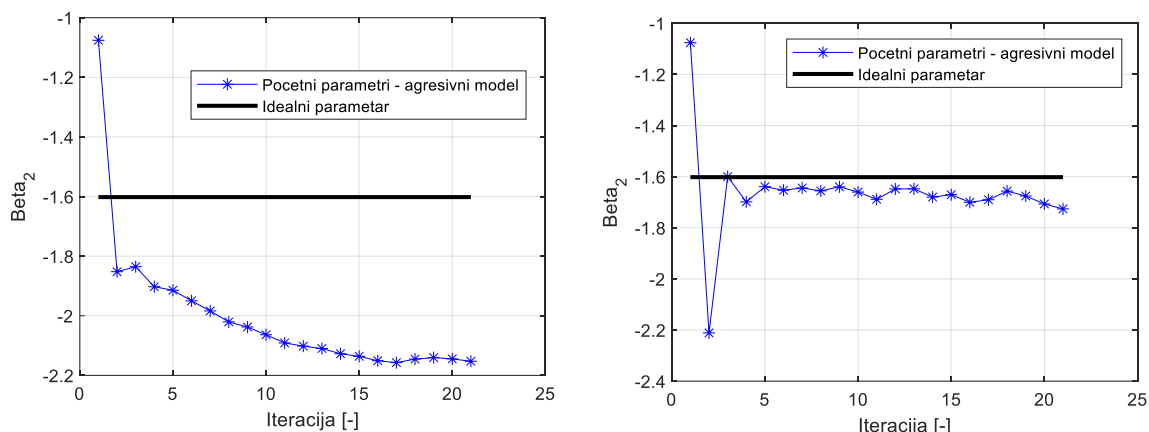
za 1000 i 10000 prolazaka. Slijedi ju Slika 52 koja prikazuje istu stvar samo za parametar  $\beta_2$ , i

Slika 53 za parametar  $\beta_3$ .

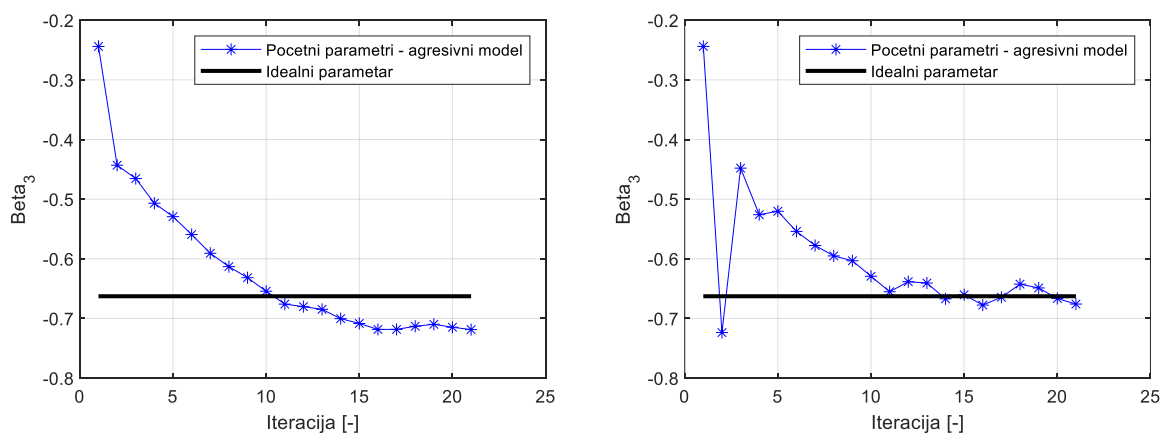


**Slika 51 Usporedba  $\alpha + \beta_1$  parametara nakon 1000 (lijevo) i 10000 (desno) prolazaka**

Primjećuje se da ni nakon 10000 prolazaka,  $\alpha + \beta_1$  parametri neće doseći idealnu vrijednost, iako joj se dosta približe nakon 15-ak iteracija.



**Slika 52 Usporedba  $\beta_2$  parametara nakon 1000 (lijevo) i 10000 (desno) prolazaka**



**Slika 53 Usporedba  $\beta_3$  parametara nakon 1000 (lijevo) i 10000 (desno) prolazaka**

Za parametre  $\beta_2$  i  $\beta_3$  jasno se vidi promjena između lijevih i desnih grafova te se vidi da je za njih bilo dovoljno 10000 prolazaka kako bi se jako dobro približili idealnoj vrijednosti.

Sada se cijela ova analiza može sažeti u jednu tablicu (v. Tablica 7), koja će opisati za koje vrijednosti broja prolazaka i za koliku količinu podataka određeni model daje najbolje rezultate točnosti i funkcije troška.

**Tablica 7 Tablični pregled analize grafova**

		BROJ PROLAZAKA	BROJ PODATAKA
TOČNOST	Jako perturbirani	10000	600
	Konzervativni	1000	250
	Umjereni	1000	200
	Agresivni	1000	350
FUNKCIJA TROŠKA	Jako perturbirani	10000	600
	Konzervativni	1000	250
	Umjereni	1000	250
	Agresivni	1000	100

#### 4.2.2. Filtracija podataka

Nakon analize svih tipova pješaka, može se zaključiti da je modelu i dalje potrebno relativno mnogo točaka kako bi konvergirao prema idealnom modelu. Zato je glavna motivacija ovoga rada pokušati smanjiti taj broj podataka putem filtracije, čime bi se omogućilo da model zadrži samo one točke za koje daje neke značajne informacije i koje odstupaju od očekivanog predviđanja te da se dalje u iteracijama trenira upravo na takvim točkama. Tako bi se skup podataka za treniranje smanjio pa bi se posljedično smanjio i broj unesenih podataka u sustav autonomnog vozila. Filtracija je zamišljena na način da model iterativno prolazi kroz podatke u grupama po 50 i da ih filtrira grupu po grupu, a u iduću iteraciju šalje samo filtrirane podatke.

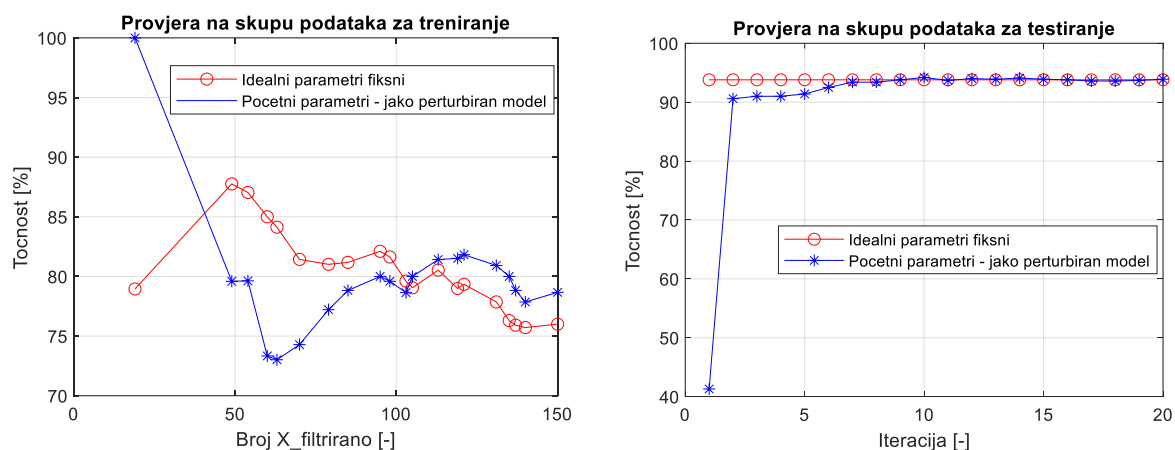
Filtriranje započinje računanjem vjerojatnosti pomoću izraza iz (25) pomoću trenutnih parametara. Pomoću nje se računa vjerojatnost da je realizirani ishod jednak 1 (znači, da pješak prelazi cestu). Uvjeti za filtriranje, tj. propuštanje razmatrane točke u skup za treniranje, su postavljeni temeljem izraza:

- kada je stvarni/realizirani ishod  $y = 0$  i nasumično generiran broj u rasponu od 0 do 1 veći od vrijednosti (1-vjerojatnost) iz (25), ili
- kada je stvarni/realizirani ishod  $y = 1$  i nasumično generiran broj u rasponu od 0 do 1 veći od vrijednosti vjerojatnost iz (25).

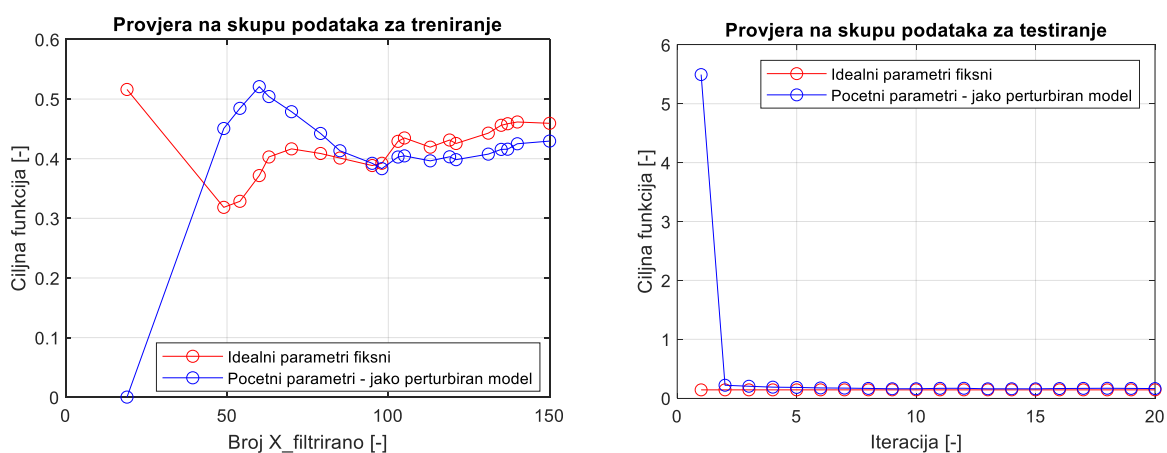
Ako je barem jedan od tih uvjeta zadovoljen, vrijednost logičkog vektora nazvan „indeksi\_za\_filtriranje“ postavlja se na vrijednost *true*, te se pomoću njega promatrani podatak dodaje u skup za treniranje modela. U suprotnom se dodaje u skup odbačenih podataka. Ovaj postupak ponavlja se za svaku iteraciju kroz podatke. Na kraju iteracija, rezultati treniranja i testiranja modela prikazuju se i analiziraju. Važno je napomenuti da se radi o procesu uzorkovanja uz pomoć slučajnih brojeva i da će rezultati biti različiti svaki put kada se algoritam pokrene. U nastavku slijedi spomenuta analiza grafova, a filtriranje će se pokrenuti prvo za jako perturbiranog pješaka te za agresivnog, kako bi se lakše usporedili rezultati.

#### 4.2.2.1. Analiza filtriranih podataka

Prvo će se prikazati graf točnosti (Slika 54) i funkcija troška (Slika 55) za treniranje i testiranje za jako perturbiranog pješaka. Početna inicijalizacija postavljena je na 5000 prolazaka gradijentnog spusta kroz funkciju troška, na stopu učenja  $\alpha = 0.005$ .

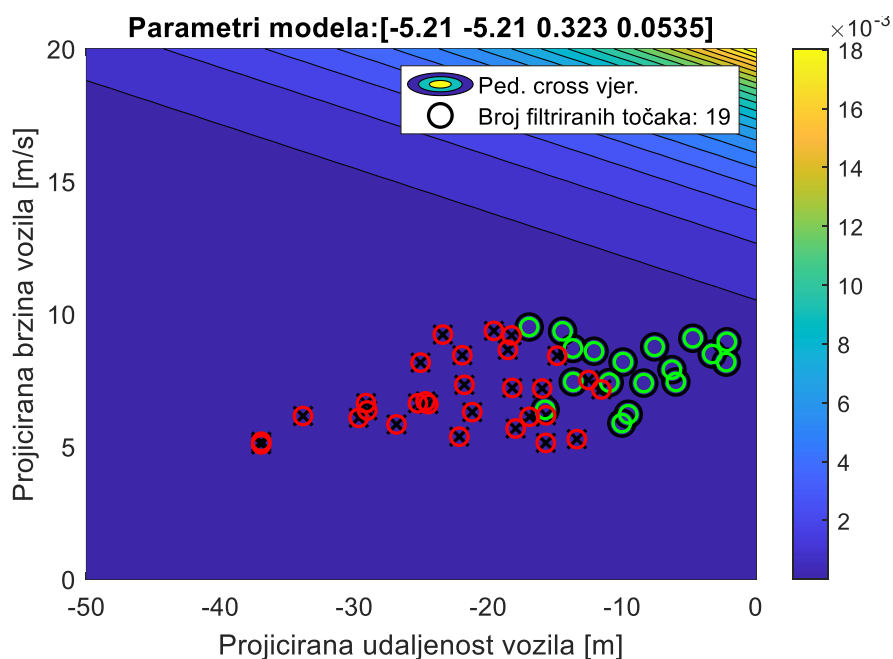


Slika 54 Točnost treniranje (lijevo) u odnosu na testiranje (desno)

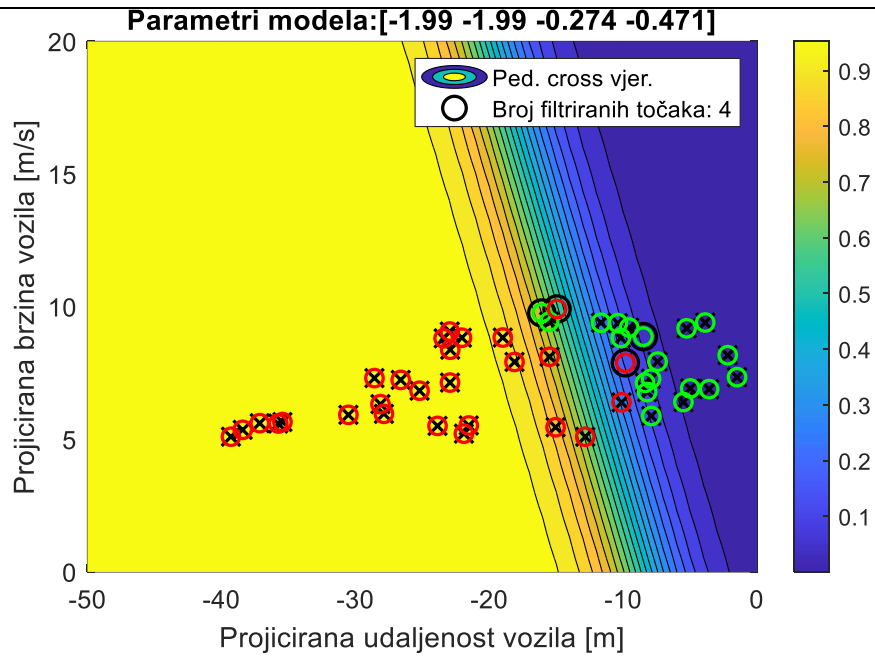


Slika 55 Funkcija troška treniranje (lijevo) u odnosu na testiranje (desno)

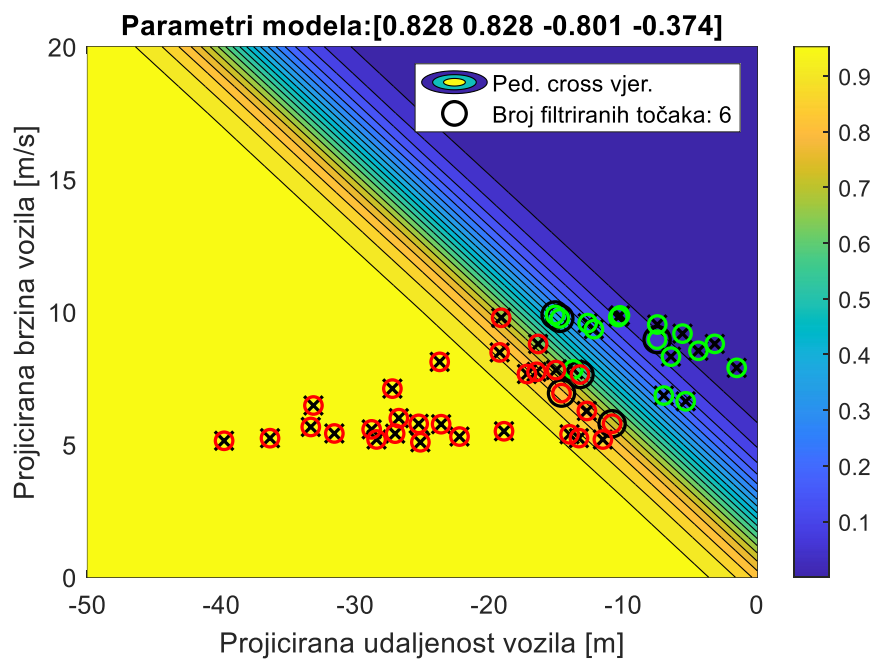
Iz grafa na Slika 54 može se vidjeti da je idealna točnost prilikom treniranja dosta nestabilnija i niža nego kod testiranja. To objašnjava činjenica da model sada uzima samo najkritičnije (nesigurne) točke u obzir i pomoću njih računa idealnu točnost i funkciju troška. S druge strane, prilikom testiranja se uzima svih 1000 točaka, iz skupa podataka za testiranje, bez filtracije te su iz tog razloga njihove vrijednosti iste kao i one bez filtriranja. Činjenica da se već nakon druge iteracije tijekom testiranja vrijednosti odmah približe idealnoj, znači da su parametri dobiveni tijekom treniranja na filtriranim podacima jako dobri i primjenjivi. Jedna neobična anomalija koja se događa na skupu za treniranje za jako perturbiranog pješaka je ta da on u početku postiže točnost 100% i odmah postiže idealnu funkciju troška (Slika 55). To se događa iz razloga što ovaj jako pogrešan model u početku za sve točke (prvih 50) daje vjerojatnost 1 i prema tome odbacuje one stvarne ishode koji su uistinu bili 1 ( $y=1$ ) jer mu to nije ništa novo, a sigurno uzima one suprotne. Znači te vrijednosti koje je zadržao sve imaju isti ishod te na njima provodi treniranje. Za takav skup točaka koje imaju isti ishod, model savršeno uči. Međutim, čim se parametri modela počinju mijenjati, mijenja se i točnost. Tijekom filtracije se ukupno zadržala 152 točka, što je otprilike 6.5 puta manje nego bez filtriranja. Ilustracija točaka koje su se zadržale tijekom filtracije bit će prikazane na idućim grafovima (Slika 56, Slika 57, Slika 58, Slika 59 i Slika 60). One će biti obilježene crnim krugom, a one koje nisu zadržane u filtraciji (odbačene), bit će prikazane crnim križićima. Crveni kružići na grafu predstavljaju slučaj kada pješak prelazi ( $y=1$ ), a zeleni predstavljaju slučaj kada pješak ne prelazi ( $y=0$ ).



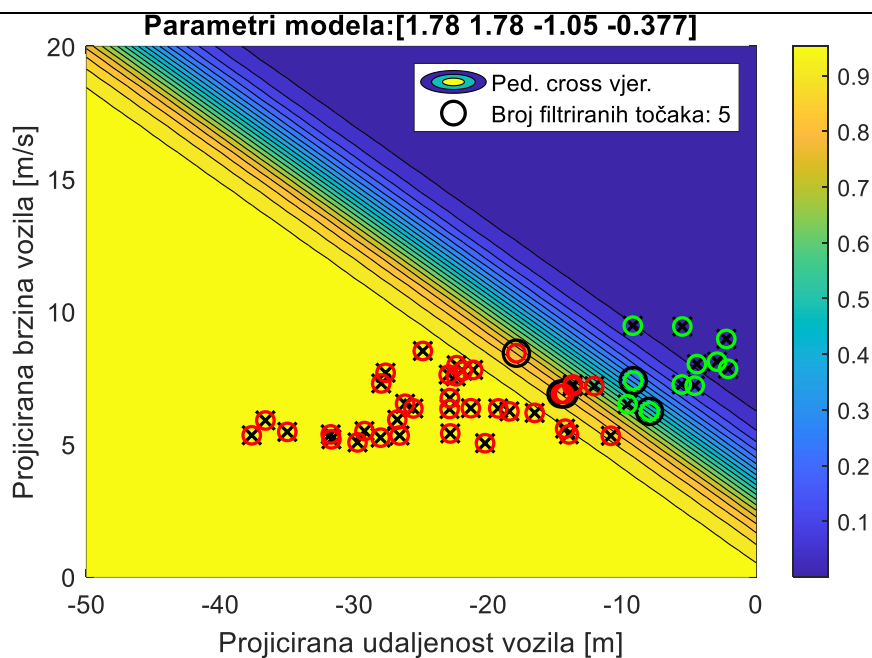
Slika 56 Broj filtriranih točaka nakon 1. iteracije



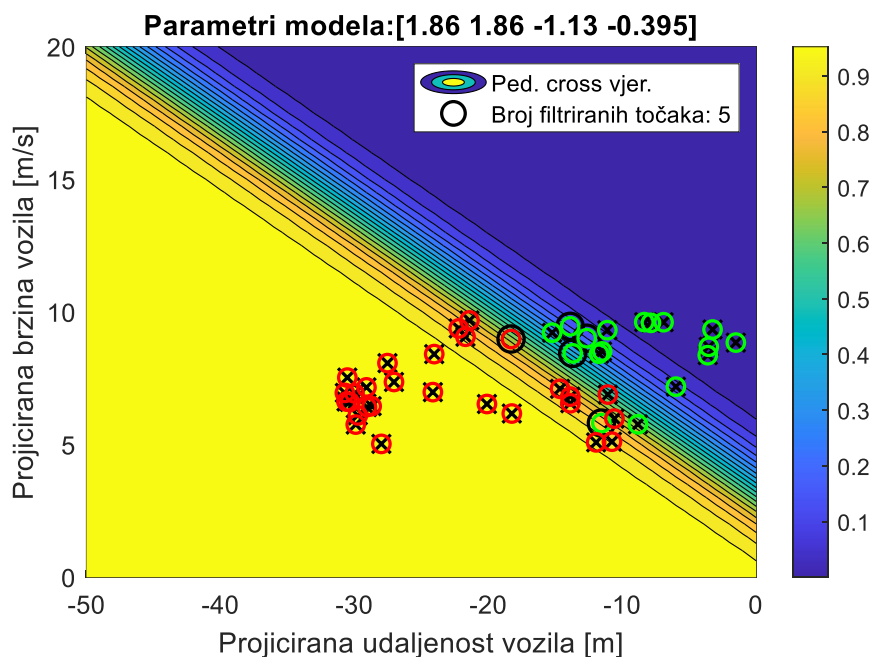
Slika 57 Broj filtriranih točaka nakon 5. iteracije



Slika 58 Broj filtriranih točaka nakon 10. iteracije



Slika 59 Broj filtriranih točaka nakon 15. iteracije

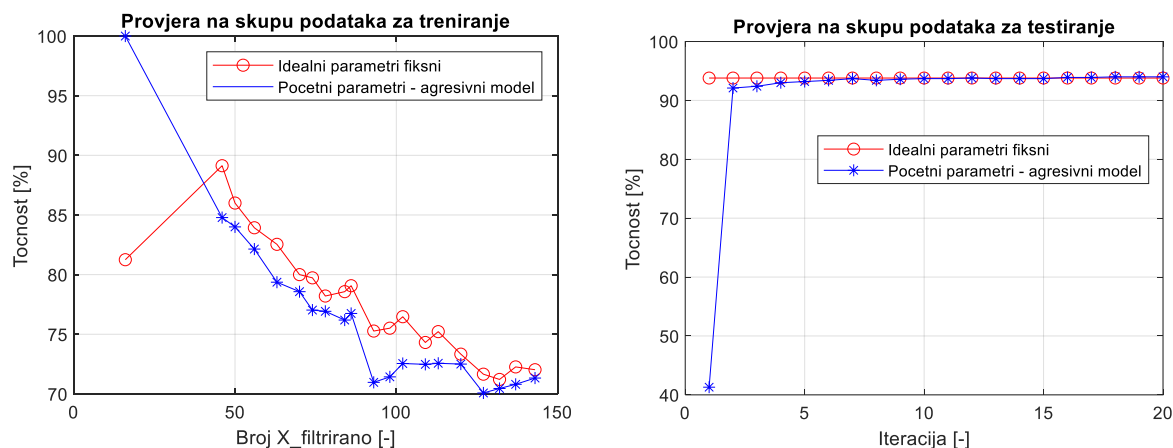


Slika 60 Broj filtriranih točaka nakon 20. iteracije

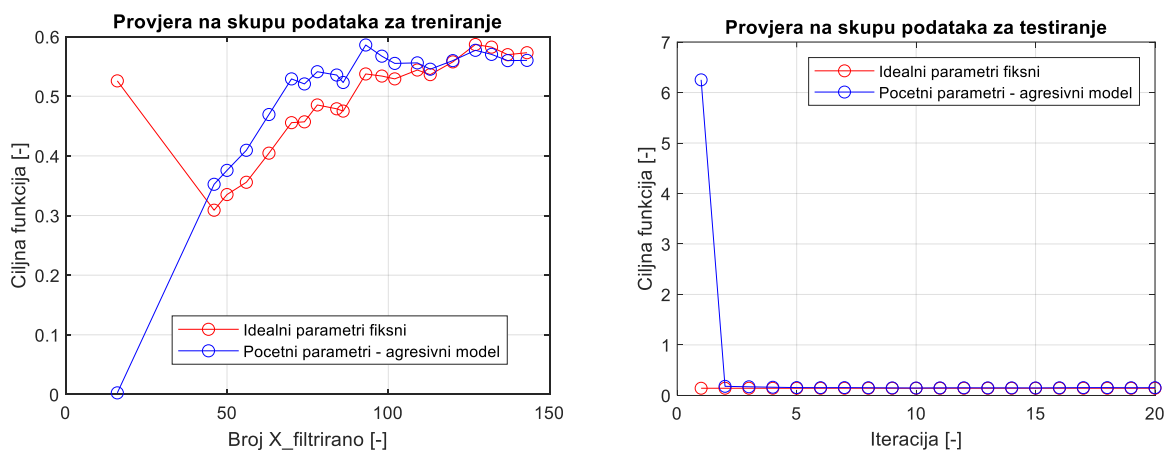
Proučavanjem grafova, može se vidjeti da su točke koje su uzete u obzir nakon filtriranja upravo one koje se nalaze na graničnom dijelu vjerojatnosti prelaska pješaka (u vrijednosti od 0.3 i 0.8). Upravo su te točke najkritičnije i doprinose uspješnom treniranju modela. Isto tako, može se vidjeti da povećanjem broja iteracija, broj točaka koje se zadržavaju nakon filtriranja pada. Započeo je s 19, a nakon toga se kretao oko 5 po iteraciji. Sada slijedi analiza kada je tip pješaka postavljen na agresivnog. Inicijalna konfiguracija sada je postavljena na 1000



prolazaka, a stopa učenja je ponovo  $\alpha = 0.005$ . Prvo će se prikazati graf točnosti (Slika 61) i funkcija troška (Slika 62) za treniranje i testiranje.



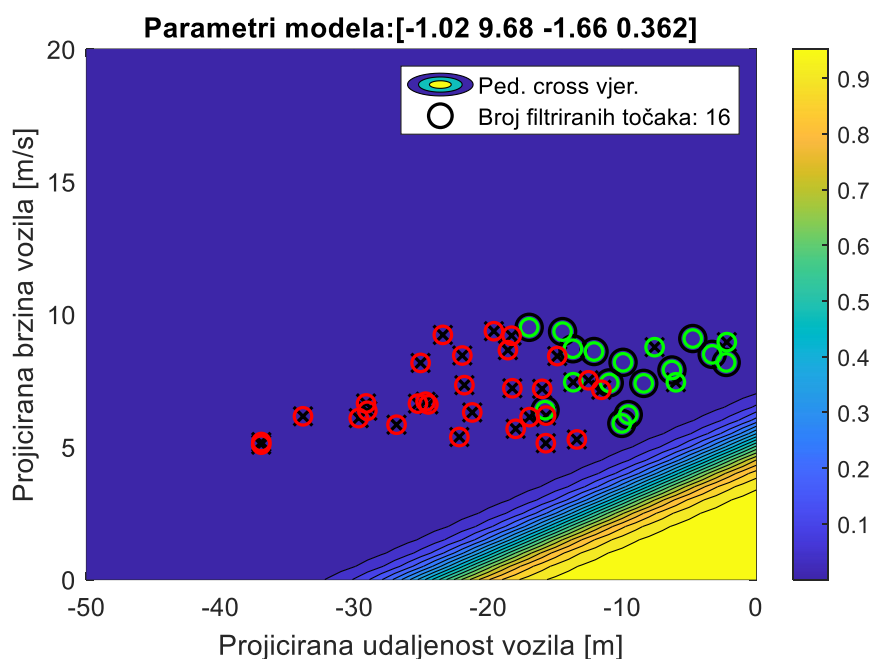
Slika 61 Točnost treniranje (lijevo) u odnosu na testiranje (desno)



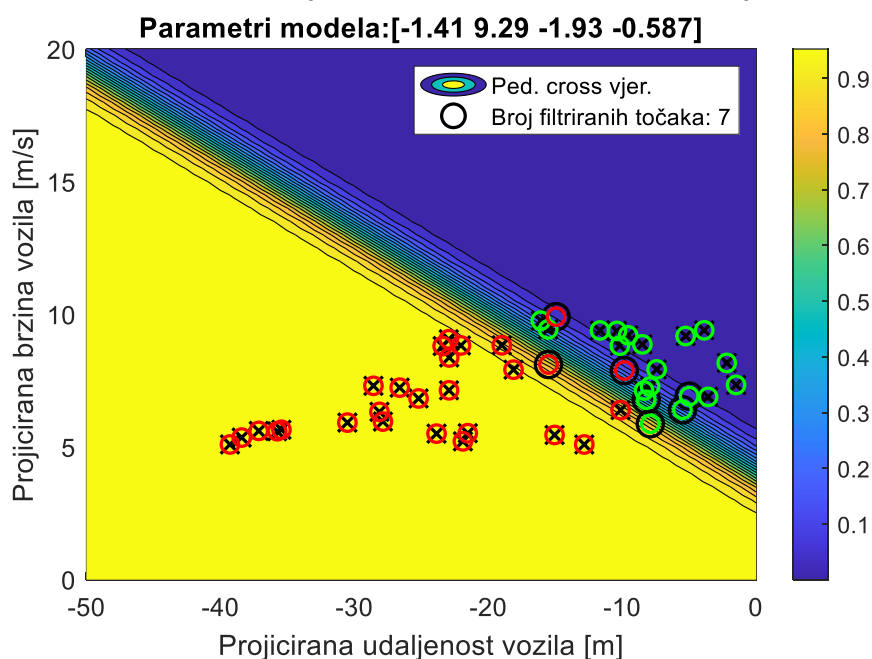
Slika 62 Funkcija troška treniranje (lijevo) u odnosu na testiranje (desno)

Iz grafa na Slika 61 (lijevo) se može vidjeti da je idealna točnost prilikom treniranja dosta nestabilnija i niža nego kod testiranja (Slika 61 (desno)) te se može zaključiti da se isti efekt primijećen na prethodnom slučaju može zamijetiti i ovdje. Prilikom testiranja se ponovo vidi da se već nakon druge iteracije vrijednosti odmah približe idealnoj, znači da su parametri dobiveni tijekom treniranja na filtriranim podacima jako dobri i primjenjivi. Na početku se opet vidi ista anomalija kao i kod jako perturbiranog pješaka, a do nje je došlo iz sličnih razloga, zato što će agresivni model pješaka također u početku sve točke „smješati“ u vjerojatnost 1, jer su mu tako namješteni početni parametri. Ali, tijekom učenja, točnost se počinje mijenjati isto kao i kod jako perturbiranog pješaka. Tijekom ove filtracije se ukupno zadržalo 143 točaka, što je skoro sedam puta manje nego bez filtriranja. Ilustracija točaka koje su se zadržale tijekom filtracije bit će prikazane na idućim grafovima (Slika 63, Slika 64, Slika 65, Slika 66 i Slika

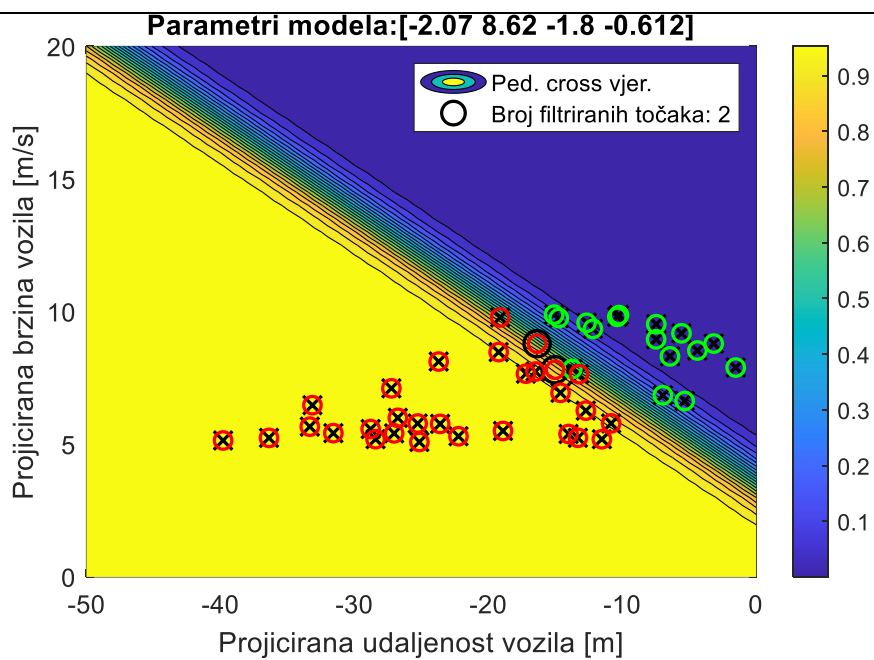
67). One će ponovo biti obilježene crnim krugom, a one koje nisu zadržane u filtraciji (odbačene), bit će prikazane crnim križićima. Crveni kružići na grafu predstavljaju slučaj kada pješak prelazi ( $y = 1$ ), a zeleni predstavljaju slučaj kada pješak ne prelazi ( $y = 0$ ).



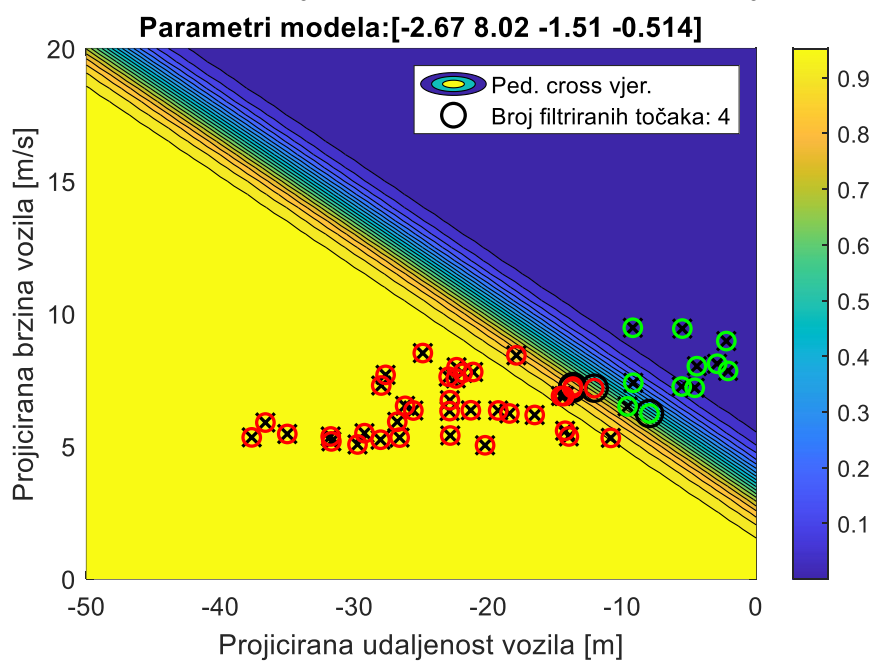
**Slika 63 Broj filtriranih točaka nakon 1. iteracije**



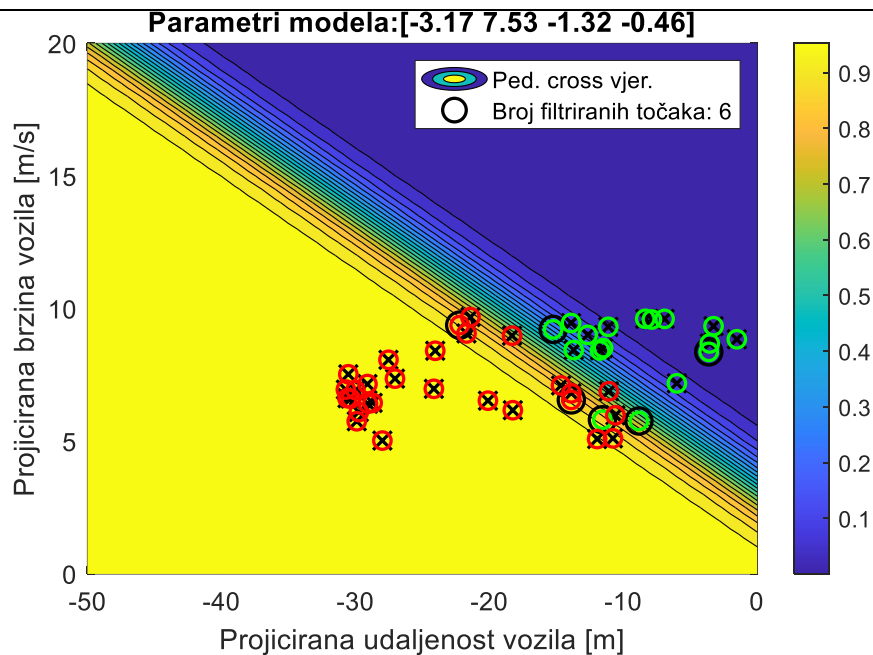
**Slika 64 Broj filtriranih točaka nakon 5. iteracije**



Slika 65 Broj filtriranih točaka nakon 10. iteracije



Slika 66 Broj filtriranih točaka nakon 15. iteracije

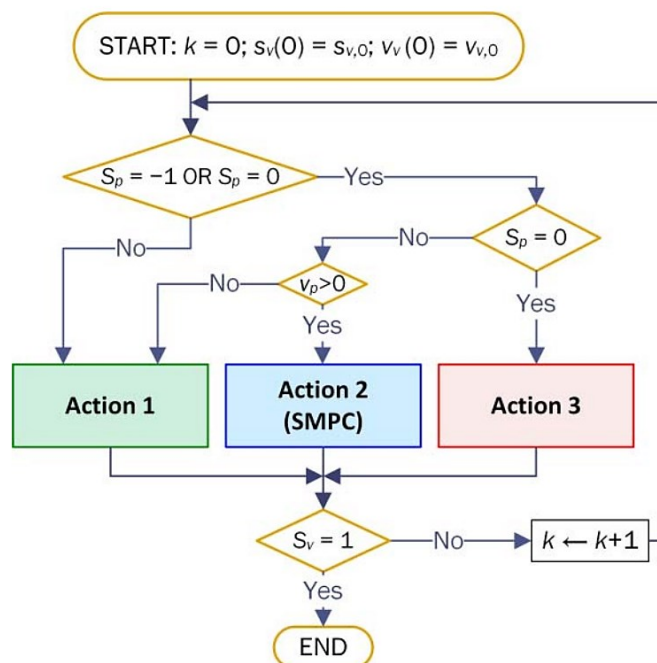


**Slika 67 Broj filtriranih točaka nakon 20. iteracije**

Proučavanjem grafova, ponovo se može vidjeti da su točke koje su uzete u obzir nakon filtriranja upravo one koje se nalaze na graničnom dijelu vjerojatnosti prelaska pješaka (u vrijednosti od 0.3 i 0.8). Iznimka je prvi graf na kojemu još nisu dovoljno naučeni parametri kako bi prikazali stvaran udio vjerojatnosti jer agresivni pješak u početku reagira na način da sve točke smješta u vrijednost prelaska. Isto tako, može se vidjeti da povećanjem broja iteracija, broj točaka koje se zadržavaju nakon filtriranja pada. Započeo je s 16, a nakon toga se kretao oko 5 po iteraciji, slično kao i kod jako perturbiranog pješaka. Iz analize se može zaključiti da se filtracijom postigao željeni cilj, a to je da se što više smanji broj podataka koje će autonomno vozilo koristiti prilikom vožnje. Ti podaci će biti samo oni koji nose najvrijednije informacije.

## 5. INTEGRACIJA PREDIKCIJSKOG MODELA S POSTOJEĆOM UPRAVLJAČKOM STRATEGIJOM

Predstavljeni predikcijski model može se integrirati u modul sustava koje je odgovoran za sigurnu kontrolu brzine, opisan u [21]. Model organizacije strategije sigurne kontrole brzine autonomnih vozila prikazan je na Slika 68.



Slika 68 Organizacija strategije sigurne kontrole brzine [21]

Strategija se temelji na modelskom prediktivnom upravljanju (eng. MPC- *Model predictive control*). Model je oblikovan na način da je rad vozila organiziran oko tri upravljačke akcije, na slici označene kao *Action 1*, *Action 2* i *Action 3*. Akcija 1 je aktivna kada pješaka nema ili je već prošao pješački prijelaz, Akcija 3 se aktivira kada pješak odluči prijeći pješački prijelaz i već je na njemu te se koristi da se izbjegne sudar. Akcija 2 je najkritičnija i najizazovnija za projektiranje. Aktivna je kada pješak prilazi pješačkom prijelazu, nalazi se na rubu, i tada nastupa nesigurnost s obzirom na njegovu odluku. Akcija 2 ima funkciju troška koja se minimizira na sljedeći način:

$$J_{Act2} = J_s + \sum_{i=0}^{N_2} \left[ (v_{v,pr} - v_v(i|k))^2 + w_a a_v^2(i|k) \right], \quad (26)$$

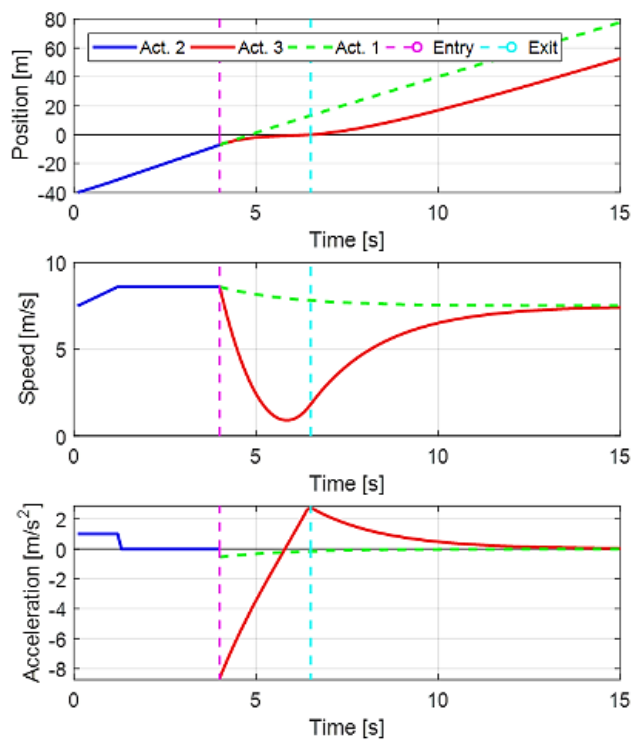
gdje  $J_{Act2}$  označava funkciju troška Akcije 2,  $v_{v,pr}$  predstavlja preferiranu brzinu vozila,  $i$  predstavlja vremenski korak horizonta predviđanja u odnosu na stvarni vremenski korak  $k$ , pa tako  $v_v(i|k)$  predstavlja brzinu vozila na predikcijskom horizontu,  $w_a$  je težinski faktor koji „kažnjava“ velike iznose akceleracije  $a_v$ .

Isto tako, sadrži stohastički član  $J_S$ :

$$J_S = \begin{cases} p_{yield}J_{Act1} + p_{cross}J_{Act3}, & \text{ako je } \hat{t}_{v,en} > \hat{t}_{p,pd} \\ J_{Act1}, & \text{ako je } \hat{t}_{v,en} \leq \hat{t}_{p,pd} \end{cases} \quad (27)$$

$J_S$  ovisi o odluci pješaka. Ako se pješak odluči podrediti  $p_{yield}$  posljedična ciljna funkcija je funkcija troška Akcije 1  $J_{Act1}$ , a ako pješak odluči prijeći  $p_{cross}$ , funkcija troška je ona Akcije 3  $J_{Act3}$ . Njihov zbroj ( $p_{yield}J_{Act1} + p_{cross}J_{Act3}$ ) predstavlja očekivanje. Vremenski korak  $\hat{t}_{v,en}$  predstavlja vrijeme ulaska vozila na pješački prijelaz, a  $\hat{t}_{p,pd}$  vrijeme odluke pješaka (kada je na rubu pješačkog prijelaza). Upravo u sklopu upravljačke strategije Akcije 2 potrebno je uključiti predstavljeni predikcijski model temeljen na logističkoj regresiji. On je u ovome koraku vrlo relevantan jer su njegovi izlazi upravo vjerojatnosti korištene u jednadžbi (27),  $\hat{p}_{yield}$  i  $\hat{p}_{cross}$ . Te vjerojatnosti ovise o brzini i udaljenosti vozila u trenutku kada je pješak došao do ruba.

Na Sliku 69 ilustrirani su mogući scenariji koji se mogu dogoditi na vremenskom horizontu i koji se u Akciji 2 uzimaju u obzir kroz (27). Naime, vozilo vozi do trenutka dolaska pješaka do ruba pješačkog prijelaza (trajektorije označene plavo), te u tom trenutku pješak može odlučiti prijeći ili stati. U slučaju odluke o prelasku Akcija 3 se mora aktivirati kako bi se kočenjem izbjegao sudar (trajektorije označene crveno). U suprotnom slučaju, ako pješak odluči stati, vozilo se može nastaviti gibati vođeno Akcijom 1 koja ga polako približava preddefiniranoj željenoj brzini (u ovom primjeru postavljena na 7.5 m/s). Kako je i navedeno, vjerojatnost ovih dviju odluka koje se koriste u (27) vozilo dobiva od predikcijskog modela (za više detalja v. [21]). Stoga je za brzinu izvođenja optimizacijskog algoritma od značajne važnosti računalna efikasnost predikcijskog modela.



Slika 69 Ilustracija potencijalnih scenarija preko vremenskog horizonta u ovisnosti o pješakovoj odluci [28]

---

## **6. ZAKLJUČAK**

U sklopu ovog diplomskog rada analiziran je računalno efikasan predikcijski model ponašanja pješaka u blizini pješačkih prijelaza bez semafora temeljen na binarnoj logističkoj regresiji. Predložena je metoda stohastičkog filtriranja podataka interakcije vozila i pješaka, s ciljem probira samo informativnijih podataka koji mogu doprinijeti poboljšanju točnosti predikcijskog modela. Predikcijski model te spomenuta metoda filtriranja ispitani su na simulacijskim podacima. Pokazano je da je u prosjeku potrebno od 250 do 600 točaka za konvergenciju modela kad se ne koristi filtracija podataka. U slučaju filtracije podataka potreban broj podataka i iteracija značajno opada i kreće se oko vrijednosti od 150 točaka. Na kraju je opisana primjena analiziranog predikcijskog modela u sklopu postojeće upravljačke strategije autonomnog vozila.

Potencijalna nadopuna ovoga diplomskog rada bila bi upravo integracija predstavljenog predikcijskog modela u postojeću upravljačku strategiju. Također, mogli bi se razmotriti i kompliciraniji slučajevi, kada se pješak ne kreće konstantnom već promjenjivom brzinom, što je slučaj sličniji stvarnom okruženju.



**LITERATURA**

- [1] <https://enciklopedija.hr/natuknica.aspx?id=71368> (pristupljeno 5.11.2023.)
- [2] <https://www.tomorrowstoday.com/production/history-of-autonomous-cars/> (pristupljeno 5.11.2023.)
- [3] <https://spectrum.ieee.org/tesla-autopilot-data-deluge> (pristupljeno 24.11.2023.)
- [4] <https://www.cinch.co.uk/guides/car-maintenance/over-the-air-car-updates> (pristupljeno 24.11.2023.)
- [5] <https://www.tesla.com/support/software-updates> (pristupljeno 24.11.2023.)
- [6] Ana Rodríguez Palmeiro, Sander van der Kint, Luuk Vissers, Haneen Farah, Joost C.F. de Winter, Marjan Hagenzieker, Interaction between pedestrians and automated vehicles: A Wizard of Oz experiment, Transportation Research Part F: Traffic Psychology and Behaviour, Volume 58, 2018, p. 1005-1020, ISSN 1369-8478, <https://doi.org/10.1016/j.trf.2018.07.020>.
- [7] <https://www.synopsys.com/automotive/what-is-autonomous-car.html> (pristupljeno 5.11.2023.)
- [8] <https://www.orangemantra.com/blog/how-connected-vehicles-are-using-cloud-technology/> (pristupljeno 24.11.2023.)
- [9] S. Hasanat-E-Rabbi, O.F. Hamim, M. Debnath, M.S. Hoque, R.C. McIlroy, K.L. Plant, N.A. Stanton, „Exploring the Relationships between Demographics, Road Safety Attitudes, and Self-Reported Pedestrian Behaviours in Bangladesh“, 2021, 13(19), 10640, <https://doi.org/10.3390/su131910640>
- [10] G. Jacobs, D. G. Wilson, “A study of pedestrian risk in crossing busy roads in four towns.”, London, U.K, 1967.
- [11] W. Andrew Harrell „Factors Influencing Pedestrian Cautiousness in Crossing Streets, The Journal of Social Psychology, 131:3, 1991., p. 367-372, DOI: 10.1080/00224545.1991.9713863, <https://www.tandfonline.com/doi/abs/10.1080/00224545.1991.9713863> [iz 7 u 8]
- [12] E. Moreno, P. Denny, E. Ward, J. Horgan, C. Eising, E. Jones, M. Glavin, A. Parsi, D. Mullins, B. Deegan, „Pedestrian Crossing Intention Forecasting at Unsignalized Intersections Using Naturalistic Trajectories.“, *Sensors* 2023; 23(5):2773. <https://doi.org/10.3390/s23052773>

- [13] J. Bock<sup>1</sup>, R. Krajewski<sup>1</sup>, T. Moers<sup>2</sup>, S. Runde<sup>1</sup>, L. Vater<sup>1</sup>, L. Eckstein, „The inD Dataset: A Drone Dataset of Naturalistic Road User Trajectories at German Intersections.“ Bock, Julian et al. “The inD Dataset: A Drone Dataset of Naturalistic Road User Trajectories at German Intersections.” 2020 IEEE Intelligent Vehicles Symposium (IV) (2019): p. 1929-1934.
- [14] <https://builtin.com/data-science/random-forest-algorithm> (pristupljeno 21.11.2023.)
- [15] <https://www.kdnuggets.com/2020/01/decision-tree-algorithm-explained.html> (pristupljeno 21.11.2023.)
- [16] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al. „Scikit-learn: Machine learning in Python.“ Journal of Machine Learning Research. 12., 2012.
- [17] B. Völz, H. Mielenz, I. Gilitschenski, R. Siegwart and J. Nieto, "Inferring Pedestrian Motions at Urban Crosswalks," in IEEE Transactions on Intelligent Transportation Systems, vol. 20, no. 2, 2019., p. 544-555, , doi: 10.1109/TITS.2018.2827956.A.
- [18] C. J. Burges, “A tutorial on support vector machines for pattern recognition,”, appeared in Data Mining and Knowledge Discovery, vol. 2, 1998., p. 121–167,
- [19] <https://medium.com/@evertongomede/recursive-feature-elimination-a-powerful-technique-for-feature-selection-in-machine-learning-89b3c2f3c26a> (pristupljeno 21.11.2023.)
- [20] B. Škugor, J. Topić, J. Deur, V. Ivanović, H. E. Tseng, “Analysis of a Game Theory-Based Model of Vehicle-Pedestrian Interaction at Uncontrolled Crosswalks,” in Proc. of 2020 SST Conference, Osijek, Croatia, 2020, pp. 73–81
- [21] B. Škugor, J. Deur, V. Ivanovic and H. E. Tseng, "Stochastic Model Predictive Control of an Autonomous Vehicle Interacting with Pedestrians at Unsignalized Crosswalks," 2023 European Control Conference (ECC), Bucharest, Romania, 2023, pp 2, doi: 10.23919/ECC57647.2023.10178144.
- [22] P. Chen, C. Wu, and S. Zhu, “Interaction between vehicles and pedestrians at uncontrolled mid-block crosswalks”, Safety Science, Vol. 82, pp. 68-76, 2016
- [23] <https://www.collinsdictionary.com/dictionary/english/logistic-regression> (pristupljeno 6.11.2023.)
- [24] <https://www.analyticsvidhya.com/blog/2021/08/conceptual-understanding-of-logistic-regression-for-data-science-beginners/> (pristupljeno 6.11.2023.)
- [25] <https://medium.com/analytics-vidhya/logistic-regression-and-its-mathematical-implementation-722434ed01a5> (pristupljeno 6.11.2023.)

- [26] <https://amalj7.medium.com/logistic-regression-eb2903251107>(pristupljeno 6.11.2023.)
- [27] <https://towardsdatascience.com/introduction-to-logistic-regression-66248243c148>  
(pristupljeno 6.11.2023.)
- [28] B. Škugor, J. Deur, V. Ivanovic and H. E. Tseng, "Safety Manifold for Stochastic Model Predictive Control of an Autonomous Vehicle Approaching Unsignalized Crosswalks with Pedestrians," 2024 European Control Conference (ECC), Stockholm, Sweden, 2024, submitted for review