

dr Đorđe Petrović, Akademija strukovnih studija Zapadna Srbija, Odsek Valjevo.

Branko Čebić, MA, Akademija strukovnih studija Zapadna Srbija, Odsek Valjevo.

Dejan Beljić, MA, Akademija strukovnih studija Zapadna Srbija, Odsek Valjevo.

DOI 10.5937/turpos0-31845

UDK 005.521:338.485.2]:519

PREDVIĐANJE BROJA TURISTA PRIMENOM MAŠINSKOG UČENJA

PREDICTING THE NUMBER OF TOURISTS USING MACHINE LEARNING

Apstrakt: U ovom radu je dat pregled aktuelnih modela mašinskog učenja iz vremenskih serija i njihova primena za potrebe predviđanja broja turističkih poseta u narednom periodu. Pojava virusa Covid-19 je generalno imala veliki uticaj na turizam i unela veliku neizvesnost u ovu oblast. Primena mašinskog učenja i pokušaj predviđanja broja turističkih poseta u narednom periodu, može da bude korisna onima koji se bave ponudom u ovoj oblasti.

Ključne reči: mašinsko učenje, vremenske serije, turizam, posete.

Abstract: This paper provides an overview of current models of machine learning from time series and their application for the purpose of predicting the number of tourist visits in the coming period. The emergence of the Covid-19 virus has generally had a major impact on Tourism and has introduced great uncertainty in this area. The application of machine learning and an attempt to predict the number of tourist visits in the coming period, can be useful to those who deal with the offer in this area.

Keywords: machine learning, time series, tourism, visits.

Uvod

Predviđanje vremenskih serija je teško. Za razliku od jednostavnijih problema klasifikacije i regresije, problemi vremenskih serija imaju dodatnu složenost zbog redosleda ili vremenske zavisnosti između posmatranih podataka. Prilikom izrade modela, ova vremenska struktura može da se iskoristi tako što pruža dodatne strukture poput trendova,

sezonskih podataka, a koji mogu da se iskoriste za poboljšanje sposobnosti modela. Tradicionalno, prilikom predviđanja vremenskih serija dominiraju linearne metode poput ARIMA, koje su dobro razumljive i efikasne u rešavanju mnogih problema. U odnosu na klasične metode, metode mašinskog učenja mogu da budu efikasnije u složenijim problemima predviđanja vremenskih serija (Brownlee, 2018).

Mašinsko učenje koristi algoritme koji iterativno uče iz podataka i omogućavaju da se pronađu skrivene informacije ili odnosi između podataka. Veštačke neuronske mreže su najpopularnija i jedna od najčešće primenjena metoda mašinskog učenja, koja je sposobna da automatski nauči složene odnose između ulaznih i izlaznih podataka. One su moćan alat koji može da se primeni za predviđanje vremenskih serija, posebno za probleme sa složenim zavisnostima između podataka, kao i za predviđanje više koraka unapred.

Povezani radovi

Postoji više radova koji se bave predviđanjem broja turista. U istraživanju (Ghalekhondabi, et al., 2019), autori su se bavili predviđanjem potražnje za turizmom i putničkim prevozom, kroz pregled „klasičnih“ metoda za predviđanje vremenskih serija, a koje ne spadaju u mašinsko učenje i u veštačke neuronske mreže. U svom radu, autori (Sun, et al., 2019) su predložili okvir predviđanja koji koristi mašinsko učenje i pretraživanje putem Interneta, za predviđanje dolazaka turista u popularne destinacije u Kini, konkretno, u Peking. Oni su dokazali da postoji uzročnost između indeksa pretraživanja Intereneta i broja turista u Peking. Njihova studija se odnosi na period pre pojave virusa Covid-19. Postoji nekoliko radova koji su se bavili predviđanjem oporavka turizma usred epidemije izazvane virusom Covid-19, kao što su (Zaborovskaia, et al., 2020), (Zhang, et al.,

2021) i (Fotiadis, et al., 2021). Navedeni radovi su, svaki sa svog aspekta, se bavili predviđanjem u pojedinim oblastima, za koje su pravili predviđanje.

U ovom istraživanju, mi smo se bavili analizom podataka koji se odnose na Republiku Srbiju, kao i primenom mašinskog učenja za potrebe predviđanja, uzimajući u obzir podatke koji su nastali nakon pojave virusa Covid-19, kao i pre pojave ovog virusa.

Skup podataka o turističkim posetama

U ovom istraživanju smo koristili javno dostupne podatke Republičkog zavoda za statistiku o dolascima i noćenjima turista po regionima – mesečni podaci (Republički zavod za statistiku,

2021). U trenutku preuzimanja podataka, bili su dostupni podaci o turističkim posetama domaćih i stranih turista po regionima, kao i na nivou cele teritorije Republike Srbije, za period od januara 2010. godine do februara 2021. godine. Ovaj skup podataka sadrži karakteristike, kao što su:

- Poreklo turista (domaći turisti, strani turisti)
- Vrsta korišćene turističke usluge (dolasci turista, noćenje turista)
- Teritorija (region koji je posećen)

Podaci su na mesečnom nivou, pa ćemo se u ovom istraživanju baviti predviđanjima na mesečnom nivou. U sledećoj tabeli su prikazani podaci o broju domaćih turista na teritoriji Republike Srbije u navedenom periodu (Republički zavod za statistiku, 2021):

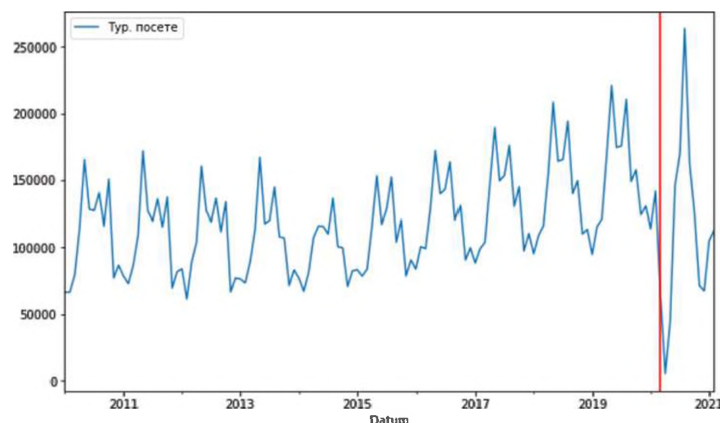
Tabela 1 Broj domaćih turista na teritoriji Republike Srbije, mesečni podaci

| Vreme | Broj turističkih poseta |
|---------------|-------------------------|
| Januar 2010. | 66370 |
| Februar 2010. | 66468 |
| ... | ... |
| Januar 2021. | 104909 |
| Februar 2021. | 112698 |

Izvor: Republički zavod za statistiku

Na sledećoj slici su grafički prikazani ovi podaci:

Slika 1 Broj domaćih turista na mesečnom nivou na teritoriji Republike Srbije



Izvor: Autori



Na grafičkom prikazu broja poseta domaćih turista, a za posmatrani period, postoji podela na period pre pojave virusa Covid-19 i pre primene pratećih mera za sprečavanje širenja ovog virusa i period nakon pojave virusa. Period pre pojave virusa karakterišu svojstva koja se odnose na sve vremenske serije, kao što su trend, sezonske promene i slično. U periodu nakon pojave virusa, jasno se uočavaju dve lokalne ekstremne vrednosti. Prva ekstremna vrednost se odnosi na april 2020. godine kada je počela epidemija virusa Covid-19 i kada je zbog te epidemije i zbog primenjenih mera za sprečavanje širenja virusa, naglo smanjen broj turističkih poseta. Druga ekstremna vrednost se odnosi na avgust 2020. godine kada domaći turisti nisu mogli da putuju van zemlje, kada su mere za sprečavanje širenja virusa ublažene i kada je povećan broj njihovih turističkih poseta u Republici Srbiji. Ovi podaci su u daljem istraživanju korišćeni za obučavanje veštačkih neuronskih mreža, sa ciljem predviđanja broja turističkih poseta u narednom periodu.

Priprema podataka

Posmatrani skup podataka najpre treba podeliti na podatke za obučavanje modela, podatke za validaciju i podatke za testiranje, npr. u odnosu 70%:20%:10%. Ovo se radi iz dva razloga:

1. Osigurava se da je i dalje moguće deliti podatke u uzorke uzastopnih prozora sa podacima;
2. Osigurava se da su podaci za validaciju i testiranje realni i da se evaluacija primenjuje na podacima koji su prikupljeni nakon obučavanja modela.

Pre obučavanja veštačkih neuronskih mreža je potrebno i da se skaliraju vrednosti, a postupak normalizacije je uobičajeni način za ovo skaliranje. Ovde se primenjuje postupak po kome se od svake posmatrane vrednosti oduzima izračunata prosečna vrednost, a zatim se to podeli sa standardnom devijacijom. To se primenjuje za svaku posmatranu karakteristiku. Prosečna vrednost i standardna devijacija se računaju samo koristeći skup podataka za obuku modela, tako da model koji se obučava nema pristup vrednostima iz skupa za validaciju i skupa za testiranje.

Modeli mašinskog učenja iz vremenskih serija prave skup predviđanja na osnovu „prozora“ uzastopnih uzoraka iz podataka, tako što se na osnovu liste ulaznih podataka ti podaci pretvaraju u ulazni prozor sa podacima i u prozor sa oznakama.

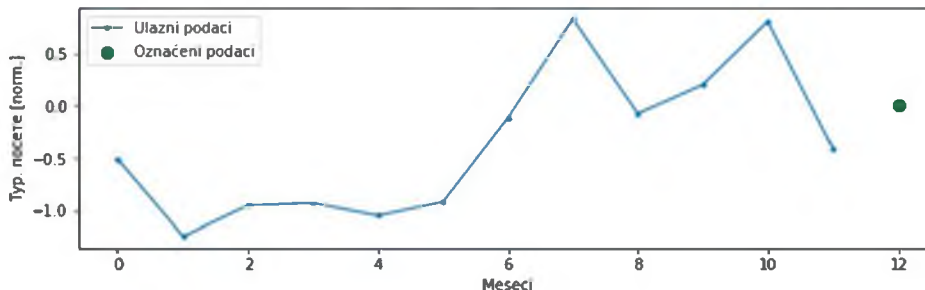
Glavne karakteristike ulaznih prozora sa podacima su:

- Širina (broj vremenskih koraka) prozora sa ulaznim podacima i oznakama
- Vremenski razmak između njih
- Koje karakteristike se koriste kao ulazi, kao oznake ili kao jedno i drugo

U zavisnosti od istraživanja i vrste modela, moguće je generisati različite prozore sa podacima. Evo nekoliko primera:

- Moguće je napraviti prozor sa podacima za model koji će da pravi predviđanje 1 mesec unapred na osnovu podataka iz prethodnih 12 meseci;
- Moguće je napraviti prozor sa podacima za model koji će da pravi predviđanje 12 meseci unapred, na osnovu podataka iz prethodnih 12 meseci.

Slika 2 Primer ulaznog prozora sa podacima i prozor sa oznakama



Izvor: Autori

Na prethodnoj slici je prikazan grafički prikaz primera ulaznog prozora sa podacima i prozora sa oznakama za model koji će da pravi predviđanje 1 mesec unapred (zelena tačkica) na osnovu podataka iz prethodnih 12 meseci (plave tačkice), za podatke koji su prethodno normalizovani.

Za potrebe analize i istraživanja koji su prikazani u ovom radu, korišćena je platforma otvorenog koda, TensorFlow (TensorFlow, 2015), kao biblioteka koja omogućava obavljanje višestrukih zadataka mašinskog učenja, kao i biblioteka za rad sa veštačkim neuronskim mrežama, Keras (Chollet F. et al., 2015), koja radi u TensorFlow okruženju. Prilagođavanje modela i njegovo izvršavanje je rađeno u programskom jeziku Python (Chollet, 2018).

Aktuelni modeli mašinskog učenja iz vremenskih serija

Veštačke neuronske mreže mogu da budu napravljene primenom različitih modela, koji uključuju konvolucione i rekurentne neuronske mreže. Prema uputstvu (TensorFlow, 2021) koje će biti primenjeno u nastavku, kada se vrši predviđanje vremenskih serija, to predviđanje može da se radi:

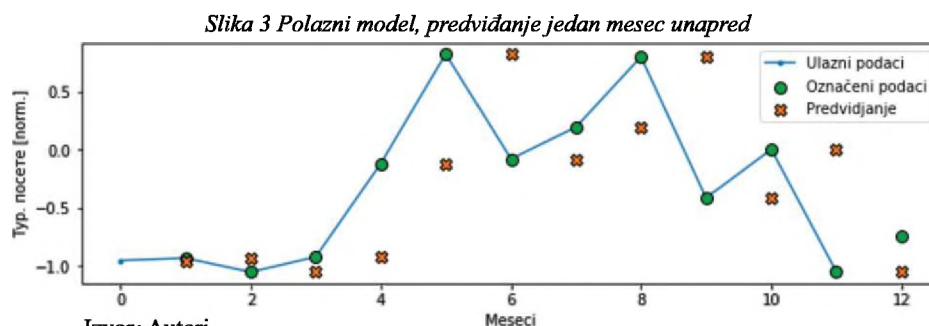
- Jedan vremenski korak unapred
 - Predviđanje jednog svojstva
 - Predviđanje više svojstava
- Više vremenskih koraka unapred
 - Pravljenje svih predviđanja odjednom
 - Pravi se jedno predviđanje i izlaz se vraća natrag u model (Autoregresija)

Modeli sa predviđanjem jednog vremenskog koraka unapred

Modeli sa predviđanjem jednog vremenskog koraka su najjednostavniji modeli koji mogu da se naprave. To je model koji predviđa vrednost posmatranog svojstva za jedan vremenski korak unapred, a na osnovu trenutnih uslova, na primer, predviđa broj turističkih poseta za jedan mesec unapred. U odnosu na broj svojstava koja se predviđaju, modeli mogu da budu sa predviđanjem jednog svojstva i modeli sa predviđanjem više svojstava. U nastavku će biti opisano nekoliko tipova ovih modela.

Polazni model (en. Baseline)

Pre izgradnje i obučavanja ostalih modela, korisno je imati „Polazni“ model, kao tačku za poređenje sa kasnijim složenijim modelima. Ako je zadatak da se vrši predviđanje turističkih poseta jedan mesec unapred, polazni model samo vraća trenutni broj turističkih poseta kao svoje predviđanje za naredni mesec, bez promena te vrednosti. Naravno, ovaj model će raditi lošije ako se pravi predviđanje dalje u budućnosti. Na grafičkom prikazu predviđanja Polaznog modela se može videti da su to jednostavno oznake koje su pomerene za jedan vremenski korak unapred:



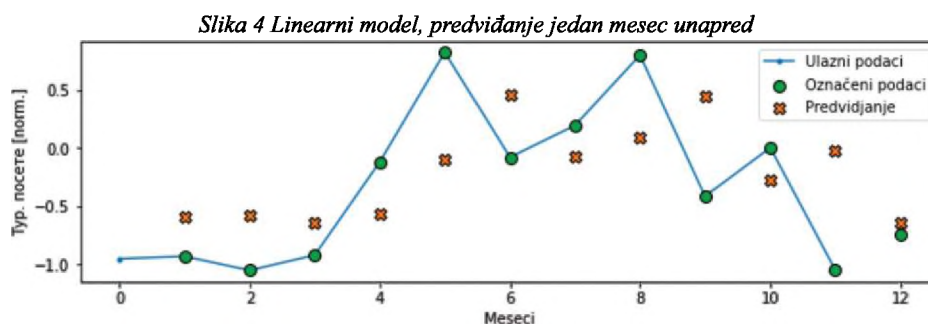
Prethodni model, koji vrši predviđanje jednog vremenskog koraka unapred, je više puta pokrenut tokom posmatranih 12 meseci, a elementi dobijenog grafikona su:

- Plava linija prikazuje ulazne podatke, odnosno broj turističkih poseta po mesecima, za posmatrani period;
- Zelene oznake prikazuju ciljane vrednost predviđanja. Te tačke su prikazane u vremenu predviđanja, odnosno, jedan mesec unapred, u odnosu na ulazne podatke. Zbog toga se ospeg oznaka pomera za jedan vremenski korak u odnosu na ulazne podatke. Ciljane vrednosti su poznate i sa njima će se vršiti poređenje podataka dobijenih predviđanjem;

- Narandžasti krstići su vrednosti dobijene predviđanjem modela za svaki vremenski korak. Ako bi model savršeno predviđao, predviđanja bi se poklopila sa zelenim oznakama.

Linearni model

Najjednostavniji obučavani model, koji može da se primeni na ovaj zadatak je umetanje linearne transformacije između ulaza i izlaza. U ovom slučaju izlazni podaci iz nekog vremenskog koraka zavise samo od tog koraka. Poput Polaznog modela, Linearni model poziva seriju širokih prozora sa podacima. Na taj način model koristi skup nezavisnih predviđanja u uzastopnim vremenskim koracima, pri čemu ne postoji interakcija između predviđanja u svakom vremenskom koraku.



Izvor: Autori

Na prethodnom grafikonu se vidi predviđanje sa jednim vremenskim korakom unapred, dobijeno primenom Linearnog modela nad posmatranim segmentom podataka. Model je više puta pokrenut tokom posmatranih 12 meseci. Prilikom svakog pokretanja modela, kao ulaz se koriste podaci iz posmatranog vremenskog koraka i vrši se predviđanje u sledećem vremenskom koraku. Prozor sa podacima se zatim pomera za jedan vremenski korak i ponavlja se postupak predviđanja. U poređenju sa Polaznim modelom, predviđanje je uglavnom bolje, ali mogu da se dese i slučajevi kada je lošije.

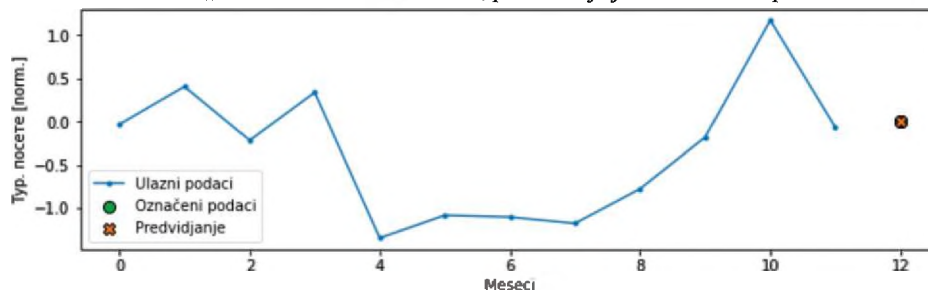
„Gusti“ model u više koraka (en. Multi-step dense)

Prethodna dva modela, obrađuju svaki vremenski korak nezavisno. To modeli

posmatraju samo trenutnu vrednost svog ulaza i ne uzimaju u obzir kako se ulazne vrednosti menjaju tokom vremena. Ovaj problem se prevazilazi pravljenjem i primenom modela koji prilikom predviđanja koristi pristup podacima iz više vremenskih koraka da bi proizveo jedan izlaz. Za tu namenu se koristi generator prozora sa podacima, koji će proizvesti serije ulaznih podataka o turističkim posetama od npr. 12 uzastopnih meseci, sa jednom oznakom.

„Gusti“ model ima arhitekturu koja je slična Linearnom modelu, osim što se ovaj model sastoji od nekoliko kompletno povezanih slojeva („layers.Dense“) između ulaza i izlaza.

Slika 5 „Gusti“ model u više koraka, predviđanje jedan mesec unapred



Izvor: Autori

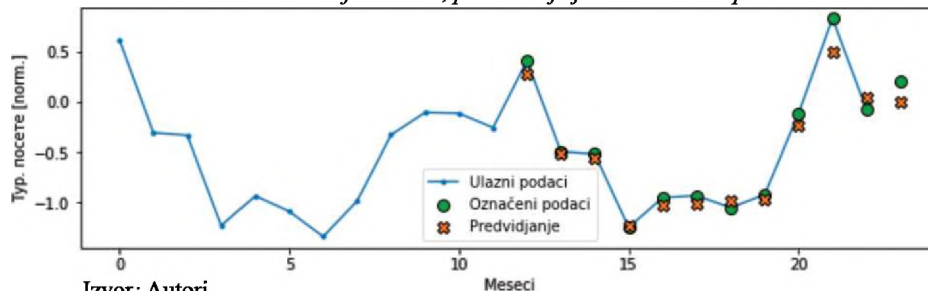
Na prethodnom grafikonu je prikazano predviđanje sa jednim vremenskim korakom unapred, dobijeno primenom „Gusti“ model u više koraka, a nad posmatranim segmentom podataka Model je pokrenut jednom i dobijeno je predviđanje turističkih poseta u narednom mesecu, na osnovu podataka iz prethodnih 12 meseci. U prethodnom primeru se predviđanje u značajnoj meri poklopilo sa ciljanom vrednošću predviđanja.

ulaznim prozorima sa podacima koji su tačno tog oblika. Konvolucijske neuronske mreže (en. Convolutional Neural Networks, CNN) rešavaju ovaj problem. Sloj konvolucije („layers.Conv1D“) takođe uzima više vremenskih koraka kao ulaz, za svako predviđanje, a razlika u odnosu na „Gusti“ model u više koraka je u tome što se Konvolucijski model može pokretati sa ulazima bilo koje dužine.

Konvolucijske neuronske mreže

Glavna negativna strana prethodnog pristupa je da se rezultujući model može izvršiti samo na

Slika 6 Konvolucijski model, predviđanje jedan mesec unapred



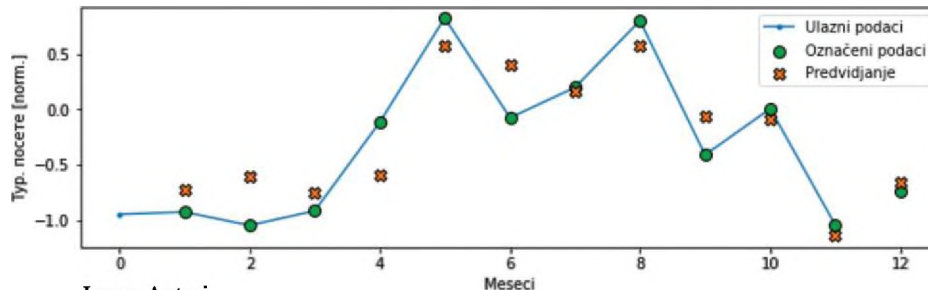
Izvor: Autori

Na prethodnom grafikonu je prikazano predviđanje sa jednim vremenskim korakom unapred, dobijeno primenom Konvolucijskog modela nad posmatranim segmentom podataka. Model je pokrenut sa širokim prozorom sa podacima. Pre prvog predviđanja, kao ulaz se koristi serija ulaznih podataka o turističkim posetama tokom 12 meseci i svako sledeće predviđanje se zasniva na prethodnih 12 vremenskih koraka.

Rekurentne neuronske mreže

Rekurentna neuronska mreža (Recurrent Neural Network, RNN) je vrsta veštačke neuronske mreže koja je pogodna za podatke vremenskih serija. Ove neuronske mreže obrađuju vremensku seriju korak po korak, pri tom uspostavljajući unutrašnje stanje od vremenskog koraka do vremenskog koraka. Modeli ovih mreža mogu odjednom da se obučavaju na seriji ulaznih podataka.

Slika 7 Rekurentna neuronska mreža, predviđanje jedan mesec unapred



Izvor: Autori

Na prethodnom grafikonu se vidi predviđanje sa jednim vremenskim korakom unapred, dobijeno primenom Rekurentnih neuronskih mreža nad posmatranim skupom podataka. Model je obučavan na seriji ulaznih podataka, tako da daje predviđanje turističkih poseta na osnovu podataka iz prethodnih 12

Modeli za predviđanje više vremenskih koraka unapred (Multi-step models)

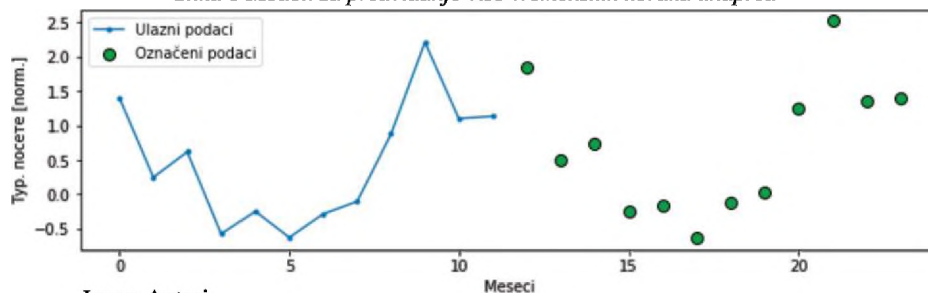
Modeli koji su prethodno opisani su pravili predviđanja za jedan vremenski korak unapred, kao u pokazanim primerima, konkretno za jedan mesec unapred. U nastavku će biti pokazano kako mogu da se prošire mogućnosti modela mašinskog učenja iz vremenskih serija, tako da mogu da se primene za potrebe predviđanja više vremenskih koraka unapred. Prilikom predviđanja vrednosti više

vremenskih koraka unapred, model treba da nauči da predviđa opseg budućih vrednosti. Prema tome, za razliku od modela koji predviđaju jedan vremenski korak unapred gde se predviđa samo jedna buduća vrednost, ovi modeli predviđaju redosled budućih vrednosti. Postoje dva pristupa za predviđanje više vremenskih koraka unapred:

- Modeli „jednog stanja“ (en. Single-shot models), u kojima se na osnovu jednog stanja vrši predviđanje čitave sekvence odjednom;
- Autoregresivno predviđanje koje pravi jedno predviđanje, a zatim se izlaz vraća natrag u model i koristi za sledeće predviđanje.

U nastavku, u primerima za predviđanje više vremenskih koraka unapred, modeli će biti obučavani da predviđaju 12 meseci unapred na osnovu podataka iz prošlih 12 meseci, kao što je prikazano na sledećoj slici:

Slika 8 Modeli za predviđanje više vremenskih koraka unapred

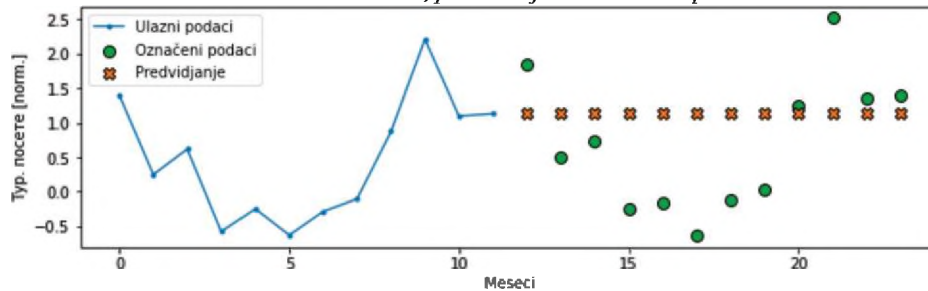


Izvor: Autori

Polazni model (Baseline)

Jednostavan polazni model ponavlja poslednji ulazni korak vremena za potreban broj izlaznih koraka vremena, kao što je prikazano na sledećoj slici:

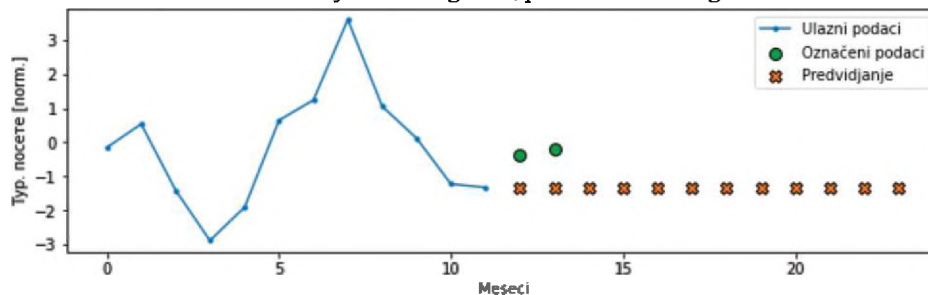
Slika 9 Polazni model, predviđanje 12 meseci unapred



Izvor: Autori

Budući da je bio zadatak da se predviđaju vrednosti za 12 meseci unapred, a na osnovu vrednosti iz prošlih 12 meseci, jednostavan Polazni model je za neki segment podataka ponovio vrednosti iz prethodnog meseca, uz pretpostavku da će u sledećim mesecima biti slična vrednost. Na sledećem grafikonu je prikazana primena ovog modela za predviđanje broja turističkih poseta u 2021. godini, na osnovu podatka iz 2020. godine:

Slika 10 Predviđanje za 2021. godinu, primenom Polaznog modela



Izvor: Autori

Elementi ovog grafikona su:

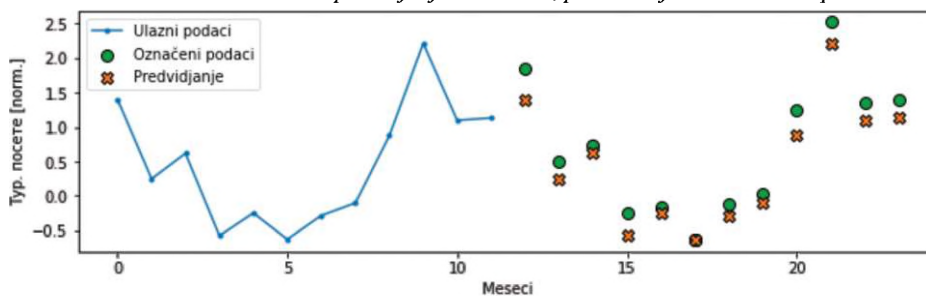
- Plava linija prikazuje ulazne podatke, odnosno podatke iz 2020. godine;
- Zelene oznake prikazuju ciljanu vrednost predviđanja. Kako su u trenutku pisanja ovog rada već poznati podatci o broju poseta za mesece januar i februar 2021. godine, i ti podatci su prikazani na grafikonu;
- Narandžasti krstići su vrednosti dobijene predviđanjem modela za 2021. godinu. Polazni model je jednostavno ponovio podatak iz decembra 2020. godine i te podatke prikazao kao podatke dobijene „predviđanjem“ modela.



Polazni model sa ponavljanjem sekvenci (Repeat Baseline)

Ovaj model ponavlja poslednju sekvencu vremena za potreban broj izlaznih koraka vremena.

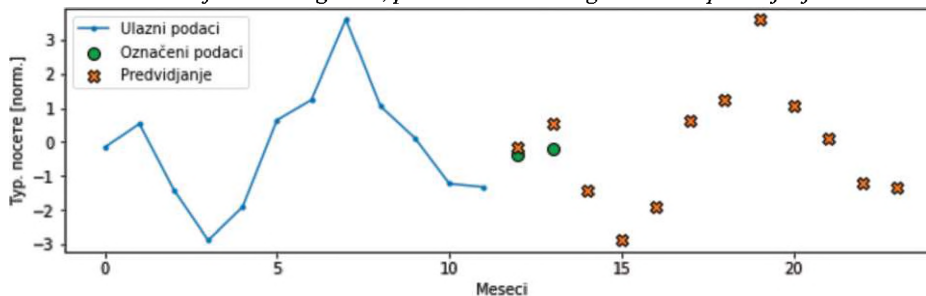
Slika 11 Polazni model sa ponavljanjem sekvenci, predviđanje 12 meseci unapred



Izvor: Autori

Kao što se može videti na prethodnoj slici, ovaj model je za neki segment podataka kao predviđanje vrednosti za narednih 12 meseci, ponovio seriju podataka iz prošlih 12 meseci. Na sledećem grafikonu je prikazana primena ovog modela za predviđanje broja turističkih poseta u 2021. godini, na osnovu podatka iz 2020. godine, gde je model za 2021. godinu jednostavno ponovio podatke iz 2020. godine:

Slika 12 Predviđanje za 2021. godinu, primenom Polaznog modela sa ponavljanjem sekvenci

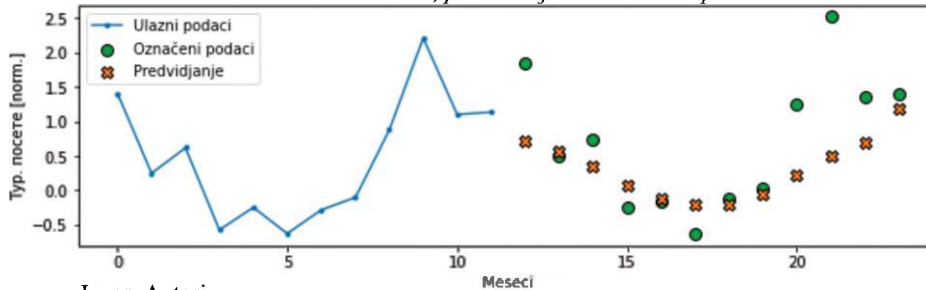


Izvor: Autori

Linearni model

Jednostavan linearni model spada u modele „jednog stanja“. Zasnovan je na poslednjem ulaznom vremenskom koraku iz kojih treba da predvidi buduće vremenske korake sa linearnom projekcijom.

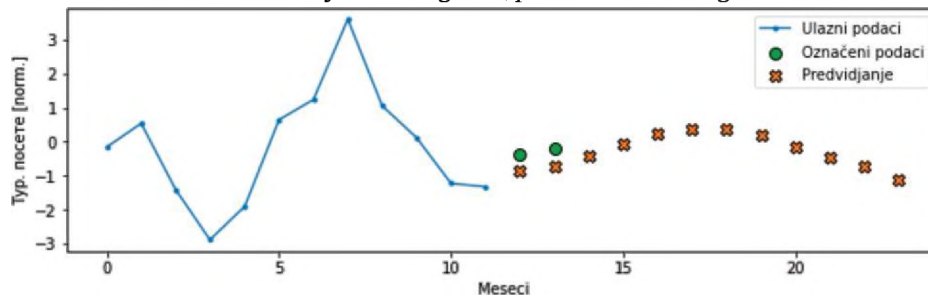
Slika 13 Linearni model, predviđanje 12 meseci unapred



Izvor: Autori

Kao što se može videti sa prethodne slike, ovaj model se čini boljim od Polaznog modela, ali to i dalje nije dovoljno moćan model. Na sledećem grafikonu je prikazana primena ovog modela za predviđanje broja turističkih poseta u 2021. godini, na osnovu podatka iz 2020. godine:

Slika 14 Predviđanje za 2021. godinu, primenom Linearnog modela

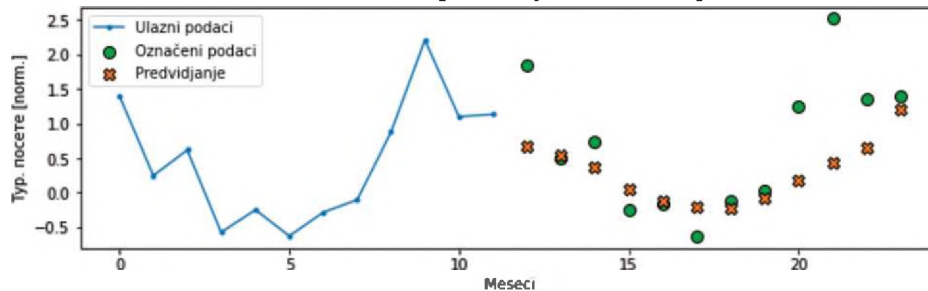


Izvor: Autori

„Gusti“ model (Dense)

Kao što je to već rečeno, ovaj model nastaje dodavanjem kompletno povezanog sloja (layers.Dense) između ulaza i izlaza.

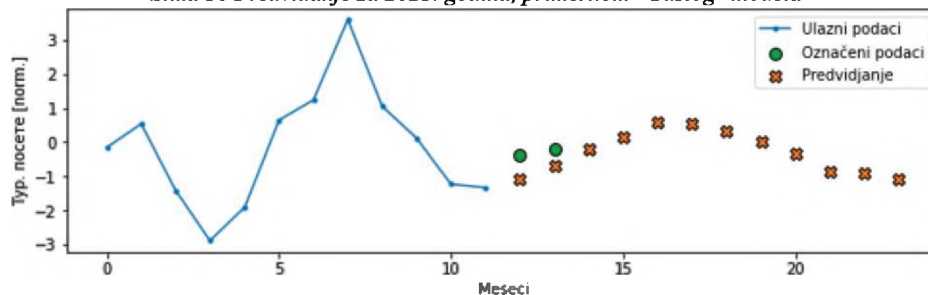
Slika 15 "Gusti" model, predviđanje 12 meseci unapred



Izvor: Autori

U poređenju sa prethodnim, ovaj model ima veću snagu, ali se model i dalje zasniva samo na jednom ulaznom vremenskom koraku. Na sledećem grafikonu je prikazana primena ovog modela za predviđanje broja turističkih poseta u 2021. godini, na osnovu podatka iz 2020. godine:

Slika 16 Predviđanje za 2021. godinu, primenom "Gustog" modela

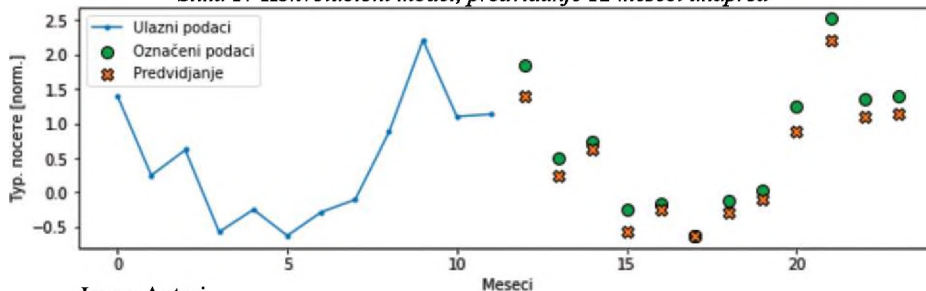


Izvor: Autori

Konvolucijske neuronske mreže

Ovaj model daje predviđanja na osnovu istorije fiksne širine i na taj način uzima u obzir kako se podaci vremenom menjaju.

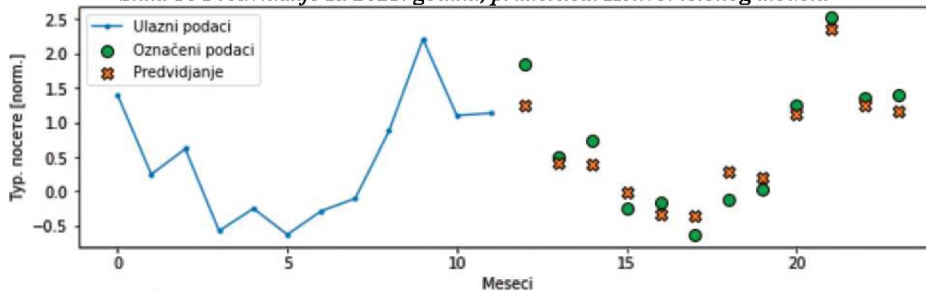
Slika 17 Konvolucioni model, predviđanje 12 meseci unapred



Izvor: Autori

Prilikom predviđanja više vremenskih koraka unapred, na osnovu svojih karakteristika Konvolucioni model može da dovede do boljih performansi od „Gustog“ modela. Na sledećem grafikonu je prikazana primena ovog modela za predviđanje broja turističkih poseta u 2021. godini, na osnovu podatka iz 2020. godine:

Slika 18 Predviđanje za 2021. godinu, primenom Konvolucionog modela

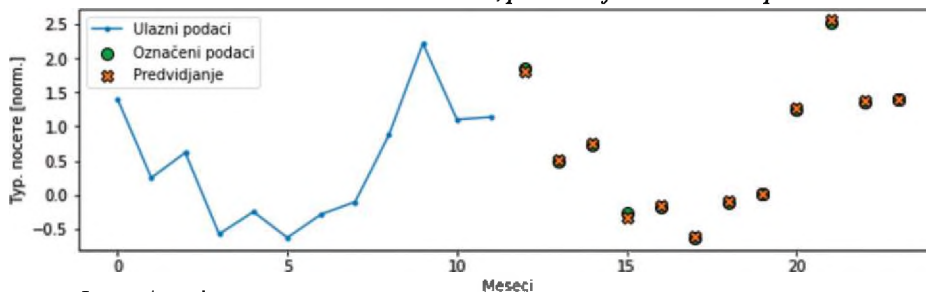


Izvor: Autori

Rekurentne neuronske mreže

Model Rekurentnih neuronskih mreža može da uči tako što koristi dugu istoriju ulaznih podataka, ako je to relevantno za predviđanja koja model pravi. Ovde model akumulira unutrašnje stanje za 12

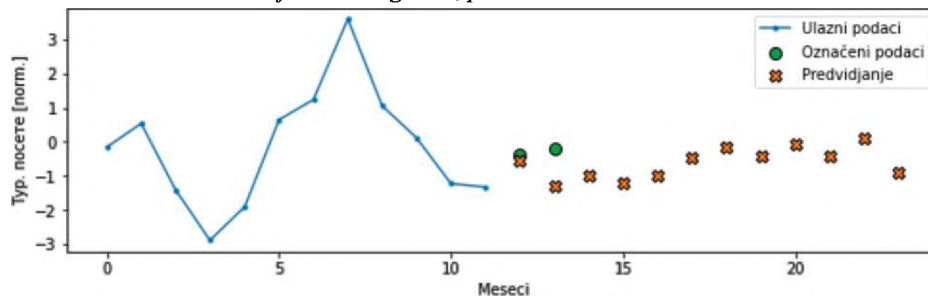
Slika 19 Rekurentne neuronske mreže, predviđanje 12 meseci unapred



Izvor: Autori

Na sledećem grafikonu je prikazana primena ovog modela za predviđanje broja turističkih poseta u 2021. godini, na osnovu podatka iz 2020. godine:

Slika 20 Predviđanje za 2021. godinu, primenom Rekurentne neuronske mreže



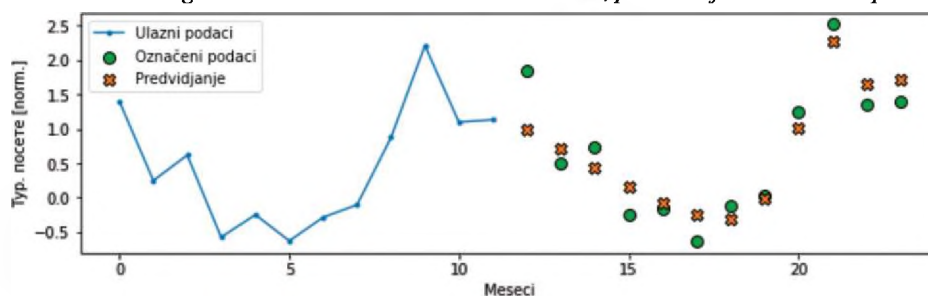
Izvor: Autori

Autoregresivni model Rekurentne neuronske mreže

Svi prethodno pomenuti modeli za predviđanje više vremenskih koraka unapred, predviđaju celu izlaznu sekvencu u jednom koraku. U nekim slučajevima može da bude korisno za model da se ovo predviđanje razloži na pojedinačne vremenske korake. Tada svaki izlaz modela može da bude vraćen nazad u model u svakom koraku i predviđanja mogu da se uslove prethodnim korakom (Graves, 2013). Jedna jasna prednost ovog stila modela je ta što model može da se podesi tako da daje izlaz različite dužine.

Za model autoregresivne Rekurentne neuronske mreže, prva metoda koja je potrebna je metoda „zagrevanja“ (en. warmup), kako bi se na osnovu ulaza inicijalizovalo unutrašnje stanje modela. Jednom obučeno stanje će obuhvatiti relevantne delove istorije ulaza. U ovoj metodi se predviđanje jednog vremenskog koraka vraća u unutrašnje stanje. Sa inicijalizovanim unutrašnjim stanjem i sa početnim predviđanjem, sada može da se nastavi sa iteracijom modela, koji daje predviđanja na osnovu svakog prethodnog vremenskog koraka, kao ulazom.

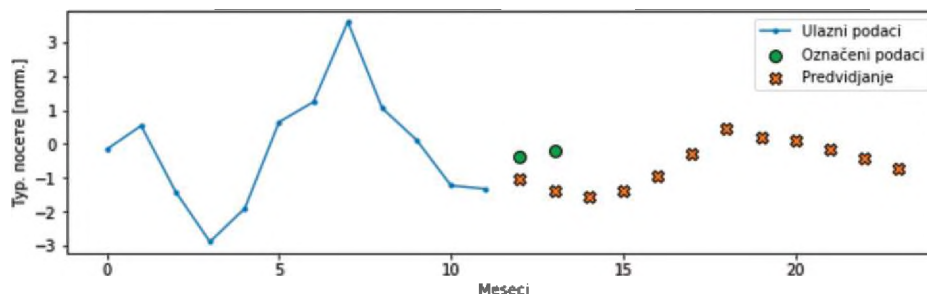
Slika 21 Autoregresivni model Rekurentne neuronske mreže, predviđanje 12 meseci unapred



Izvor: Autori

Na prethodnom grafikonu je za neki segment podataka prikazano predviđanje 12 meseci unapred, primenom autoregresivnog modela Rekurentne neuronske mreže. Na sledećem grafikonu je prikazana primena ovog modela za predviđanje broja turističkih poseta u 2021. godini, na osnovu podatka iz 2020. godine:

Slika 22 Predviđanje za 2021. godinu, primenom Autoregresivnog modela Rekurentne neuronske mreže



Izvor: Autori

Posmatrani skup podataka je relativno mali - sadrži „samo“ 134 zapisa na mesečnom nivou o broju domaćih turista na području Republike u periodu od januara 2010. godine do februara 2021. godine. Za potrebe mašinskog učenja iz podataka je važno da se raspolaže sa što većim brojem podataka. Da bi se postigli bolji rezultati predviđanja, poželjno je da posmatrani skup podataka pokriva neki veći vremenski period, kako i da u tom skupu podataka, budu podaci o broju turističkih poseta npr. na nedeljnom ili na dnevnom nivou. Prikupljanje takvih podataka u narednom periodu može da doprinese boljim rezultatima za potrebe predviđanja broja turista u narednom periodu.

U ovom radu su prikazani aktuelni modeli mašinskog učenja iz vremenskih serija, sa ciljem predviđanja broja turističkih poseta u narednom periodu. Pokazana su dva tipa modela. Prva grupa modela vrši predviđanje jedan vremenski korak unapred, odnosno jedan mesec unapred, a druga grupa modela vrši predviđanje više vremenskih koraka unapred, odnosno, nekoliko meseci unapred. Modeli koji su pokazani se kreću od onih koji su najjednostavniji, do složenijih modela „dubokog mašinskog učenja“. Takođe, u zavisnosti od modela koji je korišćen, i izlazni podaci koji se dobijaju se razlikuju. Kao što je napomenuto u (Ghalekhondabi, et al., 2019), istraživanju koje se bavilo predviđanjem potražnje za turizmom i putničkim prevozom, autori su naveli nekoliko primera u kojima je neka metoda nadmašila ostale metode, čime su pokazali da ni jedna pojedinačna metoda ne može univerzalno da se prihvati u ovakvim istraživanjima, kao poželjnija od ostalih metoda. Tako, predmet ovog istraživanja nije bio poređenje različitih metoda i performansi, već samo prikaz tih metoda i prikaz moguće primene u domenu koji je bio predmet njihovih

istraživanja. Pored toga, svaki pojedini model se sastoji od većeg broja različitih parametara modela. Jedan od parametara koji može da se dodatno podešava je parametar za definisanje Funkcije gubitka. Tokom obučavanja modela, vrednost za gubitak je vrednost koja treba da se smanji na minimum. Ova funkcija može da se podesi na različite načine. U ovom istraživanju je korišćena „Huberova“ funkcija za gubitak (Huber, 1964), a koja se koristi u robusnoj regresiji, i koja je manje osetljiva na odstupanja (en. Outliers) u podacima, od uobičajene funkcije za kvadratni gubitak greške. Takođe, jedan od neophodnih parametara za optimizaciju modela mašinskog učenja su optimizatori, kojih takođe ima nekoliko na raspolaganju. U ovom istraživanju je korišćen „Nadam“ optimizator, odnosno to je optimizator „Adam“ uz „Nesterov momentum“ (Dozat, 2016). U daljem radu, moguće je vršiti istraživanje koje se odnosi na podešavanje ovih i ostalih parametara, kao i podešavanje arhitekture svakog pojedinog modela da bi se postigli bolji rezultati predviđanja.

Modeli koji su opisani u ovom radu mogu da se iskoriste za neku drugu grupu turista, za neku drugu grupu usluga koje se nude turistima i naravno za istraživanje broja turističkih poseta na nekoj drugoj teritoriji.

Literatura

Brownlee, J., 2018. *Deep Learning for Time Series Forecasting: Predict the Future with MLPs, CNNs and LSTMs in Python*. s.l.:Machine Learning Mastery. Chollet F. et al., 2015. *Keras*. [Na mreži] Available at: <https://keras.io> [Poslednji pristup2021].

- Chollet, F., 2018. *Deep Learning with Python*. s.l.:Manning Publications Co..
- Dozat, T., 2016. *Incorporating Nesterov Momentum into Adam*. s.l., ICLR 2016.
- Fotiadis, A., Polyzos, S. & Huan, T.-C. T., 2021. The good, the bad and the ugly on COVID-19 tourism recovery. *Annals of tourism research*, Tom 87.
- Ghalekhondabi, I., Ardjmand, E., Young, W. A. & Weckman, G. R., 2019. A review of demand forecasting models and methodological developments within tourism and passenger transportation industry. *Journal of tourism futures*, 5(1), pp. 75-99.
- Graves, A., 2013. Generating Sequences With Recurrent Neural Networks. *arXiv:1308.0850*.
- Huber, P. J., 1964. Robust Estimation of a Location Parameter. *The Annals of Mathematical Statistics*, 35(1), pp. 73-101.
- Republički zavod za statistiku, 2021. *Dolasci i noćenja turista po regionima - mesečni podaci*. [Na mreži]
Available at:
<https://data.stat.gov.rs/Home/Result/220203?languageCode=sr-Cyrl>
- [Poslednji pristup 04 2021].
- Sun, S., Wei, Y., Tsui, K.-L. & Wang, S., 2019. Forecasting tourist arrivals with machine learning and internet search index. *Tourism Management*, Tom 70, pp. 1-10.
- TensorFlow, 2015. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. [Na mreži]
Available at: <http://tensorflow.org/>
[Poslednji pristup 2021].
- TensorFlow, 2021. *Time series forecasting*. [Na mreži]
Available at:
https://www.tensorflow.org/tutorials/structured_data/time_series
[Poslednji pristup 2021].
- Zaborovskaia, O., Sharafanova, E. & Maksanova, L., 2020. Scenario Forecasting Tourist Flows during the COVID-2019 Pandemic. *International Journal of Technology*, 11(8), pp. 1570-1578.
- Zhang, H., Song, H., Wen, L. & Liu, C., 2021. Forecasting tourism recovery amid COVID-19. *Annals of Tourism Research*, 87(4).