



PREDIKCIJA INDEKSA KORISNOSTI IGRAČA U EVROLIGI

PREDICTION OF THE PERFORMANCE INDEX RATING OF BASKETBALL PLAYERS IN THE EUROLEAGUE

Bakir Nikšić, *Fakultet tehničkih nauka, Novi Sad*

Oblast – ELEKTROTEHNIKA I RAČUNARSTVO

Kratak sadržaj – Problem kojim se bavi ovaj rad jeste predikcija indeksa korisnosti igrača u evroligi. Ovaj rad nastoji da iskoristi veliku količinu podataka, koji se nalaze na oficijalnom sajtu evrolige, obradi ih i pripremi za upotrebu algoritama mašinskog učenja. Algoritmi koji su razmatrani u ovom radu su Naive Bayes, SVM (Support Vector Machine) i KNN (K nearest neighbours). Za svaki algoritam izvršena je optimizacija parametara koji najviše utiču na samu klasifikaciju. Osim navedenih algoritama mašinskog učenja, razmatrana su još dva jednostavna algoritma za predikciju indeksa korisnosti koje je implementirao autor. Najbolju tačnost od 92% postigao je algoritam SVM, dok su algoritmi Naive Bayes i KNN imali manju tačnost od 87%, odnosno 72%. Namenski razvijeni algoritmi imali su značajnije manju tačnost od algoritama mašinskog učenja.

Abstract – This paper presents and evaluates machine learning (ML) as well as simple custom made approaches for the prediction of the Performance Index Rating of basketball players in the Euroleague. The approaches rely on the large amounts of data, available on the official website of the Euroleague. The data are processed and prepared in order to create a training and evaluation datasets for the application of ML algorithms. The algorithms which were considered are: Naive Bayes, SVM (Support Vector Machines) and KNN (K nearest neighbours). Parameters of each of the algorithms were optimized using a validation set. Besides the ML models, two simple custom made algorithms were designed and implemented by the author of this paper. The best accuracy of 92% was achieved by the SVM algorithm, while the Naive Bayes and KNN algorithms had a lower accuracy of 87% and 72% respectively. The accuracy of the custom made algorithms was significantly lower when compared to the machine learning models.

Keywords: prediction in sport, machine learning

1. UVOD

U današnje vreme, veliki procenat populacije priklonjen je praćenju sporta. Kao jedan od najpopularnijih sportova današnjice, izdvaja se košarka. Kako bi praćenje košarke učinilo uzbudljivijim, određeni broj ljudi se upušta u pojedine vidove takmičenja. Fantazi (eng. Fantasy), predstavlja jedan od najpopularnijih vidova ovakve vrste takmičenja.

NAPOMENA:

Ovaj rad proistekao je iz master rada čiji mentor je bio dr Aleksandar Kovačević, vanr.prof.

U ovom takmičenju, učesnici pokušavaju da sastave što bolji tim uz ograničeni budžet. Takmičari kupuju igrače za koje smatraju da će ostvariti veliki indeks korisnosti. Ostvareni indeks korisnosti svih igrača se sabiraju i dobija se ukupni broj bodova koje je takmičar osvojio u trenutnom kolu. Ovo takmičenje traje kroz čitavu sezonu. Svaki učesnik ima svoj tim, i na kraju sezone, pobjednik je učesnik koji je ostvario najviše bodova u sezoni. Ovaj rad bavi se predikcijom ostvarenog indeksa korisnosti igrača u Evroligi, na osnovu statistika igrača sa prethodno odigranih utakmica. U narednoj sekciji, biće opisani neki od radova koji su se bavili sličnim problemom. Nakon toga, u trećoj sekciji, biće opisana definicija problema i metodologije za njegovo rešavanje. Četvrta sekcija se bavi eksperimentalnom evaluacijom dobijenih rezultata, dok će u petoj sekciji biti reči o sumarnizaciji rada i pravcima budućeg razvoja.

2. PRETHODNA REŠENJA

Veoma je mali broj radova koji se bavi rešavanjem problema predikcije indeksa korisnosti. Međutim, postoje radovi koji rešavaju slične probleme, koji mogu da posluže i koji predstavljaju na neki način vodilju, odnosno, put kojim bi trebalo ići, kako bi se došlo do rešenja ovog problema.

Prvo rešenje, čiji je problem rešavanja bio najbliži rešavanju problema predikcije indeksa korisnosti u Evroligi, bio je rad [1]. Tema ovog rada bila je predikcija indeksa korisnosti košarkaša u NBA ligi. Na osnovu podataka već odigranih utakmica i metoda mašinskog učenja, prediktuje se indeks korisnosti igrača. Takođe, sabiranjem prediktovanih indeksa, vrši se predikcija i pobjednika utakmice. Ovaj rad koristio je skup podataka u kojem su se nalazili statistički podaci igrača u prethodnim sezonama NBA lige. Sam skup podataka sadržao je podatke o broju pogođenih šuteva, promašenih šuteva, skokova, asistencija, blokadi, ukradenih lopti, izgubljenih lopti i postignutih poena. Ovo rešenje koristilo je metode mašinskog učenja, linearnu regresiju, slučajne šume i mašine potpornih vektora (Support vector machine, SVM). U radu se došlo do zaključka da se u predikciji najbolje pokazao metod SVM, koji je u najboljem slučaju pogađao indeks u opsegu +/- 3.2 indeksna poena.

Sledeće rešenje koje je razmatrano bio je rad [2]. Tema ovog rada je predikcija indeksa najboljeg tima koji se sastoji od 8 igrača, sa ograničenjem da postoji ukupan budžet u koji se cena tih igrača mora uklopiti. Skup podataka koji je ovde korišćen, formiran je na osnovu

podataka iz sezona 2014/2015 i 2015/2016. Obeležja koja se nalaze u ovom dataset-u su statističke brojke koje je igrač ostvario tokom odigravanja utakmice. Tu spadaju procenti šuteva, broj skokova, asistencija, kao i mnogobrojni parametri koji utiču na način računanja indeksa korisnosti. Ovaj rad koristio je metode mašinskog učenja linearnu regresiju i naivni Bajesov klasifikator. U radu su se dobijeni rezultati upoređivali sa rezultatima koje je predvideo DraftKings [3] sajt i kao najbolji algoritam, pokazao se naivni Bajesov klasifikator. Rad [4], bavi se predikcijom košarkaša koji će biti MVP (Most Valuable player). Model čija je tačnost 69%, izdvaja 5 najvažnijih atributa, odnosno atributa koji po autoru najviše utiču na indeks korisnosti. Najvažnijim atributima se dodeljuju težine, koje su izračunate na osnovu medijana prethodnih pet dobitnika ove nagrade. Takođe, model u obzir uzima i težinu koja se odnosi na tim za koji igrači nastupaju. Dodeljivanje težina se u ovom radu pokazalo kao izuzetno poboljšanje prethodnih modela čija je preciznost bila ispod 50%. Na osnovu toga, ustanovljeno je da i u rešavanju ovog problema takođe postoji prostor gde bi se određenim stvarima dodeljivalje težine.

3. METODOLOGIJA I ALATI

U ovoj sekciji predstavljene su primenjene metodologije za određivanje indeksa korisnosti igrača u Evroligi. Izložen je pristup u procesu dobavljanja i pretprocesiranja podataka, kao i svaki od 3 korišćena algoritma za mašinsko učenje. Pored algoritama mašinskog učenja, predstavljena su i dva namenski napisana algoritma od strane autora ovog rada. Dobijeni skup podataka sadržao ~70.000 redova. Ovaj skup podeljen je tako da 70% čini trening skup podataka, dok ostalih 30% čine testni skup podataka.

3.1. Prikupljanje podataka

Prvi i osnovni problem koji je bilo potrebno rešiti jeste, prikupiti dovoljno podataka, koji će se dalje obrađivati i pripremati, i na osnovu kojih će se dolaziti do prediktovanih vrednosti ostvarenog indeksa korisnosti igrača u Evroligi. Nije bilo moguće dobiti skup podataka u već donekle pripremljenom formatu i kao takvog ga preuzeti sa interenta. Međutim, na oficijalnom sajtu Evrolige [5], nalazi se statistika igrača sa 17 prethodnih sezona (počevši od sezone 2000/01, pa do sezone 2017/18). S obzirom na to, bilo je potrebno iskoristiti tehnike web scraping-a (tehnik skidanja podataka sav eb stranica), kako bi se željeni podaci dobavili. Za ovako nešto, potrebno je bilo pronaći odgovarajući alat, sa kojim je moguće izvršiti *web scraping*. Istraživajući, pronađen je alat *DataToolbar*. Ovaj alat radi tako što se definiše skript u kojoj se specificira, koji podaci se kojim redosledom dobavljaju sa veb stranice.

3.2. Predprocesiranje podataka

Nakon uspešno preuzetih podataka, naredni korak koji je bilo potrebno uraditi, jeste obraditi podatke i spremati ih za dalju upotrebu. Potrebno je bilo primeniti razne tehnike *data wranglinga*, i pomoću njih doći do skupa podataka koji je spreman za dalje korišćenje. *Data wrangling* tehnike, su tehnike čišćenja, strukturiranja i prenošenja sirovih podataka u format koji je poželjan za dalje

korišćenje podataka. Za ovako nešto, razvijen je alat koji je automatski rešavao sve probleme obrade podataka. Nakon automatske obrade skupa podataka, primećeni su neki specifični propusti prilikom obrade podataka i ovi problemi rešeni su ručno naknadno.

3.3 Redukcija dimenzionalnosti

Nakon završetka procesa opisanog u sekciji 3.3, skup podataka sadržao je 28 kolona, od kojih određeni broj kolona nije bio numeričkog tipa, poput npr. imena i prezimena igrača, dok veći deo jeste bio numeričke prirode. S obzirom da indeks korisnosti predstavlja numeričku vrednost, u formulu indeksa korisnosti ulaze samo numerički atributi, kojih je u ovom slučaju velik broj. Indeks korisnosti računa se po sledećoj formuli preuzetoj iz [6]:

$$\text{INDEX} = Pts + OReb + DReb + As + St + Rv + Fv - To - Pf - Ag - (FGA - FG) - (FTA - FT)$$

Za redukciju dimenzionalnosti iskorištena je linearna regresija, pomoću koje se tražila zavisnost određenih numeričkih atributa. Pomoću linearne regresije, jasno je mogla da se vidi zavisnost između atributa i da se pronađu atributi koji ne utiču na vrednost indeksa. Od 22 numerička atributa, linearna regresija je pokazala da 7 atributa nema nikakvu zavisnost sa ciljnom varijablom - indeksom, odnosno koeficijent korelacije iznosi 0 te su te kolone redukovane.

3.4 Diskretizacija ciljnog atributa

Nakon redukcije dimenzionalnosti i smanjenja broja atributa, dodat je novi atribut – grupa. Grupa je kolona koja predstavlja u kom opsegu se nalazi indeks koji je ostvario igrač na utakmici. Ovaj atribut je uveden kako bi se diskretizovao ciljni atribut, odnosno kako bi se mogli primeniti klasifikacioni algoritmi. Postoji 8 grupa:

- Grupa 1 [-50, 0)
- Grupa 2 [0, 5)
- Grupa 3 [5, 10)
- Grupa 4 [10, 15)
- Grupa 5 [15, 20)
- Grupa 6 [20, 25)
- Grupa 7 [25, 30)
- Grupa 8 [30, 35)
- Grupa 9 [35, 100)

3.5 Algoritmi za mašinsko učenje

Za potrebe primene algoritama mašinskog učenja, potrebno je bilo napraviti novi skup podataka, odnosno modifikovati već pripremljen skup podataka. Ova modifikacije bila je neophodna jer algoritmi mašinskog učenja rade sa istorijskim podacima. U skupu ne postoje ulazni parametri na osnovu kojih se prediktuje indeks korisnosti, s obzirom da indeks korisnosti zavisi od vrednosti parametara koje su ostvarene na istoj utakmici. U ovom radu je odabrano da se kao ulazni parametri koriste srednje vrednosti tih parametara ostvarenih na prethodnim utakmicama. Ovakav pristup očuvava najviše informacija i ne favorizuje davne ili skorije utakmice. Primenjena su 3 algoritma mašinskog učenja: Naive Bayes, SVM i KNN.

3.5.1 Naive Bayes

Klasifikator naivnog Bajesa (*Naive Bayes*) spada u klasu statističkih klasifikatora. Zasniva se na Bajesovoj teoremi. Međutim, ovaj klasifikator zanemaruje uslovnu nezavisnost atributa u Bajesovoj teoremi - on

podrazumeva da vrednost jednog atributa u okviru jedne klase ne zavisi od vrednosti ostalih atributa u toj klasi. Ovo se naziva *naivna pretpostavka* [6] [7].

3.5.2 SVM (Support Vector Machines)

Mašine potpornih vektora (*Support Vector Machines - SVM*) su skup algoritama za mašinsko učenje, koji se koriste za klasifikaciju i regresiju. Ideja algoritma jeste da se u prostoru u kome su podaci predstavljeni, pronađe razdvajajuća hiper-ravan, tako da su svi podaci iz date klase sa iste strane ravni [8]. Prilikom evaluacije SVM klasifikatora isprobane su različite vrednosti optimizacionih parametara. U ovom istraživanju najbolje se pokazala *C-SCV* implementacija klasifikatora SVM sa radijalnom (*RBF*) kernel funkcijom. Klasifikator je davao različite tačnosti u zavisnosti od odabira parametra *C*. Na osnovu validacionog skupa, eksperimentalno je ustanovljeno da se najveća tačnost dobija kada je parametar *C=80*.

3.5.2 KNN (K Nearest Neighbour)

K najbližih komšija (*K nearest neighbour - KNN*) je algoritam mašinskog učenja koji se zasniva na primerima, tzv. lenjo učenje. Objekat se klasifikuje na osnovu broja najbližih komšija. K najbližih komšija nekog objekta, su k tačaka koje su od objekta udaljene manje od ostalih tačaka [7]. Prilikom evaluacije KNN klasifikatora isprobane su različite vrednosti optimizacionog parametra *k*. Na osnovu validacionog skupa, eksperimentalno je ustanovljeno da se najveća tačnost dobija kada je parametar *k=5*.

3.6 Namenski razvijeni algoritmi

U ovom poglavlju, biće opisani algoritmi koji su razvijeni od strane autora ovog master rada, za rešavanje problema predikcije ostvarenog indeksa korisnosti košarkaša u Evroligi. Motivacija za razvoj ovih algoritama je, pored postizanja velike tačnosti, i istraživačke prirode. Algoritmi mašinskog učenja mogu biti veoma kompleksni i vremenski zahtevni, kako za obučavanje, tako i za primenu. Iz tog razloga, izvršeni su eksperimenti sa dva jednostavna algoritma za predikciju indeksa korisnosti koji se oslanjaju na srednju vrednost ostvarenih statistika i potragu za već postojećom pojavom sekvence u skupu podataka. U ovom poglavlju biće opisana dva algoritma koja su razvijena, a to su:

1. Algoritam srednje vrednosti
2. Algoritam slične sekvence

3.6.2 Algoritam srednje vrednosti

Algoritam srednjih vrednosti se zasniva na računanju srednjih vrednosti svih parametara zasebno, ostvarenih na prethodnim utakmicama. Osim toga, ideja ovog algoritma jeste da posebno favorizuje poslednjih pet utakmica, kao i utakmice koje su odigrane protiv ekipe sa kojom igrač igra utakmicu u kojoj se prediktuje njegov indeks korisnosti. Sasvim je očekivano da forma košarkaša varira tokom sezone. Zbog toga je poslednjih pet utakmica potrebno posebno favorizovati, u odnosu na ostatak utakmica, jer su to utakmice koje najobjektivnije opisuju trenutnu formu igrača. Isto tako, utakmica protiv ekipe sa kojom je već igrao ima poseban značaj, te bi i njih na neki način trebalo favorizovati.

3.6.2 Algoritam slične sekvence

Ideja algoritma slične sekvence je da pokuša da pronađe sličnu sekvencu ostvarenih indeksa korisnosti. Cilj je

prediktovati *n+1* utakmicu jednog igrača, tako što će se u skupu podataka naći *n* sličnih ostvarenih indeksa koje je ostvario na prethodnih *n* utakmica igrač. Ukoliko je pronađenih sekvenci više, računa se srednja vrednost svih dobijenih grupa pete utakmice iz pronađenih sekvenci i ona se prediktuje. Eksperimentalno je utvrđeno da je ovaj algoritam bilo moguće primeniti (samim tim i evaluirati) samo ako se predviđaju indeksi korisnosti do prvih 9 utakmica. Bilo je i pojedinačnih slučajeva kod kojih je predikcija bila moguća i posle 9 utakmica, ali je takvih slučajeva bilo vrlo malo. Algoritam je davao rezultate dakle samo za prvih ~9 utakmica.

4. EVALUACIJA I REZULTATI

Eksperimentalna evaluacija se vrši kako bi se odredile performanse svih algoritama. Postoje različite tehnike evaluacije u sistemima analize podataka. U ovom istraživanju performanse klasifikatora evaluirane su upotrebom test skupa podataka. Modeli su se obučavali na 70% podataka iz skupa, dok se testiranje vršilo na ostalih 30%. Ovih 30% podataka nije bilo korišćeno ni na koji način tokom obučavanja. Metrike evaluacije performansi klasifikatora koje su korišćene u ovom radu su preciznost (eng. *precision*), povrat (eng. *recall*) i tačnost (eng. *accuracy*) [9]. Preciznost i povrat su metrike koje se određuju zasebno za svaku klasu. Tačnost je globalna metrika za određivanje mere uspešnosti klasifikatora [9].

Tabela 1. Preciznost povrat i tačnost klasifikatora Naive Bayes po grupama

Grupa	Preciznost	Povrat	Tačnost
Grupa1	85.23%	93.91%	85.23%
Grupa2	85.52%	78.57%	85.52%
Grupa3	87.21%	85.47%	87.21%
Grupa4	88.92%	89.45%	88.92%
Grupa5	91.41%	89.91%	91.41%
Grupa6	94.47%	94%	94.47%
Grupa7	92.94%	90.29%	92.94%
Grupa8	90%	91.30%	90%
Grupa9	71.86%	91.67%	71.86%

Ukupna tačnost klasifikatora naivnog Bajesa iznosi 87.16%. Preciznost je u opsegu 85%-95%, sa izuzetkom grupe 9, dok je povrat je takođe u opsegu 85%-95%, sa izuzetkom grupe 2.

Tabela 1. Preciznost povrat i tačnost SVM klasifikatora po grupama

Grupa	Preciznost	Povrat	Tačnost
Grupa1	99.68%	97.38%	99.68%
Grupa2	97.49%	99.70%	97.49%
Grupa3	94.44%	99.95%	94.44%
Grupa4	90.05%	91.46%	90.05%
Grupa5	75.87%	82.85%	75.87%
Grupa6	43.91%	47.75%	43.91%
Grupa7	0%	0%	0%
Grupa8	0%	0%	0%
Grupa9	0%	0%	0%

Ukupna tačnost klasifikatora mašina potpornih vektora iznosi 91.99%. Preciznost i povrat su u opsegu 44%-100%, sa izuzecima grupa 7, 8, i 9.

Ukupna tačnost klasifikatora k najbližih komšija iznosi 72.77%. Preciznost i povrat variraju od grupe do grupe i to u opsegu od 24%-100% sa izuzetkom grupe 9.

Tačnost je izračunata i za namenski napisane algoritme. Za algoritam srednje vrednosti tačnost je merena za dva slučaja. Prvi slučaj obuhvata predikovanje indeksa

korisnosti igrača na osnovu njegovih rezultata u svim sezonama unazad. Drugi slučaj obuhvata prediktovanje indeksa na osnovu samo trenutno aktuelne sezone.

Tabela 2. Preciznost povrat i tačnost KNN klasifikatora po grupama

Grupa	Preciznost	Povrat	Tačnost
Grupa1	99.65%	99.42%	99.65%
Grupa2	92.31%	78.08%	92.31%
Grupa3	63.66%	63.31%	63.66%
Grupa4	51.44%	56.98%	51.44%
Grupa5	41.56%	49.05%	41.56%
Grupa6	32.76%	38.25%	32.76%
Grupa7	26.34%	36.57%	26.34%
Grupa8	23.38%	26.09%	23.38%
Grupa9	8%	8.33%	8%

Tabela 4. Tačnost namenskog algoritma srednjih vrednosti

Broj sezona	Tačnost
Sve sezone	31%
Trenutna sezona	42%

Tačnost je izračunata i za namenski napisan algoritam slične sekvence, kako bi se on mogao porediti sa rezultatima algoritama mašinskog učenja. Dobijena tačnost u slučaju kada ovaj algoritam vraća rešenje, iznosi 36%. Analizirajući dobijene rezultate svih modela, ustanovilo se da se tačnost modela kreće u opsegu od 36% do 92%. U tabeli 6, može se videti tačnost svih iskorišćenih algoritama u rešavanju problema predikcije indeksa korisnosti igrača u Evroligi ponaosob.

Tabela 6. Uporedni rezultati evaluacije svih algoritama

Algoritam	Tačnost
Naive Bayes	87%
SVM	92%
KNN	72%
Algoritam srednjih vrednosti	42%
Algoritam slične sekvence	36%

5. ZAKLJUČAK

Problem koji je rešavan u ovom master radu je predikcija indeksa korisnosti igrača koji nastupaju u najkvalitetnijem evropskom takmičenju – Evroligi. U rešavanju ovog problema, ispitan je rad algoritama za mašinsko učenje: Naive Bayes, SVM i KNN. Korišćene su implementacije ovih algoritama u RapidMiner alatu. Osim ovih algoritama mašinskog učenja, namenski su implementirana dva algoritma: algoritam srednjih vrednosti i algoritam slične sekvence. Tačnost koja je dobijena u predikciji indeksa korisnosti igrača u Evroligi varira od algoritma do algoritma. Najveću tačnost ostvario je klasifikator mašina potpornih vektora, koji je prediktovao sa tačnošću od 92%, dok su se ostali algoritmi pokazali lošijim. Najmanju tačnost ostvario je namenski napisan algoritam slične sekvence.

Na osnovu prethodno dobijenih rezultata, može se zaključiti da ovaj rad predstavlja dobar korak ka rešavanju sličnih problema, koji se mogu rešavati. U svakom sportu postoji mera koja određuje koliko je neki igrač dobar, te bi ovaj rad mogao da bude slično rešenje takvih problema.

Ono što treba imati na umu, jeste to da na osnovu svega navedenog, idalje postoji prostora za unapređenje i pronalaženje boljeg i preciznijeg rešenja. Samo povećanje skupa podataka, utiče na rad pojedinih algoritama, te se očekuje da se njihov kvalitet rešenja poboljšava iz godinu u godinu, kako novi podaci pristižu. Osim toga, postoji prostora za unapređivanje poput upotrebe još nekih od algoritama mašinskog učenja. Takođe, u budućnosti bi moglo da se pokuša sa rigoroznijim opsegom prihvatljivosti onoga šta je tačno. Trenutne grupe pravljene su u opsezima od 5, dok bi u budućnosti bilo poželjno da se opsezi grupe smanje na recimo 3. Ovako nešto bi definitivno smanjilo tačnost trenutnih algoritama, te bi bilo poželjno razviti ili iskoristiti neki algoritam koji bi za ovakve uslove davao veliku tačnost. Osim toga, sama tema bi mogla da se proširi i da se prediktuje odabir čitavog tima za fantasy takmičenje. Ovo bi uključilo i određena ograničenja poput budžeta i pozicija igrača. Takođe, bilo bi poželjno napraviti web stranicu koju bi budući korisnici mogli da iskoriste za svoje potrebe.

6. LITERATURA

- [1] Chan-Hu-Shivakumar (2015) *Learning to Turn Fantasy Basketball Into Real Money Introduction to Machine Learning* <https://shreyasskandan.github.io/files/report-ChanHuShivakumar.pdf>
- [2] Eric Hermann and Adebina Ntoso (2015), *Machine Learning Applications in Fantasy Basketball* http://cs229.stanford.edu/proj2015/104_report.pdf
- [3] Draft kings official website <https://www.draftkings.com/>
- [4] Mason Chen (2017) *Predict NBA Regular Season MVP Winner* <http://ieomsociety.org/bogota2017/papers/9.pdf>
- [5] Euroleague official website <https://euroleague.net/>
- [6] K. M. Leung, "Naive Bayesian Classifier", Polytechnic University, Department of Computer Science, Finance and Risk Engineering, November 2007.
- [7] A. Kovacevic, Predavanja iz predmeta "Sistemi za istraživanje i analizu podataka", školska 2015/2016., Fakultet tehničkih nauka, Univerzitet u Novom Sadu
- [8] "Support Vector Machines", SciKit, <http://scikit-learn.org/stable/modules/svm.html>
- [9] D. Petrović, "Preprocesiranje podataka i generisanje skupa atributa za sentiment analizu Tviter poruka", Fakultet tehničkih nauka, Univerzitet u Novom Sadu, decembar 2016.

Kratka biografija



Bakir Nikšić rođen je 17.04.1994. u Prištini. Osnovnu školu „Miloš Crnjanski“ u Subotici završio je 2009. godine. Tehničku školu „Ivan Sarić“ u Subotici, završio je 2013. godine. Iste godine upisao je osnovne akademske studije na smeru Računarstvo i automatiku, Fakulteta tehničkih nauka u Novom Sadu. Osnovne studije je završio 2017. godine, nakon čega upisuje master akademske studije na Fakultetu tehničkih nauka, smer Elektronsko poslovanje. Položio je sve ispite propisane planom i programom.