



PREDIKCIJA CENE FINANSIJSKE AKTIVE ANALIZOM SENTIMENTA  
FINANSIJSKIH TVITOVA

PRICE PREDICTION OF FINANCIAL ASSETS USING SENTIMENT ANALYSIS OF  
FINANCIAL TWEETS

Dušan Stević, *Fakultet tehničkih nauka, Novi Sad*

**Oblast – ELEKTROTEHNIKA I RAČUNARSTVO**

**Kratka sadržaj** – U ovom radu predstavljen je sistem za predikciju cena finansijske aktive analizom sentimenta finansijskih tvitova i vesti. Cilj razvijenog sistema je ispitivanje sledeće hipoteze: postoji korelacija između sentimenta finansijskih tvitova i cene finansijske aktive. To jest, proveramo da li će rezultati sentiment analize doprineti boljim rezultatima predikcije. Razvijeni sistem se sastoji od dva podsistema. Prvi podsistem je zadužen za sentiment analizu, dok je drugi podsistem zadužen za predikciju cena finansijske aktive. Studija slučaja izvršena je na primeru tvitova i akcija Apple korporacije u vremenskom periodu od januara 2010. godine do decembra 2020. godine. Predloženi su koraci za poboljšanje i dalji razvoj sistema.

**Ključne reči:** sentiment analiza, NLP, transfer učenja, DistilBERT, predikcija cena, LSTM, fintech

**Abstract** – This paper presents a system for predicting the prices of financial assets by analyzing the sentiment of financial tweets and news. The developed system aims to test the following hypothesis: there is a correlation between the sentiment of financial tweets and the price of financial assets. The developed system consists of two subsystems. The first subsystem is responsible for sentiment analysis, while the second subsystem is responsible for predicting the prices of financial assets. The case study was conducted on the example of tweets and stocks of Apple Corporation in the period from January 2010 to December 2020. Steps for further system improvement are discussed.

**Keywords:** sentiment analysis, NLP, transfer learning, DistilBERT, price prediction, LSTM, fintech

**1. UVOD**

Domen finansija pokazao se kao izuzetno privlačan i izazovan za primenu tehnika veštačke inteligencije. Osnovna motivacija za izradu ovog rada je produbljanje znanja iz oblasti veštačke inteligencije u domenu finansija, ali i aplikativnost softverskog rešenja u procesu donošenja investicionih odluka. Multidisciplinarni pristup u izradi rada omogućio je primenu *state of the art* - SOTA algoritama veštačke inteligencije na polju finansija.

**NAPOMENA:**

Ovaj rad proistekao je iz master rada čiji mentor je bila dr Jelena Slivka, vanredni profesor.

Rad se sastoji od tri tematske celine. Prva tematska celina bavi se problemima rudarenja i obrade podataka (engl.

*data mining*). Kako podaci predstavljaju sirovine za modele veštačke inteligencije posebna pažnja se posvećuje raznorodnosti podataka. Na performanse ovih modela direktno utiče kvalitet obrađenih podataka. Raznorodnost podataka, koja se ogleda u različitim izvorima i tipovima podataka, obezbeđuje da se u radu pored tradicionalnih strukturiranih tipova podataka koriste i nestrukturirani tipovi podataka. U poslednjih nekoliko godina beleži se dominantnost nestrukturiranih tipova podataka nad strukturiranim.

Eksplorzija nestrukturiranih podataka doprinela je da se u radu prilikom predikcije cene finansijske aktive pored tradicionalnih tabelarnih finansijskih izveštaja koriste i finansijski tvitovi (engl. *tweets*). Na ovaj način u radu je postignuta raznorodnost podataka koja se manifestuje u simbiozi strukturiranih (tabelarni berzanski izveštaji) i nestrukturiranih (finansijski tvitovi i vesti) podataka.

Druga tematska celina bavi se klasifikacionim problemima konkretno klasifikacijom sentimenta finansijskih tvitova i vesti. Sentiment analiza (engl. *sentiment analysis* - SA) ili ekstrakcija mišljenja i stavova (engl. *opinion mining*) za cilj ima da utvrdi da li neki tekst ima negativan, neutralan ili pozitivan prizvuk. Na donošenje investicionih odluka pored kretanja berzanskih indeksa utiče i investiciono raspoloženje javnog mnjenja. Iz tog razloga neophodno je prilikom predikcije cene finansijske aktive uključiti i analizu sentimenta finansijskih tvitova i vesti.

Treća tematska celina bavi se regresionim problemima, konkretno, pronalaženjem korelacije između kretanja cene finansijske aktive i sentimenta finansijskih tvitova. Osnovni zadatak ove celine je da ispita uticaj koji finansijski tvitovi i vesti, kao izraz javnog mnjenja, imaju na berzanska kretanja.

U nastavku rada će detaljnije biti objašnjeni različiti aspekti rešavanja problema, kao i način realizacije samog rešenja. U poglavlju 2 je napravljen pregled prethodnih radova koji se bave istom ili sličnom problematikom. Poglavlje 3 opisuje proces formiranja ciljnih skupova podataka korišćenih u daljim analizama. U poglavlju su opisani skupovi podataka o finansijskim tvitovima i kretanju cene finansijske aktive. U nastavku poglavlja predstavljen je postupak obrade sakupljenih podataka u nastojanju da se dobiju ciljni skupovi podataka. Poglavlje 4 daje osvrt na metodološke pristupe korišćene u radu. Poglavlje 5 prezentuje dobijene rezultate nakon sprovedenog istraživanja. Poglavlje 6 je zaključno poglavlje u kojem se sumarizuje celokupan rad.

## 2. PRETHODNA REŠENJA

U ovom poglavlju predstavljena su prethodna rešenja koja su se bavila sentiment analizom finansijskih tvitova i predikcijom cene finansijske aktive.

Autori rada [1] su predstavili model za sentiment analizu zasnovan na pravilima (engl. *rule-based approach*) – VADER (*Valence Aware Dictionary and sEntiment Reasoner*<sup>1</sup>). VADER se zasniva na jezičkim pravilima na osnovu kojih se određuje polaritet i intenzitet sentimenta tekstualnog sadržaja. Autori su grupisali jezička pravila u pet opštih heuristika (engl. *generalizable heuristics*) na osnovu kojih su formirali VADER model. Koristeći VADER model za sentiment analizu tvitova autori su postigli bolje rezultate od ljudskih anotatora. U radu se kao klasifikaciona metrika koristila F1 mera tačnosti. VADER model je postigao 0.96, dok su ljudski anotatori postigli 0.84 F1 mere tačnosti.

U radu [2] autori su predstavili model za sentiment analizu finansijskih tvitova zasnovan na pravilima - *TextBlob*<sup>2</sup>. Kao i VADER, *TextBlob* spada u modele nenadgledanog učenja. Autori rada su izvršili ispitivanje korelacije između sentimenta finansijskih tvitova i kretanja cene akcija. *TextBlob* se zasniva na jezičkim pravilima na osnovu kojih se određuje polaritet i subjektivnost sentimenta tekstualnog sadržaja. Autori su na raspolaganju imali 15 000 javno dostupnih tvitova koji imaju tag “\$AAPL”. Za evaluaciju rešenja koristili su koeficijent korelacije. U radu koeficijent korelacije iznosi 0.611639 što nam govori da između prosečne vrednosti sentimenta i cene akcije postoji jaka pozitivna korelacija tj. objave na *Twitter*-u utiču na tržišna kretanja.

Autori rada [3] razvili su specijalizovani finansijski rečnik za potrebe sentiment analize iz domena finansija (engl. *dictionary-based approach*). Rečnik je dobio ime po autorima koji su ga kreirali - *Loughran-McDonald*. Pristup sentiment analizi baziran na rečniku podrazumeva postojanje unapred pripremljenog rečnika u kojem je dostupan polaritet reči. *Loughran-McDonald* rečnik<sup>3</sup> je nastao kao finansijska specijalizacija *Harvard IV-4* rečnika<sup>4</sup> opšte namene. Pristup sentiment analizi baziran na rečnicima spada u kategoriju nenadgledanog učenja.

U radu [4] autori su predstavili pristup sentiment analizi baziran na korpusima (engl. *corpus-based approach*). Korpus-bazirana sentiment analiza koristi prikupljene anotirane tvitove na osnovu kojih se formira korpus. Korpus-bazirana sentiment analiza spada u pristupe rukovođene podacima (engl. *data-driven approach*). Prikupljeni podaci su neformalnog karaktera i često su domen specifični. Zbog različite prirode polazišnih podataka korpusu-bazirana sentiment analiza može bolje da prepozna žargon i neformalne izraze. Pristup sentiment analizi baziran na korpusu spada u kategoriju nenadgledanog učenja.

Autori rada [5] su ispitivali hipotezu da li objave na *Twitter*-u imaju uticaja na finansijska tržišta. U radu je detaljno opisan postupak pretprocesiranja prikupljenih tvitova. Za potrebe sentiment analize korišćen je pristup zasnovan na mašinskom učenju (engl. *ML-based approach*) konkretno SVM (engl. *Support Vector Machine*) klasifikator. Pristupi sentiment analizi zasnovani na mašinskom učenju spadaju u nadgledano učenje. Autori su SVM model obučavali koristeći labelirane (anotirane) tvitove. Kao rezultat istraživanja potvrđena je polazna hipoteza da postoji pozitivna korelacija između objava na *Twitter*-u i tržišnih kretanja.

Rad [6] se bavi predikcijom cena akcija koristeći sentiment analizu podataka sa *Twitter*-a. U radu je opisan postupak prikupljanja i obrade podataka sa berze i *Twitter*-a. Sentiment analiza je sprovedena koristeći SVM i NB (engl. *Naive Bayes*) klasifikatore. Za evaluaciju dobijenih rezultata sentiment analize koristila se tačnost. NB je postigao tačnost od 0.80609, dok je SVM postigao tačnost od 0.79308.

Cilj rada [7] je bio da se izvrši analiza sentimenta finansijskih tvitova upotrebom neuronskih mreža. U radu su korišćene CNN (engl. *Convolutional Neural Network*), RNN (engl. *Recurrent Neural Network*) i LSTM (engl. *Long Short-Term Memory*). Pored analize sentimenta finansijskih tvitova ideja je da se izvrši i predikcija kretanja cena *Apple* korporacije u budućnosti. Rađena je komparativna analiza više modela zasnovanih na neuronskim mrežama. Metodološki se problem klasifikacije tvitova uliva u problem regresije. Kada se dobije sentiment tvitova on predstavlja dodatno obeležje za regresioni problem. Autori su na raspolaganju imali 334K javno dostupnih tvitova koji imaju cashtag “\$AAPL”. Evaluacija rešenja vršna je podelom na trening i test skup. Za potrebe evaluacije klasifikacije koristi se tačnost i F1 mera. Dok se za potrebu regresije koristi  $R^2$  (engl. *coefficient of determination*) i MSE (engl. *mean squared error*). Doprinos ovog rada se manifestuje u komparativnom pristupu izučavanju modela i jedinstvenom pristupu disekcije realnog problema na probleme klasifikacije i regresije. Primena sveobuhvatnog metričkog sistema za evaluaciju, omogućava komparaciju različitih modela.

U radovima [8][9] autori su koristili DistilBERT model za sentiment analizu tekstualnog sadržaja iz domena finansija. DistilBERT model spada u modele zasnovane na transferu učenja (engl. *transfer learning*). DistilBERT model je pre sentiment analize fino podešen (engl. *fine-tuning*) za finansijsku terminologiju. Fino podešavanje DistilBERT modela je neophodno kako bi se generički model bolje prilagodio rešavanom problemu. Postupak finog podešavanja DistilBERT-a spada u nadgledano učenje. U oba rada DistilBERT model je postigao superiorne performanse u poređenju sa modelima mašinskog učenja.

Radovi [8][10] na sistematičan način daju komparativni pregled trenutnih modela i tehnika u oblasti sentiment analize.

## 3. FORMIRANJE SKUPOVA PODATAKA

Za potrebe testiranje hipoteze prikupljeni su strukturirani i nestrukturirani tipovi podataka (engl. *data mining*). Strukturirani tipovi podataka obuhvataju tabelarne

<sup>1</sup> Implementacija VADER modela se može pronaći na <https://github.com/cjhutto/vaderSentiment>

<sup>2</sup> Implementacija TextBlob modela se može pronaći na <https://github.com/sloria/TextBlob>

<sup>3</sup> Lista pozitivnih i negativnih reči Loughran-McDonald rečnika može se pronaći na <https://sraf.nd.edu/textual-analysis/resources/>

<sup>4</sup> Lista pozitivnih i negativnih reči Harvard IV-4 rečnika može se pronaći na <http://www.wjh.harvard.edu/~inquirer/homecat.htm>

berzanske izveštaje. Nestrukturirani tipovi podataka obuhvataju dve grupe finansijskih podataka. Prvu grupu predstavljaju labelirani (anotirani) finansijski tvitovi i vesti, dok drugu grupu predstavljaju nelabelirani (neanotirani) finansijski tvitovi i vesti. Labelirani finansijski tvitovi i vesti koriste se za potrebe obučavanja algoritama nadgledanog učenja. Nelabelirani tvitovi i vesti koriste se za potrebe predikcije cena finansijske aktive.

Akvizicija labeliranih (anotiranih) finansijskih tvitova i vesti vršena je manuelno sa sajtova specijalizovanih za deponovanje skupova podataka. Prikupljeno je ukupno 6 skupova labeliranih podataka: [Apple Twitter Sentiment](#), [FinancialPhraseBank](#), [Stock Market Tweet](#), [Stock market sentiment](#), [Daily News for Stock Market Prediction](#) i [Two Sigma](#). Akvizicija nelabeliranih (neanotiranih) tvitova vršena je pomoću *Twint API-a*<sup>5</sup>. Listu korisnika čiji tvitovi se sakupljaju čine najeminentiji ljudi, televizijske kuće i korporacije iz sveta finansija (engl. *influencers*). Listu korisnika čiji tvitovi se sakupljaju sastavljena je manuelno na osnovu domen specifičnog znanja iz oblasti finansija.

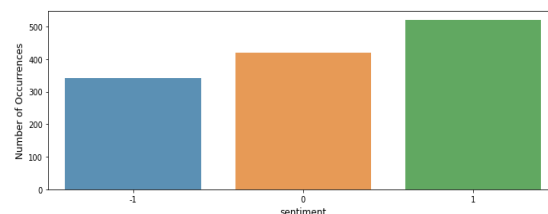
Prilikom sakupljanja tvitova, filtrirani su se samo oni tvitovi koji u sebi sadrže termine iz liste termina. Listu termina čine cashtag \$AAPL, hashtag #AAPL i mention @apple. Akvizicija tabelarnih berzanskih izveštaja vršena je pomoću *Yahoo Finance API-a*<sup>6</sup>. Prikupljeni su podaci o korporaciji Apple od januara 2010. godine do decembra 2020. godine.

Nakon prikupljanja podataka pristupljeno je različitim tehnikama *data wrangling-a* i pretprocesiranja podataka. Neke od korišćenih tehnika za pretprocesiranje tekstualnih podataka su<sup>7</sup>: eliminacija stop reči (engl. *stop word removal*), morfološka normalizacija reči upotrebom steminga ili lematizacije.

Pored pretprocesiranja tvitova, potrebno je izvršiti i pretprocesiranje numeričkih vrednosti u tabelarnim berzanskim izveštajima. Neke od korišćenih tehnika za pretprocesiranje numeričkih podataka su<sup>8</sup>: rukovanje nedostajućim vrednostima i autlajerima (engl. *outliers*), numerička normalizacija vrednosti atributa upotrebom standardizacije ili normalizacije. Primenjene tehnike za cilj imaju konverziju sirovih podataka u oblik pogodan za dalju analizu. Eksplorativna analiza podataka je iskorišćena za potrebe boljeg upoznavanja sa prikupljenim podacima. Slika 1 predstavlja distribuciju sentimenta finansijskih tvitova i vesti po labelama.

#### 4. METODOLOGIJA

Nakon što su prikupljeni i obrađeni podaci, pristupilo se dizajniranju arhitekture sistema kako bismo testirali polaznu hipotezu. Kompletan arhitektura sastoji se iz dva podsistema.



Slika 1. Distribucija sentimenta tvitova po labelama

Prvi podsistem se bavi sentiment analizom. Sentiment analiza u osnovi predstavlja klasifikacioni problem. U okviru ovog podsistema korišćene su nenadgledane i nadgledane tehnike za sentiment analizu.

U okviru nenadgledanih tehnika korišćene su tehnike zasnovane na pravilima (*VADER* i *TextBlob*) i tehnike zasnovane na rečnicima (*Harvard IV-4* opšti rečnik i *Loughran-McDonald* finansijski rečnik). U okviru nadgledanih tehnika korišćene su tehnike zasnovane na ML (*Support-vector machine*, *Random forest*, *Logistic regression* i *Naive Bayes*) i tehnike zasnovane na DL (*DistilBERT*).

Da bi se prikupljeni finansijski tvitovi i vesti mogli propustiti kroz ML i DL algoritme potrebno je prvo izvršiti ekstrakciju osobina iz teksta. Postupak vektorizacije vrši se ekstrakcije feature-a iz teksta. Za potrebe ML algoritama korišćene su dve tehnike za vektorizaciju teksta: *Bag of words – BOW* i *term frequency-inverse document frequency – TF-IDF*. Za potrebe DL algoritama korišćen je kontekstualni embedding kojim se vrši vektorizacija teksta.

Drugi podsistem bavi se predikcijom buduće cene finansijske aktive. Predikcija buduće cene finansijske aktive u osnovi predstavlja regresioni problem. Pored istorijskih podataka o kretanju cene finansijske aktive, ulaz u podsistem za predikciju cene finansijske aktive predstavljaju i anotirani finansijski tvitovi i vesti iz podsistema za sentiment analizu. Nelabelirani finansijski tvitovi i vesti nakon ulaska u podsistem za sentiment analizu postaju anotirani i kao takvi predstavljaju ulaz u podsistem za predikciju.

Ovim ulančavanjem obezbeđuje se dodatni izvor informacija prilikom predikcije cene finansijske aktive. Anotirani tvitovi se postupkom temporalne fuzije<sup>9</sup> spajaju sa istorijskim podacima o kretanju cene finansijske aktive. Cilj postupka temporalne fuzije jeste da se izvrši obogaćivanje istorijskih podataka o kretanju cene finansijske aktive dodatnim izvorima informacija. Ti dodatni izvori informacija zapravo predstavljaju anotirane finansijske tvitove i vesti koji su rezultat rada podsistema za sentiment analizu.

Regresioni problem predikcije budućih cena finansijske aktive posmatra se kroz prizmu finansijskih vremenskih serija.

Zbog sekvencijalne prirode finansijskih vremenskih serija, projekcija budućih cena finansijske aktive vrši se korišćenjem tehnike klizećeg prozora (engl. *sliding window technique*) i LSTM neuronske mreže. Postupakom klizećeg prozora vrši se ekstrakcije obeležja iz vremenske serije.

<sup>5</sup> Implementacija Twint API-a može se pronaći na

<https://github.com/twintproject/twint>

<sup>6</sup> Implementacija Yahoo Finance API-a može se pronaći na

<https://python-yahoofinance.readthedocs.io/en/latest/api.html>

<sup>7</sup> Korišćena biblioteka za postupak pretprocesiranja tekstualnog sadržaja može se pronaći na <https://www.nltk.org/howto/sentiment.html>

<sup>8</sup> Korišćena biblioteka za postupak pretprocesiranja numeričkog sadržaja može se pronaći na [https://scikit-](https://scikit-learn.org/stable/modules/preprocessing.html)

[learn.org/stable/modules/preprocessing.html](https://scikit-learn.org/stable/modules/preprocessing.html)

<sup>9</sup> Spajanje heterogenih podataka po vremenu nastanka

## 5. REZULTATI I DISKUSIJA

Formiranjem arhitekture sistema može se pristupiti testiranju hipoteze. Evaluacija oba podsistema vršena je podelom odgovarajućeg skupa podataka na trening, validacioni i test skup podataka. Kao mera performansi u podsistemu za sentiment analizu koristi se tačnost. Najbolje performanse od svih korišćenih tehnika za sentiment analizu u pogledu tačnosti postigao je DistilBERT model (Tabela 1).

Tabela 1 *Rezultati podsistema za sentiment analizu*

Tip učenja	Pristup baziran na	Model	Tačnost
Nenadgledano	Pravilima	VADER	0.54
		TextBlob	0.49
	Rečnicima	Harvard IV-4	0.38
		Loughran-McDonald	0.57
Korpusima	Stock market sentiment korpus	0.41	
Nadgledano	ML	SVM	0.70
		RF	0.74
		LR	0.73
		NB	0.68
	DL	DistilBERT	0.89

Superiorne performanse u rešavanju različitih NLP zadataka, DistilBERT dugujem ogromnim obučavajućim skupovima podataka i transferu znanja. Za potrebe testiranja hipoteze koristila su se dva pristupa: Pearsonov koeficijent korelacije i komparacija regresionih modela. Visoke vrednosti Pearsonovog koeficijenta korelacije indiciraju snažan uticaj koji javno mnjenje ima na cenu finansijske aktive.

Najviše vrednosti Pearsonovog koeficijenta korelacije postigao je DistilBERT (0.88). Visoke vrednosti Pearsonovog koeficijenta korelacije indiciraju da je pristup baziran na DL uspeo da detektuje povezanost između sentimenta finansijskih tvitova i tržišnih kretanja. Nakon što je detektovan klasifikator koji može najviše da doprinese poboljšanju predikcije cene finansijske aktive pristupa se ispitivanju hipoteze. Da bismo ispitali hipotezu, potrebno je da se razviju dva prediktivna LSTM modela čiji performanse će se upoređivati (Tabela 2).

Tabela 2 *Rezultati podsistema za predikciju*

Podaci	Model	R <sup>2</sup>	RMSE
Berza	LSTM1	0.75	4.7447
Berza + SA	LSTM2	0.87	3.1642

Prvi LSTM je razvijen samo na osnovu istorijskih podataka sa berze, dok je drugi LSTM razvijen koristeći pored istorijskih podataka sa berze i rezultate iz podsistema za sentiment analizu. Kao mere performansi u podsistemu za predikciju koriste se koeficijent determinacije – R<sup>2</sup> i koren srednje kvadratne greške - RMSE. Dodavanje rezultata sentiment analize u regresioni model doprinelo je smanjenju greške regresionog modela (RMSE je pala sa 4.7447 na 3.1642) i poboljšanju prediktivnosti regresionog modela (R<sup>2</sup> je porastao sa 0.75 na 0.87), čime je potvrđena polazna hipoteza.

## 6. ZAKLJUČAK

Sprovedeno istraživanje potvrdilo je hipotezu da postoji pozitivna korelacija između sentimenta finansijskih tvitova i kretanja cene finansijske aktive. Pozitivan sentiment uslovljava rast cena finansijske aktive, dok negativan sentiment uslovljava pad cena finansijske aktive. Potvrđivanjem polazne hipoteze zaokružena je priča o predikciji cene finansijske aktive analizom sentimenta finansijskih tvitova.

## 7. LITERATURA

- [1] Hutto, C., & Gilbert, E. (2014, May). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the International AAAI Conference on Web and Social Media* (Vol. 8, No. 1).
- [2] Ao, S. (2018, July). Sentiment analysis based on financial tweets and market information. In *2018 International Conference on Audio, Language and Image Processing (ICALIP)* (pp. 321-326). IEEE.
- [3] Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of finance*, 66(1), 35-65.
- [4] Darwich, M., Mohd, S. A., Omar, N., & Osman, N. A. (2019). Corpus-Based Techniques for Sentiment Lexicon Generation: A Review. *J. Digit. Inf. Manag.*, 17(5), 296.
- [5] Smailović, J., Grčar, M., Lavrač, N., & Žnidarišič, M. (2013, July). Predictive sentiment analysis of tweets: A stock market application. In *International workshop on human-computer interaction and knowledge discovery in complex, unstructured, big data* (pp. 77-88). Springer, Berlin, Heidelberg.
- [6] Kordonis, J., Symeonidis, S., & Arampatzis, A. (2016, November). Stock price forecasting via sentiment analysis on Twitter. In *Proceedings of the 20th Pan-Hellenic Conference on Informatics* (pp. 1-6).
- [7] Chen, C. C., Huang, H. H., & Chen, H. H. (2018, April). Fine-grained analysis of financial Tweets. In *Companion Proceedings of the The Web Conference 2018* (pp. 1943-1949).
- [8] Mishev, K., Gjorgjevikj, A., Vodenska, I., Chitkushev, L. T., & Trajanov, D. (2020). Evaluation of sentiment analysis in finance: from lexicons to transformers. *IEEE Access*, 8, 131662-131682.
- [9] Dogra, V., Singh, A., Verma, S., Jhanjhi, N. Z., & Talib, M. N. (2021). Analyzing DistilBERT for Sentiment Classification of Banking Financial News. In *Intelligent Computing and Innovation on Data Science* (pp. 501-510). Springer, Singapore.
- [10] Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, 5(1), 1-167.

### Kratka biografija:



**Dušan Stević** rođen je 1991. godine. Osnovne akademske studije završio je 2020. godine na Fakultetu tehničkih nauka, na kom brani i master rad 2021. godine iz oblasti Elektrotehnike i računarstva – Softversko inženjerstvo i informacione tehnologije.

Kontakt: [stevicdule@gmail.com](mailto:stevicdule@gmail.com)