

Određivanje tipa osobnosti autora teksta temeljeno na dubokom učenju za klasifikaciju sa složenim oznakama klasa

Čerkez, Ninoslav

Doctoral thesis / Disertacija

2022

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Electrical Engineering and Computing / Sveučilište u Zagrebu, Fakultet elektrotehnike i računarstva**

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:168:980172>

Rights / Prava: [In copyright](#) / [Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2024-04-26**



Repository / Repozitorij:

[FER Repository - University of Zagreb Faculty of Electrical Engineering and Computing repository](#)





Sveučilište u Zagrebu

FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

NINOSLAV ČERKEZ

**ODREĐIVANJE TIPA OSOBNOSTI AUTORA
TEKSTA TEMELJENO NA DUBOKOM UČENJU
ZA KLASIFIKACIJU SA SLOŽENIM OZNAKAMA
KLASA**

DOKTORSKI RAD

Zagreb, 2022.



Sveučilište u Zagrebu

FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

NINOSLAV ČERKEZ

**ODREĐIVANJE TIPRA OSOBNOSTI AUTORA
TEKSTA TEMELJENO NA DUBOKOM UČENJU
ZA KLASIFIKACIJU SA SLOŽENIM OZNAKAMA
KLASA**

DOKTORSKI RAD

Mentori:

prof. dr. sc. Boris Vrdoljak

doc. dr. sc. Sandro Skansi

Zagreb, 2022.



University of Zagreb

FACULTY OF ELECTRICAL ENGINEERING AND COMPUTING

NINOSLAV ČERKEZ

**DETERMINING PERSONALITY TYPE OF TEXT
AUTHOR BASED ON DEEP LEARNING FOR
CLASSIFICATION WITH COMPOUND CLASS
LABELS**

DOCTORAL THESIS

Supervisors:

Professor Boris Vrdoljak, PhD

Assistant Professor Sandro Skansi, PhD

Zagreb, 2022

Doktorski rad izrađen je na Sveučilištu u Zagrebu, Fakultetu elektrotehnike i računarstva, na Zavodu za primijenjeno računarstvo.

Mentori: Prof. dr. sc. Boris Vrdoljak i doc. dr. sc. Sandro Skansi

Doktorski rad ima: 138 stranica

Doktorski rad br.: _____

O mentoru

Boris Vrdoljak rođen je 1969. godine u Osijeku. Diplomirao je, magistrirao i doktorirao u polju elektrotehnike na Sveučilištu u Zagrebu Fakultetu elektrotehnike i računarstva (FER), 1995., 1999. odnosno 2004. godine.

U veljači 1996. godine zaposlio se na Zavodu za osnove elektrotehnike i električka mjerenja FER-a, a od ožujka 2006. godine radi na Zavodu za primijenjeno računarstvo FER-a. 2001. godine proveo je tri mjeseca na znanstvenom usavršavanju na Sveučilištu Bologna, Italija. Od listopada 2004. do rujna 2005. bio je na postdoktorskom usavršavanju na institutu INRIA u Francuskoj. Izabran je 2018. godine u zvanje redovitog profesora. Sudjelovao je na više znanstvenih projekata kao istraživač ili voditelj. Objavio je više od 40 radova u časopisima i zbornicima konferencija u području skladištenja podataka, analitičke obrada podataka, elektroničkog poslovanja, automatizacije pronalaženja podudarnosti u ontologijama te strojnog učenja. Prof. Vrdoljak član je stručne udruge IEEE, Znanstvenog centra izvrsnosti za znanost o podacima i kooperativne sustave te istraživačkog Laboratorija za informacijsku sigurnost i privatnost.

About the Supervisor:

Boris Vrdoljak was born in Osijek in 1969. He received his B.Sc., M.Sc. and Ph.D. degrees in electrical engineering from the University of Zagreb, Faculty of Electrical Engineering and Computing (FER), Zagreb, Croatia, in 1995, 1999 and 2004, respectively.

From February 1996 he was working at the Department of Fundamentals of Electrical Engineering and Measurements at FER, and from March 2006 he has been working at the Department of Applied Computing at FER. In 2001, he spent 3 months as a visiting researcher at the University of Bologna, Italy. From October 2004 to September 2005 he was a postdoctoral researcher at INRIA institute, France. In 2018 he was promoted to Full Professor. He participated in several scientific projects as a researcher or project manager. He published more than 40 papers in journals and conference proceedings in the area of data warehousing, analytical data processing, e-business, automation of ontology matching, and machine learning.

Prof. Vrdoljak is a member of IEEE, Centre of Research Excellence for Data Science and Cooperative Systems, and research Laboratory for information security and privacy.

O komentoru

Sandro Skansi rođen je 1985. godine u Zagrebu. Diplomirao je 2009. godine Filozofiju i Kroatologiju na Hrvatskim studijima Sveučilišta u Zagrebu, i doktorirao disertacijom iz logike na Filozofskom fakultetu Sveučilišta u Zagrebu. Član je Association for Symbolic Logic i Association for the Advancement of Artificial Intelligence. Od 2017. godine radi kao docent iz logike na Fakultetu hrvatskih studija Sveučilišta u Zagrebu. Njegovi istraživački interesi uključuju: zadovoljivost propozicijske logike, teoriju grafova, neizrazitu logiku, kibernetiku i umjetne neuronske mreže.

About the Co-advisor:

Sandro Skansi was born in 1985 in Zagreb. He attained his MA in Philosophy and Croatian Culture in 2009 at the University of Zagreb and completed a thesis in mathematical logic at the University of Zagreb in 2013. He has been a member of the Association for Symbolic Logic since 2006 and a life member the Association for the Advancement of Artificial Intelligence from 2016. From 2017 he is an Assistant professor in Logic at the University of Zagreb. His research interests include the satisfiability of propositional logic, graph theory, fuzzy logic, cybernetics, and artificial neural networks.

Ovaj rad posvećujem svojim roditeljima koji su mi pružili ljubav i dali potporu kroz život te svojim sinovima Tiboru i Jakovu da im bude poticaj za njihov rast i razvoj.

Zahvala

Najveću zahvalnost želim izraziti svojoj supruzi Lei i sinovima Tiboru i Jakovu na stalnoj potpori, strpljenju i razumijevanju za moju čestu odsutnost duhom i tijelom za vrijeme istraživačkog rada i pisanja doktorske disertacije.

Hvala mojim roditeljima i bratu na moralnoj podršci i razumijevanju.

Hvala mojoj punici Snježani za svu pomoć u logistici kako bih nesmetano mogao raditi na doktoratu.

Zahvaljujem prof. dr. sc. Borisu Vrdoljaku i doc. dr. sc. Sandru Skansiju na mentorstvu, stručnom vodstvu i savjetima tijekom istraživačkog rada.

Sažetak

Ova doktorska disertacija rezultat je istraživanja u području višeklasne klasifikacije tipova osobnosti autora teksta sa složenim oznakama klasa. Predikcija tipa osobnosti autora teksta ima dobro poznatu uporabu u psihologiji te praktične primjene u poslovnom okruženju. Iz perspektive znanosti o podacima, na problem predikcije osobnosti autora teksta može se gledati kao problem tekstualne klasifikacije te se može rješavati uporabom metoda obrade prirodnog jezika (engl. *natural language processing – NLP*) te metodama dubokog učenja.

U disertaciji je dan pregled rezultata prijašnjih poznatih istraživanja klasifikacije MBTI na način da su prezentirani rezultati primjenom standardnih algoritama strojnog učenja te rezultati istraživanja s primjenom dubokog učenja. Dosadašnja istraživanja višeklasne klasifikacije sa složenim oznakama klasa nisu uzimala u obzir mogućnost uključivanja informacija u komponentama složenih oznaka klasa. Disertacija predlaže novi algoritam za određivanje tipa osobnosti autora teksta temeljen na dubokom učenju s funkcijom gubitka za višeklasnu klasifikaciju prema indikatoru tipa po Myers-Briggsovoj (engl. *Myers-Briggs Type Indicator - MBTI*) koji poboljšava postojeće rezultate višeklasne klasifikacije MBTI budući da uzima u obzir komponente složenih oznaka klasa kao potporu za bolju klasifikaciju prema instrumentu MBTI. Ovakav pristup je važan jer rješava problem uključivanja komponenti složenih oznaka klasa u višeklasnu klasifikaciju te poboljšava rezultate prijašnjih poznatih istraživanja.

Eksperimentalni su rezultati demonstrirani nad dvije arhitekture umjetnih neuronskih mreža: LSTM i CNN a korišteni su i standardni algoritmi strojnog učenja. Rezultati istraživanja s novim algoritmom i funkcijom gubitka uspoređeni su sa standardnim rezultatima pristupa kategoričke unakrsne entropije i s rezultatima prijašnjih poznatih istraživanja. Napravljena je usporedba sa standardnim modelima za višeklasnu klasifikaciju, sličnim istraživanjima za višeklasnu klasifikaciju te istraživanjima s četiri binarna pristupa klasifikaciji MBTI. Algoritam se može koristiti i za ostale probleme višeklasne klasifikacije u situacijama kada postoje binarno isključive komponente složenih oznaka klasa. Na kraju je predstavljen prototip aplikacije s implementiranim predloženim algoritmom za višeklasnu klasifikaciju MBTI, a koji uzima u obzir složene oznake klasa MBTI.

Ključne riječi: strojno učenje, obrada prirodnog jezika, umjetne neuronske mreže, Indikator Tipa osobnosti po Myers-Briggsovoj - MBTI, višeklasna klasifikacija, binarna klasifikacija, složene oznake klasa, gubitak unakrsne entropije.

Extended Abstract

This dissertation results from research in the multi-class classification of the author's personality types with complex class labels. Psychologists define personality as a relatively stable set of traits and unique characteristics that determine the consistency and differences in the individual psychological behavior of people (thoughts, feelings, and actions).

The prediction of the author's personality type has a well-known use in psychology and practical applications in the business environment. From the perspective of data science, the problem of predicting the personality of the text author can be seen as a text classification problem and can be solved using natural language processing (NLP) and deep learning methods.

A selected assessment instrument that relies on the appropriate personality theory determines the personality type assessment. Theories of human personality provide a framework for understanding personality and are based on different assumptions of human nature, depending on what the individual researcher has emphasized. For example, the factors by which theories of human personality differ are free will vs. determinism, nature vs. upbringing, the importance of the influence of the past vs. the importance of the present, and uniqueness vs. universality. Throughout history, different researchers have had different approaches to personality analysis:

Researchers had begun to explore ideas about personality type in a way that is important for the dissertation, especially since 1921 when the Swiss psychiatrist Carl G. Jung proposed the foundations of personality typology in the book *Psychological Types* and introduced the concepts of function and attitude as well as specific names for the four functions of conscious orientation. Carl Jung was a student of Sigmund Freud and explained human nature with an approach he called analytical psychology. In addition, Jung introduced the term psyche for psychic energy that manifests itself as an individual's personality through psychological activities such as thinking, perceiving, feeling, and wishing.

Psychologists have developed reliable and valid personality assessment instruments; the most famous are the MBTI, Big Five (OCEAN), Enneagram, and DiSC. In predicting the personality type of the author of the text, the most significant number of studies conduct research on data obtained using the MBTI and Big Five instruments. However, The Big Five model has less popularity in the business environment than the MBTI model (e.g., in

recruitment processes), and in general, the MBTI model is widely prevalent among laypeople. Additionally, the Big Five model does not allow exploring complex class labels with exclusive binary values, which is the focus of this dissertation.

The Myers-Briggs Type Indicator (MBTI) is a personality assessment instrument according to Jung's personality theory. The questionnaire aims to distinguish applicants according to two categories on each of the four key dimensions. According to the MBTI instrument, more than two million people take the test every year for employment and business promotion.

Jung defined functions as an example of the dichotomy of opposites, which is a significant fact for research in this paper. Jung developed the idea of a hierarchy of functions since a person cannot have each of the eight functions equally developed. However, there is a dominant function, as the most developed, and an auxiliary function that supports the dominant function so that the attitudes or orientations of the leading and auxiliary functions differ. The dichotomies are E-I, S-N, T-F, and J-P. The Judging-Perceiving dichotomy was developed not to label people or human behavior but to help approach the dynamics of cognitive processes. Practitioners often view the MBTI model as a tool to 'classify' people, neglecting the model's dynamic and developmental aspects. Analyzing the deeper dimensions of the MBTI model, the first dimension is Extrovert (E) or Introvert (I). The dimension indicates that the person is outgoing, talkative, or reserved.

In other words, the dimension defines the person's orientation towards the external or internal world and constitutes the primary energy motivation. The second dimension is sensitivity (S) or intuition (N). This dimension defines how a person receives information. For example, a person with a more sensory approach processes more facts, while a more intuitive approach tries to interpret information and find deeper meaning. The third dimension is thinking (T) and feeling (F). This dimension describes how a person makes decisions. For example, a person with a primary thinking approach uses logic and consistency in reasoning and decision-making, while an approach with more emotions uses empathy and focuses on people and specific circumstances. The last dimension is judgment (J) and perception (P). This dimension describes personal orientation towards the outside world and how a person spends everyday life; in other words, it describes personal lifestyle. For example, a judgmental person tends to have an organized daily life, while a perceptive person prefers flexibility. This structure leads us to the 16 possible combinations of personality types of the MBTI model. Since each class consists of 4 components, it is evident that these are complex class labels.

According to the MBTI model, researchers mainly view the personality classification problem as a quadruple binary classification problem or a multiclass classification problem of compound class labels of the MBTI model. Deep learning models try, from the prepared input data, to classify the elements of individual dichotomies of the complex MBTI label, and the final classification result according to the complex class labels is obtained by combining the individual results. The deep learning model tries to classify the prepared input data according to the 16 complex MBTI class labels when approaching the problem as a multi-class classification.

The advantage of the approach with four binary classifiers is that the subsets used for binary classification are better balanced than the set with all 16 complex classes. An additional advantage is that such an approach provides flexibility, as the researchers can combine different classifiers for individual binary classifications. The disadvantage of the approach with four binary classifiers is that individual binary classifications cannot consider the influence of other components of a particular compound class. The disadvantage of the multi-class classification approach is that the set is highly unbalanced, which is the main reason for the poor results of the multi-class classification.

The dissertation approaches this problem as a multi-class classification with complex class labels, aiming to demonstrate how to include the influence of individual components of misclassified complex classes in the multi-class classification.

The dissertation presents an overview of previous well-known MBTI classification research in such a way that gives the results using standard machine learning algorithms and the results of research using deep learning. Previous research on multiclass classification with complex class labels has not considered the possibility of including information in the components of complex class labels.

The research used the MBTI data set from the Kaggle platform and the MBTI9k data set from the Reddit social network.

The first selected data set is from the Kaggle platform and contains 8,675 lines collected on the Personality Cafe forum in 2017. The data consists of posts by forum users in English, labeled with personality types according to the MBTI model. According to the MBTI model, the personality type label is not the result of a survey or some test, but each forum user

entered his personality type label along with the text. The dataset consists of 11.2 million words and more than 420,000 tags. Each row represents the last 50 posts of each user.

Another of the data sets used is the MBTI9k version of the data set based on discussion groups dedicated to the MBTI model from the Reddit platform. The post authors themselves created the MBTI-type labels of the individual post. The dataset contains 9252 rows of user comments containing more than 1000 words in the comment. When creating this dataset, the authors omitted comments with topics related to the topic of personality prediction using the MBTI model. Also, they omitted explicit occurrences of complex MBTI class labels when creating the data set.

The dissertation proposes a new algorithm for determining the author's personality type based on deep learning with a loss function for multi-class classification according to the Myers-Briggs Type Indicator (MBTI). The proposed approach improves the existing results of the multi-class MBTI classification since it considers components of complex class labels as support for better classification according to the MBTI instrument. This approach is important because it solves the problem of including components of complex class labels in multi-class classification and improves the results of previously known research.

The idea of the algorithm that uses multi-class classification, and includes individual components that are incorrectly determined, is to penalize the classification of a complex class label for the classification error of an individual component. For example, suppose the correct class label is INFJ, and the model predicts the ENFJ class label. So, we want to use the similarity of the compound class labels and penalize the model for the error made on the E/I dichotomy. In addition to the classification of the compound class, there was an error on the component as well. In this way, we enable the model to learn better to classify the component with an error in the classification of the compound class label.

The steps of the deep learning-based author personality type algorithm with a loss function that considers the similarity between complex class labels are: data cleaning and pre-processing, encoding of complex class labels in ascending order according to the English alphabet, probability calculation for the individual component of the dichotomy, access to individual components of complex class labels, loss function derivation from similarity-based evaluation measures of complex class labels, creating a deep learning model, training and validation of deep learning models, and implementation of the deep learning model in the application prototype.

Since the usual MBTI multi-class classification procedure implies, during the preparation of the data set, the conversion of MBTI type labels into integer values from 0 to 15, ensuring uniform access to individual components is necessary. This problem is solved by encoding sorted compound class labels.

The procedure for deriving a loss function from similarity-based evaluation measures of complex class labels includes the following steps: determining the categorical cross-entropy of the classification of a complex class label and determination of the categorical cross-entropy of an individual component on which the model made a classification error with the inclusion of weighting factors. An incorrect MBTI multiclass classification result includes an error in one of the 16 compound classes and at least one component of the composite class label. For example, for a true class label INTJ and a false prediction INTP, the multiclass classification model makes not only an error in classes with complex labels but also an error on individual components. In the given an example, it is an error on the last component, the dichotomy P/J.

The idea of the MBTI multi-class classification improvement procedure involves the fact that the complex class labels are similar; in this particular example, they are similar in that they have the first three components equal, but the last component is different. The similarity of the complex class labels will consider including in the calculation of the model's classification error, not only the cross-entropy of the complex class label but also the cross-entropy of the misclassified component.

Experiments demonstrate the results on two architectures of artificial neural networks: LSTM and CNN and standard machine learning algorithms. The research compares results with the new algorithm and loss function with the standard results of the categorical cross-entropy approach and the results of previously known research. Furthermore, it compares the results with standard models for multi-class classification, similar research for multi-class classification, and research with four binary approaches to MBTI classification.

The algorithm can also be used for other multiclass classification problems when exclusive binary components of complex class labels exist. In the end, the dissertation presents an application prototype with the implemented proposed algorithm for multi-class MBTI classification, which considers the complex labels of the MBTI classes.

Finally, the dissertation demonstrates the development of the application prototype using the Python programming language with supported programming libraries and frameworks.

Keywords: machine learning, natural language processing, artificial neural networks, Myers-Briggs Indicator Type - MBTI, multiclass classification, binary classification, compound class labels, cross-entropy loss

Sadržaj

1.	UVOD.....	1
1.1.	Opis problema i motivacija.....	1
1.2.	Izvorni znanstveni doprinos	3
1.3.	Struktura disertacije.....	4
2.	PRIMJENA DUBOKOG UČENJA U PREDIKCIJI TIPRA OSOBNOSTI AUTORA TEKSTA	5
2.1.	Procjena tipa osobnosti.....	6
2.2.	Psihološki modeli za procjenu tipa osobnosti	6
2.3.	Model MBTI.....	10
2.4.	Strojno učenje	14
2.4.1.	LSTM.....	18
2.4.2.	CNN.....	30
2.4.3.	Primjena klasifikacije u obradi prirodnog jezika	35
3.	PRIJAŠNJA ISTRAŽIVANJA – PRIMJENA STROJNOG UČENJA U KLASIFIKACIJI MBTI.....	40
3.1.	Istraživanja s primjenom standardnih algoritama strojnog učenja	41
3.2.	Istraživanja s primjenom dubokog učenja	42
4.	ALGORITAM ZA ODREĐIVANJE TIPRA OSOBNOSTI AUTORA TEKSTA TEMELJEN NA DUBOKOM UČENJU	52
4.1.	Metoda pristupa komponentama složenih oznaka klasa.....	56
4.2.	Izračun vjerojatnosti za pojedinu komponentu.....	58
4.3.	Određivanje ispravne komponente i položaja u složenoj klasi.....	60
4.4.	Postupak izvođenja funkcije gubitka iz mjera vrednovanja temeljenih na sličnosti složenih oznaka klasa	62
4.4.1.	Kategorička unakrsna entropija - CE.....	62
4.4.2.	Unakrsna entropija s utjecajem komponenti složenih oznaka klasa - CECI.....	63
5.	POSTAVKE I REZULTATI TRENIRANJA PRIMJENOM PREDLOŽENOG ALGORITMA	66
5.1.	Postavke treniranja	66
5.2.	Skupovi podataka	66
5.3.	Rezultati treniranja	74
5.3.1.	Model LSTM	74
5.3.2.	Model CNN	88
5.4.	Usporedbe s rezultatima drugih modela klasifikacije.....	104
6.	PROTOTIP PROGRAMSKOG SUSTAVA.....	106
6.1.	Mrežni okvir prototipa aplikacije	106
6.2.	Arhitektura prototipa aplikacije.....	108

6.3.	Korisničko sučelje prototipa aplikacije	111
7.	ZAKLJUČAK	116
8.	LITERATURA	119
9.	POPIS KRATICA	129
10.	POPIS SLIKA	131
11.	POPIS TABLICA	134
12.	ŽIVOTOPIS	136
13.	BIOGRAPHY	138

1. UVOD

1.1. Opis problema i motivacija

Osobnost (engl. *personality*) se može definirati kao relativno stabilan skup osobina i jedinstvenih karakteristika koje određuju konzistentnost i razlike u individualnom psihološkom ponašanju ljudi (misli, osjećaji, i akcije). Iako ne postoji općenita definicija osobnosti, postoji konsenzus da skup osobina koje čine osobnost imaju kontinuitet u vremenu te da se stoga ne mogu jednostavno razumjeti kao odvojeni rezultat socijalnog i biološkog utjecaja trenutka [1].

Procjena tipa osobnosti ima praktičnu primjenu osobito u poslovnom okruženju u sklopu procesa kao što su evaluacija i odabir kandidata za zaposlenje, odabir kandidata za projektne zadatke, upravljanje karijerom zaposlenika. Analiza tipa osobnosti može pomoći i u zdravstvenoj skrbi, budući da omogućava rano otkrivanje mentalnih poremećaja. Klinički psiholozi koriste procjenu osobnosti kao dodatni atribut u razumijevanju simptoma pacijenata i određivanju odgovarajuće terapije [2]. Dodatno, školski psiholozi procjenu osobnosti koriste za savjetovanje u odabiru daljnjeg profesionalnog razvoja ili otkrivanja problema u učenju ili ponašanju djece. Procjena osobnosti se koristi i za potrebe oružanih snaga u okviru kojih su tehnike procjene osobnosti i nastale, budući da je prva procjena osobnosti kreirana za potrebe vojske Sjedinjenih Američkih Država tijekom I svjetskog rata. Također, Taibi Kahler je, uz financijsku podršku NASA-e kreirao jedan od najčešće korištenih modela do današnjih dana. Suvremeni pristupi modeliranju osobnosti funkcioniraju na način da klasificiraju osobnost u određen broj dimenzija, a psiholozi razvijaju i koriste odgovarajući upitnik kao mjerni instrument [3] [4].

Standardni pristupi procjene tipa osobnosti podrazumijevaju primjenu tradicionalnih instrumenata kao što su samo-procjene, projekcije, upitnici te uključuju angažman visoko-obrazovanih stručnjaka psihologa. Time se standardni pristupi mogu ocijeniti sporima i skupima. Iz perspektive strojnog učenja, predikcija tipa osobnosti autora teksta je primjer obrade prirodnog jezika (engl. *natural language processing* - *NLP*) te se primjenom tehnika strojnog učenja proces predikcije tipa osobnosti može poboljšati i ubrzati [5]. Važno je istaknuti da primjena metoda strojnog učenja na problemima predikcije osobnosti otvara i kvalitativno drugačije nove oblike analize teksta, koji su temeljeni na velikoj količini podataka, a koji nisu usmjeravani od strane ispitivača. Ta količina podataka je načelno

prevelika da bi ju čovjek mogao analizirati i time metode strojnog učenja dobivaju specifičnu važnost.

Predmet istraživanja doktorske disertacije čini algoritam za višeklasnu klasifikaciju koji nadilazi ograničenja postojećih pristupa višestruke klasifikacije u slučajevima kada klase imaju složene oznake. Cilj algoritma je pokazati da u složenim oznakama postoje dodatne informacije i da je moguće dodatne informacije složenih oznaka klasa uključiti u okviru modela dubokog učenja. Dosadašnji algoritmi iz poznatih prijašnjih istraživanja nisu uključivali dodatne informacije složenih oznaka klasa.

Također, algoritam je važan jer treba omogućiti i poboljšanje rezultata višeklasne klasifikacije upotrebom informacija koje složene oznake klasa dodatno sadrže, u odnosu na jedino informaciju u složenoj oznaci klase. Stoga, dodatna motivacija je pokazati da je moguće unaprijediti rezultate dosadašnje strojne klasifikacije osobnosti uz pomoć složenih oznaka klasa. Glavne smjernice u istraživanju čine pristupi kojima se osigurava uključivanje pojedinih komponenti složenih oznaka klasa i koriste mjere sličnosti složenih oznaka klasa. Posredno, cilj je i pokazati da algoritam može ublažiti problem preklapanja te problem neuravnoteženosti složenih oznaka klasa skupa podataka MBTI. U sklopu istraživanja, uz razvijeni novi algoritam, koji čini izvorni znanstveni doprinos doktorske disertacije, cilj je u okviru algoritma pokazati i kako funkcija gubitka može uzeti u obzir oznake klasa koje su složene, slične i međuzavisne.

U disertaciji je fokus na uporabu strojnog učenja primjenom klasifikacije sa složenim oznakama klasa za određivanje tipa osobnosti koristeći indikator tipa osobnosti po Myers-Briggsovoj (engl. *Myers-Briggs Type Indicator – MBTI*), koji je vrlo popularan u poslovnom okruženju i među laicima. Detaljno pojašnjenje tipova osobnosti po Myers-Briggsovoj daje [6].

Osobna motivacija je proizašla iz dosadašnjeg profesionalnog iskustva, koje je uključivalo dugogodišnji rad sa suvremenim informacijskim tehnologijama u razvoju poslovnih informacijskih sustava i stalnu poslovnu potrebu za različitim oblicima profiliranja ljudi prema njihovoj osobnosti. Dodatna motivacija disertacije je bila i kreiranje osnove za buduće eksperimentalna istraživanja na osnovu dubljeg značenja tipova osobnosti prema MBTI instrumentu, a koji su vezani uz kognitivne funkcije.

U nastavku su izneseni i ukratko opisani izvorni znanstveni doprinosi doktorske disertacije, a zatim je opisana i struktura disertacije.

1.2. Izvorni znanstveni doprinos

Izvorni znanstveni doprinos ove doktorske disertacije se odnosi na algoritam koji uključuje komponente složenih oznaka klasa u višeklasnoj klasifikaciji autora teksta:

1. Algoritam za određivanje tipa osobnosti autora teksta temeljen na dubokom učenju s funkcijom gubitka koja u obzir uzima sličnost između složenih oznaka klasa.

U okviru prvog dijela doprinosa predložen je algoritam koji omogućava pristupanje pojedinačnim komponentama složenih oznaka klasa. Razvijeni algoritam uzima u obzir komponente po kojima se razlikuju slične složene oznake klasa te ih uključuje u predloženu funkciju gubitka.

2. Postupak izvođenja funkcije gubitka iz mjera vrednovanja temeljenih na sličnosti složenih oznaka klasa.

Drugi dio doprinosa se odnosi na postupak izvođenja funkcije gubitka. Mjere vrednovanja, na kojima se temelji funkcija gubitka, su izvedene iz postupaka kodiranja složenih oznaka klasa i izračuna vjerojatnosti pojedinih komponenti složenih oznaka klasa te su iste iskorištene nad komponentama na kojima se razlikuju slične složene oznake klasa.

3. Prototip programskoga sustava za određivanje tipa osobnosti sa složenim oznakama klasa temeljen na dubokom učenju.

Kao treći dio doprinosa kreiran je prototip programskog sustava u kojem je implementiran algoritam za određivanje tipa osobnosti autora teksta temeljen na dubokom učenju s korištenom funkcijom gubitka koja uzima u obzir sličnost između složenih oznaka klasa.

1.3. Struktura disertacije

Poglavlje 2 – Ovo poglavlje daje uvod u primjenu dubokog učenja u procjeni tipa osobnosti. Opisan je problem procjene tipa osobnosti te je dan pregled psiholoških modela za procjenu tipa osobnosti. Poseban naglasak je stavljen na MBTI instrument, budući da je MBTI u fokusu ove disertacije kao model na kojem su demonstrirani izvorni znanstveni doprinosi. Na kraju je predstavljen pregled obrade prirodnog jezika i modela dubokog učenja koji se koriste u praktičnom dijelu disertacije.

Poglavlje 3 – Ovo poglavlje daje pregled prijašnjih istraživanja u primjeni algoritama strojnog učenja za klasifikaciju MBTI. U prvom dijelu je napravljen pregled istraživanja u kojima su korišteni standardni algoritmi strojnog učenja. U drugom dijelu su prezentirani rezultati istraživanja u kojima su korišteni pristupi dubokog učenja.

Poglavlje 4 – Ovo poglavlje je vezano uz prvi i drugi znanstveni doprinos. Prvo je opisana metoda pristupa komponentama složenih oznaka klasa. Potom je prikazan postupak izračuna vjerojatnosti za pojedinu komponentu. Nakon toga je pokazan postupak određivanja ispravne komponente i položaja u složenoj klasi. Na kraju je pokazan postupak izvođenja funkcije gubitka iz mjera sličnosti složenih oznaka klasa.

Poglavlje 5 – Ovo poglavlje daje pregled postavki istraživanja u okviru ove doktorske disertacije. Potom, poglavlje prezentira rezultate treniranja primjenom metode višeklasne klasifikacije MBTI koja se temelji na uključivanju komponenti klasa. U svrhu ispunjavanja očekivanih znanstvenih doprinosa te dokazu provedbe ideja i mjerenja rezultata višeklasne klasifikacije MBTI postavljena su i provedena istraživanja uporabom modela s dugom kratkoročnom memorijom (engl. *long short-term memory* – *LSTM*) i modela s konvolucijskom neuronskom mrežom (engl. *convolutional neural network* – *CNN*). Na kraju su demonstrirani rezultati treniranja za odabrane LSTM i CNN modele te je napravljena usporedba rezultata.

Poglavlje 6 - Ovo poglavlje je vezano uz treći izvorni znanstveni doprinos. U ovom poglavlju je predstavljen prototip programskog sustava koji prikazuje uporabu višeklasne klasifikacije MBTI temeljenoj na uključivanju komponenti složenih oznaka klasa. Prvo je prikazano kreiranje mrežnog okvira prototipa aplikacije. Potom je pokazana arhitektura prototipa aplikacije. Na kraju je pokazano kreirano korisničko sučelje prototipa aplikacije.

Poglavlje 7 – Ovo poglavlje daje komentare na dobivene rezultate te zaključuje doktorsku disertaciju.

2. PRIMJENA DUBOKOG UČENJA U PREDIKCIJI TIPa OSOBNOSTI AUTORA TEKSTA

Područje obrade prirodnog jezika je područje u kojem je zadnjih nekoliko godina vrlo intenzivna primjena dubokog učenja. Klasične lingvističke i statističke metode za obradu prirodnog jezika zahtijevaju jezične eksperte za definiranje pravila u specifičnim situacijama. Mogućnosti nelinearnog modeliranja modela dubokog učenja su se pokazale vrlo uspješnim u poslovnim primjenama u NLP-u. Neka od primjena dubokog učenja u NLP-u su automatsko prevođenje teksta, prepoznavanje i generiranje naslova slike te automatska klasifikacija teksta. Predikcija tipa osobnosti autora teksta predstavlja problem koji spada u kategoriju problema automatske klasifikacije teksta pomoću dubokog učenja.

Iako i u području psihologije, preciznije kognitivne znanosti, postoje primjene i istraživanja u području neuronskih mreža, kao što je [7], u kojem se predlaže model neuronske mreže osobnosti temeljene na motivaciji, a na osnovu istraživanja o strukturi i neurobiologiji ljudske osobnosti, u svrhu premošćivanja udaljenosti između dinamičkog i strukturnog pristupa osobnosti, takve vrsta istraživanja nisu u fokusu istraživanja u okviru ove disertacije, budući da je u disertaciji naglasak na istraživanju pomoću modela dubokog učenja. Ipak, istraživanje [7] obrađuje zanimljive kognitivne modele koje se možda moći iskoristiti u budućim istraživanjima prilagodbi funkcije gubitka.

Primjer istraživanja koje je doprinijelo unapređenju oba područja je istraživanje [8] u kojem je pokazano kako se duboke neuronske mreže mogu koristiti kao računalni model za simulaciju ljudskog učenja riječi. Uočena je mogućnost povezivanja alata i tehnika iz područja razvojne psihologije i neuronskih mreža na problemu učenja riječi za objekte kod djece. Eksperimentalni dio je napravljen uzimajući u obzir skupove podataka koji su na raspolaganju iz područja psihologije te treniranjem modela na skupu podataka na ImageNet [9] bazi podataka. Uočena je identična pristranost prilikom učenja neuronske mreže, kao kod ljudi - preferencija kategorizacije objekata sukladno obliku u odnosu na boju. Zaključak je da tehnike kognitivne psihologije mogu pomoći boljem razumijevanju dubokih neuronskih mreža, pri čemu osiguravaju računalni model za učenje riječi kod ljudi. Istraživanje je pomoglo boljem razumijevanju dvaju tipova dubokih neuronskih mreža: početne (engl. *inception*) i podudarajuće mreže (engl. *matching networks*). Istraživanje je otvorilo i

moгуćnosti za daljnju primjenu tehnika kognitivne psihologije u razumijevanju i poboljšanju neuronskih mreža, ali se ipak ograničilo na dva tipa neuronskih mreža i područje učenja riječi.

Prije opisa algoritma za određivanje tipa osobnosti autora teksta koji se temelji na dubokom učenju, u nastavku je prvo opisan problem procjene tipa osobnosti. Nakon toga su predstavljeni modeli koje su psiholozi razvili u svrhu predikcije tipa osobnosti. Zatim je objašnjen model MBTI, budući da je algoritam koji predstavlja izvorni znanstveni doprinos disertacije demonstriran na modelu MBTI. Na kraju su opisani modeli dubokog učenja koji su se koristili u praktičnom dijelu disertacije.

2.1. Procjena tipa osobnosti

Psiholozi su za određivanje tipa osobnosti kreirali mnoštvo instrumenata za procjenu tipa osobnosti koji moraju ispunjavati uvjete da se radi o pouzdanim i valjanim instrumentima. Standardne tehnike procjene u psihologiji uključuju npr. samo-procjenu, projekcije, metode uzorkovanja. Ako se verificira konzistentnost mjerenih veličina s prihvatljivom varijancom, tehnika se označava pouzdanom. Dodatno, kada postoji potvrda da tehnika mjeri ciljanu osobinu, tehnika je valjana. S tom svrhom psiholozi su razvili tehnike i psihološke modele za procjenu osobnosti čija primjena rezultira predikcijom tipa osobnosti.

2.2. Psihološki modeli za procjenu tipa osobnosti

Procjena tipa osobnosti temelji se na odabranom instrumentu za procjenu koji se oslanja na odgovarajuću teoriju osobnosti. Teorije ljudske osobnosti osiguravaju okvir (engl. *framework*) za razumijevanje osobnosti i temelje se na različitim pretpostavkama ljudske prirode, ovisno o tome na što je pojedini istraživač stavio naglasak. Neki od čimbenika po kojima se razlikuju teorije ljudske osobnosti su: slobodna volja vs. determinizam, priroda vs. odgoj, važnost utjecaja prošlosti vs. važnost utjecaja sadašnjosti, jedinstvenost vs. univerzalnost itd. Kroz povijest su različiti istraživači imali različite pristupe analizi osobnosti: Sigmund Freud je utemeljio psihoanalitički pristup, Erik Erikson je kroz teoriju identiteta proširio psihoanalitički pristup, Alfred Adler se fokusirao na jedinstvenost svake osobe i razvio individualnu psihologiju, Karen Horney je stavila naglasak na utjecaj socijalne interakcije na razvoj osobnosti, Abraham Maslow je razvio humanistički pogled na osobnost itd. Disertacija se oslanja na instrument za procjenu tipa osobnosti koji je proizašao iz rezultata analitičkog

pristupa osobnosti koji je razvio Carl Jung, uz naglasak na računalni pristup procjene tipa osobnosti [10] [11].

Ideje o tipu osobnosti su se na način koji je bitan za disertaciju posebno počele istraživati od 1921., kada je švicarski psihijatar Carl G. Jung u djelu *Psychological Types* postavio osnove tipologije osobnosti te uveo pojmove funkcije i stava, kao i specifične nazive za četiri funkcije svjesne orijentacije. Carl Jung je bio učenik Sigmunda Freuda i ljudsku prirodu je objasnio pristupom koji je nazvao analitička psihologija. Jung je uveo pojam psihe za psihičku energiju koja se ispoljava kao osobnost pojedinca kroz psihološke aktivnosti kao što su razmišljanje (engl. *thinking*), percepcija (engl. *perceiving*), osjećanje (engl. *feeling*) i kreiranje želje (engl. *wishing*). Kada osoba uloži veću količinu psihičke energije u određenu ideju ili osjećanje, smatra se da ideja ili osjećanje imaju veliku psihičku vrijednost u životu pojedinca. Npr., ako je osoba motivirana da ostvari poziciju moći, posvetit će veliki dio psihičke energije pronalasku načina da ju ostvari. Jung je uveo i tri principa po kojima se ponaša psihička energija: suprotnosti, ekvivalentnost i entropija [12]. Princip suprotnosti naglašava da svaka želja ili osjećaj ima svoju suprotnost, a razina psihičke energije raste s većim konfliktom suprotnosti. Princip ekvivalentnosti govori o konzervaciji psihičke energije na način da se slabljenjem psihičke energije u nekom području energija samo transferira u neko drugo područje. Princip entropije naglašava da postoji tendencija prema održavanju balansa ili ekvilibrijuma u osobnosti. Ekvilibrijum se ostvaruje na način da npr. kod postojanja jače i slabije želje, psihička energija ima tendenciju tijeka prema slabijoj želji [13].

Jung je uveo termine stava (engl. *attitude*) i funkcije u opisu osobnosti. Stav definira orijentaciju kao internu ili eksternu. Kognitivne funkcije su esencijalne u Jungovoj teoriji u razvoju tipova osobnosti. Ipak, njihov utjecaj na MBTI nije u fokusu disertacije. Danas je moguće pronaći sinonime za termine funkcije u mentalnim procesima, kognitivnim procesima i kognitivnim funkcijama. Ključno je za model MBTI da svaka funkcija može imati eksterni i interni aspekt. Konačno, u modelu MBTI, postoje četiri dimenzije ili dihotomije, svaka se sastoji od dvije međusobno isključive kategorije.

Jung je razvio temeljne koncepte ekstrovertiranog i introvertiranog stava nakon analize radova i pristupa kolega Sigmunda Freuda i Alfred Adlera, koji su se u svom radu fokusirali na eksterni svijet odnosno unutarnji svijet pacijenta prilikom promatranja njihova ponašanja. Jung je istaknuo da je dio ljudi orijentiran ka interakciji prema vanjskom svijetu te da su ti ljudi ekstrovertirani u svojoj prirodi. Ljudi mogu biti orijentirani i prema unutarnjem svijetu,

na način da ih više ispunjavaju samostalne, reflektivne aktivnosti i po svojoj prirodi su takvi ljudi introvertirani. Pored ekstrovertiranosti i introvertiranosti, koje je nazvao stavom, Jung je smatrao da je bitno promatrati mentalne aktivnosti, koje je nazvao funkcijama. Danas se uglavnom za funkcije upotrebljavaju nazivi: mentalni procesi, kognitivni procesi ili kognitivne funkcije. Budući da se kognitivni procesi mogu koristiti na ekstrovertirani ili introvertirani način te da se svaka mentalna aktivnost sastoji od najmanje od jednog od četiri kognitivna procesa, postoji ukupno osam različitih procesa [14] [15].

Jung je vjerovao da se kompletna osobnost pojedinca, ili psiha, sastoji od dijelova ili aspekata koji mogu utjecati jedan na drugi:

- Ego, centar svjesnosti, dio psihe zadužen za percepciju, razmišljanje, osjećanje i pamćenje.
- Stavovi: ekstrovertiranost i introvertiranost. Jung je smatrao da se psihička energija može usmjeriti eksterno, prema vanjskom svijetu ili interno, prema sebi. Tako se ekstrovertirane osobe smatra otvorenima, socijalno aktivnima, orijentiranim prema drugim ljudima i vanjskom svijetu. Introvertirane osobe, s druge strane, su više povučene, često sramežljive, s tendencijom fokusa na sebe, svoje misli i osjećaje. Jung je smatrao da su u svakome od nas zastupljena oba stava, ali da samo jedan postaje dominantan i postaje dijelom osobne podsvijesti te djeluje na ponašanje osobe.
- Psihološke funkcije. Jung je smatrao da kako postoje različite vrste ekstrovertiranih i introvertiranih osoba, predložio je dodatne kriterije razlikovanja, koje je nazvao psihološkim funkcijama. Psihološke funkcije se odnose na različite načine percipiranja vanjskog - realnog i unutarnjeg-subjektivnog, svijeta. Jung je definirao četiri funkcije psihe: senzibilnost (engl. *sensing*), intuiciju (engl. *intuiting*), razmišljanje (engl. *thinking*) i osjećanje (engl. *feeling*). Senzibilnost i intuicija su grupirane zajedno, jer ne koriste procese rezoniranja, kao ne-racionalne funkcije. Ove funkcije prihvataju iskustva i ne evaluiraju ih. Senzibilnost reproducira iskustvo pomoću osjetilnih kanala, na način kako fotograf kopira objekt. Intuicija se ne aktivira preko vanjskih poticaja. Npr., ako vjerujemo da je još netko u tamnoj prostoriji, naše uvjerenje se može temeljiti na intuiciji a ne na nekom senzornom iskustvu. Drugi par funkcija, razmišljanje i osjećanje su racionalne funkcije i uključuju kreiranje procjene i evaluacije iskustva. Iako su suprotstavljene, obje se odnose na organizaciju i kategorizaciju iskustva. Funkcija razmišljanja uključuje svjesnu evaluaciju istinitosti iskustva. Evaluacija pomoću osjećanja je izražena u terminima kao što su

dopadanje/nedopadanje (engl. *like/dislike*), ugoda, stimulacija ili dosada. Jedan par funkcija je dominantan – racionalni ili iracionalni. Osoba ne može biti vođena istovremeno i razmišljanjem i osjećajima ili senzibilnošću i intuicijom budući da su to suprotstavljene funkcije.

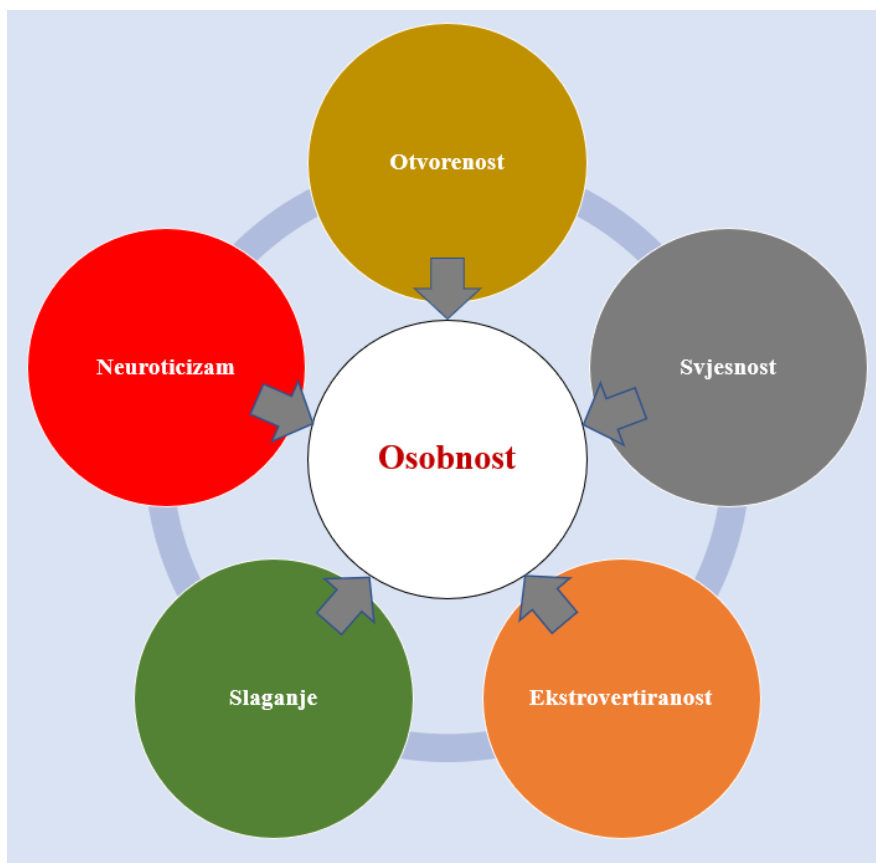
Jung je predložio osam psiholoških tipova, prema interakciji dva stava i četiri psihološke funkcije.

Procjena (engl. *assessment*) osobnosti podrazumijeva evaluaciju osobnosti uporabom pouzdanih i valjanih tehnika. Tehnika podrazumijeva način na koji se dolazi do podataka koji se mogu koristiti i u podatkovnoj znanosti, a uobičajene tehnike za procjene osobnosti su: samo-procjena, tehnike projekcije, klinički intervjui, procedure procjene ponašanja te procedure uzorkovanja mišljenja i iskustva ispitanika. Tehnike za procjenu osobnosti se smatraju pouzdanim ako rezultati koji se primjenom tehnika prikupljaju pokazuju konzistentnost vrijednosti unutar prihvatljive varijance prilikom ponavljanja testa. Valjanost tehnike podrazumijeva da je potvrđeno da odabrana tehnika mjeri karakteristike osobnosti za koje se koristi. Npr., ako se procjenjuje emocionalna inteligencija, valjanost označava da je potvrđeno da primijenjena tehnika doista mjeri emocionalnu inteligenciju [10] [16].

Upravo zbog važnosti praktične primjene procjene osobnosti te činjenice da su klasični postupci procjene i vremenski zahtjevni ima smisla razmisliti o primjeni tehnika podatkovne znanosti. S obzirom da psiholozi smatraju da je osobnost relativno stabilan i predvidljiv skup karakteristika koji se može promijeniti kao odgovor na različite situacije uporaba tehnika strojnog učenja kojima bi se prepoznali obrasci pojedinih tipova je pristup koji može olakšati i pojednostaviti postupak procjene osobnosti. Kao dodatni argument za primjenu tehnika strojnog učenja svakako treba uzeti u obzir činjenicu da su tradicionalni postupci spori i skupi jer zahtijevaju dugotrajniji angažman kvalificiranih stručnjaka.

Psiholozi su razvili pouzdane i valjane instrumente za procjenu osobnosti, a među najpoznatijim su MBTI, Big Five (OCEAN) [17], Enneagram [18] i DiSC [19]. U području predikcije tipa osobnosti autora teksta najveći broj istraživanja je proveden nad podacima koji su dobiveni primjenom instrumenata MBTI i Big Five. Kvantitativna usporedba broja istraživanja osobnosti autora iz teksta pokazuje da se veći broj istraživanja odnosi na Big Five instrument [20] [21]. Big Five model definira sljedeće dimenzije osobnosti: ekstrovertiranost (engl. *extroversion*), slaganje (engl. *agreeableness*), svjesnost (engl. *conscientiousness*), neuroticizam (engl. *neuroticism*), i otvorenost (engl. *openness*) [22]. Slika 1 prikazuje

dimenzije osobnosti prema Big Five modelu. Big Five model ima manju popularnost u poslovnom okruženju od modela MBTI (npr. u procesima zapošljavanja), i općenito je model MBTI široko popularan među laicima. Dodatno, Big Five model ne pruža mogućnost istraživanja složenih oznaka klasa s binarno isključivim vrijednostima, što je u fokusu ove disertacije.



Slika 1 Dimenzije osobnosti prema Big Five modelu

Ipak, Big Five je dotaknut zbog toga jer, uz model MBTI, istraživači najviše promatraju taj instrument te postoje istraživanja koja uspoređuju instrumente Big Five i MBTI i jer postoje istraživanja koja su pokazala određenu korelaciju dimenzija dva instrumenta. Disertacija stavlja fokus na instrument MBTI. Naglasak je isključivo na računalnom pristupu i ne razmatraju se psihološke analize instrumenta u predikciji autora teksta.

2.3. Model MBTI

Indikator tipa po Myers-Briggsovoj (MBTI) je instrument procjene osobnosti prema Jungovoj teoriji osobnosti [23] [24] [14]. Razvile su ga 1940-ih Katharine Cook Briggs i Isabel Briggs Myers i prvu put je publiciran 1962. Svrha upitnika je razlikovati pristupnike prema dvije

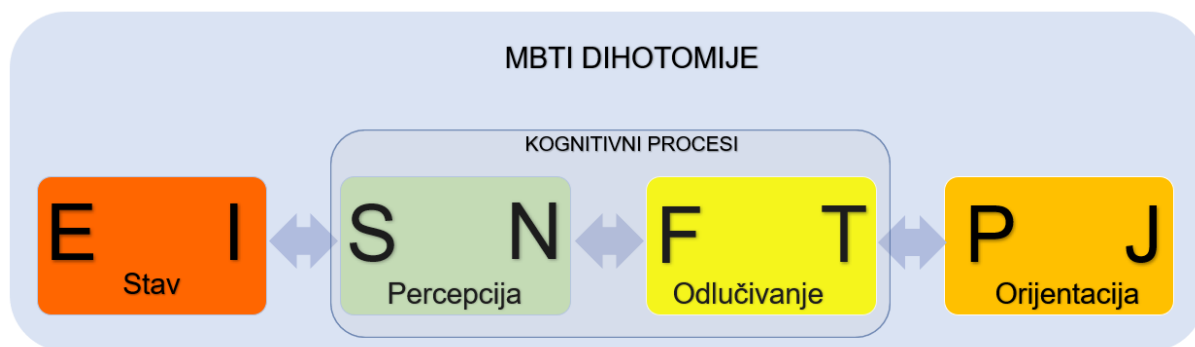
kategorije u svakoj od četiri ključne dimenzije. Svake godine testu prema instrumentu MBTI, u svrhu zapošljavanja i poslovne promocije, pristupi više od dva milijuna ljudi [13].

Među psiholozima postoje dvojbe s obzirom na instrument MBTI [25] [26], budući da postoje prigovori da mu nedostaje stabilnost i mjera neuroticizma kao karakteristike osobnosti. Dodatno, postoje studije koje potvrđuju korelaciju između modela MBTI i modela Big Five, gdje dimenzije ekstrovertiranosti jako koreliraju, a dihotomija J/P korelira s dimenzijom svjesnosti. Također, studije pokazuju da su komponente modela MBTI složenije za predikciju od komponenti modela Big Five [27]. Istraživanje [28] pokazuje kako je moguće ostvariti bolje performanse s algoritmima treniranim na MBTI nego Big Five i da Big Five daje više informacija i značajnu varijabilnost ovisno o korištenom algoritmu.

Jung je grupirao kognitivne funkcije prema načinu percepcije (engl. *perception*) i procjene (engl. *judgment*) jer ih je smatrao glavnim tipovima mentalnih procesa. Percepcija se odnosi na način prikupljanja ili pristupa informacijama. Prema tome odnosi li se percepcija na svijest o opipljivoj (engl. *tangible*) informaciji ili konceptualnoj informaciji (engl. *conceptual*), Jung je podijelio tipove percepcije na osjetilnu (engl. *sensation*) i intuitivnu (engl. *intuition*), koji mogu biti dio ekstrovertiranog i introvertiranog stava. Procjena je vrsta mentalnog procesa koji se odnosi na organizaciju, evaluaciju, i zaključivanje. Procjena se može podijeliti na dva tipa: razmišljanje (engl. *thinking*), koje je temeljeno na objektivnim kriterijima ili principima, i osjećanje (engl. *feeling*) koje je temeljeno na osobnim, interpersonalnim, ili univerzalnim vrijednostima. Funkcije razmišljanja i osjećanja također mogu biti dio ekstrovertiranog i introvertiranog stava.

U opisivanju kognitivnih procesa Jung je opisao svih osam tipova osobnosti, od kojih je svaki osoban prema dominantnosti jedne od četiri funkcije. Jung je definirao funkcije kao primjer dihotomije suprotnosti, što je značajna činjenica za istraživanje u ovom radu. To znači da je npr. procjena na osnovu vrijednosti i kriterija u isto vrijeme nemoguća. Naravno, u različitim situacijama suprotstavljene funkcije mogu biti zastupljene, ali ne u isto vrijeme. Npr., osoba može uzeti u obzir različite kriterije za procjenu, a nakon toga može uzeti u obzir vrijednosti kod procjene važnosti. Jung je razvio ideju hijerarhije funkcija budući da osoba ne može imati podjednako razvijenu svaku od osam funkcija nego postoji dominantna funkcija, kao najrazvijenija, te pomoćna koja je potpora dominantnoj funkciji na način da su stavovi ili orijentacije glavne i pomoćne funkcije različite [29]. Npr. ako je dominantni proces unutarnji svijet (introvertirani stav), tada je pomoćni proces fokusiran na vanjski svijet (ekstrovertirani

stav). Dodatno, vrsta procesa, procjena ili percepcija su različite. Ako je dominantni proces percepcija, tada je pomoćni proces procjena, i obratno. Dihotomije su E-I, S-N, T-F i J-P. Dihotomija Judging-Perceiving je razvijena ne da bi se ljudima ili ljudskom ponašanju postavljale oznake nego pomogne pristupu dinamikama kognitivnih procesa. Na model MBTI se često gleda kao alat za 'klasificiranje' ljudi te se zanemaruju dinamični i razvojni aspekti modela. Slika 2 prikazuje način konstruiranja kôda tipa osobnosti, prema modelu MBTI, pomoću četiri oznake, zajedno s procesom koji definira svaka od dihotomija [30].



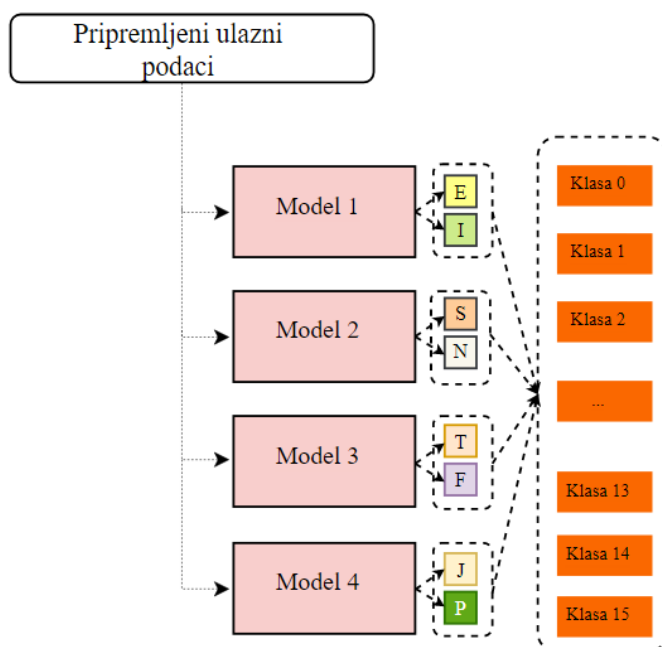
Slika 2 Model MBTI osobnosti

Analizirajući dublje dimenzije modela MBTI, prva dimenzija je Ekstrovert (E) ili Introvert (I). Dimenzija ukazuje da je osoba okrenuta prema van, razgovorljiva ili rezervirana. Drugim riječima, dimenzija definira kakva je orijentacija osobe, prema vanjskom svijetu ili unutarnjem svijetu te čini primarnu energetske motivaciju. Druga dimenzija je osjećajnost (S) ili intuicija (N). Ova dimenzija definira kako osoba prima informacije. Npr., osoba s više osjetilnim pristupom obrađuje više činjenica, dok osoba s više intuitivnim pristupom pokušava interpretirati informacije i naći dublje značenje. Treća dimenzija je mišljenje (T) i osjećanje (F). Ova dimenzija opisuje kako osoba donosi odluke. Npr., osoba s pristupom koji uključuje primarno mišljenje koristi logiku i konzistentnost u rezoniranju i donošenju odluka, dok osoba s više osjećajnim pristupom koristi empatiju i fokus na ljude i specifične okolnosti. Posljednja dimenzija je prosuđivanje (J) i percepcija (P). Ova dimenzija opisuje osobnu orijentaciju prema vanjskom svijetu i kako osoba provodi svakodnevicu, drugim riječima, opisuje osobni životni stil. Npr., osoba koja ima sklonost prosuđivanja teži organiziranom svakodnevnom životu, dok osoba sklona percepciji preferira fleksibilnost. To nas vodi do 16 mogućih kombinacija tipova osobnosti modela MBTI. Budući da je svaka klasa sačinjena od 4 komponente, očito je da se radi o složenim oznakama klasa. Npr., osoba koja preferira osamljenost (I), vjeruje svojoj intuiciji u procjeni i interpretiranju informacija (N), koristi

logiku u rezoniranju (T) i preferira spontani svakodnevni život (P) bi većinski pripadala tipu INTP modela MBTI.

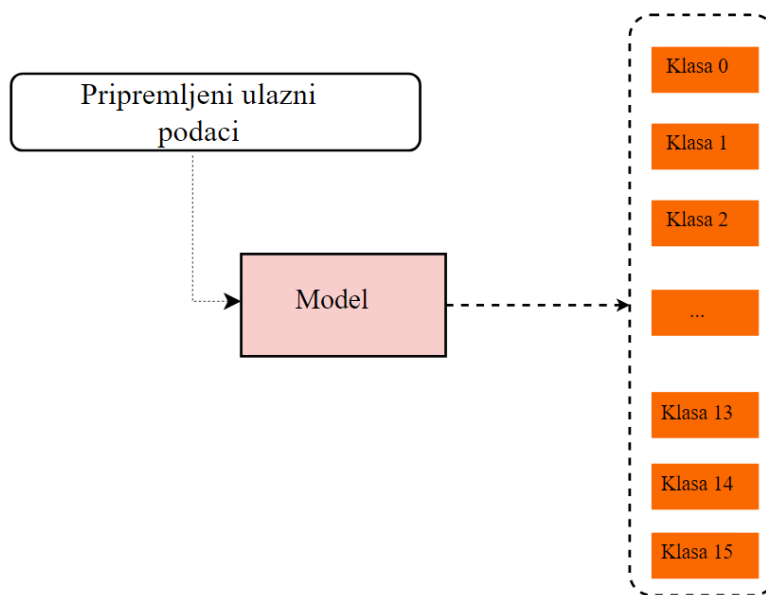
Problem klasifikacije osobnosti prema modelu MBTI se uglavnom promatra kao problem četverostruke binarne klasifikacije ili problem višeklasne klasifikacije složenih oznaka klasa modela MBTI.

Problem binarne klasifikacije prikazuje slika 3. Modeli dubokog učenja nastoje, iz pripremljenih ulazni podataka, klasificirati elemente pojedinih dihotomija složene oznake MBTI, a konačni rezultat klasifikacije prema složenim oznakama klasa se dobiva udruživanjem pojedinačnih rezultata.



Slika 3 MBTI – pristup binarne klasifikacije

Problem višeklasne klasifikacije demonstrira slika 4. Model dubokog učenja nastoji klasificirati pripremljene ulazne podatke prema 16 složenih oznaka klasa MBTI.



Slika 4 MBTI - pristup višeklasne klasifikacije

Prednost pristupa s četiri binarna klasifikatora je u činjenici da su pod skupovi koji se koriste za binarnu klasifikaciju bolje balansirani od skupa sa svih 16 složenih klasa. Dodatna prednost je da se takvim pristupom dobiva na fleksibilnosti jer se mogu kombinirati različiti klasifikatori za pojedine binarne klasifikacije. Nedostatak pristupa s četiri binarna klasifikatora je u činjenici da se pojedinim binarnim klasifikacijama ne može uzeti u obzir i utjecaj ostalih komponenti pojedine složene klase. Očiti nedostatak pristupa višeklasne klasifikacije je u činjenici da je skup izrazito nebalansiran, što je osnovni razlog loših rezultata višeklasne klasifikacije. U disertaciji je ovaj problem promatran kao višeklasna klasifikacija sa složenim oznakama klasa, s namjerom da se demonstrira način kako se u višeklasnoj klasifikaciji može uključiti utjecaj pojedinačnih komponenti pogrešno klasificiranih složenih klasa.

2.4. Strojno učenje

Metode strojnog učenja (engl. *machine learning*) nastoje, bez eksplicitnog programiranja, naučiti odnose između podataka u odabranom skupu podataka. Primjena metoda strojnog učenja podrazumijeva proces u kojem računalni model nastoji poboljšati svoje performanse stalnim ugrađivanjem novih informacija u okviri primijenjenog statističkog modela.

U podjeli strojnog učenja uobičajeno se ističu nenadzirano učenje (engl. *unsupervised learning*), podržano/ojačano učenje (engl. *reinforcement learning*) i nadzirano učenje (engl.

supervised learning). Nenadzirano strojno učenje nastoji pronaći predloške unutar podataka te ih prema predlošcima grupirati. Pojačano strojno učenje podrazumijeva metode strojnog učenja koje se temelje na postupcima nagrađivanja željenog ponašanja algoritma te kažnjavanja ponašanja koja nisu poželjna.

Nadzirano strojno učenje podrazumijeva pristupe u kojima se postupak učenja nadzire na način da se koriste skupovi podataka s poznatim vrijednostima, tj. ciljanim kategorijama koje model treba naučiti u zadacima klasifikacije i vrijednostima koje model strojnog učenja treba što bolje interpolirati u zadacima regresije. Ulazni skup sadrži poznate činjenice, a izlazni skup ciljane informacije. Postupcima nadziranog učenja se mogu rješavati dvije osnovne vrste problema: klasifikacija i regresija. U postupcima klasifikacije primjeru pridružujemo diskretnu ili nominalnu varijablu koja predstavlja klasu (razred) kojoj primjer pripada. Kod postupaka regresije ciljna varijabla je kontinuirana budući da se primjeru pridružuje kontinuirana vrijednost. Problem određivanja tipa osobnosti prema instrumentu MBTI se razmatra kao klasifikacijski problem.

Zadatak klasifikacije je da, za primjer $\mathbf{x}$, koji se definira kao vektor značajki $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)^T$, a pripada skupu svih mogućih primjera $\mathcal{X}$, odredi pripadajuću klasu $\mathcal{C}$. Dimenzija vektora $\mathbf{x}$ je određena s n . Svaki primjer se može interpretirati kao točka u n -dimenzijskom vektorskom prostoru koji nazivamo prostorom primjera (engl. *instance space*) ili ulaznim prostorom (engl. *input space*). Izuzetno je bitna pretpostavka algoritama strojnog učenja da su primjeri iz $\mathcal{X}$ uzorkovani nezavisno te iz iste zajedničke distribucije $P(\mathbf{x}, \mathbf{y})$. Simbolom $\mathbf{y}$ definiramo oznaku klase kojoj pripada primjer $\mathbf{x}$ iz $\mathcal{X}$, a skup svih N primjera za učenje s $\mathcal{D} = \{(\mathbf{x}^{(i)}), \mathbf{y}^{(i)})\}, i \in (1, \dots, N)$.

Cilj višeklasnog klasifikacijskog algoritma je naučiti hipotezu $h: \mathcal{X} \rightarrow \{0, 1, \dots, K-1\}$, pri čemu se oznaka klase primjera $\mathbf{x}^{(i)}$ može prikazati kao K -dimenzijski vektor, gdje je K ukupan broj klasa kojima primjeri mogu pripadati. Indeksne vrijednosti oznaka se mogu označiti s $j = 0, 1, 2, \dots, K-1$ te vrijedi $\mathbf{y}^{(i)} = (\mathbf{y}_0^{(i)}, \mathbf{y}_1^{(i)}, \dots, \mathbf{y}_{K-1}^{(i)})^T$ gdje je

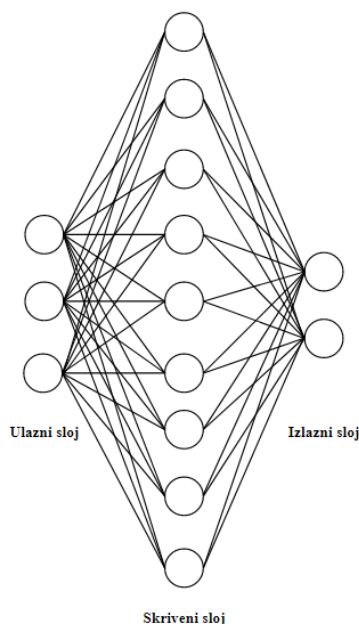
$$y_j^{(i)} = \begin{cases} 1, & \text{ako } \mathbf{x}^{(i)} \in C_j \\ 0, & \text{inače} \end{cases} \quad (1)$$

U primjeru klasifikacije MBTI vrijedi $y \in \{0, 1, 2, \dots, 15\}$ te će svaki primjer x biti označen s jednom od 16 mogućih vrijednosti. Npr., za skup podataka MBTI bi mogla postojati mogućnost da za $y^{(0)} = (1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)^T$ i primjer $x^{(0)}$ primjer pripada klasi C_1 te bi ta klasa, u ovisnosti o odabranom redoslijedu složenih oznaka klasa MBTI, mogla označavati klasu ENFJ. Stoga, klasifikacija MBTI spada u grupu višeklasnih klasifikacija. Za razliku od navedenog primjera, u slučajevima kada klasifikator uči kako pridružiti svaki primjer jednoj od dvije moguće klase, govorimo o binarnom klasifikatoru.

Učenje modela predstavlja traženje najbolje hipoteze h iz prostora hipoteza H na skupu označenih primjera D . Mjeru kvalitete hipoteze h predstavlja funkcija empirijske pogreške ili pogreška učenja (engl. *training error*), označava se s $E(h|D)$. Empirijska pogreška predstavlja numeričku procjenu kvalitete hipoteze h za zadatke klasifikacije kvantificira koliko hipoteza griješi kada klasificira primjere a za zadatke regresije koliko su vrijednosti daleko od ciljnih vrijednosti može se predstaviti izrazom (2). L predstavlja funkciju gubitka (engl. *loss function*) koja kvantificira u obliku nenegativnog realnog broja koliko je model za svaki označeni primjer izgubio na točnosti (tj. dobio na ukupnoj pogrešci). Također, važno je za uspješnost klasifikacije da model ima dovoljno kapaciteta za učenje klase C [31].

$$E(h|D) = \frac{1}{N} \sum_{i=1}^N L(y^{(i)}, h(x^{(i)})) \quad (2)$$

Duboko učenje (engl. *deep learning*) obuhvaća podskup metoda strojnog učenja. Duboko učenje podrazumijeva uporabu umjetnih neuronskih mreža (engl. *artificial neural networks*) koji su kategorija algoritama strojnog učenja čiji razvoj je inspiriran željom za oponašanjem rada životinjskog neurona [32]. Slika 5 prikazuje primjer arhitekture jednostavne višeslojne unaprijedne umjetne neuronske mreže s ulaznim slojem koji čine tri neurona, izlaznim slojem s dva neurona te jednim skrivenim slojem s devet neurona. Ulazni sloj prima podatke i prenosi informacije direktno skrivenom sloju. Potom, svaki neuron u skrivenom sloju izvršava izračun težinskom linearnom kombinacijom te prenosi informaciju prema sljedećem sloju nakon primjene aktivacijske funkcije. Proces se ponavlja do izlaznog sloja. Neuronska mreža se može smatrati nelinearnim klasifikatorom koji dijeli ulazni vektorski prostor primjera (ulaznih podataka) u klase koje imaju nelinearne granice.



Slika 5 Primjer umjetne neuronske mreže

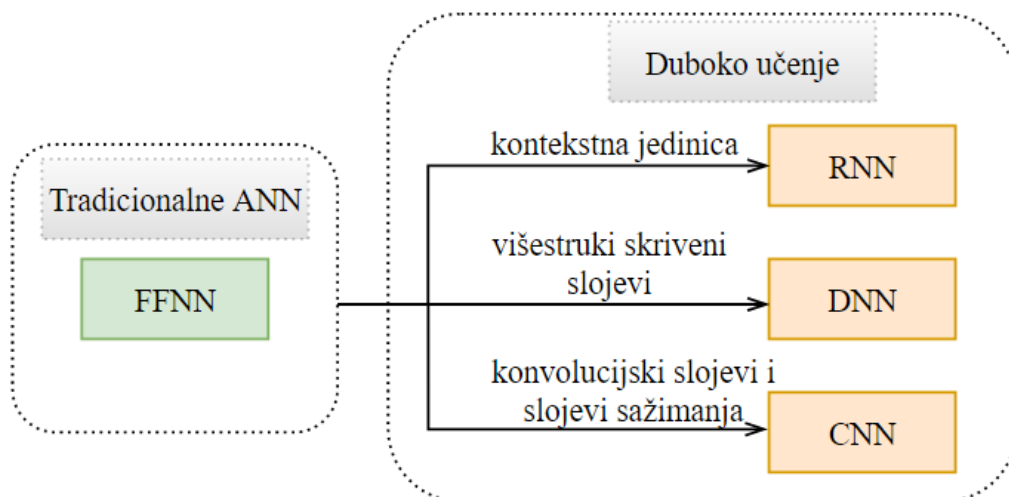
Učenje modela dubokog učenja predstavlja proces procjene parametara unutar umjetne neuronske mreže. Vrlo važnu ulogu u postupku nadziranog učenja ima funkcija gubitka L koja za parametre modela θ izračunava razliku ciljne vrijednosti $y^{(i)}$ i aproksimacije modela $h(x^{(i)} | \theta)$ te predstavlja i način postavljanja prioriteta jer usmjerava korake postupka treniranja koje je potrebno napraviti kako bi model greške treniranja minimizirao ažuriranjem vrijednosti parametara. Izraz (3) daje uobičajenu definiciju funkcije gubitka u zadacima klasifikacije kao negativnu (log-)izglednost [33].

$$L(y^{(i)}, h(x^{(i)} | \theta)) = - \sum_i \log p(h(x^{(i)} | \theta) = y^{(i)} | \theta) \quad (3)$$

Pored modela i funkcije gubitka, značajno mjesto u postupcima nadziranog učenja zauzima i optimizacijski postupak koji definira način pronalaska vrijednosti θ parametara modela za koje je empirijska pogreška (očekivanje funkcije gubitka primjere x iz $\mathcal{D}$) najmanja.

Uobičajena podjela umjetnih neuronskih mreža je na tradicionalne umjetne neuronske mreže, kao što je unaprijedna neuronska mreža (engl. *feed-forward neural network* - *FFNN*) koju prikazuje slika 5, jer sadrži jedan skriveni sloj te duboke neuronske mreže (engl. *deep neural networks*) koje imaju više skrivenih slojeva. Slika 6 prikazuje podjelu na tradicionalne

umjetne neuronske mreže i duboke neuronske mreže. U okviru dubokih neuronskih mreža uobičajeno razlikujemo mreže koje sadrže kontekstnu jedinicu, kao što su rekurentne neuronske mreže (eng. *recurrent neural networks* – *RNN*), mreže s višestrukim skrivenim slojevima (engl. *deep neural networks* - *DNN*) te duboke neuronske mreže s konvolucijskim slojevima i slojevima sažimanja (eng. *convolutional neural networks* – *CNN*).



Slika 6 Duboko učenje i tradicionalne umjetne neuronske mreže

U praktičnom dijelu disertacije su korištena dvije arhitekture dubokog učenja koje do sada prevladavaju u odabiru arhitektura za obradu prirodnog jezika: LSTM i CNN.

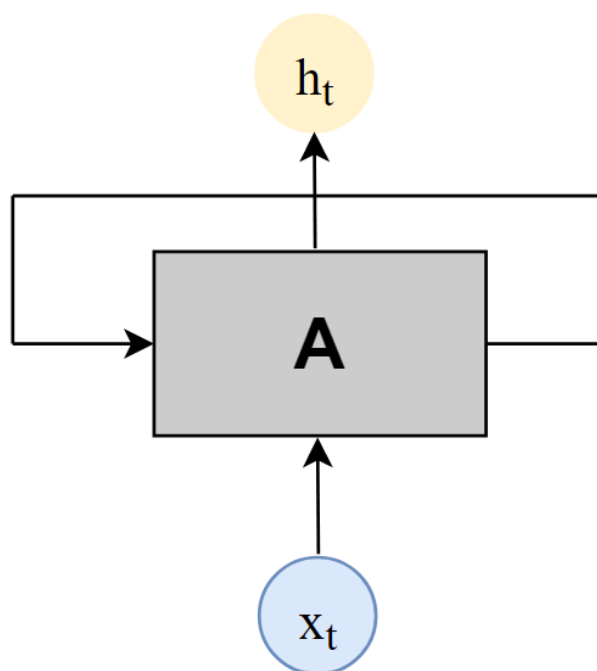
2.4.1. LSTM

Duga kratkoročna memorija (engl. *Long Short-Term Memory* - *LSTM*) je vrsta rekurentnih neuronskih mreža (engl. *Recurrent Neural Network* - *RNN*), osmišljena kako bi se riješio problem nestajućeg gradijenta (engl. *vanishing gradient*) ili eksplodirajućeg gradijenta (engl. *exploding gradient*) [34]. Postoje mnoge nadogradnje rekurentnih mreža, kao što su npr. modeli bazirani na pozornosti (engl. *attention models*) i rekurentna neuronska mreža slijed-u-slijed (engl. *sequence-to-sequence*).

Slika 7 prikazuje rekurentnu neuronsku mrežu. Element neuronske mreže je označen s A , input s x_t , a izlaz s h_t . Petlja omogućava da se informacije prenose kroz elemente neuronske mreže te je na taj način omogućena perzistencija informacije kroz mrežu. Mreže RNN mogu pratiti privremene ovisnosti kroz vrijeme. Kako trenutni output ovisi o novim ulaznim vrijednostima i o prethodnim izlaznim vrijednostima, karakteristika se može promatrati kao oblik memorije jer postoji ovisnost kroz vrijeme i trenutni ulaz ovisi također o ulazima u

prošlosti. Modeliranje privremenih ovisnosti je značajno za predikciju modela jer tekst, govor i slike imaju vremenski promjenljive karakteristike te imaju ovisnosti kroz vrijeme.

Stoga se mreže RNN koriste za obradu sekvencijalnih podataka, kao što je tekst. Mreže RNN u teoriji mogu upravljati dugoročnim ovisnostima (engl. *long-term dependencies*), no u praksi mreže RNN imaju problema s dugim nizom ovisnosti [35] [36].



Slika 7 Rekurentna neuronska mreža

Mreže LSTM mogu učiti dugoročne ovisnosti te poboljšavaju rekurentne mreže na način da zamjenjuju neuron u neuronskim mrežama složenijom jedinicom za obradu podataka te implementiraju pristup pod nazivom vrtuljak konstantne pogreške (engl. *constant error carousel* – CEC) [37], koji je jedan od pristupa za ublažavanje problema s gradijentom. Uobičajeni naziv za LSTM jedinicu obrade je ćelija (engl. *cell*) te predstavlja napredni koncept jedinice s vratima (engl. *gated unit*).

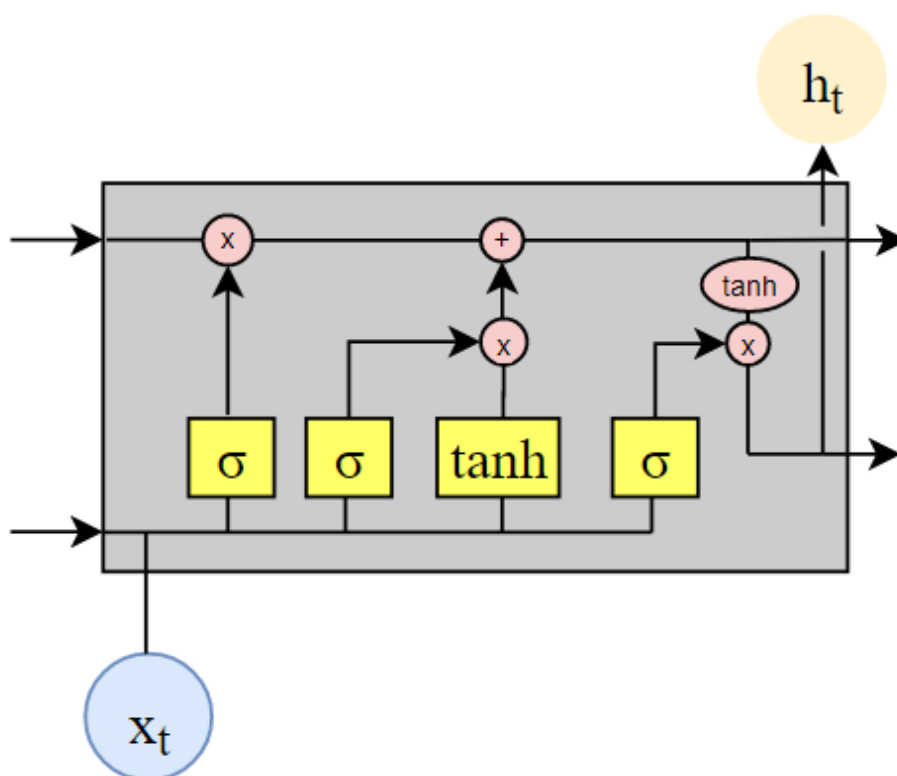
Tablica 1 daje pregled tehnika za ublažavanje s nestajućim gradijentom ili gradijentom koji poprima ekstremno velike vrijednosti. Vrtuljak konstantne pogreške se primjenjuje samo na rekurentnim slojevima i koristi se unutar mreže LSTM. Nezasićenje neurona kao ReLU (engl. *no saturating neurons*) se primjenjuje na svim neuronima osim onima iz izlaznog sloja mreže. Grupna normalizacija (engl. *batch normalization*) [38] se primjenjuje na skrivenim

neuronima. Glorot ili He inicijalizacija težina [39] se primjenjuje na svim neuronima. Isječak gradijenta (engl. *gradient clipping*) se primjenjuje na svim neuronima.

Tablica 1 Pregled tehnika za ublažavanje problema s gradijentima

Pristup	Ublažava ekstremne vrijednosti	Ublažava nestajući gradijent
Vrtuljak konstantne pogreške	Da	Da
Nezasićenje neurona	Ne	Da
Grupna normalizacija	Ne	Da
Glorot ili He inicijalizacija	Ne	Da
Isječak gradijenta	Da	Ne
Preskakanje poveznica	Ne	Da

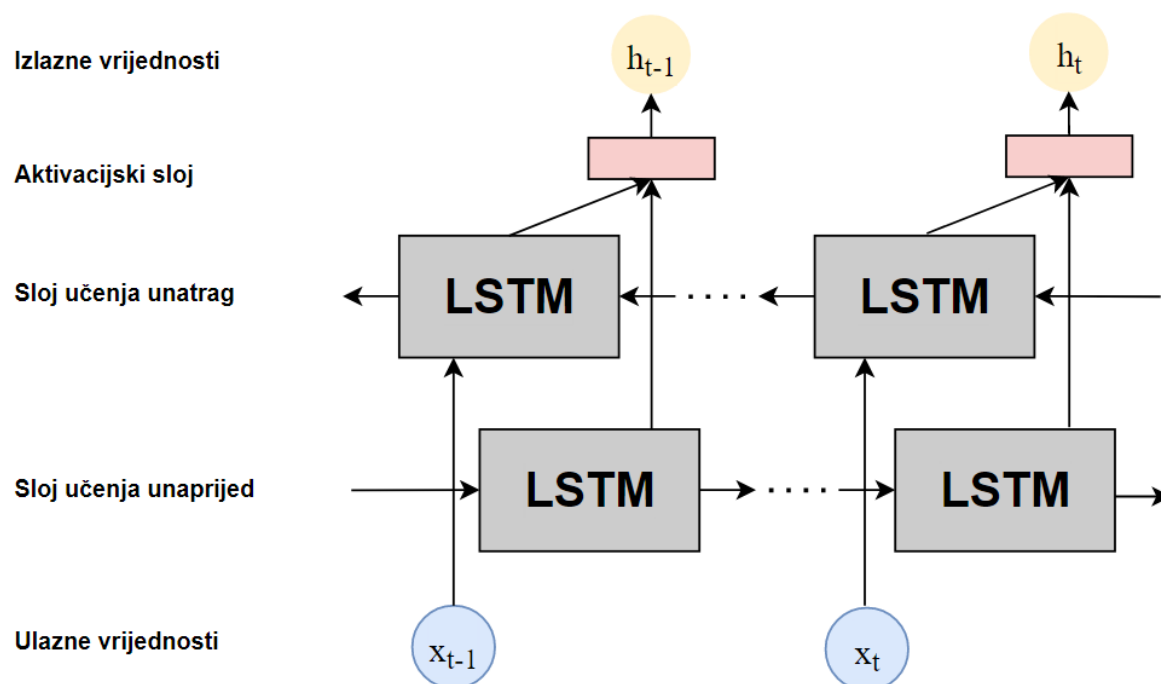
LSTM ćelija koristi tri implementacije logističke sigmoidne funkcije kao vrata u ćeliji te dvije implementacije drugih aktivacijskih funkcija, a najčešće $\tanh$ i ReLU . Ćelija također sadrži četiri težinske sume, za razliku od jedne u klasičnim mrežama RNN. Slika 8 prikazuje ćeliju neuronske mreže LSTM. Mreža LSTM može ukloniti ili dodati informaciju u stanje ćelije strukturom koja se nazivaju vrata (engl. *gates*). Vrata su mehanizam propuštanja informacija kroz ćeliju. Vrata se sastoje od sloja neuronske mreže sa sigmoidnom funkcijom koja kao izlaz daje vrijednost između 0 i 1, čime se određuje količina propuštene informacije i operacije množenja.



Slika 8 Čelija LSTM neuronske mreže

Važna karakteristika mreže LSTM je da pojedini signali, koji imaju ime varijable stanja, mogu imati uporabom vrata fiksne vrijednosti do trenutka kada se ponovno uključe u obradu podataka u mreži. To je način implementacije vremenskog intervala i praćenja vremenske ovisnosti. Varijante mreže LSTM, kao što je GRU (engl. *Gated Recurrent Network*) dodatno poboljšavaju mreže LSTM.

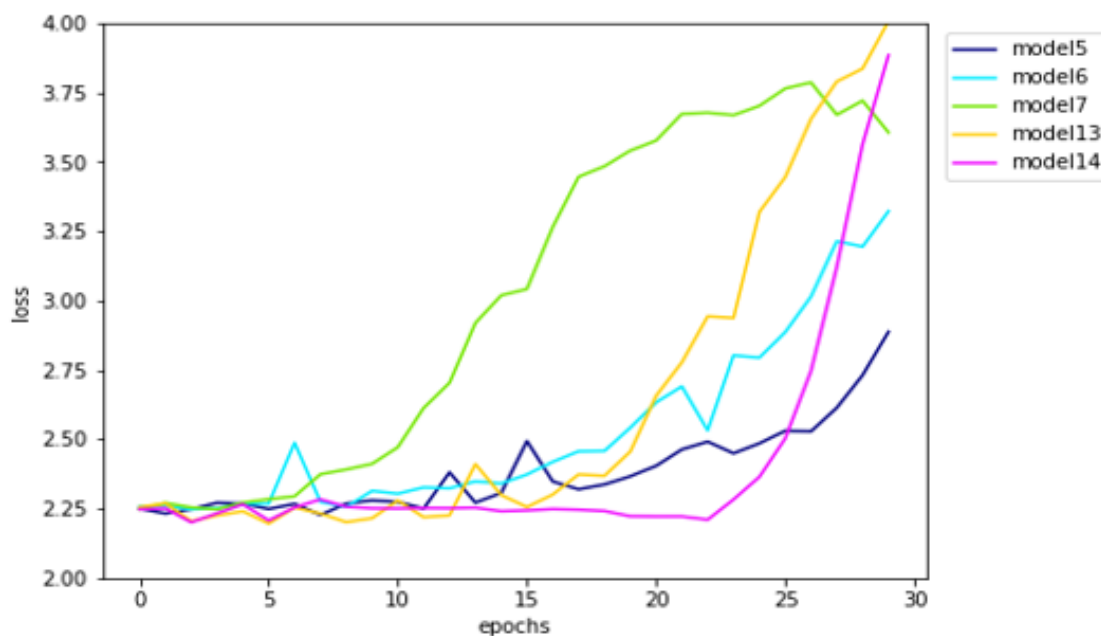
Dvosmjerna duga kratkoročna memorija (engl. *Bi-directional long short-term memory* - *Bi-LSTM*) obrađuje informacije u dva smjera mreže LSTM. Mreža Bi-LSTM je verzija dvosmjerne rekurentne neuronske mreže [40]. Dva smjera predstavljaju tijek od prošlosti prema budućnosti te od budućnosti prema prošlosti. Dvosmjerna mreža LSTM se obično koristi kod sekvencijalnih zadataka te su takve mreže opcija arhitekture npr. za tekstualnu klasifikaciju ili prepoznavanje govora.



Slika 9 Mreža bi-LSTM

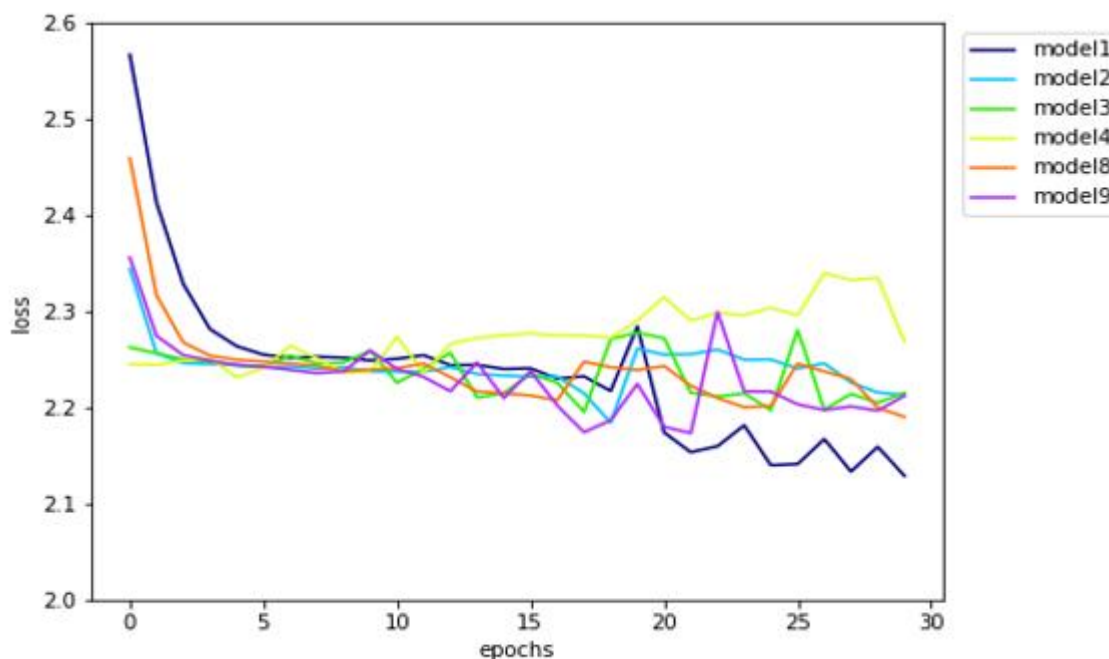
Budući da postoji mnoštvo hiperparametara koji mogu utjecati na rezultate klasifikacije LSTM mreže, analiza utjecaja hiperparametara na rezultate treniranja modela može dati smjernice za odabir hiperparametara koji će osigurati bolje rezultate neuronske mreže. Strukturalni hiperparametri se definiraju prije treniranja, a istraživanje [41] promatra broj skrivenih slojeva i broj neurona u skrivenim slojevima LSTM ćelija kao tip strukturalnih parametara. Eksperimenti klasifikacije tipa osobnosti MBTI promjenom strukturalnih parametara su pokazali da mreže LSTM s manjim brojem neurona i slojeva daju najbolje F1 rezultate, dok modeli s većim brojem neurona imaju tendenciju prenaučivosti (engl. *overfitting*). Također, modeli s više slojeva i rastućim brojem neurona pokazuju bolje performanse ako je ciljna metrika točnost. Kada se promatraju srednja makro-preciznost, odziv i F1-mjera, mreže s 3 sloja daju slične ili nešto lošije rezultate od mreža s 1 ili 2 sloja. Kod analize broja neurona po sloju, u mrežama LSTM s jednim slojem, manji broj neurona daje bolje rezultate. U biti, modeli s 100, 150 i 500 neurona po sloju pokazuju znakove prenaučivosti. Slika 10, slika 11, slika 12 i slika 13 pokazuju dijagrame gubitka na validacijskom skupu za funkciju gubitka unakrsne entropije. Npr., moguće je uočiti da za sve modele koje pokazuje slika 10, rast gubitka na validacijskom skupu, što je očiti pokazatelj da navedene arhitekture negativno utječu na kvalitetu klasifikacije. Također, npr. model 10 čije ponašanje pokazuje slika 12 pokazuje bolje rezultate gubitka na validacijskom skupu.

Tablica 4 daje pregled rezultata eksperimenata sa strukturnim hiperparametrima mreže LSTM pri višeklasnoj klasifikaciji MBTI. Ti su rezultati bili smjernica u odabiru arhitekture za eksperimentalni dio istraživanja u disertaciji te su za arhitekturu odabrana dva skrivena sloja s manjim broj neurona.



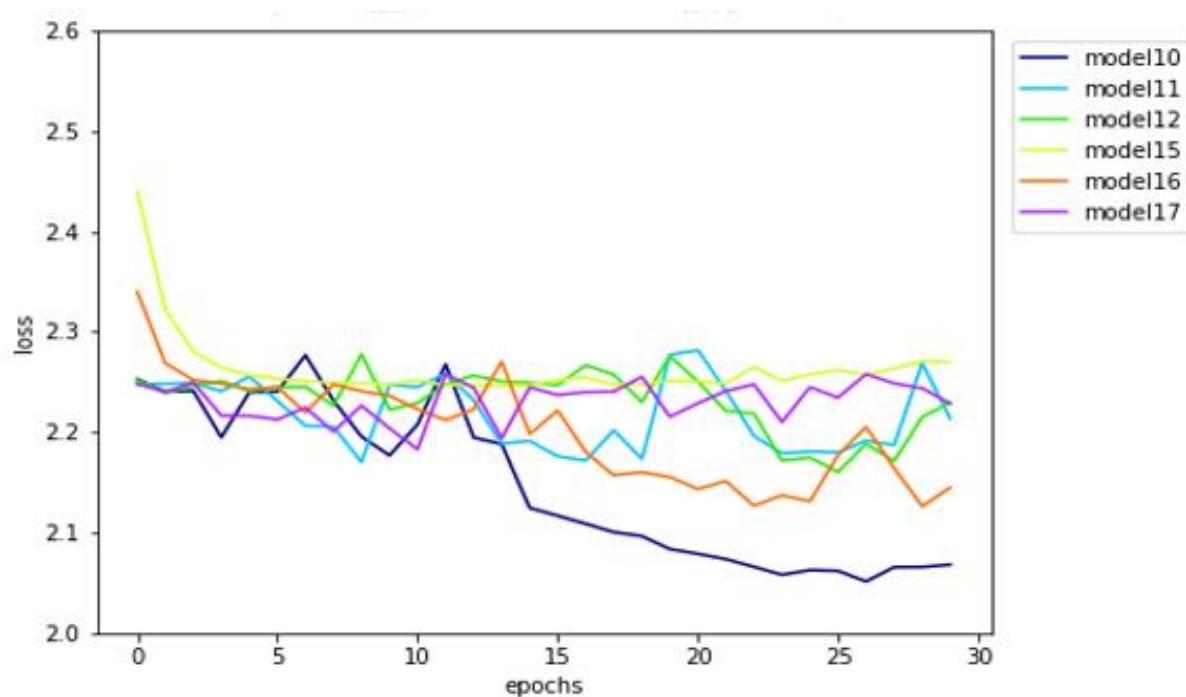
a) Validation loss of models 5, 6, 7, 13, 14

Slika 10 Gubitak validacije za funkciju gubitka unakrsne entropije [33] a)



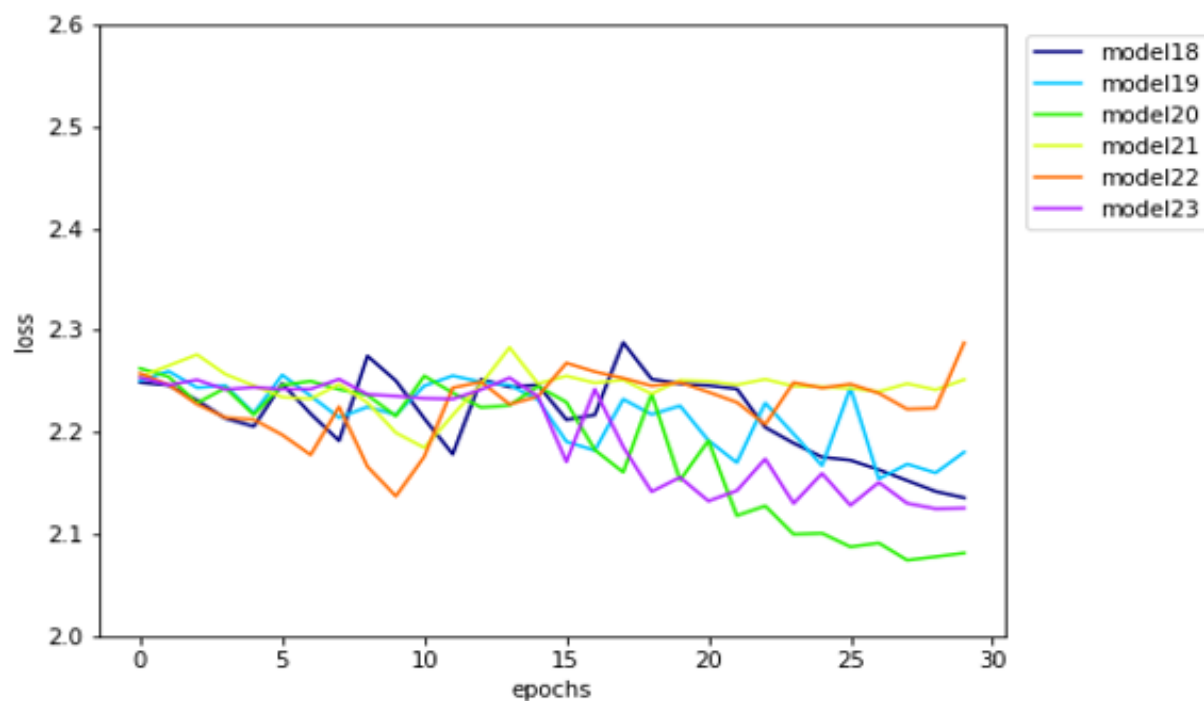
b) Validation loss of models 1, 2, 3, 4, 8, 9

Slika 11 Gubitak validacije za funkciju gubitka unakrsne entropije [33] b)



c) Validation loss of models 10, 11, 12, 15, 16, 17

Slika 12 Gubitak validacije za funkciju gubitka unakrsne entropije [33] c)



d) Validation loss of models 18, 19, 20, 21, 22, 23

Slika 13 Gubitak validacije za funkciju gubitka unakrsne entropije [33] d)

Dodatno, istraživanje [42] ukazuje na tehnike uravnoteženja skupa podataka MBTI koje bi mogle pomoći u ostvarenju boljih rezultata s modelom neuronske mreže LSTM, pri čemu je metoda slučajnog višestrukog naduzorkovanja (engl. *random oversampling*) pokazala najbolje rezultate u poboljšanju svih metrika, a osobito u poboljšanju F1 – mjere.

Klasifikacija nad neuravnoteženim skupom podataka predstavlja važan problem u strojnom učenju jer neuravnoteženost skupa podataka utječe na rezultate klasifikacije, ukupnu točnost i odluke modela pri klasifikaciji, budući da standardni modeli klasifikacije pretpostavljaju ravnomjernu distribuciju klasa te imaju sklonost prema negativnim klasama, što ima za posljedicu pogrešnu klasifikaciju primjera pozitivnih klasa (ciljanih primjera).

Tablica 2 daje pregled tehnika uravnoteženja za višeklasnu klasifikaciju MBTI te ostvarenih rezultata za pojedine pristupe [43] [44]. Kappa mjera ukazuje kakve su performanse klasifikatora u odnosu na klasifikator koji bi slučajno pogađao predviđenu vrijednost prema frekvenciji svake klase. Kappa vrijednosti mogu biti u rasponu od 0 do 1, a klasifikator je bolji ako su vrijednosti bliže vrijednosti 1 [45]. Modeli strojnog učenja za koje su upotrijebljeni težinski faktori, kako bi se slabije zastupljenim klasama dao veći značaj prilikom treniranja, imaju oznaku „W“.

Tablica 2 Primjena tehnika uravnoteženja za višeklasnu klasifikaciju MBTI [42]

Pristup	Preciznost	Odziv	F-mjera	Podrška	Kappa
Base	0.023	0.062	0.033	64.062	-0.0023
SMOTE [46]	0.155	0.158	0.144	216.625	0.107
Borderline-SMOTE [47]	0.148	0.190	0.144	190.188	0.159
ROS [48]	0.271	0.345	0.287	216.625	0.305
RUS [48]	0.028	-0.109	0.038	4.812	0.004
SVC+W	0.121	0.068	0.069	64.062	0.003
SVC	0.035	0.066	0.040	64.062	0.016
DTC	0.049	0.047	0.048	64.062	-0.018
DTC+W	0.060	0.058	0.059	64.062	0.005
Logreg	0.125	0.069	0.064	64.062	-0.001
Logreg+W	0.062	0.052	0.035	64.062	-0.001
Ridge	0.041	0.062	0.046	64.062	0.001
Ridge+W	0.041	0.044	0.016	64.062	-0.007
MNB	0.061	0.058	0.047	64.062	-0.005
RFC	0.051	0.059	0.043	64.062	-0.008
RFC+W	0.068	0.063	0.045	64.062	0.005
XGBOOST	0.055	0.068	0.050	64.062	0.024

Pristup	Preciznost	Odziv	F-mjera	Podrška	Kappa
Bag	0.055	0.058	0.052	64.062	-0.008
Calibration SVC	0.035	0.063	0.028	64.062	0.004
Calibration RFC	0.033	0.063	0.026	64.062	0.003

Problem neuravnoteženosti skupova podataka se rješava različitim pristupima koji imaju prednosti i nedostatke, kao što pokazuje tablica 3.

Jedan od pristupa se primjenjuje na razini podataka i odnosi se na različite pristupe uzorkovanju podataka. Pristupi se uglavnom primjenjuju u trenutku predobrade podataka s idejom da se ili smanji broj podataka klasa koje su najzastupljenije u neuravnoteženom skupu podataka, što je poznato kao pristup poduzorkovanja (engl. *under-sampling*) ili da se poveća broj podataka klasa koje su najmanje zastupljene u neuravnoteženom skupu podataka, što je poznato kao pristup naduzorkovanja (engl. *over-sampling*). Za razliku od pristupa poduzorkovanja koji se uglavnom oslanjaju na tehnike grupiranja, među pristupima naduzorkovanja su popularni pristupi generiranja umjetnih primjera kojima je osnova metoda naduzorkovanja umjetnih manjinskih primjera (engl. *synthetic minority oversampling technique – SMOTE*).

Također, moguće je pristupiti uravnoteženju podataka postupcima koji se fokusiraju na pristupe učenja koji su troškovno osjetljivi. Takvi pristupi koriste modele učenja koji manjinskim klasama daju veći značaj prilikom klasifikacije.

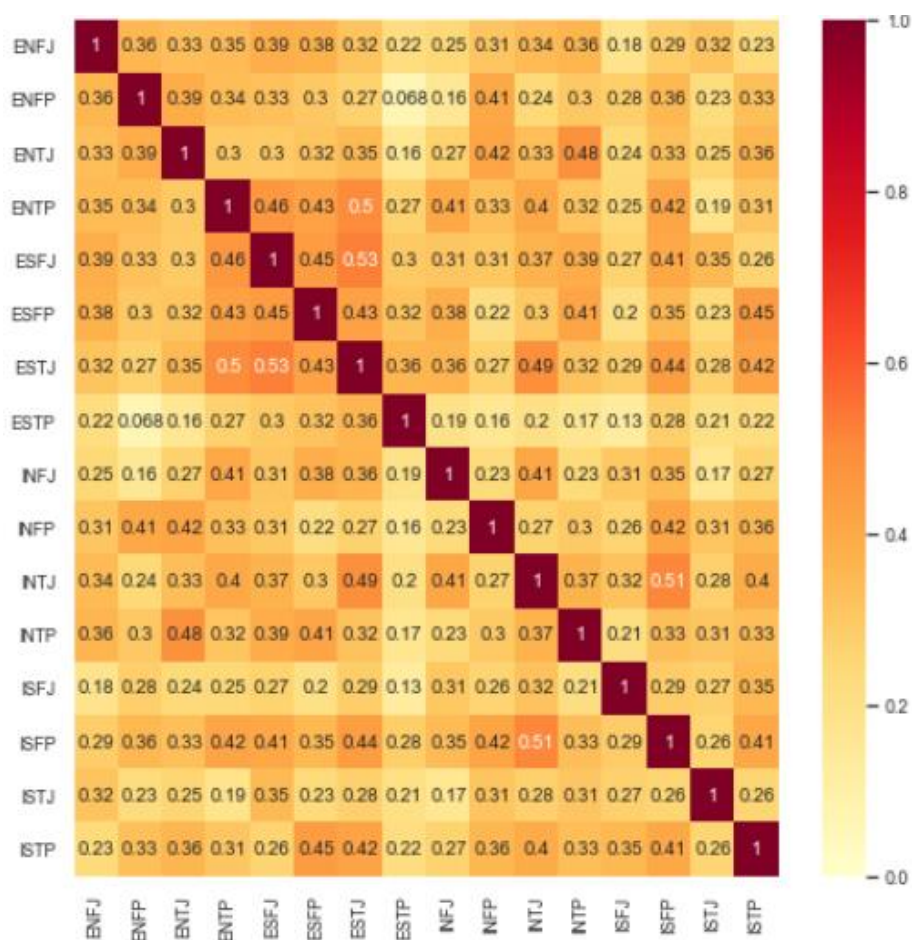
Dodatni primjer pristupa uravnoteženju podataka uključuje klasifikatore koji uče samo na primjerima ciljane klase (primjeri manjinske klase). Na taj način se poboljšava efikasnost klasifikacije za evaluacijske podatke manjinske klase.

Ansambl pristupi uključuju više klasifikatora kako bi se poboljšali rezultati klasifikacije. Najpoznatiji ansambl pristupi su treniranje na pod uzorcima skupa za učenje (engl. *bagging*) i slijedno učenje algoritama na pogreškama prethodnih algoritama (engl. *boosting*).

Tablica 3 Pregled prednosti i nedostataka izabranih pristupa za uravnoteženje podataka

Pristup	Prednosti	Nedostaci
Naduzorkovanje	Izostanak gubitka informacija	Rizici prenaučnosti Vremenski zahtjevniji Dodatna računalna obrada
Poduzorkovanje	Jednostavan pristup	Rizik pristranosti Rizik gubitka korisnih informacija Neupotrebljiv na manjim skupovima podataka
Ansambli	Bolji rezultati klasifikacije	Povećana složenost Rizik prenaučnosti Vremenski zahtjevniji
Učenje jedne klase	Jednostavnost i brzina	Nije za sve vrste algoritama Rizik lošije efikasnosti predikcije
Troškovno osjetljivo učenje	Jednostavnost i efikasnost	Ne-efikasnost ako je trošak pogreške nepoznat

Potrebno je istaknuti i na činjenicu da istraživanje [42] ukazuje i na mogućnosti poboljšanja klasifikacije MBTI pomoću mreže LSTM uzevši u obzir i semantičke aspekte pogreške. Slika 14 daje rezultate semantičke tekstualne sličnosti klasa MBTI. Izračunata je korelacija između tekstne semantičke sličnosti između klasa MBTI i Hammingove udaljenosti između klasa MBTI. Najveću sličnosti imaju klase s najmanjom Hamming udaljenosti između klasa, što je i očekivano. Kao rezultat, koeficijent korelacije je -0.516, a p vrijednost se može smatrati 0. Ovi rezultati ukazuju da je informaciju semantičke sličnosti moguće uključiti u poboljšanje klasifikacije MBTI, a jedan od načina uključivanja rezultata istraživanja [42] bi mogao biti da se dobiveni koeficijenti korelacije iskoriste kao dodatni težinski faktori prilikom klasifikacije.



Slika 14 Semantička tekstualna sličnost klasa MBTI

Tablica 4 Analiza strukturnih hiperparametara mreže LSTM pri višeklasnoj klasifikaciji MBTI [41]

Model	Broj slojeva	Broj neurona	Točnost [%]	Preciznost	Odziv	F1-mjera	Vrijeme treniranja [s]
model1	1	5	56,605	0,559	0,559	0,559	80,22
model2	1	10	66,052	0,656	0,653	0,653	74,35
model3	1	25	64,900	0,644	0,641	0,641	76,65
model4	1	50	53,763	0,509	0,506	0,469	84,01
model5	1	100	55,991	0,555	0,554	0,555	99,37
model6	1	150	58,986	0,675	0,621	0,567	135,18
model7	1	500	72,043	0,720	0,711	0,712	295,60
model8	2	5, 5	62,366	0,628	0,629	0,624	123,39
model9	2	10, 10	70,276	0,700	0,700	0,700	127,06
model10	2	25, 25	56,912	0,584	0,526	0,447	130,03
model11	2	50, 50	70,353	0,704	0,706	0,703	143,91
model12	2	100, 100	55,223	0,526	0,500	0,359	173,72
model13	2	150, 150	68,356	0,695	0,693	0,683	256,33
model14	2	500, 500	58,909	0,585	0,563	0,545	607,40
model15	3	5, 5, 5	66,206	0,680	0,675	0,661	177,29
model16	3	10, 10, 10	52,995	0,678	0,571	0,470	181,15
model17	3	25, 25, 25	55,376	0,777	0,502	0,360	183,98
model18	3	50, 50, 50	71,813	0,716	0,717	0,716	199,23
model19	3	100, 100, 100	56,068	0,547	0,541	0,532	255,29
model20	3	150, 150, 150	66,743	0,670	0,651	0,649	377,69
model21	3	500, 500, 500	70,353	0,722	0,683	0,681	945,07
model22	3	25, 50, 100	55,223	0,276	0,500	0,356	217,40
model23	3	100, 50, 25	59,601	0,652	0,555	0,492	234,87

2.4.2. CNN

Konvolucijska neuronska mreža (engl. *convolutional neural network* - *CNN*) predstavlja arhitekturu dubokog učenja koja je inicijalno razvijena za obradu slika. Konvolucijski slojevi su se prvi put koristili u okviru mreže AlexNet s osam slojeva za klasifikaciju slika, a uvedeni su u radu [49] te potom korišteni u LeNet [50].

Konvolucijski modeli su specijalizirani za podatke sa strukturom rešetke koja definira relaciju susjedstva. Za razliku od tradicionalnih FFNN, u kojima se elementima dodjeljuju težinski faktori, mreža CNN ima najmanje jedan konvolucijski sloj umjesto potpuno povezanih slojeva i koristi skup jezgri (engl. *kernels*) koji imaju svrhu filtriranja podataka.

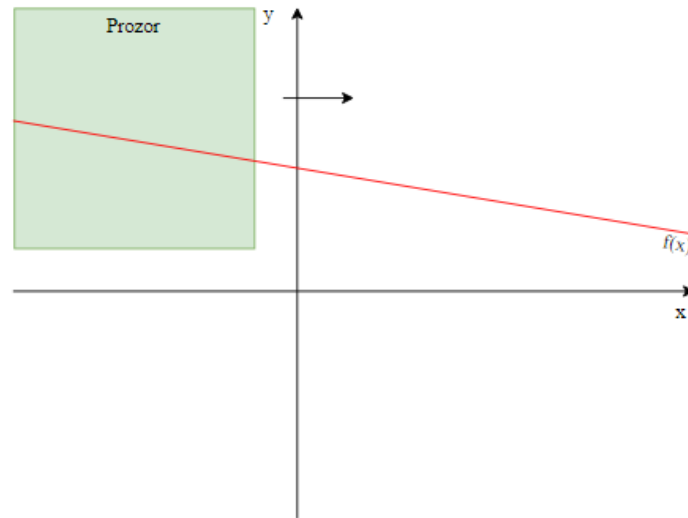
Jedna od ključnih karakteristika konvolucijske neuronske mreže je invarijanca translacije (engl. *translation invariance*) koja podrazumijeva da mreža može prepoznati objekt i ako je objekt transliran horizontalno ili vertikalno na drugu poziciju. Druga ključna karakteristika mreže CNN je mogućnost kompozicije (engl. *compositionality*) te predstavlja sposobnost prepoznavanja sve složenijih predložaka kroz hijerarhijski pogled, tj. kombinacijom elementarnih značajki u složenije predloške više razine. Upravo zahvaljujući kompoziciji i invarijanci translacije mreže CNN se mogu koristiti i u NLP zadacima. Mreže CNN su se pokazale vrlo efikasne u klasifikaciji tekstualnih sadržaja [51] [52].

U usporedbi s mrežama RNN, mreže CNN ne mogu kodirati duge lance ovisnosti te za zadatke gdje su važni dugi lanci ovisnosti treba dati prednost RNN arhitekturama, kao što je LSTM. S druge strane, mreže CNN prvo kreiraju vektor s elementima promatranog teksta za sve pod-fraze teksta, ne samo za gramatički korektne, kao što je slučaj s mrežama RNN. Na taj način mreže CNN grupiraju sadržaj za odgovarajući zadatak.

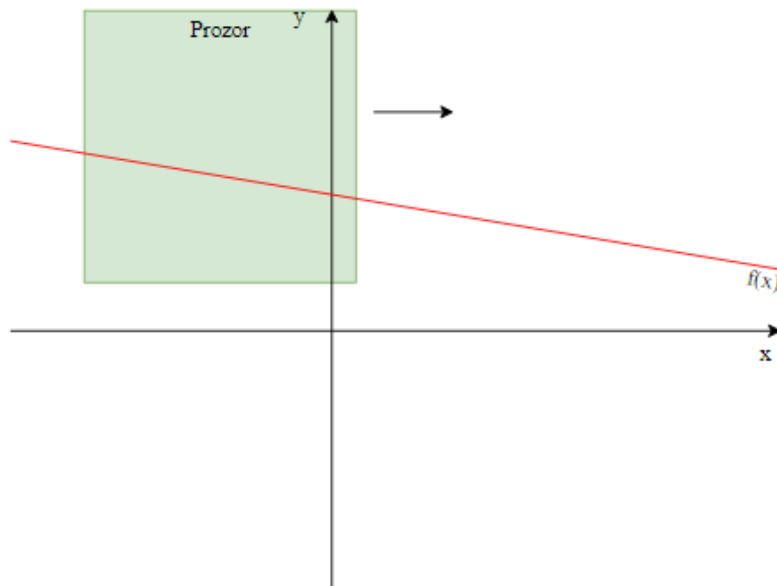
Za razliku od primjene na obradi slike, mreže CNN za obradu prirodnog jezika koriste kao ulaz, umjesto vrijednosti piksela slika, vrijednosti vektora riječi u ugradbenoj matrici (engl. *word embeddings*). U primjeni mreža CNN u NLP zadacima, dokument se predstavlja matricom realnih brojeva $A \in \mathbb{R}^{s \times d}$, gdje je s duljina dokumenta, a d dimenzija ugradbene matrice vektora riječi. Vektori riječi se mogu inicijalizirati slučajnim procesom ili mogu biti odabrani iz grupe predtreniranih rječnika.

Princip rada mreže CNN u obradi slike uključuje konvoluciju tenzora (engl. *tensor*) slike skupom jezgri (engl. *kernel*) dimenzije $m \times m$. Drugim riječima, procesom konvolucije se

pomiče prozor preko uzorka podataka (čuvajući fokus na podskup koji je dio veće cjeline). Ideju konvolucije prikazuju slika 15 i slika 16.



Slika 15 Ideja konvolucije prozora nad funkcijom $f(x)$ (a)



Slika 16 Ideja konvolucije prozora nad funkcijom $f(x)$ (b)

Formalno, konvolucija između vektora f i g , nad ulazom n je određena izrazom (4). Vektor je predstavljen s g .

$$(f * g)[n] = \sum_{m=-M}^M f[n-m] g[m] \quad (4)$$

Konvoluciju u strojnom učenju se može smatrati ekvivalentnom unakrsnoj korelaciji [33], što je prikazano relacijom (5), čime je istaknuto da se jezgra w može koristiti za ekstrakciju lokalnih značajki iz signala x .

$$h(t) = (w * X)(t) = \int_{D(w)} W(\tau)X(t + \tau)d\tau \quad (5)$$

Ulaz je određen funkcijom x , jezgra je određena funkcijom w , a rezultat koji predstavlja mapu značajki je određen s $h(t)$.

Konvolucijski slojevi su mape značajki i predstavljaju filtere ulaznih vrijednosti. Filteri imaju dva ključna elementa: težinske vrijednosti i aktivacijsku funkciju. Konvolucijski sloj predstavlja linearnu operaciju koju slijedi nelinearna transformacija i sličan je potpuno povezanom sloju. Ako je veličina prozora označena s m , matrica parametara s W , a korak (engl. *stride*) je 1, tada broj polja koji će se obuhvatiti konvolucijom određuje izraz $s-m+1$. Izlazna vrijednost konvolucijskog sloja za pojedini filter je vektor čiji su elementi određeni izrazom (6).

$$o_i = W \cdot A[i:i + m - 1]; o_i \in \mathbb{R}^{s-m+1} \quad (6)$$

$A[i:i + m - 1] \in \mathbb{R}^{m \times d}$ predstavlja i -ti dio matrice a operator $\cdot$ vraća sumu skalarnog umnoška dvije matrice prema retku. Mapa značajki, koja je pridružena filteru i je određena izrazom (7), gdje je nelinearna aktivacijska funkcija označena s f a realna vrijednost pristranosti označena s b .

$$c_i = f(o_i) + b \quad (7)$$

Pojednostavljenje računalnih izračuna u okviru mreža CNN se osigurava uporabom tehnike sažimanja (engl. *pooling*) koja smanjuje izlazne vrijednosti iz jednog sloja prije predaje vrijednosti u novi sloj. Dakle, postupak sažimanje se koristi za reduciranje dimenzionalnosti u mreži CNN. S obzirom na ciljane značajke mogu se koristiti različite tehnike udruživanja

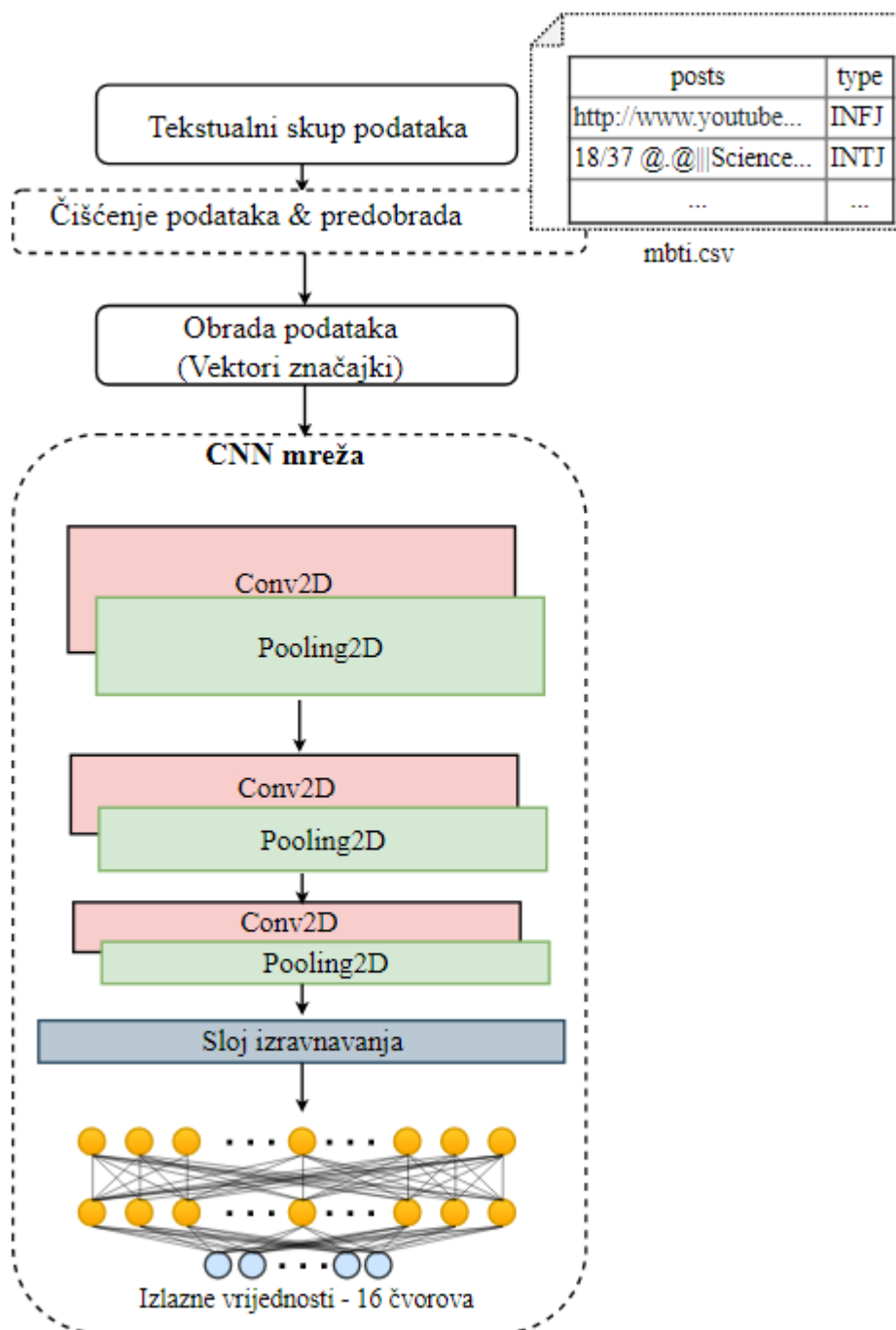
[53], a najčešće korištena je maksimalno sažimanje (engl. *max pooling*) u kojoj se odabire element s najvećom vrijednosti i promatranom prozoru sažimanja (engl. *window pooling*).

Pored maksimalnog, često se koristi i sažimanje koje se zasniva na prosječnoj vrijednosti. Iako je prosječna vrijednost intuitivniji odabir, odabir sažimanja prema maksimalnoj vrijednosti osigurava da odabirom najveće aktivacijske vrijednosti za određeno područje, model može prepoznati značajku koja je najistaknutija za područje. Na taj način konvolucijska mreža modelira lokalne interakcije te model ima smjernicu za učenje ciljane značajke, bez obzira na položaj značajke. CNN modele karakterizira dijeljenje parametara, čime se postiže reprezentacija ekvivarijantnosti s obzirom na pomak.

Kako bi se vrijednosti prenijele nakon operacije udruživanja u sljedeći sloj, mape značajki se poravnavaju u jednu kolonu (engl. *flatten*). Završni slojevi u mreži CNN su uobičajeno potpuno povezani slojevi. Za zadatke klasifikacije, u mrežama CNN se na kraju dodaje funkcija softmax, s konačnim rezultatom distribucije vjerojatnosti klasa. Tijekom učenja konvolucijske mreže, u koraku povratne sprege (engl. *backpropagation*) ažuriraju se vrijednosti težina i filteri za pronalazak značajki. Točna pozicija neke značajke nije bitna nego je bitno postoji li određena značajka ili ne.

Slika 17 daje primjer arhitekture konvolucijske neuronske mreže za klasifikaciju teksta na primjeru skupa podataka koji je označen za klasifikaciju MBTI. Ulazne vrijednosti čini tekstualni skup podataka koji je u ovom primjeru prilagođen klasifikaciji tipa osobnosti prema modelu MBTI. Primjer tekstualne datoteke je datoteka `.csv` formata s tekstualnim vrijednostima i označenim tipom podataka prema modelu MBTI.

Nakon čišćenja i predobrade podataka, kreiraju se vektori značajki. Primjer pokazuje tri sloja konvolucije i udruživanja, sloj izravnavanja te završava s dva sloja potpuno povezane mreže i izlaznim vrijednostima čiji broj odgovara vrijednostima skupa izlaznih vrijednosti. Na primjeru skupa podataka MBTI, za višeklasnu klasifikaciju, moguće je 16 izlaznih vrijednosti.



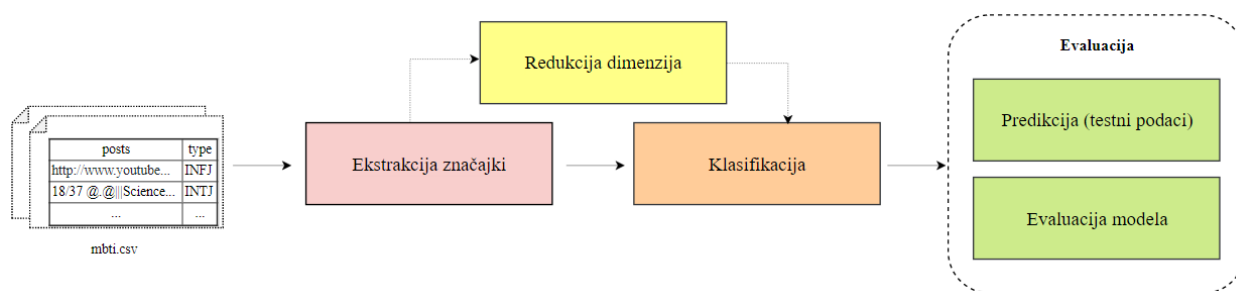
Slika 17 Arhitektura konvolucijske neuronske mreže za klasifikaciju teksta

Svaki neuron u konvolucijskom sloju implementira operaciju poznatu kao konvolucijska jezgra (engl. *convolutional kernel*). Težine su posložene u 2D predlošku i formiraju konvolucijsku matricu. Neuroni su u konvolucijskom sloju spojeni rijetko (engl. *sparsely*). Konvolucijski sloj se sastoji od više kanala ili mapi značajki. Svi neuroni u istom kanalu dijele težinske vrijednosti (engl. *weights*).

2.4.3. Primjena klasifikacije u obradi prirodnog jezika

Obrada prirodnog jezika predstavlja interdisciplinarno područje koje povezuje umjetnu inteligenciju (engl. *artificial intelligence*), računalnu znanost i lingvistiku s ciljem obrade prirodnog jezika u govoru i pismu. Ciljevi obrade prirodnog jezika uključuju otkrivanje korisnih informacija, razumijevanje odnosa između riječi u specifičnim problemima te na osnovu spoznaja iz obrađenog teksta automatiziranje zadataka kao što su npr. klasifikacija tekstualnih dokumenata i generiranje teksta. U fokusu disertacije su zadaci klasifikacije.

Većina pristupa klasifikaciji teksta i kategorizaciji dokumenata se sastoji od sljedećih koraka: ekstrakcija značajki (engl. *feature extraction*), redukcija dimenzija (engl. *dimensions reduction*), primjena modela klasifikacije (engl. *classifier selection*) i evaluacija (engl. *evaluation*). Slika 18 daje prikaz standardnog procesa klasifikacije teksta na primjeru tekstualnih datoteka za klasifikaciju MBTI.



Slika 18 Standardni proces klasifikacije teksta

Ekstrakcija značajki se koristi da se tekst i dokumenti, koji su nestrukturirani skupovi podataka transformiraju matematičkim modeliranjem u strukturirani prostor značajki. Prvo se podaci čiste od nepotrebnih znakova i riječi, a nakon toga se primjenjuje formalna ekstrakcija značajki. Primjeri tehnika ekstrakcije značajki su Frekvencija izraza – Inverzna frekvencija dokumenta (engl. *Term Frequency-Inverse Document Frequency – TF-IDF*), Frekvencija izraza (engl. *Term Frequency – TF*) [54], Riječ u vektor (engl. *Word2Vec*) [55], Globalni vektori za predstavljanje riječi (engl. *Global Vectors for Word Representation – GloVe*) [56]. U eksperimentalnom dijelu je korišten FastText. Kao vektorski model FastText je razvijen u okviru Facebook AI Research laboratorija kako bi se riješio problem da dotadašnji modeli ignoriraju morfologiju riječi i pridružuju različite vektore svakoj riječi [57]. FastText je razvijen s vektorskim modelima s dimenzijom vektora 300, za 294 jezika te su modeli trenirani na Wikipedia dokumentima.

Redukcija dimenzija se može koristiti kako bi se reduciralo vrijeme i memorijska kompleksnost algoritma klasifikacije uzrokovana velikim brojem jedinstvenih riječi u tekstualnom skupu podataka. Primjeri tehnika redukcije dimenzija su Analiza glavne komponente (engl. *Principal Component Analysis – PCA*), Linearna razlikovna analiza (engl. *Linear Discriminant Analysis – LDA*), ne-negativna matrična faktorizacija (engl. *non-negative matrix factorization – NMF*), slučajna projekcija (engl. *random projection*), autoenkoderi (engl. *autoencoders*), t-distribuirano stohastičko ugrađivanje susjeda (engl. *t-distributed stochastic neighbor embedding – t-SNE*).

Korak klasifikacije u procesu uključuje odabir modela klasifikacije. Jedan od pristupa klasifikaciji može biti uporaba modela klasičnog strojnog učenja kao što su: tehnike glasanja (engl. *ensemble*) kao što su pojačavanje i pakiranje (engl. *boosting and bagging*), logistička regresija (engl. *logistic regression – LR*), naivni Bayesov klasifikator (engl. *naïve Bayes Classifier – NBC*), *k-nearest neighbor – KNN*, metoda potpornih vektora (engl. *support vector machine – SVM*) itd. U fokusu disertacije su metode dubokog učenja budući da modeli dubokog učenja zbog svoje strukture imaju kapacitet modelirati složene nelinearne odnose između podataka [51].

Postupak evaluacije omogućava ocjenu performansi modela u klasifikaciji teksta. Primjena odgovarajuće metrike je bitan korak u ocjeni performansi modela. Pri tome je bitan oprez kod odabira jer npr. izračun točnosti predstavlja jednostavnu metriku, ali metrika ne vodi ispravnim zaključcima kada se koristi nad neuravnoteženim skupovima podataka [58].

Metrike koje se mogu koristiti su klasifikaciji su preciznost, točnost, F_β mjera (engl. *F_β score*), Matthews koeficijent korelacije (engl. *Matthews Correlation Coefficient – MCC*), operativne karakteristike primatelja (engl. *receiver operating characteristics – ROC*) [59], površina ispod ROC krivulje (engl. *area under ROC curve – AUC*) [60]. U nastavku su detaljnije objašnjene mjere koje su korištene u disertaciji [61]. Oznaka M se odnosi na makro-uprosječivanje, a μ na mikro-uprosječivanje.

Matrica zabune (engl. *confusion matrix*) osigurava usporedbu stvarnih oznaka i oznaka koje su predikcija modela. Slika 19 daje strukturu matrice zabune.

		Oznake modela	
		1	0
Stvarne oznake	1	TP	FN
	0	FP	TN

Slika 19 Matrica zabune za binarnu klasifikaciju

Za slučaj višeklasne klasifikacije, iz matrice $K \times K$ ($K > 2$), za svaku klasu i se elementi matrice zabune određuju na sljedeći način:

- $TP_i = i$ -ti element dijagonale
- $FP_i =$ zbroj ne dijagonalnih elemenata i -tog stupca
- $FN_i =$ zbroj ne dijagonalnih elemenata i -tog retka
- $TN_i = N - TP_i - FP_i - FN_i$ (zbroj po elementima izvan retka i te stupca i)

Prosječna točnost višeklasne klasifikacije se izračunava prema izrazu (8) te predstavlja mjeru klasifikatora i izračunava udio točno klasificiranih primjera u skupu svih primjera.

$$Acc_M = ACC_\mu = \frac{\sum_{i=1}^K \frac{TP_i + TN_i}{TP_i + FN_i + FP_i + TN_i}}{K} = \frac{1}{K} \sum_{i=1}^K ACC_{i_\mu} \quad (8)$$

TP_i predstavlja vrijednost istinski pozitivne (engl. *true positive*) vrijednosti za C_i , TN_i predstavlja istinski negativne (engl. *true negative*) vrijednosti za C_i , FN_i predstavlja lažno negativne (engl. *false negative*) vrijednosti za C_i , a FP_i predstavlja lažno pozitivne (engl. *false positive*) vrijednosti za C_i .

Izraz (9) prikazuje kako se izračunava preciznost za višeklasnu klasifikaciju skupa s K klasa i mikro-uprosječivanje, a izraz (10) je izraz za preciznost za makro-uprosječivanje. Preciznost daje udio pozitivno klasificiranih primjera u skupu pozitivno klasificiranih primjera. S obzirom na činjenicu da je MBTI neuravnotežen skup podataka, za preciznost je bitno da makro-uprosječivanje ima što veću vrijednost jer makro-uprosječivanje tretira sve klase ravnopravno dok mikro-uprosječivanje daje prevagu većinski zastupljenim klasama.

$$P_{\mu} = \frac{\sum_{i=1}^K TP_i}{\sum_{i=1}^K (TP_i + FP_i)} \quad (9)$$

$$P_M = \frac{\sum_{i=1}^K \frac{TP_i}{TP_i + FP_i}}{K} = \frac{1}{K} \sum_{i=1}^K P_{i_{\mu}} \quad (10)$$

Odziv (engl. *recall*, *true positive rate*, *sensitivity*) predstavlja mjeru koja pokazuje koliko model ispravno identificira TP i prikazuje udio pozitivno klasificiranih primjera u skupu svih pozitivnih primjera. Izraz (11) prikazuje izračun odziva kod mikro-uprosječivanja, a izraz (12) prikazuje izračun kod makro-uprosječivanja. I za ovu mjeru je značajno da kod neuravnoteženih skupova podataka makro-uprosječivanje, koje daje jednaku važnost svim klasama, ima veći značaj.

$$R_{\mu} = \frac{\sum_{i=1}^K TP_i}{\sum_{i=1}^K (TP_i + FN_i)} \quad (11)$$

$$R_M = \frac{\sum_{i=1}^K \frac{TP_i}{TP_i + FN_i}}{K} = \frac{1}{K} \sum_{i=1}^K R_{i_{\mu}} \quad (12)$$

Harmonijska srednja vrijednost preciznosti i odziva se predstavlja F_1 mjerom koja uzima u obzir FP i FN. Ova mjera daje precizniji odgovor o kvaliteti klasifikatora od točnosti, osobito za neuravnotežene skupove podataka. Izraz (13) daje izračun za mikro-uravnoteženu mjeru, a izraz (14) za makro-uravnoteženu mjeru. Izraz (15) je standardni način izračuna mjere makro- F_1 i daje jednaku težinu svim klasama, bez obzira na njihovu frekvenciju pojavljivanja u skupu podataka te predstavlja robusniji izraz u odnosu na izraz (14). Izraz (14) upravo u nebalansiranim skupovima podataka može dati pogrešne vrijednosti, koje su više od ispravnih vrijednosti, budući da izračun uključuje umnožak vrijednosti preciznosti svake klase s vrijednosti odziva za sve klase. Izraz (15) uzima za vrijednost preciznosti svake klase samo vrijednost odziva za istu klasu [62]. U disertaciji je korišten izraz (15) za izračun mjere

makro-F1. Ako se želi naglasiti značaj odziva u odnosu na preciznost može se koristiti i općenitiji izraz (16) za F mjeru s mikro-uprosječivanjem te izraz (17) s makro-uprosječivanjem. Uobičajene vrijednosti za β su 2, čime je za odziv stavljen veći značaj u odnosu na preciznost te $\frac{1}{2}$ kada se želi dati veći značaj za preciznost.

$$F_{1\mu} = 2 \cdot \frac{P_\mu \cdot R_\mu}{P_\mu + R_\mu} \quad (13)$$

$$F_{1M} = 2 \cdot \frac{P_M \cdot R_M}{P_M + R_M} \quad (14)$$

$$F_1 = \frac{1}{K} \sum_{i=1}^K F_{1i} \quad (15)$$

$$F_{\beta\mu} = (1 + \beta^2) \cdot \frac{P_\mu \cdot R_\mu}{(\beta^2 \cdot P_\mu) + R_\mu} \quad (16)$$

$$F_{\beta M} = (1 + \beta^2) \cdot \frac{P_M \cdot R_M}{(\beta^2 \cdot P_M) + R_M} \quad (17)$$

Kada je vrijednost metrike F_{1M} visoka, to znači da klasifikator dobro klasificira svaku individualnu klasu. Stoga, metrika F_{1M} je indikativna za kvalitetnu procjenu uspješnosti klasifikatora nad neuravnoteženim skupovima podataka, kao što je i skup podataka MBTI.

3. PRIJAŠNJA ISTRAŽIVANJA – PRIMJENA STROJNOG UČENJA U KLASIFIKACIJI MBTI

Istraživanje u okviru ove disertacije se dotiče rezultata istraživanja u području tipova osobnosti, ali izvorni znanstveni doprinos nije u području psihologije ličnosti. Predloženi pristup podliježe istim kriterijima i ograničenjima kao i svaki psihometrijski instrument iz psihologije ličnosti. Jedno od značajnih ograničenja proizlazi iz načina na koji je skup podataka označen, budući da na valjanost psihometrijskog instrumenta veliki značaj ima činjenica je li skup podataka samo označen ili je oznaku tipa osobnosti dala stručna osoba. Na primjeru klasifikacije osobnosti prema Big Five metodi je naznačena korelacija od 88% između samo označenih tipova osobnosti i tipova osobnosti procijenjenih od strane stručnih osoba. Interesantno je zapažanje da i ručna i automatizirana procjena tipa osobnosti na osnovu sadržaja s društvenih mreža imaju nisku razinu razlikovne valjanosti, koja se odnosi na mjerenje osobina ličnosti koje jasno razlikuje pojedine osobine ličnosti koje se promatraju. Dodatno, istaknuta je sumnja da i ručna i automatizirana procjena tipa osobnosti mogu dati kompletnu sliku osobnosti. U cilju poboljšavanja valjanosti automatizirane procjene tipa osobnosti na osnovu sadržaja na društvenim mrežama, sugerira se fokus na praćenje procesa u kojem karakteristike osobnosti utječu na angažman osobe na društvenim mrežama [63].

S obzirom na fokus istraživanja, disertacija se prvenstveno dotiče istraživanja u području tipova osobnosti po Myers-Briggsovoj, ali se istraživanja u području klasifikacije tipova osobnosti često bave i tzv. Big Five-klasifikacijom tipova osobnosti. Detaljno pojašnjenje Big Five tipova osobnosti daje [64], a usporedbu klasifikacije tipova osobnosti prema modelima Myers-Briggs i Big Five je moguće dobiti analizom [6] i [64].

Najčešći pristupi pri klasifikaciji MBTI uključuju višeklasnu klasifikaciju na 16 klasa ili kombinaciju četiri binarne klasifikacije. Istraživači najviše koriste kombinaciju binarnih klasifikatora iz više razloga. Primjenom binarnih klasifikatora skup podataka dijelimo na pod skupove koji su mnogo bolje uravnoteženi od skupa koji promatramo kroz svih 16 složenih klasa. Time se osiguravaju uvjeti za bolju točnost pojedinačnih binarnih klasifikatora u širem smislu. Binarni klasifikatori pružaju i veću fleksibilnost, budući da osiguravaju mogućnosti kombiniranja klasifikatora, za razliku od višeklasne klasifikacije.

3.1. Istraživanja s primjenom standardnih algoritama strojnog učenja

Ako se promatra standardne algoritme strojnog učenja, isti su se iz povijesnih razloga, nametali kao prva opcija za klasifikaciju MBTI. Tako istraživanje [65] prikazuje rezultate primjene algoritma ekstremnog podizanja gradijenta (engl. *extreme gradient boosting*) za binarnu klasifikaciju parova suprotnosti modela MBTI. Kao metrika validacije je korištena samo točnost (engl. *accuracy*). Prezentirani su rezultati točnosti od 86.06% za dihotomiju N/S kao najbolji te 65.70% za dihotomiju J/P kao najlošiji rezultati. Autori su koristili i rekurentne neuronske mreže s najvećom točnošću od 77.8% za dihotomiju F/T te najmanjom točnošću 62% za dihotomiju N/S.

Uporabu jednostavnog neighbor klasifikatora prezentira rad [66]. Prezentirani su najbolji rezultati odziva (engl. *recall*) i preciznosti (engl. *precision*) između 80% i 95% za dihotomiju E/I, a ostale metrike su u rasponu od 40% do 70%. Najlošija preciznost od 40% je ostvarena za dihotomiju J/P.

Rad [67] predlaže modele učenja glasanja (engl. *ensemble*) za binarnu klasifikaciju MBTI: pakiranje (engl. *bagging*), pojačavanje (engl. *boosting*) i slaganje (engl. *stacking*). Najbolji rezultati su ostvareni slaganje pristupom i to 97.53% točnosti za dihotomiju S/N. Isti pristup je dao i najviši F1-score 97.42%. Kod ostalih metrika, najvišu vrijednost za odziv je postigao pristup slaganja te pristup pojačavanja za metriku odziva.

Uporabu SVM, naivnog Bayesovog klasifikatora i neuronskih mreža za binarnu klasifikaciju MBTI obrađuje rad [68]. Najbolja točnost je postignuta uporabom SVM za E/I od 84.9%, S/N od 88.4%, T/F od 87% te J/P od 78.8%. Autori su koristili za semantičku reprezentaciju teksta rječnik LIWC (engl. *linguistic inquiry and word count - LIWC*), EmoSentNet (Emolex) i ConceptNet u kombinaciji s TF-IDF za svaki redak i dekompozicijom singularne vrijednosti (engl. *singular value decomposition - SVD*).

Rad [69] prikazuje binarnu klasifikaciju MBTI na skupu podataka koji je kreiralo 40 diplomiranih studenata na bazi pisanih uzoraka u učionici. Autori su koristili naivan Bayesov klasifikator i metodu potpornih vektora. Najbolje rezultate je pokazao pristup s naivnim Bayesovim klasifikatorom za preciznost i odziv od 75%.

Pojedini istraživači su koristili slučajnu šumu kao klasifikator glasanja za binarni pristup klasifikaciji MBTI [70]. Za vektorsku prezentaciju riječi je korišten Word2Vec te dodatne značajke, konkretno broj riječi po komentaru. Predstavljena je točnost 100% za sve

dihotomije. Ipak, metrike koje su bitne za neuravnotežene skupove podataka nisu predstavljene.

Pojačivač gradijenta (engl. *gradient boosting*) za predikciju te K-means klasteriranje s tradicionalnim TF-IDF za klasteriranje je pristup koji predlaže rad [71] za binarnu klasifikaciju MBTI. Za predloženu arhitekturu je predstavljena najbolja točnost 89.01% za E/I, a najlošija za dihotomiju F/T 81.19%.

U radu [72] autori su koristili klasifikator dodatnih stabala (engl. *extra trees classifier*), naivni Bayesov klasifikator, logističku regresiju i SVM za klasifikaciju MBTI. Predstavljen je najbolji rezultat za logističku regresiju, gdje su točnosti i F-mjera bili 66.59%. Autori su krenuli s pretpostavkom da je točnost najvažnija metrika, što je upitno s obzirom na neuravnoteženost skupa podataka, što nije uzeto u obzir. Nakon pokušaja poboljšanja parametara, predstavljeno je poboljšanje od 1%. Rad ne citira kvantitativne detalje vezano uz poboljšanje parametara.

3.2. Istraživanja s primjenom dubokog učenja

Rad [4] daje pregled trendova u pristupima dubokog učenja, pri čemu su korišteni različiti izvori podataka, tekstualni, audio, video i multimedijalni izvori. Moguće je uočiti da prevladavaju istraživanja nad Big Five instrumentom za određivanje tipa osobnosti. Prezentiran je samo jedan pristup klasifikaciji MBTI uporabom dubokog učenja. Autori su potencirali potrebu efikasnijeg načina označavanja skupova podataka, što bi moglo poboljšati broj i kvalitetu skupova podataka.

Modeli koji su se koristili su nenadzirano klasteriranje s dekompozicijom na singularne vrijednosti (engl. *singular value decomposition - SVD*), skup ili vreća riječi (engl. *bag of words*) neuronska mreža s pomakom unaprijed (engl. *feed forward*) te rekurentna neuronska mreža (RNN) s dugom kratkotrajnom memorijom (LSTM) kao glavni model korišten u radu. Ostvarena je prosječna točnost od 37% pri klasifikaciji u jednu od 16 mogućih klasa. Izrada preciznog modela bi imala značajan utjecaj na poslovnu inteligenciju, analize kompatibilnosti odnosa i druga srodna pitanja, s posebnim naglaskom na dodatne primjene u društvenim znanostima. U istraživanju se istaknulo dobro zapažanje da nedostatak predloženog modela predstavlja činjenica da je ograničenih kapaciteta te da ne uzima u obzir sekvencijalnu prirodu rečenice. U istraživanju se nagovještava da bi možda bila korisna ideja razrade funkcije

gubitka na način da penalizira na različite načine sličnu klasifikaciju od potpuno različite klasifikacije. U radu je otvoreno pitanje razrade funkcije gubitka, a u istraživanju u okviru disertacije se nastoji dati odgovor na to pitanje [73]. No, rad nije recenziran i objavljen u okviru časopisa ili na konferenciji.

Rad [74] istražuje pristupe dijeljenja tekstualnih informacija za potrebe RNN, što je u kontekstu ovog istraživanja bitno zbog prijedloga potencijalno srodnog pristupa u analizi informacijski zavisnih elemenata, što je relevantno za analizu komponenata klasa. Predložen je model neuronske mreže, arhitekture RNN, s tri mehanizma dijeljenja informacija za modeliranje teksta sa specifičnim zadacima i dijeljenim slojevima.

Također, postoji istraživanje [75] u kojem se uporabom plitkih neuronskih mreža radi na prepoznavanju osobnosti programera iz zadanog fragmenta njihovog programskog kôda. Predloženo rješenje predviđa pet karakteristika ličnosti (engl. *Big Five*). Za razliku od njihovog pristupa koji analizira polustrukturirane podatke, u disertaciji se razvija pristup za analizu nestrukturiranih podataka.

Radovi [76] i [77] se također bave klasifikacijom osobnosti pomoću umjetnih neuronskih mreža, ali se u bitnom razlikuju od istraživanja u disertaciji. Rad [76] se također bavi primjenom umjetnih neuronskih mreža na određivanje tipa osobnosti, po čemu je srodno ovom istraživanju, ali njihov pristup koristi značajke iz područja računalnog vida za analizu grafoloških elemenata rukopisa, dok je pristup u istraživanju u okviru disertacije temeljen na obradi prirodnog jezika. U radu [77] se istražuje primjena RNN kao alata za praćenje promjena emocionalnog stanja analizom filmskog scenarija, pri čemu je promjena emocionalnog stanja predstavljena kroz pet faktora (engl. *Big Five*) osobnosti. U istraživanju je naglasak na mjerenje promjene stanja, a ne na apsolutnu mjeru osobnosti kroz klasifikaciju uporabom Big Five modela.

Društvene mreže kreiraju volumen podataka sa stopom rasta 20 % - 30% [78] [79]. Npr., Twitter mreža ima 3.62 milijardi korisnika. Okruženje društvene mreže potiče korisnike na interakciju članova te na taj način postaju izvori enormne količine podataka za predikciju tipa osobnosti na osnovu otvorenog rječnika (engl. *open vocabulary*). U radu [80] je istaknuto da prilikom preuzimanja podataka sa društvenih mreža treba uzeti u obzir i ograničenja kao što su GDPR unutar EU te da kod primjene strojnog učenja u automatiziranoj predikciji tipa osobnosti treba uključiti stručnjake psihologe i lingviste.

Doprinos klasifikaciji MBTI kroz osiguravanje označenog Reddit skupa podataka daje rad [81]. Korištena su tri klasifikatora: troslojni perceptron (engl. *multilayer perceptron MLP*), logistička regresija (LR), metoda potpornih vektora (engl. *support vector machine - SVM*). Rješavanje problema je postavljeno kao binarni problem klasifikacije. Najbolji rezultat je ostvaren uporabom LR i MLP. Najbolja makro F1-mjera je ostvarena za dimenziju E/I 82.8%, za S/N 79.2%, T/F 67.2%, J/P 74.8%. Napravljena je i višeklasna klasifikacija MBTI s najboljom makro F1-mjerom 41.7% za MLP. Istraživanje ukazuje na teškoće razlikovanja INTP od INTJ, da je INTJ sličniji INFJ, a INTP sličniji INFP. Istraživanje ukazuje na grupiranje sličnih tipova MBTI, slično teorijskim postavkama modela MBTI.

Uporabu Twitter društvene mreže za predikciju osobnosti obrađuje rad [82], u kojem su autori koristili instrumente Big Five i MBTI i pristup je jezično neovisan. Rezultati su se temeljili na brojanju riječi i pokazali su najveće rezultate za dihotomiju S/N. Potom, istraživanje pokazuje značajne razlike u rezultatima između odabranih jezika. Npr., za dihotomiju E/I je najbolja predikcija ostvarena za njemački, talijanski i španjolski jezik. Dodatno, rad je prezentirao potencijalne izvore grešaka predikcije: strukturne greške algoritma predikcije, greška uzrokovana promjenom autora teksta kroz vrijeme i greška vezana s uporabom eseja kao polazne osnove.

U radu [83] su autori koristili grupe riječi (engl. *word n-grams*) i spol za predikciju binarnog tipa MBTI autora objave sa samostalno označenim tvitovima (engl. *tweets*). Kao meta-značajke (engl. *meta-features*) koristili su sljedbenike (engl. *followers*), tvitove i odgovore na tvitove, kao i broj omiljenih (engl. *favorite*) tvitova. Koristili su logističku regresiju kao model i zaključak autora je da dihotomije F/T i E/I imaju poprilično dobro razlikovanje, u usporedbi s drugim dimenzijama gdje je učenje bilo složenije i sa slabijim uspjehom. Najviši prijavljeni rezultat je točnost predikcije dihotomije E/I od 77%.

I rad [84] koristi Twitter informacije za predikciju osobnosti prema Big Five instrumentu, gdje su autori koristili ZeroR i Gaussove procese kao algoritme strojnog učenja i ostvareni rezultati za svaku karakteristiku osobnosti su između 11% i 18%.

Rad [85] prezentira eksperimente na Twitter skupu podataka za binarnu klasifikaciju MBTI i s 12 različitih algoritama: stohastički gradijentni spust (engl. *stochastic gradient descent - SGD*), slučajna šuma (engl. *random forest - RF*), LR, *K-nearest neighbors (KNN)*, naivni Bayesov klasifikator (NB), multinomialni naivni Bayesov klasifikator (MNB), GaussBayesov klasifikator (GNB), SVM, MLP, stablo odlučivanja (engl. *decision tree -DT*), pojačavanje, i

ET. Za dihotomiju E/I, najveća točnost je ostvarena od 78.6% za LR i MLP, ali najviša F1-mjera i odziv su bili 38% i 40%, za SGD. Za drugu dimenziju S/N, najviša točnost od 86.2% za MLP i najviša F1-mjera i odziv su bili 17% i 18%, za DT. Za treću dimenziju F/T najviša točnost je ostvarena 64.7% za MLP, a najviša F1-mjera i odziv 69% i 100%. Konačno, MLP je osigurao najveću točnost za P/J dimenziju, a s BNB klasifikatorom najvišu F1-mjeru i odziv.

U radu [86] autori koriste novi skup podataka za različite eksperimente uporabom instrumenata Big Five, MBTI i Enneagram za procjenu tipa osobnosti. Značajna činjenica je da skup podataka uključuje demografske podatke (starost, spol, lokacija, i jezik). Za trening prema modelu MBTI ostvareni postotak točnosti je 45%. U eksperimentima, autori su koristili binarnu klasifikaciju, linearnu/logističku regresiju i neuronske mreže. Pristup neuronskim mrežama je pokazao značajan prostor za unapređenje jer postoji mnogo komentara po korisniku.

Društvene mreže u jezicima drukčijim od engleskog jezika imaju interesantne semantičke izazove. Npr., semantička analiza kineskog jezika je kompleksnija nego analiza engleskog jezika. Sina Weibo je jedna od najpopularnijih društvenih mreža u Kini i vodeći mikro-blogging pružatelj usluga u Kini. Sina Weibo je bogati izvor za istraživanja predikcije osobnosti. Ipak, broj Sina Weibo korisnika koji su uključeni je relativno mali (131 od 589 korisnika). Autori su istraživali predikciju osobnosti prema Big Five dimenzijama. Pearsonova korelacijska analiza je korištena za usporedbu rezultata za dimenzije osobnosti i sve značajke. Dodatno, korišteni su rječnik LIWC za analizu sadržaja, LR, i naivni Bayesov klasifikator. Algoritam naivni Bayesov klasifikator je imao bolje rezultate preciznosti, i oba algoritma su imali slične rezultate odziva. Srednja vrijednost točnosti pet osobina ličnosti bila je 70.7%. Imajući na umu korelaciju između Big Five i MBTI, značajno je zapažanje da je najteže predvidjeti neuroticizam. Dodatno, otvorenost i slaganje je bilo lako za predvidjeti, te je u značajnoj korelaciji s dimenzijom S/N modela MBTI [87].

Istraživanje [88] se fokusira na binarni pristup binarne klasifikacije MBTI u Bahasa jeziku u Indoneziji, uporabom pristupa otvorenog rječnika. Izvor podataka je Twitter društvena mreža. Istraživanje je koristilo tri statistička modela, i pokazalo se da je naivni Bayesov klasifikator dao bolje rezultate od pristupa temeljenih na leksikonu i gramatičkim pravilima. Najviša točnost je ostvarena 80% za dihotomiju E/I te 60% za ostale dihotomije. Dodatno zapažanje istraživača je da je naivni Bayesov klasifikator zahtijevao i najkraće vrijeme izvršavanja.

U [89] se demonstrira kako uravnoteženje skupa podataka MBTI može utjecati za rezultate klasifikacije. Korišteni su nasumično prekomjerno uzorkovanje (engl. *random over-sampling*) i TF-IDF za selekciju značajki. Eksperimenti su napravljeni s nizom algoritama strojnog učenja: KNN, stablo odlučivanja, slučajna šuma, MLP, LR, SVM, XGBoost, MNB i SGDC. XGBoost algoritam klasifikacije je pokazao najbolje rezultate – više od 99% za preciznost i točnost. Prijavljeni su najniži rezultati za dihotomiju P/J.

Budući da istraživači uglavnom prijavljuju najniži rezultat klasifikacije za dihotomiju J/P, neki istraživači se fokusiraju na poboljšanje predikcije zadnje dihotomije [90]. Napravljene su usporedbe za TF-IDF, znakovnu razinu (engl. *character level*) TD i razinu riječi (engl. *word level*) TF. Korišten je također skup podataka MBTI s Personality Cafe foruma. Zaključak autora je da su prethodna istraživanja na ovom skupu podataka previše optimistična. Korišteno je pet algoritama tijekom eksperimentiranja i konačno je sugerirana uporaba LightGBM modela s TF sa znakovnom razinom kao najboljim modelom za predikciju dihotomije P/J zbog robusnosti koju je model pokazao. Rezultati su uspoređeni s rezultatima za SVM, koji je pokazao slične rezultate. Istraživanje je koristilo rječnik LIWC. Autori su prijavili najbolji rezultat za P/J i F1- mjeru 80.77% za Kaggle i 65% za Kaggle pročišćeni skup podataka. Autori sugeriraju da dihotomija P bolje korelira nego dihotomija J prema jezičnim markerima u komunikaciji u društvenim mrežama.

Piedboeuf et al. [91] demonstriraju rezultate određivanja tipa osobnosti prema DISC ili modelu MBTI na osnovu skupa podataka na poslovnoj društvenoj mreži LinkedIn. Ostvarena je preciznost između 73.7% i 80.5% za DISC model i između 80.7% i 86.2% za binarni model MBTI osobnosti. Autori su koristili SVM klasifikator u kombinaciji s algoritmom rangiranja značajki i optimizacije (engl. *feature ranking algorithm and optimization*). Korišteni su također slučajna šuma (engl. *random forest*) s AdaBoost i naivni Bayesov klasifikator. Rezultati su iznenađujući zbog profesionalne prirode mreže LinkedIn i očekivanja da će profesionalni i formalni stil pisanja utjecati na klasifikaciju osobnosti. Rezultati su slični rezultatima na drugim društvenim mrežama gdje korisnici upotrebljavaju neformalni stil pisanja i izražavanja.

Neka istraživanja u predikciji osobnosti autora teksta razmatraju da sve objave na društvenim mrežama nisu jednako bitne i predstavljaju model koji daje naglasak na razinu poruke u cilju učenja relativne težine. Istraživanje implementira koncept nad Big Five skupom podataka.

Autori zaključuju da zadnja dihotomija predstavlja ključni element u predikciji tipa osobnosti prema modelu MBTI [92].

U [93] autori predlažu novi skup podataka MBTI s oznakom tipa osobnosti i spola za nizozemski, njemački, francuski, talijanski, portugalski i španjolski jezik. Eksperimenti su rađeni s LinearSVC i 10-strukom unakrsnom provjerom (engl. *10-fold cross-validation*). Upotrijebljena je logistička regresija. Najbolji rezultati su ostvareni za nizozemski jezik, gdje je prijavljeno poboljšanje za WRB (engl. *weighted random baseline*) u metrici F1-mjera s 50.04% na 82.61% za predikciju spola. Ipak, najviši rezultat je F1-mjera 79.21% za dihotomiju S/N na talijanskom jeziku. Istraživanje ponovno prijavljuje da model daje najbolje rezultate za predikciju dihotomija F/T i E/I u odnosu na ostale dihotomije.

Moguće je tretirati tekstualni sadržaj na način da se kreira hijerarhijska struktura vektora riječi i rečenica u modelima dubokog učenja. S tim pristupom je moguće modelirati predikciju osobnosti u više-jezičnom pristupu i postići visoku performansu. Autori su koristili Big Five skup podataka i tri jezika: engleski, španjolski i talijanski. Bilo bi dobro vidjeti pristup nad skupom podataka MBTI, kako je obećano u [94].

Kako postoji specifična korelacija između instrumenata MBTI i Big Five, moguće je predvidjeti Big Five dimenzije na osnovu označenog skupa podataka MBTI. Autori su usporedili šest algoritama nadziranog strojnog učenja i tri metode ekstrakcije značajki: TF-IDF (engl. *term frequency and inverse document frequency*), BOW (engl. *bag of words*) i GloVe (engl. *global vector for word representation*). Korišten je binarni pristup i najviša točnost je ostvaren za TF-IDF i slučajne šume. Za eksperimente s BOW, najviša točnost je ostvarena s XGB. Konačno, autori su ostvarili najvišu točnost s GloVe i XGB od 99.99% [85].

U [95] autori koriste naivni Bayesov klasifikator, KNN i SVM nad BigFive skupom podataka i prema rezultatima naivni Bayesov klasifikator je dao najbolji rezultat od 60% točnosti. Autori su utvrdili da eksperiment nije uspio poboljšati prethodne rezultate i da sustav ima 65% točnosti u usporedbi s testovima koji provedeni kao ankete. Istraživanje je uključeno zbog ukupnog rezultata točnosti.

U [96] su koristili CNN i Mairesse značajke i ostvarili su najbolju točnost od 62.68% za Big Five skup podataka. Mairesse značajke [97] su skup značajki na razini dokumenta, a sastoje se od 84 značajke za predikciju osobnosti. Ipak, ne postoji diskusija glede uravnoteženosti

skupa podataka, što je važno kod odabira metrike. Ipak, istraživanje je uključeno je prezentira rezultate višeklasnog pristupa.

Pojedini pristupi redukcije, kao što su PCA (engl. *principal component analysis*) pokazuju poboljšanje, s dobitkom od manje od 2% za Big Five skup podataka [98].

Predikcija osobnosti može biti dodatni alat za analizu sentimenta u analizi sadržaja emaila i kreiranje spam filtera. Ovaj pristup može biti koristan jer se broj neželjenih elektronskih poruka povećava. Ova istraživanja su primjeri istraživanja u kojima model uključuje predikciju tipa osobnosti MBTI kao web servis postavljen na *uClassify* [99] [100].

Pojedina istraživanja analiziraju informacijske udaljenosti. U dijelu analize udaljenosti, postoje istraživanja u kojima je predložen model temeljen na metodi kNN koja klasificira podatke uz očuvanje semantičke sličnosti. Predložen je i algoritam za praćenje gubitka čija kompleksnost je kvadrat binarne duljine ulaznog koda [101].

U istraživanju [102] je predloženo uvođenje poboljšane arhitekture općenite Hammingove mreže, koja koristi neizrastu logiku. Poboljšana neuronska mreža koristi neizrastu logiku i osnovu općenite Hamming udaljenosti koja se izračunava na izlazima neurona. Pokazano je da je poboljšane linearne jedinice (ReLU) teorijski nisu esencijalne kod jednostavnih zadataka, kao što je MNIST klasifikacija. Kod složenijih klasifikacijskih problema, dvostruka struktura praga unapređuje učenje značajno.

Društvene mreže su izvori ljudske interakcije s obzirom da korisnici mogu kreirati i dijeliti svoje aktivnosti, razmišljanja i misli po različitim temama. Na taj način su društvene mreže i izvor velike količine podataka o ljudskom ponašanju i društvenoj interakciji a isti se mogu iskoristiti u različite svrhe - od poslovnih do znanstvenih.

Postoje istraživanja koja su se bavila predikcijom prema Big Five modelu analizom kratkih sadržaja na društvenim mrežama (engl. *micro blogging*) aktivnosti kineskih korisnika koji su sudjelovali u istraživanju. Korišteni su SVM i Pave Regression model. Zaključak je da se osobine ličnosti za osobe koje koriste mikro-blogging mogu predvidjeti uporabom strojnog učenja. Uspješnost klasifikacije je od 84% do 92%, a Pearsonov korelacijski koeficijent je od 0.48 do 0.54 [103]. Moguće je analizom društvenih medija raditi predikciju osobnosti prema Big Five modelu [104]. Zbog toga što se istraživanje u ovoj disertaciji tiče nestrukturiranog teksta, kakav je prisutan na društvenim mrežama, rezultati istraživanja i prototip se mogu

direktno primijeniti na klasifikaciju tipova osobnosti autora sadržaja na društvenim mrežama prema klasifikaciji tipova osobnosti po Myers-Briggsvoj.

Pristupi nadziranog strojnog učenja klasifikaciji MBTI uključuju višeklasnu klasifikaciju u 16 klasa ili četiri binarne klasifikacije. Upravo zbog dva osnovna pristupa rezultati dosadašnjih istraživanja su prikazani u dvije tablice: jednu koja prikazuje rezultate klasifikacije MBTI binarnim pristupom te drugu koja prikazuje rezultate višeklasne klasifikacije.

Tablica 5 daje prikaz istraživanja i pripadajuće algoritme u binarnoj klasifikaciji MBTI. Većina istraživanja klasifikacije MBTI koristi binarni pristup jer binarni pristup osigurava više fleksibilnosti te daje bolje rezultate klasifikacije nego višeklasna klasifikacija sa standardnom funkcijom gubitka unakrsne entropije. Dodatno, klase za binarni pristup imaju bolju uravnoteženost, što je preduvjet boljih rezultata klasifikacije. Iz kuta disertacije, binarna klasifikacija može osigurati uvide u vezi težinskih faktora u predloženoj metodi klasifikacije. Istraživači nisu imali jedinstven pristup primijenjenim metrikama za interpretaciju rezultata istraživanja, što je potrebno imati na umu s obzirom na neuravnoteženost skupa podataka MBTI.

U prikazu rezultata višeklasne klasifikacije ostavljeno je nekoliko rezultata BigFive višeklasne klasifikacije jer je broj istraživanja višeklasne klasifikacije MBTI mali, a postoje istraživanja koja pokazuju određenu korelaciju instrumenata. Za svaki pristup su predstavljeni najbolji rezultati te primijenjeni algoritam.

Tablica 6 daje rezultate istraživanja i upotrijebljene algoritme kod višeklasne klasifikacije MBTI ili ukupni rezultat kod binarnog pristupa.

Tablica 5 Rezultati binarne klasifikacije MBTI

Rad	Pristup	Model	Metrika	Najbolji rezultat
[105]	Binarni	XGBoost RNN	Točnost Točnost	86.06% (N/S) 77.8% (F/T)
[106]	Binarni	KNN	Odziv Preciznost	80%-95% (E/I) 84% - 90% (E/I)
[68]	Binarni	NB SVM Neural Net	Točnost	86.2% (S/N) 88.4% (S/N) 86.3% (S/N)

Rad	Pristup	Model	Metrika	Najbolji rezultat
[69]	Binarni	NB SVM	Preciznost Odziv Preciznost Odziv Preciznost Odziv	> 75% (S/N) > 75% (S/N) > 75% (F/T) > 75% (F/T) 31% - 60% (S/N) 45%
[81]	Binarni	LR, MVP	F1-mjera	82.8% (E/I)
[82]	Binarni	MCCV	Srednja F1	92% (S/N) 92% (S/N) 55% (E/I)
[83]	Binarni	LR	Srednja	77% (E/I)
[86]	Binarni	NN Linear/LR	Makro srednji F1 Makro srednji F1	63.4% (T) 54.6% (I) 52.8% (N) 56.6% (P) 73.9% (T) 64.2% (P) 65.4% (I) 60.6% (N)
[85]	Binarni	MLP, LR SVM, MLP MLP MLP SGD DT SGD SGD, BNB, MNB, SVM	Točnost Točnost Točnost Točnost F1-mjera F1-mjera F1-mjera F1-mjera	78.6% (E/I) 86.2% (S/N) 64.7% (F/T) 59.6% (P/J) 38% (E/I) 17% (S/N) 69% (F/T) 74% (P/J)
[107]	Binarni	NB	Točnost Točnost	80% (E/I) 60% (S/N, F/T,
[93]	Binarni	LR, Linear SVC	F1-mjera	79.21% (S/N)
[90]	Binarni	LightGBM	makro F1-mjera Točnost AUROC	80.77% (P/J) 82.77% (P/J) 90.08% (P/J)
[89]	Binarni	XGBoost	Točnost F1-mjera	99.92% (S/N) 99.75% (S/N)
[70]	Binarni	Slučajna šuma	Točnost	100% (E/I) (S/N) 100% (F/T) (P/J)
[67]	Binarni	Stacking Boosting	Točnost F1-mjera Odziv	95.79% (S/N) 97.42% (S/N) 96.91% (S/N)

Rad	Pristup	Model	Metrika	Najbolji rezultat
[71]	Binarni	K-Means, Clustering, XGBoost	Točnost	89.01% (E/I)
[91]	Binarni	SVM	Točnost	86.2% (N/S)

Tablica 6 Rezultati višeklasne klasifikacije MBTI

Rad	Pristup	Model	Metrika	Najbolji rezultat
[81]	Binarni	MLP	F1-mjera	47%
[86]	Binarni	Linear/LR	Točnost	45%
[96]	Višeklasni (Big Five)	CNN	Točnost	62.68%
[95]	Višeklasni (Big Five)	NB	Točnost	60%
[72]	Višeklasni	LR	Točnost	66.59%

4. ALGORITAM ZA ODREĐIVANJE TIPA OSOBNOSTI AUTORA TEKSTA TEMELJEN NA DUBOKOM UČENJU

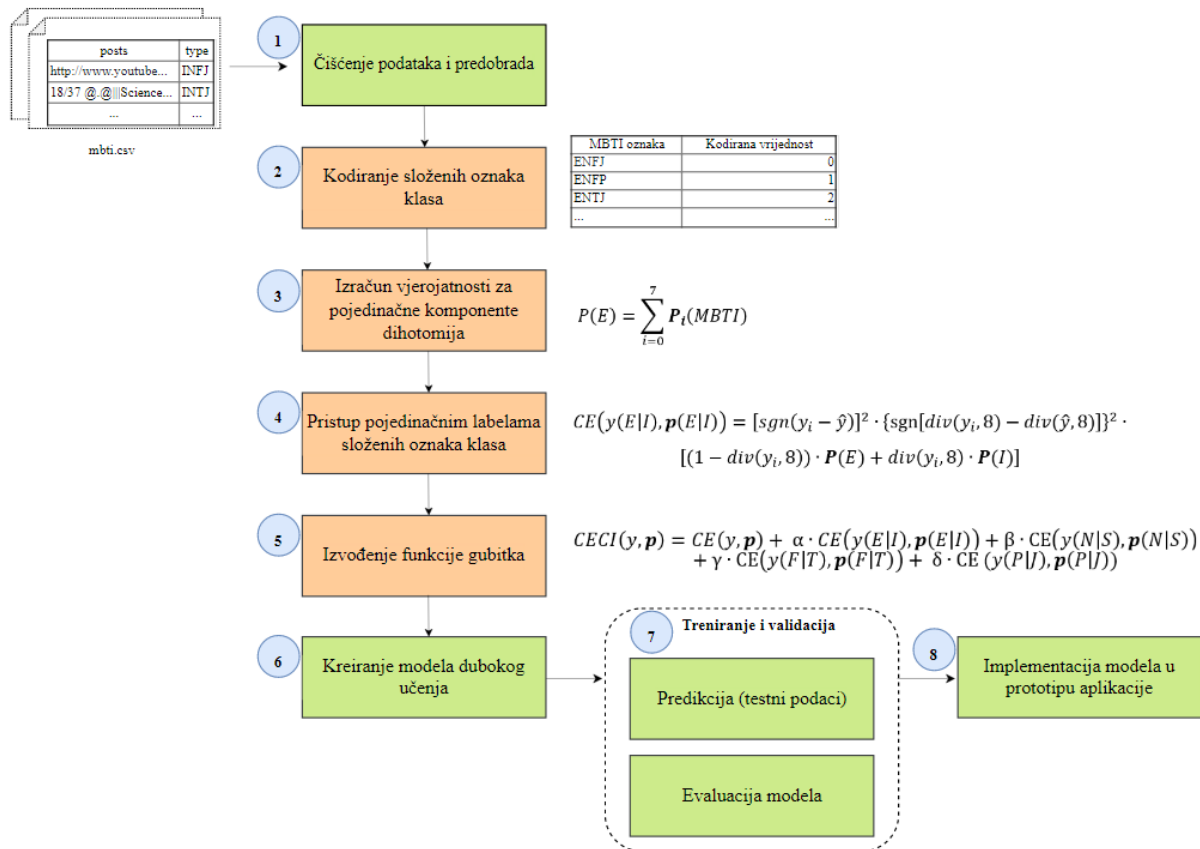
Ovo poglavlje opisuje algoritam za određivanje tipa osobnosti autora teksta temeljen na dubokom učenju s funkcijom gubitka koja u obzir uzima sličnost između složenih oznaka klasa. Ideja algoritma koja koristi višeklasnu klasifikaciju, a uključuje i pojedine komponente koje su pogrešno određene, je penalizirati klasifikaciju složene oznake klase za načinjenu pogrešku klasifikacije pojedine komponente. Npr., ako je ispravna oznaka klase INFJ a model predviđa oznaku klase ENFJ, želimo iskoristiti sličnost složenih oznaka klasa te dodatno penalizirati model za učinjenu pogrešku na dihotomiji E/I, budući da je pored pogreške u klasifikaciji složene klase učinjena pogreška i na toj komponenti. Na taj način omogućavamo modelu da bolje nauči klasificiranje komponente na kojoj je napravljena pogreška u klasifikaciji složene oznake klase.

Koraci algoritma za određivanje tipa osobnosti autora teksta temeljenog na dubokom učenju s funkcijom gubitka koja uzima u obzir sličnost između složenih oznaka klasa su:

1. Čišćenje podataka i predobrada
2. Kodiranje složenih oznaka klasa rastućim slijedom prema engleskoj abecedi
3. Izračun vjerojatnosti za pojedinačnu komponentu dihotomije
4. Pristup pojedinačnim komponentama složenih oznaka klasa
5. Izvođenje funkcije gubitka iz mjera vrednovanja temeljenih na sličnosti složenih oznaka klasa
6. Kreiranje modela dubokog učenja
7. Treniranje i validacija modela dubokog učenja
8. Implementacija modela dubokog učenja u prototipu aplikacije

Slika 20 daje prikaz algoritma za određivanje tipa osobnosti MBTI autora teksta temeljenog na dubokom učenju s funkcijom gubitka koja u obzir uzima sličnost između složenih oznaka klasa. Uobičajeni koraci u procesima klasifikacije u obradi prirodnog jezika su označeni zeleno (1,6,7 , 8). Koraci koji predstavljaju dijelove koji su znanstveni doprinos su označeni narančasto (2, 3, 4, 5) te su detaljno objašnjeni u nastavku poglavlja.

4. ALGORITAM ZA ODREĐIVANJE TIPa OSOBNOSTI AUTORA TEKSTA TEMELJEN NA DUBOKOM UČENJU



Slika 20 Algoritam za određivanje tipa MBTI osobnosti autora teksta

Rješavanje problema klasifikacije MBTI uključuje dva osnovna principa nadziranog strojnog učenja: binarni i višeklasni. Uobičajeni pristup uključuje rješavanje binarnog problema klasifikacije, pri čemu se problem podijeli u četiri binarne klasifikacije. Prvo, uključuje se dodatna značajka za prvu dihotomiju i pridružuje se vrijednost ciljne vrijednosti 0 ili 1. Na taj se način mapiraju dihotomije E i I za implementaciju binarne klasifikacije za prve dvije dihotomije. Ovakav pristup pojednostavljuje problem budući da svako opažanje pripada ili E ili I klasi. Na sličan način se proces ponavlja za ostale tri dihotomije. Konačno, uspjeh ukupne klasifikacije za četiri binarne klasifikacije se računa kombinacijom rezultata pojedinačnih binarnih komponenti.

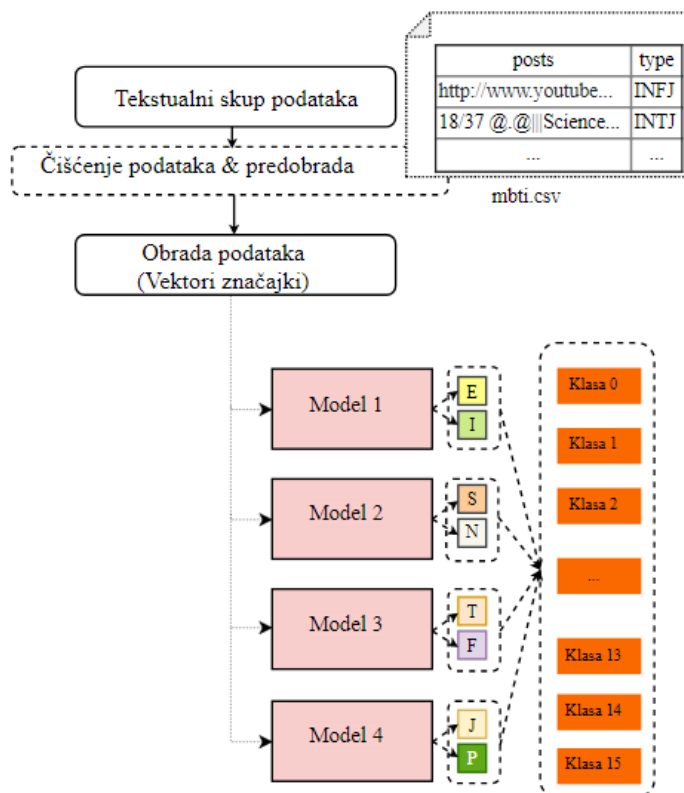
Upravo zbog toga što standardni pristup višeklasne klasifikacije ne daje fleksibilnost kao što daje binarni pristup i daje lošije rezultate u klasifikaciji MBTI, binarni pristup je jednostavniji odabir. Binarni pristup podrazumijeva izdvajanje dihotomija kao dodatnih značajki. Također, budući da se u svakoj klasifikaciji tretiraju dvije dihotomije, tada se i klasifikaciju može lakše prilagoditi kako bi se ostvarili bolji rezultat u ciljanoj metrici. Značajno je da ovakav pristup osigurava i bolje uravnotežene skupove podataka. Binarna klasifikacija nije u fokusu

4. ALGORITAM ZA ODREĐIVANJE TIPA OSOBNOSTI AUTORA TEKSTA TEMELJEN NA DUBOKOM UČENJU

disertacije budući da se veća predikcijska točnost želi postići promatranjem i modeliranjem odnosa između složenih oznaka klasa tijekom treniranja modela kod višeklasne klasifikacije.

Slika 21 daje pregled procesa binarnog pristupa klasifikaciji MBTI. Oznake modela su označene indeksima od 1 do 4 ali se može raditi i o istom modelu koji se koristi za klasifikaciju pojedinih dihotomija. Na početku je potrebno napraviti čišćenje skupa podataka i predobradu te iz riječi generirati vektorski prostor (vektore značajki), koji će biti ulazne vrijednost u modele strojnog učenja. Rezultat binarne predikcije svakog modela je dihotomija za koju je model dao najveću vjerojatnost, a konačni rezultat predstavlja po pitanju vjerojatnosti umnožak vjerojatnosti za pojedinačne binarne predikcije, a po pitanju oznake klase, klasu koja ima oznake prema pojedinačnim binarnim predikcijama.

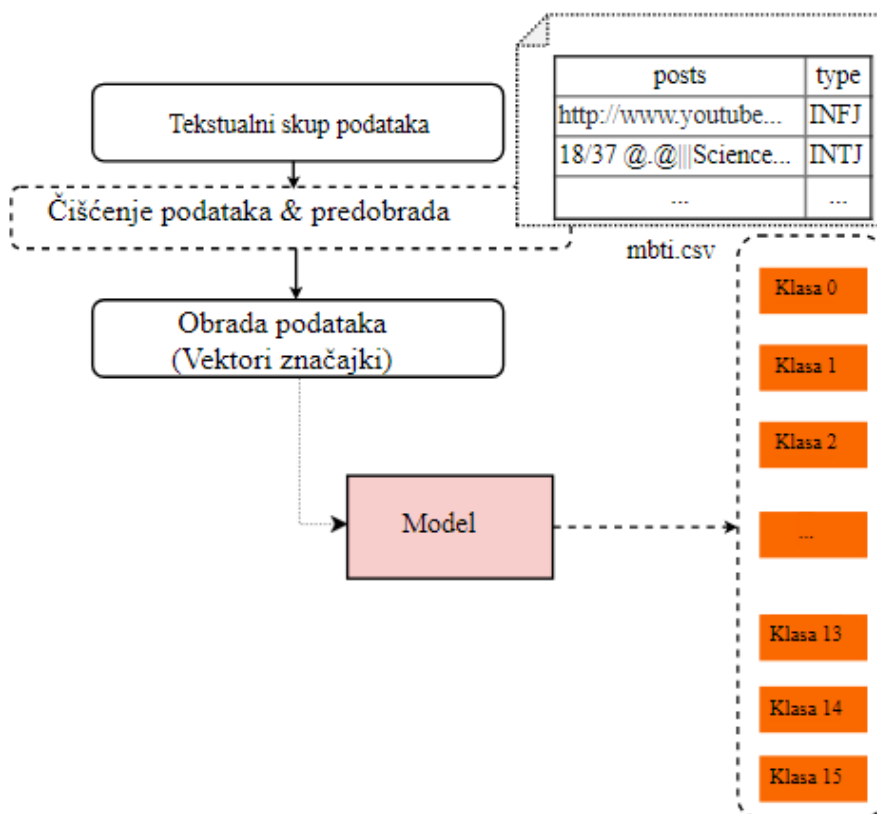
Npr., ako je Model 1 dao predikciju dihotomije E 85%, Model 2 predikciju dihotomije S 25%, Model 3 predikciju dihotomije T 95% a Model 4 predikciju dihotomije J 80%, tada je predikcija klase sa složenim oznakama ENTJ, a vjerojatnost predikcije klase ENTJ je 48.45% ($0.85 \cdot 0.75 \cdot 0.95 \cdot 0.80$).



Slika 21 Binarni pristup klasifikaciji MBTI

Drugi način podrazumijeva višeklasnu klasifikaciju svakog opažanja u jednu od 16 mogućih klasa MBTI sa složenim oznakama. Pristup višeklasne klasifikacije mora riješiti višestruke probleme skupa podataka MBTI, kao što su neuravnoteženost skupa podataka i preklapanje između klasa. Npr., očekivanje klasifikacije je da odabrani model razlikuje klase ESTP i ESTJ kao različite klase, iako se većina vrijednosti klasa preklapaju i razlikuju se samo na zadnjoj dihotomiji. Dodatno, klasifikacija mora uzeti u obzir i ograničenje malog broja primjeraka za obje klase. Ovaj slučaj je odličan primjer motivacije za metodu koja je izložena u disertaciji, a koja može pristupiti dijelovima složenih oznaka klasa. Predložena metoda uključuje komponente klasa u procesu klasifikacije te na taj način efikasno pomaže poboljšanju klasifikacije.

Slika 22 daje pregled procesa višeklasnog pristupa klasifikaciji MBTI. I u ovom pristupu je na početku potrebno napraviti čišćenje skupa podataka i predobradu te iz riječi generirati vektorski prostor (vektore značajki), koji će biti ulazne vrijednost u modele strojnog učenja. Rezultat višeklasne predikcije modela je jedna od 16 klasa za koju je model dao najveću vjerojatnost. Npr., ako je model dao najvišu predikciju za klasu ENTJ tada ta klasa predstavlja predikciju modela višeklasne klasifikacije.



Slika 22 Višeklasni pristup klasifikaciji MBTI

Motivacija predloženog pristupa je uključiti utjecaj komponenti složenih oznaka klasa u algoritam za višeklasnu klasifikaciju MBTI. Na taj način je moguće ublažiti ili pojačati efekte komponenti za koje je napravljena pogreška u klasifikaciji te posljedično poboljšati klasifikaciju složene oznake klase. Dodatno, ovakav pristup osigurava potencijal za dodatna istraživanja, uključujući istraživanje utjecaja kognitivnih funkcija na složene oznake klasa, ali i pojedinačne komponente. Metoda je objašnjena kroz dva dijela. U prvom dijelu je objašnjena tehnika za pristup komponenti složenih oznaka klasa, budući da je to prvi korak koji je potrebno osigurati. U drugom dijelu je opisan postupak uključivanja izračuna vjerojatnosti za ciljanu komponentu te uporaba u predloženoj funkciji gubitka [108].

4.1. Metoda pristupa komponentama složenih oznaka klasa

Budući da uobičajeni postupak višeklasne klasifikacije MBTI podrazumijeva, tijekom pripreme skupa podataka, konverziju oznaka tipova MBTI u cjelobrojne vrijednosti od 0 do 15, potrebno je osigurati uniforman pristup pojedinim komponentama. Taj problem je riješen

pomoću kodiranja sortiranih složenih oznaka klasa. Rezultat takvog pristupa pokazuje Tablica 7.

Tablica 7 Kodiranje oznaka MBTI

Oznaka MBTI	Kodirana vrijednost
ENFJ	0
ENFP	1
ENTJ	2
ENTP	3
ESFJ	4
ESFP	5
ESTJ	6
ESTP	7
INFJ	8
INFP	9
INTJ	10
INTP	11
ISFJ	12
ISFP	13
ISTJ	14
ISTP	15

Složene oznake klasa su sortirane rastućim slijedom prema engleskoj abecedi te je sortiranim klasama pridružena cjelobrojna vrijednost slijedom od 0 do 15. Moguće je uočiti pravilnosti rasporeda pojedinih komponenti oznaka. Npr., za prvu komponentu složene oznake 'E' je na prvih osam mjesta, a 'I' na preostalim osam mjesta. Za drugu komponentu složene oznake moguće je uočiti da se 'N' nalazi na prva četiri mjesta te na mjestima od oznake 8 do 11. Na sličan način je moguće prepoznati predloške pojavljivanja svake komponente. Ovi predlošci ponavljanja imaju dvije ključne uloge: izračun vjerojatnosti za svaku komponentu i određivanje gubitka u odnosu na ispravnu vrijednost komponente te poziciju u složenoj oznaci klase. Predloženi način kodiranja oznaka MBTI predstavlja novost u pristupu u višeklasnoj klasifikaciji MBTI, i koliko je poznato iz dostupnih istraživanja, do sada nije korišten.

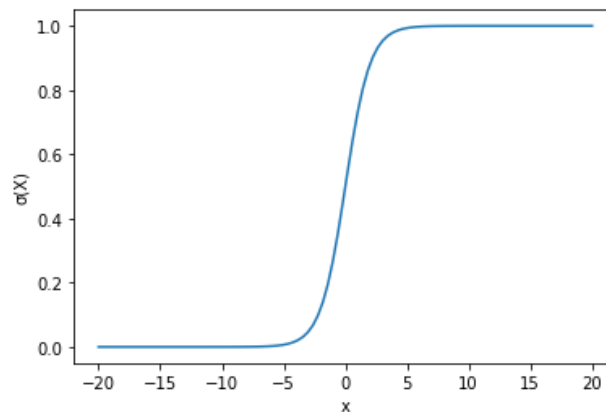
4.2. Izračun vjerojatnosti za pojedinu komponentu

Neuronska mreža u zadacima klasifikacije uobičajeno ima izlazni rezultat u obliku nenormaliziranih, tzv. sirovih vrijednosti (engl. *raw values*), koji se još nazivaju i `logit` vrijednosti. Logit vrijednosti pojedinih neurona u izlaznom sloju kod višeklasne klasifikacije se često označavaju s z_i i određene su izrazom (18). Logit se prosljeđuje sigmoidnoj ili funkciji `softmax`.

Standardnu logističku sigmoidnu funkciju (19) predstavlja slika 23.

$$z_i = \text{logit}(x_i) = w \cdot x_i + b \quad (18)$$

$$\text{sigmoid}(z_i) = \frac{1}{1 + e^{-z_i}} \quad (19)$$



Slika 23 Graf standardne logističke sigmoidne funkcije

Za problem višeklasne klasifikacije MBTI neuronska mreža daje 16 logit vrijednosti, tj. za svaku klasu model daje predviđenu vrijednost. Sljedeći korak u klasifikaciji je uobičajeno uporaba funkcije `softmax` $\sigma(x)$, kojom se logit vrijednosti normaliziraju i pretvaraju u oblik vjerojatnosti. Za višeklasnu klasifikaciju MBTI vrijedi izraz iskazan formulom (20).

$$\hat{y}_i = \sigma(\vec{z})_i = \frac{e^{z_i}}{\sum_{j=0}^{15} e^{z_j}} \quad (20)$$

Logit vrijednosti su predstavljene vektorom $\vec{z}$, a vjerojatnost i komponente vektora s $\sigma(\vec{z})_i$. Suma vjerojatnosti za svih 16 elemenata je jednaka 1.

$$\sum_{i=0}^{15} \sigma(\vec{z})_i = 1 \quad (21)$$

Kako neuronska mreža klasificira 16 klasa MBTI, do vjerojatnosti pojedinih komponenti se može doći na sljedeći način koji uključuje sumiranje svih vjerojatnosti funkcije softmax s pojavljivanjem komponente za koju želimo izračunati vjerojatnost. Npr., za komponente 'E' i 'I' vrijedi:

$$P(E) = P(ENFJ) + P(ENFP) + P(ENTJ) + P(ENTP) + P(ESFJ) + P(ESFP) + P(ESTJ) + P(ESTP) \quad (22)$$

$$P(E) = \sum_{i=0}^7 P_i(MBTI) \quad (23)$$

$$P(I) = P(INFJ) + P(INFP) + P(INTJ) + P(INTP) + P(ISFJ) + P(ISFP) + P(ISTJ) + P(ISTP) \quad (24)$$

$$P(I) = \sum_{i=8}^{15} P_i(MBTI) \quad (25)$$

Izrazi (22) – (25) odgovaraju marginalizaciji združene vjerojatnosti četiri binarne (Bernoullijeve) varijable.

Dodatno, suma vjerojatnosti za binarno isključive komponente mora biti 1.

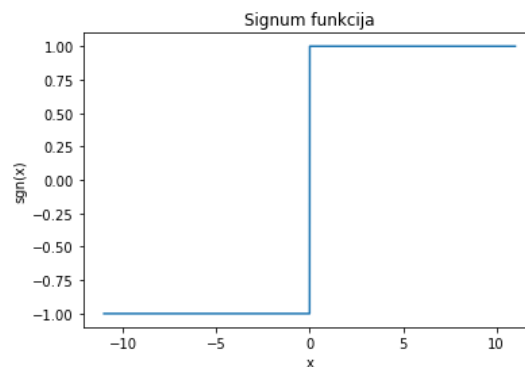
$$P(E) + P(I) = 1 \quad (26)$$

Na sličan način se mogu izračunati vjerojatnosti ostalih komponenti složenih klasa. Način izračuna vjerojatnosti za pojedine dihotomije, prema dostupnim istraživanjima, do sada nije korišten na ovaj način.

4.3. Određivanje ispravne komponente i položaja u složenoj klasi

Prvo je potrebno ocijeniti je li potrebno dodatno penalizirati klasifikaciju na način da se ocijeni je li potrebno uzeti u obzir vjerojatnost funkcije softmax ili ne. U tu svrhu koristimo usvojeni sustav kodiranja oznaka klasa. Izraz (27) daje definiciju, a slika 24 daje graf funkcije Signum koja se koristi u izrazu (28), koji daje kvadratnu vrijednost funkcije Signum (sgn) za razliku vrijednosti koju predviđa model ($\hat{y}$) i ciljane vrijednosti (y_i).

$$\text{sgn}(x) := \begin{cases} -1, & x < 0, \\ 0, & x = 0, \\ 1, & x > 0 \end{cases} \quad \forall x \in \mathbb{R} \quad (27)$$



Slika 24 Graf Signum funkcije

$$[\text{sgn}(y_i - \hat{y})]^2 \quad (28)$$

Izraz će kod svake razlike između polazne klase i predviđene klase dati vrijednost 1. Ako je model ispravno odredio oznaku, vrijednost će biti 0 i neće biti dodatne penalizacije.

Nakon toga je potrebno ocijeniti za svaku oznaku ako postoji razlika. Npr., za procjenu prve komponente sljedeći izraz daje 1 ako postoji razlika i 0 ako ne postoji razlika. U izrazu se koristi oznaka div za cjelobrojno dijeljenje.

$$\{sgn[\text{div}(y_i, 8) - \text{div}(\hat{y}, 8)]\}^2 \quad (29)$$

Vrijednost 8 se pojavljuje u izrazu jer se prema shemi kodiranja prvih 8 složenih oznaka počinje s E a drugih 8 s I. Na kraju je potrebno ocijeniti, ako postoji razlika na pojedinoj komponenti koju vjerojatnost od dvije moguće uzeti za penalizaciju. U tu svrhu koristimo ponovno shemu kodiranja koju prikazuje tablica 7.

Sljedeći izraz (30) daje vjerojatnost za prvu komponentu.

$$[1 - \text{div}(y_i, 8)] \cdot P(E) + \text{div}(y_i, 8) \cdot P(I) \quad (30)$$

Očito je da za prvih osam vrijednosti izraz daje $P(E)$, a za preostalih osam $P(I)$. Na taj način se dolazi do punog izraza za izračun entropije za pogrešno klasificiranu komponentu. Sljedeći izraz daje izračun za prvu dihotomiju E/I:

$$CE(y(E|I), p(E|I)) = [sgn(y_i - \hat{y})]^2 \cdot \{sgn[\text{div}(y_i, 8) - \text{div}(\hat{y}, 8)]\}^2 \cdot [(1 - \text{div}(y_i, 8)) \cdot P(E) + \text{div}(y_i, 8) \cdot P(I)] \quad (31)$$

Važno je naglasiti da je aritmetički izraz za pojedinačne dihotomije nužan budući da funkcija gubitka mora biti diferencijabilna. Na sličan način, polazeći od inicijalnog način kodiranja oznaka, može se doći do odgovarajuće vjerojatnosti za svaku komponentu složene oznake klase MBTI. Način izračuna entropije za pojedine komponente složenih oznaka klase MBTI, prema dostupnim istraživanjima, nije do sada korišten kao pristup. Konačno, sljedeći korak podrazumijeva uključivanje izračunate vjerojatnosti u funkciju gubitka.

4.4. Postupak izvođenja funkcije gubitka iz mjera vrednovanja temeljenih na sličnosti složenih oznaka klasa

Postupak izvođenja funkcije gubitka iz mjera vrednovanja temeljenih na sličnosti složenih oznaka klasa uključuje sljedeće korake:

- određivanje kategoričke unakrsne entropije klasifikacije složene oznake klase
- određivanje kategoričke unakrsne entropije pojedine komponente na kojoj je model napravio pogrešku klasifikacije s uključivanjem težinskih faktora

Pogrešan rezultat višeklasne klasifikacije MBTI ne uključuje samo pogrešku jedne od 16 složenih klasa, nego i pogrešku na minimalno jednoj komponenti složene oznake klase. Npr., za istinitu oznaku klase INTJ i pogrešnu predikciju INTP, model višeklasne klasifikacije ne čini samo pogrešku u klasama sa složenim oznakama nego i pogrešku na pojedinim komponentama. U navedenom primjeru, radi se o pogrešci na zadnjoj komponenti, dihotomiji P/J.

Ideja postupka poboljšanja višeklasne klasifikacije MBTI uključuje činjenicu da su složene oznake klasa slične, u konkretnom primjeru su slične na način da imaju prve tri komponente jednake, ali se zadnja komponenta razlikuje. Sličnost složenih oznaka klasa će se iskoristiti na način da se uključi u izračun pogreške klasifikacije modela ne samo unakrsna entropija složene oznake klase nego i unakrsna entropija pogrešno klasificirane komponente.

4.4.1. Kategorička unakrsna entropija - CE

Standardni pristup višeklasnoj klasifikaciji uključuje izračun kategoričke unakrsne entropije (eng. *categorical cross-entropy* - CE) kao funkciju gubitka. Unakrsna entropija ima korijene u teoriji informacija, a procjenjuje razliku dvije distribucije vjerojatnosti. Koncept entropije pokušava kvantificirati neizvjesnost povezanu sa slučajnim događajem [109]. Dobro uravnoteženi skupovi podataka imaju višu entropiju te je niža vjerojatnost predikcije pojedine klase, a neuravnoteženi skupovi podataka imaju nižu entropiju te izrazito zastupljene klase imaju višu vjerojatnost za predikciju modela.

Entropija mjeri nesigurnost koja je vezana uz distribuciju vjerojatnosti. Za diskretni univarijantni primjer, za slučajnu varijablu X , koja može poprimiti jednu od $x_1, x_2, \dots, x_n$

diskretnih vrijednosti, s vjerojatnostima $p(x_1), p(x_2), \dots, p(x_n)$, entropija je definirana izrazom (32).

$$\mathbb{H}(X) = - \sum_{i=1}^n p(x_i) \cdot \log p(x_i) \quad (32)$$

Entropija ukazuje na ponašanje funkcije gustoće vjerojatnosti te će kod uravnoteženog skupa podataka funkcija gustoće vjerojatnosti pokazati uniformnost vrijednosti, a kod neuravnoteženog skupa podataka funkcija gustoće vjerojatnosti će pokazivati stršće vrijednosti.

Za diskretne distribucije vjerojatnosti, kao što je slučaj i s klasama MBTI, CE se izračunava pomoću izraza:

$$CE(y, \mathbf{p}) = - \sum_{i=1}^N y_i \log(\mathbf{p}_i) = - \log(\mathbf{p}_k) \quad (33)$$

Očekivana distribucija vjerojatnosti je y_i , kao istinita oznaka, p_i je distribucija vjerojatnosti funkcije softmax za i^{th} klasu, a p_k je distribucija vjerojatnosti funkcije softmax za k^{th} klasu. Vjerojatnost koja se odnosi na istinitu oznaku je jednaka 1 za one-hot encoding vektor. Drugim riječima, kodiramo ciljanu distribuciju vjerojatnosti s vrijednosti 1 za indeks k i 0 za ostale. Model klasifikacije aproksimira ciljanu distribuciju vjerojatnosti i kategorička unakrsna entropija daje ukupnu entropiju između dvije distribucije.

4.4.2. Unakrsna entropija s utjecajem komponenti složenih oznaka klasa - CECI

Neuravnoteženi skupovi podataka, kao što je skup podataka MBTI imaju nejednaku distribuciju vjerojatnosti i nisku entropiju jer najzastupljenije klase imaju prevagu. Modeli višeklasne klasifikacije implementiraju CE zbog brzog i jednostavnog izračuna, a bitno je imati na umu da CE uzima u obzir samo distribuciju aktualne klase. Drugim riječima, CE ne uzima u obzir distribuciju drugih klasa. Funkcija gubitka koja uzima u obzir i druge distribucije za koje postoji neispravno klasificirana komponenta, nazvana unakrsna entropija s utjecajem komponenti složenih oznaka klasa (engl. *cross-entropy compound class-label*

impact - *CECI*), uključuje težinske faktore za neispravno klasificirane komponente klasa i predstavljena je izrazom:

$$CECI(y, \mathbf{p}) = CE(y, \mathbf{p}) + \alpha \cdot CE(y(E|I), \mathbf{p}(E|I)) + \beta \cdot CE(y(N|S), \mathbf{p}(N|S)) + \gamma \cdot CE(y(F|T), \mathbf{p}(F|T)) + \delta \cdot CE(y(P|J), \mathbf{p}(P|J)) \quad (34)$$

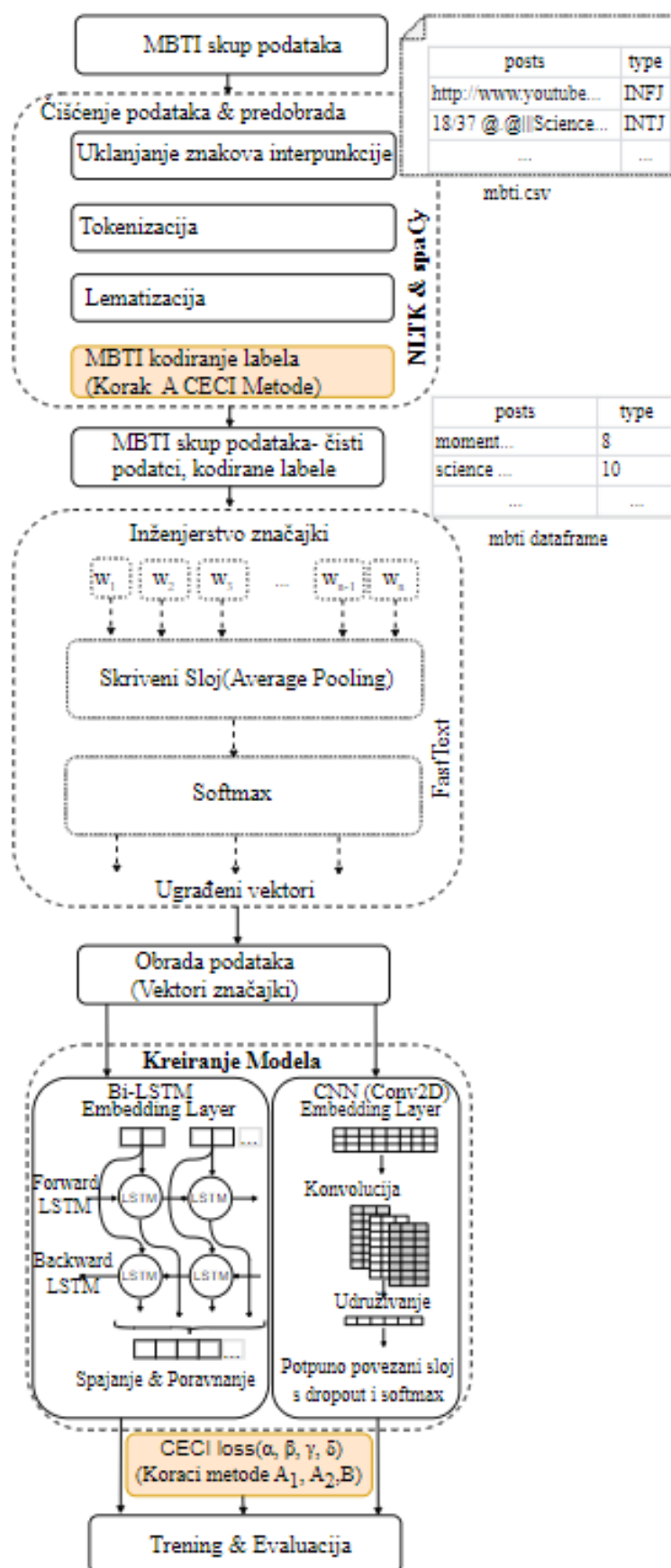
α , β , γ , i δ su težinski faktori s odgovarajućim funkcijama unakrsne entropije za svaku komponentu, ovisno o odgovarajućoj dihotomiji, a time i poziciji. Težinske faktore je moguće trenirati kako bi se pronašle vrijednosti koje poboljšavaju rezultate višeklasne klasifikacije. S obzirom na vrijednosti težinskih faktora, napravljeno je intenzivno testiranje i odabrani su rezultati s obzirom na najviše vrijednosti F1-mjere vrijednosti i odziva, s vrijednostima težinskih faktora između 0 i 1.

Izraz (22) se sastoji iz dva ključna dijela. Prva dio je izraz (33) s tradicionalnom formulom kategoričke unakrsne entropije, koja je u prijašnjim istraživanjima korištena za višeklasnu klasifikaciju pomoću dubokog učenja. Drugi dio izraza s α , β , γ , i δ težinskim faktorima za odgovarajuće funkcije unakrsne entropije za svaku komponentu, ovisno o odgovarajućoj dihotomiji, a time i poziciji, se predlaže u okviru disertacije.

Slika 25 daje pregled implementacije metode CECI u postupku višeklasne klasifikacije MBTI. U prvom dijelu se čiste i predprocesiraju podaci u skupu podataka. Važan korak metode CECI je korak kodiranja oznaka MBTI. Nakon toga se procesiraju značajke, a rezultat su ugrađeni vektori riječi. Potom se kreiraju dva modela arhitekture bi-LSTM i CNN.

Cilj disertacije nije bio pronaći idealnu arhitekturu za klasifikaciju, kao u [41], nego dokazati da metoda CECI poboljšava rezultate s različito odabranim arhitekturama. Dodatno, kako mreže LSTM mogu prepoznati predloške kroz vrijeme, a mreže CNN kroz prostor, težinski faktori mogu dati uvid u ponašanje komponenti složenih oznaka klasa MBTI. Na kraju su trenirani i evaluirani modeli, uporabom funkcije gubitka CECI.

4. ALGORITAM ZA ODREĐIVANJE TIPa OSOBNOSTI AUTORA TEKSTA TEMELJEN NA DUBOKOM UČENJU



Slika 25 Pregled procesa implementacije metode CECI

5. POSTAVKE I REZULTATI TRENIRANJA PRIMJENOM PREDLOŽENOG ALGORITMA

Postavkama istraživanja su definirani hardverska i softverska okruženja koja su se koristila za predobradu i treniranje modela. Rezultati istraživanja obuhvaćaju rezultate za obje razmatrane arhitekture umjetne neuronske mreže: LSTM i CNN, klasičnu funkciju gubitka kategoričke unakrsne entropije, funkciju gubitka CECI te usporedbu prethodnih istraživanja.

5.1. Postavke treniranja

Za istraživanje su korištena dva okruženja. Prvo okruženje je namijenjeno kreiranju prototipa i preliminarnom testiranju, a činili su ga: Windows 10 operativni sustav, Python 3.8.5, Jupyter Notebook i Python skripte, PyTorch 1.8.1, cuda 10.2 i torchtext 0.9.1. Kao grafička procesorska jedinica (engl. *graphical processing unit - GPU*) korišten je GeForce GTX 1050. Drugo okruženje je Google Colab Pro+ i to je okruženje korišteno za konačno testiranje i rezultate koji te su prezentirani u disertaciji.

Za pripremu skupova za testiranje i validaciju su korištene opcije PyTorch okvira. Korištena je i opcija stratifikacije. Dodatno, kako bi se osigurala ponovljivost rezultata na platformi za treniranje, korištena je opcija postavka seed vrijednosti, opcije deterministic i benchmark opcija. Vrijednost za batch opciju je bila 256 za treniranje te 64 za validaciju. Kao iterator je upotrijebljena ugrađena opcija BucketIterator s vrijednosti False za opciju sortiranja i vrijednosti True za opciju sort_within_batch_option.

5.2. Skupovi podataka

Dostupno je nekoliko skupova podataka MBTI, kao što je korpus od 1.2 milijuna tvitova na engleskom jeziku te su anotirani od strane svakog sudionika [83], skupovi podataka MBTI9k i PANDORA s Reddit društvene mreže [81] [86] te korpus s tipovima autora prema modelu MBTI za šest zapadnoeuropskih jezika [93]. Potrebno je istaknuti skup podataka PANDORA, kao prvi skup podataka koji pokriva više modela procjene osobnosti (Big Five, MBTI, Enneagram) te uključuje demografske podatke, što skupovi podataka redovito nemaju kao značajke. Istraživanje je napravljeno na skupu podataka MBTI s Kaggle platforme [110] te skupu podataka MBTI9k s Reddit društvene mreže [81].

Prvi od odabranih skupova podataka se nalazi na Kaggle platformi i sadrži 8.675 redaka, sakupljenih na Personality Cafe forumu 2017. Podaci čine objave korisnika foruma na engleskom jeziku te su označeni tipovima osobnosti prema modelu MBTI. Oznaka tipa osobnosti prema modelu MBTI nije rezultat anketiranja ili nekog testa nego je svaki korisnik foruma unio uz tekst i oznaku svog tipa osobnosti. Skup podataka čini korpus od 11.2 milijuna riječi i više od 420.000 oznaka. Svaki redak predstavlja zadnjih 50 objava svakog korisnika. Slika 26 prikazuje izvadak nekoliko redaka iz skupa podataka MBTI koji je korišten za eksperimentalna istraživanja. Skup podataka čini datoteka u .csv formatu i sadrži dvije kolone: tekst objave na forumu i složenu oznaku tipa MBTI. Očekivani izlazni rezultat treniranog modela je predikcija tipa MBTI za predani tekst. Diskusija korisnika na forumu određuje tip osobnosti MBTI [66]. Slika 26 prikazuje izvadak nekoliko redaka iz ovog skupa podataka MBTI koji je korišten za eksperimentalna istraživanja. Skup podataka čini datoteka u .csv formatu i sadrži dvije kolone: tekst objave na forumu i složenu oznaku tipa MBTI. Očekivani izlazni rezultat treniranog modela je predikcija tipa MBTI za predani tekst.

type	posts
INFJ	'http://www.youtube.com/watch?v=qsXHcwe3krw ...
ENTP	'I'm finding the lack of me in these posts ver...
INTP	'Good one ____ https://www.youtube.com/wat...
INTJ	'Dear INTP, I enjoyed our conversation the o...
ENTJ	'You're fired. That's another silly misconce...
INTJ	'18/37 @.@ Science is not perfect. No scien...
INFJ	'No, I can't draw on my own nails (haha). Thos...

Slika 26 Izvadak iz skupa podataka MBTI s Personality Cafe foruma

Drugi od korištenih skupova podataka je MBTI9k verzija skupa podataka na osnovu diskusijskih grupa posvećenih modelu MBTI s Reddit platforme. Oznake tipa MBTI pojedinog autora objave su kreirali sami autori objave. Skup podataka sadrži 9252 retka s komentarima korisnika koji sadrže više od 1000 riječi u komentaru. Prilikom kreiranja ovog skupa podataka, izostavljeni su komentari s temama koje su povezane s temom predikcije

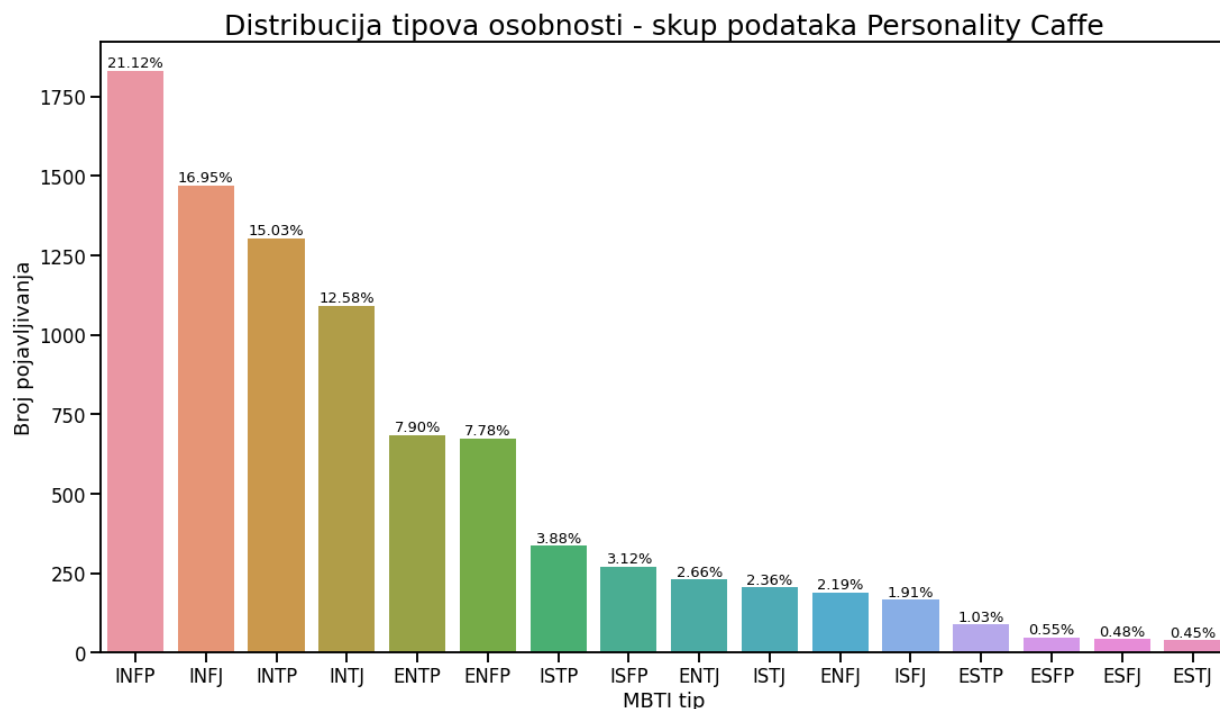
osobnosti pomoću modela MBTI. Također, prilikom kreiranja skupa podataka su izostavljena eksplicitna pojavljivanja složenih oznaka klasa MBTI [81].

Slika 27 prikazuje izvadak nekoliko redaka iz skupa podataka MBTI s Reddit platforme koji je korišten za eksperimentalna istraživanja. Skup podataka čini datoteka u .csv formatu i sadrži šest kolona (`author`, `comment`, `type`, `subreddits_commented`, `mbti_subreddits_commented`, `wc`, `comments_num`) od kojih su upotrijebljene dvije kolone, identične kao u prethodnom skupu: tekst objave na forumu i složena oznaka tipa MBTI. Očekivani izlazni rezultat treniranog modela je predikcija tipa MBTI za predani tekst.

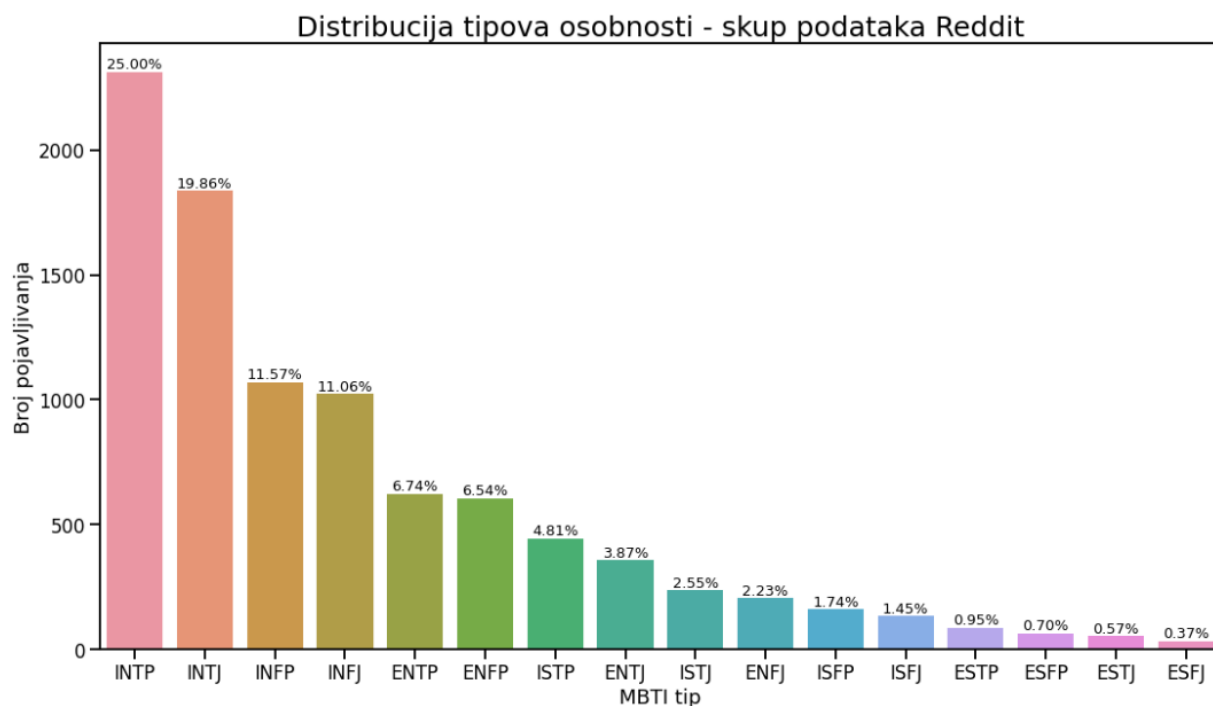
type	posts
ISTP	My PC has been running this game perfectly for...
ENTJ	Where in Utah are you living now? I also have ...
INTP	Guy sitting behind the rail up top is not impr...
ENTP	Oh I know, he just didn't seemed to find his I...
ISTP	I would say I hate the cite. There are very fe...
INFP	That's true but maybe they could breed once re...
INTJ	This happens in many families. It's not right,...

Slika 27 Izvadak iz skupa podataka MBTI s Reddit platforme

Slika 28 daje prikaz distribucije klasa MBTI unutar skupa podataka s Personality Cafe foruma, a Slika 29 daje prikaz distribucije klasa MBTI unutar skupa podataka MBTI s Reddit platforme. Vizualni prikaz ukazuje izrazitu neuravnoteženost skupa podataka. Ta činjenica dodatno otežava klasifikaciju klasičnim pristupima klasifikacije budući da takvi pristupi preferiraju uravnotežene skupove podataka.



Slika 28 Distribucija klasa u skupu podataka MBTI – Personality Cafe



Slika 29 Distribucija klasa u skupu podataka MBTI – Reddit

Tablica 8 daje broj pojavljivanja za svaki tip osobnosti i postotak pojavljivanja svakog tipa u odnosu na ukupan broj redaka u oba korištena skupa podataka. Npr., prva četiri tipa MBTI čine 65.67% ukupnog broja redaka u skupu podataka s Personality Cafe foruma a iste klase

čine 67.49% ukupnog broja redaka u skupu s Reddit platforme te se ove klase može smatrati većinskim klasama (engl. *majority*). Prikazana je i relativna frekvencija tipova MBTI u populaciji stanovništva u Sjedinjenim Američkim Državama [111].

Tablica 8 Distribucije klasa skupova podataka s Reddit i Personality Cafe platformi

MBTI	Personality Cafe		Reddit		Procjena/ populacija SAD-a
	#	%	#	%	
INFP	1832	21.12	2313	25.00	4.4% (4-5%)
INFJ	1470	16.95	1837	19.86	1.5% (1-3%)
INTP	1304	15.03	1070	11.57	3.3% (3-5%)
INTJ	1091	12.58	1023	11.06	2.1% (2-4%)
ENTP	685	7.90	624	6.74	3.2% (2-5%)
ENFP	675	7.78	605	6.54	8.1% (6-8%)
ISTP	337	3.89	445	4.81	5.4% (4-6%)
ISFP	271	3.12	358	3.87	8.8% (5-9%)
ENTJ	231	2.66	236	2.55	1.8% (2-5%)
ISTJ	205	2.36	206	2.23	11.6% (11-14%)
ENFJ	190	2.19	161	1.74	2.5% (2-5%)
ISFJ	166	1.91	134	1.45	13.8% (9-14%)
ESTP	89	1.03	88	0.95	4.3% (4-5%)
ESFP	48	0.55	65	0.70	8.5% (4-9%)
ESFJ	42	0.48	53	0.57	12.3% (9-13%)
ESTJ	39	0.45	34	0.37	8.7% (8-12%)

Može se uočiti mala razlika neuravnoteženosti između odabranog skupa podataka za pojedine dihotomije (najviše su izražene za I, E i N) i procjene distribucije klasa populacije SAD-a koja se u velikoj mjeri može smatrati reprezentativnom za populaciju SAD-a kao takvu. Ova razlika je prirodna jer se upravo zbog tipova osobnosti može očekivati da će među korisnicima foruma biti najzastupljenije introvertirane osobe a najmanje ekstrovertirane osobe. Tome govori u prilog činjenica da su četiri najzastupljenija složena tipa osobnosti introverti a četiri najmanje zastupljena tipa osobnosti ekstroverti. Promatrajući dihotomije na cijelom skupu složenih oznaka klasa, dihotomija E je višestruko više zastupljena u odnosu na dihotomiju I. Stoga, usprkos neuravnoteženosti, možemo smatrati da su odabrani skupovi reprezentativni i dobro oslikavaju stvarnu distribuciju osobnosti za populaciju SAD-a kada se uzme u obzir metodologija njihovog prikupljanja.

Moguće je uočiti razliku procjene i vrijednosti postotka pojedinačnog tipa MBTI u odnosu na postotak u skupu podataka, što može biti predmet budućih istraživanja. Također, činjenica da su podaci ograničeni na određeni forum može utjecati na uzorkovanje aktualne populacije. Korisno je promatrati i distribucije četiri dihotomije budući da takva usporedba može dati uvide koji mogu usmjeriti eksperimente s težinskim faktorima za pojedinačne komponente. Npr., moguće je razmotriti korelaciju težinskog faktora dihotomije i frekvencije dihotomije u populaciji. Dodatno, na samu distribuciju klasa MBTI u promatranom skupu podataka mogu utjecati i sklonosti pojedinih tipova MBTI osobnosti prema forumima kao što je Personality Cafe forum.

Tablica 9 daje pregled distribucija dihotomija u skupu podataka MBTI. Moguće je uočiti da su skupovi dihotomija mnogo bolje uravnoteženi, što je značajan argument za dosadašnja istraživanja koja su se temeljila prvenstveno na metodi binarne klasifikacije pojedinih dihotomija. Informacije će biti smjernica za buduća istraživanja u uporabi metode CECI. U usporedbi raspodjele dihotomija unutar dva odabrana skupa mogu se uočiti razlike u raspodjeli dihotomija T i F. Preostale dihotomije imaju vrlo male razlike u raspodjelama i među njima prevladavaju dihotomije I, N i P, što ukazuje da su korisnici i Personality Cafe foruma te Reddit platforme predominantno introvertirani, intuitivni i perceptivni. Npr., skup podataka s Personality Cafe foruma sadrži 54.11% uzoraka u kojima prevladava dihotomija F te 45.89% uzoraka u kojima prevladava dihotomija T. Skup podataka s Reddit platforme ima 35.65% uzoraka u kojima prevladava dihotomija F i 64.35% uzoraka u kojima prevladava dihotomija T.

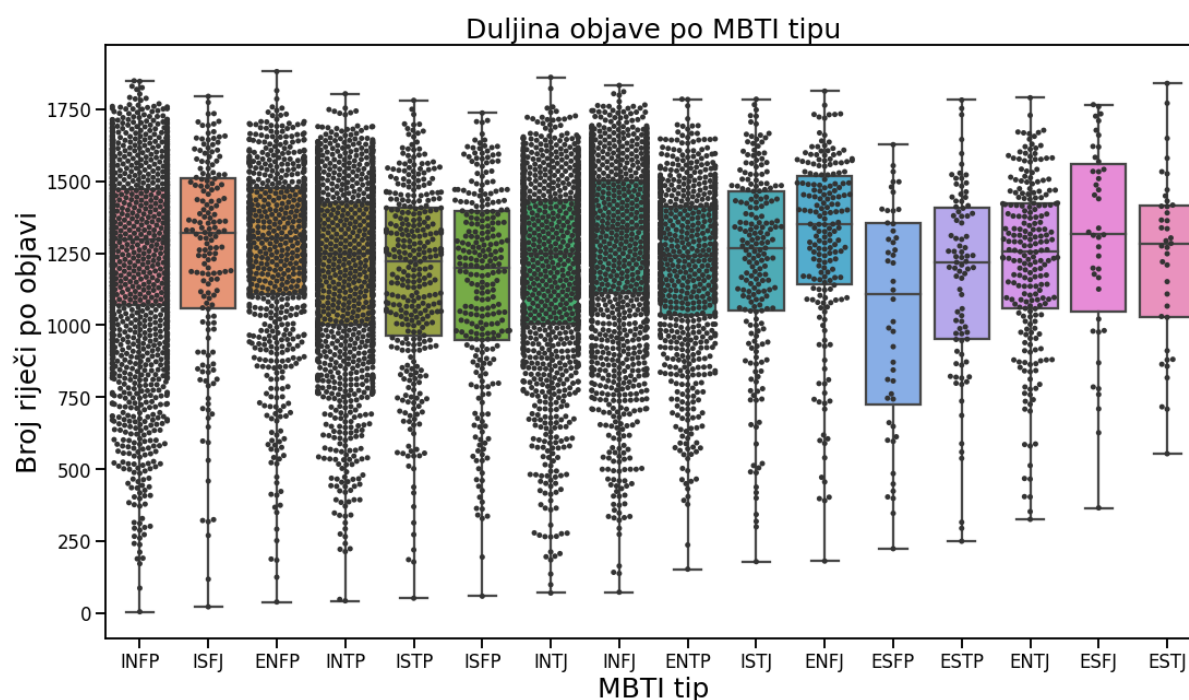
Tablica 9 Distribucija dihotomija u skupu podataka MBTI

Dihotomija	Personality Cafe		Reddit		Procjena/populacija SAD-a
	#	%	#	%	
E	1999	23.04	2033	21.97	49.3% (45-53%)
I	6676	76.96	7219	78.03	50.7% (47-55%)
S	1197	13.80	1216	13.14	73.3% (66-74%)
N	7478	86.20	8036	86.86	26.7% (26-34%)
T	3981	45.89	5954	64.35	40.2% (40-50%)
F	4694	54.11	3298	35.65	59.8% (50-60%)
J	3434	39.59	3881	41.95	54.1% (54-60%)
P	5241	60.41	5371	58.05	45.9% (40-46%)

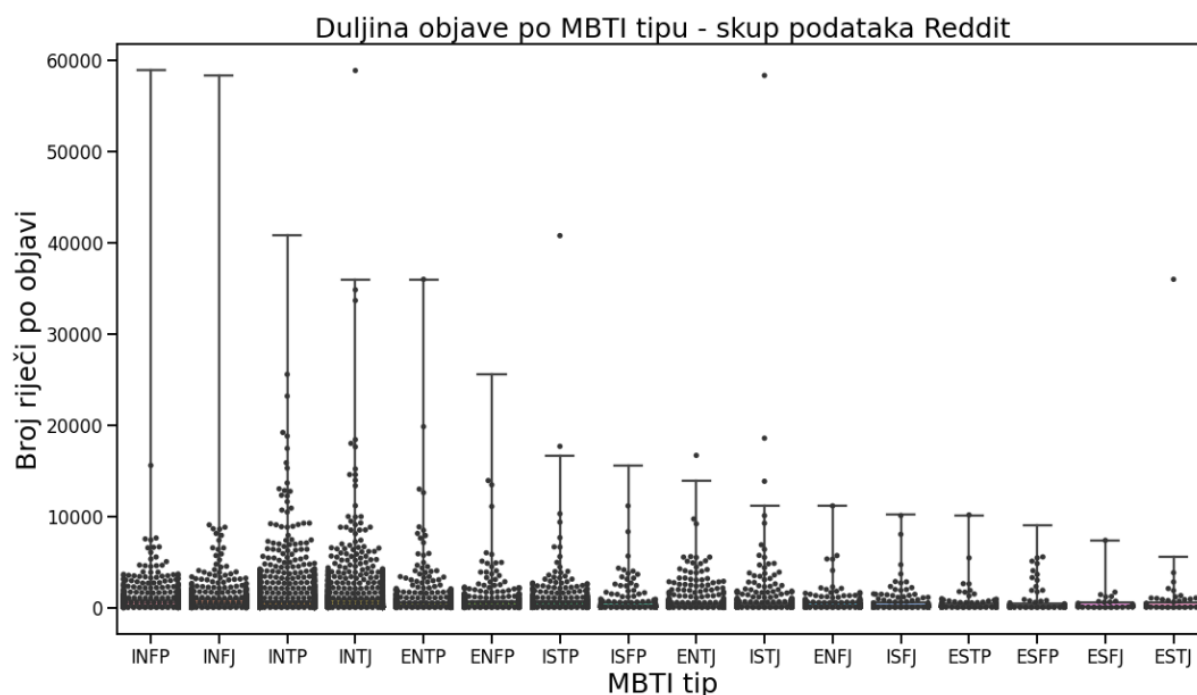
Pored analize frekvencije tipova MBTI korisno je analizirati broj riječi po objavi i tipu osobnosti MBTI. Tablica 10, slika 30 i slika 31 daju pregled statističkih vrijednosti skupa podataka te pregled duljine objava po tipu MBTI. Pearsonova korelacijska analiza između broja riječi po objavi i elipsi po komentaru ukazuje da postoji korelacija između riječi po komentaru i elipsi po komentaru za cijeli skup podataka. Najveća korelacija je za ENFP, INFJ i INTP tipove MBTI [70]. Interesantno je uočiti da je tip MBTI koji ima maksimalni prosječan broj riječi po objavi u skupu podataka s Personality Cafe foruma ESFJ (drugi najniži broj pojavljivanja) a u u skupu podataka s Reddit platforme ESTJ (najniži broj pojavljivanja).

Tablica 10 Statistika skupova podataka

Statistika	Personality Cafe		Reddit	
	Vrijednost	Klasa	Vrijednost	Klasa
Maksimalni broj riječi po objavi	1881	ENFP	58845	INTJ
Minimalni broj riječi po objavi	4	INFP	20	INFJ
Maksimalni prosječan broj riječi po objavi	1287	ESFJ	1214	ESTJ
Minimalni prosječan broj riječi po objavi	1022	ESFP	517	INFJ
Prosječan broj riječi po objavi	1226		712	
Prosječna varijanca broja riječi po objavi	137.21		481.19	



Slika 30 Duljina objave po tipu MBTI– Personality Cafe



Slika 31 Duljina objave po tipu MBTI– Reddit

Instrument MBTI za procjenu tipa osobnosti sadrži 16 različitih složenih oznaka. Svaka oznaka se sastoji od 4 komponente koje predstavljaju jednu od dvije vrijednosti dihotomije za svaku poziciju. Tako prva pozicija tipa MBTI odgovara vrijednostima E(extrovert) ili I(introvert), druga pozicija odgovara vrijednostima N(intuitive) ili S(sensitive), treća pozicija odgovara vrijednostima T(thinking) ili F(feeling), a četvrta odgovara vrijednostima J(judging) ili P(perceive). S obzirom na strukturu tipova MBTI, problemu klasifikacije se može pristupiti kao problemu četiri binarne klasifikacije, klasifikacije više oznaka ili višeklasne klasifikacije. S obzirom na cilj disertacije, odabrana je višeklasna klasifikacija.

Nad oba skupa podataka su napravljeni identični standardni koraci predobrade prije kreiranja modela neuronske mreže. Npr., uklonjeni su brojevi, specijalni znakovi, poveznice i znakovi interpunkcije. Osigurano je da su sve riječi navedene malim slovima. Zaustavne riječi (engl. *stop words*) nisu isključene, budući da su zaustavne riječi dobri prediktori tipa osobnosti. Iz teksta su uklonjeni nazivi klasa po modelu MBTI, kako bi se povećala generalizacijska moć modela dubokog učenja. Tekst je u konačnici transformiran pomoću FastTexta u ugradbene matrice vektora riječi.

5.3. Rezultati treniranja

Istraživanje je napravljeno s dva modela dubokog učenja: LSTM i CNN. Korištena je bi-LSTM opcija mreže LSTM te 2-dimenzionalni konvolucijski model. Nad svakim modelom je napravljeno treniranje sa standardnom CE kao funkcijom gubitka, a nakon toga s CECI kao funkcijom gubitka te je napravljena usporedba. Prilikom uporabe funkciju gubitka CECI napravljeni su eksperimenti s vrijednostima za težinske faktore. Eksperimenti su postavljeni na način da je lako napraviti usporedbu rezultata treniranja uporabom CE i funkcija gubitka CECI. Uz pristup u radu [108], gdje su prilikom odabira težinskih faktora α , β , γ , δ eksperimentalne postavke određene za vrijednosti između 0 i 1 uz korak promjene vrijednosti 0.05, kako bi se ograničila računalna kompleksnost, disertacija obrađuje i pristup u kojem su faktori α , β , γ , δ parametri koje model uči tijekom postupka treniranja. Potom je napravljena usporedba rezultata istraživanja s rezultatima dosadašnjih istraživanja u ovom području. Skup podataka je podijeljen na trening, validacijski i testni skup podataka u omjeru 70% : 15%: 15%.

Rezultati su evaluirani uporabom različitih metrika, kao što su F1- mjera, preciznost, točnost, odziv i matrica zabune. Pri odabiru metrika se vodilo računa da se radi o neuravnoteženom skupu podataka te da je izuzetno bitno odabrati metrike koje daju ispravno tumačenje za takav skup podataka. Npr., F1-mjera pokazuje ravnotežu između odziva i preciznosti, što je izuzetno važno za neuravnotežene skupove podataka.

5.3.1. Model LSTM

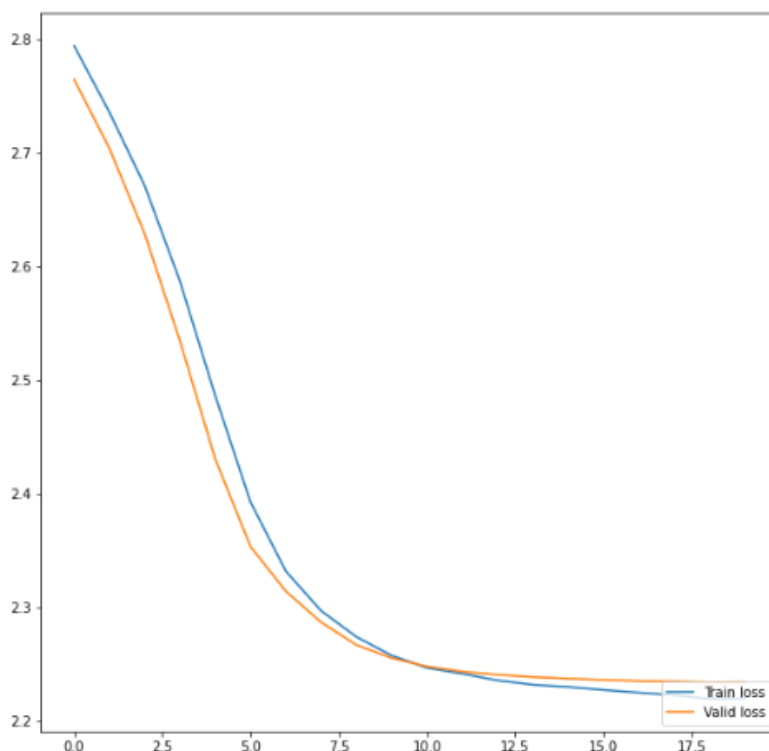
Slika 32 prikazuje arhitekturu mreže LSTM koja je korištena za treniranje. Može se uočiti da su korištena dva sloja u mreži s 25 neurona. Također, kao mehanizam regularizacije je korišteno nasumično ispuštanje (engl. *dropout*) s vrijednosti 0.4. Dimenzija ugrađenog vektora je 25002. Model je treniran s 20 epoha. Broj epoha je određen eksperimentalno, promatrajući ponašanje funkcije gubitka na validacijskom skupu, u odnosu na ponašanje funkcije gubitka na skupu za treniranje.

```
LSTM(  
  (embedding): Embedding(25002, 25, padding_idx=1)  
  (rnn): LSTM(25, 25, num_layers=2, dropout=0.4, bidirectional=True)  
  (fc1): Linear(in_features=50, out_features=25, bias=True)  
  (fc2): Linear(in_features=25, out_features=16, bias=True)  
  (dropout): Dropout(p=0.4, inplace=False)  
)
```

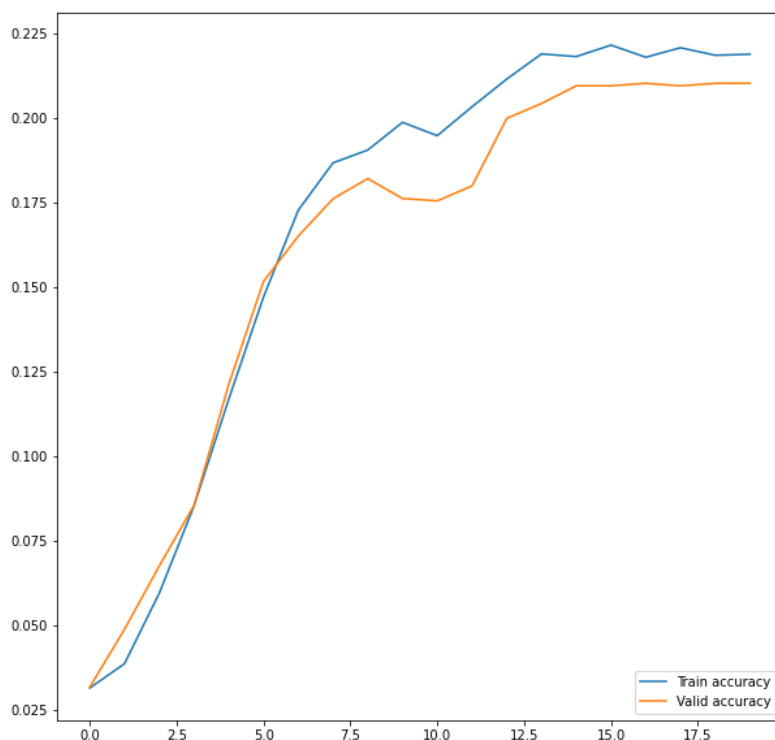
Slika 32 Model LSTM

Model LSTM je prvo treniran sa standardnom funkcijom gubitka CE, kako bi se dobio rezultat koji predstavlja osnovu za usporedbu kod višeklasne klasifikacije s modelom LSTM. Nakon toga je isti model treniran s funkcijom CECI. Treniranje modela LSTM s funkcijama CE i CECI je ponovljeno nad oba skupa podataka.

Model LSTM je prvo treniran na skupu s Personality Cafe foruma, koji je preuzet s Kaggle platforme. Gubitak treniranja i validacije te točnost treniranja i validacije za funkciju gubitka CE daju slika 33 i slika 34. Rezultat višeklasne klasifikacije uporabom mreže LSTM i klasične funkcije gubitka CE daju tablica 11 i slika 35.



Slika 33 Gubitak treniranja/validacije za LSTM-CE na skupu MBTI s Personality Cafe foruma



Slika 34 Točnost treniranja/validacije za LSTM-CE na skupu MBTI s Personality Cafe foruma

Budući da se radi o izrazito neuravnoteženom skupu podataka, rezultati klasifikacije standardnom CE kao funkcijom gubitka su očekivano niski. Model uči vrlo malo i brzo ulazi u prenaučenosť. Također, rezultati su usporedivi s rezultatima postojećih istraživanja višeklasne klasifikacije s CE funkcijom gubitka.

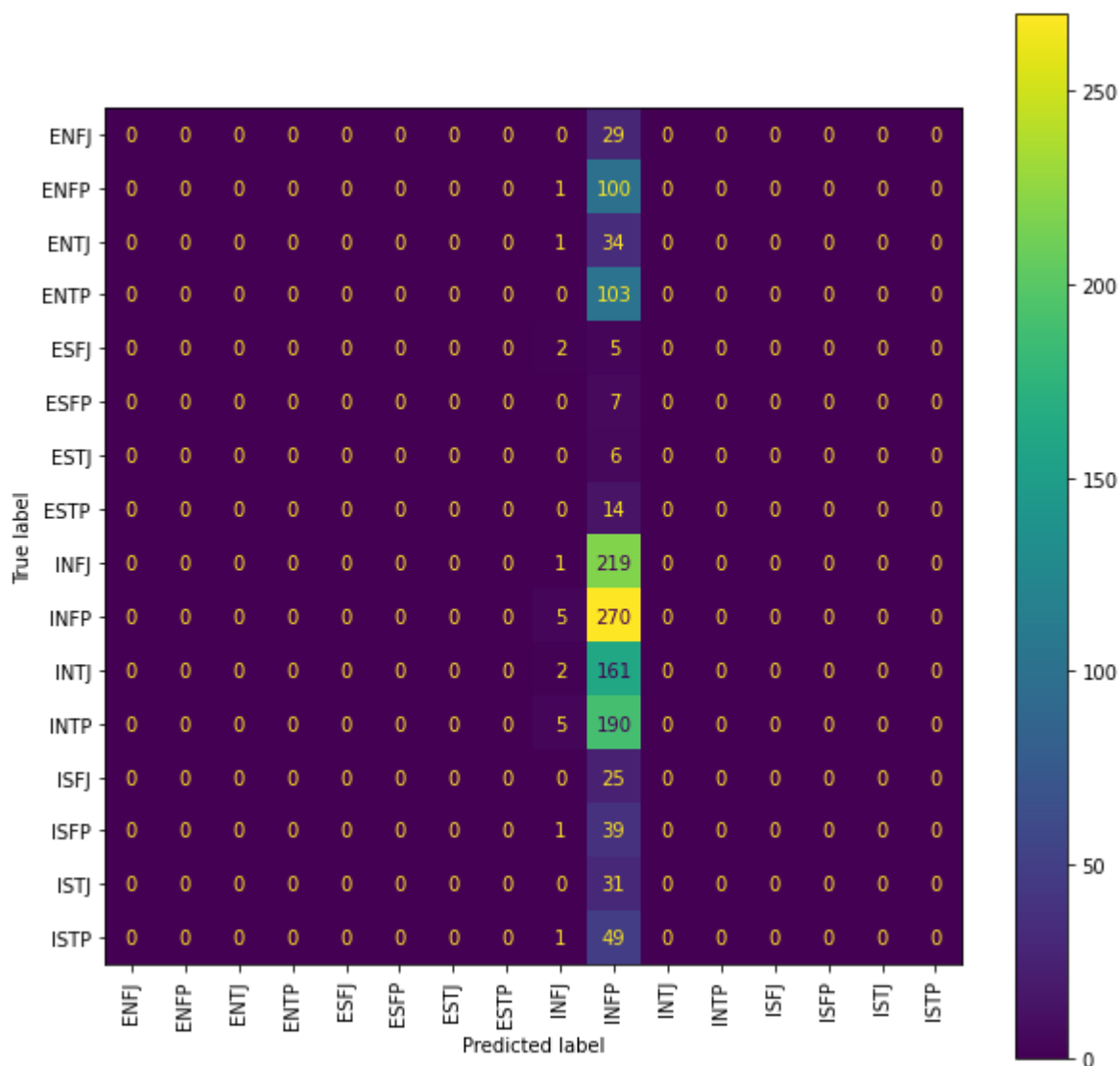
Klasifikacijski izvještaj prikazuje tablica 11. Kako se radi o neuravnoteženom skupu podataka, F1-mjera je važna metrika za usporedbu jer ukazuje da klasifikator dobro klasificira svaku individualnu klasu, a ne samo većinski zastupljene klase. Težinska srednja vrijednost F1- mjere s 0.07 te makro srednja vrijednosti F1-mjere 0.02 su očekivano niske vrijednosti, budući da se radi o izrazito nebalansiranom skupu podataka a model preferira klase koje su najzastupljenije u skupu podataka.

Tablica 11 Klasifikacijski izvještaj za LSMT-CE na skupu MBTI s Personality Cafe foruma

MBTI	preciznost	odziv	F1-mjera	podrška
ENFJ	0.00	0.00	0.00	29
ENFP	0.00	0.00	0.00	101
ENTJ	0.00	0.00	0.00	35
ENTP	0.00	0.00	0.00	103
ESFJ	0.00	0.00	0.00	7

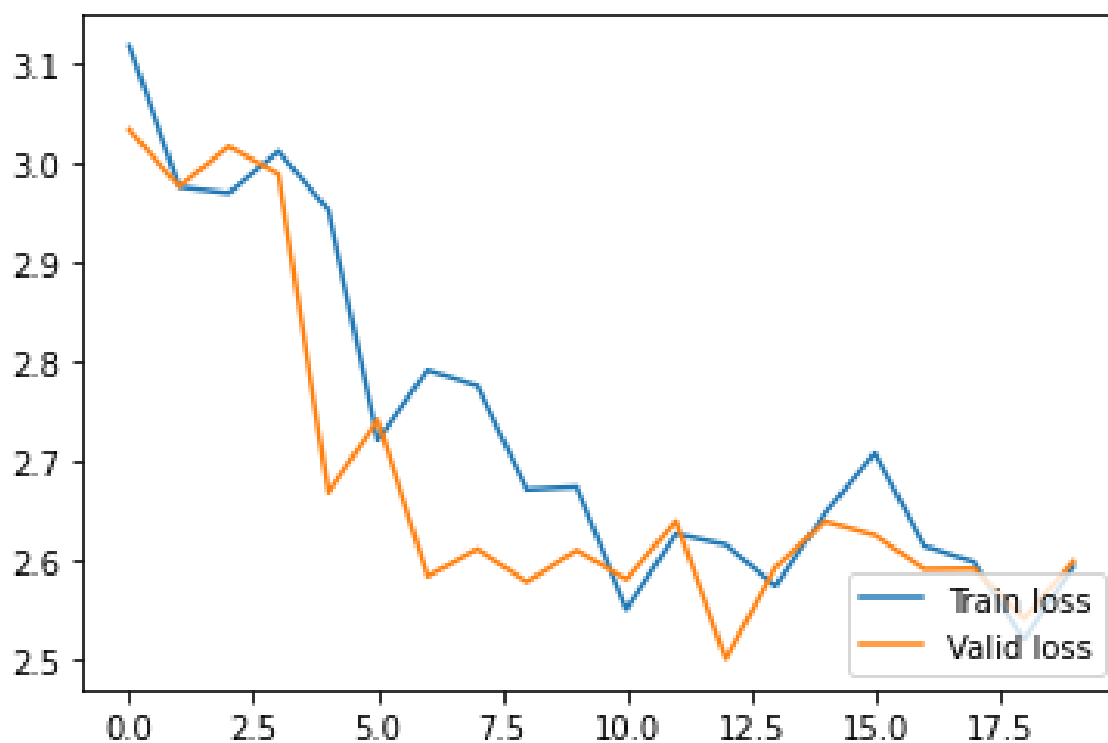
MBTI	preciznost	odziv	F1-mjera	podrška
ESFP	0.00	0.00	0.00	7
ESTJ	0.00	0.00	0.00	6
ESTP	0.00	0.00	0.00	14
INFJ	0.05	0.00	0.01	220
INFP	0.21	0.98	0.35	275
INTJ	0.00	0.00	0.00	163
INTP	0.00	0.00	0.00	195
ISFJ	0.00	0.00	0.00	25
ISFP	0.00	0.00	0.00	40
ISTJ	0.00	0.00	0.00	31
ISTP	0.00	0.00	0.00	50
točnost			0.21	1301
makro avg	0.02	0.06	0.02	1301
težinski avg	0.05	0.21	0.07	1301

Matricu zabune za model LSTM-CE nad skupom podataka s Personality Cafe foruma prikazuje slika 35 te matrica pokazuje da model najbolje uči većinski zastupljene klase INFP i INFJ, što je također očekivano s obzirom da klasična CE funkcija gubitka preferira većinski zastupljene klase u neuravnoteženom skupu podataka. Moguće je zaključiti da LSTM CE model na neuravnoteženom skupu MBTI postiže skromne rezultate po svim metrikama i da model ostvaruje nizak rezultat generalizacije klasifikacije MBTI, jer preferira većinske klase te da je model naučio predikciju svega dvije najviše zastupljene klase. Model nije uspio naučiti niti jednu od deset najmanje zastupljenih klasa od ISTP do ESTJ (slika 28). Interesantno je i da je model ostvario svega 1% za F1–mjeru za klasu INFJ, koja je druga najzastupljenija klasa u skupu podataka. Očekivano, model je ostvario najbolji rezultat u klasifikaciji najzastupljenije klase INFP (F1 – mjera je 35%).

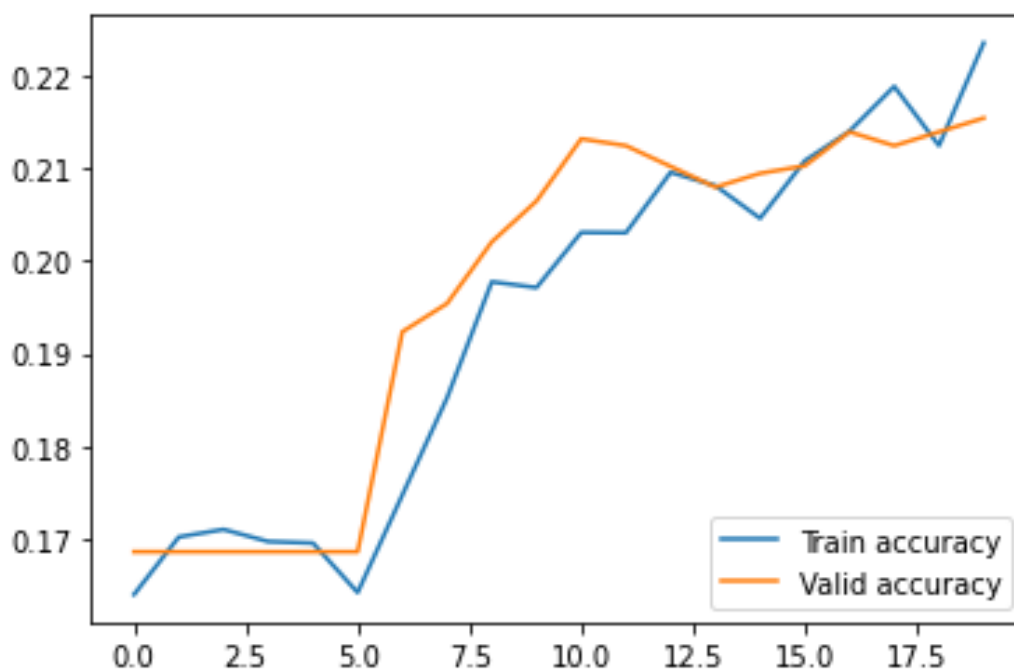


Slika 35 Matrica zabune za LSTM-CE na skupu MBTI s Personality Cafe foruma

Treniranje je ponovljeno uporabom pristupa LSTM-CECI. Gubitak treniranja i validacije te točnost treniranja i validacije daju slika 36 i slika 37.



Slika 36 Gubitak treniranja/validacije za LSTM-CECI na skupu MBTI s Personality Cafe foruma



Slika 37 Točnost treniranja/validacije za LSTM-CECI na skupu MBTI s Personality Cafe foruma

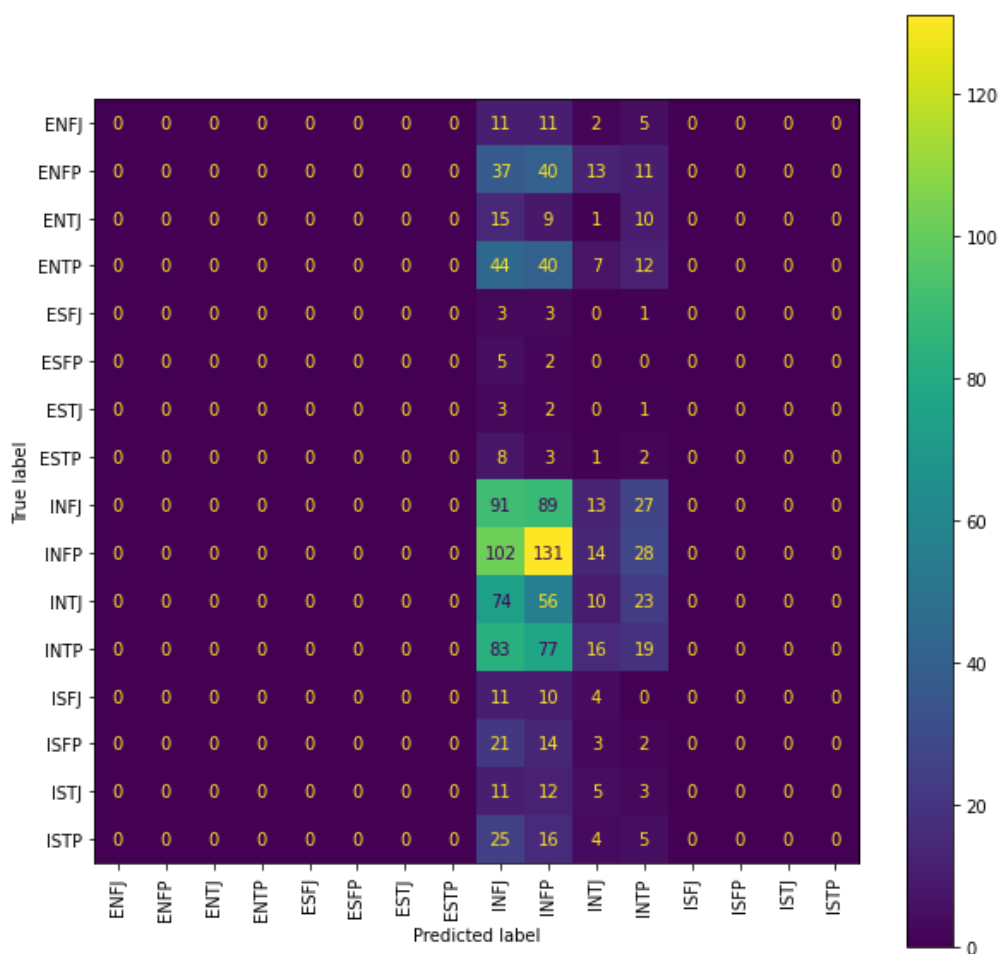
Budući su težinski faktori bili parametri koje je model učio, model je optimizirao težinske faktore α , β , γ , δ funkcije CECI na vrijednosti 0.61, 0.0, 0.24 i 0.19. Tablica 12 i slika 38 daju

klasifikacijski izvještaj i matricu zabune za LSTM-CECI na skupu MBTI s Personality Cafe foruma. Usporedbom rezultata s rezultatima dobivenim treniranjem primjenom funkcije CE, moguće je uočiti povećanje rezultata makro F1-mjere s 2% na 5%, rezultata težinski srednje F1-mjere sa 7% na 14% te da je model naučio klasificirati i klase INTJ i INTP.

Tablica 12 Klasifikacijski izvještaj za LSMT-CECI na skupu MBTI s Personality Cafe foruma

MBTI	preciznost	odziv	F1-mjera	podrška
ENFJ	0.00	0.00	0.00	29
ENFP	0.00	0.00	0.00	101
ENTJ	0.00	0.00	0.00	35
ENTP	0.00	0.00	0.00	103
ESFJ	0.00	0.00	0.00	7
ESFP	0.00	0.00	0.00	7
ESTJ	0.00	0.00	0.00	6
ESTP	0.00	0.00	0.00	14
INFJ	0.17	0.41	0.24	220
INFP	0.25	0.48	0.33	275
INTJ	0.11	0.06	0.08	163
INTP	0.13	0.10	0.11	195
ISFJ	0.00	0.00	0.00	25
ISFP	0.00	0.00	0.00	40
ISTJ	0.00	0.00	0.00	31
ISTP	0.00	0.00	0.00	50
točnost			0.19	1301
makro avg	0.04	0.07	0.05	1301
težinski avg	0.11	0.19	0.14	1301

Model je postigao najbolji rezultat za F1-mjeru za klasu INFP (33%) te za klasu INFJ (24%). Također, za razliku od pristupa s funkcijom CE, model je naučio klasificirati i klase INTJ i INTP.



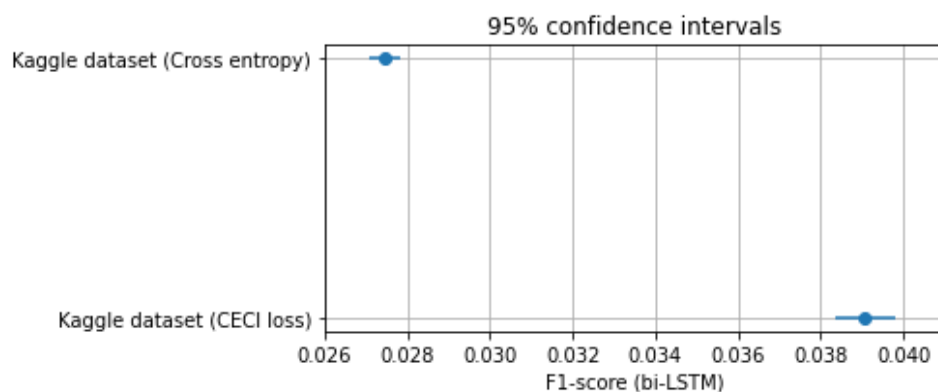
Slika 38 Matrica zabune za LSTM-CECI na skupu MBTI s Personality Cafe foruma

Tablica 13 daje usporedbu rezultata pristupa LSTM-CE i LSTM-CECI na skupu MBTI s Personality Cafe foruma. Može se uočiti da pristup LSTM-CECI ostvaruje bolji rezultat za ključnu metriku F1-mjera te da može klasificirati veći broj klasa.

Tablica 13 Usporedba rezultata pristupa LSTM-CE i LSMT-CECI na skupu MBTI s Personality Cafe foruma

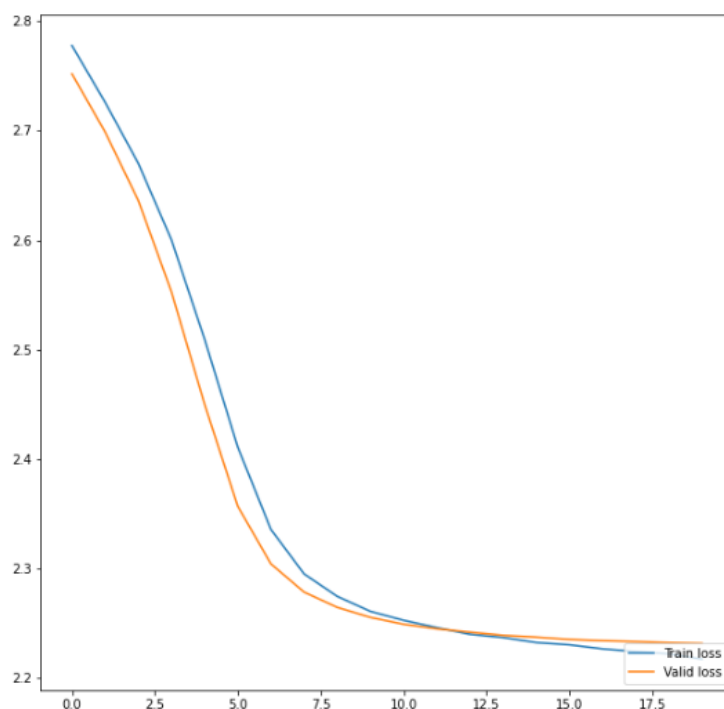
	LSTM-CECI	LSTM-CE
MBTI	F1-mjera	F1-mjera
INFP	0.33	0.35
INFJ	0.24	0.01
INTJ	0.08	0.00
INTP	0.11	0.00
točnost	0.19	0.21
makro avg	0.05	0.02
težinski avg	0.14	0.07

Slika 39 pokazuje rezultat provjere statističke signifikantnosti rezultata usporedbe pristupa LSTM-CECI u odnosu na LSTM-CE nad skupom MBTI s Personality Cafe foruma (skup podataka Kaggle). Moguće je zaključiti da je razlika rezultata F1-mjere ova dva pristupa statistički signifikantna, iako se može uočiti da je prosječna razlika unutar 2%.

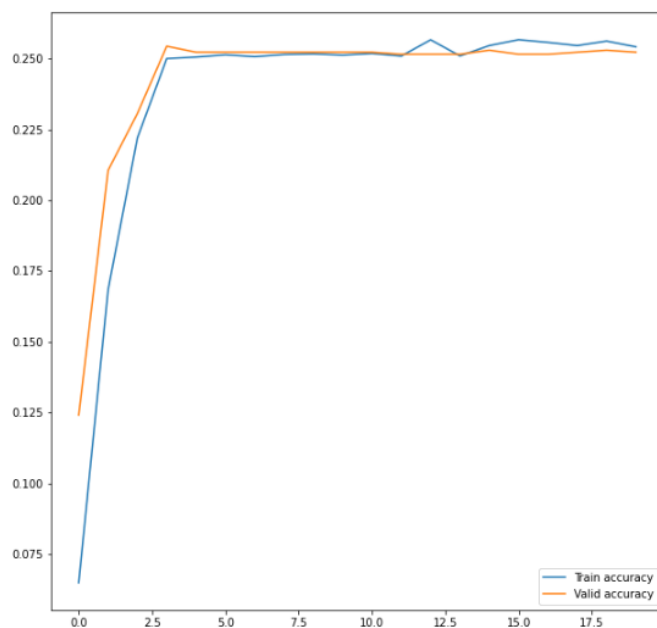


Slika 39 Interval 95% pouzdanosti bi-LSTM CE/CECI na skupu MBTI s Personality Cafe foruma

Nakon treniranja nad skupom podataka MBTI s Personality Cafe foruma, treniranje pomoću modela LSTM je ponovljeno nad skupom podataka s Reddit platforme. Slika 40 i slika 41 prikazuju gubitak treniranja i validacije te točnost treniranja i validacije za LSTM-CE na skupu MBTI s Reddit platforme.

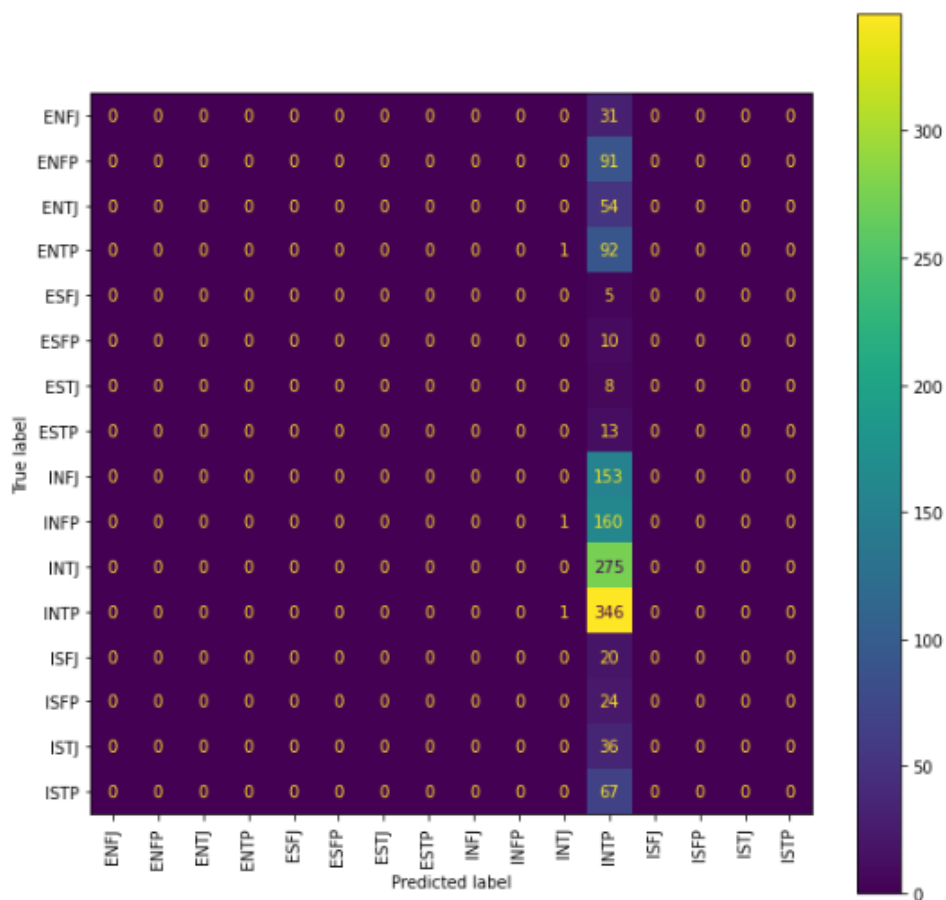


Slika 40 Gubitak treniranja/validacije za LSTM-CE na skupu MBTI s Reddit platforme



Slika 41 Točnost treniranja/validacije za LSTM-CE na skupu MBTI s Reddit platforme

Tablica 14 Klasifikacijski izvještaj za LSTM-CE na skupu s Reddit platforme i slika 42 prikazuju klasifikacijski izvještaj i matricu zabune za LSTM-CE na skupu s Reddit platforme.



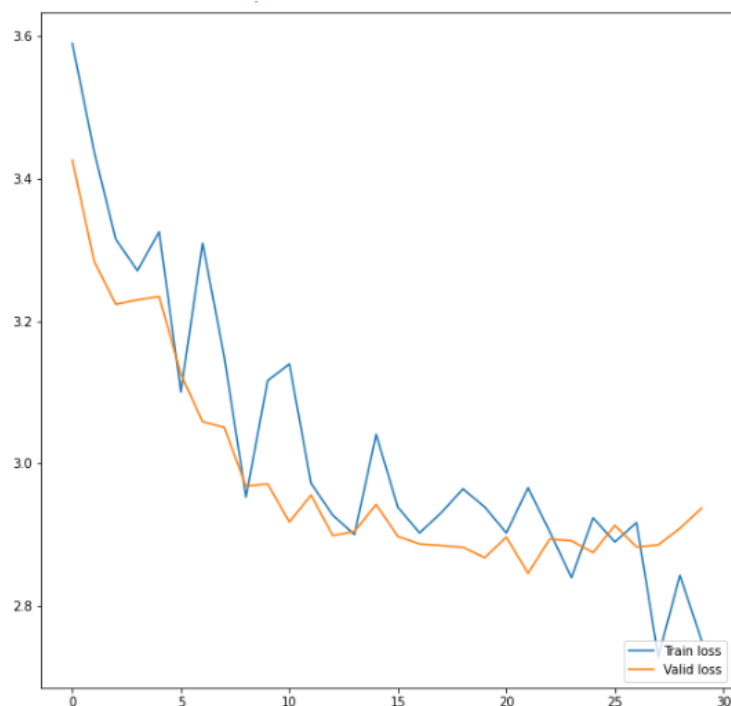
Slika 42 Matrica zabune za LSTM-CE na skupu MBTI s Reddit platforme

Tablica 14 Klasifikacijski izvještaj za LSTM-CE na skupu s Reddit platforme

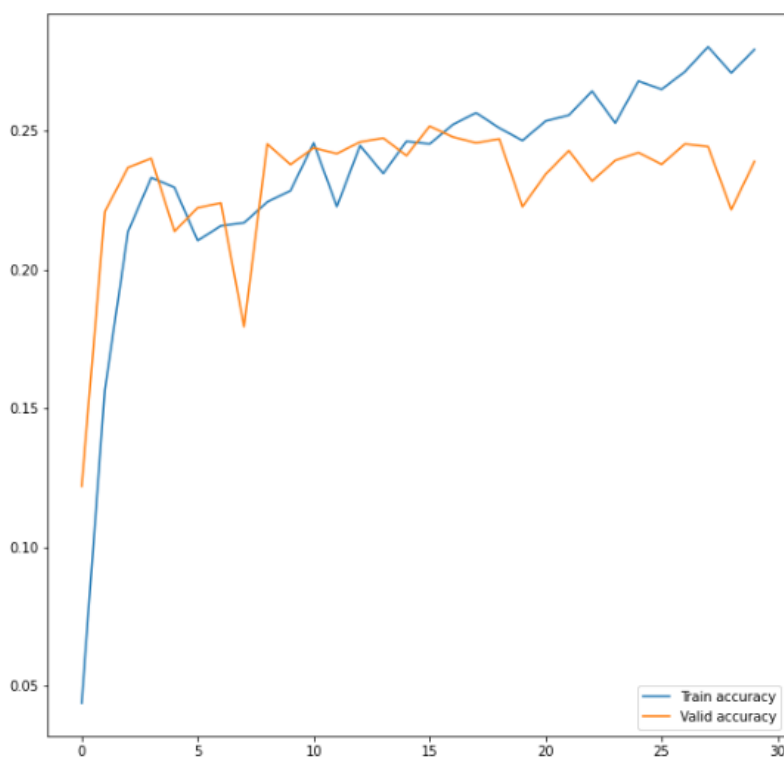
MBTI	preciznost	odziv	F1-mjera	podrška
ENFJ	0.00	0.00	0.00	31
ENFP	0.00	0.00	0.00	91
ENTJ	0.00	0.00	0.00	54
ENTP	0.00	0.00	0.00	93
ESFJ	0.00	0.00	0.00	5
ESFP	0.00	0.00	0.00	10
ESTJ	0.00	0.00	0.00	8
ESTP	0.00	0.00	0.00	13
INFJ	0.00	0.00	0.00	153
INFP	0.00	0.00	0.00	161
INTJ	0.00	0.00	0.00	275
INTP	0.25	1.00	0.40	347
ISFJ	0.00	0.00	0.00	20
ISFP	0.00	0.00	0.00	24
ISTJ	0.00	0.00	0.00	36
ISTP	0.00	0.00	0.00	67
točnost			0.25	1388
makro avg	0.02	0.06	0.02	1388
težinski avg	0.06	0.25	0.10	1388

Model s pristupom CE pokazuje loše rezultate, s rezultatom F1-mjere od 2% i uz činjenicu da je model klasificiramo samo klasu INTP (F1-mjera 40%). Ostvarena je točnost od 25% i težinski prosjek F1-mjere 10%.

Treniranje na skupom MBTI s Reddit platforme je ponovljeno s uporabom funkcije CECI. Slika 43 i slika 44 prikazuju gubitak treniranja i validacije za LSTM-CECI te točnost treniranja i validacije na skupu MBTI s Reddit platforme.



Slika 43 Gubitak treniranja/validacije za LSTM-CECI na skupu MBTI s Reddit platforme

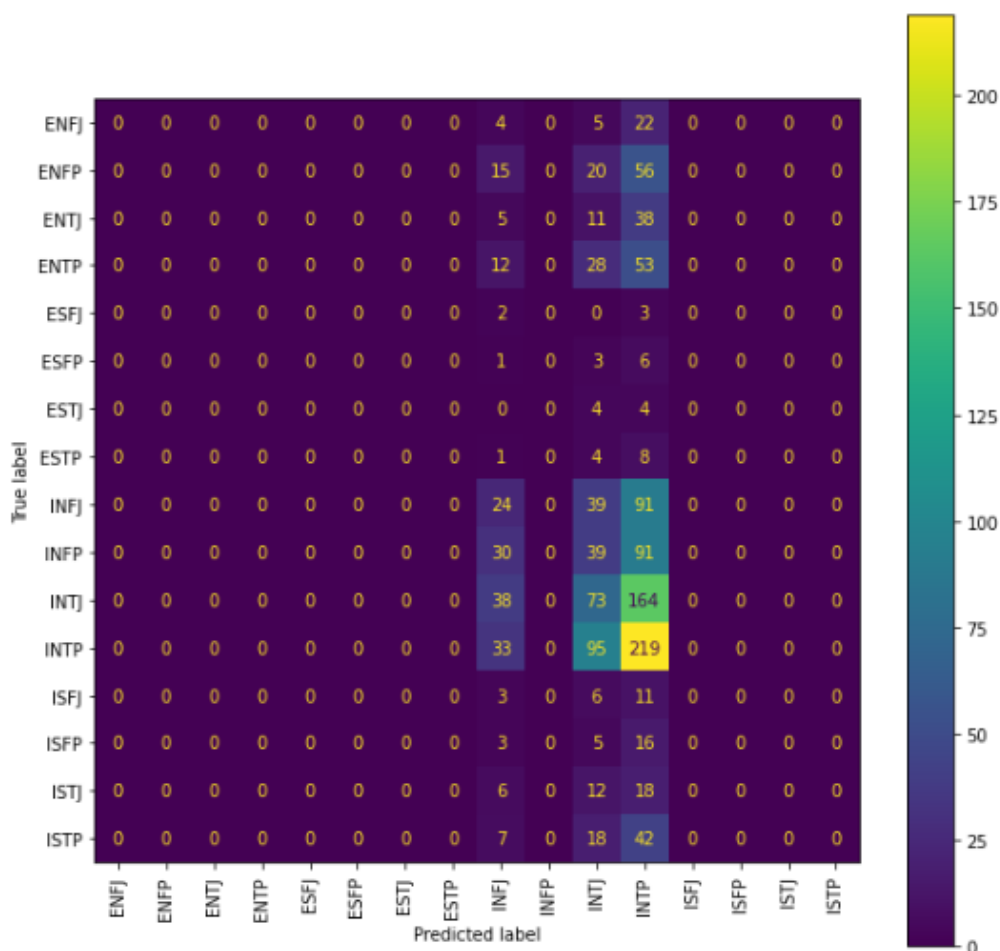


Slika 44 Točnost treniranja/validacije za LSTM-CECI na skupu MBTI s Reddit platforme

Tablica 15 i Slika 45 prikazuju klasifikacijski izvještaj i matricu zabune za LSTM-CECI na skupu s Reddit platforme. Moguće je uočiti da je vrijednost F1-mjere veća u odnosu na pristup s funkcijom CE nad istim skupom podataka i iznosi 5%. Također, model je naučio klasificirati i klase INFJ i INTJ. Vrijednosti težinskih faktora α , β , γ , δ , koje je model naučio su 0.49, 0.24, 0.68 i 0.93.

Tablica 15 Klasifikacijski izvještaj za LSTM-CECI na skupu s Reddit platforme

MBTI	preciznost	odziv	F1-mjera	podrška
ENFJ	0.00	0.00	0.00	31
ENFP	0.00	0.00	0.00	91
ENTJ	0.00	0.00	0.00	54
ENTP	0.00	0.00	0.00	93
ESFJ	0.00	0.00	0.00	5
ESFP	0.00	0.00	0.00	10
ESTJ	0.00	0.00	0.00	8
ESTP	0.00	0.00	0.00	13
INFJ	0.13	0.16	0.14	153
INFP	0.00	0.00	0.00	161
INTJ	0.20	0.27	0.23	275
INTP	0.26	0.63	0.37	347
ISFJ	0.00	0.00	0.00	20
ISFP	0.00	0.00	0.00	24
ISTJ	0.00	0.00	0.00	36
ISTP	0.00	0.00	0.00	67
točnost			0.23	1388
makro avg	0.04	0.07	0.05	1388
težinski avg	0.12	0.23	0.15	1388



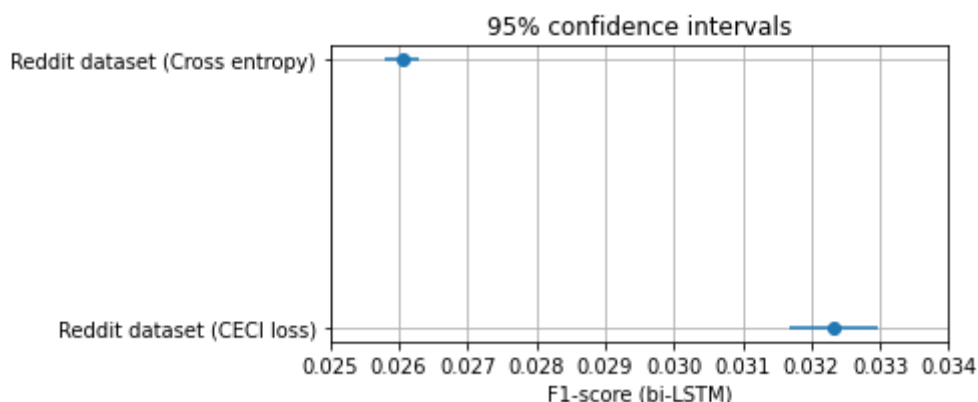
Slika 45 Matrica zabune za LSTM-CECI na skupu MBTI s Reddit platforme

Tablica 16 daje usporedbu rezultata pristupa LSTM-CE i LSTM-CECI na skupu MBTI s Reddit platforme. Može se uočiti da pristup LSTM-CECI ostvaruje bolji rezultat za ključnu metriku F1-mjera te da može klasificirati veći broj klasa.

Tablica 16 Usporedba rezultata pristupa LSTM-CE i LSMT-CECI na skupu MBTI s Reddit platforme

	LSTM-CECI	LSTM-CE
MBTI	F1-mjera	F1-mjera
INFJ	0.14	0.00
INTJ	0.23	0.00
INTP	0.37	0.40
točnost	0.23	0.25
makro avg	0.05	0.02
težinski avg	0.15	0.10

Slika 46 pokazuje rezultat provjere statističke signifikantnosti rezultata usporedbe pristupa LSTM-CECI u odnosu na LSTM-CE nad skupom MBTI s Reddit platforme. Moguće je zaključiti da je razlika rezultata i ova dva pristupa statistički signifikantno različita, iako se može uočiti da je prosječna razlika i ovdje unutar 2%.



Slika 46 Interval 95% pouzdanosti bi-LSTM CE/CECI na skupu MBTI s Reddit platforme

5.3.2. Model CNN

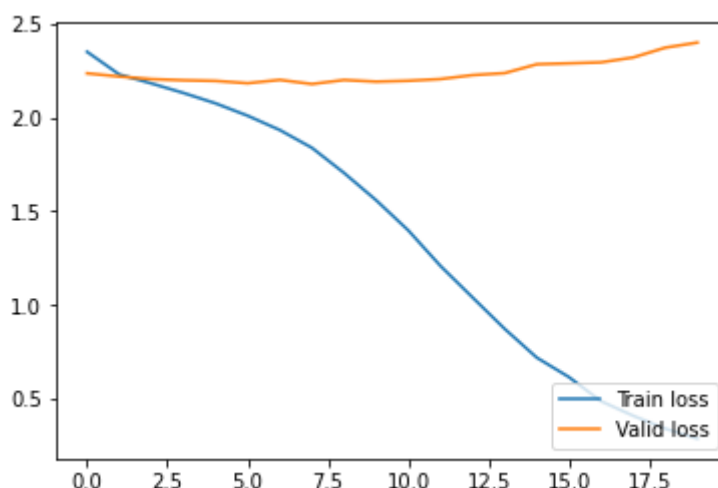
Plan istraživanja uključuje da se efikasnost metode CECI pokaže na više od jednog modela te je odabrana arhitektura CNN za provođenje dodatnih eksperimenata. Slika 47 prikazuje model CNN 2D arhitekture koja je korištenja za treniranje. Model je također treniran s 20 epoha, a vrijednost za nasumično ispuštanje je 0.2. Broj epoha je određen eksperimentalno, promatrajući ponašanje funkcije gubitka na validacijskom skupu, u odnosu na ponašanje funkcije gubitka na skupu za treniranje. I model CNN je treniran nad oba odabrana skupa podataka za istraživanje.

```
CNN(
    (embedding): Embedding(25002, 300)
    (convs): ModuleList(
      (0): Conv2d(1, 100, kernel_size=(2, 300), stride=(1, 1))
      (1): Conv2d(1, 100, kernel_size=(3, 300), stride=(1, 1))
      (2): Conv2d(1, 100, kernel_size=(4, 300), stride=(1, 1))
    )
    (fc): Linear(in_features=300, out_features=16, bias=True)
    (dropout): Dropout(p=0.2, inplace=False)
)
```

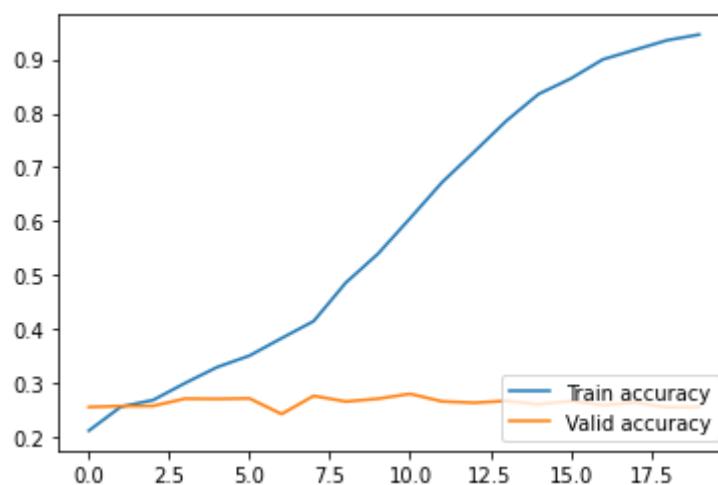
Slika 47 Model CNN

Model CNN je prvo treniran sa standardnom funkcijom gubitka CE, kako bi se dobio rezultat koji predstavlja osnovu za usporedbu kod višeklasne klasifikacije s modelom CNN. Nakon toga je isti model treniran s funkcijom CECI. Treniranje modela CNN s funkcijama CE i CECI je ponovljeno nad oba skupa podataka. Model CNN je prvo treniran na skupu MBTI s Reddit platforme.

Slika 48 i slika 49 prikazuju trening i validacijski gubitak te trening i validacijsku točnost za CNN-CE na skupu s Reddit platforme. Tablica 17 i slika 50 daju klasifikacijski izvještaj i matricu zabune za CNN-CE na skupu MBTI s Reddit platforme. Može se uočiti da je model ostvario rezultat makro srednje F1-mjere 7% te da je naučio klasificirati klase INFJ, INFP, INTJ i INTP.



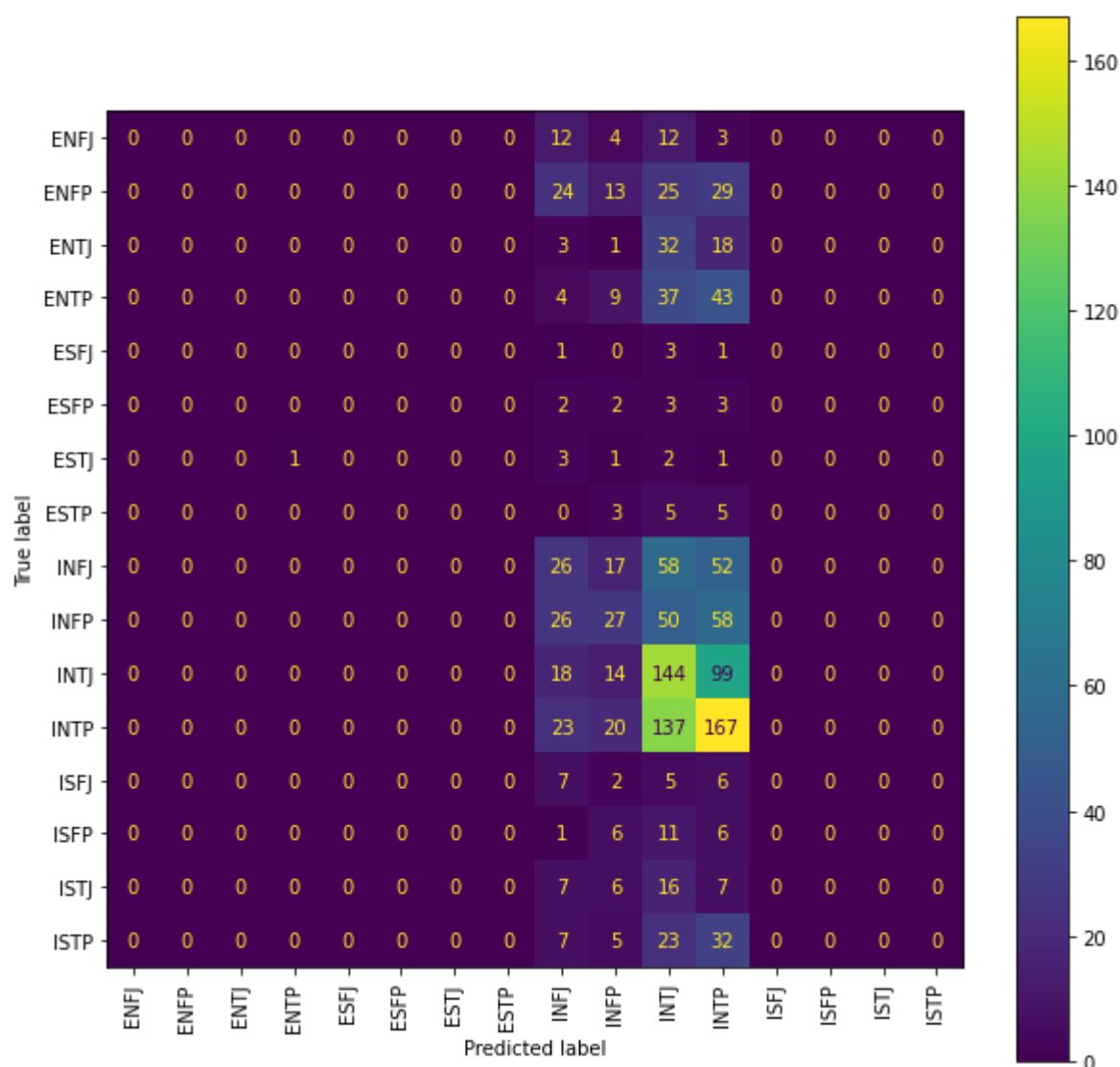
Slika 48 Trening/validacijski gubitak za CNN-CE na skupu MBTI s Reddit platforme



Slika 49 Trening/validacijska točnost za CNN-CE na skupu MBTI s Reddit platforme

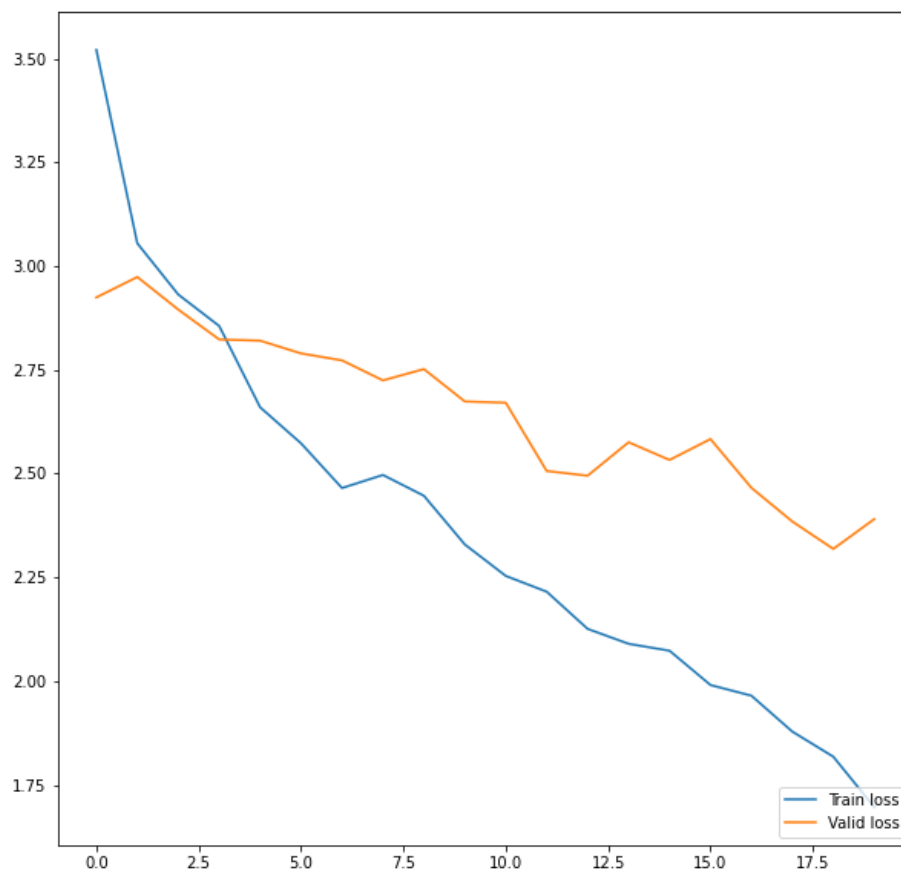
Tablica 17 Klasifikacijski izvještaj za CNN-CE na skupu MBTI s Reddit platforme

MBTI	preciznost	odziv	F1-mjera	podrška
ENFJ	0.00	0.00	0.00	31
ENFP	0.00	0.00	0.00	91
ENTJ	0.00	0.00	0.00	54
ENTP	0.00	0.00	0.00	93
ESFJ	0.00	0.00	0.00	5
ESFP	0.00	0.00	0.00	10
ESTJ	0.00	0.00	0.00	8
ESTP	0.00	0.00	0.00	13
INFJ	0.16	0.17	0.16	153
INFP	0.21	0.17	0.19	161
INTJ	0.26	0.52	0.34	275
INTP	0.32	0.48	0.38	347
ISFJ	0.00	0.00	0.00	20
ISFP	0.00	0.00	0.00	24
ISTJ	0.00	0.00	0.00	36
ISTP	0.00	0.00	0.00	67
točnost			0.26	1388
makro avg	0.06	0.08	0.07	1388
težinski avg	0.17	0.26	0.20	1388

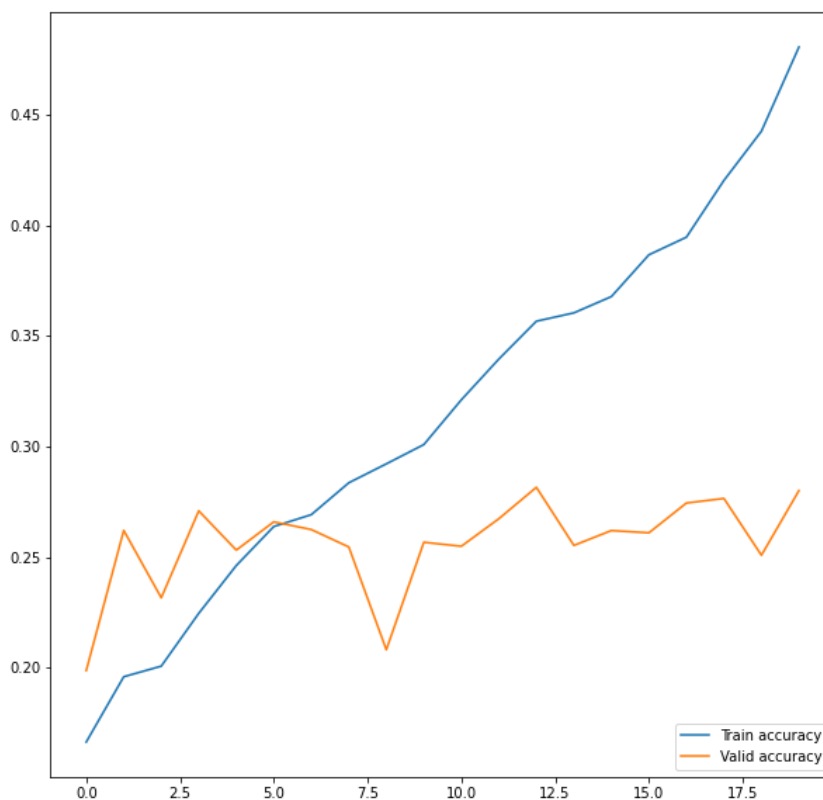


Slika 50 Matrica zabune za CNN-CE na skupu MBTI s Reddit platforme

Potom je nad skupom MBTI s Reddit platforme treniran model primjenom funkcije CECI. Slika 51 i slika 52 i prikazuju trening i validacijski gubitak te trening i validacijsku točnost za CNN-CECI na skupu s Reddit platforme.



Slika 51 Trening/validacijski gubitak za CNN-CECI na skupu s Reddit platforme

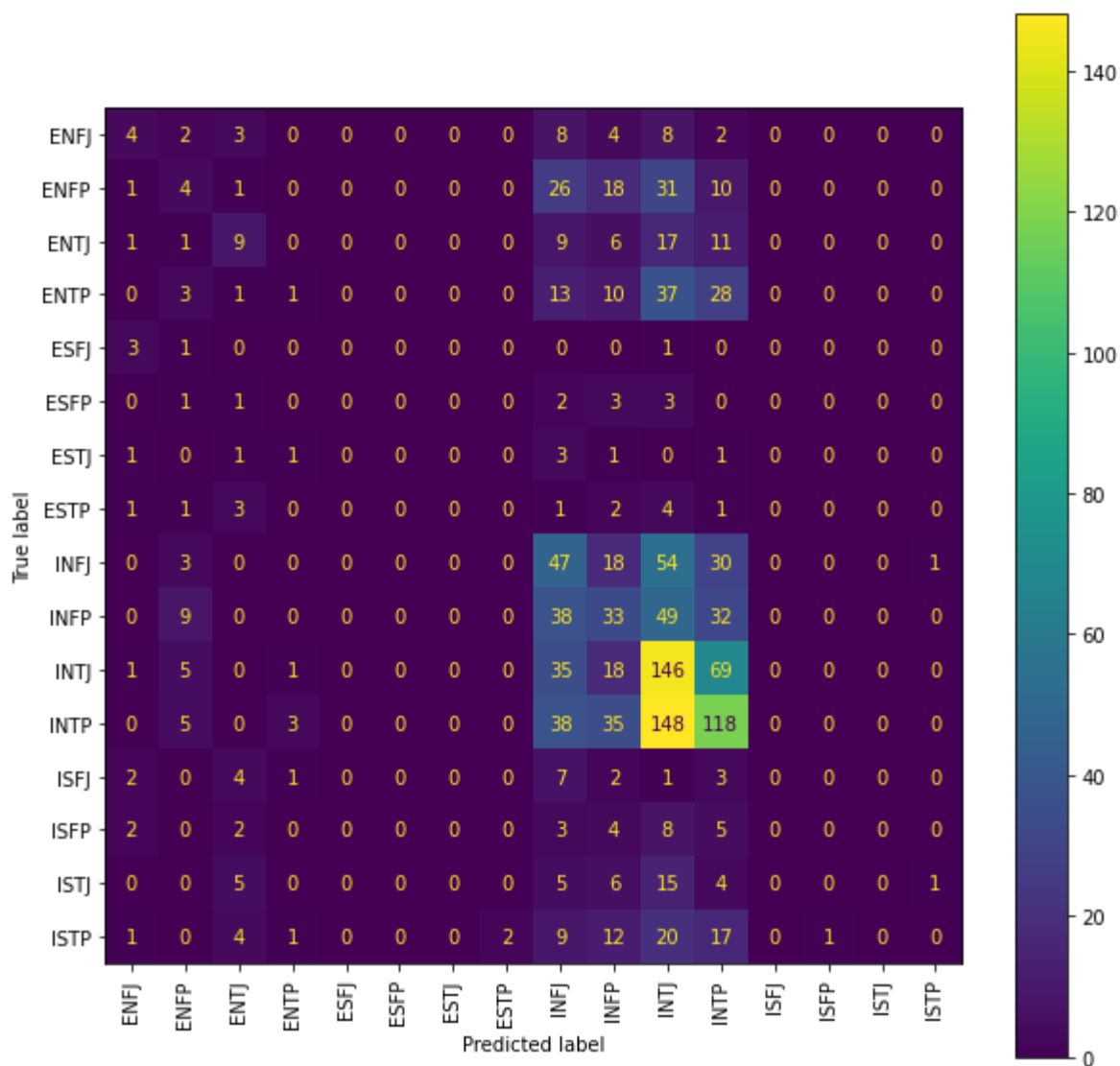


Slika 52 Trening/validacijska točnost za CNN-CECI na skupu s Reddit platforme

Tablica 18 i Slika 50 daju klasifikacijski izvještaj i matricu zabune za CNN-CECI na skupu MBTI s Reddit platforme. Može se uočiti da je model ostvario rezultat makro srednje F1-mjere 10% te da je, u odnosu na pristup CE, naučio klasificirati i klase ENFJ, ENFP, ENTJ i ENTP. Vrijednosti težinskih faktora α , β , γ , δ , koje je model naučio su 0.16, 0.02, 0.22 i 0.04.

Tablica 18 Klasifikacijski izvještaj za CNN-CECI na skupu s Reddit platforme

MBTI	preciznost	odziv	F1-score	podrška
ENFJ	0.24	0.13	0.17	31
ENFP	0.11	0.04	0.06	91
ENTJ	0.26	0.17	0.20	54
ENTP	0.12	0.01	0.02	93
ESFJ	0.00	0.00	0.00	5
ESFP	0.00	0.00	0.00	10
ESTJ	0.00	0.00	0.00	8
ESTP	0.00	0.00	0.00	13
INFJ	0.19	0.31	0.24	153
INFP	0.19	0.20	0.20	161
INTJ	0.27	0.53	0.36	275
INTP	0.36	0.34	0.35	347
ISFJ	0.00	0.00	0.00	20
ISFP	0.00	0.00	0.00	24
ISTJ	0.00	0.00	0.00	36
ISTP	0.00	0.00	0.00	67
točnost			0.26	1388
makro avg	0.11	0.11	0.10	1388
težinski avg	0.22	0.26	0.22	1388



Slika 53 Matrica zabune za CNN-CECI na skupu MBTI s Reddit platforme

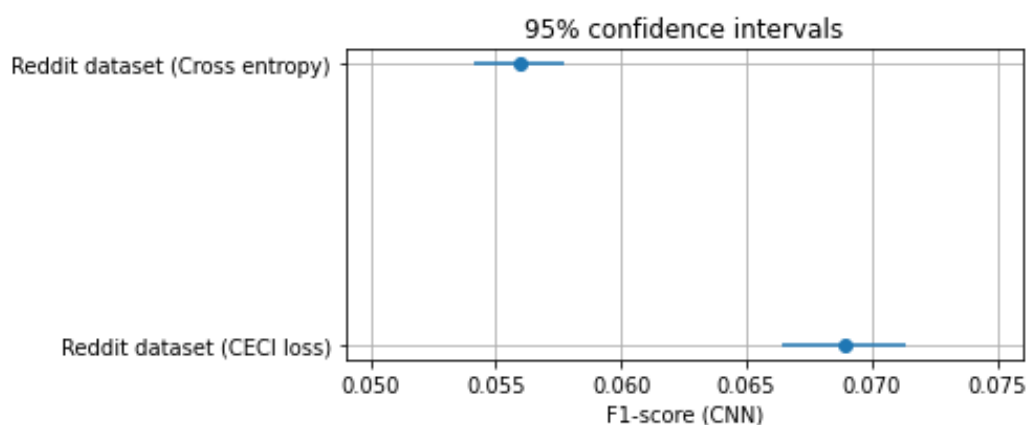
Tablica 19 daje usporedbu rezultata pristupa CNN-CE i CNN-CECI na skupu MBTI s Reddit platforme. Može se uočiti da pristup CNN-CECI ostvaruje bolji rezultat za ključnu metriku F1-mjera te da može klasificirati veći broj klasa.

Tablica 19 Usporedba rezultata pristupa CNN-CE i CNN-CECI na skupu MBTI s Reddit platforme

	CNN-CECI	CNN-CE
MBTI	F1-mjera	F1-mjera
ENFJ	0.17	0.00
ENFP	0.06	0.00
ENTJ	0.20	0.00
ENTP	0.02	0.00
INFJ	0.24	0.16

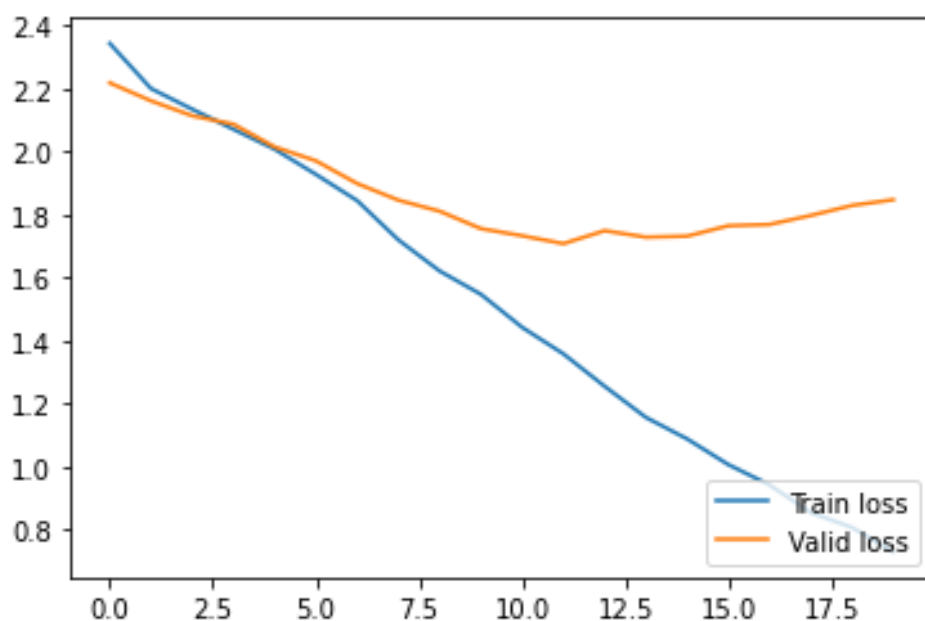
	CNN-CECI	CNN-CE
MBTI	F1-mjera	F1-mjera
INFP	0.20	0.19
INTJ	0.36	0.34
INTP	0.35	0.38
točnost	0.26	0.26
makro avg	0.10	0.07
težinski avg	0.22	0.20

Slika 54 pokazuje rezultat provjere statističke signifikantnosti rezultata usporedbe pristupa CNN-CECI u odnosu na CNN-CE nad skupom MBTI s Reddit platforme. Moguće je zaključiti da je razlika rezultata i ova dva pristupa statistički signifikantna, iako se može uočiti da je prosječna razlika i ovdje unutar 2%.

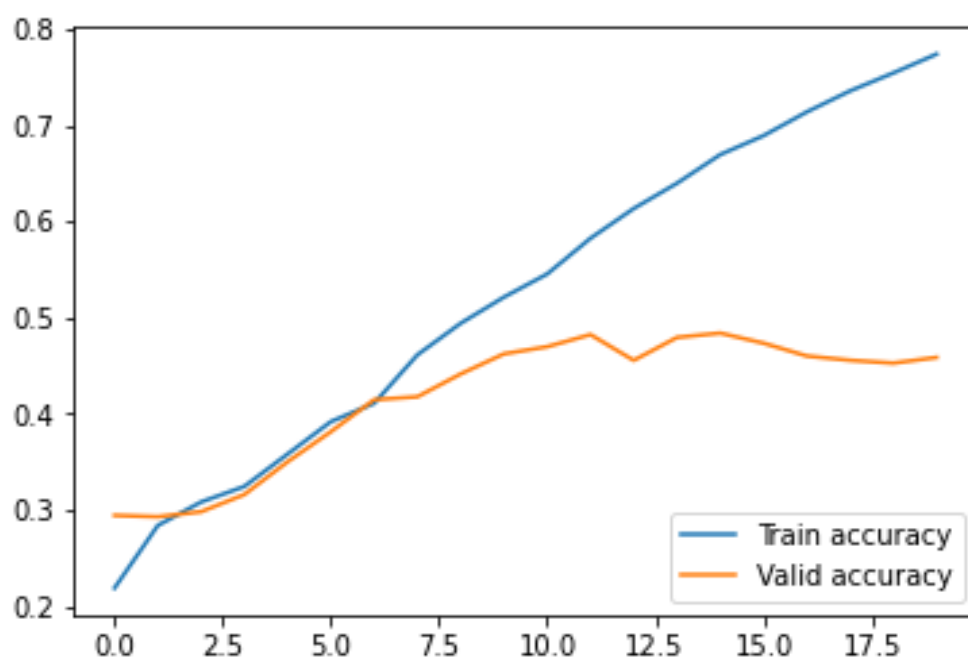


Slika 54 Interval 95% pouzdanosti CNN CE/CECI na skupu MBTI s Reddit platforme

Treniranje pomoću mreže CNN je ponovljeno i na skupu MBTI s Personality Cafe foruma. Slika 55 i Slika 56 prikazuju trening i validacijski gubitak te trening i validacijsku točnost za CNN-CE na skupu MBTI podataka s Personality Cafe foruma.



Slika 55 Trening/validacijski gubitak za CNN-CE na skupu MBTI podataka s Personality Cafe foruma

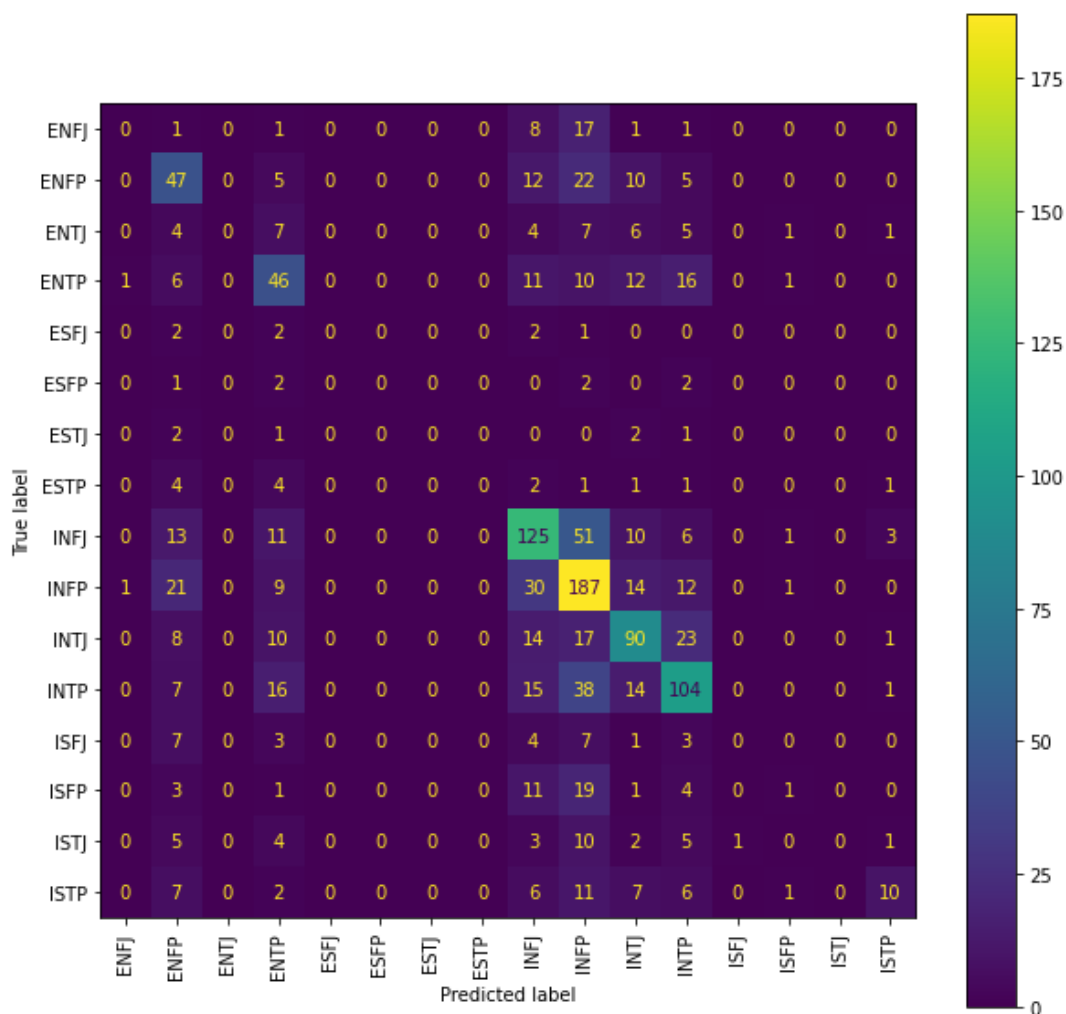


Slika 56 Trening/validacijska točnost za CNN-CE na skupu MBTI podataka s Personality Cafe foruma

Tablica 20 i slika 57 prikazuju klasifikacijski izvještaj i matricu zabune za CNN-CE na skupu MBTI podataka s Personality Cafe foruma. Ostvaren je rezultat makro F1-mjere od 21%. Rezultati su ponovno očekivano niski jer se radi o izrazito neuravnoteženom skupu podataka, ali su dobiveni rezultati značajno bolji od rezultata dobivenih pomoću modela LSTM, primjenom funkcije gubitka CE nad istim skupom podataka.

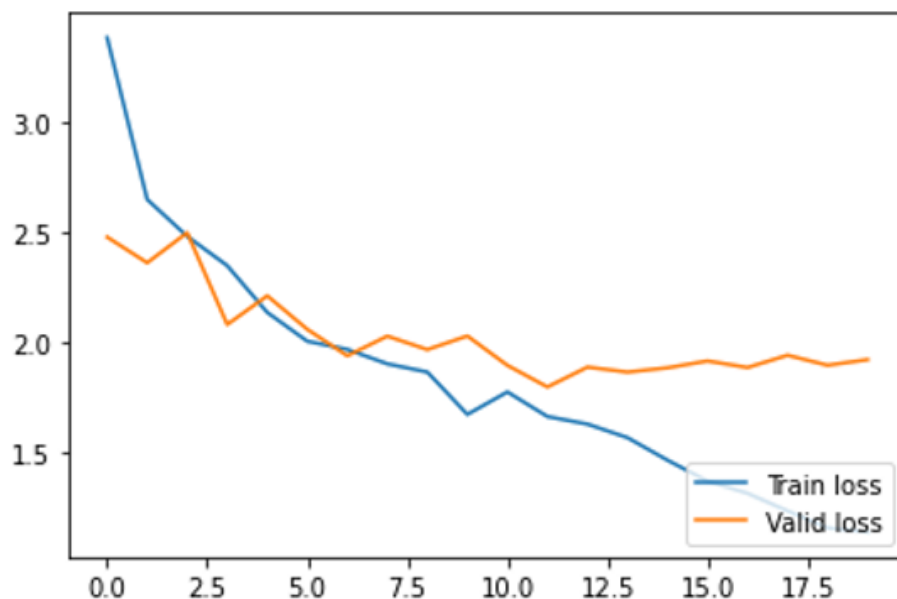
Tablica 20 Klasifikacijski izvještaj za CNN-CE na skupu MBTI podataka s Personality Cafe foruma

MBTI	preciznost	odziv	F1-mjera	podrška
ENFJ	0.00	0.00	0.00	29
ENFP	0.34	0.47	0.39	101
ENTJ	0.00	0.00	0.00	35
ENTP	0.37	0.45	0.41	103
ESFJ	0.00	0.00	0.00	7
ESFP	0.00	0.00	0.00	7
ESTJ	0.00	0.00	0.00	6
ESTP	0.00	0.00	0.00	14
INFJ	0.51	0.57	0.54	220
INFP	0.47	0.68	0.55	275
INTJ	0.53	0.55	0.54	163
INTP	0.54	0.53	0.53	195
ISFJ	0.00	0.00	0.00	25
ISFP	0.17	0.03	0.04	40
ISTJ	0.00	0.00	0.00	31
ISTP	0.56	0.20	0.29	50
točnost			0.47	1301
makro avg	0.22	0.22	0.21	1301
težinski avg	0.41	0.47	0.43	1301

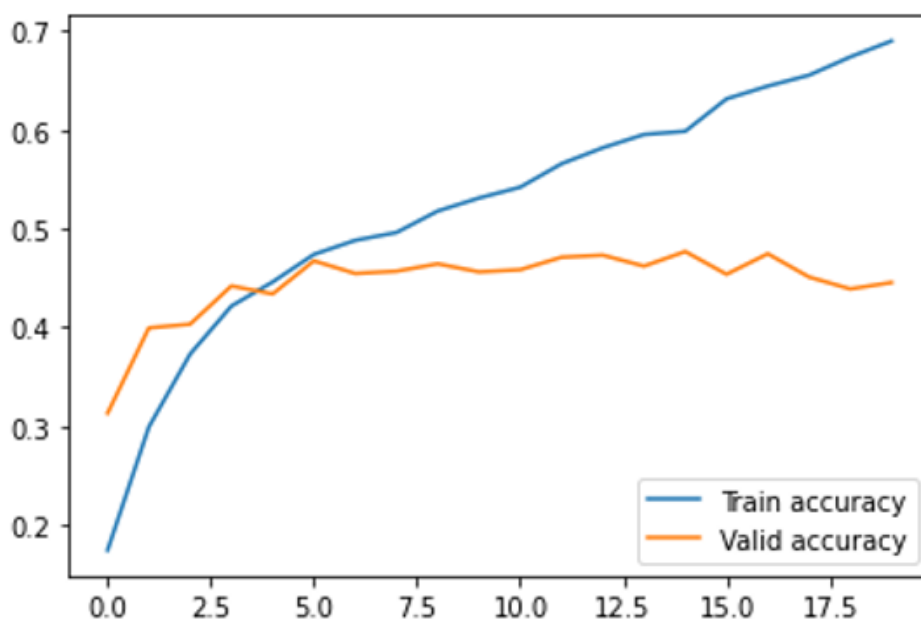


Slika 57 Matrica zabune za CNN-CE na skupu MBTI podataka s Personality Cafe foruma

Treniranje je ponovljeno pomoći funkcije CECI. Slika 58 i slika 59 prikazuju trening i validacijski gubitak te trening i validacijsku točnost za CNN-CECI na skupu MBTI podataka s Personality Cafe foruma.



Slika 58 Trening/validacijski gubitak za CNN-CECI na skupu podataka s Personality Cafe foruma

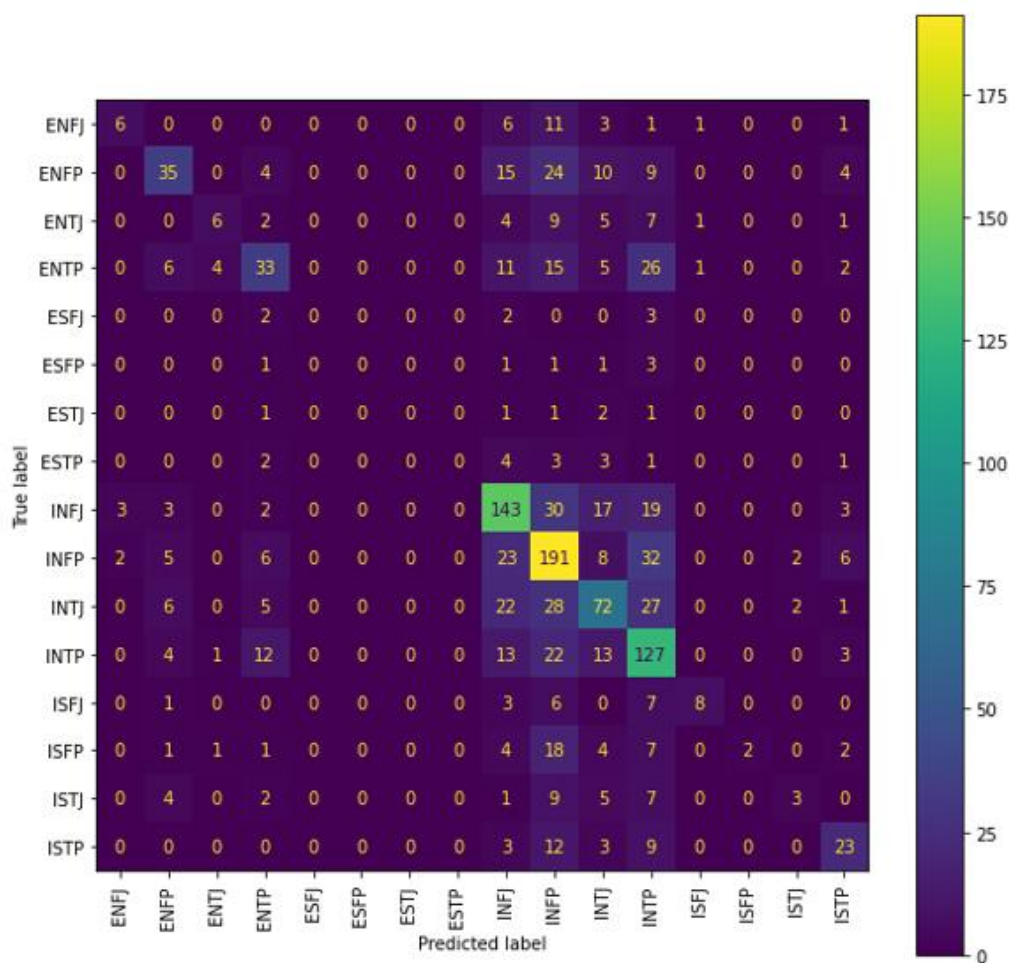


Slika 59 Trening/validacijska točnost za CNN-CECI na skupu podataka s Personality Cafe foruma

Tablica 21 i slika 60 prikazuju klasifikacijski izvještaj i matricu zabune za CNN-CECI na skupu MBTI podataka s Personality Cafe foruma. Ostvaren je rezultat makro F1-mjere od 29% a model je naučio klasificirati i klase ENFJ, ENTJ, ISFJ i ISTJ. Vrijednosti težinskih faktora α , β , γ , δ , koje je model naučio su 0.11, 0.01, 0.68 i 0.79.

Tablica 21 Klasifikacijski izvještaj za CNN-CE na skupu podataka s Personality Cafe foruma

MBTI	preciznost	odziv	F1-score	podrška
ENFJ	0.55	0.21	0.30	29
ENFP	0.54	0.35	0.42	101
ENTJ	0.50	0.17	0.26	35
ENTP	0.45	0.32	0.38	103
ESFJ	0.00	0.00	0.00	7
ESFP	0.00	0.00	0.00	7
ESTJ	0.00	0.00	0.00	6
ESTP	0.00	0.00	0.00	14
INFJ	0.56	0.65	0.60	220
INFP	0.50	0.69	0.58	275
INTJ	0.48	0.44	0.46	163
INTP	0.44	0.65	0.53	195
ISFJ	0.73	0.32	0.44	25
ISFP	1.00	0.05	0.10	40
ISTJ	0.43	0.10	0.16	31
ISTP	0.49	0.46	0.47	50
točnost			0.50	1301
makro avg	0.42	0.28	0.29	1301
težinski avg	0.50	0.50	0.47	1301



Slika 60 Matrica zabune za CNN-CECI na skupu podataka s Personality Cafe foruma

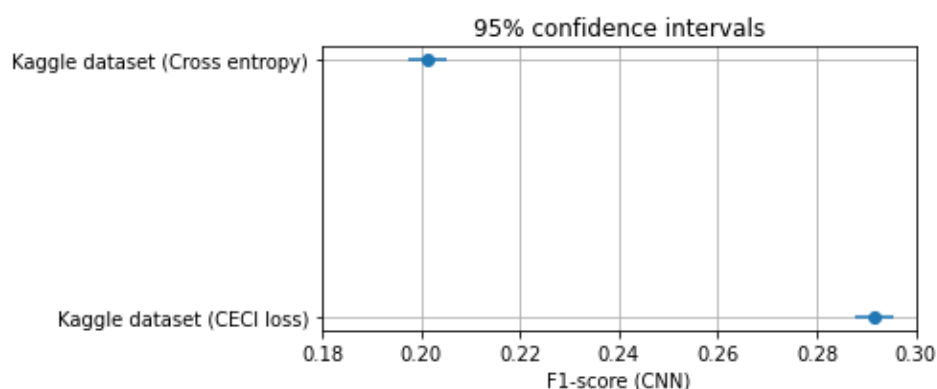
Tablica 22 daje usporedbu rezultata pristupa CNN-CE i CNN-CECI na skupu MBTI s Personality Cafe foruma. Može se uočiti da pristup CNN-CECI ostvaruje bolji rezultat za ključnu metriku F1-mjera te da može klasificirati veći broj klasa.

Tablica 22 Usporedba rezultata pristupa CNN-CE i CNN-CECI na skupu MBTI s Personality Cafe foruma

	CNN-CECI	CNN-CE
MBTI	F1-mjera	F1-mjera
ENFJ	0.30	0.00
ENFP	0.42	0.39
ENTJ	0.26	0.00
ENTP	0.38	0.41
INFJ	0.60	0.54
INFP	0.58	0.55

	CNN-CECI	CNN-CE
MBTI	F1-mjera	F1-mjera
INTJ	0.46	0.54
INTP	0.53	0.53
ISFJ	0.44	0.00
ISFP	0.10	0.04
ISTJ	0.16	0.00
ISTP	0.47	0.29
točnost	0.50	0.47
makro avg	0.29	0.21
težinski avg	0.47	0.43

Slika 61 pokazuje rezultat provjere statističke signifikantnosti rezultata usporedbe pristupa CNN-CECI u odnosu na CNN-CE nad skupom MBTI s Kaggle platforme (Personality Cafe). Moguće je zaključiti da je razlika rezultata i ova dva pristupa statistički signifikantno različita, iako se može uočiti da je prosječna razlika ovdje unutar 10%.



Slika 61 Interval 95% pouzdanosti CNN CE/CECI na skupu MBTI s Personality Cafe foruma

U usporedbi rezultata dobivenih s dva modela uporabom funkcije gubitka CE, rezultati dobiveni CNN pristupom su bolji od rezultata s LSTM pristupom. Npr., za skup podataka s Kaggle platforme, težinska prosječna F1-mjera za CNN CE pristup je 43%, u odnosu na 7% (LSTM CE). Također, za skup podataka s Reddit platforme, težinska prosječna F1-mjera za CNN CE pristup je 20% u odnosu na 10% (LSTM CE). Makro F1-mjera za CNN CE pristup nad skupom s Kaggle platforme je 21%, u odnosu na 2%, (LSTM CE) što je mnogo bolji rezultat. Također, makro F1-mjera za CNN CE pristup na skupom podataka s Reddit platforme iznosi 7% u odnosu na 2% za LSTM CE pristup.

Primjenom funkcije gubitka CECI na CNN modelu ostvarena su poboljšanja u odnosu na oba LSTM pristupa te CNN CE pristup. Npr., za CNN CECI skup podataka s Kaggle platforme makro F1-mjera je 29% u odnosu na 5% za LSTM CECI pristup. Također, za skup podataka s Reddit platforme, CNN CECI pristup je postigao makro F1-mjeru 10% u odnosu na 5% za LSTM CECI. Za skup podataka s Kaggle platforme, poboljšanje je iznosilo s 5% (LSTM CECI) na 29% (CNN CECI). Model CNN CECI pokazuje značajan potencijal za modeliranje ovog tipa klasifikacijskih problema jer je ostvario i najveći broj klasa koje je uspio klasificirati, uz bolje rezultate i prema ostalim metrikama.

Može se uočiti da je kod primjene pristupa CECI nad LSTM modelom najbolji rezultat ostvaren s najvećom penalizacijom za četvrtu dihotomiju (0.93) za skup podataka s Reddit platforme, odnosno za prvu dihotomiju (0.61) za skup podataka s Kaggle platforme. Kod primjene funkcije CECI s modelom CNN, za skup podataka s Reddit platforme najveća je penalizacija na trećoj dihotomiji (0.7), odnosno za skup podataka s Kaggle platforme, na zadnjoj dihotomiji (0.79). To zapažanje može biti osnova dodatnih budućih istraživanja. Prikazani su i dodatni rezultati za model CNN CECI, za druge težinske vrijednosti parametara.

Sumirani rezultati za modele LSTM i CNN daje Tablica 23.

Tablica 23 Pregled rezultata modelima LSTM i CNN

Model	CE/CECI	Skup podataka	$\alpha, \beta, \gamma, \delta$	Makro-F1	Težinski F1
LSTM	CE	Personality Cafe	0.0, 0.0, 0.0, 0.0	0.02	0.07
LSTM	CECI	Personality Cafe	0.61, 0.0, 0.24, 0.19	0.05	0.14
LSTM	CE	Reddit	0.0, 0.0, 0.0, 0.0	0.02	0.10
LSTM	CECI	Reddit	0.49, 0.24, 0.68, 0.93	0.05	0.15
CNN	CE	Personality Cafe	0.0, 0.0, 0.0, 0.0	0.21	0.43
CNN	CECI	Personality Cafe	0.11, 0.01, 0.68, 0.79	0.29	0.47
CNN	CE	Reddit	0.0, 0.0, 0.0, 0.0	0.07	0.20
CNN	CECI	Reddit	0.1, 0.2, 0.7, 0.1	0.10	0.22

Istraživanje je provedeno na dvije tipične arhitekture neuronskih mreža, LSTM i CNN, i na obje arhitekture metoda CECI je pokazala poboljšanja u odnosu na rezultate dobivene klasičnom funkcijom gubitka CE. LSTM model s funkcijom gubitka CE je imao niske performanse, makro F1-mjera je bila 2%, a težinska F1-mjera 7% na skupu podataka s Kaggle

platforme. LSTM CECI model je ostvario F1-mjeru 5%, a težinsku F1-mjeru 20%. CNN CE model je pokazao F1-mjeru 27% i težinsku F1-mjeru 14%. I na skupu podataka s Reddit platforme se može uočiti poboljšanje s 2% na 5% za makro F1-mjeru te s 10% na 15% za težinsku F1-mjeru.

Konačno, CNN CECI je pokazao F1-mjeru 29% i težinsku F1-mjeru 47 u odnosu na 21% za makro F1-mjeru i 43% za težinsku F1-mjeru. Dodatno, primjena metode CECI je omogućila bolju predikciju za pojedine klase MBTI. Npr., LSTM CECI je naučio predikciju klasa INTJ i INPJ za skup podataka s Kaggle platforme. Model CNN CECI je naučio predikciju klasa ENFJ, ENFP, ENTJ i ENTP. Dakle, metoda CECI je unaprijedila i model LSTM i model CNN.

5.4. Usporedbe s rezultatima drugih modela klasifikacije

Kod usporedbe rezultata je bitno istaknuti važnost primijenjene metrike na osnovu koje se donose zaključci. Kako je skup podataka MBTI neuravnotežen, uporaba točnosti kao metrike je dvojbeni i može voditi pogrešnim zaključcima, osobito kod binarne ili višeklasne klasifikacije ako se radi o izrazito neuravnoteženom skupu podataka [106]. Na osnovu podataka o distribuciji klasa MBTI u populaciji SAD (Tablica 8) se može zaključiti da su skupovi s uzorcima klasa MBTI u populaciji očekivano neuravnoteženi. Stoga su i skup podataka s Personality Cafe foruma i s Reddit platforme, koji su korišteni u disertaciji, izrazito neuravnoteženi. Tablica 9 pokazuje da je neuravnoteženost prisutna i na razini dihotomije te je npr. na prvoj poziciji udio dihotomije I 76.96%, a na drugoj 86.20% za dihotomiju N. Tablica 8 i slika 28 pokazuju kako je neuravnoteženost još izraženija za višeklasnu klasifikaciju. Iz tog razloga postignuće visoke točnosti ne potvrđuje model kao uspješan model klasifikacije jer visoka točnost ukazuje da model ostvaruje visoku predikciju većinski zastupljenih klasa, a izostavlja manjinski prisutne klase.

Preciznost ili pozitivna vrijednost predikcije predstavlja količnik istinski pozitivnih vrijednosti (engl. *true positive* – *TP*) i broja pozitivno predviđenih klasa. Na taj način, preciznost govori kolika je mjera pouzdanosti u rezultat kada klasifikator predvidi klasu kao pozitivnu. Iz tog razloga je drugi naziv pozitivna predviđena vrijednost (engl. *positive predictive value*). Odziv ili osjetljivost mjeri kompletnost klasifikatora izračunom količnika *TP* i ukupnog broja pozitivno klasificiranih klasa. Mjera je poznata i kao stopa *TP* (engl. *true-positive rate*). F1-mjera ili F-score predstavlja mjeru ravnoteže preciznosti i odziva kroz

težinsku srednju vrijednost. Budući da maksimalna prosječna vrijednost uzima u obzir sve klase, predstavlja pouzdaniju mjeru u neuravnoteženom skupu podataka.

Dodatno, kako bi se osigurala sumjerljivost rezultata, neophodno je osigurati da usporedba rezultata drugih modela bude napravljena tako da su ti modeli trenirani nad skupom podataka nad kojim su primijenjeni identični koraci lingvističke predobrade kao na modelu s kojim se radi usporedba.

Imajući na umu zapažanja o pouzdanim metrikama za neuravnotežene skupove podataka, i potrebu treniranja modela na isti način i na potpuno identičnim skupovima podataka, tablica 24 daje pregled usporedbe rezultata makro F1-mjere za modele logističke regresije, slučajne šume i naivan Bayesov klasifikator. Za te modele su korišteni samo polazni parametri za treniranje a nisu ni korištene verzije klasifikatora koje su prilagođene za klasifikaciju nad nebalansiranim skupom podataka. Može se uočiti da pristup CNN CECI nadmašuje ostale pristupe nad oba skupa podataka. Među dodatnim modelima strojnog učenja za usporedbu, slučajne šume daju najbolji rezultat za makro F-1 mjeru. Pristup CNN CECI je ostvario makro F1-mjeru 29%, u odnosu na 21% za CNN CE te 6% za RF kod skupa podataka s Personality Cafe foruma (Kaggle). Za skup podataka s Reddit platforme, pristup CNN CECI je ostvario 10%, u odnosu na 7% za CNN CE te 6% za RF.

Tablica 24 Usporedba rezultata s drugim modelima - F1-mjera

Model	Skup podataka	$\alpha, \beta, \gamma, \delta$	Makro F1-mjera
LSTM	Personality Cafe	0.61, 0.0, 0.24, 0.19	0.05
LR			0.02
RF			0.06
MNB			0.02
CNN		0.11, 0.01, 0.68, 0.79	0.29
LSTM		0.0, 0.0 0.0, 0.0	0.02
CNN		0.0, 0.0, 0.0, 0.0	0.21
LSTM	Reddit	0.49, 0.24, 0.68, 0.93	0.05
LSTM		0.0, 0.0, 0.0, 0.0	0.02
CNN		0.16, 0.02, 0.22, 0.04	0.10
CNN		0.0, 0.0, 0.0, 0.0	0.07
LR			0.03
MNB			0.02
RF			0.06

6. PROTOTIP PROGRAMSKOG SUSTAVA

Jedan od doprinosa disertacije je prototip programskoga sustava za određivanje tipa osobnosti sa složenim oznakama klasa temeljen na dubokom učenju. Prototip aplikacije je razvijen pomoću programskog jezika Python s podržanim programskim bibliotekama i okvirima. U poglavlju je prvo opisan način razvoja mrežnog okvira, koji je korišten kako bi se prototip razvio sustavnim pristupom modeliranja aplikativnog rješenja. Potom je opisana arhitektura programskog rješenja. Opisani su načini korištenja treniranih modela i rječnika kreiranih tijekom treniranja te popratnih programskih datoteka za razvoj sučelja i poslovne logike. Nakon toga je demonstriran prototip programskog rješenja.

6.1. Mrežni okvir prototipa aplikacije

Uobičajeni pristup u suvremenom razvoju softverskih rješenja uključuje kreiranje mrežnog okvira (engl. *wireframing*) s ciljem skiciranja očekivanog izgleda i ponašanja korisničkog sučelja. Mrežni okvir može biti dinamični ili statični (engl. *mockup*). Statični mrežni okvir omogućava skiciranje vizualnog sučelja, a dinamični mrežni okvir omogućava i simulaciju ponašanja aplikacije. Prilikom razvoja prototipa u disertaciji kreiran je statični mrežni okvir. Na taj se način pomoću mrežnog okvira razvoja plan aplikacije i osiguravaju uvjeti za manje naknadnih izmjena aplikacije. Slika 62 daje mrežni okvir prototipa aplikacije koja demonstrira primjenu metode CECI za klasifikaciju osobnosti autora teksta prema modelu MBTI. Za razvoj mrežnog okvira je upotrijebljen alat Balsamiq Mockup¹.

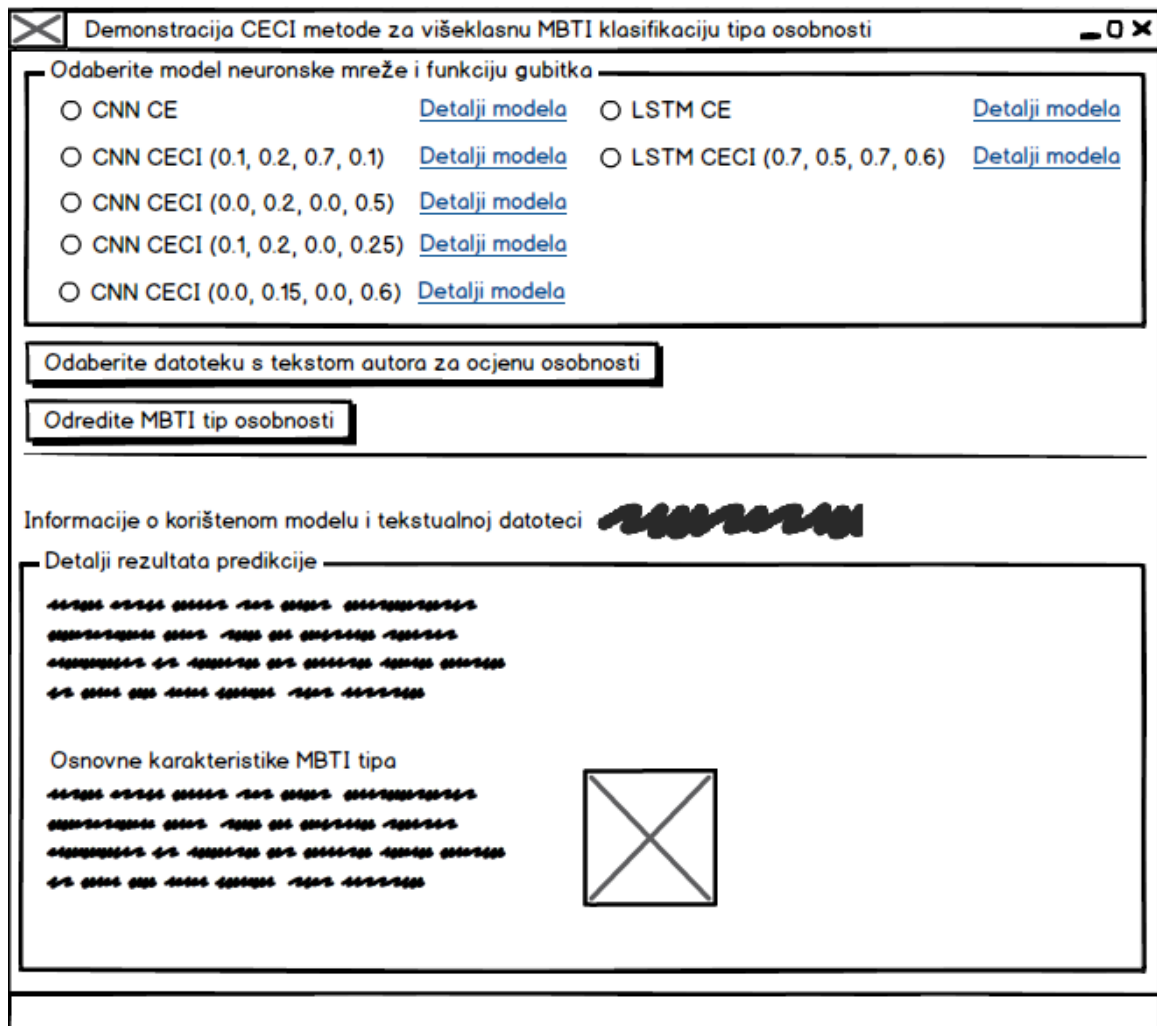
U prvom dijelu sučelja aplikacije predviđeno je da korisnik aplikacije može odabrati željeni model između ponuđenih opcija LSTM i CNN modela. Korisnik može odabrati LSTM ili CNN model sa CE funkcijom gubitka ili funkcijom gubitka CECI, pri čemu je omogućeno korisniku da kod funkcije CECI odabere i različite opcije težinskih faktora.

Dodatno, aplikacija treba omogućiti korisniku pregled detalja svakog modela. Slika 63 daje pregled očekivanih detalja. Detalji svakog modela uključuju slike gubitka i preciznosti prilikom treniranja odabranog modela, osnovne informacije o modelu klasifikacije, ostvarene rezultate prilikom treniranja te dodatne informacije kao što su težinski faktori u funkciji gubitka i optimizator koji je korišten prilikom treniranja modela.

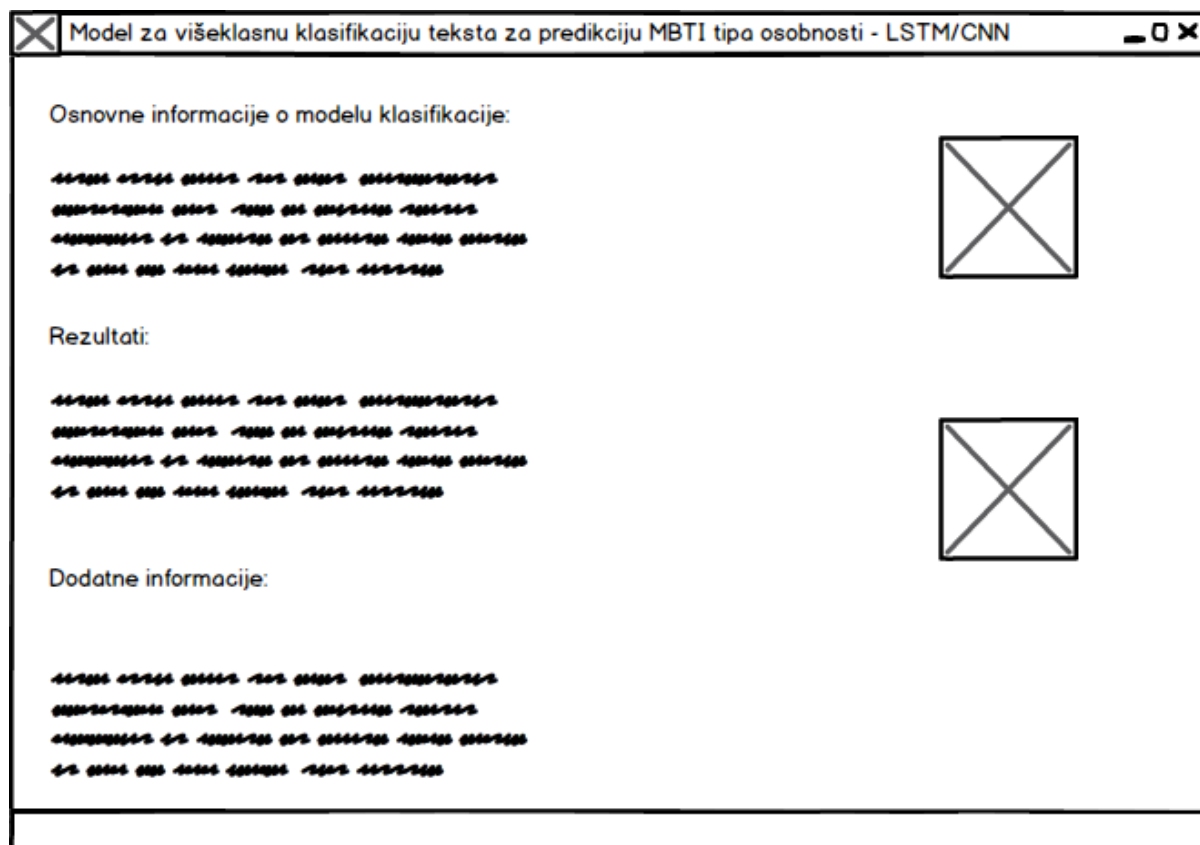
¹ <https://balsamiq.com/>

Polazni ekran aplikacije treba omogućiti korisniku da odabere tekstualnu datoteku za čiji sadržaj korisnik aplikacije želi napraviti predikcija tipa osobnosti MBTI. Konačno, aplikacija treba omogućiti pokretanje predikcije osobnosti na osnovu prethodno odabranih parametara: modela neuronske mreže i tekstualne datoteke.

Rezultat predikcije je potrebno prikazati s detaljima rezultata tipa osobnosti MBTI. U okviru rezultata je potrebno dati i osnovne karakteristike tipa MBTI koji je rezultat predikcije.



Slika 62 Mrežni okvir prototipa aplikacije – naslovna strana aplikacije



Slika 63 Mrežni okvir prototipa aplikacije - Detalji modela

Mrežni okvir aplikacije se lako može prilagoditi i koristiti za različite vrste aplikacija. S obzirom na tip i svrhu aplikacije, za prototip je odabran razvoj klijentske aplikacije u Python programskom jeziku i podržanim bibliotekama za razvoj korisničkog sučelja.

6.2. Arhitektura prototipa aplikacije

Python programski jezik pruža podršku nekoliko biblioteka za razvoj korisničkog sučelja kao što su PyQt5, Kivy, wxPython i PyGUI, a za razvoj prototipa je odabrana biblioteka Tkinter². U osmišljavanju arhitekture prototipa aplikacije ključnu ulogu imaju modeli strojnog učenja koji su kreirani prije treniranja modela, kako je opisano u poglavljima 5.3.1 i 5.3.2. Dodatno, potrebno je osigurati obradu teksta na isti način kako je obrađen tekst skupa podataka tijekom treniranja modela, kako bi se osiguralo vektorsku reprezentaciju riječi u istom formatu i time se ispunio preduvjet za uporabu modela treniranja u predikciji nepoznatog teksta. Zbog toga su korištene identične Python funkcije za predobradu teksta kao prilikom pripreme skupa podataka za treniranje modela. Dodatno, potrebno je iskoristiti

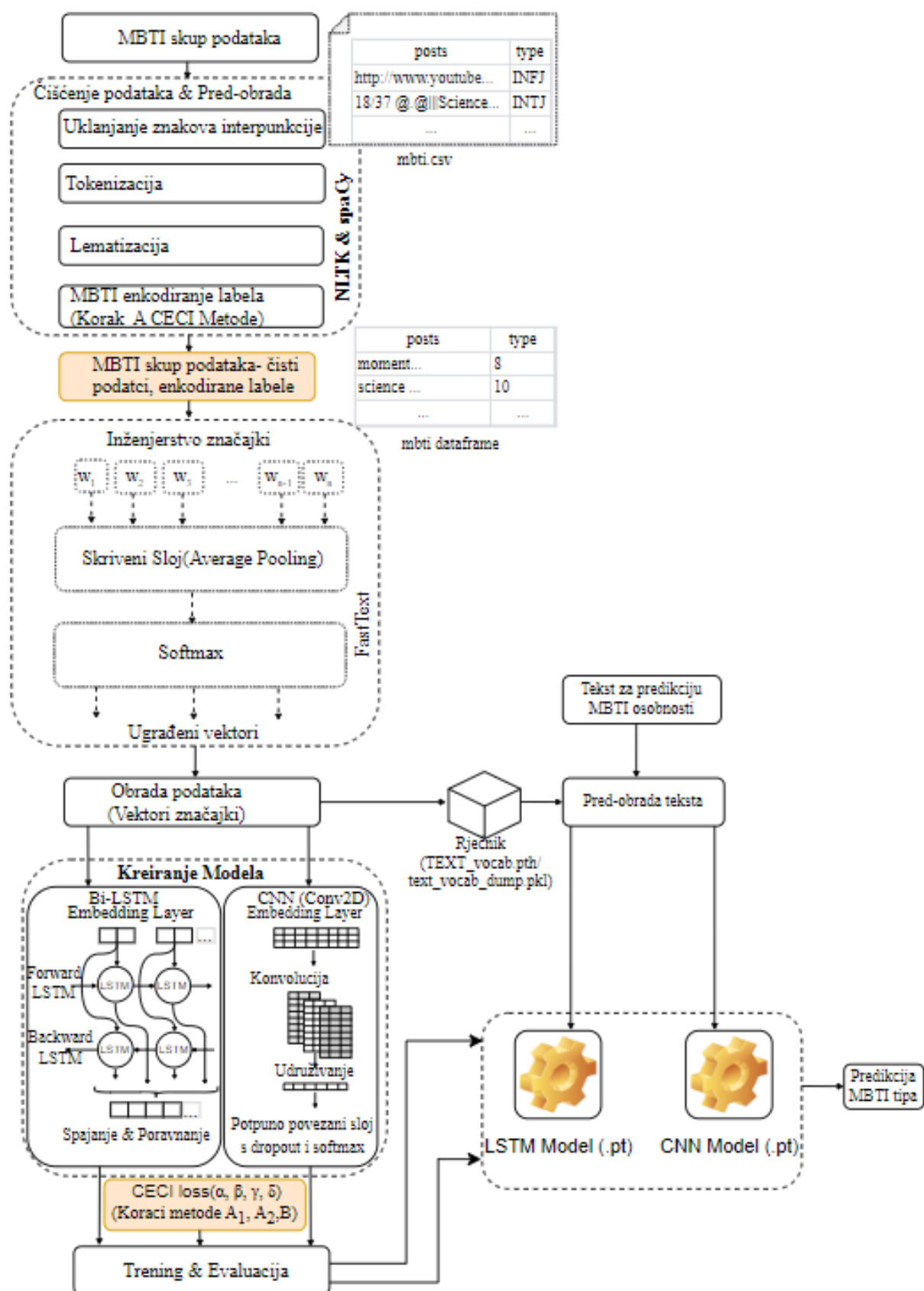
² <https://docs.python.org/3/library/tkinter.html#module-tkinter>

datoteku s vektorskom reprezentacijom riječi koja je kreirana tijekom pripreme skupa podataka za treniranje.

Slika 64 daje pregled postupka koji povezuje proces treniranja modela i implementacije modela za predikciju tipa MBTI iz nepoznatog tekstualnog sadržaja. Tijekom treniranja je potrebno spremići arhitekturu modela neuronske mreže LSTM i CNN. Rezultat se sprema odvojeno za svaki tip mreže i svaki trenirani model u datotečnom formatu koji je ovisan o okviru koji se koristi za treniranje. Budući da je za treniranje modela korišten okvir PyTorch, datoteka su spremljene u `.pt` format. Za CNN arhitekturu je spremljeno pet modela, a za LSTM arhitekturu dva modela.

Tijekom predobrade teksta, koji će predstavljati ulazne podatke za model treniranja neuronske mreže, potrebno je iz ulaznog skupa riječi kreirati rječnik vektora. Format datoteke rječnika vektora ovisi o upotrijebljenom pristupu i okviru. Kako je za potrebe treniranja upotrijebljen okvir PyTorch te biblioteka `torchtext`, rječnik vektora se sprema u datoteku formata `.pkl`.

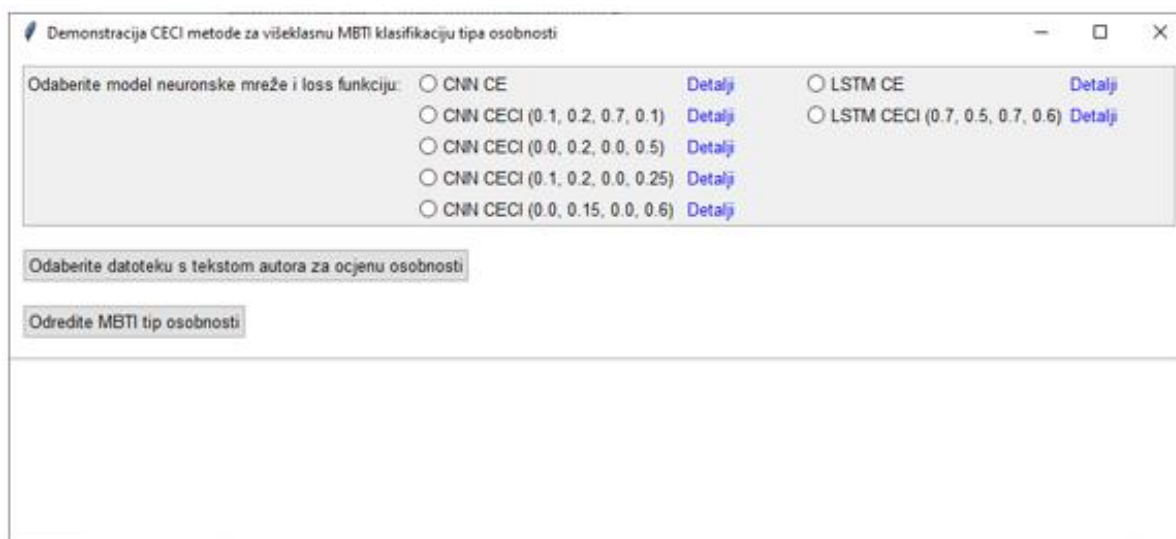
Dodatno, budući da je prilikom predikcije osobnosti iz nepoznatog teksta, isti potrebno predobraditi na isti način kao prilikom predobrade teksta polaznog skupa podataka, potrebno je spremići funkcije koje se koriste u fazi eksperimentiranja i uključiti ih u prototip aplikacije. Kako se radi o Python funkcijama, kôd je spremljen u datoteke s `.py` formatom.



Slika 64 Implementacija modela u prototipu aplikacije

6.3. Korisničko sučelje prototipa aplikacije

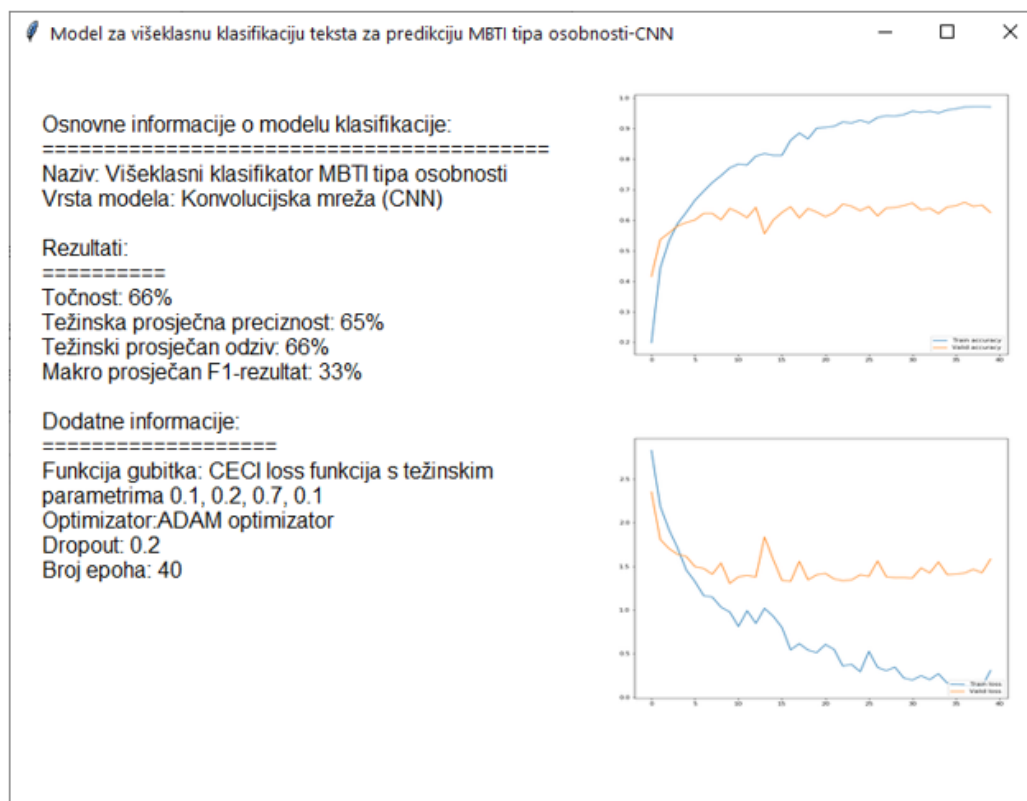
Korisničko sučelje je kreirano na osnovu smjernica statičnog mrežnog okvira te pregleda arhitekture za implementaciju modela u prototipu aplikacije. Slika 65 daje prikaz naslovne strane prototipa aplikacije. U prvom dijelu naslovne strane korisnik odabire, unutar CNN ili LSTM grupe arhitektura neuronske mreže, jedan od sedam istreniranih modela. Također, podržana je mogućnost odabira datoteke s tekstom autora za ocjenu osobnosti te mogućnost pokretanja predikcije iz odabranih parametara. Preostali dio naslovne strane sučelja je namijenjen prikazu rezultata predikcije.



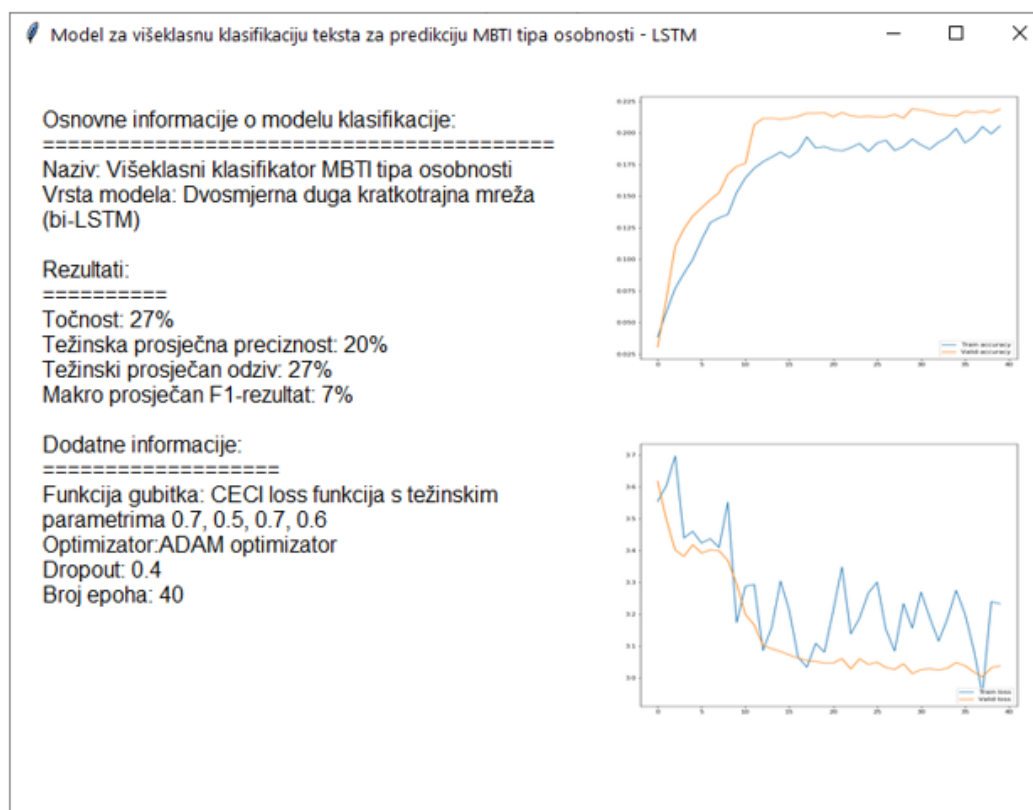
Slika 65 Naslovna strana prototipa aplikacije

Slika 66 daje prikaz detalja odabranog CNN CECI (0.1, 0.2, 0.7, 0.1) modela za predikciju tipa osobnosti iz proizvoljnog teksta. Slika 67 daje prikaz detalja odabranog LSTM (0.7, 0.5, 0.7, 0.6) modela. Prikazane su osnovne informacije o modelu klasifikacije, rezultati koje je model ostvario tijekom treniranja modela te dodatne informacije kao što su funkcija gubitka, broj epoha treniranja, vrijednost za nasumično ispuštanje i korišteni optimizator tijekom treniranja.

Dodatno, prikazani su dijagram ponašanja funkcije gubitka na skupu za treniranje i validaciju te dijagram točnosti na skupu treniranja i validaciju.

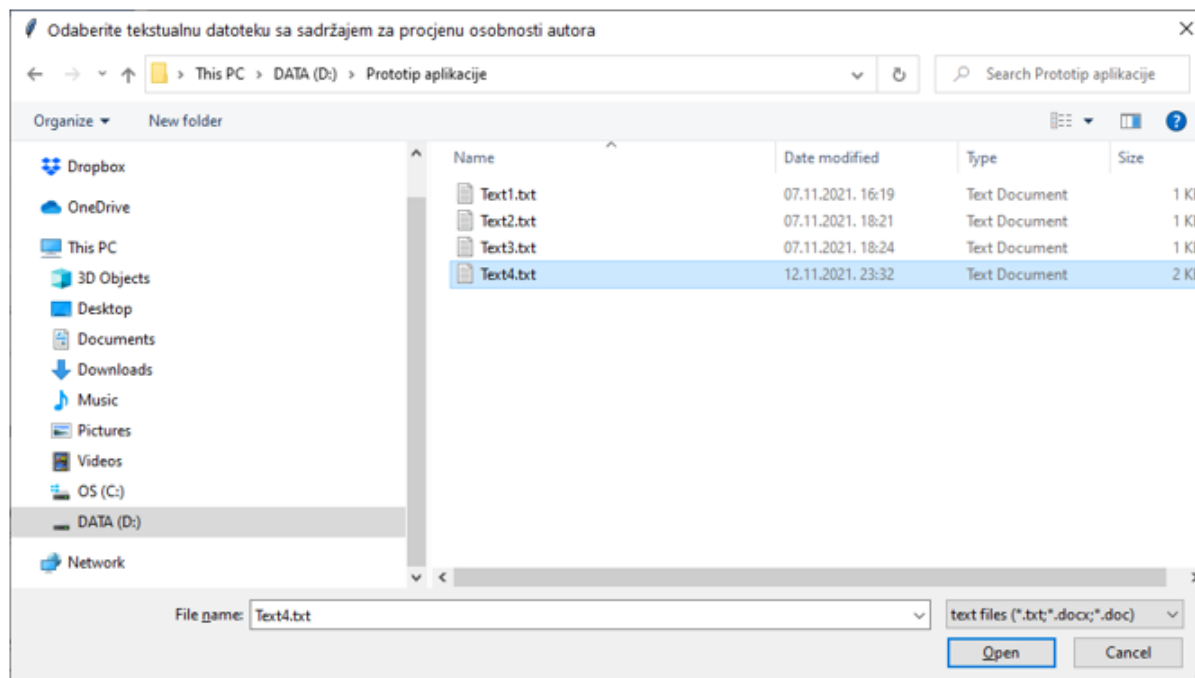


Slika 66 Detalji CNN modela s funkcijom gubitka CECI (0.1, 0.2, 0.7, 0.1)



Slika 67 Detalji LSTM modela s funkcijom gubitka CECI (0.7, 0.5, 0.7, 0.6)

Sljedeći korak prema predikciji tipa MBTI osobnosti je odabir tekstualne datoteke za čiji sadržaj želimo napraviti predikciju. Slika 68 daje prikaz uporabe prototipa aplikacije u ovom koraku.



Slika 68 Odabir tekstualne datoteke za predikciju MBTI

Konačno, prikazani su rezultati predikcije tipa osobnosti prema modelu MBTI za parametre: odabrani trenirani model, funkcija gubitka i tekstualna datoteku za čiji sadržaj želimo napraviti predikciju tipa osobnosti MBTI. Slika 69 daje primjer rezultata predikcije tipa osobnosti MBTI za CNN CECI (0.1, 0.2, 0.7, 0.1). Slika 70 daje rezultate predikcije tipa osobnosti MBTI za LSTM CECI (0.7, 0.5, 0.6, 0.7). Za tip MBTI su navedene osnovne karakteristike tipa, slikovito je dočaran arhetip koji opisuje tip te su navedene poznate ličnosti kojima je to dominantni tip osobnosti MBTI.

CNN CECI (0.1, 0.2, 0.7, 0.1) model je dao predikciju tipa osobnosti INFJ, a LSTM CECI (0.7, 0.5, 0.6, 0.7) je dao predikciju tipa osobnosti INFP. Razlika u predikciji tipa osobnosti se može očekivati s obzirom na razliku u rezultatima treniranja CNN i LSTM modela uz funkciju gubitka CECI.

Demonstracija CECI metode za višeklasnu MBTI klasifikaciju tipa osobnosti

Odaberite model neuronske mreže i loss funkciju:
☐ CNN CE
☒ CNN CECI (0.1, 0.2, 0.7, 0.1)
☐ CNN CECI (0.0, 0.2, 0.0, 0.5)
☐ CNN CECI (0.1, 0.2, 0.0, 0.25)
☐ CNN CECI (0.0, 0.15, 0.0, 0.6)

☐ LSTM CE
☐ LSTM CECI (0.7, 0.5, 0.7, 0.6)

Odaberite datoteku s tekстом autora za ocjenu osobnosti
Text4.txt

Odredite MBTI tip osobnosti

Model CNN CECI(0.1, 0.2, 0.7, 0.1) za tekstualni sadržaj u datoteci Text4.txt daje predikciju MBTI tipa osobnosti:

=====

INFJ - SAVJETNIK

=====

Osnovne karakteristike:

=====

Traže značenje i povezanost u idejama, odnosima i materijalnom posjedovanju. Žele razumjeti što motivira ljude i pronicljivi su o drugima. Savjesni i predani svojim čvrstim vrijednostima. Razvijaju jasnu viziju o tome kako najbolje služiti općem dobru. Organizirani i odlučni u provedbi svoje vizije.

Arhetip: Savjetnik

=====

INFJ su vizionari i idealisti. Ovaj tip osobnosti odiše kreativnom maštom i briljantnim idejama iz svake pore. Imaju drugačiji, često dubok, pogled na svijet koji nije uvijek shvaćen. INFJ preferiraju suštinu i dubinu u načinu na koji razmišljaju. Ovaj tip osobnosti nikada neće prihvatiti ništa na površinskoj razini ili odbiti prihvatiti bolji način pristupa problemima. Drugi mogu percipirati INFJ kao čudan ili zabavan zbog ovog promjenjivog pogleda na život.

Poznate osobe s INFJ tipom osobnosti:

=====

Peter Gabriel, Al Pacino, Noam Chomsky, Carl Jung, Platon, Leo Tolstoj

Izvori:

=====

<https://www.myersbriggs.org/>
<https://www.personalityclub.com/>
<https://www.personalityperfect.com/>



Slika 69 Predikcija tipa osobnosti MBTI autora teksta – CNN CECI

Demonstracija CECI metode za višeklasnu MBTI klasifikaciju tipa osobnosti

Odaberite model neuronske mreže i loss funkciju:
☐ CNN CE
☐ CNN CECI (0.1, 0.2, 0.7, 0.1)
☐ CNN CECI (0.0, 0.2, 0.0, 0.5)
☐ CNN CECI (0.1, 0.2, 0.0, 0.25)
☐ CNN CECI (0.0, 0.15, 0.0, 0.6)
☒ LSTM CE
☒ LSTM CECI (0.7, 0.5, 0.7, 0.6)

Odaberite datoteku s tekстом autora za ocjenu osobnosti
Text4.txt

Odredite MBTI tip osobnosti

Model LSTM CECI(0.7, 0.5, 0.7, 0.6) za tekstualni sadržaj u datoteci Text4.txt daje predikciju MBTI tipa osobnosti:

=====

INFP - IDEALIST

=====

Osnovne karakteristike:

=====

Idealisti, odani svojim vrijednostima i ljudima koji su im važni. Žele vanjski život koji je u skladu s njihovim vrijednostima. Znatiželjni, koji brzo vide mogućnosti, mogu biti katalizatori za provedbu ideja. Nastoje razumjeti ljude i pomoći im da ostvare svoj potencijal. Prilagodljivi, fleksibilni i prihvaćajući osim ako vrijednosti nisu ugrožene.

Arhetip: Idealist

=====

INFP-ovi su, kao i većina introverta, tihi i rezervirani. Ovaj tip osobnosti radije ne govori o sebi, osobito nakon prvog susreta s novom osobom. Više vole provoditi vrijeme sami na tihim mjestima. To daje priliku INFP-u da shvati svijet oko sebe. INFP osobe vole analizirati znakove i simbole, često ih smatraju metaforama koje imaju dublje značenje vezano za život. Ovaj tip osobnosti može se izgubiti u svojoj mašti i sanjarenjima. To povremeno dovodi do toga da se INFP utapa u dubini svojih misli, fantazija i ideja.

Poznate osobe s INFJ tipom osobnosti:

=====

William Shakespeare, George Orwell, J.R.R. Tolkien, Vincent Van Gogh, Andy Warhol

Izvori:

=====

<https://www.myersbriggs.org/>
<https://www.personalityclub.com/>
<https://www.personalityperfect.com/>



Slika 70 Predikcija tipa osobnosti MBTI autora teksta – LSTM CECI

7. ZAKLJUČAK

Određivanje tipa osobnosti predstavlja važnu aktivnost u poslovnim procesima, u kojima je cilj odabir osobe s tipom osobnosti koji najbolje odgovara idealnom kandidatu za ciljane poziciju ili aktivnost te odabir najboljeg kandidata može uštediti vrijeme i novac. Standardni pristupi uključuju procese koji su relativno spori jer nisu automatizirani te uključuju visoko obrazovane stručnjake koji moraju analizirati dobivene rezultate. Primjena metoda strojnog učenja u predikciji tipa osobnosti autora teksta može automatizirati postupak procjene osobnosti i doprinijeti konzistentnim rezultatima, stalnom poboljšanju rezultata te financijskim uštedama.

U doktorskoj disertaciji predložen je pristup višeklasnoj klasifikaciji prema modelu MBTI na način da se uzima u obzir komponente složenih oznaka klasa. Kreiran je algoritam za određivanje tipa osobnosti autora teksta temeljen na dubokom učenju s funkcijom gubitka koja uzima u obzir ~~uzima~~ mjere sličnosti između složenih oznaka klasa na način da se promatraju komponente složenih oznaka klasa koje su različite između složene oznake uzorka i složene oznake predikcije modela. U okviru algoritma demonstriran je pristup u kojem se koristi shema kodiranja kako bi se osigurao predložak ponavljanja pojedinih komponenti složenih oznaka klasa MBTI. Na taj način je osiguran konzistentan pristup komponentama složenih klasa. Nakon toga je pokazan način izračuna pojedinačnih vjerojatnosti za složene oznake klasa MBTI.

Dodatno, pokazana je metoda kojom se može odrediti ispravna komponenta i položaj komponente u složenoj oznaci klase MBTI u slučaju da postoji razlika između ciljane vrijednosti i predikcije modela. Također, pokazan je način izračuna vjerojatnosti za pojedinu komponentu složene oznake klasa. U sklopu algoritma je predložen i postupak izvođenja funkcije gubitka iz mjera vrednovanja temeljenih na sličnosti složenih oznaka klasa. Disertacija pokazuje kako se pojedinačne vjerojatnosti komponenti složenih oznaka klase mogu uključiti u funkciji gubitka duboke neuronske mreže. Prikazani su rezultati klasifikacije koji su usporedivi i nadmašuju rezultate za točnost, odziv i F1 – mjeru u dosadašnjim istraživanjima.

Nakon demonstracije algoritma za određivanje tipa osobnosti i postupka izvođenja funkcije gubitka demonstriran je prototip programskog sustava za određivanje tipa osobnosti sa

složenim oznakama klasa temeljen na dubokom učenju. Za kreiranje prototipa je korišten Python programski jezik s podržanim programskim bibliotekama i okvirima.

Kroz disertaciju se odgovorilo na pitanja u postavci znanstvenih očekivanja, budući da metoda CECI omogućava višeklasnu MBTI klasifikaciju uz uključivanje složenih oznaka klasa, ublažava problem preklapanja komponenti te ublažava problem neuravnoteženosti složenih klasa MBTI. Dodatno, napravljen je trening i evaluacija za dvije arhitekture umjetnih neuronskih mreža, LSTM i CNN te je napravljena usporedba tri kategorije rezultata: bazična koju su činili rezultati dobiveni uporabom klasične kategorične unakrsne entropije, rezultati dobiveni uporabom metode CECI te rezultati prijašnjih istraživanja. Korišteni su i standardni algoritmi strojnog učenja logističke regresije, slučajne šume i naivan Bayesov klasifikator.

Pristup CECI je pokazao poboljšanja u svim metrikama u usporedbi s osnovnim LSTM CE i CNN CE pristupima. Također, pristup CECI je poboljšao rezultate višeklasne klasifikacije MBTI prijašnjih istraživanja i dao usporedive rezultate s prijašnjim binarnim pristupima u klasifikaciji MBTI. Pojedini binarni pristupi su prezentirali nešto bolje rezultate. Pristup CNN CECI je pokazao bolje rezultate i u usporedbi s korištenim standardnim algoritmima strojnog učenja.

Ograničenja koja se trebaju uzeti u obzir za daljnja istraživanja uključuju činjenicu da su eksperimenti napravljeni na dva skupa podataka: skupu podataka s Reddit platforme te skupu podataka s Personality Cafe foruma. Dodatno, primijenjeni skupovi podataka su kreirani na dvije društvene mreže i sadrže samo tekstualne podatke. Stoga, eksperimenti na drugim skupovima podataka MBTI i s drugim tipovima podataka bi mogli dati nove poglede na primijenjenu metodu CECI i odnose između složenih oznaka klasa.

Naravno, moguća su daljnja unapređenja i primjene. Moguće je pristup CECI primijeniti na sličnim problemima klasifikacije sa složenim oznakama klasa i binarnim vrijednostima za svaku komponentu. Eksperimenti su napravljeni na skupovima podataka na engleskom jeziku i pristup koji bi uključio više jezika bi mogao dati nove perspektive. Također, koncept je pokazan na dva modela dubokih mreža: bi-LSTM i 2-dimenzionalni CNN te bi bili korisni eksperimenti s drugim modelima umjetnih neuronskih mreža. Promatranje klasifikacije osobnosti prema modelu MBTI pristupom klasifikacije s višestrukim (engl. *multilabel*) oznakama je također moguće područje budućih istraživanja.

Buduća istraživanja s metodom CECI će uključivati i eksperimente na bolje uravnoteženim skupovima podataka te uporabu različitih tehnika za upravljanje uravnoteženosti u skupu podataka MBTI. Dodatni smjer istraživanja će uključiti i pronalazak metode koja bi dodatno poboljšala automatizirano određivanje težinskih vrijednosti funkcije gubitka CECI. Na kraju, istraživanja će uključivati kognitivne funkcije i druge oblike relacija između komponenti tipova MBTI i težinskih faktora u implementaciji metode CECI nad skupom podataka MBTI.

Istraživanje provedeno u disertaciji je opravdalo očekivanja jer je demonstrirani algoritam uspješno implementirao komponente složenih oznaka klasa u višeklasnoj klasifikaciji prema modelu MBTI. Također, utjecaj komponenti složenih oznaka klasa je ugrađen u funkciju gubitka te je funkcija gubitka utjecala na poboljšavanje rezultata višeklasne klasifikacije MBTI na odabranim arhitekturama dubokog učenja. Na kraju, predloženi algoritam je implementiran unutar prototipa programskog sustava.

8. LITERATURA

- [1] S. R. Maddi, *Personality Theories : A Comparative Analysis*, Waveland Pr Inc; 6th edition, 2001.
- [2] M. Mitchell, K. Hollingshead and G. and Coppersmith, "Quantifying the language of schizophrenia in social media," in *In Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*, Denver, Colorado, 2015.
- [3] M. Papurt, "A study of the woodworth psychoneurotic inventory with suggested revision," *The Journal of Abnormal and Social Psychology*, vol. 25, no. 3, p. 335, 1930.
- [4] Y. Mehta, N. Majumder, A. Gelbukh and a. E. Cambria, "Recent trends in deep learning-based personality detection," *Artificial Intelligence Review*, vol. 53, p. 2313–2339, 2020.
- [5] D. Liu, Y. Li and a. M. A. Thomas, "A Roadmap for Natural Language Processing Research in Information Systems," in *Hawaii International Conference on System Sciences*, Honolulu, 2017.
- [6] N. L. Quenk, *Essentials of Myers-Briggs Type Indicator Assessment*, New Jersey: Wiley, 2009.
- [7] S. J. Read, B. M. Monroe, A. L. Brownstein, G. C. Yu Yang and L. Miller, "A neural network model of the structure and dynamics of human personality," *Psychological Review*, vol. 117, no. 1, pp. 61-92, 2010.
- [8] S. Ritter, D. G. Barrett, A. Santoro and M. M. Botvinick, "Cognitive Psychology for Deep Neural Networks: A Shape Bias Case Study," in *Proceedings of the 34th International Conference on Machine Learning*, Sydney, Australia, 2017.
- [9] "ImageNet," [Online]. Available: <http://www.image-net.org/>. [Accessed 18 03 2018].
- [10] R. J. Larsen and D. M. Buss, *Personality Psychology: Domains of Knowledge about Human Nature*, New York: McGraw-Hill Education, 2018.
- [11] R. B. Ewen, *An Introduction to Theories of Personality*, New York: Psychology Press, 2010.
- [12] C. G. Jung, *Contribution to Analytical Psychology*, Hesperides Press, 2008.
- [13] D. P. Schultz and S. E. Schultz, *Theories of Personality*, 11th ed., Boston: Cengage

- Learning, 2016.
- [14] L. V. Berens, Dynamics of Personality Type : Understanding and Applying Jung's Cognitive Processes, Telos Publications, 2000.
- [15] C. Jung, Psychological Types, Routledge; 1 edition (October 1, 2016), 2016.
- [16] I. B. Weiner and R. L. Greene, Handbook of Personality Assessment, New Jersey: John Wiley & Sons, 2017.
- [17] L. R. Goldberg, "An Alternative "Description of Personality": The Big-Five Factor Structure," *Personality and Personality Disorders: The Science of Mental Health*, vol. 7, no. 34, 2013.
- [18] A. M. Bland, "The Enneagram: A Review of the Empirical and Transformational Literature," *Journal of Humanistic Counseling, Education and Development*, vol. 49, no. 1, pp. 16-31, 2010.
- [19] D. Shaffer, M. Schwab-Stone, P. Fisher, P. Cohen, J. Piacentini, M. Davies, C. K. Conners and D. Regier, "The Diagnostic Interview Schedule for Children-Revised Version (DISC-R): I. Preparation, Field Testing, Interrater Reliability, and Acceptability," *Journal of the American Academy of Child & Adolescent Psychiatry (J Am Acad Child Adolesc Psychiatry)*, vol. 32, no. 3, pp. 643-650, 1993.
- [20] H. Ahmad, D. M. Z. Asghar, D. M. Z. Asghar, A. S. Khan and A. Habib, "A Systematic Literature Review of Personality Trait Classification from Textual Content," *Open Computer Science*, vol. 10, no. 1, pp. 175-193, 2020.
- [21] B. Agarwal, "Personality detection from the text: A review," *International Journal of Computer System I*, vol. 1, no. 1, pp. 1-4, 2014.
- [22] C. Soto, "Big Five personality traits," in *The SAGE Encyclopedia of Lifespan Human Development; Borstein, M.H., Arterberry, M.E., Fingerman, K.L., Lansford, J.E.*, CA, USA, SAGE Publications, Thousand Oaks, 2018, p. 240–241.
- [23] M. P. Myers I., Gifts Differing, Palo Alto: Consulting Psychologists Press, Incorporated, 1990.
- [24] C.G.Jung, Psychological Types (The Collected Works of C. G. Jung, Vol. 6), Princeton University Press, 1976.
- [25] T. L. Bess and R. J. Harvey, "Bimodal score distributions and the Myers–Briggs type indicator. Fact or artefact?," *Journal of Personality Assessment*, vol. 78, no. 1, p. 176–186, 2002.

- [26] P. Costa, R. R. McCrae and D. A. Dye, "Facet scales for agreeableness and conscientiousness: A revision of the neo personality inventory," *Personality and Individual Differences*, vol. 12, no. 9, p. 887–898, 1991.
- [27] A. Furnham, "The big five versus the big four: the relationship between the Myers-Briggs Type Indicator (MBTI) and NEO-PI five-factor model of personality," *Personality and Individual Differences*, vol. 1, no. 2, p. 303–307, 1996.
- [28] F. Celli and B. Lepri, "Is Big Five better than MBTI? A personality computing challenge using Twitter data," in *Proceedings of the Fifth Italian Conference on Computational Linguistics CLiC-it*, Torino, Italy, 2018.
- [29] J. Beebe, *Energies and Patterns in Psychological Type*, Routledge, 2016.
- [30] L. Haas and M. Hunziker, *Building Blocks of Personality Type: A Guide to Using the Eight-Process Model of Personality Type*, Huntington Beach: Unite Business Press, A Division of Telos Publications, 2006.
- [31] J. Snajder and B. D. Bašić, "www.fer.hr - Skripta za kolegij Strojno učenje," 2014. [Online]. Available: https://www.fer.unizg.hr/_download/repository/StrojnoUcenje.pdf. [Accessed 2 10 2021].
- [32] I. Goodfellow, Y. Bengio and A. Courville, *Deep Learning*, Cambridge: The MIT Press, 2017.
- [33] S. Šegvić, "Neslužbene stranice predmeta Duboko učenje," [Online]. Available: <http://www.zemris.fer.hr/~ssegvic/du/>. [Accessed 10 11 2021].
- [34] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation Archive*, vol. 9, no. 8, p. 1735–1780, 1997.
- [35] Y. Bengio, P. Simard and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," *IEEE Transactions on Neural Networks*, vol. 5, no. 2, pp. 157-166, 1994.
- [36] J. Hochreiter, "Untersuchungen zu dynamischen neuronalen Netzen," DIPLOMARBEIT IM FACH INFORMATIK - Technische Universität München, <https://people.idsia.ch/~juergen/SeppHochreiter1991ThesisAdvisorSchmidhuber.pdf>, München, 1991.
- [37] X. Wang, Y. Liu, C. Sun, B. Wang and X. Wang, "Predicting Polarities of Tweets by Composing Word Embeddings with Long Short-Term Memory," in *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*; DOI: 10.3115/v1/P15-1130, Beijing, China, 2015.

- [38] S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," in *ICML*, Lille, 2015.
- [39] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *AISTATS*, Sardinia, Italy, 2010.
- [40] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," *IEEE Transactions on Signal Processing*, vol. 45, no. 11, pp. 2673-2681, 1997.
- [41] M. Frković, N. Čerkez, B. Vrdoljak and S. Skansi, "Evaluation of Structural Hyperparameters for Text Classification with LSTM Networks," in *Mipro*, Opatija, 2020.
- [42] N. Cerkez and V. Vareškić, "Machine Learning Approaches to Personality Classification on Imbalanced MBTI Datasets," in *44th International Convention on Information, Communication and Electronic Technology (MIPRO)*, 2021, pp. 1259-1264, doi: 10.23919/MIPRO52101.2021.9596742, Opatija, 2021.
- [43] Y. Ma and H. He, *Imbalanced Learning: Foundations, Algorithms, and Applications*, New Jersey: Wiley-IEEE Press, 2013.
- [44] *Learning from Imbalanced Data Sets*, Springer, 2018.
- [45] J. R. Landis and G. G. Koch, "The Measurement of Observer Agreement for Categorical Data," *Biometrics*, pp. 159-174, 3 1977.
- [46] N. V. Chawla, K. W. Bowyer, L. O. Hall and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *Journal of Artificial Intelligence Research* 16 , p. 321–357, 2002.
- [47] H. Han, W.-Y. Wang and B.-H. Mao, "Borderline-SMOTE: A New Over-Sampling Method in Imbalanced Data Sets Learning," in *international conference on Advances in Intelligent Computing*, Hefei China, 2005.
- [48] J. V. Hulse, T. Khoshgoftaar and A. Napolitano, "Experimental Perspectives on Learning from Imbalanced Data," in *Proceedings of the Twenty-Fourth International Conference (ICML 2007)*, DOI: 10.1145/1273496.1273614, Corvallis, Oregon, USA, 2007.
- [49] K. Fukushima, "Neocognitron: A Self-Organizing Neural Network Model," *Biological*, vol. 36, no. 4, p. 193–202, 1980.
- [50] Y. LeCun, B. Boser, J. S. H. Denker, D. Howard and W. R. E. Hubbard, "Handwritten Digit Recognition with a Back-Propagation," in *CO: Morgan Kaufmann*, 396–404, Denver, 1990.

- [51] Y. LeCun, Y. Bengio and G. Hinton, "Deep learning," *Nature*, no. 521, pp. 436-444, 2015.
- [52] Y. LeCun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-based learning applied to document recognition," *Proc IEEE*, no. 86, pp. 2278-2324, 1998.
- [53] D. Scherer, A. Müller and S. Behnke, "Evaluation of pooling operations in convolutional architectures for object recognition," in *In Proceedings of the Artificial Neural Networks–ICANN 2010, 15–18 September 2010; pp. 92–101.*, Thessaloniki, Greece,, 2010.
- [54] G. Salton and C. Buckley, "Term-weighting approaches in automatic text retrieval," *Inf. Process. Manag.*, no. 24, pp. 513-523, 1988.
- [55] Y. Goldberg and O. Levy, "Word2vec explained: Deriving Mikolov et al.'s negative-sampling word-embedding method," 2014. [Online]. Available: arXiv:1402.3722.. [Accessed 18 11 2021].
- [56] J. Pennington, R. Socher and C. Manning, "Glove: Global Vectors for Word Representation," in *In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Doha, Qatar, 25–29 October 2014; Volume 14, pp. 1532–1543.
- [57] P. Bojanowski, E. Grave, A. Joulin and T. Mikolov, "Enriching word vectors with subword information," arXiv:1607.04606., 2016.
- [58] J. Huang i C. Ling, »Using AUC and accuracy in evaluating learning algorithms,« *IEEE Trans. Knowl. Data Eng.*, br. 17, p. 299–310, 2005.
- [59] J. Hanley and B. McNeil, "The meaning and use of the area under a receiver operating characteristic (ROC) curve," *Radiology*, no. 143,, p. 29–36, 1982.
- [60] M. Pencina, R. D'Agostino and R. Vasan, "Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond," *Stat. Med.*, no. 27, p. 157–172, 2008.
- [61] J. Šnajder, "Strojno učenje, predavanja, ak. god. 2021./2022.," 2021. [Online]. Available: https://www.fer.unizg.hr/_download/repository/SU1-2021-P21-VrednovanjeModela.pdf. [Accessed 22 03 2022].
- [62] J. Opitz and S. Burst, "Macro F1 and Macro F1," 8 11 2019. [Online]. Available: <https://arxiv.org/abs/1911.03347>. [Accessed 21 05 2021].
- [63] L. Tay, S. E. Woo, L. Hickman and R. M. Saef, "Psychometric and Validity Issues in Machine Learning Approaches to Personality Assessment: A Focus on Social Media Text Mining," *European Journal of Personality*, DOI: 10.1002/per.2290, vol. 34, no. 5,

- pp. 826-844, 2020.
- [64] C. S. Cloninger, *Theories of Personality*, 4th edition, New Jersey: Pearson Prentice Hall, 2004.
- [65] H. K. Mohammad Hossein Amirhosseini, "Machine Learning Approach to Personality Type Prediction Based on the Myers–Briggs Type Indicator," *Multimodal Technol. Interact.*, vol. 4(1), no. 9, 2020.
- [66] C. Li, M. Hancock, B. Bowles, O. Hancock, L. Perg, P. Brown, A. Burrell, G. Frank, F. Stiers, S. Marshall, G. Mercado, A.-W. Ahmed, P. Beckelheimer, S. Williamson and R. Wade, "Feature extraction from social media posts for psychometric typing of participants," *Augmented Cognition: Intelligent Technologies Lecture Notes in Computer Science*, pp. 267-286, 2018.
- [67] H. P. Kishan Das, "Personality identification based on MBTI dimensions using natural language processing," *International Journal of Creative Research Thoughts (IJCRT)*, vol. 8, no. 6, pp. 1653-1657, 2020.
- [68] S. Bharadwaj, S. Sridhar, R. Choudhary and R. Srinath, "Persona Traits Identification based on Myers-Briggs Type Indicator(MBTI) - A Text Classification Approach," in *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, Bangalore, 2018.
- [69] M. Komisin and C. Guinn, "Identifying personality types using document classification methods," in *25th International Florida Artificial Intelligence Research Society Conference*, pp. 232–237, Marco Island, FL, USA, 23–25 May 2012.
- [70] N. H. Z. Abidin, M. A. Remli, N. M. Ali, D. N. E. Phon, N. Yusoff, H. K. Adli and A. H. Busalim, "Improving Intelligent Personality Prediction using Myers-Briggs Type Indicator and Random Forest Classifier," (*IJACSA*) *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 11, pp. 192-199, 2020.
- [71] Z. Mushtaq, S. Ashraf and N. Sabahat, "Predicting MBTI Personality type with K-means Clustering and Gradient Boosting," in *23rd International Multitopic Conference (INMIC)*, Bahawalpur, Pakistan, 2020.
- [72] S. Chaudhary¹, R. Singh, S. T. Hasan and M. I. Kaur, "A Comparative Study of Different Classifiers for Myers-Brigg Personality Prediction Model," *International Research Journal of Engineering and Technology - IRJET*, vol. 05, no. 05, pp. 1410-1413, 2018.
- [73] G. L. Anthony Ma, "Neural Networks in Predicting Myers Briggs Personality Type from Writing Style," Stanford, Palo Alto, 2017.

- [74] X. Q. X. H. Pengfei Liu, "Recurrent Neural network for Text Classification with Multi-Task learning," Fundan University, Shanghai, 2016.
- [75] Y. Doval, C. Gomes-Rodriguez and J. Vilares, "Shallow recurrent NN for personality recognition in source code," in *FIRE 2016 - Forum for Information Retrieval Evaluation*, Kolkata, India, 2016.
- [76] C. H. N and D. K. R. AnandaKuma, "Artificial-Neural-Network-for-Human-Behavior-Prediction-through-Handwriting-Analysi," *International Journal Of Computer Applications* (0975-8887), vol. 2, no. 2, 2010.
- [77] T. Ichimura, K. Tanabe and T. Yamashita, "An Adaptive Learning Method of Personality Trait Based Mood in Mental State Transition Network by Recurrent Neural Network," in *Computational Intelligence and Applications (IWCIA), 2014 IEEE 7th International Workshop on Computational Intelligence and Applications (IWCIA)*, Hiroshima, Japan, 2014.
- [78] S. Barrett, "How Much Data Is Produced Every Day in 2021," 30 05 2021. [Online]. Available: <https://the-tech-trend.com/reviews/how-much-data-is-produced-every-day/>. [Accessed 28 07 2021].
- [79] N. Patel, "17 charts that show where social media is heading," 2021. [Online]. Available: <https://neilpatel.com/blog/social-media-trends/>. [Accessed 26 07 2021].
- [80] N. Cerkez, B. Vrdoljak and S. Skansi, "The Present Situation and the Prospect of Determining the Personality Type of Text Author with Machine Learning," in *Softcom, Split*, 2019.
- [81] M. Gjurković and J. Šnajder, "Reddit: A Gold Mine for Personality Prediction," in *Proceedings of the Second Workshop on Computational Modeling of People's Opinions, Personality, and Emotions in Social Media*, New Orleans, 2018.
- [82] D. S. Nasser Alsadhan, "Estimating Personality from Social Media Post," in *2017 IEEE International Conference on Data Mining Workshops (ICDMW)*, New Orleans, LA, USA, 2017.
- [83] B. Plank and D. Hovy, "Personality traits on Twitter or how to get 1,500 personality tests in a week," in *Proceedings of the 6th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, Lisboa, Portugal, 2015.
- [84] J. Golbeck, C. Robles, M. Edmondson and K. Turner, "Predicting Personality from Twitter," in *2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing*, Boston, MA, USA, 2011.

- [85] S. Garg, P. Kumar and A. Garg, "Comparison of Machine Learning Algorithms for Content-Based Personality Resolution of Tweets," *Social Sciences & Humanities Open*, no. <http://dx.doi.org/10.2139/ssrn.3626306>, 2020.
- [86] M. Gjurković, M. Karan, I. Vukojević, M. Bošnjak and J. Šnajder, "PANDORA Talks: Personality and Demographics on Reddit," in *Proceedings of the Ninth International Workshop on Natural Language Processing for Social Media, NAACL 2021*, Online, 2021.
- [87] D. Wan, C. Zhang, M. Wu and Z. An, "Personality Prediction Based on All Characters of User Social Media Information," in *Chinese National Conference on Social Media Processing*, Beijing, China, 2014.
- [88] L. C. Lukito, A. Erwin, J. Purnama and W. Danoekoesoemo, "Social media user personality classification using computational linguistic," in *8th International Conference on Information Technology and Electrical Engineering*, Yogyakarta, Indonesia, 2016.
- [89] A. S. Khan, H. Ahmad, M. Z. Asghar, F. K. Saddozai, A. Arif and H. A. Khalid, "Personality Classification from Online Text using Machine Learning Approach," (*IJACSA*) *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 3, pp. 460-476, 2020.
- [90] E. J. Choong and K. D. Varathan, "Predicting judging-perceiving of Myers-Briggs Type Indicator (MBTI) in an online social forum," *PeerJ* 9:e11382, <https://doi.org/10.7717/peerj.11382>, 2021.
- [91] F. Piedboeuf, P. Langlais and L. Bourg, "Personality Extraction Through LinkedIn," in *Advances in Artificial Intelligence, 32nd Canadian Conference on Artificial Intelligence, Canadian AI 2019*, Kingston, ON, Canada, May 28-31, 2019.
- [92] V. Lynn, N. Balasubramanian and a. H. A. Schwartz, "Hierarchical modelling for user personality prediction: The role of message-level attention," in *In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 5306–5316, *Association for Computational Linguistics*, Online, 2020.
- [93] B. Verhoeven, W. Daelemans and a. B. Plank, "TwiSty: a multilingual twitter stylometry corpus for gender and personality profiling," in *10th edition of the Language Resources and Evaluation Conference, LREC*, pages 1632–1637, Portoroz, Slovenia, 2016.
- [94] F. Liu, J. Perez and S. Nowson, "A language-independent and compositional model for personality trait recognition from short texts," in *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1*,

- Long Papers, pages 754–764, 2017, Valencia, Spain, 2017.*
- [95] B. Y. Pratama and R. Sarno, "Personality classification based on Twitter text using Naïve Bayes, KNN and SVM," in *2015 International Conference on Data and Software Engineering (ICoDSE)*, Yogyakarta, Indonesia, 2015.
 - [96] N. Majumder, S. Poria, A. Gelbukh and a. E. Cambria, "Deep Learning-Based Document Modeling for Personality Detection from Text," *IEEE Intelligent Systems*, vol. 32, no. 2, pp. 74-79, Mar.-Apr. 2017.
 - [97] F. Mairesse, M. A. Walker, M. R. Mehl and R. K. Moore, "Using Linguistic Cues for the Automatic Recognition of Personality in Conversation and Text," *J. Artificial Intelligence Research*, vol. 30, p. 457–500, 2007.
 - [98] E. P. Tighe, J. C. Ureta, B. A. L. Pollo, C. K. Cheng and a. R. d. D. Bulos, "Personality Trait Classification of Essays with the Application of Feature Reduction," in *Proceedings of the 4th Workshop on Sentiment Analysis where AI meets Psychology (SAAIP 2016), IJCAI 2016*, New York, USA, 2016.
 - [99] E. Ezpeleta, I. V. d. Mendizabal, J. M. G. Hidalgo and U. Zurutuza, "Novel email spam detection method using sentiment analysis and personality recognition," *Logic Journal of IGPL*, vol. 28, no. 1, p. 83–94, 202.
 - [100] E. Ezpeleta, U. Zurutuza and J. M. Gómez Hidalgo, "Using Personality Recognition Techniques to Improve Bayesian Spam Filtering," *Procesamiento del Lenguaje Natural*, vol. 57, pp. 125-132, 2016.
 - [101] D. J. F. R. S. Mohammad Noroui, "Hamming distance metric learning," in *Advances in Neural Information Processing Systems 25 (NIPS 2012)*, Toronto, Canada, 2012.
 - [102] L. Fan, "Revisit Fuzzy Neural Network: Demystifying Batch Normalization and ReLU with Generalized Hamming Network," in *Advances in Neural Information Processing Systems 30 (NIPS 2017)*, Toronto, Canada, 2017.
 - [103] L. Li, A. Li, B. Hao, Z. Guan and T. Zhu, "Predicting Active Users' Personality Based on Micro Blogging Behaviors," *plos.org*, 2014.
 - [104] M. R. D. N. S. S. Mayuri Pundlik Kalghatgi, "A Neural Network Approach to Personality Prediction based on the Big-Five Model," *International Journal of innovative Research in Advanced Engineering (IJIRAE)*, vol. 2, no. 8, August 2015.
 - [105] H. K. Mohammad Hossein Amirhosseini, "Machine Learning Approach to Personality Type Prediction Based on the Myers–Briggs Type Indicator," *Multimodal Technol. Interact.*, vol. 4(1), no. 9, 220.

- [106] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Information Processing & Management*, vol. 45, no. 4, pp. 427-437, 2009.
- [107] L. C. Lukito, A. Erwin, J. Purnama and W. Danoekoesoemo, "Social media user personality classification using computational linguistic," in *8th International Conference on Information Technology and Electrical Engineering*, pp. 1-6, Yogyakarta, Indonesia, Oct. 2016.
- [108] N. Cerkez, B. Vrdoljak and S. Skansi, "A Method for MBTI Classification based on Impact of Class Components," *IEEE Access*; doi:10.1109/access.2021.3121137, vol. 9, pp. 146550-146567, 2021.
- [109] C. E. Shannon, "A Mathematical Theory of Communication," *The Bell System Technical Journal*, vol. 27, no. July, October, pp. 379–423, 623–656, 1948.
- [110] "Kaggle - MBTI dataset," [Online]. Available: <https://www.kaggle.com/datasnaek/mbti-type/download>. [Accessed 01 03 2021].
- [111] T. M. & B. Foundation, "The Myers & Briggs Foundation," 2021. [Online]. Available: <https://www.myersbriggs.org/my-mbti-personality-type/my-mbti-results/how-frequent-is-my-type.htm>. [Accessed 08 07 2021].

9. POPIS KRATICA

KRATICE

AUC: površina ispod ROC krivulje (engl. area under ROC curve).....	36
Bi-LSTM: Dvosmjerna duga kratkoročna memorija (engl. Bi-directional long short-term memory)	21
CE: kategorička unakrsna entropija (eng. categorical cross-entropy).....	62
CECI: unakrsna entropija s utjecajem komponenti složenih oznaka klasa (engl. cross-entropy compound class-label impact).....	64
CNN: konvolucijska neuronska mreža (engl. convolutional neural network)	4
DNN: mreže s višestrukim skrivenim slojevima (engl. deep neural networks).....	18
DT: stablo odlučivanja (engl. decision tree)	44
FFNN: unaprijedna neuronska mreža (engl. feed-forward neural network).....	17
GloVe: Globalni vektori za predstavljanje riječi (engl. Global Vectors for Word Representation)	35
GPU: grafička procesorska jedinica (engl. graphical processing unit)	66
GRU: Rekurentna mreža s vratima (engl. Gated Recurrent Network)	21
LDA: Linearna razlikovna analiza (engl. Linear Discriminant Analysis)	36
LIWC: rječnik LIWC (engl. linguistic inquiry and word count)	41
LR: logistička regresija (engl. logistic regression).....	36
LSTM: duga kratkoročna memorija (engl. long short-term memory)	4
MBTI: indikator tipa osobnosti po Myers-Briggsovoj (engl. Myers-Briggs Type Indicator) ...	2
MCC: Matthews koeficijent korelacije (engl. Matthews Correlation Coefficient).....	36
NBC: Naivni Bayesov klasifikator (engl. Naïve Bayes Classifier)	36
NLP: obrada prirodnog jezika (engl. natural language processing).....	1
NMF: ne-negativna matrična faktorizacija (engl. non-negative matrix factorization)	36
PCA: Analiza glavne komponente (engl. Principal Component Analysis)	36
RF: slučajna šuma (engl. random forest)	44
RNN: rekurentne neuronske mreže (eng. recurrent neural networks)	18
ROC: operativne karakteristike primatelja (engl. receiver operating characteristics)	36
SGD: stohastički gradijentni spust (engl. stochastic gradient descent).....	44
SMOTE: metoda naduzorkovanja umjetnih manjinskih primjera (engl. synthetic minority oversampling technique)	26

SVD: dekompozicija na singularne vrijednosti (engl. singular value decomposition).....	42
SVD: dekompozicija singularne vrijednosti (engl. <i>singular value decomposition</i>).....	41
SVM: metoda potpornih vektora (engl. support vector machine).....	36
TF: Frekvencija izraza (engl. Term Frequency).....	35
TF-IDF: Frekvencija izraza – Inverzna frekvencija dokumenta (engl. Term Frequency-Inverse Document Frequency).....	35
t-SNE: t-distribuirano stohastičko ugrađivanje susjeda (engl. t-distributed stochastic neighbor embedding).....	36
Word2Vec: Riječ u vektor.....	35

10. POPIS SLIKA

Slika 1 Dimenzije osobnosti prema Big Five modelu.....	10
Slika 2 Model MBTI osobnosti.....	12
Slika 3 MBTI – pristup binarne klasifikacije	13
Slika 4 MBTI - pristup višeklasne klasifikacije	14
Slika 5 Primjer umjetne neuronske mreže.....	17
Slika 6 Duboko učenje i tradicionalne umjetne neuronske mreže	18
Slika 7 Rekurentna neuronska mreža	19
Slika 8 Čelija LSTM neuronske mreže	21
Slika 9 Mreža bi-LSTM	22
Slika 10 Gubitak validacije za funkciju gubitka unakrsne entropije [33] a).....	23
Slika 11 Gubitak validacije za funkciju gubitka unakrsne entropije [33] b).....	23
Slika 12 Gubitak validacije za funkciju gubitka unakrsne entropije [33] c).....	24
Slika 13 Gubitak validacije za funkciju gubitka unakrsne entropije [33] d).....	24
Slika 14 Semantička tekstualna sličnost klasa MBTI	28
Slika 15 Ideja konvolucije prozora nad funkcijom $f(x)$ (a).....	31
Slika 16 Ideja konvolucije prozora nad funkcijom $f(x)$ (b).....	31
Slika 17 Arhitektura konvolucijske neuronske mreže za klasifikaciju teksta	34
Slika 18 Standardni proces klasifikacije teksta	35
Slika 19 Matrica zabune za binarnu klasifikaciju	37
Slika 20 Algoritam za određivanje tipa MBTI osobnosti autora teksta	53
Slika 21 Binarni pristup klasifikaciji MBTI.....	54
Slika 22 Višeklasni pristup klasifikaciji MBTI.....	56
Slika 23 Graf standardne logističke sigmoidne funkcije.....	58
Slika 24 Graf Signum funkcije.....	60
Slika 25 Pregled procesa implementacije metode CECI.....	65
Slika 26 Izvadak iz skupa podataka MBTI s Personality Cafe foruma.....	67
Slika 27 Izvadak iz skupa podataka MBTI s Reddit platforme.....	68
Slika 28 Distribucija klasa u skupu podataka MBTI – Personality Cafe	69
Slika 29 Distribucija klasa u skupu podataka MBTI – Reddit	69
Slika 30 Duljina objave po tipu MBTI– Personality Cafe	72
Slika 31 Duljina objave po tipu MBTI– Reddit	73

Slika 32 Model LSTM.....	75
Slika 33 Gubitak treniranja/validacije za LSTM-CE na skupu MBTI s Personality Cafe forumu	75
Slika 34 Točnost treniranja/validacije za LSTM-CE na skupu MBTI s Personality Cafe forumu	76
Slika 35 Matrica zabune za LSTM-CE na skupu MBTI s Personality Cafe foruma	78
Slika 36 Gubitak treniranja/validacije za LSTM-CECI na skupu MBTI s Personality Cafe forumu	79
Slika 37 Točnost treniranja/validacije za LSTM-CECI na skupu MBTI s Personality Cafe forumu	79
Slika 38 Matrica zabune za LSTM-CECI na skupu MBTI s Personality Cafe foruma	81
Slika 39 Interval 95% pouzdanosti bi-LSTM CE/CECI na skupu MBTI s Personality Cafe forumu	82
Slika 40 Gubitak treniranja/validacije za LSTM-CE na skupu MBTI s Reddit platforme	82
Slika 41 Točnost treniranja/validacije za LSTM-CE na skupu MBTI s Reddit platforme	83
Slika 42 Matrica zabune za LSTM-CE na skupu MBTI s Reddit platforme	83
Slika 43 Gubitak treniranja/validacije za LSTM-CECI na skupu MBTI s Reddit platforme ..	85
Slika 44 Točnost treniranja/validacije za LSTM-CECI na skupu MBTI s Reddit platforme ..	85
Slika 45 Matrica zabune za LSTM-CECI na skupu MBTI s Reddit platforme	87
Slika 46 Interval 95% pouzdanosti bi-LSTM CE/CECI na skupu MBTI s Reddit platforme .	88
Slika 47 Model CNN.....	88
Slika 48 Trening/validacijski gubitak za CNN-CE na skupu MBTI s Reddit platforme.....	89
Slika 49 Trening/validacijska točnost za CNN-CE na skupu MBTI s Reddit platforme.....	89
Slika 50 Matrica zabune za CNN-CE na skupu MBTI s Reddit platforme	91
Slika 51 Trening/validacijski gubitak za CNN-CECI na skupu s Reddit platforme	92
Slika 52 Trening/validacijska točnost za CNN-CECI na skupu s Reddit platforme.....	92
Slika 53 Matrica zabune za CNN-CECI na skupu MBTI s Reddit platforme	94
Slika 54 Interval 95% pouzdanosti CNN CE/CECI na skupu MBTI s Reddit platforme.....	95
Slika 55 Trening/validacijski gubitak za CNN-CE na skupu MBTI podataka s Personality Cafe foruma.....	96
Slika 56 Trening/validacijska točnost za CNN-CE na skupu MBTI podataka s Personality Cafe foruma.....	96
Slika 57 Matrica zabune za CNN-CE na skupu MBTI podataka s Personality Cafe foruma ..	98

Slika 58 Trening/validacijski gubitak za CNN-CECI na skupu podataka s Personality Cafe foruma	99
Slika 59 Trening/validacijska točnost za CNN-CECI na skupu podataka s Personality Cafe foruma	99
Slika 60 Matrica zabune za CNN-CECI na skupu podataka s Personality Cafe foruma	101
Slika 61 Interval 95% pouzdanosti CNN CE/CECI na skupu MBTI s Personality Cafe foruma	102
Slika 62 Mrežni okvir prototipa aplikacije – naslovna strana aplikacije	107
Slika 63 Mrežni okvir prototipa aplikacije - Detalji modela.....	108
Slika 64 Implementacija modela u prototipu aplikacije.....	110
Slika 65 Naslovna strana prototipa aplikacije	111
Slika 66 Detalji CNN modela s funkcijom gubitka CECI (0.1, 0.2, 0.7, 0.1).....	112
Slika 67 Detalji LSTM modela s funkcijom gubitka CECI (0.7, 0.5, 0.7, 0.6).....	112
Slika 68 Odabir tekstualne datoteke za predikciju MBTI	113
Slika 69 Predikcija tipa osobnosti MBTI autora teksta – CNN CECI	114
Slika 70 Predikcija tipa osobnosti MBTI autora teksta – LSTM CECI.....	115

11. POPIS TABLICA

Tablica 1 Pregled tehnika za ublažavanje problema s gradijentima	20
Tablica 2 Primjena tehnika uravnoteženja za višeklasnu klasifikaciju MBTI [42]	25
Tablica 3 Pregled prednosti i nedostataka izabranih pristupa za uravnoteženje podataka.....	27
Tablica 4 Analiza strukturnih hiperparametara mreže LSTM pri višeklasnoj klasifikaciji MBTI [41]	29
Tablica 5 Rezultati binarne klasifikacije MBTI.....	49
Tablica 6 Rezultati višeklasne klasifikacije MBTI	51
Tablica 7 Kodiranje oznaka MBTI.....	57
Tablica 8 Distribucije klasa skupova podataka s Reddit i Personality Cafe platformi	70
Tablica 9 Distribucija dihotomija u skupu podataka MBTI.....	71
Tablica 10 Statistika skupova podataka	72
Tablica 11 Klasifikacijski izvještaj za LSMT-CE na skupu MBTI s Personality Cafe foruma	76
Tablica 12 Klasifikacijski izvještaj za LSMT-CECI na skupu MBTI s Personality Cafe foruma	80
Tablica 13 Usporedba rezultata pristupa LSTM-CE i LSMT-CECI na skupu MBTI s Personality Cafe foruma.....	81
Tablica 14 Klasifikacijski izvještaj za LSTM-CE na skupu s Reddit platforme	84
Tablica 15 Klasifikacijski izvještaj za LSTM-CECI na skupu s Reddit platforme.....	86
Tablica 16 Usporedba rezultata pristupa LSTM-CE i LSMT-CECI na skupu MBTI s Reddit platforme	87
Tablica 17 Klasifikacijski izvještaj za CNN-CE na skupu MBTI s Reddit platforme.....	90
Tablica 18 Klasifikacijski izvještaj za CNN-CECI na skupu s Reddit platforme.....	93
Tablica 19 Usporedba rezultata pristupa CNN-CE i CNN-CECI na skupu MBTI s Reddit platforme	94
Tablica 20 Klasifikacijski izvještaj za CNN-CE na skupu MBTI podataka s Personality Cafe foruma	97
Tablica 21 Klasifikacijski izvještaj za CNN-CE na skupu podataka s Personality Cafe foruma	100
Tablica 22 Usporedba rezultata pristupa CNN-CE i CNN-CECI na skupu MBTI s Personality Cafe foruma.....	101

Tablica 23 Pregled rezultata modelima LSTM i CNN.....	103
Tablica 24 Usporedba rezultata s drugim modelima - F1-mjera.....	105

12. ŽIVOTOPIS

Ninoslav Čerkez je rođen u Zenici, Bosna i Hercegovina, 1971. Ondje pohađa osnovnu školu i matematičku gimnaziju, tijekom kojih sudjeluje na republičkim natjecanjima iz fizike te saveznim natjecanjima iz matematike. Diplomira 1998. i magistrira 2005. na Sveučilištu u Zagrebu, Fakultet Elektrotehnike i računarstva te završava M.B.A. na IGBS, Indiana University 2011. Trenutno pohađa doktorski studij iz primijenjene računalne znanosti na Sveučilištu u Zagrebu, Fakultet Elektrotehnike i računarstva.

Od 1999 do 2018, zaposlen je u različitim tehničkim i voditeljskim ulogama u IT industriji, a osobito u području tehničke obuke novih zaposlenika te HR procesa. Također, radi kao honorarni Predavač na IT veleučilištima u Zagrebu. Od 2018, radi kao Predavač i Viši predavač na IT veleučilištu u Zagrebu. Dodatno, od 2018 do 2020, obnaša dužnost Prodekana na IT veleučilištu. Trenutno je zaposlen na zadacima koji uključuju primjenu umjetne inteligencije u automobilskoj industriji. Znanstveni interesi također uključuju primijenjenu znanost o podacima na područje psihologije osobnosti.

Član je stručnih udruga SHRM od 2014, HRCI od 2014 i PMI od 2007. Ostvario je PMP, SHRM-SCP i HRCI profesionalne certifikate.

Popis radova i aktivnih sudjelovanja na znanstvenim skupovima

Popis radova

1. Cerkez, N., Vrdoljak, B., Skansi, S., "A Method for MBTI Classification based on Impact of Class Components," IEEE Access; doi:10.1109/access.2021.3121137, vol. 9, pp. 146550-146567, 2021.
2. Čerkez, N.; Čorić, R., Đumić, M., Matijević, D., "Finding an optimal seating arrangement for employees traveling to an event", Croatian operational research review, 6 (2015), 2; 419-427 doi:10.17535/corr.2015.0032

Sudjelovanje na znanstvenim skupovima

1. Cerkez, N., Vrdoljak, B., Skansi, S., "The Present Situation and the Prospect of Determining the Personality Type of Text Author with Machine Learning," in Softcom, Split, 2019.

2. Čerkez, N., Vareškić, V., "Machine Learning Approaches to Personality Classification on Imbalanced MBTI Datasets," in 44th International Convention on Information, Communication and Electronic Technology (MIPRO), 2021, pp. 1259-1264, doi: 10.23919/MIPRO52101.2021.9596742, Opatija, 2021.
3. Frković, M., Čerkez, N., Vrdoljak, B., Skansi, S., "Evaluation of Structural Hyperparameters for Text Classification with LSTM Networks," in Mipro, Opatija, 2020.
4. Čerkez, N., Skočir, Z., "Holistic Approach to DBMS Tuning", 12th International Conference on Software, Telecommunications and Computer Networks SoftCOM 2004 and Workshop on Information and Communication Technologies & Posters: proceedings, SoftCOM 2004 Split: Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture, 2004. str. 46-50

13. BIOGRAPHY

Ninoslav Čerkez was born in Zenica, Bosna and Hercegovina, in 1971. He received the B.S. and M.S. degrees in engineering from the University of Zagreb, Croatia, in 1998 and 2005, respectively, and the M.B.A. degree from IGBS, Indiana University in 2011. He is currently pursuing the Ph.D. degree in computer science with the University of Zagreb.

From 1999 to 2018, he had different technical and HR roles in the IT industry. Also, he was working as a part-time Lecturer at IT colleges in Zagreb. Since 2018, he has been a Lecturer at the IT College in Zagreb. In addition, from 2018 to 2020, he was the Vice-Dean at the IT College. He works in applied AI in automotive industry. His research interest also includes applied data science to the psychology of personality.

Mr. Čerkez became a member of SHRM, in 2014; HRCI, in 2014; and PMI, in 2007. He holds PMP, SHRM-SCP, and HRCI professional certificates.