

# Prikupljanje i analiza podataka o interakciji čovjeka i robota

---

Lovrić, Irena

Master's thesis / Diplomski rad

2023

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

*Permanent link / Trajna poveznica:* <https://urn.nsk.hr/urn:nbn:hr:235:823801>

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

*Download date / Datum preuzimanja:* **2024-12-18**

*Repository / Repozitorij:*

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



SVEUČILIŠTE U ZAGREBU  
FAKULTET STROJARSTVA I BRODOGRADNJE

# **DIPLOMSKI RAD**

**Irena Lovrić**

Zagreb, 2023.

SVEUČILIŠTE U ZAGREBU  
FAKULTET STROJARSTVA I BRODOGRADNJE

# DIPLOMSKI RAD

Mentori:

Izv. prof. dr. sc. Tomislav Stipančić, dipl. ing.

Student:

Irena Lovrić

Zagreb, 2023.

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

Zahvaljujem se mentoru izv. prof. dr. sc. Tomislavu Stipančiću na pristupačnosti i pruženoj pomoći tijekom izrade rada i svim ostalima koji su sudjelovali i vodili istraživanje na kojem je baziran ovaj diplomski rad.

Irena Lovrić



SVEUČILIŠTE U ZAGREBU  
FAKULTET STROJARSTVA I BRODOGRADNJE



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

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

## DIPLOMSKI ZADATAK

Student: **Irena Lovrić** JMBAG: 0035213799

Naslov rada na hrvatskom jeziku: **Prikupljanje i analiza podataka o interakciji čovjeka i robota**

Naslov rada na engleskom jeziku: **Collection and analysis of data on human-robot interaction**

Opis zadatka:

PLEA je afektivni robot razvijen na Fakultetu strojarstva i brodogradnje koji se koristi za proučavanje interakcije ljudi i robota. Od tri razvijene osjetilne modalnosti PLEA trenutno koristi vizualnu da bi donijela zaključak o trenutnom emocionalnom stanju osobe u interakciji. PLEA se može shvatiti kao emocionalno ogledalo tako da tijekom interakcije koristi neverbalne komunikacijske znakove kako bi izrazila iste emocije kao osoba u interakciji.

U radu je potrebno definirati i provesti prikupljanje podataka kroz anketiranje studenata koji su u prethodnom koraku proveli u interakciji s robotom najmanje četiri minute. Potom je te podatke potrebno obraditi koristeći alate za analizu sentimenta temeljene na umjetnoj inteligenciji.

Sumarno, u okviru rada je potrebno:

- isplanirati tijek istraživanja,
- definirati pitanja za anketu tako da odgovori ispitanika odražavaju informacije o prihvaćenosti tehnologije,
- kroz metode oglašavanja prikupiti kandidate koji su voljni sudjelovati u istraživanju te provesti proces anketiranja,
- analizirati i normirati sakupljene podatke,
- primijeniti tehnike temeljene na umjetnoj inteligenciji za očitavanje sentimenta iz prikupljenih podataka,
- dati osvrt na rezultate istraživanja.

U istraživanju je potrebno voditi se etičkim načelima osiguravši pritom anonimnost ispitanika te poštivanje njihovih ljudskih prava.

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

Zadatak zadan:

Datum predaje rada:

Predviđeni datumi obrane:

4. svibnja 2023.

6. srpnja 2023.

17. – 21. srpnja 2023.

Zadatak zdao:

Predsjednik Povjerenstva:

Izv. prof. dr. sc. Tomislav Stipančić

Prof. dr. sc. Ivica Garašić

**SADRŽAJ**

|  |     |
|--|-----|
| POPIS SLIKA .....  | III |
| POPIS TABLICA.....   | V   |
| SAŽETAK.....   | VI  |
| SUMMARY .....  | VII |
| 1. UVOD.....   | 1   |
| 2. PLEA .....  | 2   |
| 3. OGLAŠAVANJE .....   | 5   |
| 4. PRIKUPLJANJE PODATAKA .....   | 7   |
| 4.1. PLEA proces prikupljanja podataka .....   | 8   |
| 4.2. Zaštita podataka sudionika.....   | 9   |
| 5. UPITNIK .....   | 10  |
| 5.1. PLEA upitnik .....  | 11  |
| 6. PODACI .....  | 13  |
| 7. DESTILACIJA ZNANJA ( eng. <i>KNOWLEDGE DISTILLATION</i> ).....                            | 15  |
| 7.1. Znanje (eng. <i>Knowledge</i> ) .....   | 16  |
| 7.2. Destilacijski sheme (eng. <i>Distillation Schemes</i> ).....                            | 18  |
| 7.3. Arhitektura učitelj-student (eng. <i>Teacher-Student Architecture</i> ).....            | 19  |
| 7.4. Destilacijski algoritmi (eng. <i>Distillation Algorithms</i> ).....                     | 20  |
| 7.5. Primjene destilacije znanja (eng. <i>Applications of knowledge distillation</i> ) ..... | 22  |
| 7.5.1. Vizijsko prepoznavanje.....   | 22  |
| 7.5.2. Prepoznavanje govora .....  | 22  |
| 7.5.3. Natural Language Processing (NLP) .....   | 23  |
| 8. DistilBERT .....  | 24  |
| 8.1. Kod za sentiment analizu .....  | 25  |
| 8.2. Transformers biblioteka i transformeri .....  | 26  |

|  |    |
|--|----|
| 8.3. Model i pipeline .....                            | 28 |
| 8.4. Izlazi dobiveni pomoću DistilBERT-a.....          | 31 |
| 9. STATISIKA.....                                      | 34 |
| 9.1. Izlazi koda za prebrojavanje riječi/emocija ..... | 35 |
| 10. ZAKLJUČAK.....                                     | 38 |
| LITERATURA.....  | 39 |
| PRILOZI.....   | 42 |

**POPIS SLIKA**

|   |    |
|---|----|
| Slika 1. Razvoj PLEA robot glave [1].....   | 3  |
| Slika 2. Sustav projekcije svjetla [1].....   | 4  |
| Slika 3. Slika oglašena na društvenim mrežama .....   | 5  |
| Slika 4. Raspodjela dolaska sudionika u određenim terminima.....                            | 6  |
| Slika 5. Odgovori sudionika pretvoren u tekst.....  | 13 |
| Slika 6. Prikaz početka CSV datoteke za pitanje 5 .....                                     | 14 |
| Slika 7. Shema destilacije znanja između učitelj i student modela [10] .....                | 15 |
| Slika 8. Shematski prikaz izvora znanja [10].....   | 17 |
| Slika 9. Primjer različitih destilacija [10].....   | 19 |
| Slika 10. Veze između učitelj i student modela u procesu destilacije znanja [10] .....      | 20 |
| Slika 11. Multi-Teacher Distillation [10].....  | 21 |
| Slika 12. Cross-Model Distillation [10].....  | 21 |
| Slika 13. Broj parametra istreniranih modela za jezik [16].....                             | 24 |
| Slika 14. Sentiment analiza pomoću DistilBERT modela #1 .....                               | 25 |
| Slika 15. Sentiment analiza pomoću DistilBERT modela #2 .....                               | 25 |
| Slika 16. Pojednostavljena arhitektura transformera [16].....                               | 26 |
| Slika 17. Arhitektura transformera [27] .....   | 27 |
| Slika 18. Broj članaka koji su koristili BookCorpus od 2019.godine do 2023. godine [40] ... | 29 |
| Slika 19. Pie grafikon za pitanje 3.....  | 31 |
| Slika 20. Pie grafikon za pitanje 4.....  | 32 |
| Slika 21. Pie grafikon za pitanje 5.....  | 33 |



|   |    |
|---|----|
| Slika 22. Kod za prebrojavanje riječi/odgovora #1 ..... | 34 |
| Slika 23. Kod za prebrojavanje riječi/odgovora #2 ..... | 34 |
| Slika 24. Bar grafikon za pitanje 3 .....               | 35 |
| Slika 25. Bar grafikon za pitanje 4 .....               | 36 |
| Slika 26. Bar grafikon za pitanje 5 .....               | 37 |

## POPIS TABLICA

|   |    |
|---|----|
| Tablica 1 Tablica emocija i PLEA-inih izraza .....              | 2  |
| Tablica 2 Hiperparametri modela [31] .....                      | 28 |
| Tablica 3 Statistički sažetak podatka o BookCorpus-u [40] ..... | 29 |
| Tablica 4 Usporedba modela [31] .....                           | 30 |

## SAŽETAK

U ovom radu je objašnjen tijek istraživanja vezan za afektivni PLEA robot. Nastoji se razumjeti potencijalna primjena PLEA robota za dobrobit i socijalnu skrb. Objašnjeno je što je PLEA robot, kako su se prikupili podaci, kako se definiraju pitanja i upitnik, način oglašavanja, kako se obrađuju podaci i rezultat analize podataka.

PLEA je afektivni robot koji se koristiti za proučavanje interakcije ljudi i robota. Komunikacija između ljudi i PLEA robota je neverbalna. Trenutno se smatra kao emocionalno zrcalo. Razvijen na Fakultetu strojarstva i brodogradnje.

Prvi korak je bio prikupljanje sudionika preko društvenih mreža i službenih stranica fakulteta. Zatim slijedi prikupljanje podataka koje se sastoji od dva dijela: interakcije sudionika s PLEA robotom i naknadnog intervjua koji se bazirao na upitniku od 12 pitanja. Intervjui su spremljeni u obliku audio zapisa i kasnije pretvoreni u tekstualni format u obliku CSV datoteke. Podaci dobiveni u drugom dijelu prikupljanja podataka na pitanja 3, 4 i 5 su obrađeni u ovom radu.

Rezultati istraživanja su dobiveni pomoću sentiment analize koristeći model baziran na DistilBERT-u i pomoću klasične statistike.

Ključne riječi: Prikupljanje podataka, DistilBERT, sentiment analiza, upitnik

## **SUMMARY**

This paper explains the course of research related to affective PLEA robot. The aim is to understand the potential application of PLEA robot for wellbeing and social care. It explains what the PLEA robot is, how the data was collected, how the question and questionnaire were defined, the advertising method, data processing and the results of data analysis.

PLEA is an affective robot used for studying human-robot interaction. Communication between humans and PLEA robot is non-verbal. It is currently considered an emotional mirror. It was developed at the Faculty of Mechanical Engineering and Naval Architecture.

The first step was gathering participants through social media and the official faculty webpage. Followed by data collection, which consisted of two parts: participants' interaction with PLEA robot and subsequent interview based on 12 question questionnaire. The interviews were recorded as audio files and later converted into a textual format as CSV file. The data processed in the research was obtained in the second part of data collection from questions 3, 4 and 5.

The research results were obtained using sentiment analysis with a model based on DistilBERT and classical statistics.

**Key words:** Data collection, DistilBERT, sentiment analysis, questionnaire

## 1. UVOD

U ovom radu je analiziran skup podataka koji predstavlja interakciju između čovjeka i PLEA robota kako bi dobili informacije o upotrebljivosti, prihvatljivosti, pouzdanosti i privlačnosti PLEA robota. Cilj je stvoriti temelj za daljnju analizu obrasca interakcije.

Podaci se prikupljaju koristeći PLEA robot koji je osjetljiv na emocije i čija je sposobnost reagiranja na ljudske emocije. PLEA je virtualno biće koje se nalazi u cyber prostoru i kao virtualno biće može biti bilo gdje u svijetu gdje postoji internet. PLEA radi u kaotičnoj i nedefiniranoj okolini ili drugim riječima u prirodnoj okolini, gdje je jedina konstanta kontinuirana promjena. U budućnosti će PLEA biti u mogućnosti previđati odgovore za vrijeme interakcije sa ljudima i to će joj čak omogućiti da kontrolira neverbalni razgovor. PLEA se trenutno smatra emocionalnim zrcalom.

Istražuju se sposobnost PLEA robota da pametno okruženje učini ugodnijim i prijateljskim za korisnike. Savršeno sučelje bi trebalo popraviti interakciju između korisnika i PLEA robota. U istraživanju na kojem je rađen ovaj rad fokusira se na populaciju starijih osoba u domovima za njegu i vlastitim domovima i kako bi ulogu PLEA robot imao u takvom okruženju. Pametne tehnologije su podijeljene u tri glavne kategorije:

- I. One koje pomažu u održavanju zdravlja
- II. One koje pomažu kod održavanja sigurnosti
- III. One koje pomažu u osiguravanju sreće







Treća kategorija je kategorija u kojoj će PLEA robot najviše doprinijeti. Te je glavni cilj ove analize otkriti načine i u kojoj mjeri PLEA može pomoći pri stvaranju pametnog okruženja koji će biti ugodno i gostoljubivo.

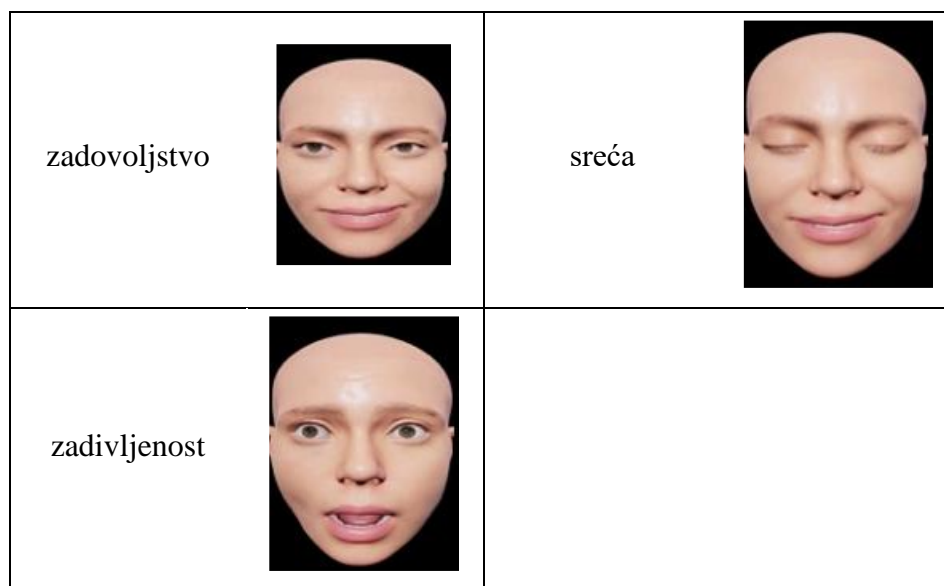
U ovom radu će se analizirati informacije dobivene na pitanja vezana za sposobnosti PLEA robota kod izvođenja svoje uloge tj. razmjena emocija sa sudionicima.

## 2. PLEA

PLEA je afektivno virtualno biće koje se nalazi u fizičkom serveru i kibernetičkom prostoru. S obzirom da je server povezan s internetom, PLEA je dostupna svugdje u svijetu putem interakcijskog sučelja kao što su AR i VR naočale, ekran mobitela, računalo, robot i sl. PLEA uspostavlja dvosmjernu komunikaciju s posjetiteljima na neverbalnoj razini temeljenu na emocijama koje posjetitelj pokazuje. Za sad PLEA zrcali emocije koje zamjećuje na licu posjetitelja. U sljedećoj tablici [Tablica 1] su prikazane emocije i kako ih PLEA izražava.

**Tablica 1 Tablica emocija i PLEA-inih izraza**

|                  |   |             |   |
|------------------|---|-------------|---|
| ljutnja          |   | dosada      |   |
| cool             |  | razočaranje |  |
| entuzijastičnost |  | srdačnost   |  |



Fizički produžetak virtualnog bića PLEA robota je robotska glava i sučelje. Robotska glava je dizajnirana da bude realistična i izražajna. Na sljedećoj slici [ Slika 1 ] prikazan je razvoj glave robota u 4 faze i krajnji rezultat.

Faze razvoja su:

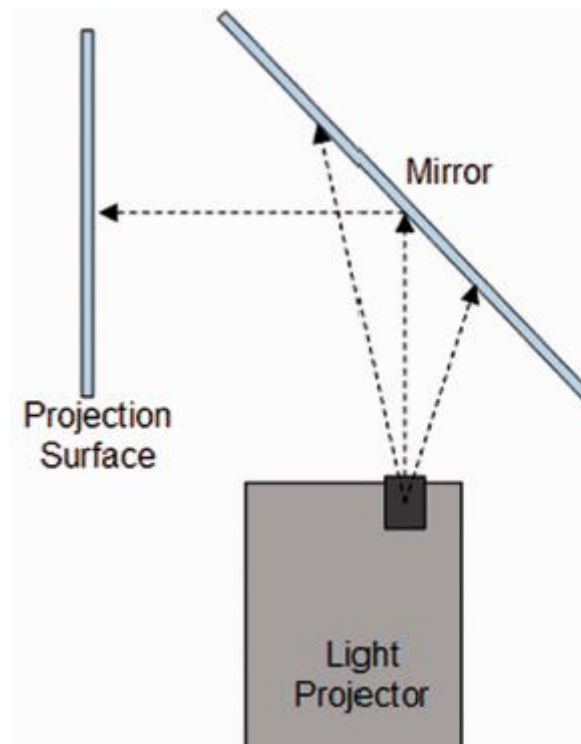
- 1) Dizajn glave
- 2) Faza u kojoj se koristi ravna projekcijska površina
- 3) Faza u kojoj se koristi 3D projekcijska površina
- 4) 3D CAD model glave



Slika 1. Razvoj PLEA robot glave [1]

Projekcija na ravnu površinu je imala problem kod vizualizacije jer pati od fenomena Mona Lise u kojem se čini da oči portreta prate promatrača dok prolazi. Realizacijom 3D glave taj se problem riješio.

Sustav projekcije svjetla je ugrađen u glavu robota i njegova shema je prikazana na sljedećoj slici [Slika 2].



**Slika 2. Sustav projekcije svjetla [1]**

Svjetlosni projektor baca svjetlo na ogledalo koje je postavljeno pod kutom od  $45^\circ$ , zatim se svjetlo reflektira na površinu lica. Dizajn glave nije potpuno originalan. Dizajn je usporediv sa robotom FURHAT koji je razvijen u KTH, Stockholm.



### 3. OGLAŠAVANJE

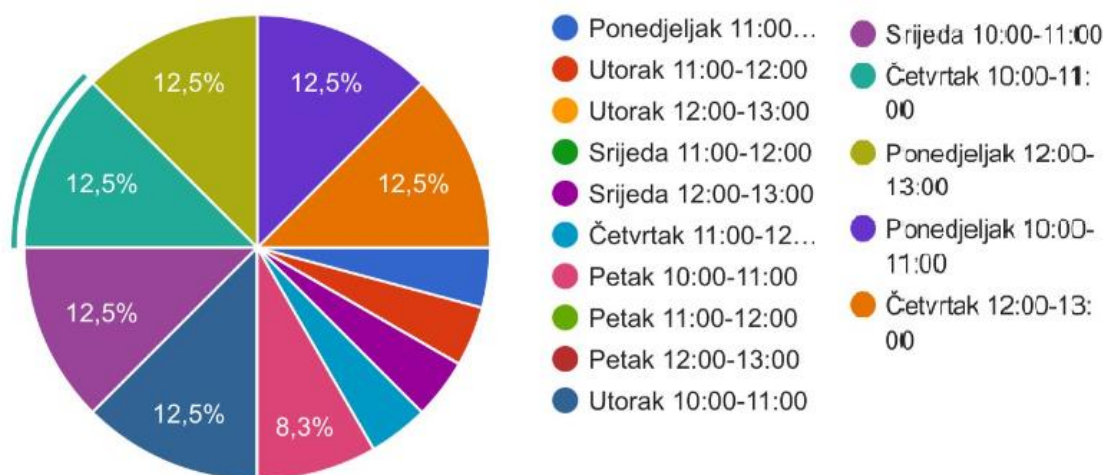
Za prikupljanje podataka sudionici su saznali preko obavijesti na stranicama Fakulteta strojarstva i brodogradnje ili preko društvenih mreža. Upravo je to razlog zašto su sudionici bili isključivo studenti. Sljedeća slika [Slika 3] je objavljena na Instagram i Facebook profilima istraživača koji su vodili proces prikupljanja podataka. Objava je također bila podijeljena na jednom od Instagram profila Fakulteta strojarstva i brodogradnje (@fsb\_unizg).



**Slika 3. Slika oglašena na društvenim mrežama**

Na objavi je postavljen link za Google obrazac preko kojeg su sudionici mogli rezervirati vrijeme dolaska. Na sljedećoj slici [Slika 4] se može vidjeti raspodjela dolaska sudionika u ponuđenim terminima:

24 odgovora



**Slika 4. Raspodjela dolaska sudionika u određenim terminima**

Na ulazu u prostoriju gdje su se prikupljali podaci je bila postavljena prijašnja slika [Slika 3] i bilo tko je mogao doći i sudjelovati u prikupljanju podataka bez popunjavanja Google obrasca ako je bilo slobodan termin u to vrijeme.

## 4. PRIKUPLJANJE PODATAKA

Prikupljanje podataka je proces prikupljanja i analiziranja podataka iz različitih relevantnih izvora kako bi se našli odgovori na istraživačke probleme, trendove, vjerojatnosti itd. kako bi se evaluirali mogući ishodi [4]. Za vrijeme prikupljanja podataka moraju se identificirati vrste podataka, izvori podataka i metode koje će se koristiti. Podaci su informacije koje su formatirane na poseban način.

Postoje dva pristupa za prikupljanje podataka a to su prikupljanje primarnih i sekundarnih podataka.

Prikupljanje primarnih podataka iz glavnog/primarnog izvora. Odnosi se na podatke koji nisu nikad prije bili korišteni u prošlosti i smatraju se najboljom vrstom podataka.

Metode za prikupljanje primarnih podataka su:

1. Intervju – izravna metoda prikupljanja podataka u kojoj ispitivač postavlja pitanja a sudionik odgovara na njih.
2. Promatranja – Istraživači promatraju situaciju oko sebe i bilježe ono što vide. Jednostavna i učinkovita metoda koja izravno ne ovisi o drugi sudionicima.
3. Ankete – mogu se provoditi licem u lice, slati poštom, objaviti na internetu. Pitanja u anketi mogu biti otvorenog i zatvorenog tipa.
4. Fokusne grupe – Metoda slična intervjuu ali se provodi sa grupom ljudi koji imaju nešto zajedničko.
5. Usmena povijesti – metoda koja uključuje postavljanje pitanja kao što su intervju i fokusna grupa, ali je ova metoda preciznije definirana. Prikupljeni podaci su povezani sa jednom pojavom. Prikupljaju se mišljenja i osobna iskustva ljudi za vrijeme određenog događaja u kojem su sudjelovali.

Prikupljanje sekundarnih podataka podrazumijeva korištenje podataka koje je netko drugi prije prikupio za neku drugu svrhu. Sekundarne podaci se mogu prikupiti iz sljedećih izvora:

1. Javni izori – knjige, akademski časopisi, novine, državna izvješća itd.

2. Online baze podataka – postoje brojne online baze podataka koje omogućuju pristup širokom rasponu sekundarnih podataka od akademskih članaka, statističkih informacija itd.
3. Državna izvješća – Vladine agencije, istraživačke institucije i organizacije održavaju baze podataka i zapise koji se mogu koristiti u istraživačke svrhe
4. Javno dostupni podaci – Podaci koji su podijeljeni od strane pojedinaca, organizacija na javnim platformama i mogu se koristiti za istraživanje
5. Prošla istraživanja – Istraživači pregledavaju i analiziraju podatke kako bi stekli uvid ili nadogradili već postojeće znanje

#### **4.1. PLEA proces prikupljanja podataka**

Prikupljanje podataka u ovom istraživanju se odvijao pomoću intervjua i kasnije analize snimki interakcija između sudionika i PLEA robota. Prikupljanje podataka se odvijalo u prostorijama Fakulteta strojarstva i brodogradnje u periodu od 5 radnih dana.

Prvi dio prikupljanja podataka je interakcija sudionika sa PLEA robotom. Sudionici su proveli 4 minute ispred PLEA robota kako bi se uspostavilo međusobno razumijevanje. Vrijeme je bilo isto za sve sudionike. Dok su bili ispred PLEA robota, sudionici su iskazivali svoje osjećaje pomoću raznih izraza lica. Komunikacija je bila isključivo neverbalna. Interakcija između sudionika i PLEA robota je bila snimana. Snimke omogućuju kasnije promatranje i analizu interakcije sudionika i PLEA robota, što daje uvid u trenutne tehničke mogućnosti PLEA robota što će omogućiti daljnja unaprijeđenja.

Drugi dio prikupljanja podataka je bio intervju koji se sastojao od unaprijed definiranog upitnika od 12 pitanja. Prikupljeni podaci su primarnog tipa. U ovom radu će se pažnja obratiti na podatke prikupljene pomoću intervjua, pa je u poglavlju 5 opisano šta je to upitnik i kako se formiraju pitanja. U poglavlju 5.1 su zapisana sva pitanja iz upitnika koji je korišten u intervju i koje informacije dobivamo od njih.

## 4.2. Zaštita podataka sudionika

Podaci koji su dani od strane sudionika su povjerljivi i samo uz dopuštenje sudionika se mogu koristiti anonimizirani citati. Anonimizirani citati su izmijenjeni citati ili su iz citata uklonjeni podaci koji mogu otkriti identitet sudionika. Informacije koje daju sudionici se smatraju kao izvor pozadinskih informacija uz istraživanje temeljeno na literaturi i intervjuima sa drugom sudionicama. Dani podaci se neće koristiti u dodatne svrhe osim onih koje se naglašene sudionicima i neće se bilježiti više od onog što je potrebno za istraživanje.

Svakom sudioniku je dan informacijski list, prije početka prikupljanja podataka, u kojem su zapisani osnovni podaci o istraživanju kao što su:

- Cilj prikupljanja podataka i istraživanja
- Kratko objašnjenje procesa prikupljanja podataka i kako će se podaci spremiti
- Objašnjenje kako je moguće povući svoje podatke iz istraživanja
- Kratko objašnjenje o čuvanju povjerljivosti podataka i tko će imat pristup podacima

Niti u jednoj publikaciji istraživačkog rada neće biti objavljeno ime sudionika kao i ni ostali osobni podaci.

## 5. UPITNIK

Upitnik je istraživačko sredstvo za prikupljanje podataka i sastoji se od seta pitanja. Podaci koji se prikupe pomoću upitnika mogu biti kvantitativni i kvalitativni. Za prikupljanje kvantitativnih podataka koriste se strukturirani upitnici. Strukturirani upitnici su isplanirani i formirani za prikupljanje preciznih informacija. Za prikupljanje kvalitativnih podataka koriste se nestrukturirani upitnici. Nestrukturirani upitnici koriste osnovnu strukturu i razgranata pitanja ali to ne limitira odgovore na pitanja. Pitanja u upitniku mogu biti različitih tipova. U jednom upitniku se mogu koristiti više tipova što zadržava sudionike angažiranima. Tipovi pitanja su :

- a) Pitanja otvorenog tipa – sudionik na pitanja odgovara bez restrikcija u slobodnoj formi. Ovaj tip pitanja se koristi kod prikupljanja kvalitativnih podataka.
- b) Pitanja zatvorenog tipa- sudionik na pitanja odgovara odabirom na ponuđene odgovore.
  - a. Dihotomna pitanja- pitanja koja imaju samo dva ponuđena odgovora DA ili NE
  - b. Pitanja s ponuđenim višestrukim odgovorima- pitanja koja imaju više ponuđenih odgovora, sa opcijama da se može odabrati samo jedan odgovor ili više njih
  - c. Pitanja sa skalom– pitanja na koja se odgovara pomoću mjernih skala.
  - d. Slikovna pitanja– slična pitanjima sa višestrukim ponuđenim odgovorima. Razlika je u tome što su ponuđeni odgovori u formatu slike pa je odgovaranje na pitanje brže i sudionici manje propituju svoje odgovore što daje točnije podatke.

Kako bi pitanja bila dobro napisana potrebno je pratiti određena pravila:

1. Koristiti jednostavan jezik
2. Izbjegavati žargon
3. Preciznost kada se pitanja odnose na čimbenike poput učestalosti
4. Izbjegavati negative u pitanjima
5. Izbjegavati dvostruka pitanja
6. Izbjegavanje pitanja koja mogu navoditi sudionika na odgovor
7. Provjeriti je li formulacija pitanja prirodna (čitanje pitanja na glas) i može li se poboljšati
8. Provjeriti jesu li pitanja jasna sa osobama koje nisu uključene u projekt.

Kreiranje upitnika je proces u više faza gdje se treba obratiti pažnja na više detalja od jednom. Za vrijeme kreiranja upitnika treba se voditi sljedećim pitanjima:

- a) Što trebamo saznati iz upitnika? – Odlučujemo koja pitanja će biti u upitniku, a koja ne prema istraživačkim ciljevima.
- b) Koliko dug upitnik treba biti? – Postaviti samo pitanja na koja se dobivaju točno oni podaci koji su potrebni. Predug upitnik može odbiti ljude od sudjelovanja.
- c) Tko odgovara na pitanja? – Potrebno je prilagoditi pitanja tako da svaki sudionik posve jasno razumije pitanje bez obzira na njihovo znanje, pismenost i jezične vještine
- d) Koji će biti redoslijed pitanja? – Nakon što su pitanja osmišljena, potrebno je paziti na redoslijed kojim će se pitanja postavljati. Pitanja na početku upitnika ne smiju utjecati na pitanja koja dolaze poslije njih.
- e) Što trebamo saznati o sudionicima? – Kod analize dobivenih podataka nekad je potrebno znat neke osobne podatke o pojedincu kao što su rasa, rod, dob itd. kako bi se moglo utvrditi kako ti osobni podaci utječu na njihova razmišljanja.

### **5.1. PLEA upitnik**

Upitnik korišten u ovom istraživanju sadrži 12 pitanja otvorenog tipa. Sudionici daju odgovore na pitanja usmeno i odgovori na pitanja se spremaju u obliku audio zapisa. Dobivaju se informacije o izgledu PLEA robota, rasponu emocija koje su bile izmijenjene, međusobnom razumijevanju i mogućim ulogama PLEA-e u stvarnom životu.

Pitanja pod rednim brojem 1 i 2 su se odnosila na izgled, a pitanja 3, 4 i 5 su se odnosila na mogućnost interakcije. Iz odgovora na prva dva pitanja se želi saznati koje značajke robota su se sudionicima najviše svidjele, a koje najmanje i kako bi se pojedine značajke mogle poboljšati. Prema odgovorima na pitanja od 3 do 5 saznajemo koliko dobro PLEA izvodi svoju ulogu (komunikacija emocija). Saznajemo raspon prikazanih PLEA-nih emocija i koje emocije najbolje komunicira.

Pitanja pod rednim brojem 6, 7 i 8 se odnose na ulogu koju bi PLEA mogla zauzeti u stvarnom životu. Sudionici prema svom osobnom iskustvu ili iskustvu ljudi u svom životu mogu

prepoznati neke nove uloge i iskazati svoje mišljenje o prihvatljivosti PLEA robota u različitim generacijama.

Pitanja pod rednim brojem 9, 10, 11 i 12 se odnose više na subjektivnu procjenu sudionika o PLEA-i. Saznajemo kako sudionici gledaju na PLEA-u. Dobivamo informacije o pouzdanosti, privlačnosti (izgled, želji za korištenjem isl.), rodu robota i koliko je rod bitan.

Pitanja u upitniku su:

1. Koje značajke PLEA robota Vam se najviše sviđaju?
2. Koje značajke PLEA robota Vam se najmanje sviđaju?
3. Koji raspon emocija je PLEA komunicirala s Vama?
4. Za koje osjećaje ste osjetili da PLEA najbolje komunicira s vama?
5. Koje osjećaje ste Vi dijelili s PLEA robotom?
6. Koju ulogu bi po Vama PLEA robot mogao imati kao robot koji njeguje ljude?
7. Kome bi (u ovisnosti o kontekstu uporabe) PLEA robot bio najprikladniji kao uslužni robot u okruženju skrbi?
8. Da imate člana obitelji u ustanovi skrbi, biste li rado upotrijebili PLEA robota – zašto ili zašto ne?
9. U Vašoj današnjoj interakciji s PLEA robotom, smatrate li da je PLEA robot pouzdan? (Što za Vas znači pouzdanost?)
10. U Vašoj današnjoj interakciji s PLEA robotom, smatrate li da je PLEA robot privlačan? (Što za Vas znači privlačnost?)
11. Kojeg je spola PLEA robot i kako to znate? Kako Vam je PLEA priopćila spol?
12. Je li spol PLEA robota važan za Vas kako biste bili sigurni u njegovu korisnost kao uslužnog robota u okruženju skrbi? Ukoliko ne, zašto ne?



## 6. PODACI

Kao što je spomenuto u poglavlju 5.1 odgovori na pitanja se spremila u obliku audio zapisa. Kasnije su odgovori pretvoreni u tekst kako bi se mogli koristiti za daljnju analizu. Na sljedećoj slici [Slika 5] su prikazani kako su odgovori jednog ispitanika zapisani:

| Positive features(1.)              | Negative features(2.)          | Emotions(3.)                     | Best comm emotions(4) | Shared emotions(5) | Roles(6.)                                  |
|------------------------------------|--------------------------------|----------------------------------|-----------------------|--------------------|--|
| mimicking facial expressions       | slowly adapts to given emotion | happiness                        | happiness             | happiness          | therapy robot                              |
|                                    |                                | disgust                          |                       | disgust            | company                                    |
|                                    |                                |                                  |                       | sadness            |  |
|                                    |                                |                                  |                       | confusion          |  |
| Appropriate as a service robot(7.) | Use in care setting(8.)        | Trustworthy(9.)                  | Attractive(10.)       | Gender(11.)        | Gender important(12.)                      |
| elderly                            | NO                             | NO                               | NO                    | F                  | NO   |
| children                           |                                | needs to speed up mimicry        | skin has no color     | facial features    | gender doesnt define emotinal intelligency |
| children with speacial needs       |                                | facial features seem undone      | unhumanly             |                    |  |
|                                    |                                | didnt recognize all the emotions |                       |                    |  |

**Slika 5. Odgovori sudionika pretvoren u tekst**

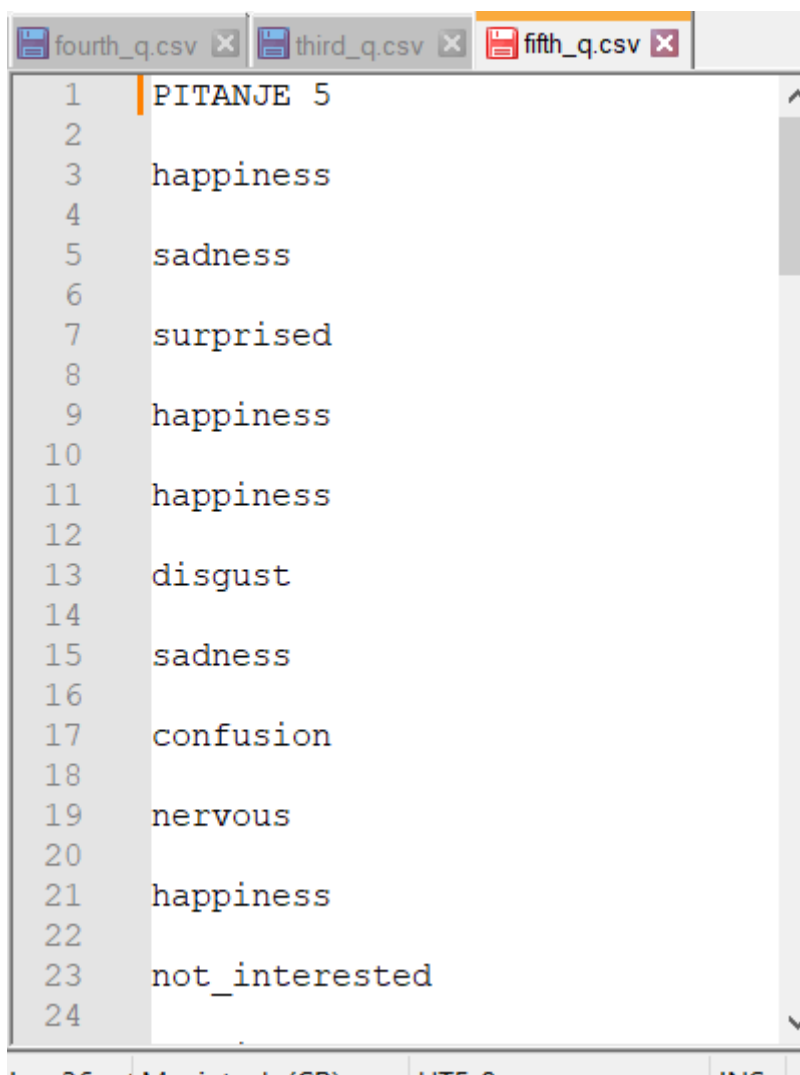
Analiza se sastoji od dva dijela:

- Sentiment analiza pomoću DistilBERT metode koja će biti objašnjena u daljnjim poglavljima
- Statistička analiza

Analize će se obavljati na pitanjima 3, 4 i 5.

Za obavljane analize za svako pitanje smo napravili CSV datoteku kako bi je kod mogao čitati. Kod CSV (Comma-Separated values) datoteka podaci su zapisani u tekst formatu i kao što ime ukazuje podaci odvojeni su zarezom. CSV podaci se ne moraju odvajati zarezom kao separator se mogu koristiti i točka-zarez (;), razmak ili neki drugi znak. Najčešće se koriste kod baza podataka i proračunskih tablica. CSV datoteku mogu otvoriti mnogi programi što omogućuje razmjenu podataka koje inače nisu moguće.

U CSV datotekama korištenim za analizu podataka u ovom radu za separator se koristio novi red. Na sljedećoj slici [Slika 6] će biti prikazano prvih nekoliko redova CSV datoteke za peto pitanje.

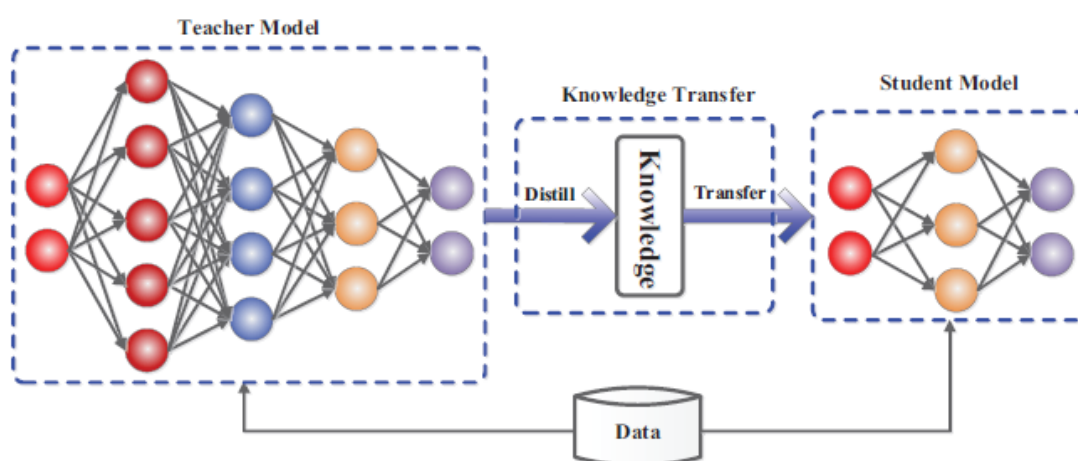


Slika 6. Prikaz početka CSV datoteke za pitanje 5

## 7. DESTILACIJA ZNANJA ( eng. *KNOWLEDGE DISTILLATION*)

Duboke neuronske mreže se dosta koriste i u industriji i u znanstvenim istraživanjima. Duboko učenje ima velik uspjeh radi mogućnosti rukovanja s velikom količinom podataka i milijardama parametara. Kako bi to bilo moguće koriste se *graphic processing units (GPU)* koji izvode operacije istodobno, a ne jednu za drugom time se skraćuje vrijeme treniranja. Nedostatak GPU je to što se ne mogu koristiti na manje jakim hardverima kao što su mobiteli, računala isl. Zbog toga što se ne mogu koristiti računala i mobiteli za izvođenje velikih modela, zbog kompleksnosti modela i zbog velike potrebe za memorijom potrebno je provesti kompresijsku metodu kao što je destilacija znanja ( eng. *Knowledge distillation*).

Destilacija znanja je proces kompresije modela u kojem se manji *student model* trenira kako bi odgovarao velikom istreniranom *učitelj modelu* (eng. *teacher model*). Znanje se prenosi sa učitelj modela na student model tako da se minimizira funkcija gubitka (eng. *loss function*) i cilj je usklađivanje softened logita i ground-truth labela. Logiti su omekšani primjenom funkcije skaliranja „temperature“ u softmax-u, učinkovito izgladujuću distribuciju vjerojatnosti i otkrivajući međurazredene odnose koje je naučio učitelj model[13]. Za vrijeme procesa destilacije znanja bitno je da se zadrži ista valjanost u student modelu kao što je bila u učitelj modelu. Učitelj model nadgleda student model u procesu destilacije znanja. Na sljedećoj slici [Slika 7] prikazana je shema destilacije znanja.



Slika 7. Shema destilacije znanja između učitelj i student modela [10]

Destilacija znanja se može podijeliti u tri komponente koje se mogu vidjeti na prethodnoj slici [Slika 7]:

- 1) Znanje (eng. *Knowledge*)
- 2) Destilacijski algoritam (eng. *Distillation Algorithm*)
- 3) Arhitektura učitelj-student (eng. *Teacher-Student Architecture*)

Koliko brzo student model uči i što uči ukazuje na faktore koji su utječu na uspješnost destilacije učitelj modela. Uspješna destilacija se zasniva na geometriji podataka, optimizacijska pristranost destilacijskog cilja i stroga monotonost studentskog klasifikatora.

Empirijski dokazi pokazuju da veliki modeli možda neće bit dobar učitelj zbog razlike u kapacitetu, a eksperimenti pokazuju da destilacija negativno utječe na učenje student modela.

Student model se može lako dodati u aplikacije za vizijsko prepoznavanje, prepoznavanje govora i *natural language processing (NLP)*.

### **7.1. Znanje (eng. *Knowledge*)**

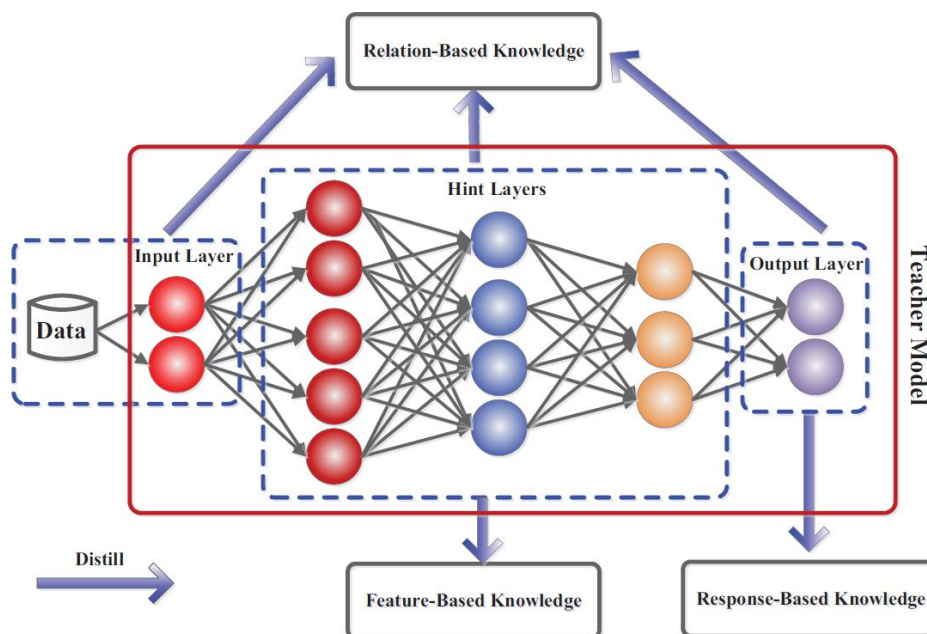
Naučene težine i *bias* se uobičajeno podrazumijeva kao znanje u neuronskim mrežama. Destilacija znanja uobičajeno koristi logite kao izvore znanja za učitelj model, uz logite za znanje učitelj modela se mogu koristiti aktivacijske funkcije, neuroni ili značajke skrivenih slojeva. Odnosi između aktivacijskih funkcija i između neurona kao i parametri učitelj modela imaju važne informacije i znanje za učitelj model.

Postoje 3 kategorije znanja:

- a) Znanje temeljeno na odgovorima (eng. *Response-Based Knowledge*) – fokusira se na izlaze zadnjeg sloja učitelj modela. Glavna ideja je da će student model oponašati finalna predviđanja učitelj modela. destilacija znanja bazirana na odgovorima je vrlo jednostavna i efektivan metoda za kompresiju modela. Nedostatak je to što znanje temeljeno na odgovorima ne uzima u obzir skrivene slojeve koji su jako važni za reprezentacijsko učenje. Ova kategorija znanja se koristi u kontekstu nadgledanog učenja.

- b) Znanje temeljeno na značajkama (eng. *Feature-Based Knowledge*) – istrenirani učitelj model prikuplja znanje iz skrivenih slojeva koje je dosta relevantno za duboke neuronske mreže. Skriveni slojevi uče razlikovati specifične značajke i to znanje će se koristiti pri učenju student modela. Cilj destilacije znanja baziranoj na značajkama je da se student model nauči iste aktivacije funkcija kao i učitelj model. Destilacijska funkcija gubitka minimizira razlike između aktivacijskih funkcija učitelj i student modela.
- c) Znanje temeljeno na odnosima (eng. *Relation-Based Knowledge*) – za razliku od znanja temeljenog na odgovorima i znanja temeljenog na značajkama, znanje temeljeno na odnosima ne prikuplja samo informacije iz skrivenih i izlaznih slojeva, nego i odnose između mapa značajki koje se također koriste za treniranje student modela. Istraživanje odnosa između mapa značajki se vrši pomoću *flow of solution process (FSP)* koristeći Gramovu matricu između dva sloja. Odnos se može opisati kao veza između dvije mape značajki, grafovi, matrice sličnosti, ugrađenih značajki ili distribucije vjerojatnosti temeljene na prikazanim značajkama.

Na sljedećoj slici [Slika 8] se može vidjeti shematski prikaz izvora znanja.



Slika 8. Shematski prikaz izvora znanja [10]

## 7.2. Destilacijski sheme (eng. Distillation Schemes)

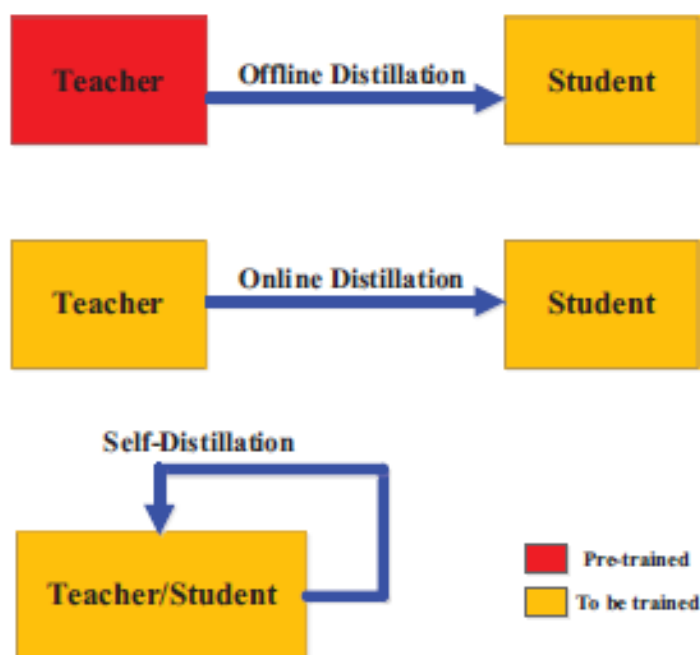
Razlikuju se 3 destilacijske sheme ovisno o tome je li se učitelj model ažurirao istodobno sa student modelom ili ne.

Tri destilacijske sheme su:

- a) Offline destilacija (eng. *Offline distillation*) – najčešća metoda i dosta je jednostavna za implementaciju. Odvija se u dvije faze i prijenos znanja je jednosmjernan (od učitelj modela do student modela). U prvoj fazi se trenira učitelj model na setu primjera za trening prije destilacije, u drugoj fazi znanje iz učitelj modela je destilirano za treniranje student modela. Prva faza u offline destilaciji se često ne smatra dijelom destilacije znanja pa se pridaje malo pažnje učitelj modelu i njegovu vezu sa student modelom, te se fokusira na poboljšavanju različitih dijelova transferu znanja.
- b) Online destilacija (eng. *Online distillation*) – se odvija u jednoj fazi i operacije se provode paralelno te je ovo jako efikasna metoda. Nastala je kako bi riješili ograničenja offline destilacije, a to je da istrenirani učitelj model nije uvijek dostupan. Za vrijeme online destilacija odvija se ažuriranje oba, učitelj i student, modela.
- c) Samostalna destilacija (eng. *Self distillation*) – može se gledati kao specijalni slučaj online destilacije. Koriste se iste mreže i u student i u učitelj modelu. Znanje iz dubljih slojeva neuronske mreže se koristi za treniranje plićih slojeva mreže.

Sve tri destilacijske sheme se mogu intuitivno usporediti sa načinima kako ljudi uče. Offline destilacija se uspoređuje kada učitelj prenosi znanje na učenika, online destilacija je kad učitelj i učenik zajedno uče, a samostalna destilacija kada učenik uči sam. Kao što kod ljudi se te vrste učenja mogu kombinirati tako se mogu kombinirati i navedeni destilacijske sheme. Međusobno se nadopunjuju zbog svojih prednosti.

Na sljedećoj slici [Slika 9] su prikazane destilacijske sheme. Crvena boja označava da je model unaprijed istreniran, a žuta boja da se model trenira tijekom procesa destilacije.



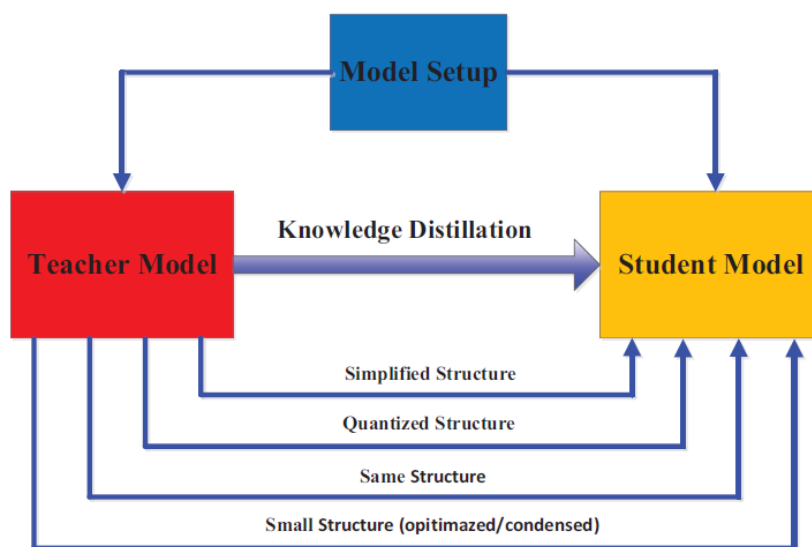
Slika 9. Primjer različitih destilacija [10]

### 7.3. Arhitektura učitelj-student (eng. *Teacher-Student Architecture*)

Dizajn arhitekture učitelj-student mreže je ključan za efikasnost destilacije i stjecanje znanja student modela od učitelj modela. Između kompleksnog učitelj modela i jednostavnijeg student modela postoji strukturna razlika i razlika u kapacitetima. Kompleksnost modela potječe od dvoje dimenzije dubine i širine modela. Transfer znanja većinom se odvija od dubljeg i šireg modela do tanjeg i užeg. Optimiziranje destilacije znanja pomoću efikasnih učitelj-student arhitektura smanjuje se strukturna razlika između modela. U najčešćim učitelj-student arhitekturama za prijenos znanja, student model ima jednu od sljedećih struktura:

- a) Pojednostavljena struktura (eng. *Simplified Structure*)- student model je isti ako i učitelj model samo sa manjim brojem slojeva (eng. *layer*) i manjim brojem informacija unutar slojeva
- b) Kvantizirana struktura (eng. *Quantized Structure*)- struktura mreže je očuvana
- c) Ista struktura (eng. *Same Structure*)- student ima istu mrežu kao i učitelj

- d) Kondenzirana mala struktura (eng. *Condensed Small Structure*)- mala mreža koja izvršava efikasno osnovne funkcije
- e) Optimirana mala struktura (eng. *Optimized Small Structure*)- mala mreža sa optimiziranom globalnom strukturom



Slika 10. Veze između učitelj i student modela u procesu destilacije znanja [10]

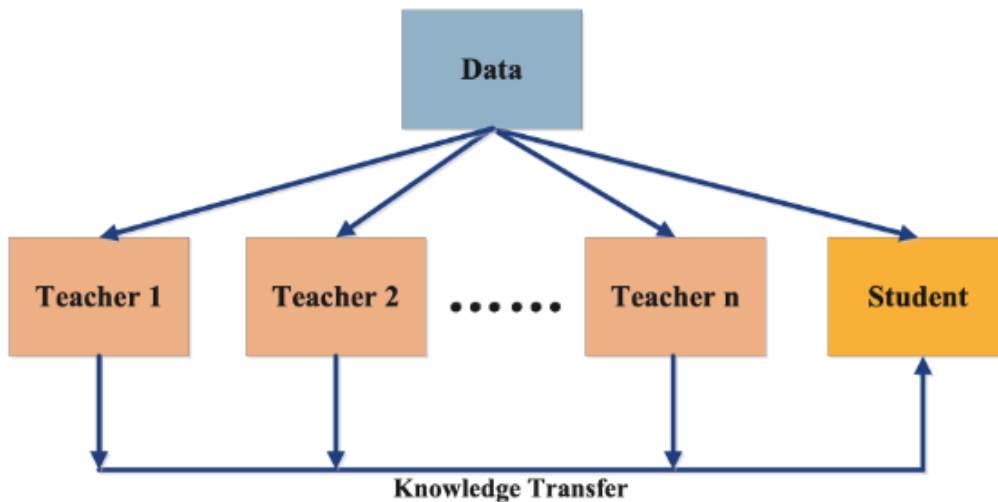
#### 7.4. Destilacijski algoritmi (eng. *Distillation Algorithms*)

Učinkovita i jednostavna metoda prijenosa znanja je izravno usklađivanje znanja (Znanje temeljeno na odgovorima, Znanje temeljeno na značajkama) ili raspodjela reprezentacija u prostoru značajki između učitelj i student modela. Predloženo je mnogo različitih modela za poboljšanje procesa prijenosa podatka u složenim okruženjima. Ovo su neki od tih algoritama [11]:

- a) Suparnička destilacija( eng. *Adversarial Distillation*)- koristi se za treniranje generator modela za sintetičke uzorke podataka koji sliče pravoj distribuciji podataka. Također se koristi za diskriminator model za razlikovanje sintetičkih i pravih uzoraka podataka

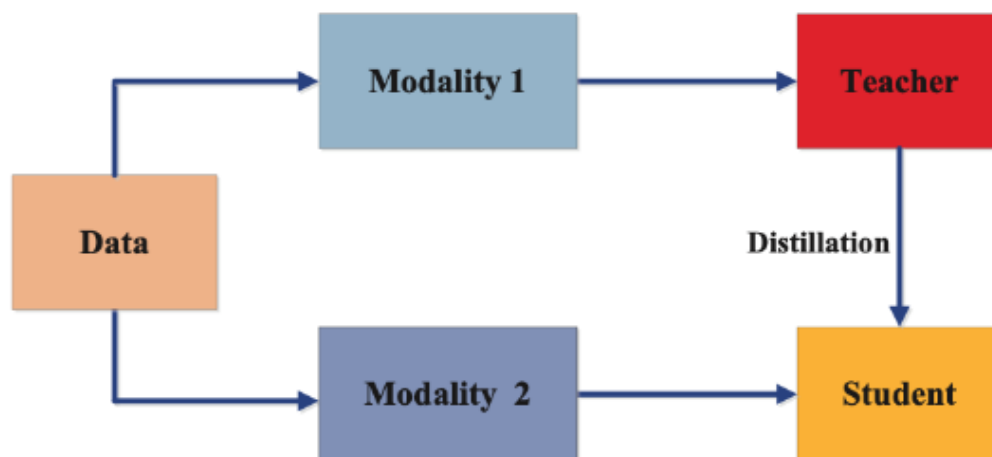


- b) Destilacija s više učitelja (eng. *Multi-Teacher Distillation*)- Student model prikuplja znanje od nekoliko različitih učitelj modela. Znanje od više učitelj modela se može kombinirati kao prosječni odgovor svih modela. Ovaj destilacijski algoritam može prenijeti više različitih vrsta znanja



Slika 11. Multi-Teacher Distillation [10]

- c) Križna destilacija (eng. *Cross-Model Distillation*)- najčešće se koristi u vizualnoj domeni. Učitelj model je treniran u jednoj modalnosti i njegovo se znanje prenosi na student model kojem je potrebno znanje iz druge modalnosti.



Slika 12. Cross-Model Distillation [10]

- d) Destilacija temeljena na grafikonima (eng. *Graph-Based Distillation*)- bilježi veze unutar podataka koristeći grafove. Grafovi se koriste za prijenos znanja sa učitelj na student model i za kontrolu znanja učitelj modela.
- e) Destilacija bez podataka (eng. *Data-Free Distillation*)- temelji se na sintetskim podacima jer set za trening nije dostupan radi privatnosti, sigurnosti isl.
- f) Kvantizirana destilacija (eng. *Quantized Distillation*)- koristi se za prijenos znanja od učitelj modela visoke preciznosti na student model niske preciznosti
- g) Cijeloživotna destilacija (eng. *Lifelong Distillation*)- temelji se na cjeloživotnom učenju, kontinuiranom učenju i meta-učenju, gdje se svo znanje akumulira i prenosi na buduća učenja
- h) Destilacija temeljena na pretraživanju neuronske arhitekture (eng. *Neural architecture search -Based Distillation*)- koristi se za identifikaciju odgovarajućih arhitektura student modela koje optimiziraju učenje iz učitelj modela

## **7.5. Primjene destilacije znanja (eng. *Applications of knowledge distillation*)**

Destilacija znanja je efektivna tehnika za kompresiju i ubrzavanje dubokih neuronskih mreža. Koristi se u više grana umjetne inteligencije kao što su vizijsko prepoznavanje, prepoznavanje govora i *natural language processing (NLP)*

### **7.5.1. Vizijsko prepoznavanje**

U posljednjih nekoliko godina razne metode destilacije znanja se koriste za kompresiju modela u različitim primjenama vizijskog prepoznavanja. Destilacija znanja kod vizijskog prepoznavanja se koristi :prepoznavanju lica, segmentaciji slika i videa, prepoznavanju pokreta, detekciji objekata, detekciji traka, identifikaciji osoba, detekciji pješaka, povratku slika itd.

### **7.5.2. Prepoznavanje govora**

Trenutno se sve više koriste sistemi za prepoznavanje govora u stvarnom vremenu ugrađeni u platforme koje se nalaze u računalima sa ograničenim mogućnostima, resursima i vremenom

odziva. Duboki kompleksni modeli ne mogu zadovoljiti zahtjeve, pa se naširoko koristi destilacija znanja u mnogim zadacima sa prepoznavanjem govora. Neke od primjena destilacije znanja u prepoznavanju govora su: identifikacija jezika, klasifikacija zvuka, prepoznavanje govornika, sinteza govora, detekcija akustičnog događaja, poboljšavanje govora, detekcija naglaska.

### 7.5.3. Natural Language Processing (NLP)

Destilacija znanja se intenzivno proučava u području NLP-a, jer je velika potreba za laganim, efektivnim i efikasnim modelima. Učitelj model sa velikim kapacitetom može prenijeti znanje bogato sa različitim podacima o jeziku ako bi trenirao manji student model. Tada manji student model može brzo obavljati razne zadatke vezane uz jezik. Znanja iz višejezičnih modela se mogu prenositi i dijeliti između modela što omogućava jednostavno i efikasno rješavanje višejezičnih zadataka. NLP zadaci koji koriste destilaciju znanja su: *neural machine translation (NMT)*, generator teksta, sistemi za odgovaranje na pitanja, povratak dokumenta, prepoznavanje teksta.

BERT (*Bidirectional Encoder Representations from Transformers*) kao višejezični model je privukao pažnju u području NLP-a. Problem kod BERT-a je to što je glomazno dubok model koji nije lako primijeniti, pa kako bi se riješio taj problem nastalo je nekoliko lakših varijanti BERT-a koristeći destilaciju znanja. Neki od tih su BERT-PKD, CLS, TinyBERT, BiLSTIM i DistilBERT.

U ovom istraživanju, za obradu informacija dobivenih iz pitanja 3, 4 i 5 tijekom intervjua, će se koristiti DistilBERT koji će se detaljnije objasniti u sljedećem poglavlju.

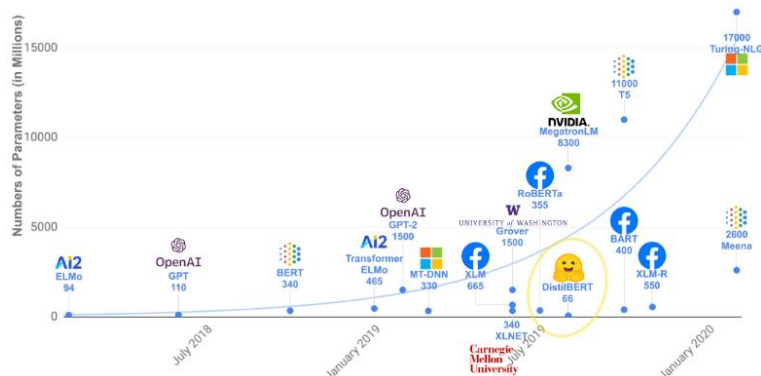
## 8. DistilBERT

Student model DistilBERT ima istu opću arhitekturu kao i učitelj model BERT. Za razliku od BERT-a DistilBERT ima upola manji broj slojeva i uklonjeni su *pooler*-i i *token-type embeddings*. Za vrijeme treniranja student modela uvodi se trostruka funkcija gubitka koja kombinira:

1. Gubitak modeliranja jezika (eng. *Language modeling loss*) – specifična za jezične modele i u slučaju BERT-a i DistilBERT-a koristi se *Masked language modeling loss (MLML)*. Koristi se za treniranje modela kod predviđanja skrivenih riječi ili tokena u rečenici pomoću konteksta.
2. Gubitak destilacije (eng. *Distillation loss*) – procjenjuje sličnost između BERT-ovih i DistilBERT-ovih output-a
3. Gubitak kosinus udaljenosti (eng. *Cosine-distance loss*) – uspoređuje skrivena stanja unutar BERT-a i DistilBERT-a. Dodavanjem gubitka kosinus udaljenosti usklađuju se smjerovi skrivenih vektora učitelj i student modela.

Trostruka funkcija gubitka se koristi kako bi se iskoristile prednosti induktivnih *biases* koje su naučili veliki učitelj modeli tijekom treniranja.

DistilBERT ima 40% manje parametara nego BERT, brži je za 60%, a pri tome zadržava oko 97% BERT-ovih mogućnosti kod razumijevanja jezika. Na sljedećoj slici [Slika 13] se može vidjeti usporedbu broja parametara DistilBERT-a i drugih modela



Slika 13. Broj parametara istreniranih modela za jezik [16]

## 8.1. Kod za sentiment analizu

Kod se nalazi u prilogu pod nazivom KOD ZA SENTIMENT ANALIZU i može se vidjeti na slikama [Slika 14 , Slika 15] koristi *Transformes* biblioteku za stvaranje *pipeline* za klasifikaciju teksta. Unaprijed istrenirani DistilBERT model koji se koristi za inicijalizaciju klasifikacije teksta je *bhadresh-savani/distilbert-base-uncased-emotion*. Također se koristi *pandas* biblioteka za manipulaciju podacima i kao alat za analizu. *Pandas* se koristi za obradu strukturiranih podataka, a u ovom slučaju se koristi CSV datoteka. Kao što je spomenuto u poglavlju 6 podaci u CSV datoteci su zapisani u prvom stupcu i kod ih konvertira u listu pomoću funkcije *tolist()* koja se sprema u varijablu *df\_positive*. Klasifikacija teksta upućuje da je podatke potrebno konvertirati u *string* vrstu podataka i to se radi pomoću funkcije *str()* svaki element liste se odvaja razmakom.

```
from transformers import pipeline

import pandas as pd

df = pd.read_csv('fifth_q.csv')
df_positive = df[df.columns[0]].values.tolist()
df_positive = ' '.join(map(str, df_positive))
```

### Slika 14. Sentiment analiza pomoću DistilBERT modela #1

Zatim se definira zadaća koda i model koji se koristi, u ovom slučaju je zadaća klasifikacija teksta *text-classification* i model je *bhadresh-savani/distilbert-base-uncased-emotion*, zatim se *top\_k* postavlja na *None* što znači da kod vraća sve moguće labele.

```
classifier = pipeline("text-classification", model='bhadresh-savani/distilbert-base-uncased-emotion', top_k=None)
prediction = classifier(df_positive, )

print(prediction)
```

### Slika 15. Sentiment analiza pomoću DistilBERT modela #2

Rezultati ovog koda se raspodjeljuju u šest kategorija. Labele tih kategorija su : sadness, joy, love, anger, fear, surprise.

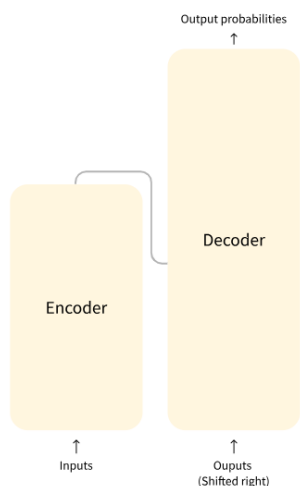
## 8.2. Transformers biblioteka i transformeri

Huggingface je razvio *transformers open-source* biblioteku koja se fokusira na modele temeljene na transformatorima. *Transformers* je biblioteka koja pruža API-ova za skidanje unaprijed istreniranih modela za rješavanje NLP zadataka, zadatke u području računalnog vid, zadatka za procesuiranje audio snimaka i govora u samo nekoliko linija koda. Modeli u ovoj biblioteci su uglavnom bazirani na transformatorima, ali ima modela koji nisu. Dva glavna cilja za dizajniranje ove biblioteke su:

1. Brzo i jednostavno korištenje
2. Pružanje modela novije generacije sa performansama što sličnijima onima originalnih modela

Arhitektura transformera je predstavljena u lipnju 2017.godine i originalno istraživanje se fokusiralo na zadatke prevođenja. Slijedno tome nastalo je nekoliko modela sa transformer arhitekturom kao bazom a neki od njih su: GPT (lipanj 2018.godine), BERT (listopad 2018.godine), GPT-2 (veljača 2019.godine), DistilBERT (listopad 2019.godine), BART (listopad 2019.godine), GPT-3 (svibanj 2020.godine).

Arhitektura tranformer modela se može podijeliti u dva dijela enkoder i dekoder, kao što se može vidjeti na sljedećoj slici [Slika 16].



**Slika 16. Pojednostavljena arhitektura transformera [16]**

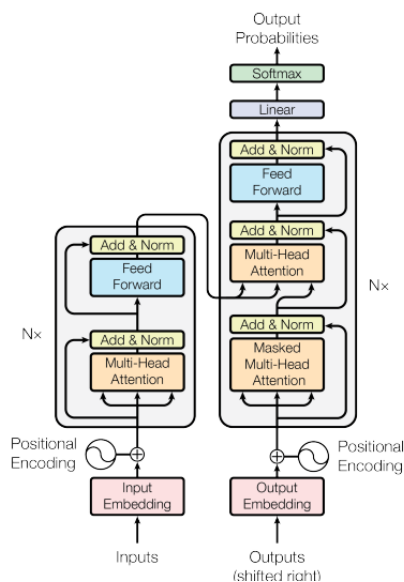
- Enkoder- prima ulaze i gradnji svojstva modela što znači da je model optimiran za stjecanje razumijevanja iz ulaza
- Dekoder-koristi svojstva, koje je izgradio enkoder, sa drugim ulazima kako bi kreirao ciljanu sekvencu što znači da je model optimiran za stvaranja izlaza

Oba dijela se mogu koristiti samostalno za rješavanje zadatka pa tako imamo:

- Enkoder modele koji su dobri kad je potrebno razumijevanje ulaza
- Dekoder modele koji su dobri za zadatke kao što je generiranje teksta
- Enkoder-dekoder modele koji se još zovu sequence-to-sequence modeli koji su dobri za zadatke generiranja za koje su potrebni ulazi kao što su prijevod ili sažimanje teksta.

Transformer modeli su građeni do specijalnih slojeva koji se nazivaju *attention layers* i ti slojevi govore modelu da obrati posebnu pažnju na određene riječi u zadanoj rečenici dok ostale više-manje ignorira. Na primjeru prevođenja, jer je to bio fokus na početku istraživanja transformera, to znači da model za vrijeme prevođenja jedne riječi promatra druge riječi koje mu u tom trenutku pomažu kod prijevoda.

Na sljedećoj slici [Slika 17] se može vidjeti originalna i detaljna arhitektura transformera.



Slika 17. Arhitektura transformera [27]

### 8.3. Model i pipeline

Model korišten u kodu za sentiment analizu na slici [Slika 15] je *bhadresh-savani/distilbert-base-uncased-emotion*. Model je *uncased* i to znači da ne pravi razliku između veliki i malih slova. Taj model je precizno podešen pomoću *emotion* baze podataka. Korišteni su sljedeći hiperparametri koji su zapisani u sljedećoj tablici [Tablica 2]:

**Tablica 2 Hiperparametri modela [31]**

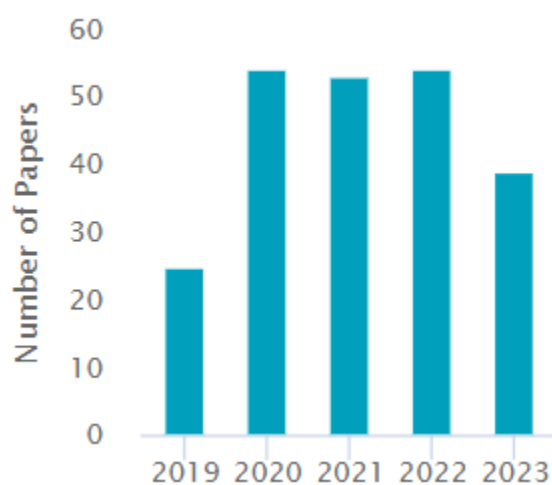
| <b>Learning rate</b> | <b>Batch size</b> | <b>Num_train_epochs</b> |
|----------------------|-------------------|-------------------------|
| 0,00002              | 64                | 8                       |

*Emotion* baza podataka ima 20 000 podataka. 16 000 podataka je odvojenu u set za trening dok se u test set i validacijski set odvojilo po 2000 podataka. Uz svaki podatak je zapisana emocija koja je prisutna u rečenici. Emocije su također označene brojevima:

- Sadness = 0
- Joy = 1
- Love = 2
- Anger = 3
- Fear = 4
- Surprise = 5

DistilBERT je treniran na istoj bazi podataka kao i BERT, a ta baza se zove *BookCorpus* koja se sastoji od 11 038 neobjavljenih knjiga i engleske Wikipedije isključujući liste, tablice i naslove poglavlja. Na sljedećoj slici [Slika 18] možemo vidjeti graf na kojem možemo vidjeti koliko članaka je koristilo BookCorpus kroz godine 2019. do 2023. Također u sljedećoj tablici [Tablica 3] se mogu vidjeti statistički podaci vezani za BookCorpus.





**Slika 18. Broj članaka koji su koristili BookCorpus od 2019.godine do 2023. godine [40]**

**Tablica 3 Statistički sažetak podatka o BookCorpus-u [40]**

| Broj knjiga | Broj rečenica | Broj riječi | Broj jedinstvenih riječi | Prosječan broj riječi po rečenici | Medijan broja riječi u rečenici |
|-------------|---------------|-------------|--------------------------|-----------------------------------|---------------------------------|
| 11 038      | 74 004 228    | 984 846 357 | 1 316 420                | 13                                | 11                              |

Emotion from Twitter je baza podataka često korištena za zadatke sentiment analize, a u ovom modelu je korištena za usporedbu modela. Baza se sastoji od tvitova prikupljenih sa Twitter-a i svakom tvitu je dodijeljena labela za odgovarajuću emocionalnu kategoriju. Iz svakog tvita su uklonjeni svi osjetljivi podaci i identifikacijske informacije kako bi se osigurala privatnost i anonimnost. Tvitovi su također očišćeni od šumova kao što su URL-ovi, emotikoni isl. .U sljedećoj tablici [Tablica 4] je prikazana usporedba modela koji su također precizno podešeni sa bazom podatak Emotin from Twitter.

Tablica 4 Usporedba modela [31]

| MODEL                           | TOČNOST | F1 REZULTAT | TESTNI UZORCI<br>PO SEKUNDI |
|---------------------------------|---------|-------------|-----------------------------|
| DISTILBERT-BASE-UNCASED-EMOTION | 93,8    | 93,79       | 398,69                      |
| BERT-BASE-UNCASED-EMOTION       | 94,05   | 94,06       | 190,152                     |
| ROBERTA-BASE-EMOTION            | 93,95   | 93,97       | 195,639                     |
| ALBERT-BASE-V2-EMOTION          | 93,6    | 93,65       | 182,794                     |

Rezultati ovog modela se raspodjeljuju u šest kategorija. Labele tih kategorija su : sadness, joy, love, anger, fear, surprise.

Za jednostavnije korištenje modela koriste se *pipeline*. To su objekti koji sažimaju većinu kompleksnog koda iz biblioteke pomoću jednostavnih API-ova namijenjenih za različite zadatke kao što su MLM, analizu sentimenta, izdvajanje značajki itd.

Primjer korištenja *pipeline* objekta se može vidjeti u kodu na slici [Slika 15]. Kad je pozvana *pipeline* funkcija u prikazanom primjeru u zagradi su zadani parametri:

*task*- parametar pomoću kojeg *pipeline* zna koji zadatak se rješava

*model*-parametar koji označava koji model će se koristiti za odabirni zadatak

*top\_k*- parametar koji govori koliko rezultata će se dobiti

Task i model parametri se zadaju u string vrsti datoteke, a top\_k se zadaje u integer vrsti datoteke.

#### 8.4. Izlazi dobiveni pomoću DistilBERT-a

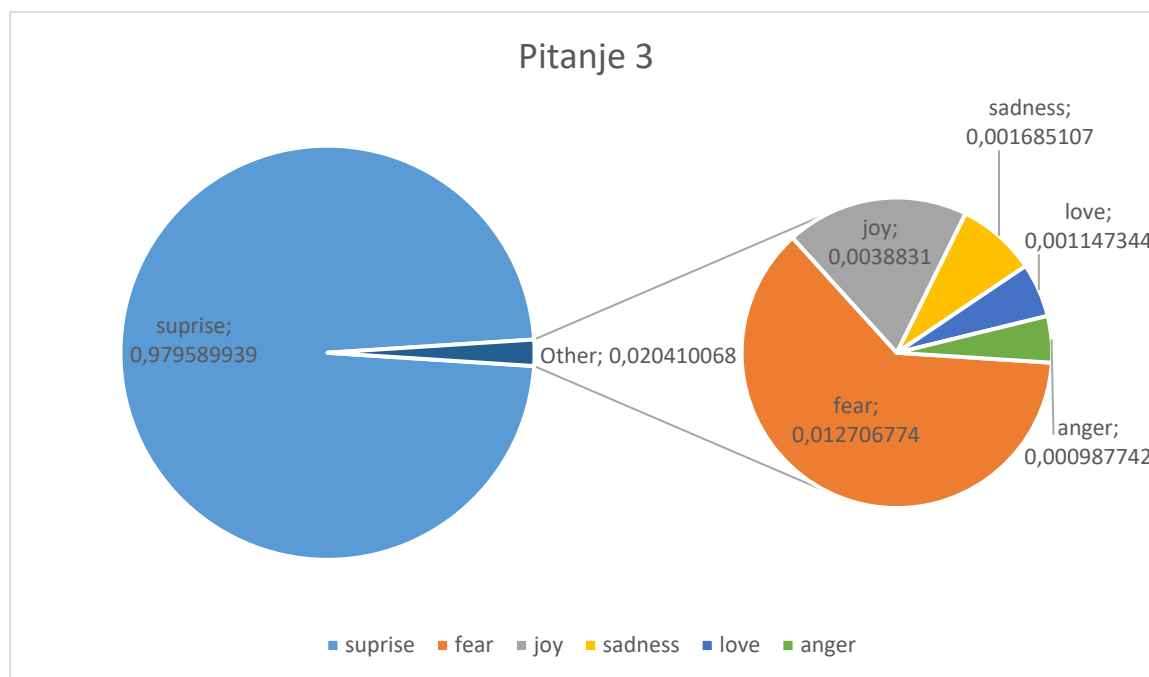
Model *bhadresh-savani/distilbert-base-uncased-emotion* je korišten za sentiment analizu pitanja 3, 4 i 5. Za svako pitanje će rezultati odgovara biti zapisani u dva oblika obliku grafikona i u obliku koji dobijemo nakon što pustimo kod.

Prije dobivenih rezultata će biti naglašeno kako glasi pitanje. Sva pitanja su napisana i ukratko objašnjena u poglavlju 5.1.

Pitanje 3. Koji raspon emocija je PLEA komunicirala s Vama?

```
[[{'label': 'surprise', 'score': 0.9795899391174316}, {'label': 'fear', 'score': 0.012706774286925793}, {'label': 'joy', 'score': 0.0038830996491014957}, {'label': 'sadness', 'score': 0.0016851071268320084}, {'label': 'love', 'score': 0.0011473441263660789}, {'label': 'anger', 'score': 0.0009877423290163279}]]
```

Na sljedećoj slici [Slika 19] je prikazan grafički prikaz dobivenog izlaza koda [KOD ZA SENTIMENT ANALIZU] za pitanje 3:

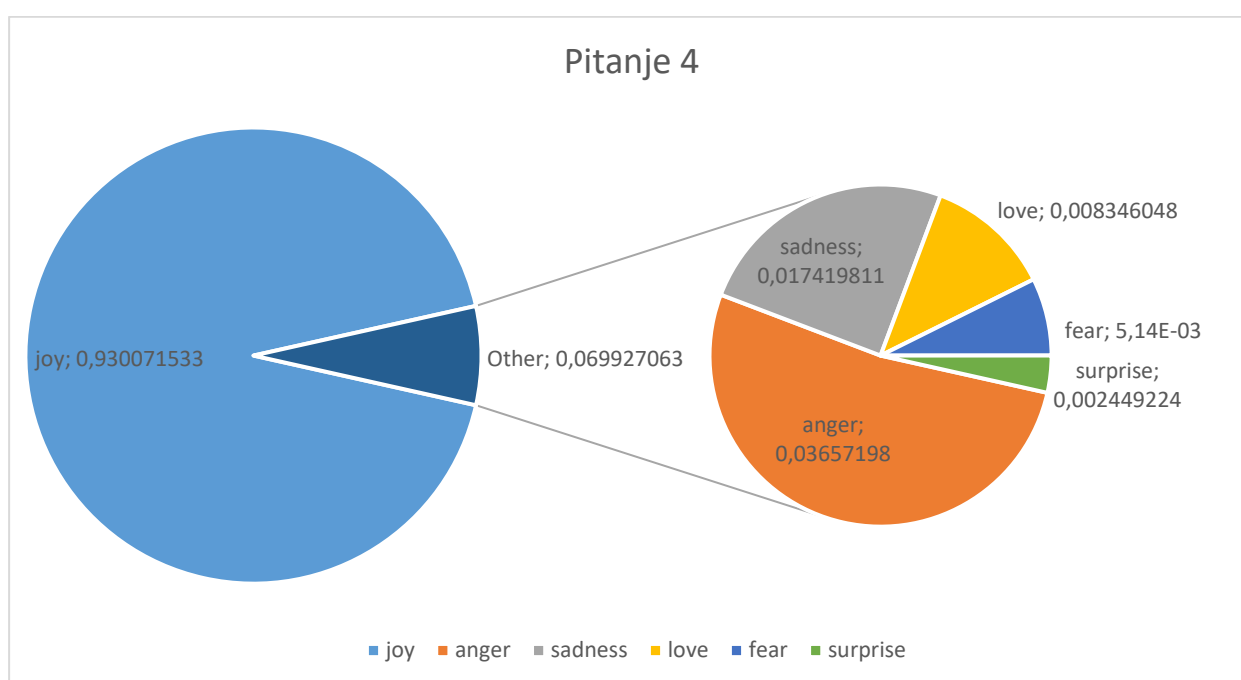


Slika 19. Pie grafikon za pitanje 3

Pitanje 4. Za koje osjećaje ste osjetili da PLEA najbolje komunicira s vama?

```
[[{'label': 'joy', 'score': 0.9300715327262878}, {'label': 'anger', 'score': 0.03657197952270508},
{'label': 'sadness', 'score': 0.017419811338186264}, {'label': 'love', 'score':
0.008346048183739185}, {'label': 'fear', 'score': 0.005141411907970905}, {'label': 'surprise',
'score': 0.0024492237716913223}]]
```

Na sljedećoj slici [Slika 20] je prikazan grafički prikaz dobivenog izlaza koda [KOD ZA SENTIMENT ANALIZU] za pitanje 4:

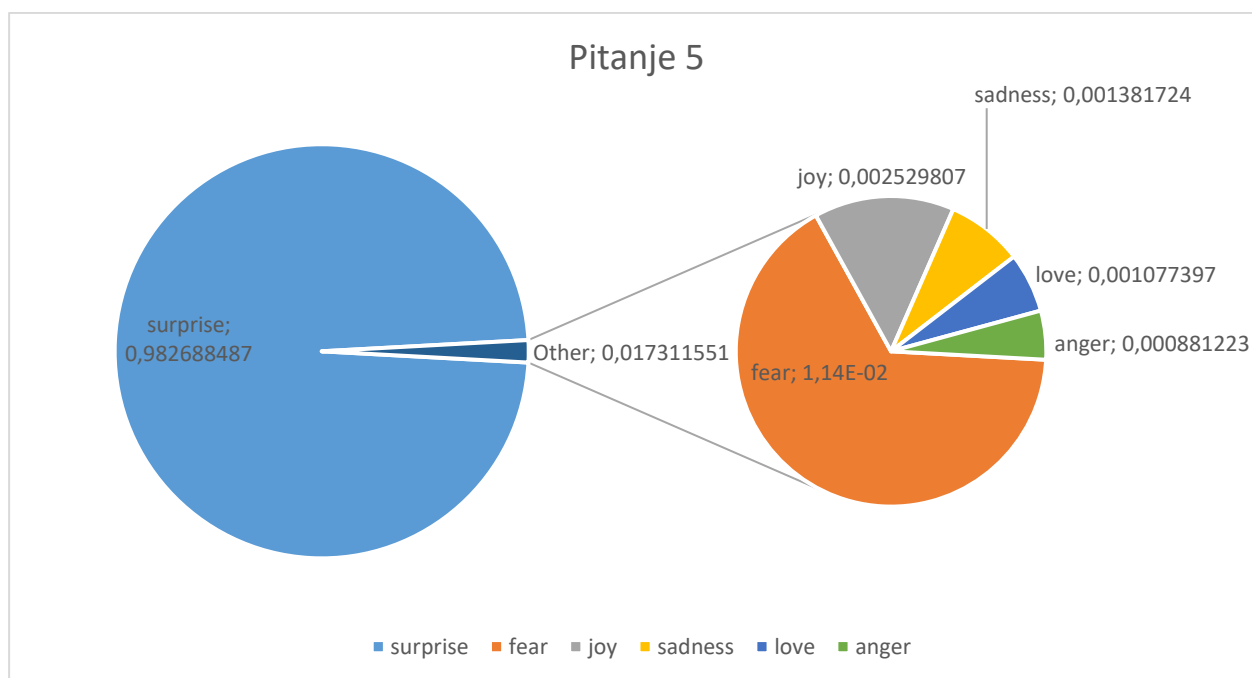


**Slika 20. Pie grafikon za pitanje 4**

Pitanje 5. Koje osjećaje ste Vi dijelili s PLEA robotom?

```
[[{'label': 'surprise', 'score': 0.9826884865760803}, {'label': 'fear', 'score':
0.011441431008279324}, {'label': 'joy', 'score': 0.002529806923121214}, {'label': 'sadness',
'score': 0.001381723559461534}, {'label': 'love', 'score': 0.0010773967951536179}, {'label':
'anger', 'score': 0.0008812227752059698}]]
```

Na sljedećoj slici [Slika 21] je prikazan grafički prikaz dobivenog izlaza koda [KOD ZA SENTIMENT ANALIZU] za pitanje 5:



**Slika 21. Pie grafikon za pitanje 5**

## 9. STATISIKA

Kako bi točno odradili statistiku za svako pitanje što efikasnije koristimo kod za prebrojavanje različitih odgovora. Kod se nalazi u prilogu pod nazivom KOD ZA PREBROJAVANJE RIJEČI/EMOCIJA. Koriste se isti CSV datoteke kao i za *bhadresh-savani/distilbert-base-uncased-emotion* model. Kao i u prijašnjem kodu [KOD ZA SENTIMENT ANALIZU] ovaj kod ih konvertira u listu pomoću funkcije *tolist()* koja se sprema u varijablu *df\_word*. Podatke potrebno konvertirati u *string* vrstu podataka i to se radi pomoću funkcije *str()* svaki element liste se odvaja razmakom.

```
import pandas as pd
df = pd.read_csv('fifth_q.csv')
df_word = df[df.columns[0]].values.tolist()
df_word = ' '.join(map(str, df_word))
```

### Slika 22. Kod za prebrojavanje riječi/odgovora #1

Definiramo funkciju *word\_count* koja uzima *string* kao ulazni parametar. *Word\_count* stvara prazni *dictionary* naziva *count* u kojem se sprema količina riječi i stvara listu riječi *words* u koju sprema riječi koje je dobila tako da ulazni string razdvaja na riječi. Pomoću *for* i *if* petlje prebrojava riječi u listi *words* i sprema količinu u *dictionary counts*. Za svaku riječ u listi *words* provjerava se je li se nalazi u *dictionary*-u kao ključ i ako se nalazi količina se povećava za jedan, a ako se ne nalazi ta riječ se dodaje u *dictionary* sa vrijednošću 1.

```
def word_count(string):
    counts = dict()
    words = string.split()

    for word in words:
        if word in counts:
            counts[word] += 1
        else:
            counts[word] = 1

    return counts

print(word_count(df_word))
```

### Slika 23. Kod za prebrojavanje riječi/odgovora #2

Sve što je odvojeno razmakom kod gleda kao zasebnu riječ. S obzirom da su neke emocije zapisane u nekoliko riječi, nisu korišteni razmaci za odvajanje riječi nego “\_” i tako smo dobili da kod koji prebrojava iste emocije.

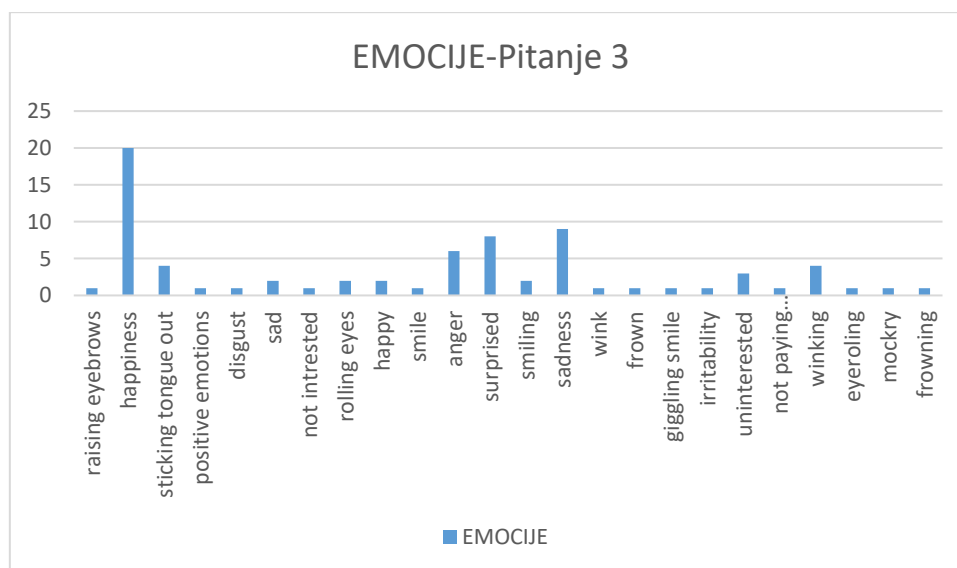
### 9.1. Izlazi koda za prebrojavanje riječi/emocija

Za svako pitanje će rezultati prebrojavanja odgovara biti zapisani u dva oblika obliku grafikona i u obliku koji dobijemo nakon što pustimo kod. Prije dobivenih rezultata će biti naglašeno kako glasi pitanje. Sva pitanja su napisana i ukratko objašnjena u poglavlju 5.1.

Pitanje 3. Koji raspon emocija je PLEA komunicirala s Vama?

```
{'happiness': 20, 'raising_eyebrows': 1, 'sticking_tongue_out': 4, 'positive_emotions': 1,
'disgust': 1, 'sad': 2, 'not_interested': 1, 'rolling_eyes': 2, 'happy': 2, 'smile': 1, 'anger': 6,
'surprised': 8, 'smiling': 2, 'sadness': 9, 'wink': 1, 'frown': 1, 'giggling_smile': 1, 'irritability': 1,
'uninterested': 3, 'not_paying_attention': 1, 'winking': 4, 'eyerolling': 1, 'mockery': 1, 'frowning':
1}
```

Na sljedećoj slici [Slika 24] je prikazan grafički prikaz dobivenog izlaza koda [KOD ZA PREBROJAVANJE RIJEČI/EMOCIJA] za pitanje 3:

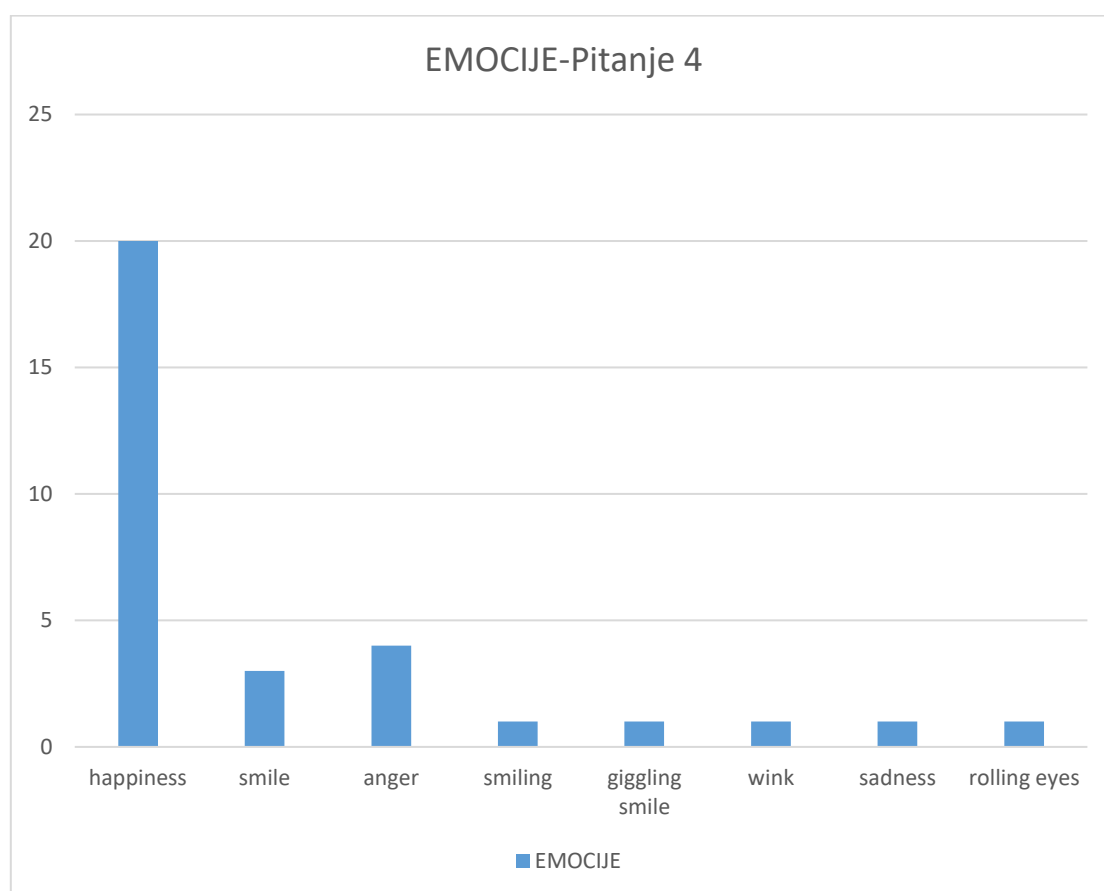


Slika 24. Bar grafikon za pitanje 3

Pitanje 4. Za koje osjećaje ste osjetili da PLEA najbolje komunicira s vama?

{'happiness': 20, 'smile': 3, 'anger': 4, 'smiling': 1, 'giggling\_smile': 1, 'wink': 1, 'sadness': 1, 'rolling\_eyes': 1}

Na sljedećoj slici [Slika 25] je prikazan grafički prikaz dobivenog izlaza [KOD ZA PREBROJAVANJE RIJEČI/EMOCIJA] za pitanje 4:



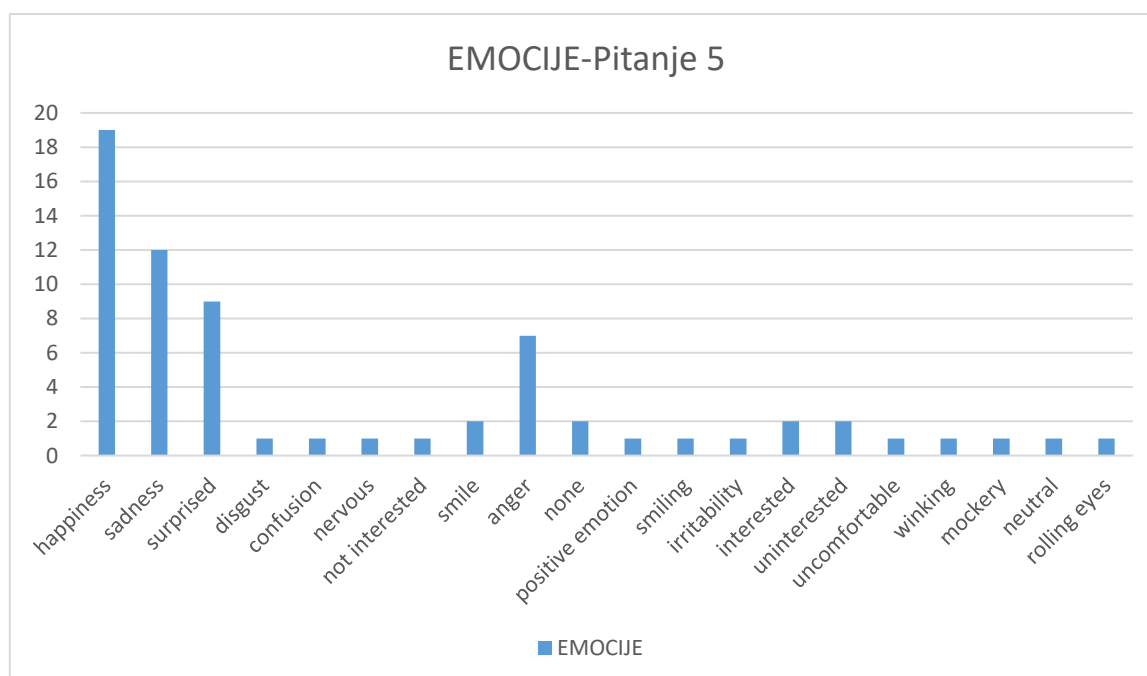
Slika 25. Bar grafikon za pitanje 4



Pitanje 5. Koje osjećaje ste Vi dijelili s PLEA robotom?

```
{'happiness': 19, 'sadness': 12, 'surprised': 9, 'disgust': 1, 'confusion': 1, 'nervous': 1, 'not_interested': 1, 'smile': 2, 'anger': 7, 'none': 2, 'positive_emotions': 1, 'smiling': 1, 'irritability': 1, 'interested': 2, 'uninterested': 2, 'uncomfortable': 1, 'winking': 1, 'mockery': 1, 'neutral': 1, 'rolling_eyes': 1}
```

Na sljedećoj slici [Slika 26] je prikazan grafički prikaz dobivenog izlaza koda [KOD ZA PREBROJAVANJE RIJEČI/EMOCIJA] za pitanje 5:



Slika 26. Bar grafikon za pitanje 5

## 10. ZAKLJUČAK

U ovom radu su obrađeni podaci prikupljeni za istraživanje o sposobnosti PLEA robota da učini pametno okruženje ugodno i gostoljubivo. Podaci su prikupljeni pomoću intervjua baziranog na upitniku od 12 pitanja. Fokus analize u ovom radu je bio na pitanjima 3, 4 i 5 koja daju podatke o tome kako PLEA robot izvodi svoju ulogu koja je komunikacija emocija sa sudionicima.

Analiza podataka se sastoji od sentiment analize i klasične statističke analize. Sentiment analiza je obavljena pomoću modela temeljenog na DistilBERT-u, a klasična analiza pomoću koda čija je osnovna zadaća prebrojavanje riječi u dokumentu.

Za pitanje 3 rezultati analize pokazuju da je najčešće komunicirana emocija sreća, a zatim tuga. Najviše emocija upada u kategoriju surprise.

Za pitanje 4 rezultati analize pokazuju da je najbolje komunicirana emocija sreća, a zatim ljutnja. Najviše emocija upada u kategoriju joy.

Za pitanje 5 rezultati analize pokazuju da je najčešća emocija koju su sudionici dijelili sa PLEA robotom sreća, a zatim tuga. Najviše emocija upada u kategoriju surprise.

## LITERATURA

- [1] Stipančić, T.; Koren, L.; Korade, D.; Rosenberg, D.; PLEA: A social robot with teaching and interacting capabilities, *Journal of Pacific Rim Psychology*, Sječanj-Prosinac 2021.god
- [2] Boyarskaya E., Sebastian A., Bauermann T., Hecht H., Tüscher O.; The Mona Lisa effect: Neural correlates of centered and off-centered gaze. *Human Brain Mapping*, 36(2), 619–632. Crossref., Veljača 2015.
- [3] Al Moubayed S., Beskow J., Skantze G., Granström B. Furhat; A back-projected human-like robot head for multiparty human–machine interaction. In *Cognitive behavioural systems* (pp. 114–130). Berlin: Springer. Crossref., 2012
- [4] <https://www.simplilearn.com/what-is-data-collection-article> [posjećeno 13.6.2023.]
- [5] <https://www.simplilearn.com/data-collection-methods-article> [posjećeno 13.6.2023.]
- [6] <https://www.lotame.com/what-are-the-methods-of-data-collection/> [posjećeno 13.6.2023.]
- [7] <https://www.pewresearch.org/our-methods/u-s-surveys/writing-survey-questions/> [posjećeno 12.6.2023.]
- [8] <https://www.england.nhs.uk/wp-content/uploads/2018/01/bitesize-guide-writing-an-effective-questionnaire.pdf> [posjećeno 12.6.2023.]
- [9] <https://www.businessinsider.com/guides/tech/what-is-csv-file> [posjećeno 1.7.2023.]
- [10] Gou, J.; Yu, B.; Maybank, S. J.; Tao, D.: Knowledge Distillation: A Survey, *International Journal of Computer Vision*, Ožujak 2021
- [11] <https://neptune.ai/blog/knowledge-distillation> [posjećeno 2.7.2023.]
- [12] <https://analyticsindiamag.com/top-10-gpus-for-deep-learning-in-2021/> [posjećeno 4.6.2023.]
- [13] [https://keras.io/examples/vision/knowledge\\_distillation/](https://keras.io/examples/vision/knowledge_distillation/) [posjećeno 5.6.2023.]
- [14] <https://analyticsindiamag.com/a-beginners-guide-to-knowledge-distillation-in-deep-learning/> [posjećeno 5.6.2023.]

- [15] Sanh, V.; Debut, L.; Chaumond, J.; Wolf, T.: DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter, 2020
- [16] <https://huggingface.co/learn/nlp-course/chapter1/4?fw=pt> [posjećeno 28.6.2023.]
- [17] <https://arunm8489.medium.com/understanding-distil-bert-in-depth-5f2ca92cf1ed> [posjećeno 10.6.2023]
- [18] <https://towardsdatascience.com/distillation-of-bert-like-models-the-theory-32e19a02641f> [posjećeno 10.6.2023.]
- [19] <https://towardsdatascience.com/masked-language-modelling-with-bert-7d49793e5d2c> [posjećeno 11.6.2023.]
- [20] <https://towardsdatascience.com/everything-you-need-to-know-about-albert-roberta-and-distilbert-11a74334b2da> [posjećeno 11.6.2023.]
- [21] <https://towardsdatascience.com/how-to-use-bert-from-the-hugging-face-transformer-library-d373a22b0209> [posjećeno 11.6.2023.]
- [22] <https://www.analyticsvidhya.com/blog/2022/11/introduction-to-distilbert-in-student-model/> [posjećeno 11.6.2023]
- [23] [https://www.youtube.com/playlist?list=PLo2EIpI\\_JMQvWfQndUesu0nPBAtZ9gP1o](https://www.youtube.com/playlist?list=PLo2EIpI_JMQvWfQndUesu0nPBAtZ9gP1o) [posjećeno 28.6.2023]
- [24] <https://github.com/huggingface/transformers> [posjećeno 28.6.2023]
- [25] <https://huggingface.co/docs/transformers/index> [posjećeno 28.6.2023.]
- [26] <https://huggingface.co/transformers/v3.0.2/index.html> [posjećeno 28.6.2023.]
- [27] Vaswani, A.; Shazeer, N.; Parmas, N.; Uszkoreit, J.; Jones, L.; Gomez, AN.; Kaiser, L.; Polosukhin, I.; Attention Is All You Need, Prosinac 2017
- [28] <https://huggingface.co/distilbert-base-uncased> [posjećeno 28.6.2023.]
- [29] [https://huggingface.co/docs/transformers/tasks/sequence\\_classification](https://huggingface.co/docs/transformers/tasks/sequence_classification) [posjećeno 28.6.2023.]
- [30] Wolf, T.; Debut, L.; Snah, V.; Chaumond, J.; Delangue, A; Mos, A.; Cistac, P.; Rault, T.; Louf, R.; Funtowicz, M.; Davison, J.; Shleifer, S.; von Platen, P.; Jernite, Y.; Plu, J.; Xu, C.; Le Scao, T.; Gugger, S.; Drame, M.; Lhoest, Q.; Rush, A.M.; Transformers: State-of-the-Art Natural Language Processing, Studeni 2020

- [31] <https://huggingface.co/bhadresh-savani/distilbert-base-uncased-emotion> [posjećeno 28.6.2023.]
- [32] <https://huggingface.co/docs/transformers/training> [posjećeno 28.6.2023.]
- [33] <https://www.section.io/engineering-education/hugging-face/> [posjećeno 28.6.2023]
- [34] <https://towardsdatascience.com/custom-transformers-in-python-part-ii-6fe111fc82e4> [posjećeno 28.6.2023]
- [35] <https://huggingface.co/datasets/dair-ai/emotion> [posjećeno 28.6.2023.]
- [36] [https://huggingface.co/docs/transformers/v4.30.0/en/main\\_classes/pipelines#pipelines](https://huggingface.co/docs/transformers/v4.30.0/en/main_classes/pipelines#pipelines) [posjećeno 28.6.2023.]
- [37] [https://huggingface.co/docs/transformers/v4.30.0/en/main\\_classes/pipelines#transformers.TextClassificationPipeline](https://huggingface.co/docs/transformers/v4.30.0/en/main_classes/pipelines#transformers.TextClassificationPipeline) [posjećeno 28.6.2023.]
- [38] [https://huggingface.co/docs/transformers/v4.30.0/en/task\\_summary](https://huggingface.co/docs/transformers/v4.30.0/en/task_summary) [posjećeno 28.6.2023.]
- [39] <https://huggingface.co/docs/transformers/v4.30.0/en/philosophy> [posjećeno 28.6.2023.]
- [40] <https://paperswithcode.com/dataset/bookcorpus> [posjećeno 28.6.2023.]
- [41] <https://huggingface.co/datasets/mteb/emotion> [posjećeno 28.6.2023]

## PRILOZI

### I. KOD ZA SENTIMENT ANALIZU

```
from transformers import pipeline

import pandas as pd

df = pd.read_csv('fifth_q.csv')
df_positive = df[df.columns[0]].values.tolist()
df_positive = ' '.join(map(str, df_positive))

classifier = pipeline("text-classification", model='bhadresh-savani/distilbert-base-uncased-emotion', top_k=None)
prediction = classifier(df_positive, )

print(prediction)
```

## II. KOD ZA PREBROJAVANJE RIJEČI/EMOCIJA

```
import pandas as pd
df = pd.read_csv('fifth_q.csv')
df_word = df[df.columns[0]].values.tolist()
df_word = ' '.join(map(str, df_word))
1 usage
def word_count(string):
    counts = dict()
    words = string.split()

    for word in words:
        if word in counts:
            counts[word] += 1
        else:
            counts[word] = 1

    return counts

print(word_count(df_word))
```