

Predmet: Veštačka inteligencija

Projektni zadatak: Praćenje aktivnosti studenata na društvenim mrežama

Profesor:  
Leonid Stoimenov

Studenti:  
Teodora Djordjević  
Lazar Kadić

## 1. Uvod

Društvene mreže predstavljaju veoma bitnu stavku u svakodnevnom životu mladih, ponekad i sa većim prioritetom nego obaveze iz "realnog" sveta. Upravo zbog tolikog značaja i razvoja društvenih mreža, moguće je i dosta toga naučiti na njima, kako o pojedincu tako i o posmatranoj ciljnoj grupi. U našem slučaju, ciljna grupa su studenti, a konkretna društvena mreža koju razmatramo je Reddit. Reddit predstavlja platformu sličnu Twitter-u ili Facebook-u, pri čemu je veći akcenat stavljen na jasnu podelu tema (subreddit-a) o kojima se diskutuje, iznosi svoje mišljenje i stavovi, ali i zatraži pomoć za konkretan problem, što može biti od velike koristi upravo studentima. Analizom takvih objava, i generalno analizom objava pojedinaca, možemo proceniti njihovo zadovoljstvo i osećanja o određenoj temi, pa samim tim korigovati dalji pristup kako bi poboljšali subjektivni osećaj o konkretnim temama, ali i celokupno radno okruženje.

## 2. Projekat

Kako smo ranije pomenuli, za realizaciju ovog projekta koristiće se platforma Reddit. Reddit je web sajt u vidu foruma zasnovan na principu postavljanja, komentarisanja i ocenjivanja web sadržaja. Korisnici ili zajednice na sajtu postavljaju sadržaj u vidu teksta ili nekog drugog formata. Takav sadržaj je podeljen u subreddit-e koji pokrivaju neku određenu temu. Kako se na ovakvom web sajtu nalazi pregršt podataka koji mogu biti prikupljeni, analizirani i obrađivani, stvorila se i potreba za aplikacijama koje će to omogućiti. Danas, uz pomoć veštačke inteligencije, a na osnovu podataka, mogu se izvoditi i određeni zaključci.

Upravo zbog prethodno navedenog, za obradu ove teme implementirana je aplikacija u Python-u koja će koristiti određene biblioteke i doći do određenih zaključaka.

### 3. Zadatak aplikacije

Zadatak aplikacije je da na osnovu navedene zajednice, u našem slučaju je to zajednica studenata pod nazivom "College", koja piše o temama koje su aktuelne među studentima a vezane su za fakultet, prikupi njen deljeni sadržaj i zaključi da li su objave pozitivne, negativne ili neutralne. Zatim, pored osnovnih informacija o svakoj objavi pojedinačno, prikazaće i u koju, od tri navedene grupe objava spada.

### 4. Primena projekta

Sada, kada znamo koji je zadatak aplikacije, možemo reći da je ovo moderan način za praćenje i proučavanje studenata. Tradicionalni pristup koji se odnosi na usmene razgovore ili pisane formulare, nosi sa sobom velika ograničenja, od kojih je jedno broj studenata koji se mogu ispitati. Upravo ovde vidimo prednost modernog načina koji daje mogućnost da se za kratko vreme obradi velika količina podataka. Bilo koja organizacija, koja želi da brzo prouči osećanja/raspoloženja/mišljenja određene grupe, može da primeni ovaj projekat.

### 5. Implementacija

U nastavku se nalaze sve biblioteke koje su korišćene u projektu.

5.1. Requests - biblioteka koja je de facto standard za pravljenje HTTP zahteva u Python-u.

5.2. Pandas - biblioteka koja pruža brže i fleksibilne strukture podataka koje se koriste za rad i manipulaciju podacima.

5.3. Textblob - biblioteka za obradu tekstualnih podataka koja nad njima može da izvodi operacije kao što su analiza osećanja i klasifikacija.

## 6. Textblob

Kako je tema projekta praćenje aktivnosti studenata, kakvo je njihovo raspoloženje, odnos prema fakultetu i njihov osećaj, zadržaćemo se malo na biblioteci textblob.

Naveli smo da ona može nad podacima da izvršava operaciju analiza osećanja, pa hajde da pojasnimo šta je to zapravo.

Analiza osećanja nam može pomoći da dešifrujemo raspoloženje i emocije pojedinca ili grupe, i prikupimo pronicljive informacije u vezi sa kontekstom. Analiza osećanje je proces analize podataka i klasifikacije na osnovu potrebe istraživanja.

Textblob je biblioteka za obradu prirodnog jezika(NLP). Za pristupe koji su zasnovani na leksici, osećanja su definisana njenom semantičkom orijentacijom i intenzitetom svake reči u rečenici. To zahteva unapred definisan rečnik koji klasifikuje negativne i pozitivne reči. Kako je recenica predstavljena kao niz reči, nakon dodeljivanja ocena svim rečima, konačni osećaj se izračunava operacijom udruživanja. Textblob vraća polaritet rečenice. Polaritet se nalazi u opsegu [-1,1]. -1 definiše negativno osećanje, 0 neutralno osećanje, a 1 pozitivno osećanje. Textblob takođe ima i semantičke oznake koje pomažu u detaljnoj analizi, kao što su emotikoni, uzvičnici,...

## 7. Kod

7.1. Neophodni kredencijali za korišćenje Reddit API-ja, koji se generišu na osnovu korisničkog imena i šifre i traju ~2h. (Ipak može da se desi da ćete moći da pokrenete i sa ovim kredencijalima, pokušajte)

```
CLIENT_ID = '7_XBLzur5GVXFz7Y25NsuQ'  
SECRET_TOKEN = 'HzdUa_9xEzxKcMSb8kJD3VB0JmZg8g'
```

7.2. Implementacija klase RedditAuth koja služi za autentifikaciju i pribavljanje access token za pristup Reddit API-ju.

```
class RedditAuth():  
  
    def authenticate(self):  
        auth = requests.auth.HTTPBasicAuth(CLIENT_ID, SECRET_TOKEN)  
        data = {'grant_type': 'password',  
                'username': 'Macupikcu',  
                'password': 'SljivovicaRakija15'}  
        headers = {'User-Agent': 'OurAPI/0.0.1'}  
  
        res = requests.post('https://www.reddit.com/api/v1/access_token', auth=auth, data=data, headers=headers)  
        TOKEN = res.json()['access_token']  
        headers['Authorization'] = f'bearer {TOKEN}'  
        return headers
```

7.3. Implementacija klase RedditAnalyze kojom je moguće pretraživanje i pribavljanje željenih objava.

```
class RedditAnalyze():  
  
    def __init__(self):  
        self.headers = RedditAuth().authenticate()  
  
    def data_collection(self):  
        res = requests.get('https://oauth.reddit.com/r/college', headers=self.headers, params={'limit': '20'})  
  
        df = pd.DataFrame()  
  
        for post in res.json()['data']['children']:  
            temp = self.analyze_sentiment(post['data']['title'])  
            df = df.append({  
                'Author': post['data']['author'],  
                'Title': post['data']['title'],  
                'Selftext': post['data']['selftext'],  
                'Sentiment': temp  
            }, ignore_index=True)  
  
        return df
```

7.4. Implementacija metode analyze\_sentiment unutar koje se vrši analiza sentimenata uz pomoć biblioteke textblob.

```
def analyze_sentiment(self, text):  
    analysis = TextBlob(text)  
  
    if analysis.polarity > 0:  
        return "Positive post"  
    elif analysis.polarity == 0:  
        return "Neutral post"  
    else:  
        return "Negative post"
```

7.5. Implementacija metode main unutar koje se koristi objekat klase RedditAnalyze i prikazuje rezultat analize.

```
if __name__ == '__main__':  
    reddit = RedditAnalyze()  
    result = reddit.data_collection()  
    print(result)
```

## 8. Instalacije

S obzirom da je projekat rađen u python-u, neophodno je na računaru imati instaliran python, kao i korišćene biblioteke (korišćenjem naredbe pip install i navođenjem imena biblioteka).

## 9. Pokretanje

Aplikacija se pokreće preko terminala, pozicioniranjem na folder gde se navodi projekat i pozivanjem naredbe 'python reddit.py'.

## 10. Rezultat

Na narednoj slici prikazan je rezultat izvršenja aplikacije. Iz navedenih rezultata se vidi da ima 6 pozitivnih, 11 neutralnih i 4 negativne objave.

```
D:\7. semestar\Vestacka inteligencija\New folder\Reddit>python reddit.py  
0 Author Title Selftext Sentiment  
1 Laurasaur28 FAUSA/financial aid questions? Get help here! **All questions about federal student aid, the... Neutral post  
2 Mango_Hero87 College freshmen getting messages from a rand... Negative post  
3 Soft_Essay_6079 Don't lose hope: I just got accepted into a 3.... Negative post  
4 Competitive_Monk2954 Anyone else already so burnout that you've wat... Didn't take time to chill over the break, and ... Neutral post  
5 MakinBaconPancakezz Should I take an extra class and finish early?... So, I was originally going to be finished in M... Positive post  
6 burntouttryhard How do I get close with my professors? Ok, all the professors and more outgoing stude... Neutral post  
7 MightOk9038 Should I tell my roommate he shakes our bunk b... My roommate and I get along very well, but we ... Neutral post  
8 Haunting-World1877 College friends First year in college. First semester made no ... Neutral post  
9 KingBlackthorn1 Does anyone else feel bad when taking a break? I know this will sound potentially weird but I... Negative post  
10 DroidLovvr6125 Top colleges aren't worth it for most people, ... Top colleges really aren't worth it for most p... Positive post  
11 OutlandishnessNo4687 I feel like everyone is looking at me as I wal... Do others have the same feelings or am I just ... Neutral post  
12 HDSage-X Is anyone else worried that they won't be able... Internship rejection, stress from exams, press... Positive post  
13 Chocolatomulatto Is pdfuni.com a safe site? I've found a book (pdf) that I really need on ... Positive post  
14 randomles123 Greek life Can someone explain what Greek Life means beca... Neutral post  
15 selena_hilland Should I email my professor about why I missed... School started Tuesday so before my classes st... Neutral post  
16 Informal_Release_744 How long are college classes a day Specifically speaking in George Brown, Seneca ... Negative post  
17 squareonementors Two Recent Grads Offering Job Help Hi everyone! We are excited to announce the la... Neutral post  
18 yourdaisyentendu Has anyone here taken Contemporary math (quant... I just started class this week. I do not need ... Positive post  
19 catradish How to write a good program application essay ... Hi, i'm going to be applying to 3-4 community ... Positive post  
20 underweather813 What is your zoom note taking set up? Do you zoom on your laptop and take notes on a... Neutral post  
margie67_ When will I hear back about my dream school's ... Hi, I am waiting on a response from Gallaudet ... Neutral post
```

## 11. Literatura

- 1) <https://towardsdatascience.com/my-absolute-go-to-for-sentiment-analysis-textblob-3ac3a11d524>
- 2) <https://www.geeksforgeeks.org/python-pandas-dataframe/>
- 3) <https://www.reddit.com/dev/api>
- 4) <https://www.reddit.com/prefs/apps>
- 5) <https://stackoverflow.com/questions/28955541/how-to-get-access-token-reddit-api>