FER Projekt 2018/19 Tehnička dokumentacija Verzija 1.5

VODITELJ: JAKOV DORONTIĆ

STUDENTSKI TIM: VIKTOR GABAJ

MATEJ JURIĆ BARBARA KLIER

FILIP PTIČEK

DORA PUŠELJ

MAGDALENA ŠIMUNEC

NINA ZUCCON

NASTAVNIK: DOC. DR. SC. MARIN VUKOVIĆ

FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

SADRŽAJ

1. Opis razvijenog proizvoda	2
2. Upute za korištenje	3
2.1 LIUM	3
2.2 Web stranica	7
2.3 Big Data	10
3. Literatura	15

TEHNIČKA DOKUMENTACIJA

Verzija:

Datum: 16.01.2019

1.5

1. OPIS RAZVIJENOG PROIZVODA

U svrhu razvoja projekta koristili smo alat LIUM. *LIUM Speaker Diarization* je set alata razvijen na Sveučilištu u Maineu u Francuskoj. Primarno je namijenjen razvrstavanju govornika u vijestima francuskih radio emisija. Razvio se iz seta alata Mistral_Seg.

Radi u nekoliko koraka. Prvo računa MFC (Mel-frequency cepstrum) koeficijente pomoću Sphinx 4 alata. Nakon toga traži granice govora pomoću matrice kovarijanci, dobivene klizećim prozorom od 5 sekundi kroz cijeli snimak, i mjera KL2, Gaussove divergencije i BIC (Bayesian information criterion) mjere. Dobivene granice, odnosno segmente, spaja u klastere opet pomoću BIC mjere. Ti se klasteri u idućem koraku ponovno segmentiraju korištenjem Viterbi algoritma koji pronalazi najvjerojatnije putove u skrivenim Markovljevim modelima klastera. U istom se koraku segmenti duži od 20 sekundi usitnjuju.

Na red dolazi otkrivanje govora gdje se uklanja pozadinska glazba i tišina, opet se koriste skriveni Markovljevi modeli, ali i GMM (mještoviti Gussovi modeli). U tom se koraku računaju modeli (GMM) svakog segmenta koji se u zadnjem koraku koriste za spajanje segmenata u klastere.

Slični modeli se spajaju u klastere, a to su zapravo govornici. Tako je svaki govornik opisan svojim modelom i moguće je pratiti tog govornika kroz više snimaka (cross-show diarization).

2. UPUTE ZA KORIŠTENJE

2.1 LIUM

U ovom poglavlju opisat ćemo rad alata LIUM zajedno sa primjerom, tj. analizom jednog audio zapisa. Prvi korak jest dohvatiti program (lium_spkdiarization-8.4.1.jer.gz) sa adrese:

http://www-lium.univ-lemans.fr/diarization/doku.php/download

Isto tako trebat će nam audio zapis razgovora (npr. *AK01032017.pm3*). Kako naš alat radi samo sa .wav audio formatom potrebno je dodatno prilagoditi datoteku prije analize pomoću:

https://audio.online-convert.com/convert-to-wav

Obavezno prilagoditi sljedeće parametre: 1) Change bit resolution: 16 Bit

2) Change sample rate: 16000 Hz

Verzija:

Datum: 16.01.2019

1.5

3) Change audio channels: mono

Na taj način ćemo osigurati visoku preciznost analize. Preimenujmo naš novi audio zapis u *showName.wav*

Program pokrećemo slijedećom naredbom (CMD naredba):

```
java -Xmx2024m -jar ./LIUM_SpkDiarization.jar \ --fInputMask=./showName.wav
--sOutputMask=./showName.seg --doCEClustering showName
```

Možemo uočiti da za izvršavanje navedene naredbe moramo imati tri datoteke u istoj mapi (Slika 2.1.1), a to su: LIUM_SpkDiarization.jar, showName.wav, showName.seg (u showName.seg će se zapisati rezultat analize)

Name	Туре	Size
LIUM_SpkDiarization.jar	Executable Jar File	26.333 KB
🚳 run.bat	Windows Batch File	1 KB
showname.seg	SEG File	8 KB
📤 showname.wav	VLC media file (.w	36.405 KB

Slika 2.1.1: Potrebne datoteke za analizu

Datoteka run.bat sadrži maloprije navedenu naredbu: java -Xmx2...

Slijedi analiza audio zapisa koja, ovisno o jačini računala i veličini datoteke, može potrajati i do nekoliko minuta. Slika 2.1.2 prikazuje nam rezultat analize.

```
;; cluster S0 [ merge HCLR 1 = S0 in S2 with 1.6808953775812936 ] [ score
showName 1 0 515 M S U S0
showName 1 923 923 M S U S0
showName 1 2501 964 M S U S0
showName 1 25437 1132 M S U S0
showName 1 26569 908 M S U S0
showName 1 27477 1930 M S U S0
showName 1 115299 1192 M S U S0
;; cluster S1 [ merge HCLR 0 = S1 in S166 with 1.6221031516837654 ] [ scc
showName 1 515 379 M S U S1
showName 1 3975 1178 M S U S1
showName 1 5153 1783 M S U S1
showName 1 18973 1297 M S U S1
showName 1 20270 754 M S U S1
showName 1 29407 1851 M S U S1
showName 1 31258 1137 M S U S1
showName 1 32395 1986 M S U S1
showName 1 45196 253 M S U S1
showName 1 45449 1797 M S U S1
showName 1 58462 1594 M S U S1
showName 1 60056 1489 M S U S1
showName 1 61545 603 M S U S1
showName 1 68337 1246 M S U S1
showName 1 69583 976 M S U S1
showName 1 70559 675 M S U S1
```

Slika 2.1.2: Sadržaj showName.seg

Proučimo redak: showName 1 2501 964 M S U S0

Polje 1: showName ime datoteke

Polje 2: 1 broj kanala (postavili smo na mono,tj. jedan)

Polje 3: 2501 početak govornikovog segmenta (u milisekundama)

Polje 4: 964 trajanje govornikovog segmenta (u milisekundama)

Polje 5: M govornikov spol (U=nepoznato, F=žensko, M=muško)

Polje 6: S vrsta snimke (T=telefon, S=studio)

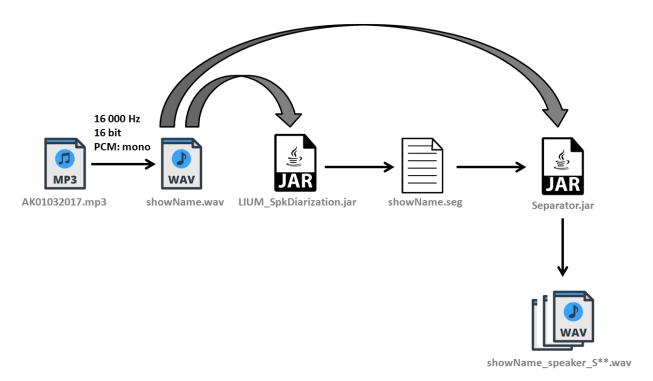
Polje7: U vrsta okruženja (glazba, samo govor, ...)

Polje 8: SØ ID_govornika

Ako želimo izdvojiti svakog govornika u zasebnu .wav datoteku tada možemo iskoristit *Separator.jar* koji na ulaz prima dvije datoteke: *showName.wav* (točno tog imena) i rezultat analize *showName.seg*

Pokretanje Separatora se izvodi lijevim dvoklikom miša, a kao rezultat će nam se generirati konačan broj .wav datoteka.

Na slika 2.1.3 prikazan je dijagram svega spomenutoga zajedno sa prikazanim vezama između datoteka.



Slika 2.1.3: Međuovisnosti datoteka

LIUM vremenski odnosi

Mjerenja su provedena na računalu sa sljedećim parametrima:

Processor: AMD FX-6300, Six-Core, 3.50GHz

RAM: 8GB

• Operating System: Windows 7 Ultimate (64-bit)

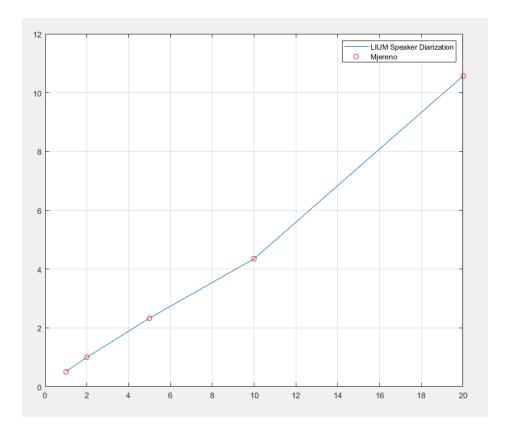
Dobiveni rezultati:

Vrijeme trajanja	Veličina audio	Vrijeme potrebno
audio zapisa	zapisa	za segmentaciju
1 min	1.83 MB	00:00:31
2 min	3.66 MB	00:01:00
5 min	9.15 MB	00:02:20
10 min	18.3 MB	00:04:21
20 min	36.6 MB	00:10:34

Verzija:

Datum: 16.01.2019

1.5



Slika 2.1.4: Graf vremenskih odnosa

2.2 WEB STRANICA

U ovom poglavlju opisat ćemo način pokretanja web stranice u lokalnom načinu rada pomoću alata MAMP (iako preporučamo alat XAMPP koji je složeniji i ima više mogućnosti. Načini konfiguriranja su slični). Prvi korak jest dohvatiti i instalirati MAMP alat:

Verzija:

Datum: 16.01.2019

1.5

https://www.mamp.info/en/

Nakon uspješne instalacije treba kopirati mapu sa izvornim kôdom web stranice na sljedeću lokaciju:

C:\MAMP\htdocs

Možemo pokrenuti MAMP alat (Slika 2.2.1)



Slika 2.2.1: Alat MAMP (pokrenut)

Pritiskom na opciju "Open WebStart page" otvara se poveznica: http://localhost/MAMP/gdje možemo kreirati MySQL bazu podatak potrebnu za login. U alatnoj traci otvorene poveznice odaberimo Tools → phpMyAdmin (Slika 2.2.2)

FER 2 - Projekt ©FER, 2019 Stranica 7 od 16





Slika 2.2.2: phpMyAdmin alat

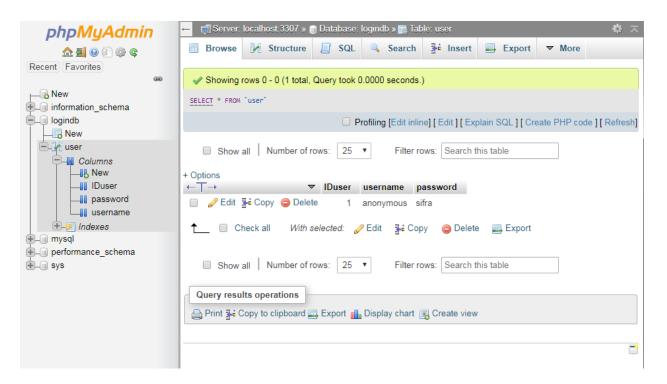
Pritiskom na "phpMyAdmin" otvara nam se pomoćni ugrađeni alat na poveznici: http://localhost/phpMyAdmin/?lang=en

Potrebno je kreirati novu bazu podatak naziva **logindb**, sa relacijom (tablicom) naziva **user** u kojoj su atributi (stupci tablice) naziva **IDuser**, **username**, **password** sljedećih tipova:

```
CREATE TABLE `logindb`.`user` (
`IDuser` SERIAL NOT NULL ,
`username` TEXT NOT NULL ,
`password` TEXT NOT NULL ,
PRIMARY KEY (`IDuser`))
ENGINE = InnoDB;
```

Gore navedeni kôd nije potrebno ručno unositi već se može koristiti korisničko sučelje

U kreiranu bazu podatak unesite korisnika/korisnike (Slika 2.2.3)



Slika 2.2.3: Ispravno konfigurirana baza podataka

Kako bi se kreirana baza podatak uspješno povezala sa našom web stranicom potrebno je provjeriti parametre u datoteci C:\MAMP\htdocs\webStranica\configDB.php sa podatcima sa Slika 2.2.2 (obratiti pažnju na parametar Port)

Ukoliko ste sve ispravno podesili web stranica je dostupna na poveznici:

http://localhost/webStranica/login.php

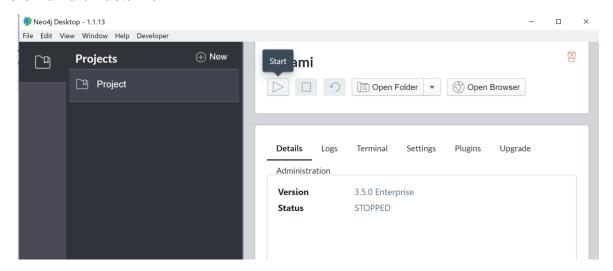
2.3 BIG DATA

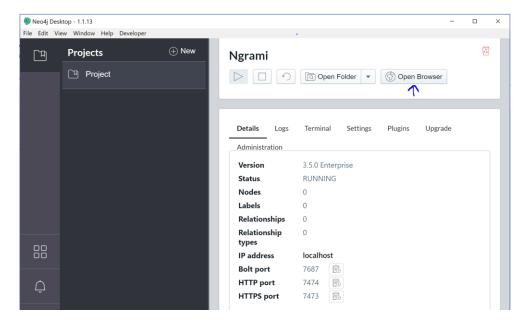
U ovom poglavlju opisat ćemo implementaciju graf baze n-grama. Prvi korak je dohvaćanje Neo4j platforme koja je dostupna na:

https://neo4j.com/download/

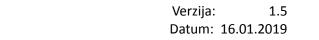
Nakon uspješne instalacije, stvaramo novi projekt klikom na tipku "New" i dodajemo graf klikom na "Add graph" pri čemu mu dajemo ime i postavljamo lozinku.

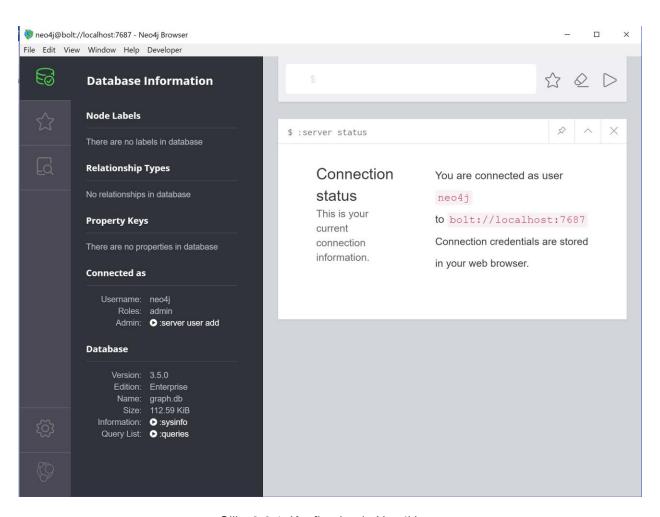
Nakon toga potrebno je pokrenuti ga i otvoriti u Neo4j pregledniku kako je prikazano na slikama u nastavku.





FER 2 - Projekt ©FER, 2019 Stranica 10 od 16





Slike 2.3.1: Konfiguriranje Neo4j baze

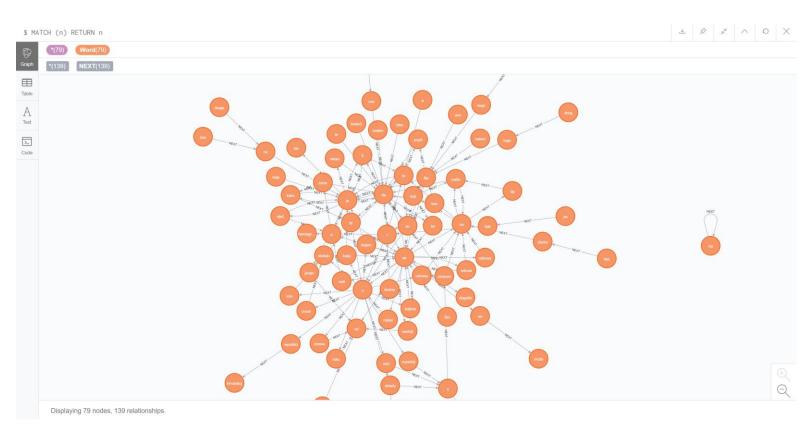
Prije pokretanja samog programa, potrebno je u konfiguraciju navesti ispravnu putanju grafa s odgovarajućim korisničkim imenom i lozinkom:

```
public static final String URL = System.getenv("NEO4J_URL") != null ?
System.getenv("NEO4J_URL") : "bolt://neo4j:password@localhost:7687"
```

Nakon toga pokreće se program implementiran Spring Data Neo4j integracijom. Pokretanjem odgovarajuće klase, program će ili spremati nove podatke u bazu ili pretraživati po postojećima.

Spremanje u bazu dug je i spor proces, ovisno o veličini datoteke i broju trigrama koje želimo spremiti u bazu. U nastavku je kratki primjer:

Pripremili smo kratak .txt dokument u kojem se nalazi 100 trigrama i stvorili novi prazan graf koji tada punimo trigramima iz pripremljenog dokumenta. Spremanje je trajalo 42.423 sekunde, pri čemu je stvoreno 79 čvorova s riječima i 139 veza među njima. Na slici je dobiveni graf:



Slika 2.3.2: Graf veza trigrama

Osim ovog ilustrativnog, proveli smo još dva umetanja u bazu s malo većim dokumentima.

Rbr.	Broj trigrama	Stvoreno čvorova	Stvoreno veza	Trajanje
1.	22388	4845	16559	1 sat
2.	162874	15513	93466	11 sati i 30 minuta

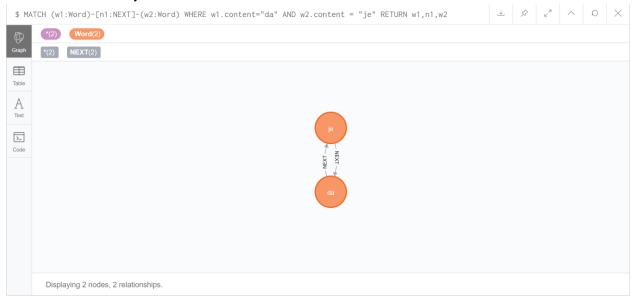
FER 2 - Projekt ©FER, 2019 Stranica 12 od 16

Pretraživanje može biti direktnim upitom na bazu ili programski.

Npr. sljedeći put vraća dva objekta tima Word povezane relacijom NEXT tako da je sadržaj prve "da", a sadržaj druge "je"

```
MATCH (w1:Word)-[n1:NEXT]-(w2:Word)
WHERE w1.content="da" AND w2.content = "je"
RETURN w1,n1,w2
```

Dobiveni rezultat je:



Slike 2.3.3: Pretraživanje baze

Programski na ulaz dajemo neki skup riječi. Za svaku se riječ provjerava je li povezana sa susjednima u slijedu u kojem je poslano na ulaz. Program staje ako neka od riječi ili veza ne postoji ili ako je došao do kraja, te ispisuje dotad pronađeni niz.

Pretraživanje je provedeno na nekoliko primjera čiji su rezultati u tablici u nastavku:

Verzija:

Datum: 16.01.2019

1.5

Opis ulaza	Trajanje
trigram kojem su sve riječi u bazi i susjedne dvije su povezane	7.809 sekundi
trigram prva riječ ne postoji u bazi	9.097 sekundi
šestogram kojem su sve riječi u bazi i susjedne dvije su povezane	7.827 sekundi
šestogram četvrta riječ nije povezana s trećom	8.115 sekundi
deveterogram sve postoje i susjedne dvije su povezane	8.972 sekundi
deveterogram sve postoje, šesta i sedma nisu povezane	7.965 sekundi
šesnaesterogram sve postoje i susjedne dvije su povezane	9.548 sekundi
šesnaesterogram, sve postoje i zadnje dvije nisu povezane	9.606 sekundi

Kratak zaključak:

- za očekivati je da kod spremanja u bazu vrijeme raste budući da za je za svaku riječ koja ne postoji potrebno proći kroz čitavu bazu, te za svaku vezu proći kroz sve veze i stvoriti novu u slučaju da već ne postoji za dani par dvije riječi
- za očekivati je da vrijeme pretraživanja raste povećanjem broja čvorova i veza u bazi, te povećanjem riječi na ulazu

FER 2 - Projekt ©FER, 2019 Stranica 14 od 16

3. LITERATURA

Mickael Rouvier, Gregor Dupuy, Paul Gay, Elie Khoury, Teva Merlin, Sylvain Meignier. An Open-source State-of-the-art Toolbox for Broadcast News Diarization. LUNAM Universite, 2013.

Verzija:

Datum: 16.01.2019

1.5

LIUM Speaker Diarization Wiki, 26.08.2013 http://www-lium.univ-lemans.fr/diarization/doku.php/welcome, 11.01.2019

Neo4j, Using Neo4j from Spring, 2019, $\underline{\text{https://neo4j.com/developer/spring-data-neo4j/}}, 15.01.2019.$

Neo4j, Cypher Basics I, 2019, https://neo4j.com/developer/cypher-query-language/, 15.01.2019.

FER 2 - Projekt ©FER, 2019 Stranica 15 od 16