

Automatsko prepoznavanje govora

Vladimir Vuksanović, Aleksa Kojadinović, Lazar Čeliković

12. decembar 2021.

Sadržaj

1	Prvi recenzent — ocena: 5	2
1.1	O čemu rad govori?	2
1.2	Krupne primedbe i sugestije	2
1.3	Sitne primedbe	2
1.4	Provera sadržajnosti i forme seminarskog rada	4
1.5	Ocenite sebe	5
2	Drugi recenzent — ocena: 3	6
2.1	O čemu rad govori?	6
2.2	Krupne primedbe i sugestije	6
2.3	Sitne primedbe	6
2.4	Provera sadržajnosti i forme seminarskog rada	6
2.5	Ocenite sebe	7
3	Dodatne izmene	8

Glava 1

Prvi recenzent — ocena: 5

1.1 O čemu rad govori?

Seminarski rad *Automatsko prepoznavanje govora* dobro uvodi u tematiku rada sa jasnom definicijom, kratkom istorijom pomoću koje objašnjava poboljšanje performansi sistema za automatsko prepoznavanje govora (poboljšanje dolazi sa korišćenjem dubokih neuronskih mreža). Predstavlja određene probleme sa kojima se susrećemo pri kreiranju modela sistema za prepoznavanje govora, detaljno i jasno ih objašnjava i daje potencijalna rešenja. Zatim, predstavlja dva konkretna modela: statistički model (model koji je ranije bio dominantan) i end-to-end model (model koji je zasnovan na rekurentnim neuronskim mrežama). Postupno, bez previše detalja i na jasan način, objašnjava način na koji rade, prikazuje funkcije koje treba da maksimizuju, kao i razlike među datim modelima. Na samo kraju, prikazuje metriku koja se koristi za evaluaciju prikazanih modela radi procene njihovog kvaliteta.

1.2 Krupne primedbe i sugestije

U trećem poglavlju smatram da bi deo o akustičkom modelu mogao biti malo detaljnije i jasnije objašnjen i da ne bi bilo loše da se doda jedan konkretan primer ako bi to omogućilo bolje razumevanje ovog dela. [Poglavlje je na više mesta prošireno detaljnijim objašnjenjima i dodata je slika sa primerom strukture jednog modela reči.](#)

U četvrtom poglavlju u delu 4.2 nije najjasnije šta tačno predstavlja stanje s_i . [Dodato je kratko objašnjenje stanja kao i prateća formula za izračunavanje istog.](#)

1.3 Sitne primedbe

Štamparske i stilske greške:

- Sažetak:

- ... približi čitaoca zadatku ... – > ... približi čitaocu zadatak ...
[Izmenjeno u skladu sa sugestijom.](#)

- Poglavlje 1:

- ... povećalo njihove performanse. – > ... poboljšalo njihove performanse.
Izmenjeno u skladu sa sugestijom.
- ... neki izazovi na koje se nailazi ... – > ... pojedini izazovi sa kojima se susrećemo ...
Prepravljeno na “najveći izazovi koji se javljaju prilikom modelovanja”. Sugestija nije upotrebljena zato što se u ostatku rada ne korsiti prvo lice.
- ... dva najpopularnija modela: statistički i end-to-end – > ... dva najpopularnija modela: statističkog i end-to-end modela...
Izmenjeno u skladu sa sugestijom.

- Poglavlje 3:

- ... mrežama opisanim u poglavlju 4 ovaj model ... – > ... mrežama (opisanim u poglavlju 4) ovaj model ...
Izmenjeno u skladu sa sugestijom.
- Sledeće rečenice su loše konstruisane, popraviti ih:
 - * Ideja je umesto modeliranja $P(W|X)$ što je teško, odvojeno modelirati verovatnoće iz prethodne formule jer za to postoje bolje tehnike.
Prepravljeno na “Direktno modeliranje veličine $P(W|X)$ je teško. Umesto toga se odvojeno modeliraju verovatnoće iz prethodne formule jer za to postoje bolje tehnike.”
 - * ... signal prvo obrađuje tako da ostanu samo ključne karakteristike i smanji šum i veličina reprezentacije ...
Prepravljeno na “signal prvo obrađuje tako da se izdvoje ključne karakteristike, smanji šum i komprimuje reprezentacija”
- Istrenirani model posle može ... – > Istrenirani model može ...
Izmenjeno u skladu sa sugestijom.
- ... ali nemaju sve semantičko značenje... – > ... ali nemaju isto semantičko značenje ... ?
Nije izmenjeno. Nije poenta da rečenice imaju isto značenje nego da svaka od njih ima neki semantički smisao.
- Kako $P(W)$ ne zavisi od zvučnog signala, može se odvojeno trenirati na samo tekstualnom skupu podataka kojih postoji dosta više i imaju veći broj primera od skupova sa transkribovanim snimcima. – > Kako $P(W)$ ne zavisi od zvučnog signala, može se odvojeno trenirati na samo tekstualnom skupu podataka. Takvih skupova postoji dosta i imaju veći broj primera od skupova sa transkribovanim snimcima.
Izmenjeno u skladu sa sugestijom.

- Poglavlje 4:

- ... end-to-end model je po strukturi prostiji. – > ... end-to-end model ima jednostavniju strukturu.
Izmenjeno u skladu sa sugestijom.

- To donosi velike prednosti, naime: više nije potreban ekspert za jezik nego se sve uči iz podataka, treniranje postaje ... – > To donosi velike prednosti. Naime, više nije potreban ekspert za jezik nego se sve uči iz podataka. Treniranje postaje ...
[Izmenjeno u skladu sa sugestijom.](#)
- Ovo je moguće na više načina. – > Ovo je moguće uraditi na više načina.
[Izmenjeno u skladu sa sugestijom.](#)
- Skušalac ... – > Slušalac ...
[Popravljeno.](#)

1.4 Provera sadržajnosti i forme seminarskog rada

1. Da li rad dobro odgovara na zadatu temu?
Da.
2. Da li je nešto važno propušteno?
Nije.
3. Da li ima suštinskih grešaka i propusta?
Nema.
4. Da li je naslov rada dobro izabran?
Jeste.
5. Da li sažetak sadrži prave podatke o radu?
Da.
6. Da li je rad lak-težak za čitanje?
Skoro ceo rad je lak za čitanje, osim pojedinih delova koji su navedeni u primedbama.
7. Da li je za razumevanje teksta potrebno predznanje i u kolikoj meri?
Rad je lakše razumeti ako ste upoznati sa uslovnom verovatnoćom i mašinskim učenjem.
8. Da li je u radu navedena odgovarajuća literatura?
Jeste.
9. Da li su u radu reference korektno navedene?
Jesu.
10. Da li je struktura rada adekvatna?
Jeste.
11. Da li rad sadrži sve elemente propisane uslovom seminarskog rada (slike, tabele, broj strana...)?
Da.
12. Da li su slike i tabele funkcionalne i adekvatne?
Jesu.

1.5 Ocenite sebe

d) malo upućeni

Glava 2

Drugi recenzent — ocena: 3

2.1 O čemu rad govori?

Може се рећи да рад представља увод у аутоматско препознавање говора кроз основне појмове, најзначајније моделе и проблеме који се могу јавити.

2.2 Krupne primedbe i sugestije

Рад је лепо урађен, повезано и комплетно као једна целина, поменути су и објашњени најзначајнији модели за аутоматско препознавање говора, нема неких крупних примедби, евентуално додати можда још слика. [Dodata je slika sa primerom skrivenog Markovljevog modela u okviru glave o akustičkom modelu.](#)

2.3 Sitne primedbe

Дешава се да речи понегде "искачу" из маргина текста, тј. нису добро поравњане. [Ovaj problem je rešen kompajliranjem pomoću drugog programa.](#)

2.4 Provera sadržajnosti i forme seminarskog rada

1. Da li rad dobro odgovara na zadatu temu?
Да, рад је комплетан одговор на задату тему.
2. Da li je nešto važno propušteno?
Не, покривени су најзначајнији појмови.
3. Da li ima suštinskih grešaka i propusta?
Не, рад суштински делује добро.
4. Da li je naslov rada dobro izabran?
Наслов је адекватно изабран.

5. Da li sažetak sadrži prave podatke o radu?
Да, садржи исправне податке о раду.
6. Da li je rad lak-težak za čitanje?
Тема је занимљива, рад није тежак за читање.
7. Da li je za razumevanje teksta potrebno predznanje i u kolikoј meri?
Неко основно предзнање је сасвим довољно.
8. Da li je u radu navedena odgovarajuća literatura?
Да, сва литература је уредно наведена.
9. Da li su u radu reference korektno navedene?
Да, коректно су наведене.
10. Da li je struktura rada adekvatna?
Структура рада је адекватна.
11. Da li rad sadrži sve elemente propisane uslovom seminarskog rada (slike, tabele, broj strana...)?
Садржи све основне ствари потребне за један семинарски рад.
12. Da li su slike i tabele funkcionalne i adekvatne?
Да, уклапају се у оно о чему се прича.

2.5 Ocenite sebe

Средње упућен, углавном преко Apple Siri и мало преко неуронских мрежа.

Glava 3

Dodatne izmene

Sklonjena je boja sa slike statističkog modela radi bolje preglednosti.