

**Univerzitet Crne Gore**

**Filološki fakultet**

**Vijeću filološkog fakulteta**

Odlukom Vijeća Filološkog fakulteta broj: 01-2545/1 od 10. 11. 2023. imenovana je Komisija za ocjenu master rada pod nazivom „Lingvistička analiza mašinskog prevoda“ kandidatkinje Sare Jovanović (broj indeksa 7/22) Studijski program za prevodilaštvo, u sljedećem sastavu: prof. dr Igor Ivanović, mentor, prof. dr Marijana Cerović, članica i prof. dr Igor Lakić, član.

Komisija Vijeću podnosi izvještaj o ocjeni master rada.

#### **Izvještaj o ocjeni master rada**

Naziv rada kandidatkinje Sare Jovanović je „Lingvistička analiza mašinskog prevoda“. Kandidatkinja se u ovom radu posvetila lingvističkoj analizi mašinskih prevodilačkih sistema, sa fokusom na semantičke, sintaksičke i pragmatičke aspekte mašinskog prevođenja. Mašinski prevod postaje sve značajniji alat u savremenoj međujezičkoj komunikaciji, što čini ovu temu posebno aktuelnom i važnom za istraživanje.

Cilj ovog rada bio je da se kroz analizu prikupljenih lingvističkih podataka sagledaju prednosti i nedostaci mašinskog prevođenja, te da se istraže mogućnosti za unapređenje ovakvih sistema. Kandidatkinja je obradila korpus koji se sastoji od tekstova iz različitih oblasti, uključujući politiku, pravo, zabavu i sport, te analizirala kvalitet mašinskog prevođenja u poređenju sa ljudskim prevodom.

## **Struktura rada i analiza poglavlja**

Master rad Sare Jovanović podijeljen je na nekoliko ključnih poglavlja, od kojih svako obrađuje određeni aspekt lingvističke analize mašinskog prevoda. U nastavku se detaljno elaboriraju pojedinačna poglavlja rada. U uvodnom poglavlju kandidatkinja predstavlja osnovne informacije o mašinskom prevođenju i njegovom značaju u savremenom društvu. Uvod daje pregled dosadašnjih istraživanja na polju mašinskog prevođenja, sa posebnim naglaskom na razvoj tehnologije i značaj lingvističke analize u ovom kontekstu. Kandidatkinja ističe razloge zbog kojih je ova tema odabrana za istraživanje, te postavlja glavna istraživačka pitanja i hipoteze rada. Osnovni ciljevi rada su definirani, a uključuju analizu kvaliteta mašinskog prevođenja i identifikaciju ključnih izazova i mogućnosti za unapređenje. Kandidatkinja takođe objašnjava razliku između različitih vrsta mašinskog prevođenja, uključujući statističko, pravilo-zasnovano i neuronsko prevođenje, te daje pregled evolucije ovih metoda kroz posljednjih nekoliko decenija. Ovaj uvod služi kao temeljan pregled razvoja tehnologije koji omogućava bolji kontekst za analizu u kasnijim poglavljima. Naredno poglavlje rada posvećeno je teorijskom okviru, koji obuhvata ključne pojmove i pristupe lingvističkoj analizi mašinskog prevođenja. Kandidatkinja daje pregled osnovnih teorijskih koncepata, uključujući semantiku, pragmatiku, morfologiju i sintaksu, te objašnjava kako se ovi koncepti primjenjuju u analizi mašinskog prevoda. Posebna pažnja posvećena je razvoju tehnologija mašinskog prevođenja, kao i različitim pristupima u njihovom razvoju, uključujući statističke i neuronske mreže. Kandidatkinja takođe razmatra ključne teorijske aspekte kvaliteta prevoda, kao što su adekvatnost i tačnost, te objašnjava kako se ove kategorije mogu koristiti za procjenu uspješnosti mašinskog prevoda. Pored toga, uključena je diskusija o tome kako lingvistički faktori poput polisemije, konteksta i kolokacija utiču na kvalitet mašinskog prevoda, te kako se ovi faktori mogu modelirati u okviru savremenih prevodilačkih sistema.

Poglavlje o metodologiji daje detaljan opis istraživačkih metoda koje su korišćene u radu. Kandidatkinja se odlučila za kombinaciju kvalitativne i kvantitativne analize kako bi dobila što potpuniji uvid u lingvističke aspekte mašinskog prevođenja. Metode koje su korišćene uključuju korpusnu analizu, diskursnu analizu i statističke procjene pomoću BLEU algoritma. Detaljno su opisani koraci prikupljanja korpusa, uključujući kriterijume za

odabir tekstova, njihovu žanrovsku raznovrsnost i jezičke specifičnosti. Kandidatkinja je koristila različite korpuse koji uključuju pravne, tehničke, publicističke i književne tekstove, kako bi obezbijedila reprezentativan uzorak za analizu. BLEU algoritam je korišćen za kvantitativnu procjenu kvaliteta prevoda, a rezultati su dopunjeni kvalitativnom diskusijom o uočenim lingvističkim problemima. U sljedećem poglavlju kandidatkinja se bavi semantičkom analizom mašinskih prevoda. Semantička analiza obuhvata proučavanje tačnosti i preciznosti u prenošenju značenja iz izvornog teksta u ciljni tekst. Kandidatkinja ističe da se u mašinskom prevođenju često javljaju problemi sa višeznačnim riječima, idiomima i kulturnim referencama, što može dovesti do nepreciznosti u prevodu. Osim navedenog, kandidatkinja analizira i konkretne primjere gdje je mašinski prevod izostavio ključne informacije ili proizveo semantički neprihvatljive ekvivalente. Kandidatkinja je ukazala na slučajeve gdje su kulturne reference bile pogrešno prevedene ili potpuno izostavljene, što ukazuje na potrebu za većom kontekstualnom svjesnošću u razvoju algoritama. Pored analize postojećih problema, data su rješenja u vidu integracije naprednijih semantičkih modela koji mogu bolje tumačiti kontekstualne faktore. Naredno poglavlje rada posvećeno je morfološkoj i sintaksičkoj analizi. Kandidatkinja objašnjava kako su mašinski prevodilački sistemi često ograničeni u prepoznavanju složenih morfoloških oblika i sintaksičkih struktura, što rezultira gramatičkim greškama i neadekvatnim rečenicama. Analizirani su primjeri u kojima su mašinski prevodi sadržali greške u upotrebi zamjenica, slaganja subjekta i predikata, te upotrebi vremena. Kandidatkinja je prepoznala kako različite vrste složenih rečenica (npr. zavisne rečenice, uslovne konstrukcije) predstavljaju izazov za mašinske sisteme, što rezultira pogrešnim redoslijedom riječi ili izostavljanjem ključnih elemenata. Takođe su analizirani problemi sa morfološkom složenošću jezika poput srpskog, gdje fleksija često dovodi do nepravilnog slaganja u mašinskom prevodu. Kandidatkinja je predložila unapređenja u vidu specifičnih pravila koja bi se mogla dodati u algoritme za bolje prepoznavanje morfoloških struktura. U sljedećem poglavlju, kandidatkinja se bavi pragmatičkim aspektima mašinskog prevođenja. Analizirano je kako mašinski prevodilački sistemi obrađuju kontekstualne informacije, te kako se nosi sa pragmatičkim izazovima kao što su kulturne reference, humor i idiomatski izrazi. Kandidatkinja navodi primjere situacija u kojima mašinski prevodi nijesu uspjeli da prenesu pragmatičko značenje iz izvornog teksta, te daje preporuke za unapređenje

algoritama kako bi se bolje razumio kontekst i prilagodila poruka ciljnom jeziku. Pored navedenog, analizirana je uloga konverzacionih implikatura i kako njihovo zanemarivanje u mašinskom prevodu može dovesti do nesporazuma. Kandidatkinja ističe da je prepoznavanje govornog konteksta od ključnog značaja za razumijevanje implicitnih značenja, posebno u formalnim i neformalnim situacijama. Predloženi su načini za unapređenje prepoznavanja pragmatičkih elemenata kroz kombinaciju lingvističkih i neuronskih modela koji mogu obraditi širi kontekst izvan pojedinačnih rečenica. Naredno poglavlje posvećeno je korpusnoj analizi i računarskoj analizi podataka. Kandidatkinja je koristila softver AntConc za analizu korpusa koji se sastoji od tekstova iz različitih oblasti. Ova analiza omogućila je identifikaciju najčešće korišćenih fraza i kolokacija u mašinskim prevodima. Takođe su prikazani rezultati statističke analize pomoću BLEU algoritma, koji pokazuje stepen preciznosti i kvaliteta mašinskog prevoda u poređenju sa ljudskim prevodom. Kandidatkinja je na osnovu rezultata identifikovala oblasti u kojima su potrebna dalja unapređenja kako bi se poboljšala efikasnost i tačnost mašinskog prevođenja. Analiza korpusa obuhvatila je i identifikaciju obrazaca grešaka, kao što su ponavljanja i nepotrebne konstrukcije, te su date preporuke za njihovo uklanjanje. Kandidatkinja je takođe istakla važnost proširenja korpusa kako bi obuhvatio više žanrova i stilova, što bi doprinijelo boljoj obuci mašinskih modela.

Šesto i sedmo poglavlje rada, pod nazivom „Diskusija“ i „Zaključak“, sažima glavne nalaze istraživanja i daje preporuke za buduća istraživanja. Kandidatkinja ističe da, iako mašinsko prevođenje može značajno olakšati rad prevodilaca, i dalje postoje značajne slabosti koje zahtijevaju unapređenje. Kandidatkinja naglašava potrebu za razvojem novih tehnika obuke mašinskih sistema koji bi bolje razumjeli kontekst i prilagodili prevod specifičnim situacijama. Takođe su predloženi načini integracije povratnih informacija od korisnika kako bi se poboljšao kvalitet prevoda kroz iterativni proces učenja. Diskusija takođe obuhvata etičke aspekte mašinskog prevođenja, uključujući moguće rizike od automatizacije i gubitka radnih mjesta za profesionalne prevodioce. Zaključeno je da su ljudska intervencija i nadzor i dalje ključni za postizanje visokog kvaliteta prevoda, posebno u situacijama koje zahtijevaju visoku preciznost i osjetljivost na kontekst.

## Zaključci i preporuke Komisije

Komisija ocjenjuje da je kandidatkinja ispunila sve zahtjeve postavljene pred master rad, te da je kroz detaljnu lingvističku analizu mašinskog prevođenja dala značajan doprinos razvoju lingvističkih istraživanja u oblasti prevođenja. Kandidatkinja je pokazala sposobnost da na temeljit način analizira složene jezičke fenomene i da izvuče relevantne zaključke koji mogu imati praktičnu primjenu u razvoju mašinskih prevodilačkih sistema. Imajući u vidu sve navedeno u ovom izvještaju, komisija predlaže Vijeću Filološkog fakulteta da prihvati izvještaj i omogući kandidatkinji odbranu master rada.

Komisija

Prof. dr Igor Lakić, član

*Igor Lakić*

Prof. dr Marijana Cerović, članica

*M. Cerović*

Prof. dr Igor Ivanović, mentor

*Igor Ivanović*

| UNIVERZITET CRNE GORE<br>FILOLOŠKI FAKULTET |       |        |            |
|---------------------------------------------|-------|--------|------------|
| 22.11.2024.                                 |       |        |            |
| Prijemno                                    | Broj  | Prilog | Vrijednost |
| 03                                          | 551/7 |        |            |