

ONLINE NOTES: sistem za razpoznavo govora in strojno prevajanje v realnem času na ravni univerzitetnih predavanj

Tjaša Šoltes, Jan Vasiljević, Marko Bajec

Fakulteta za računalništvo in informatiko Univerze v Ljubljani, Večna pot 113, Ljubljana

tjasa.soltes@fri.uni-lj.si, jv1721@student.uni-lj.si, marko.bajec@fri.uni-lj.si

Izvleček

Online Notes je sistem za avtomatsko razpoznavo govora in prevajanje v realnem času na ravni univerzitetnih predavanj, ki ga razvijamo na Univerzi v Ljubljani. Sistem je primarno namenjen študentom Univerze v Ljubljani, ki ne razumejo ali slabše razumejo slovensko, ter študentom s senzornimi oviranostmi. V sklopu projekta smo razvili dva modela za razpoznavo govora in spletno aplikacijo s štirimi portali. Do sedaj je bilo izvedenih 29 pilotnih testiranj sistema, med drugim eno semestrsko, v okviru katerega smo spremljali dve študentki na izmenjavi programa Erasmus+, ki sta tekom celotnega semestra obiskovali predmet Prostorska statistika na Fakulteti za gradbeništvo in geodezijo UL v slovenščini z uporabo sistema Online Notes. Poleg spremljanja v živo sta imeli kot učni material na voljo tudi avtomatsko generirane podnapise v angleščini, ki so nastali na podlagi popravljenih slovenskih transkriptov. Študentki sta uspešno opravili vse študijske obveznosti. Izpostavili sta, da je sistem zelo uporaben za študente, ki ne razumejo slovensko, saj omogoča sledenje predavanjem, poleg tega pa sta izpostavili tudi določene pomanjkljivosti, kot je kognitivna zahtevnost pri spremljanju sistema, predavatelja in table hkrati in mestoma slabša kvaliteta prevodov pri predavanjih v živo.

Ključne besede: avtomatska razpoznavo govora, strojno prevajanje, univerzitetna predavanja

Online notes: a real-time speech recognition and machine translation system in university lectures

Abstract

Online Notes is a system for automatic speech recognition and real-time translation at the level of university lectures which has been developed at the University of Ljubljana. The system is primarily intended for students of the University of Ljubljana who do not understand Slovenian well or do not understand it at all, as well as students with sensory impairments. Within the project, we have developed two speech recognition models and a web application consisting of four portals. So far, 29 pilots have been carried out, among them a semester pilot, where we followed two exchange students of the Erasmus+ program who, during the entire semester, attended the Spatial Statistics course at the Faculty of Civil Engineering and Geodesy at the University of Ljubljana in Slovenian using the Online Notes system. In addition to the real time use, they also had automatically generated subtitles in English available as learning material, which were created on the basis of corrected Slovenian transcripts. The two students successfully completed all their study obligations. They pointed out that the system is very useful for students who do not understand Slovene, as it makes it possible to follow lectures, and they also pointed out certain shortcomings, such as the cognitive complexity of monitoring the system, the lecturer in the table at the same time, and the poor quality of translations in lectures in alive.

Keywords: automatic speech recognition, machine translation, university lectures

1 UVOD

V letu 2022 je bilo na Univerzi v Ljubljani vpisanih 3.508 tujih študentov od skupno 37.509 (torej 9,4 %),

na študijsko izmenjavo pa jih je prišlo 2.256 [7]. Ker predavanja večinoma potekajo v slovenščini, so študentje, ki slovensko ne razumejo, omejeni le na

peščico predavanj, ki se izvajajo v angleščini. Da bi tovrstnim študentom omogočili oziroma olajšali spremljanje vseh predavanj na Univerzi v Ljubljani, je bil razvit sistem Online Notes. Gre za sistem za avtomatsko razpoznavo govora in prevajanje v realnem času. V prvi vrsti je namenjen študentom, ki ne razumejo ali slabo razumejo slovensko, predvsem so to študentje na študijskih izmenjavah ali redno vpisani študentje, katerih materni jezik ni slovenščina, poleg tega pa je sistem namenjen tudi študentom s senzornimi oviranostmi (npr. za gluhe in naglušne). Razvili smo aplikacijo s štirimi portali ter dva modela razpoznavalnika govora (en namenjen naravoslovnim in drugi namenjen družboslovnim predavanjem).

Od leta 2021 dalje je bilo na osmih različnih fakultetah Univerze v Ljubljani izvedenih 29 pilotnih predavanj z uporabo sistema Online Notes, od tega tudi dve semestrski testiranji [1], eno bomo v prispevku tudi podrobneje predstavili. Pri enem od semestrskih testiranj smo spremljali dve študentki na izmenjavi projekta Erasmus+, ki sta tekom celotnega zimskega semestra študijskega leta 2022/2023 spremljali predavanja v slovenščini z uporabo sistema Online Notes in naknadno z videoposnetki, podnaslovljeni z avtomatsko generiranimi podnapisi v angleščini, pripravljene na podlagi popravljenih slovenskih transkriptov.

Prispevek je strukturiran sledeče: v poglavju 2 na kratko predstavimo sorodne raziskave, v poglavju 3 predstavimo sam sistem in aplikacijo, in sicer arhitekturo, komponento za razpoznavo govora ter uporabniški vmesnik, v četrtem poglavju pa podrobneje predstavimo rezultate semestrskega testiranja sistema.

2 PREGLED LITERATURE

Številne univerze v državah, kjer učni proces ne poteka v angleščini, se soočajo s podobnimi težavami. Tujim študentom namreč niso na voljo predavanja, ki bi potekala v angleščini, ali pa ali pa je takšnih vsebin le omejeno število. Zaradi jezikovnih ovir so takšne univerze posledično za tuje študente manj zanimive.

Sorodni sistemi za razpoznavo govora in prevajanje v realnem času so že bili razviti in preizkušeni tudi v drugih univerzitetnih okoljih. Zelo podoben sistem je denimo Karlsruhe Institute of Technology (KIT) Lecture Translation System [2, 3], ki ponuja transkripcije in strojne prevode nemških predavanj v obliki podnapisov. Pri evalvaciji sistema so poročali, da je večina študentov, ki je uporabljala sistem, oce-

nila sistem kot koristen, posebno samo transkripcijo govora. Med negativnimi vidiki sistema so izpostavili latenco pri prevodih, poleg tega pa tudi kognitivno zahtevnost pri sočasnemu sledenju materialom (preglednicam) in prevodom transkriptov predavanja. Eden od udeležencev raziskave je še poudaril, da je bil sistem uporaben predvsem za študente, ki se učijo nemško.

Shadiev in Huang [5] sta raziskovala uporabo razpoznave govora in strojnega prevajanja na primeru tujih predavateljev, ki v angleščini predavajo študentom, ki jim je angleščina tuji jezik. Raziskovala sta kognitivno obremenitev, pozornost in stres ob poslušanju predavanja v tujem jeziku. Ugotovila sta, da so imeli študentje, ki so predavanja spremljali s strojnimi prevodi v svojem prvem jeziku, najmanjšo kognitivno obremenitev in najvišjo stopnjo zadovoljstva v primerjavi s študenti, ki so predavanje spremljali samo s prepoznavanjem govora v angleščini, in študenti, ki sploh niso imeli podpore govorne tehnologije.

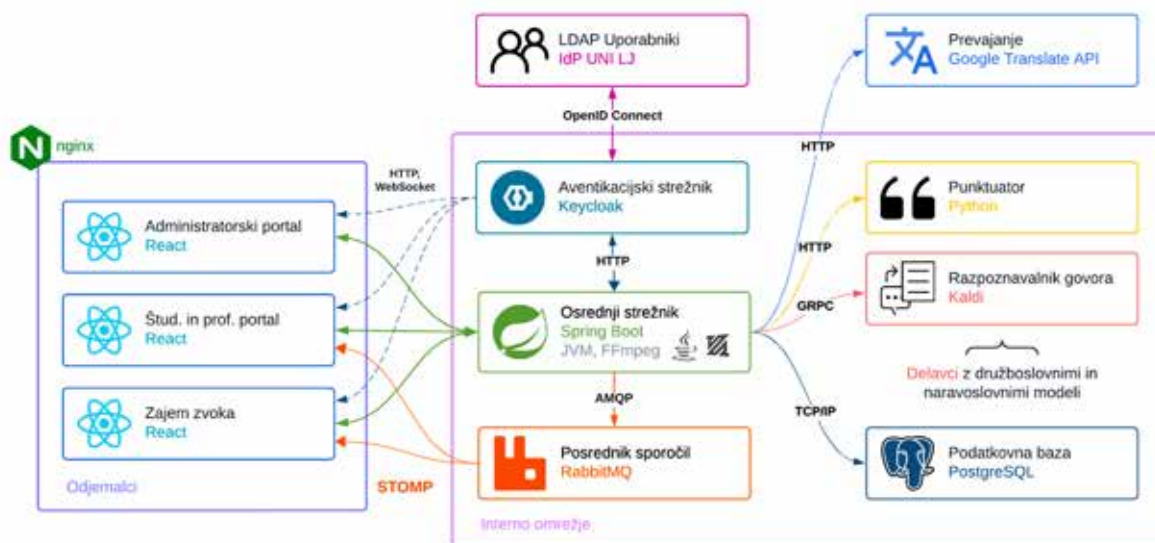
3 APLIKACIJA

Sistem Online Notes je sistem za avtomatsko razpoznavo govora in prevajanje v realnem času. V nadaljevanju predstavimo samo aplikacijo in pa komponente za razpoznavo govora in strojno prevajanje.

3.1 Arhitektura

Sistem je sestavljen iz več komponent, ki jih delimo na dva glavna dela. Čelni del (ang. frontend) sestavljajo tri spletne aplikacije, ki temeljijo na ogrodju React, za njihovo delovanje pa skrbi nginx strežnik. Odjemalske aplikacije komunicirajo z zalednim delom (ang. backend) preko več komunikacijskih kanalov, ki zahtevajo avtentikacijo. Med te kanale spadajo protokoli HTTP, Websockets in STOMP.

Osrčje zalednega dela sistema predstavlja spletni strežnik, implementiran v ogrodju Spring Boot in programskem jeziku Java 21, ki opravlja več nalog. Strežnik odjemalcem ponuja vmesnik, podoben REST-u. Prejete zahteve preverja preko avtentikacijskega strežnika Keycloak, ki pridobi tudi podatke o LDAP uporabnikih Univerze v Ljubljani. Strežnik preko protokola gRPC pošilja avdio zapis, prejet preko WebSocketov, v storitev za prepoznavanje govora. Le-ta glede na attribute predavanja dinamično alocira delo Kaldi delavcu. Prejeti transkripti so nato poslani punktualorju, ki dodaja ločila, in storitvi za prevajanje.



Slika 1: Shema arhitekture sistema

Med pretakanjem predavanja strežnik komunicira tudi s posrednikom sporočil RabbitMQ, ki skrbi za razpošiljanje transkriptov in prevodov odjemalcem. Avdio posnetki so dodatno obdelani z ogrodjem FFmpeg in shranjeni na datotečni sistem. Vsi preostali podatki kot so vsebina predavanj, podatki o predmetih in podobno, se hranijo v podatkovni bazi PostgreSQL.

3.2 Razpoznavo govora in strojno prevajanje

Ena od ključnih komponent sistema je tudi razpoznavnik govora. Za razpoznavo govora smo razvili dva ločena modela, in sicer enega za tehnična in drugega za družboslovna predavanja. Razpoznavnik govora temelji na ogrodju Kaldi [4], ki za usposabljanje in dekodiranje uporablja utežene pretvornike končnega stanja (*WFST*; ang. *weighted finite state transducers*).

Prvi korak v procesu prepoznavanja govora je branje valovne oblike signala in ekstrakcija značilnosti iz zvoka, ki je predstavljen v obliki Melovega spektrograma. Te podatke nato obdela akustični model, ki vrne verjetnostno matriko podbesednih enot skozi čas. Ta verjetnostna matrika se dekodira skupaj z jezikovnim modelom, ki vrne besedilo brez ločil. Besedilo procesira punktuator (ki je implementiran kot ločena storitev), ki v besedilo vstavi ločila.

Trenutno se znotraj sistema Online Notes uporabljata dva ločena modela za razpoznavanje govora, in sicer model za družboslovna predavanja in model za tehnične vede. Prednost ogrodja Kaldi je, da sta

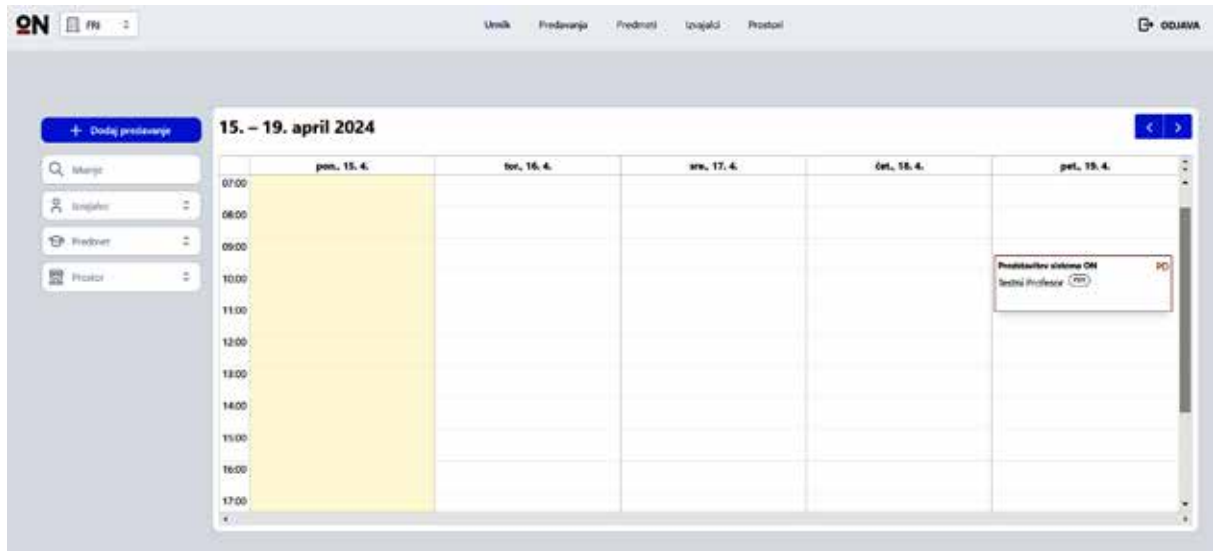
lahko akustični model in jezikovni model neodvisna, zato si oba modela za prepoznavanje govora delita skupen akustični model.

Osnova za jezikovni model je besedilni korpus. Vsebina predavanj je običajno specializirana, zato je nujno, da je jezikovni model prilagojen predavanju. Predavatelje zato prosimo, da nam pošljejo morebitne avdio ali video posnetke svojih predavanj iz preteklih let. Ti posnetki so samodejno transkribirani z uporabo razpoznavnika govora in ročno popravljani. Če gradiva ni na voljo, se za nadgradnjo modela uporabijo druga gradiva, kot so članki, revije, diplomske naloge ipd. Ko je podkorpus predavanja pripravljen, ga dodamo v enega od dveh glavnih korpusov, torej v tehnični ali družboslovni korpus. Ko je besedilni korpus sestavljen, med učenjem pripravimo leksikon besed s pripadajočimi izgovorjavami.

Akustični model je bil učen na približno 200 urah ortografsko transkribiranega besedila, jezikovna modela pa se ves čas posodabljata s transkripti preteklih predavanj. Trenutna učna množica je sestavljena iz okvirno 2 milijonov povedi in 1,3 milijona besed.

Razpoznavnik govora vrača razmeroma kratke besedilne segmente, ki se nato prevedejo v angleščino. Vsak besedilni segment se prevede samostojno, da se minimizira zamuda pri prevodu. Prevod je nato na študentskem (in kasneje na profesorskem) portalu prikazan vzporedno s slovenskim besedilom.

Za prevode se trenutno uporablja zunanji komercialni nevronske strojni prevajalnik.



Slika 2: Urnik na administratorskem portalu

3.3 Uporaba sistema in opis portalov

Sistem je sestavljen iz štirih portalov, in sicer: profesorski portal, študentski portal, portal za zajem zvoka in administratorski portal. Študentski in profesorski portal sta združena v enotno spletno aplikacijo, ki uporabnika glede na njegove pravice avtomatično preusmeri na ustrezní portal.

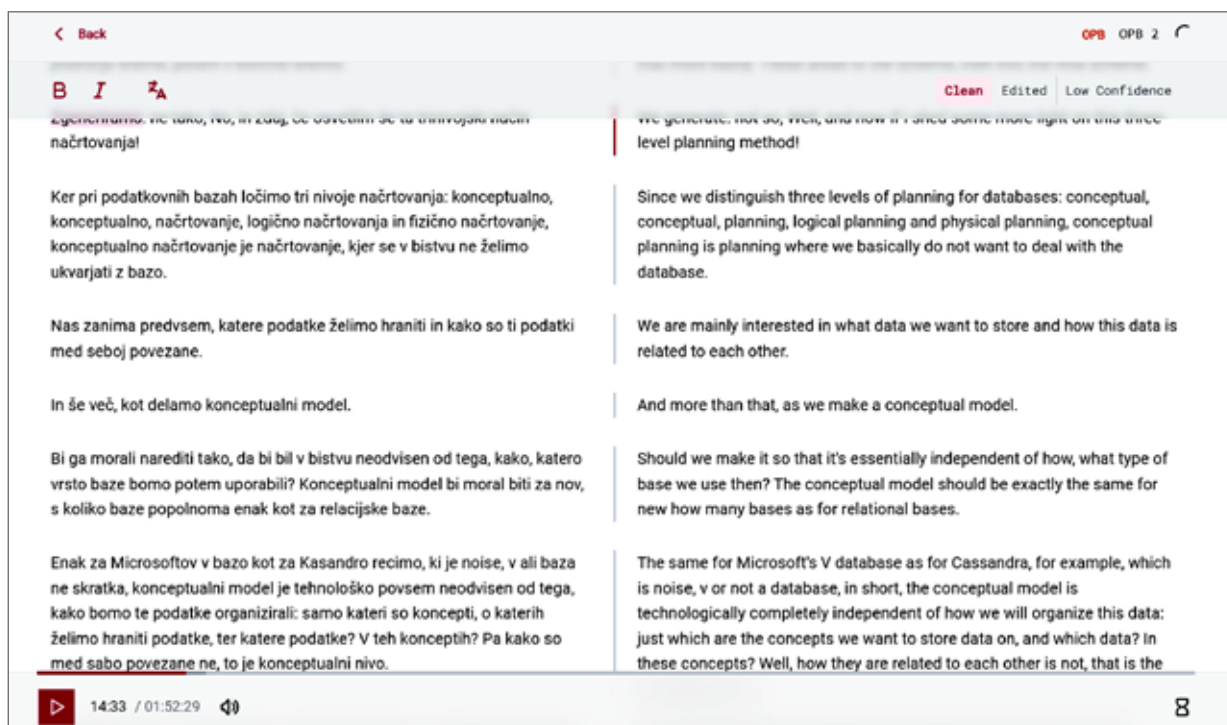
Na administratorskem portalu je možno v sistem vnesti predavatelje, predmete, posamezna predava-

nja in prostore oziroma predavalnice. Pred začetkom uporabe sistema je treba ustvariti posamezno predavanje, torej posamezno sejo, v katero se bo zapisovala vsebina in ki jo bodo študentje lahko spremljali. Ob ustvarjanju seje se določi predmet, predavatelja, prostor, prav tako pa se izbere tudi primerni model razpoznave govora (naravoslovni ali družboslovni).

Ko je seja ustvarjena, se lahko predavatelj s številko PIN vpiše v svoj račun na portalu za zajem zvoka.



Slika 3: Portal za zajem zvoka



Slika 4: Vsebina v profesorskem portalu po končanem predavanju

Pred snemanjem lahko preveri stanje vseh storitev in stanje povezave ter testira razpoznavnik govora, ne da bi se vsebina testiranja prikazovala študentom. Ko predavatelj začne s snemanjem, lahko vidi zadnje transkripte. Predavanje lahko začasno ustavi in kasneje z njim nadaljuje, ali pa ga zaključi. Za zagotavljanje dovoljšne kvalitete zvoka smo sodelujočim v pilotnih predavanjih namestili prenosne mikrofone, saj je kvaliteta opreme po različnih prostorih in članicah UL variirala.

Ko je predavanje aktivno, ga lahko študentje v živo spremljajo na študentskem portalu, kamor se vpišejo s svojimi študentskimi računi digitalne identitete UL. Ob govoru predavatelja se jim prikazujejo transkripti v slovenščini in prevodi v angleščino. Izberejo si lahko, ali naj se jim prikazuje le slovenski transkript, le angleški prevod ali oboje naenkrat, razen na manjših napravah, kot so mobilni telefoni, kjer je zaradi omejitve s prostorom možno predavanje spremljati le v enem jeziku.

Po končanem predavanju lahko predavatelj do vsebine dostopa na profesorskem portalu, kamor se vpiše s svojim računom. Tam si lahko ogleda slovenski transkript in prevod, ter popravi oziroma uredi vsebino predavanja. Predavatelj se lahko za vsako predavanje v sistemu odloči, ali ga želi deliti z na-

ročniki (tj. vsi študentje, ki so se v sistemu naročili oziroma prijavi na predavanje) ali ne. Če se odloči predavanje objaviti, postane vidno tudi na študentskem portalu.

4 SEMESTRSKO TESTIRANJE

V sklopu semestrskega testiranja smo spremljali dve študentki na študijski izmenjavi programa Erasmus+, ki sta v študijskem letu 2022/2023 cel semester spremljali predavanja pri predmetu Prostorska statistika na Fakulteti za gradbeništvo in geodezijo Univerze v Ljubljani pri prof. dr. Goranu Turku. Predavanja pri predmetu so potekala v slovenščini, študentki pa sta spremljali predavanje z uporabo sistema Online Notes s prevodi v angleščino. Poleg tega sta pri predmetu obiskovali tudi vaje, ki so potekale v angleščini.

Ena od študentk (v nadaljevanju Študentka A) je bila vpisana v tretji letnik dodiplomskega študija in prihaja iz Estonije, prav tako je estonsčina njen materni jezik. Svoje razumevanje angleščine je ocenila kot »tekoče« (C1/C2). Druga študentka (v nadaljevanju Študentka B) je bila vpisana v prvi letnik bolonjskega magistrskega študija in prihaja iz Francije. Njen materni jezik je francoščina. Svoje razumevanje angleščine je ocenila kot »dobro« (B1/B2).

Študentki smo prosili, da redno obiskujeta predavanja in jih spremljata z uporabo sistema Online Notes ter nam periodično podajata mnenje, po koncu semestra pa smo z njima opravili tudi končni intervju.

Tekom celotnega semestra sta obe študentki spremljali predavanja v živo z uporabo orodja Online Notes s svojimi študentskimi računi. Po vsakem predavanju so študentje, ki sodelujejo pri projektu, popravili transkript, da je ustrezal temu, kar je povedal predavatelj. Popravljeni transkripti so bili osnova za pripravo avtomatsko generiranih podnapisov v angleščini, profesor je namreč samostojno (brez uporabe sistema) posnel vsa predavanja in nato video-posnetke opremil z angleškimi podnapisi, ki so bili izvoženi iz sistema. Študentki sta torej imeli na voljo tudi avtomatsko generirane angleške podnapise za učenje po končanih predavanjih.

Na tem mestu je pomembno poudariti, da se kvaliteta prevodov v živo v primerjavi s prevodi v podnapisih zelo razlikuje. Za razliko od prevodov v živo, ki so prevodi posameznega končnega besedilnega segmenta v realnem času, gre tu za prevod popravljenega besedila. Vsaka napaka v slovenskem transkriptu se namreč v angleščini še potencira. Druga razlika pa je v tem, da se zaradi zagotavljanja hitre storitve v načinu v živo prevajajo posamezni odseki – segment po segment. Pri generiranju podnapisov po samem predavanju pa lahko ponovno prevedemo celotno vsebino predavanja naenkrat, kar seveda posledično zaradi več konteksta izboljša tudi kvaliteto prevoda.

4.1 Povratne informacije študentk

Študentki sta sistem uporabljali tekom enega semestra. Njuna mnenja smo zbrali ob prvi uporabi, na sredini semestra in ob koncu semestra.

Po prvem predavanju in s tem po prvi uporabi sistema Online Notes smo študentki prosili za povratne informacije v obliki odgovorov na anketni vprašalnik. Zanimala nas je splošna uporabnost sistema in kaj so glavni moteči elementi.

Študentki smo vprašali, ali se jima zdi sistem uporaben, ali bi ga priporočili svojim kolegom in ali bi ga uporabljali tudi za učenje po končanih predavanjih, pri čemer sta obe na vsa tri vprašanja odgovorili pritrdilno.

Poleg tega smo izpostavili 8 potencialnih pomanjkljivosti sistema in ju prosili, da ocenita, kako moteče so. Njuni odgovori so prikazani v Tabeli 1.

Na koncu je študentka A v razdelku splošno mnenje zapisala še, da je sistem zelo uporaben, a intenziven, študentka B pa je omenila, da je bilo težko spremljati sistem in obenem, kaj predavatelj v slovenščini piše po tabli. Poleg tega je omenila tudi, da mestoma zaradi napak v prevodu zelo težko razume, kaj je bilo mišljeno.

Ponovno smo ju prosili za mnenje v začetku decembra, tj. po dveh mesecih aktivne uporabe sistema. Zanimivo je, da sta obe študentki poročali, da se je kvaliteta angleških prevodov izboljšala, čeprav sami nismo zaznali bistvenega odstopanja v kvaliteti transkriptov ali prevodov tekom semestra, prav tako se jezikovni model v tem času ni dopolnjeval z novimi materiali s področja vsebine predmeta.

Študentka A je zapisala, da je spremljala skoraj vsa predavanja in da se ji zdi, da so se prevodi sčasoma izboljšali,¹ medtem ko je študentka B zapisala, da je besedilo bolj razumljivo kot na začetku semestra².

Študentka A je izpostavila slabše delovanje razpoznavalnika (in posledično tudi slabše prevode), kadar predavatelj piše na tablo. Izpostavila je tudi občasne neobičajne prevode, denimo »*God is uniformly distributed.*« ali »*We also go to our funeral.*«

Tabela 1: Ocene motečih elementov v prevodih v živo

| | Ni moteče | Nekoliko moteče | Zelo moteče | Ne morem določiti |
|--|-----------|-----------------|-------------|-------------------|
| Besede manjkajo ali so odveč | | A | B | |
| Napačen besedni red | A, B | | | |
| Napačno kapitalizirane besede | A, B | | | |
| Napačno postavljena ločila | A | B | | |
| Napačno črkovanje | | A | B | |
| Popolnoma napačne besede ali fraze | | A | B | |
| Nerodno oblikovane povedi | A | | B | |
| Zamik pri podajanju angleškega prevoda | | A | B | |

¹ »I have been following almost all the lectures and I do feel that the translation has improved over that time.«

² »I think that the software improved because things seem more understandable now than compare to the beginning of the semester.«

Študentka B je izpostavila težave pri razpoznavi spremenljivk, na primer „*here I have both Anič and B have different values*“, kjer je napačno razpoznana spremenljivka v prevodu posledica napačno prepoznane spremenljivke v slovenščini, gre namreč za spremenljivko *A0*.

Prav tako je izpostavila, da ji je sistem v pomoč, ko profesor ne piše, pač pa na primer razlaga koncepte, več težav pa je pri izračunih.³ Tu gre najverjetneje za več dejavnikov, ki vplivajo na razliko v kvaliteti: v korpusu besedil, ki je vključen v jezikovni model, so enačbe, pa tudi same spremenljivke zastopane v manjšem številu, prav tako se način govora (npr. dolžina premorov) ob sočasnem pisanju lahko spremeni.

Izpostavila je tudi kognitivno kompleksnost sistema, saj hkrati spremljata predavanja in predstavitve v slovenščini v živo, poleg tega pa morata ves čas spremljati še vsebino na svojem prenosniku v angleščini, kar je zahtevno.⁴

Po končanem semestru smo s študentkama opravili tudi končni intervju, kjer nas je zanimala njuna splošna izkušnja, prednosti in slabosti sistema.

Obe sta se strinjali, da jima je na splošno sistem olajšal sledenje vsebini predmeta. Moteči so bili mestoma napačni prevodi ob spremljanju v živo. Ob razlaganju teorije oziroma ob samem govoru (brez eksternih dejavnikov) so bili prevodi po njunem mnenju razumljivi in točni in sta lahko s pomočjo sistema razumeli predavatelja, medtem ko je bilo ob pisanju po tabli več napak. Po njunih izkušnjah je bilo največ težav pri razpoznavi spremenljivk in enačb. Nekoliko ju je motila tudi neobičajna struktura prevodov, kot so krajše povedi in ponavljanje besed, ki pa sicer izvira iz narave nebranega govora – pri nebranem govoru namreč govorniki pogosto ponovijo besede, premori niso le na koncu povedi in podobno. Zanimivo je, da sta se obe strinjali, da so se prevodi tekom semestra izrazito izboljšali, medtem ko naša analiza ne kaže bistvenega odstopanja v kvaliteti transkriptov in prevodov tekom semestra. Zamik pri prevodih ju v nasprotju z našimi pričakovanji ni pretirano motil, študentko B je motil proti koncu semestra, študentki A pa se ni zdel pretirano moteč. V avtomatsko generiranih podnapisih na popravljenih

slovenskih transkriptih so se obema podnapisi zdeli precej bolj pravilni in razumljivi.

Drugi problem, ki ga je izpostavila študentka B, je kognitivna zahtevnost pri uporabi – spremljati je namreč treba predavatelja in tablo v enem jeziku, poleg tega pa še prenosnik s prevodi v drugem jeziku, hkrati pa noben od teh dveh jezikov ni materni, torej je vpet še tretji jezik. Kot alternativo je predlagala, da bi se prevodi generirali v avdio obliki, kjer bi bilo treba manj vsebine spremljati vizualno. Študentke A spremljanje programa in table ni motilo in ji je v določenih primerih olajšalo razumevanje, predvsem ko so bile v prevodu napačno razpoznane spremenljivke.

Glede materialov, ki so bili na voljo (sistem za spremljanje v živo in podnaslovljeni videoposnetki), sta se strinjali, da zadostujejo za samo učenje, da pa so jima pri učenju zelo pomagali podnapisi in bi jima bilo brez tega precej težje. Med samim predavanjem sta občasno dodatne informacije poiskali na spletu, vendar sta se strinjali, da to ni nujno posledica kvalitete in točnosti prevoda, pač pa same zahtevnosti vsebine predmeta.

Na splošno sta bili s sistemom zadovoljni in bi ga priporočili tudi svojim kolegom na domači univerzi.

4.2 Intervju s predavateljem

Po koncu semestra smo opravili tudi intervju s profesorjem. Profesor je bil s sistemom zadovoljen, predvsem so se mu zdeli koristni podnapisi. Svoja predavanja snema že od študijskega leta 2017/2018 naprej, vendar bi priprava podnapisov v angleščini vzela preveč časa.

Kvaliteto podnapisov (na podlagi popravljenih transkriptov) je ocenil kot boljše od pričakovanega. Terminološko je sicer zaznal nekaj pomanjkljivosti, zato si je pripravil seznam besed, ki so bile običajno napačno prevedene, in jih popravil. Po njegovi oceni mu je priprava podnapisov vzela okvirno 10 minut na predavanje. Ocenil je, da bi sicer sam marsikateri podnapis ubesedil drugače, da pa je vsebina razumljiva in da sta se študentki lahko naučili iz videov, da sta tudi izpit naredili brez težav, iz česar je sklepal, da sta iz videov razumeli, kaj želi povedati.

Vprašali smo ga, ali bi uporabljal sistem Online Notes tudi na ostalih predavanjih, na kar je odgovoril:

³ »I would say that it is helpful when the professor is not writing on the board, just talking (explaining one concept for example). But when calculations parts arrive, it is difficult to follow because the translation of mathematic expressions is chaotic. Furthermore, we have to read at the same time the board in slovene and our computer in english so it is a lot to process at once»

⁴ »Furthermore, we have to read at the same time the board in slovene and our computer in english so it is a lot to process at once.«

ril, da bi sistem uporabljal, če bi bili v razredu tuji, saj se mu zdi to dobra rešitev, posebno v kombinaciji s podnaslovljenimi videoposnetki.

Na vprašanje, ali je opazil, da njegov način govora kakor koli vpliva na samo kvaliteto razpoznave, je odgovoril, da je pri spontanem oziroma nebranim govoru način govora drugačen, govor denimo vsebuje več premorov sredi povedi, besede se lahko ponavljajo in podobno, kar seveda vpliva na samo kvaliteto razpoznave in posledično razumljivost prevoda, da pa vendar ni namen predavanja, da je v naprej pripravljeno in brano.

Ta predmet je sicer v celoti v preteklosti že opravil študent na izmenjavi Erasmus+ iz Francije, vendar je predmet opravljal konsultacijsko. Konsenz je namreč, da se izvede ločeno predavanje za tuje študente, če je v skupini vsaj 5 tujih študentov. Zanimalo nas je, ali bi v prihodnje izvedel ločen predmet v primeru, da bi bilo nanj prijavljenih dovolj tujih študentov, na kar je odgovoril, da ne bi, saj sistem ON v kombinaciji s podnapisi omogoča sledenje pouku.

5 DISKUSIJA

V študiji smo spremljali dve tuji študentki, Francozinjo in Estonko, ki sta s pomočjo Online Notes, sistema za avtomatsko razpoznavo govora in strojno prevajanje v realnem času, v zimskem semestru 2022/2023 spremljali predavanje Prostorska statistika na Fakulteti za gradbeništvo in geodezijo Univerze v Ljubljani pri prof. dr. Goranu Turku v slovenščini. Študentki sta imeli poleg izpisov v živo med predavanji na voljo tudi videoposnetke predavanj, ki jih je predavatelj posnel samostojno (izven okvira sistema Online Notes), opremljeni pa so bili z avtomatsko generiranimi podnapisi v angleščini, ki jih je možno generirati v okviru sistema Online Notes. Osnova za podnapise so bili popravljeni slovenski transkripti. Poleg predavanj v slovenščini sta obiskovali še vaje, ki so potekale v angleščini.

Študentki sta bili na splošno s sistemom zadovoljni in sta mnenja, da jima je omogočil spremljanje predavanj. Bili pa sta enotni pri tem, da so bili pri učenju in samem razumevanju vsebine ključni tudi videoposnetki s podnapisi.

Ena od študentk je v več fazah uporabe omenila, da je spremljanje različnih enot (sistema na računalniku, table, profesorja) miselno zelo zahtevno, po-

sebnost tudi, ker gre za dva tuja jezika poleg materne jezika, o čemer so sicer poročali tudi v drugih študijah, denimo pri evalvaciji sorodnega sistema KIT v Nemčiji [3]. Druga študentka tega ni opazila. Možno je, da na percepcijo zahtevnosti vpliva tudi stopnja razumevanja angleščine (prva študentka je namreč svoje razumevanje angleščine ocenila z B1/B2, druga pa na C1/C2), vendar bi bilo treba za takšne ugotovitve študijo razširiti še na večje število udeležencev.

Poleg tega sta študentki v vseh fazah poročanja omenili slabšo kvaliteto in natančnost prevodov v živo v primerjavi s prevodi pri podnapisih. Razlika v kvaliteti oziroma točnosti prevodov izvira iz dveh pomembnih razlik v postopku priprave. Glavni razlog izboljšave je ta, da so podnapisi (prevodi po predavanju) nastali iz popravljenih in urejenih slovenskih transkriptov, medtem ko prevodi v živo seveda ohranjajo in celo potencirajo vse napake razpoznavalnika govora. Na tem mestu omenimo, da je bila sicer razpoznavna govora s takratno verzijo modela razpoznavalnika govora za tehnična predavanja pri tem predavanju nekoliko slabša od povprečja, verjetno ravno zaradi velike vključenosti eksternih elementov v govor in velike zastopanosti spremenljivk in enačb, ki so v jezikovnem modelu slabše zastopane. S standardno metriko WER (ang. *word error rate*), ki izraža razmerje dodanih, manjkajočih in napačnih besed v razmerju do vseh pojavnici⁵ v primerjavi z referenčnim transkriptom, ki je pripravljen ročno, smo ocenili vsa pilotna predavanja. Prvo predavanje v sklopu semestrskega testiranja je imelo stopnjo napake 20 %, kar je za 2 % več kot povprečna stopnja napake razpoznavalnika govora na pilotnih predavanjih v študijskem letu 2022/2023 [6].

Druga razlika v kvaliteti prevodov pa je v tem, da se pri pripravi prevodov po predavanju ponovno prevede celotno vsebino predavanja naenkrat, kvaliteta prevodov pa se običajno boljša s količino konteksta. Prevodi v živo se generirajo na ravni enega besedilnega segmenta, torej je v prevajalnik poslano minimalno konteksta – s tem se želimo izogniti pretiranemu časovnemu zamiku. Tekom predavanja, torej spremljanja v živo, želimo tujim študentom zagotoviti čim hitrejše prevode, zamik je namreč lahko zelo moteč, posebno v kombinaciji z drugimi vizualnimi informacijami (informacije na tabli oziroma predstavitvi).

⁵ WER se izračuna po naslednji enačbi: (substituti + manjkajoče besede + odvečne besede)/število vseh pojavnici.

Predavatelj je bil s samim sistemom zadovoljen. Iz njegovih odgovorov, ki smo jih prejeli tekom intervjuja, lahko sklepamo, da se mu je sistem zdel uporaben in mu je olajšal delo. Ločena predavanja običajno v angleščini izvajajo, če je v skupini več kot 5 tujih študentov, sicer pa se srečanja izvajajo konsultacijsko. Profesor vsa svoja predavanja snema že od študijskega leta 2017/2018 dalje, zato mu je posebno prav prišla opcija izvoza podnapisov iz sistema.

Prof. dr. Goran Turk je po izvedbi te študije sistem semestrsko uporabljal še pri dveh predmetih v letnem semestru študijskega leta 2022/2023 in ga trenutno uporablja v letnem semestru v študijskem letu 2023/2024.

6 NADALJNJE DELO

Ena izmed nadgradenj, ki jo želimo ponuditi v sklopu projekta in se je izkazala kot dobrodošlo tudi tekom semestrskega testiranja, opisanega v poglavju 4, je izbira ciljnega jezika prevodov. Sledenje vsebini v slovenščini v živo in branje v angleščini, poleg procesiranja v tretjem, maternem jeziku, je lahko zahtevno, zato je smiselno študentom ponuditi opcijo prevoda v njihov materni jezik. Posebno tudi zato, ker je na Univerzi v Ljubljani veliko študentov s področja Balkana, katerih materni jezik je bližje slovenščini kot angleščini, zato se zdi angleščina v takšnih primerih uporabe odvečen korak. Za prevode se trenutno uporablja zunanji strojni prevajalnik, v kratkem pa bo znotraj sistema na voljo tudi strojni prevajalnik iz slovenščine v angleščino, ki je bil razvit na Univerzi v Ljubljani. V primerih prevajanja v druge tuje jezike (poleg angleščine) pa bi bil za prevajanje še vedno na voljo eksterni prevajalnik.

Poleg uporabnosti za tuje študente je velik poudarek pri nadaljnjem razvoju same aplikacije tudi dostopnost sistema za osebe s senzornimi oviranostmi. Za izboljšanje dostopnosti in s tem večjo uporabnost sistema sodelujemo s strokovnjaki s področja dostopnosti.

Ves čas nadgrajujemo sistem tudi z vidika enostavnosti uporabe in uporabniške izkušnje. Cilj je, da je urejanju transkriptov po predavanju čim enostav-

nejše, zato poskušamo upoštevati čim več predlogov uporabnikov, ki smo jih zbrali v času pilotne faze projekta (na primer hitrejše pomikanje po vsebini predavanja, iskanje po celotni vsebini predavanja in podobno).

ZAHVALA

Raziskava je bila finančno podprta s sredstvi projekta Online Notes, ki ga financira Univerza v Ljubljani.

Hvala tudi prof. dr. Goranu Turku s Fakultete za gradbeništvo in Geodezijo na Univerzi v Ljubljani za sodelovanje v študiji in aktivno uporabo sistema ves čas pilotne faze projekta.

LITERATURA

- [1] Bajec, M., Lebar Bajec, I., Šoltes, T., Cvek, J., Čibej, J., Gantar, K., Sever, S., & Krek, S. (2023). Online Notes - a real-time speech recognition and machine translation system for Slovene university lectures. 7–10. https://is.ijs.si/wp-content/uploads/2023/11/IS2023_Volume-H.pdf
- [2] Dessloch, F., Ha, T.L., Müller, M., Niehues, J., Nguyen, T.S., Pham, N.Q., Salesky, E., Sperber, M., Stüker, S., Zenkel, T., Waibel, A.: KIT lecture translator: Multilingual speech translation with one-shot learning. In: Proceedings of the 27th International Conference on Computational Linguistics: System Demonstrations. pp. 89–93. Association for Computational Linguistics, Santa Fe, New Mexico (Aug 2018), <https://aclanthology.org/C18-2020>
- [3] Markus Müller, Sarah Fünfer, Sebastian Stüker, and Alex Waibel. 2016. Evaluation of the KIT Lecture Translation System. In Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16), str. 1856–1861, Portorož, Slovenia. European Language Resources Association (ELRA).
- [4] Povey, D., Ghoshal, A., Boulianne, G., Burget, L., Glembek, O., Goel, N., Hannemann, M., Motlicek, P., Qian, Y., Schwarz, P., Silovsky, J., Stemmer, G., Vesely, K.: The kaldi speech recognition toolkit. In: IEEE 2011 Workshop on Automatic Speech Recognition and Understanding. IEEE Signal Processing Society (Dec 2011), iEEE Catalog No.: CFP11SRW-USB
- [5] Shadiev, R., Huang, Y.M.: Investigating student attention, meditation, cognitive load, and satisfaction during lectures in a foreign language supported by speecheenabled language translation. Computer Assisted Language Learning 33(3), 301–326 (2020). <https://doi.org/10.1080/09588221.2018.1559863>
- [6] Šoltes, T., Bajec, M., Lebar Bajec, I., Gantar, K., & Žitnik, S. (2023). Online-notes system: real-time speech recognition and translation of lectures. 485–492. https://link.springer.com/chapter/10.1007/978-3-031-33080-3_29
- [7] Univerza v številkah. <https://www.uni-lj.si/univerza/o-nas/univerza-v-stevilkah>. Dostop 30. 3. 2024.

■

Tjaša Šoltes je raziskovalka v Laboratoriju za podatkovne tehnologije na Fakulteti za računalništvo in informatiko UL in doktorska študentka programa Digitalno jezikoslovje. Ukvarja se z obdelavo naravnega jezika, posebno za namene govornih tehnologij. Sodelovala je pri raznih projektih na temo digitalizacije slovenskega jezika in uporabe govornih tehnologij za slovenščino, denimo RSDO (Razvoj slovenščine v digitalnem okolju), Online Notes, razvoj sintetizatorja govora za slovenski jezik, pilotni projekt uporabe govornih tehnologij v Državnem zboru ipd.

■

Jan Vasiljević je študent prvega letnika magistrskega programa Računalništvo in informatika na Fakulteti za računalništvo in informatiko Univerze v Ljubljani. V svoji diplomski nalogi je preučeval vpliv psihometričnih lastnosti v skupinskih okoljih in napisal članek, objavljen v reviji Electronics. Kot študent dela v Laboratoriju za podatkovne tehnologije in trenutno sodeluje pri projektu ON. V sklopu projekta razvija zaledne storitve in integracijo z govornimi tehnologijami.

■

Marko Bajec je redni profesor ter vodja Laboratorija za podatkovne tehnologije na Fakulteti za računalništvo in informatiko Univerze v Ljubljani. Raziskovalno in aplikativno se ukvarja z razvojem podatkovno intenzivnih sistemov. V zadnjih letih večino časa posveča govornim tehnologijam ter digitalizaciji slovenskega jezika. Na tem področju je vodil številne projekte kot npr. Samodejno podnaslavljanje TV programov na RTV SLO, Uvedba govornih tehnologij v Telekom Slovenije ipd.