

Modul I: Inteligentni avdio in govorni sistemi

BMA-EL-AI LJ RE 2 - 2024/2025

Študijski projekti

Simon Dobrišek

TEME PREDSTAVITVE

- Predstavitev izvajanja študijskih projektov
- Predlagana teme študijskih projektov
- Časovnica izvajanja študijskih projektov
- Zaključek

PREDSTAVITEV IZVAJANJA ŠTUDIJSKIH PROJEKTOV

- Vsak študent iz seznama predlaganih tem študijskih projektov izbere svojo temo, ali pa v dogovoru s predavateljem predlaga svojo lastno temo, ki sodi v okvir vsebine predmeta.
- V primeru, da več študentov izbere isto temo in se študenti med sabo ne morejo dogovoriti o izbiri, ima prednost študent, ki je predavatelju prvi izrazil interes za predlagano temo.
- Študijski projekt se pripravi v obliki pisne pregledne študije obravnavanega področja, ki se predstavi tudi v okviru predavanj.
- Po izbiri tem študenti na predavanjih (po objavljenem rokovniku) v približno 10 minutah (10-15 drsnic) predstavijo svoje izbrane teme in katere vsebine in vire nameravajo preštudirati.

Predstavitev izvajanja študijskih projektov

- Primerne vire in literaturo za svojo izbrano študijsko področje študenti izberejo sami in svoj predlog virov in literature uskladijo s predlogi predavatelja.
- V zadnjem tednu zimskega semestra študenti na predavanjih v približno 20 minutah predstavijo rezultate svojega študijskega projekta.
- Študenti nato do konca zimskega semestra oz. najkasneje do prijave na prvi izpitni rok oddajo svojo pisno študijo.
- Po oddaji in predstavitvi študijskega projekta študenti prejmejo za svoj projekt oceno v skladu z izpitnim redom predmeta, ki se potem skupaj z oceno vaj pretvori v končno oceno.
- Skupno oceno je možno izboljšati na zaključnem pogovoru s predavateljem z dodatnim ustnim izpitom, ki pa ni obvezen.

MOŽNE TEME ŠTUDIJSKIH PROJEKTOV – STROJNI SLUH

1. Razpoznavanje okoljskih zvokov za zaznavanje onesnaženja in drugih spremenjenih stanj okolja.
2. Razpoznavanje zvokov v medicinski diagnostiki za razpoznavanje specifičnih zvokov, povezanih z zdravstvenimi stanji ljudi.
3. Slušni nadzor in diagnosticiranje naprav in strojev za razpoznavanje nenormalnih zvokov v strojih in napravah.
4. Razpoznavanje glasbenih inštrumentov in melodij
5. Inteligentni avdio nadzor za varnost in varovanje prostorov.
6. Razpoznavanje zvokov v naravi, kot so zvoki različnih bitij in drugi okoljski in vremenski zvoki.
7. Razpoznavanje zvokov v proizvodnji za nadzor kakovosti.
8. Integracija inteligentnega strojnega sluha v pametne domove

MOŽNE TEME ŠTUDIJSKIH PROJEKTOV – GOVORNE TEHNOLOGIJE

1. Primerjava kratko-časovnih govornih značilk MFCC in PLP.
2. Uporaba i-vektorjev kot govornih značilk.
3. Uporaba umetnega časovno zakasnjenege nevronskega omrežja (angl. Time-delay Neural Network – TDNN) za akustično modeliranje govora.
4. Uporaba ponavljajočega nevronskega omrežja (angl. Recurrent Neural Network - RNN) in dolgega kratkotrajnega pomnilnika (angl. Long Short-Term Memory - LSTM) za jezikovno modeliranje govora.
5. Metode in modeli, ki se uporabljajo za samodejno razpoznavanje čustvenih stanj govorcev.
6. Metode in modeli, ki se uporabljajo za samodejno razpoznavanje zdravstvenega psiho-fizičnega stanja govorcev.

Predlagana teme študijskih projektov

7. Metode in modeli, ki se uporabljajo za več-modalno razpoznavanje govora (zvok in slika), denimo, slikovno branje govora iz gibanja ustnic.
8. Tvorjenje umetnega govora s korpusno sintezo z združevanjem govornih enot (angl. corpus-based concatenative speech synthesis).
9. Tvorjenje umetnega govora z uporabo akustičnih modelov, ki temeljijo na prikritih Markovovih modelih - PMM.
10. Tvorjenje umetnega govora z globokimi nevronskimi omrežji, kot je Google WaveNet.
11. Uporaba ponavljajočega nevronskega omrežja (RNN) in dolgega kratkotrajnega pomnilnika (LSTM) za samodejno prevajanje govora.
12. Pomenska/semantična analiza govora s samodejnim tvorjenjem pomenskih/semantičnih okvirjev (angl. semantic frames).

Predlagana teme študijskih projektov

13. Pristopi h gradnji sistemov za vodenje govornega dialoga z uporabniki, ki temeljijo na uporabi končnih avtomatov in modeliranju (govornih) dejanj v dialogu.
14. Pristopi h gradnji sistemov za vodenje govornega dialoga z uporabniki, ki temeljijo na polnjenju pomenskih/semantičnih okvirjev/predalčkov.
15. Pristopi h gradnji sistemov za vodenje dialoga, ki temeljijo na usmerjanju k cilju in modeliranju prepričanj, želja in namer (angl. Beliefs, Desires, and Intentions - BDI).
16. Pristopi h gradnji sistemov za vodenje dialoga, ki temeljijo na konverzacijski teoriji iger (angl. conversational game theory)
17. Primerjava različnih scenarijev uporabniških sistemov z govornimi vmesniki in govornimi asistenti.

Viri in literatura za izvedbo študijskih projektov

- Pri iskanju primernih virov in literature na področju govornih tehnologij, se študenti sklicujejo predvsem na najpomembnejše znanstvene konference in revije na tem področju.

(<http://www.guide2research.com/topconf/computational-linguistics-speech-processing>)

- Najpomembnejša konferenca na področju govornih tehnologije je vsakoletna konferenca Interspeech in prispevki na teh konferencah so javno dostopni prek spletnega strežnika strokovne organizacije - International Speech Communication Association – ISCA, kjer so objavljeni tudi prispevki z mnogih drugih sorodnih konferenc.

(<https://www.isca-speech.org/iscaweb/index.php/archive/online-archive>)

- Pri znanstvenih in strokovnih člankih pa so na področju govornih tehnologij najpomembnejše objave v revijah, kot so:

- *Elsevier Computer Speech & Language*
- *IEEE Transactions on Audio, Speech and Language Processing*
- *Elsevier Speech Communication*

(<http://www.guide2research.com/journals/signal-processing>)

Format študijskega poročila

- Študijsko poročilo naj bo besedilno oblikovano s poljubnim urejevalnikom besedil v A4 PDF formatu v enem ali dveh stolpcih po izbiri študenta.
- Pri besedilnem oblikovanju v enem stolpcu je priporočena uporaba besedilne predloge založbe Springer (ali podobna) - <https://www.springer.com/gp/authors-editors/conference-proceedings/conference-proceedings-guidelines>
- Pri besedilnem oblikovanju v enem stolpcu se lahko zgleduje tudi po besedilni predlogi za zaključna dela na naši fakulteti - https://www.fe.uni-lj.si/izobrazevanje/studentska_pisarna/pravilniki_in_navodila/zakljucno_delo/
- Pri besedilnem oblikovanju v dveh stolpcih je priporočena uporaba besedilne predloge združenja IEEE (ali podobna) - <https://www.ieee.org/conferences/publishing/templates.html>
- Pričakovani obseg poročila pri besedilnem oblikovanju v dveh stolpcih je vsaj 4 in ne več kot 8 strani besedila, pri besedilnem oblikovanju v enem stolpcu pa vsaj 7 in ne več kot 15 strani besedila.

ČASOVNICA IZVAJANJA ŠTUDIJSKIH PROJEKTOV

- Izbira in individualna uskladitev teme študijskega projekta do predavanj v ponedeljek, **21. oktobra 2024.**
- Kratka javna predstavitev izbrane teme projekta ter virov in literature na predavanjih v ponedeljek, **4. novembra 2024.**
- Oddaja dispozicije oziroma osnutka študijskega projekta do ponedeljka, **25. novembra 2024.**
- Zaključna javna predstavitev rezultatov študijskih projektov na predavanjih v ponedeljek, **2. decembra 2024.**
- Oddaja pisne študije, ki je bila izvedena v okviru študijskega projekta najkasneje do ponedeljka, **3. februarja 2025.**

ZAKLJUČEK

- Predstavljen je bil način izvedbe študijskih projektov v okviru predmeta Inteligentni avdio in govorni sistemi.
- Predstavljene so bile različne možne teme študijskih projektov.
- Študenti lahko tudi sami predlagajo temo, ki po dogovoru s predavateljem sodi v okvir vsebine predmeta.
- Podana je bila časovnica izvajanja, predstavitve in zaključka študijskih projektov