

Vijeću Prirodno-matematičkog fakulteta

Molim Vijeće Prirodno-matematičkog fakulteta da odobri uvođenje izbornih predmeta **Vremensko-frekvencijska analiza sa primjenama** i **Automatsko prepoznavanje govora** na doktorskim studijama na smjeru Računarske nauke.

Vremensko-frekvencijska analiza je oblast od velike važnosti za sve istraživače, inženjere i studente koji se bave bilo kakvim oblikom obrade signala. Ova oblast obuhvata koncepte, tehnike i metode kojima se manipuliše signalima koji se mijenjaju u vremenu. Nalazi primjenu u velikom broju savremenih naučnih i tehnoloških oblasti, kao što su telekomunikacije, biomedicina, multimedije, informacione tehnologije itd.

Cilj kursa "Vremensko-frekvencijska analiza sa primjenama" je da studenti steknu teorijsko znanje o vremensko-frekvencijskoj analizi signala, kao i praktično znanje o primjeni ovih tehnika u obradi različitih vrsta digitalnih signala.

Automatsko prepoznavanje govora je multidisciplinarna oblast u presjeku računarskih nauka i lingvistike koja se bavi razvojem metoda i tehnika koje omogućavaju računarima da prepoznaju ljudski govor, odnosno da izvrše mapiranje govornog signala u tekst. Problemi koji su postavljeni u ovoj oblasti su predmet intenzivnog istraživanja već decenijama. U posljednjih nekoliko godina primjetan je porast interesovanja za ovu oblast praćen razvojem mobilnih uređaja i zahtjevima za postojanjem različitih aplikacija koje vrše automatsko prepoznavanje govora, kao i napretkom u oblasti mašinskog, odnosno dubokog, učenja.

Kurs "Automatsko prepoznavanje govora" će omogućiti studentima da se upoznaju sa načinima izgradnje sistema za automatsko prepoznavanje govora, počevši od najranijih modela, do savremenih modela zasnovanih na dubokom učenju.

U Podgorici 05.10.2020.

Prof. dr Igor Đurović



Naziv predmeta: Automatsko prepoznavanje govora

Šifra predmeta	Status predmeta	Semestar	Broj ECTS kredita	Fond časova
	Obavezan	I	5	

Studijski programi za koje se organizuje :

Doktorski studijski program Računarske nauke (studije traju 6 semestara, 180 ECTS kredita).

Uslovljenost drugim predmetima: Nema uslova za prijavljivanje i slušanje predmeta

Ciljevi izučavanja predmeta: Kroz ovaj predmet studenti stiču teorijska i praktična znanja o načinima izgradnje sistema za automatsko prepoznavanje govora, počevši od najranijih modela, do savremenih modela, zasnovanih na dubokom učenju.

Ime i prezime nastavnika i saradnika: prof. dr Igor Đurović

Metod nastave i savladanja gradiva: Učenje i samostalna izrada praktičnih zadataka. Konsultacije.

SADRŽAJ PREDMETA

Pripremna nedjelja	Priprema i upis semestra
I nedjelja	Uvod. Načini komunikacije između ljudi i između mašine i čovjeka. Osnovni pojmovi u fonetici i sistemima za automatsko prepoznavanje govora.
II nedjelja	Osnove arhitektura sistema za automatsko prepoznavanje govora. Mjere performansi sistema za automatsko prepoznavanje govora. Akustični model. Model jezika. Leksikon. Ekstrakcija karakteristika govornog signala. Mel-frekvencijski kepsralni koeficijenti.
III nedjelja	Gausovi mješoviti modeli. Modelovanje slučajnih procesa ponderisanom sumom Gausovih raspodjela. Primjena Gausovih mješovitih modela u modelovanju karakteristika govora.
IV nedjelja	Skriveni Markovljevi modeli. Evaluacija vjerovatnoće skrivenim Markovljevim modelom. GMM-HMM arhitektura sistema za automatsko prepoznavanje govora.
V nedjelja	EM algoritam. Primjena EM algoritma za učenje parametara skrivenog Markovljevog modela (Baum-Welch algoritam).
VI nedjelja	Viterbi algoritam za dekodiranje sekvenci stanja skrivenog Markovljevog modela.
VII nedjelja	Duboke neuronske mreže. Napredne tehnike za inicijalizaciju dubokih neuronskih mreža. Ograničene Bolcmanove mašine. Pretreniranje neuronskim mrežama vjerovanja, pretreniranje autoenkoderima. Hibridni pristupi za pretreniranje i pretreniranje primjenom dropout tehnike.
VIII nedjelja	Hibridni sistemi za automatsko prepoznavanje govora. DNN-HMM arhitektura sistema za automatsko prepoznavanje govora. Primjena konvolucionih neuronskih mreža u sistemima za prepoznavanje govora.
IX nedjelja	Rekurentne neuronske mreže. LSTM ćelije u rekurentnim neuronskim mrežama. Bidirekciona LSTM neuronska mreža.
XI nedjelja	Obučavanje modela za mapiranje sekvenci u sekvence. Programska realizacija rekurentnih neuronskih mreža korišćenjem biblioteka Keras i TensorFlow u programskom jeziku Python.
XII nedjelja	Sistemi za automatsko prepoznavanje govora zasnovani isključivo na rekurentnim neuronskim mrežama. CTC funkcija gubitka. Beam search algoritam
XIII nedjelja	Mehanizmi za pažnju kod neuronskih mreža. Modeli za prepoznavanje govora zasnovani na mehanizmima za pažnju kod neuronskih mreža.
XIV nedjelja	RNN-T model za automatsko prepoznavanje govora.
XV nedjelja	Tehnike za poboljšanje performansi sistema za automatsko prepoznavanje govora. SpecAugment tehnika za povećanje obima podataka. Planirano uzorkovanje. Ugladivanje labela. Normalizacija slojeva. Slojevi za skraćivanje vremena.
XVI nedjelja	Završni ispit
XVII nedjelja	Ovjera semestra i upis ocjena
XVIII-XXI nedjelja	Dopunska nastava i popravni ispitni rok

Opterećenje studenta u časovima:

<u>nedjeljno</u>	<u>u semestru</u>
5 kredita x 40/30 = 6 sati i 40 minuta	Nastava i završni ispit: (8 sati) x 16 = 106 sati i 20 minuta. Neophodne pripreme (administracija, upis, ovjera prije početka semestra) 2 x (6 sati i 4 minuta) = 13 sati i 20 minuta Ukupno opterećenje za predmet: 5x30 = 150 sati Dopunski rad: za pripremu ispita u popravnom ispitnom roku, uključujući i polaganje popravnog ispita od 0 do 30 sati (preostalo vrijeme od prve dvije stavke do ukupnog opterećenja za predmet 150 sati) Struktura opterećenja: 106sati i 40min.(nastava) + 13sati i 20min.(priprema) + 30 sati (dopunski rad)

Literatura:

- Dong Yu, Li Deng, „Automatic Speech Recognition: A Deep Learning Approach“, Springer Publishing Company, Incorporated, UK, 2014.
- I. Goodfellow, Y. Bengio, A. Courville, „Deep Learning“, MIT Press, USA, 2016.
- D. Povey, „Discriminative training for large vocabulary speech recognition“ Ph.D. dissertation, Cambridge University, UK, 2003.
- A. Graves, N. Jaitly and A. Mohamed, „Hybrid speech recognition with Deep Bidirectional LSTM“, IEEE Workshop on Automatic Speech Recognition and Understanding, 2013, pp. 273-278.
- O. Abdel-Hamid, A. Mohamed, H. Jiang, L. Deng, G. Penn and D. Yu, „Convolutional Neural Networks for Speech Recognition“, IEEE/ACM Transactions on Audio, Speech, and Language Processing, Vol. 22, No. 10, October 2014, pp. 1533-1545
- A. Graves, „Sequence Transduction with Recurrent Neural Networks“, arXiv preprint arXiv:1211.3711, November 2012.
- I. Sutskever, O. Vinyals, Q. Le, „Sequence to Sequence Learning with Neural Networks“, arXiv preprint arXiv:1409.3215v3, December 2014.
- A. Graves, N. Jaitly, „Towards End-to-End Speech Recognition with Recurrent Neural Networks“, Proceedings of the 31st International conference on Machine Learning (ICML), June 2014, pp. 1764-1772.
- J. Chorowski, D. Bahdanau, D. Serdyuk, K. Cho, Y. Bengio „Attention-based Models for Speech Recognition“, arXiv preprint arXiv:1506.07503v1, June 2015.
- W. Chan, N. Jaitly, Q. Le, O. Vinyals, „Listen, Attend, Spell“, arXiv preprint arXiv: 1508.01211v2, August 2015.
- D. Park, W. Chan, Y. Zhang, C. Chiu, B. Zoph, E. Cubuk, Q. Le. „SpecAugment: A simple data augmentation method for automatic speech recognition“, arXiv preprint arXiv:1904.08779v3, December 2019.
- S. Bengio, O. Vinyals, N. Jaitly, N. Shazeer, „Scheduled Sampling for Sequence Prediction with Recurrent Neural Networks“, arXiv preprint arXiv:1506.03099v3, September 2015.

Oblici provjere znanja i ocjenjivanje:

Pismeno i usmeno polaganje pred najmanje tročlanom komisijom.

Posebne naznake za predmet: Predavanja se mogu držati na engleskom jeziku.

Napomena:



Broj: 17087

Podgorica, 05-10-2020 20 god.

Vremensko-frekventna analiza sa primjenama

Naziv predmeta:

Sifra predmeta	Status predmeta	Semestar	Broj ECTS kredita	Fond časova
	Obavezan	I	5	

Studijski programi za koje se organizuje :

Doktorski studijski program Računarske nauke (studije traju 6 semestara, 180 ECTS kredita).

Uslovljenost drugim predmetima: Nema uslova za prijavljivanje i slušanje predmeta

Ciljevi izučavanja predmeta: Kroz ovaj predmet studenti stiču teorijsko znanje o vremensko-frekventnoj analizi signala, kao i praktično znanje o primjeni vremensko-frekventnih tehnika u obradi različitih vrsta digitalnih signala, sa naglaskom na govorne signale.

Ime i prezime nastavnika i saradnika: prof. dr Igor Đurović

Metod nastave i savladanja gradiva: Učenje i samostalna izrada praktičnih zadataka. Konsultacije.

SADRŽAJ PREDMETA

Pripremna nedjelja	Priprema i upis semestra
I nedjelja	Uvod u Furijeovu analizu, Furijeova transformacija. Inverzna Furijeova transformacija. Svojstva Furijeove transformacije. Veza između Furijeovog reda i Furijeove transformacije.
II nedjelja	Furijeova transformacija diskretnog signala. Teorema o odabiranju u vremenskom domenu.
III nedjelja	Diskretna Furijeova transformacija. Inverzna diskretna Furijeova transformacija. Analiza sinusoide korišćenjem diskretne Furijeove transformacije.
IV nedjelja	Kratkotrajna Furijeova transformacija. Prozori (pravougaoni, Bartlett, Hann, Hamming, Blackman i Kaiser). Spektrogram. Inverzna kratkotrajna Furijeova transformacija. Estimacija signala iz modifikovane kratkotrajne Furijeove transformacije.
V nedjelja	Realizacija kratkotrajne Furijeove transformacije i inverzne kratkotrajne Furijeove transformacije u GNU Octave-u i Python-u
VI nedjelja	Wavelet transformacija. S-transformacija (Stockwellova transformacija). Lokalna polinomijalna Furijeova transformacija.
VII nedjelja	Tehnike za kratkotrajnu Furijeovu transformaciju visoke rezolucije. Caponov i MUSIC metod.
VIII nedjelja	Wignerova distribucija. Auto-članovi i kros-članovi u Wignerovoj distribuciji. Svojstva Wignerove distribucije. Realizacija Wignerove distribucije. Diskretna pseudo Wignerova distribucija.
IX nedjelja	S-Metod. Diskretna realizacija S-Metoda.
X nedjelja	Analiza signala zahvaćenih šumom. Kratkotrajna Furijeova transformacija signala zahvaćenog šumom. Wignerova distribucija signala zahvaćenog šumom.
XI nedjelja	Adaptivni algoritam za estimaciju šuma. Estimacija šuma minimizacijom srednje kvadratne greške estimatora kratkotrajnih spektralnih amplituda.
XII nedjelja	Estimacija varijanse šuma. Ekstrakcija i rekonstrukcija signala korišćenjem kompleksnih vremensko-frekventnih filtera.
XIV nedjelja	Primjena vremensko-frekventne analize u obradi radarskih signala, seizmičkih signala, biomedicinskih signala. Detekcija signala primjenom Furijeove transformacije.
XV nedjelja	Primjena vremensko-frekventne analize u obradi govornih signala. Eliminacija šuma iz govornog signala zahvaćenog šumom primjenom tehnike spektralnog oduzimanja. Programaska realizacija tehnike spektralnog oduzimanja. Detekcija glasa u audio signalu. Statistički modeli za detekciju glasa.
XVI nedjelja	Završni ispit
XVII nedjelja	Ovjera semestra i upis ocjena
XVIII-XXI nedjelja	Dopunska nastava i popravni ispitni rok

Opterećenje studenta u časovima:

5 kredita x 40/30 = 6 sati i 40 minuta

u semestru

Nastava i završni ispit: (8 sati) x 16 = **106 sati i 20 minuta.**

Neophodne pripreme (administracija, upis, ovjera prije početka semestra) 2 x (6 sati i 4 minuta) = **13 sati i 20 minuta**

Ukupno opterećenje za predmet: 5x30 = **150 sati**

Dopunski rad: za pripremu ispita u popravnom ispitnom roku, uključujući i polaganje popravnog ispita od 0 do **30 sati** (preostalo vrijeme od prve dvije stavke do ukupnog opterećenja za predmet 150 sati)

Struktura opterećenja:

106sati i 40min.(nastava) + 13sati i 20min.(priprema) + 30 sati (dopunski rad)

Literatura:

- Lj. Stanković, M. Daković, T. Thayaparan, „*Time-Frequency Signal Analysis with Applications*“, Artech House, USA, March 2013.
- Lj. Stanković, „*Digital Signal Processing with Selected Topics*“, CreateSpace Independent Publishing Platform, An Amazon.com Company, USA, November 2015, pp. 521-664.
- H. Zhivomirov, „*On the Development of STFT-analysis and ISTFT-synthesis Routines and their Practical Implementation*“, TEM Journal Vol. 8, No. 1, February 2019, pp. 56-64.
- D. Griffin, J. Lim, „*Signal Estimation from Modified Short-Time Fourier Transform*“, IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. 32, No. 2, April 1984, pp. 236-243.
- Y. Ephraim and D. Malah, „*Speech enhancement using a minimum-mean square error short-time spectral amplitude estimator*“, IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. 32, No. 6, December 1984, pp. 1109-1121.
- D. Makovoz, „*Noise Variance Estimation In Signal Processing*“, IEEE International Symposium on Signal Processing and Information Technology, Canada, August 2006, pp. 364-369.
- W. Mack, E. Habets, „*Deep Filtering: Signal Extraction and Reconstruction Using Complex Time-Frequency Filters*“, IEEE Signal Processing Letters, Vol. 27, November 2019, pp. 61-65.
- S. Boll, „*Suppression of Acoustic Noise in Speech Using Spectral Subtraction*“, IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. 27, No. 2, April 1979, pp. 113-120.
- M. Berouti, R. Schwartz, and J. Makhoul, „*Enhancement of speech corrupted by acoustic noise*“, Proceedings of International Conference on Acoustics, Speech, and Signal Processing, USA, April 1979, pp. 208-211.
- Jongseo Sohn, Nam Soo Kim and Wonyong Sung, „*A statistical model-based voice activity detection*“, IEEE Signal Processing Letters, Vol. 6, No. 1, January 1999, pp. 1-3.

Oblici provjere znanja i ocjenjivanje:

Pismo i usmeno polaganje pred najmanje tročlanom komisijom.

Posebne naznake za predmet: Predavanja se mogu držati na engleskom jeziku.

Napomena:

