

I069	Obrada prirodnog jezika tehnikama dubinskog učenja	P	V	S	ECTS 8
		3	2	1	

Cilj predmeta. Obrada prirodnog jezika pripada jednom od temeljnih područja umjetne inteligencije stoga će studenti u ovom predmetu modelirati probleme obrade prirodnog jezika koristeći napredne tehnike iz strojnog učenja. Teorijskim pristupom bit će upoznati s različitim primjenama i jezičnim modelima, a kroz praktičan rad razvijat će vlastite dubinske modele u modernim Python bibliotekama za dubinsko učenje te vrednovati na stvarnim primjenama obrade prirodnog jezika.

Potrebna predznanja. Preddiplomski studij matematičkog, računarskog ili srodnog smjera.

Sadržaj predmeta.

1. Uvod u obradu prirodnog jezika i tehnike dubinskog učenja. Vektori riječi i pripadne metode. Modeli jezika.
2. Jednostavne reprezentacije vektora riječi: word2vec, GloVe.
3. Napredne reprezentacije vektora riječi: jezični modeli, softmax, jednoslojne mreže.
4. Neuronske mreže i povratna propagacija. NER: prepoznavanje i klasifikacija imenskih fraza.
5. Rekurentne neuronske mreže za modeliranje jezika.
6. GRU i LSTM za strojno prevođenje.
7. Rekurzivne neuronske mreže za parsiranje i analizu sentimenta.
8. Konvolucijske neuronske mreže za klasifikaciju rečenica.
9. Modeli dubinskih neuronskih mreža za prepoznavanje govora.
10. Mreže dinamičke memorije za strojno odgovaranje na pitanja.

ISHODI UČENJA

R.b.	ISHODI UČENJA
1.	Opisati temeljne primjene obrade prirodnog jezika u strojnom zaključivanju.
2.	Opisati različite reprezentacije vektora riječi i jezične modele.
3.	Definirati različite arhitekture dubinskih mreža u primjeni obrade prirodnog jezika.
4.	Odabrati odgovarajuće dubinske modele za specifične zadatke obrade prirodnog jezika.
5.	Izvesti i implementirati optimizacije algoritme za modele neuronskih mreža.
6.	Analizirati i vrednovati performanse dubinskih modela.
7.	Dizajnirati i evaluirati vlastite dubinske modele za obradu prirodnog jezika.

POVEZIVANJE ISHODA UČENJA, ORGANIZACIJE NASTAVNOG PROCESA I PROCJENA ISHODA UČENJA

ORGANIZACIJA NASTAVNOG PROCESA	ECTS	ISHOD UČENJA **	AKTIVNOST STUDENATA*	METODA PROCJENE	BODOVI	
					min	max
Pohađanje predavanja i vježbi	1	1-7	Prisutnost na nastavi, rasprava, timski rad, samostalan rad na zadacima i kratke provjere znanja	Potpisne liste, praćenje aktivnosti na nastavi, zadaci zatvorenog tipa	0	10
Domaće zadaće	2	1-7	Samostalno rješavanje programerskih zadataka	Provjera točnih rješenja (ocjenjivanje)	15	25
Provjera znanja (kolokvij)	2	1 - 6	Priprema za pisanu provjeru znanja	Provjera točnih rješenja (ocjenjivanje)	15	25

Završni projekt	3	6-7	Samostalno rješavanje problema iz primjene	Izrada postera i izlaganje (ocjenjivanje)	20	40
UKUPNO	8				50	100

Izvođenje nastave i vrednovanje znanja. Predavanja i vježbe su obvezni. Ispit se sastoji od pismenog dijela, a polaže se nakon odslušanih predavanja. Prihvatljivi rezultati postignuti na kolokvijima, koje studenti pišu tijekom semestra, zamjenjuju pismeni dio ispita. Vježbe su djelomično auditorne, a djelomično laboratorijske uz korištenje računala. Tijekom semestra studenti su obvezni pisati zadaće i izraditi završni projekt te bodovima iz zadaća i projekta utječu na ukupnu ocjenu.

Može li se predmet izvoditi na engleskom jeziku: Da

Osnovna literatura:

1. Y. Goldberg, Neural Network Models for Natural Language Processing, Morgan & Claypool Publishers, 2017.
2. P. Goyal, S. Pandey, K. Jain, Deep Learning for Natural Language Processing: Creating Neural Networks with Python, Apress, 2018.

Dopunska literatura:

1. L. Deng, Y. Liu, Deep Learning in Natural Language Processing, Springer, 2018.
2. A. Clark, C. Fox, S. Lappin, Computational Linguistics and Natural Language Processing, Wiley-Blackwell, 2010.