

O'ZBEKISTON RESPUBLIKASI
OLIY TA'LIM, FAN INNOVATSIYALAR VAZIRLIGI

ALISHER NAVOIY NOMIDAGI
TOSHKENT DAVLAT O'ZBEK TILI VA ADABIYOTI UNIVERSITETI



“NLP algoritmlari”

fanining o‘quv dasturi (sillabusi)

Bilim sohasi:	200 000 – San’at va gumanitar fanlar
Ta’lim sohasi:	230 000 – Tillar
Ta’lim yo’nalishi:	60230800 - Kompyuter lingvistikasi

Toshkent-2023

Fan/modul kodı	Fan/modul turi	Ta'lim tili	Ishlab chiqilgan o'quv yili
302NLP27	fan/tanlov	o'zbek	2023/2024

Semestr	Har bir semestrdaagi fan/modulning nomi	O'quv mashg'ulotlari (soat)				Mustaqil ta'lim (soat)	Jami yuklama (soat)
		ECTS - Kreditlar	Haftalik dars soatlari	Jami:	ma'ruza seminar amaliy laboratoriya		
4	NLP algoritmlari	6	6	90	30	0	60 0 90 180
	Jami:	6	6	90	30	0	60 0 90 180

1. FANNING MAZMUNI

“NLP algoritmlari” o'quv fannini o'rganish o'quvchilarning dasturlash tillari yordamida dasturlar tuzish, kompyuter lingvistikasi sohasida uchraydigan masalalarни yechilish ketma-ketliklari tuzish, muammolarni yechimi bo'lgidan matematik modeldar va usullarni ishlab chiqish kabi ko'nikmalarga ega bo'lismiga erishish. Usbu maqsadga erishishda qator usullarni o'rganish va algoritim tuzish usullarini mustakam o'rganish asosiy vazifalardandir;

NLP algoritmlari fannini o'zlashtirish uchun o'qitishning ilg'or va zamonaiviy grafik muharrirlarda ishlash usullari va yangi stillarni tatbiq qilish muhim ahamiyatga ega. Fanni o'zlashtirishda darslik, o'quv va ushluby qo'llanmalar, ma'ruba manzari, standartlar, tarqatma materiallар, elektron materiallар, taqdimatlar, videoinars, virtual stendlar va ilg'or pedagojik texnologiyalardan foydalaniadi.

Fanni o'qitishdan maqsad – tababalarga Tabiyiy tillar qayta ishlash sohasidagi muammolar va ularni yechilish algoritmlari, usullari, prinsipi, texnik va dasturiy ta'minoti, imkoniyatlari va ko'nikmalarni hosil qilishdan iborat.

2. FANNI O'QITISH NATIJALARI VA SHAKLLANADIGAN KASBIY KOMPETENSIYALAR

Fan bo'yicha talabalarning bilim, ko'nikma va malakalarga quyidagi talablar qo'yiladi. Talaba:

- Algoritim tushunchasi va uning turlari to'g'risida *bilinga*;
- Masalalaring qo'yilishiga qarat muammoga yechim topa olish, algoritmlar nazariyasi metodlarini qo'llay olish, uchraydigan muammolar bo'yicha yechimlar qabul qilish ko'nikmalariga;

- Algoritmlar yaratish usullari, metodlari, masalaning yechishining eng optima usullarini bilsishi va ulardan foydalananish bo'yicha *malakaga ega bo'lishi lozim*.

3. O'QUV FANI O'QITILISHI BO'YICHA USLUBIV KO'RSATMALAR

“NLP algoritmlari” fannini o'rganish davomida mashg'ulotlar paytda axborot (riyolantiruvchi ta'lim texnologiyalari, fanni to'liq o'zlashtirishga yo'nalirilgan texnologiyalar, shaxsiga yo'nalirilgan ta'lim texnologiyalari) hamda interfaol metodlar (“Aqliy hujum”, “BBB”, “Venn diagrammasi”, “T-chizma”, “Insert”, “FSMU”, “Klaster” metodi, “Bumerang”, “Skorobey”, “Tarozi”, “Epig'ich” texnologiyasi) qo'llaniladi. Bundan tashqari darsliklar, o'quv qo'llanmalar, ma'lumotnomalar, pedagogik ensiklopediyalar va lug'allar, ma'ruba manzari, videoinarslar va tarqatma materiallardan foydalaniadi.

4. TALABALARНИ BAHOLASH TARTIBI:

Talabalarning bilimini baholash tartibi O'zbekiston Respublikasi Oly va o'rta maxsus ta'lim vazirining 2018-yil 9-avgustdag'i 19-2018-sون buyrug'i bilan tasdiqlangan “Oly ta'lim muassasalarida talabalar bilimini nazorat qilish va baholash tizimi to'g'risida”gi Nizom talabari asosida belgilanadi.

1. Nazoratlarini amalgalash tartibi.

Oraliq nazorat: O'tilgan mavzular asosida yozma shakilda o'tkaziladi. Yakuniy nazorat: Semestr davomida o'tilgan mavzular bo'yicha yozma shakilda o'tkaziladi.

Izoh: Oraliq nazorat turi har bir fan bo'yicha fanning xususiyatidan kelib chiqqan holda 2 marzagacha o'kazilishi mumkin.

Talabani oraliq nazorat turi bo'yicha baholashda, uning o'quv mashg'ulotlari davomida olgan baholari inobalg'a olinadi.

Nazoratlarning o'kazilishi muddatları tuyyan o'quv yilida tasdiqlangan o'quv jarayoni/izohvali asosida tashkil etiladi.

2. Talabalarning bilimi baholash mezonlari:

talaba mustaqil xulosa va qator qabul qiladi, ijodiy fikrlay oladi, mustaqil mushohada yuritadi, olgan bilimini amalda qo'llay oladi, fanning mohiyatini tushunadi, biladi, ifoday oladi, ayrib beradi hamda fan bo'yicha tasavvurga ega deb topilganda – 5 (a'lо) baho;

talaba mustaqil mushohada yuritadi, olgan bilimini amalda qo'llay oladi, fanning mohiyatini tushunadi, biladi, ifoday oladi, ayrib beradi hamda fan bo'yicha tasavvurga ega deb topilganda – 4 (yaxshi) baho;

talaba olgan bilimini amalda qo'llay oladi, fanning mohiyatini tushunadi, biladi, ifoday oladi, ayrib beradi hamda fan bo'yicha tasavvurga ega deb topilganda – 3 (qoniqarli) baho;

talaba fan dasturini o'zlashtirnagan, fanning mohiyatini tushunmaydi hamda fan bo'yicha tasavvurga ega emas deb topilganda – 2 (qoniqarsiz) baho bilan baholandi.

5. O'QUV MASHG'ULOTLARI VA MUSTAQIL TA'LIM QISMI

1. Ma'ruba mashg'ulotlari

Nº	Ma'ruba mavzulari	Qisqacha tavsifi (kalit so'zlar)	Soatlar hajmi
1	Kirish. Regulyar ifodalar, matnni normallashirish.	Regulyar ifodalar, so'zlar, corpora, matnni normallashirish, minimal tahrirlash masofasi	2
2	Til modelari. N-gram til modelari.	Bigram, trigramm, n-gram, umumlashtirish, nollar, Kneser-Ney siliqlash	2
3	Bayes va sentiment tasniflash algoritmlari.	Matnni turkumlashirish, sentiment analiz, Naive Bayes tasnifagichlan, Bayes xulosasi, chiziqli klassifikatsiyalar, stop words	2
4	Mantiqiy regressiya algoritmлari	Regressiya, generativ model, discriminativ model, sigmoid funksiya, logistik funksiya, sentiment klassifikatsiya, Multinomial logistik regressiya	2
5	Vektorlar semantikasi	Vektor, semantika, o'matishlar, vektor semantikasi, leksik semantika, so'zlar, vektorlar, TF-IDF	2
6	Neyron tarmoqlar va til modelari	Neyron tarmoqlar, neyron tili, vektor, sigmoid, XOR muammosi, Oldinga uzaqish neyron tarmoqlari	2
7	So'z turkumlari va NER uchun teglash algoritmlari	POS (part of speech), teglash, NER (Name entity recognition)	2
8	Sharlti tasodifiy maydonlar	Sharlti tasodifiy maydon, POS tegger, qoida, metodlar	2
9	"Chuquur" o'rganish texnologiyalari	Ketma-ketlik jarayonlari, chuquur o'rganish, arxitekturasi	2
10	Mashina tarjimasi va EDM	Mashina tarjimasi, encoder, decoder, tillar tipologiyasi	2
11	Kontekstiz grammatikalar	Kontekst, grammatika, qoidalari, lexicon, formal tarif, terminal belgilari	2
12	Noaniqliklar (aniqmasliklar)	Noaniqliklar, aniqlash, semantik analiz, sintaktik ma'nosini aniqlash	2
13	Bog'liqlarni tahlil qilish	Bog'liqliklar, munosabatlari, bog'liqlarning turlari, so'z taribi	2
14	Gap ma'nosining mantiqiy ifodalari	Semantik parsing, hisoblash semantikasi, logistik ifodalar.	2

Nº	Amaliy mashg'ulot	Qisqacha tavsifi (kalit so'zlar)	Soatlar hajmi
1	Regulyar ifodalar. Ajralish, guruhlash va ustunlik	Regular ifodalar, ajralish, guruhlash, ustunlik, operatorlar	2
2	Mantarni normalashirish. Tokenizatsiya	Tokenizatsiya, word tokenizatsiya, Unix vostitali, Penn Treebank tokenizatsiya, encoding, Porter stemmer	2
3	N-gram modellar. Maksimal ehitimollikni baholash	Unigram, bigram, trigram, n-gram modellar, maksimal ehitimollik, sharlti ehitimollik	2
4	Markov modeldarining qo'llanilishi. Hidden Markov modeldar	Ehitimollik, sharlti ehitimollik, Markov modeldar, birlanchi va ikkilamchi	2
5	Siliqlash algoritmlari	Laplas siliqlash algoritmi, k-qo'shish siliqlash algoritmi	2
6	Kneser-Ney siliqlash	Kneser-Ney, Interpolyatsiya qilingan algortimi	2
7	Logistik regression modellar. Generativ model	Regressiya, logistic regressiya, generative model	2
8	Sigmoida funksiyasi va sentiment klassifikatsiya	Sigmoida funksiyasi, sentiment klassifikatsiya, logistic funksiyasi, multinomial regressiya	2
9	Kosinuslar teoremasi yordamida so'zlar o'rtaida o'xshashliklarni aniqlash.TF-IDF	O'xshashlik, kosinuslar teoremasi, semantic vektorlar. TF-IDF	2
10	TF-idf yoki PPMI vector modellarining ilovalari	TFidif, PPMI (positive pointwise mutual information), vektor	2
11	Neyron til modelari	Neyron, neyron tarmoqlar, til, model, neyron til modelari	2
12	Nomlangan obyektlar va nomlangan obyektlarning teglari	Obyektlar, nomlangan obyektlar, teglash	2
13	So'z turkumlarini teglash usullari: Markov zanjiri, Markov modelari, Yashirin Markov modeli	So'z turkumlar, teglash, metod, Markov zanjiri, Markov modelari, Yashirin Markov modeli	2

Yashirin Markov modelining teglovchi komponentlari. YMM yordamida dekodlash	Markov modeli, komponentlari, decode, encode	2
Viterbi algoritimi	Viterbi algoritimi, shartli ehtiymollik	2
Qoidalarga asoslangan metodlar va ularning algoritmлari	Qoidalalar, metodlar, semantika, algoritmлar	2
Shartli tasodifiy maydonlar.	NER, CRF, maydonlar, tasodifiy maydonlari	2
Nomlangan obyektlar uchun shartli tasodifiy maydonlar	Neyron tarmoqlar, recurrent, Encode, decode, recurrent neyron tarmoqlari	2
Rekurrent neyron tarmoqlar Encoder-decoder modeli	Mashina tajimasi tizimini qurishdagi amaly	2
O'zbek til uchun grammatic qoidalalar	Mashina tarjinmasi, tokenizatsiya, MT tavsilotlar	2
O'zbek tilidagi gap modellarining ijerasiyalari ishlab chiqish	Grammatik qoidalalar, ot brikmlar, fe'l Gap, so'z brikmlari, ijerasiya	2
CKY tahili: Dinamik dasturlash yondoshuvlari	Xomskiy normal formasi, parsing, chunking algoritmлari	2
Bog'iqlik munosabatlari va formalizmlari	So'zlar orasidagi bog'iqlik munosabatlari, formalizmlar, bog'iqlik daraxtlari	2
Bog'iqlikni tahlii qilishga asoslangan transitsiya	Transitsiya, tahlii qilish, so'zlarining bir birig bog'iqligini tahlii qilishi	2
Mantiqiy ifodalalar uchun hisoblash talablari. Nazariy semantik model	Aniq ifodalalar, tekshirish, kanonik shakl, mantiqiy va mantiqiy bo'llmagan lug'atlar, haqiqiy shartli semantika	2
Brinchchi tartibli mantiq va uning asosiy elementlari	Brinchchi tartibli mantiq, funksiya, const, o'zgaruvchi, mijor ko'rsatkichlari, lambda notatsiyasi	2
So'z ma'nolari orasidagi munosabatlari	Sinonim, ononim, antonym, paronym, meronym, giponim, polisemantik so'zlar, polifunktional so'zlar	2
Ontologiyalar	Wordnet, leksik munosabatlari, ma'lumotlar bazasi	2
30 munosabatlarning ma'lumotlar bazasi	Jami:	60 s.

3. Mustaqil ta'lim ishlari

Nº	Ma'ruba navzulari	Qaysi o'quv mashg'ulot turiga mos kelishi	Qisqacha tafsifi (kalit so'zlar)	Soattar hajmi
1	Yirik til modelari va ahmoqona qaytish	amatly	Yirik tillar, modellar, siliqlash	4
2	Matlarni normallashtirish va masofalarini minimal o'gartirish	amatly	Matn, matlarni normallashtirish, masofa, matnlar orasidagi masofalar	4
3	Regulyarlashtirish	amatly	Regulyarlashtirish, modellar, til modelari	4
4	Multinomial logistic regressiya va modelarni sharhlash(interpreting)	amatly	Multinomial, logistik regressiya, regression modellar	4
5	O'rnatishlarning semantik xususiyatlari	amatly	O'matishlar, semantik xususiyatlar, parallelogram modeli	4
6	Neyron til modelari	amatly	Neyron tarmoq, neyron til, neyron til modelari	4
7	Til modellarining potensial zararları	amatly	Til modelari, potensial zararlar	4
8	Grammatik ekvivalentlik va normal forma	amatly	Ekvivalent va normal forma	4
9	Leksikallashtirilgan grammatikalar	amatly	Leksikallashtirilgan grammatikalar	4
10	Tematic modellashtirish	amatly	IDF, TF-IDF, Okapi BM25	4
11	TF-IDF vektorlari va temmatizatsiya	amatly	TF-IDF, vektorlar, lemma, temmatizatsiya	4
12	Singulyar (yagona) qiymatlar	amatly	U, S, V ^T -singulyar vektorlari	4
13	Asosiy component usuli	amatly	PCA, vektor, uch o'lchovli vektor.	4
14	CKY tahili: Dinamik dasturlash yondoshuvlari	amatly	Xomskiy normal formasi, parsing, chunking algoritmлari	4
15	So'z vektorlari	amatly	Vektor, so'z vektorlari	4
16	Skip-gamma	amatly	Skip-gamma, semantika, vektorlar	4
17	Ko'p o'lchovli mantiqiy funkisiylar	amatly	Vektor, ko'p o'lchovli fazo, mantiqiy funkisiylar	4
18	Word2vec hisoblash	amatly	Word2vec hisoblashni amalga oshirish	4

19	Semantik qidiruv	amaliy	Semantik qidiruv, semantik analiz	4
20	Latent semantik tahlil	amaliy	Latent, semantik tahlil, omonim, ko'p ma'nolilik	4
21	Semantik rollarning belgilovchi neyron algoritmlari	amaliy	Semantik rollar, neyron tarmoqlar, neyron algoritmlari	10
			Jami:	90 s.

6. FANNING KURS ISHI MAZMUNI

Algoritmlar nazariyasi o'quv kursini o'zlashtirish davomida kurs ishi rejulashtirilmagan.

7. O'QUV ADABIYOTLARI VA AXBOROT MANBALARI

- 1 Daniel Jurafsky, James H. Martin. Speech and Language Processing. Draft of December 30, 2020.
- 2 Хобсон Лейн, Ханнес Хапке, Коул Ховард. Обработка естественного языка в действии. — СПб.: Питер, 2020. — 576 с.: ил. — (Серия «Для профессионалов»). ISBN 978-5-4461-1371-2
- 3 Natural Language Processing. Copyright 2019 by Tutorials Point (I) Pvt. Ltd.
- 4 Anita Wasilewska. Natural Language Processing
- 5 Ann Copestake. Natural Language Processing. 2004, 8
- 6 R. Kibble. Introduction to natural language processing. University of London C. Department of Computing, Goldsmiths 2013

8. O'QUV DASTURI HAQIDA MA'LUMOT

O'quv dasturi Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti tomonidan ishlab chiqilgan va universitet Kengashining 202 ___ yil “___” “___” dagi “___” -sonli bayoni bilan tasdiqlangan.

“Kompyuter lingvistikasi va raqamli texnologiyalar” kafedrasи mudiri, t.f.f.d, dotsent B.B.Elov _____ (imzo)

Tuzuvchi:

B.B.Elov _____ (imzo)

“Kompyuter lingvistikasi va raqamli texnologiyalar” kafedrasи dotsenti, PhD.

Taqrizchilar:

TATU “Sun'iy intellekt kafedrasи” kafedrasи dotsenti, t.f.d.

M.F.Rahimov _____ (imzo)