

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

O'ZBEKISTON DAVLAT JAHON TILLARI UNIVERSITETI



**NLP ALGORITMLARI FANIDAN
SILLABUS
(Kunduzgi ta'limg)**

Bilim sohasi:	200000	—	San'at va gumanitar fanlar
Ta'lim sohasi:	230000	—	Tillar
Ta'lim yo'nalishi:	60230400	—	Kompyuter lingvistikasi

Toshkent - 2025



**“NLP algoritmlari” sillabusi
Kompyuter lingvistikasi ta’lim yo‘nalishi
(Kunduzgi ta’lim)**

Fan nomi:	NLP algoritmlari
Fan turi	Majburiy
Fan kodi:	NA 1406
Yil:	2025-2026
Semestr:	4
Ta’lim shakli:	Kunduzgi
Mashg’ulotlar shakli va semestrga ajratilgan soatlar:	180
Ma’ruza	36
Amaliy mashg’uloti	-
Labaratoriya	-
Mustaqil ta’lim	108
Seminar	36
Kredit miqdori:	6
Baholash shakli:	Test sinov va imtihon
Fan tili:	o‘zbek/rus
Fanning maqsadi (FM)	
FM1	NLP (Natural Language Processing) algoritmlari fanining asosiy maqsadi - talabalarga tabiiy tilni qayta ishlash usullari va algoritmlarini o‘rgatishdir. Bu maqsadga erishish orqali talabalar inson va kompyuter o‘rtasida samarali til aloqasini tashkil etuvchi algoritmlarni ishlab chiqish va amaliyotda qo‘llash ko‘nikmalariga ega bo‘ladilar. Shuningdek, ularni tabiiy til matnlaridan foydali ma’lumotlarni olish, tahlil qilish va qo‘llash uchun zamonaviy texnologiyalar bilan tanishtirish.

Fanni o‘zlashtirish uchun zarur boshlang‘ich bilimlar	
1	Informatika
2	Kompyuterdan foydalanish ko‘nikmalar

Ta’lim natijalari (TN)	
Bilimlar jihatidan:	
TN1	NLP asoslari fani bo‘yicha bilimga ega bo‘lishlari kerak;
TN2	NLP asoslari bo‘yicha nazariy bilimlarga ega bo‘lishlari kerak;
TN3	Amaliy dasturlardan foydalanish bo‘yicha bilimga ega bo‘lishlari kerak;
Ko‘nikmalar jihatidan:	
TN4	NLP asoslari fani to‘g‘risidagi nazaryalarni biladi;
TN5	NLP asoslari va algoritmlari bo‘yicha boshlang‘ich ko‘nikmaga ega bo‘ladi;

TN6

NLP algoritmlarini ishlab chiqish va ularni amaliyotda qo'llash bo'yicha ko'nikmalar hosil bo'ladi;

Fan mazmuni

Mashg'ulotlar shakli: ma'uza (M)

M1	Kirish va NLP algoritmlari asoslari.
M2	Matn manbalari va korpus yaratish
M3	Matnni tayyorlash: tokenlash, normalizatsiya
M4	Til modellari: N-gram, perplexity
M5	Stemming, lemmatizatsiya.
M6	So'z turkumlarini belgilash (POS tagging) va nomuhim so'zlar (stop-words).
M7	Matn xususiyatlari va TF-IDF modeli.
M8	Matn o'xshashligi.
M9	NER va sentiment tahlili.
M10	Matn turlarini aniqlash, mashinali o'qitish va neyron tarmoqlar.
M11	Mavzuni modellashtirish (Topic modeling) asoslari.
M12	Stilometriya va matn uslubi.
M13	So'z vektorlari va klasterlash.
M14	Vektorli tahlil va matn semantikasi.
M15	Sintaktik bog'liqlik va analiz.
M16	Text Summarization va Transformer modellariga kirish.
M17	Matnni tuzatish texnikasi.
M18	Mashinali tarjima algoritmlari.

Mashg'ulotlar shakli: seminar mashg'ulot (S)

S1	Tabiiy tilni qayta ishlash (NLP) algoritmlariga kirish: Asosiy tushunchalar va ilmiy rivojlanish tarixi.
S2	Matn manbalari va korpuslar: Korpus lingvistikasi asoslari va korpus tuzish tamoyillari.
S3	Tokenlash va matn normalizatsiyasi: Kompyuter lingvistikadagi o'mni va muammolari.
S4	Til modellari va ehtimollik yondashuvi: N-gram modellar va Perplexity ko'rsatkichining tahlili.
S5	So'z shakllarini soddalashtirish: Stemming va Lemmatizatsiya nazariyasi.
S6	So'z turkumlarini avtomatik belgilash (POS tagging) va nomuhim so'zlar bilan ishlash nazariyasi
S7	Matn xususiyatlarini aniqlash va TF-IDF modelining matematik asoslari.
S8	Matnlar o'xshashligini tahlil qilish: Lingvistik va matematik yondashuvlar.
S9	Nomlangan entitetlarni aniqlash (NER) va matn kayfiyatini baholash (Sentiment Analysis) metodologiyasi.
S10	Matn turlarini aniqlash va tasniflash: Mashinali o'rghanish va neyron modellar nazariyasi.
S11	Mavzu modellashtirish (Topic Modeling): Lingvistik tahlil va ehtimollik modellar.
S12	Stilometriya: Matn uslubini statistik va lingvistik metodlar bilan aniqlash.

S13	So‘z vektorlari: Semantik makon va vektor ifodalash nazariyasi (Word2Vec, GloVe).
S14	Matn semantikasi va vektorli tahlil: Semantik sathdagi tahlil yondashuvlari.
S15	Sintaktik bog‘liqlik va sintaktik tahlil: Konstitutsiya va bog‘liqlik grammatickasi nazariyasi.
S16	Matn qisqartirish (Summarization) va Transformer arxitekturalari: Asosiy tushunchalar va lingvistik asoslari.
S17	Matnni avtomatik tuzatish texnikalari: Tabiiy tilda xatolarni aniqlash va tuzatish nazariyalari.
S18	Mashinali tarjima algoritmlari: Neyron tarjima tizimlari va til juftliklari o‘rtasidagi farqlar nazariyasi.

Mustaqil ta’lim (MT)		
Mustaqil ta’lim uchun berilgan mavzularga tayorgarlik ko’rish va berilgan vazifalarni bajarish		90 soat
1. Tabiiy tilni qayta ishlash (NLP) algoritmlarining rivojlanish tarixi va sun’iy intellektdagi o‘rni.	2 soat	
2. Korpus lingvistikasi va matn manbalarini tayyorlash tamoyillari.	2 soat	
3. Web-ma’lumotlarni yig‘ish va korpus tuzishda API va Web scraping texnikalarining nazariyasi asoslari.	2 soat	
4. Tokenlash va matnni normalizatsiya qilish: nazariy va lingvistik tamoyillarini taqqoslash.	2 soat	
5. Tokenizatsiya metodlari: Qoidalarga asoslangan va statistik yondashuvlarning qiyosiy tahlili.	2 soat	
6. N-gram modellar va perplexity: ehtimollik asosidagi til modellari nazariyasi.	2 soat	
7. Markov modellar: Kompyuter lingvistikada ehtimollik asosida ketma-ketliklarni modellashtirish.	2 soat	
8. So‘z shakllarini normallashtirish: stemming va lemmatizatsiya metodlarining nazariy tahli.	2 soat	
9. So‘z turkumlarini belgilash (POS tagging) va nomuhim so‘zlar bilan ishslash usullari.	2 soat	
10. Matn ifodalanishi: Bag-of-Words va TF-IDF modellarining matematik asoslari.	4 soat	
11. Nomlangan entitetlarni aniqlash (NER): Lingvistik belgilar va entitet modellashtirish	4 soat	
12. Sentiment tahlil nazariyasi: Matn kayfiyatining statistik va semantik baholash metodlari..	4 soat	
13. POS tagging va NER uchun Sequence Labeling: HMM va CRF yondashuvlari nazariyasi.	4 soat	
14. Matnni mashinali o‘qitish asosida tasniflash: Naive Bayes, Logistik Regressiya va SVM metodlarining qiyosiy tahlili.	4 soat	
15. Matnni tasniflashni baholash mezonlari: Confusion Matrix, Accuracy, Precision, Recall va F1 Score nazariyasi.	4 soat	

16.	Sequential Data uchun RNN, LSTM va GRU modellarining arxitektura nazariyasi.	4 soat
17.	Ketma-ketlik asosida matn tasniflash (Sequence classification)ning ilmiy asoslari.	4 soat
18.	Topic Modeling asoslari: LDA va Probabilistik mavzu modellarining tahliliy o'rganilishi.	4 soat
19.	Stilometriya: Matn uslubini aniqlash uchun statistik va klasterlash yondashuvlari (PCA, k-means)..	4 soat
20.	Word2Vec va Doc2Vec modellar: Semantik vektor makonini nazariy asosda o'rganish.	4 soat
21.	Sintaktik parsing: Dependency parsing va Constituency parsing nazariy farqlari va qo'llanilishi.	4 soat
22.	Matn o'xshashligini baholovchi metrikalar: Cosine Similarity, Jaccard, DTW tahliliy taqqoslash.	4 soat
23.	Matn qisqartirish (Summarization) metodlari: Extractive vs Abstractive usullarning nazariy taqqoslash.	4 soat
24.	Transformer arxitekturasi: Self-Attention va Encoder-Decoder modellarining tahliliy ko'rib chiqilishi..	4 soat
25.	Matnni avtomatik tuzatish nazariyasi: Grammatik va imlo xatolarini aniqlash algoritmlarini lingvistik tahlili..	4 soat
26.	Mashinali tarjima tizimlari: SMT va NMT modellarining nazariy tahlili va O'zbek tiliga moslashtirish.	4 soat
27.	BERT va GPT modellarining NLPdagi o'rni: Fine-tuning va transfer learning nazariyasi asosida tahlil.	4 soat

**O'quv fani o'qitilishi bo'yicha uslubiy ko'rsatmalar
Fan bo'yicha talabalar bilimini baholash va nazorat qilish mezonlari**

Yozma shakldagi nazoratda baholash mezonlari	
5 – “a'lo”	
Tavsif	Talabaning ko'rsatkichlari
Berilgan savol va topshiriq to'liq va aniq bajarilgan.	<ul style="list-style-type: none"> - NLP mavzulari bo'yicha savol va topshiriqlar yuzasidan mustaqil mushohada yuritgan, tanqidiy nuqtayi nazarini ifodalagan. - NLP algoritmlari va texnologiyalari mohiyatini ijodiy yondashuv asosida to'liq ochib bergan. - Nazariy va amaliy tushunchalarni birlashtirib, grammatick va uslubiy xatolarga yo'l qo'yмаган.
4 – “yaxshi”	
Tavsif	Talabaning ko'rsatkichlari

Berilgan savol va topshiriqlarni bajarishda xatolar mavjud.	<ul style="list-style-type: none"> - NLP algoritmlari bo'yicha berilgan savol va topshiriqlar mohiyatini to'la yoritmagan. - Savol va topshiriqlarning aksariyat qismini to'g'ri bajargan, ammo izohlashda muayyan xatolik va chalkashliklarga yo'l qo'ygan. - Grammatik va uslubiy xatolar bilan bajargan.
---	--

3 – “qoniqarli”

Tavsif	Talabaning ko'rsatkichlari
Berilgan savol va topshiriqni bajarishda qo'pol xatolarga yo'l qo'yilgan, savollarga mujmal va chalkash javoblar berilgan, lekin asosiy fikr ifodalangan.	<ul style="list-style-type: none"> - NLP algoritmlari yuzasidan yuzaki bilinga ega. - Savol va topshiriqlarga mujmal va chalkash javob bergen. - NLPga oid asosiy tushunchalarini kamchiliklar bilan bajargan. - Qo'pol grammatik yoki uslubiy xatolar mavjud.

2 – “qoniqarsiz”

Tavsif	Talabaning ko'rsatkichlari
Berilgan savol va topshiriq bajarilmagan, savol va topshiriqlar mohiyatini tushunmagan	<ul style="list-style-type: none"> - NLP fanining asosiy tushunchalariga ega emas. - NLP algoritmlari va metodlari bo'yicha tasavvurga ega emas. - Topshiriqlarni noto'g'ri yoki to'liq bajarmagan. - Qo'pol grammatik yoki uslubiy xatolarga yo'l qo'ygan.

Seminar mashg'ulotlar shakldagi nazoratda baholash mezoni

5 – “a'lo”

Tavsif	Talabaning ko'rsatkichlari
Seminar topshiriqlari va savollari to'liq, aniq va ijodiy tarzda bajarilgan.	<ul style="list-style-type: none"> - Seminar mavzusi (masalan, korpus tuzish, til modellari, POS tagging, NER, Transformerlar va h.k.) bo'yicha mustaqil fikr bildirgan, tanqidiy va ijodiy yondashuv qo'llagan. - Muammoni nazariy va amaliy jihatdan chuqr yoritib bergen. - Matnni tushunarli, mantiqiy izchil, grammatik va uslubiy xatolarsiz taqdim etgan.

4 – “yaxshi”

Tavsif	Talabaning ko'rsatkichlari
--------	----------------------------

Seminar topshiriqlarini bajarishda ayrim xatolar mavjud.	<ul style="list-style-type: none"> - Seminar mavzulari mazmunini to‘liq ochib bermagan, ayrim muhim jihatlarni e’tibordan chetda qoldirgan. - Savol va topshiriqlarning asosiy qismini to‘g’ri bajargan, ammo tushuntirishlarda chalkashliklarga yo‘l qo‘ygan. - Ta’rif va misollarni berishda grammatic va uslubiy kamchiliklarga yo‘l qo‘ygan.
--	---

3 – “qoniqarli”

Tavsif	Talabaning ko‘rsatkichlari
Seminar savollari yuzasidan mujmal va chalkash javob bergen, ammo asosiy mazmunni qisman tushungan.	<ul style="list-style-type: none"> - NLP bo‘yicha yuzaki bilim ko‘rsatgan. - Seminar topshiriqlarini bajargan, lekin ko‘plab kamchiliklar bilan. - Savollarga umumiy, yuzaki javob qaytargan. - Qo‘pol grammatic yoki uslubiy xatolar mavjud.

2 - “qoniqarsiz”

Tavsif	Talabaning ko‘rsatkichlari
Seminar savollari va topshiriqlari bajarilmagan yoki mutlaqo noto‘g’ri bajarilgan.	<ul style="list-style-type: none"> - Seminar mavzulari bo‘yicha tushunchaga ega emas. - NLP algoritmlariga oid asosiy bilim va ko‘nikmalarni namoyon etmagan. - Berilgan topshiriqlarni noto‘g’ri yoki to‘liq bajarmagan. - Qo‘pol grammatic va uslubiy xatolarga yo‘l qo‘ygan.

MUSTAQIL TA’LIMNI BAHOLASH MEZONLARI

5 – “a’lo”

Tavsif

Mustaqil ta’lim topshiriqlari to‘liq, aniq va ijodiy yondashuvda bajarilgan.

Talabaning ko‘rsatkichlari

- Mustaqil ta’lim mavzusi (masalan, N-gram modellar, Tokenlash usullari, Word2Vec, Transformer arxitekturasi, BERT va GPT modellar va h.k.) bo‘yicha mustaqil izlanish olib borgan.
- Berilgan topshiriqni ijodiy yondashuv asosida to‘liq yoritib bergen.
- Ilmiy adabiyotlar, onlayn manbalar yoki dasturlash vositalaridan foydalanganini ko‘rsatgan.
- Javoblar grammatic va uslubiy jihatdan toza, tushunarli va mantiqiy izchil bo‘lgan.

4 – “yaxshi”

Tavsif

Mustaqil ta’lim topshiriqlarini bajarishda ayrim xatolar mavjud.

Talabaning ko‘rsatkichlari

- Mavzuni yoritishda muhim jihatlarni qisman e’tibordan chetda qoldirgan.
- Topshiriqlarni asosan to‘g‘ri bajargan, ammo tushuntirishlarda chalkashliklarga yo‘l qo‘ygan.
- Amaliy yoki nazariy asoslarni yetarlicha chuqur tahlil qilmagan.
- Grammatik va uslubiy kamchiliklarga yo‘l qo‘ygan.

3 – “qoniqarli”

Tavsif

Mustaqil ta’lim topshiriqlarida mujmal va yuzaki yondashuv mavjud, ammo asosiy fikr qisman olib berilgan.

Talabaning ko‘rsatkichlari

- NLP bo‘yicha yuzaki bilim namoyon qilgan.
- Berilgan mustaqil ta’lim topshiriqlarining asosiy qismini kamchiliklar bilan bajargan.
- Javoblar mujmal, chalkash va yetarlicha izchil bo‘lmagan.
- Qo‘pol grammatik yoki uslubiy xatolar mavjud.

2 - “qoniqarsiz”

Tavsif

Mustaqil ta’lim topshiriqlari bajarilmagan yoki mutlaqo noto‘g‘ri bajarilgan.

Talabaning ko‘rsatkichlari

- Mustaqil ta’lim mavzulari bo‘yicha tushunchaga ega emas.
- Topshiriqlarni noto‘g‘ri yoki to‘liq bajarmagan.
- Amaliyat va nazariyani bog‘lay olmagan.
- Qo‘pol grammatik va uslubiy xatolarga yo‘l qo‘ygan.

Baholash mezonlari

O'zbekiston tizimi	Evropa kredit transfer tizimi (ECTS — European Credit Transfer System)	O'zbekiston tizimi
«5» (a'llo)	«A»	90 — 100
«4» (yaxshi)	«B»	70 — 89
	«C»	
«3» (qoniqarli)	«D»	60 — 69
	«E»	
«2» (qoniqarsiz)	«F»	0 — 59

Asosiy adabiyotlar

- | | |
|---|---|
| 1 | Jurafsky, D., & Martin, J. H. <i>Speech and Language Processing</i> (3rd Edition Draft). — Stanford: Stanford University, 2024/2025. — 800 p. |
| 2 | Vasiliev Y. — <i>Обработка естественного языка. Python и spaCy на практике</i> , SPb: Piter, 2021. |

Qo'shimcha adabiyotlar

- | | |
|---|---|
| 1 | Stanford University CS224n: Natural Language Processing with Deep Learning
http://web.stanford.edu/class/cs224n/ |
| 2 | Большакова Е.И., Клышинский Э.С. и др. Автоматическая обработка текстов на естественном языке и компьютерная лингвистика: учеб. пособие. — Москва: МИЭМ. 2011. — 272 с. |
| 3 | Yoav Goldberg. Neural Network Methods in Natural Language Processing. Morgan & Claypool Publishers. (2017). ISBN: 978-1681732446 |
| 4 | Bender, E. M. Linguistic Fundamentals for Natural Language Processing: 100 Essentials from Morphology and Syntax. Morgan & Claypool. (2013). ISBN: 978-1608454422 |
| 5 | Леонтьева Н.Н. Автоматическое понимание текста: системы, модели, ресурсы: учеб. пособие. — Москва: Академия, 2006. — 304 с. |

Delip Rao & Brian McMahan. Natural Language Processing with PyTorch: Build Intelligent Language Applications Using Deep Learning. O'Reilly Media. (2019). ISBN: 978-1491978238

Fan o'qituvchisi to'g'risida ma'lumot

Muallif:	Mallayev Ravshan Qo'ziboyevich, pedagogika fanlari bo'yicha falsafa doktori (PhD)
E-mail:	ravshan.mallayev@ru
Tashkilot:	O'zbekiston Davlat jahon tillari universiteti, "Zamonaviy axborot texnologiyalari" kafedrasi
Taqrizchilar:	1. B.Axrarov – O'zDJTU, Zamonaviy axborot texnologiyalari kafedrasi mudiri, pedagogika fanlari bo'yicha PhD, dotsent 2.A.Haydarov – M.Ulug'bek nomidagi O'zbekiston Milliy universiteti "Amaliy matematika va kompyuter tahlili" kafedrasi dotsenti, f.-m.f.n., dotsent.

Mazkur sillabus universitet o'quv-uslubiy Kengashining 2025 yil 28-avgustdagি 1- sonli yig'ilish bayoni bilan tasdiqlangan.

Mazkur Sillabus "Zamonaviy axborot texnologiyalari" kafedrasining 2025- yil 29 avgustdagи 1- sonli yig'ilishi bayoni bilan maqullangan

Sharq filologiyasi fakulteti dekani

Zamonaviy axborot texnologiyalari
kafedrasi mudiri

Tuzuvchi:

Zamonaviy axborot
texnologiyalari kafedrasi mudiri:



wish✓