

**O‘ZBEKISTON RESPUBLIKASI  
OLIV TA‘LIM, FAN VA INNOVATSIYALAR VAZIRLIGI  
CHIRCHIQ DAVLAT PEDAGOGIKA UNIVERSITETI**



**AXBOROTLARNI IZLASH VA AJRATIB OLISH  
O‘QUV DASTURI**

Bilim sohasi:	600000 – Axborot-kommunikatsiya texnologiyalari
Ta‘lim sohasi:	610000 – Axborot-kommunikatsiya texnologiyalari
Mutaxassislik:	70610105 – Ta‘limda axborot texnologiyalari

Fan/modul kodi AIA1104		O'quv yili 2024-2025	Semestr 1	ECTS - Kreditlar 4	
Fan/modul turi Majburiy		Ta'lim tili O'zbek/rus		Haftadagi dars soatlari 4	
1.	Fanning nomi	Auditoriya mashg'ulotlari (soat)		Mustaqil ta'lim (soat)	Jami yuklama (soat)
	Algoritmlarni loyihalashtirish va tahlil qilish	80		40	120

2. **I. Fanning mazmuni.**

Fanni o'qitishdan maqsad –Axborotni izlash va ajratib olish jarayonlari, axborot olish usullarini takomillashtirish, axborotni izlash algoritmlarini ishlab chiqish, zamonaviy veb-qidiruv tizimlarining algoritmlari asosida o'zbek tili lingvistikasini ishlab chiqishning nazariy va amaliy bilimlarini hamda ulami amaliyotga tadbiq etishdan iborat.

Fanning vazifasi - Axborotni ajratish va izlash tizimlarini loyihalash, axborot izlash tizimlarini boshqarishning nazariy asoslarini shakllantirish, matn modullarining tuzilish usullari va algoritmlari, axborotni ajratish va izlash konsepsiyasini texnik jihatdan amalga oshirish bo'yicha yondashuvlar hamda boshqa turdagi ma'lumot manbalaridan axborot izlash masalalarini amaliy ahamiyatini yoritib berishdan iborat.

**II. Nazariy qism (ma'ruza mashg'ulotlari)**

**III. Fan tarkibiga quyidagi mavzular kiradi:**

**1-mavzu. Axborotlarni izlash va ajratib olish tushunchalari.**

Axborotlarni izlash va ajratish haqida asosiy tushunchalar. Axborotlarni ajratish va izlashning nazariy mezonlari. Axborot izlash strategiyasini loyihalash bosqichlari. Axborotni ajratish prinsiplari.

**2-mavzu. Mantiqiy qidiruv asoslari.**

Axborot olish muammosi, invert indeksni yaratish, mantiqiy so'rovlarni qayta ishlash, kengaytirilgan mantiqiy model yaratish usuli.

**3-mavzu. Lug'at va e'lonlar ro'yxatini tashkil qilish asoslari.**

Hujjat chegarasi va belgilar ketma-ketligini dekodlash, ketma-ketligini olish, birligini tanlash, terminlarning so'z boyligini aniqlash, tokenizatsiya, umumiy atamalarni tashlash: so'zlarni to'xtatish, normalizatsiya (terminlarning ekvivalentlik tasnifi), urg'u va diakritika, stemming va lemmatizatsiya, o'tkazib yuborish ko'rsatkichlari orqali tezroq e'lonlar ro'yxatini kesishish, pozitsion xabarlar va iboralar so'rovlari, ikki so'zli so'rov indekslari, Pozitsion indekslar va ularning hajmi.

#### **4-mavzu. Lug'atlarni qayta tiklash asoslari.**

Lug'atlar uchun tuzilmalarni qidirish, joker belgili so'rovlar, vaqtinchalik indekslar, joker belgilar so'rovlari uchun  $\kappa$ -gram indekslari, imlo tuzatish va uning shakllari, masofani tahrirlash, imloni tuzatish uchun  $k$ -gramm.

#### **5-mavzu. Axborot izlash tizimlarida indekslar.**

Bloklangan tartib asosida indekslash, xotirada bir martalik indekslash, tarqalgan indekslash, dinamik indekslash, boshqa turdagi indekslar.

#### **6-mavzu. Axborot izlash tizimlarida indekslarni siqish.**

Axborot izlashda atamalarning statistik xususiyatlari, Heaps qonuni: atamalar sonini hisoblash, Zipf qonuni: atamalarning taqsimlanishini modellashtirish, lug'atni siqish, lug'at qator sifatida, bloklangan xotira, e'lonlar faylini siqish, o'zgaruvchan bayt kodlari, Gamma kodlari.

#### **7-mavzu. Muddatli vazn va vektor makon modeli.**

Parametrik va zona indekslari, og'irlangan zona reytingi, og'irliklarni o'rganish, optimal vazn, terminning chastotasi va vazni, teskari hujjat chastotasi, Tf-idf vazni, skorlash uchun vektor fazo modeli, so'rovlar vektor sifatida, vektor ballarini hisoblash, variant tf-idf funktsiyalari, sublinear tf masshtablash, maksimal tf normalizatsiyasi, hujjat va so'rovlarni tortish sxemalari, normalangan hujjat uzunligi.

#### **8-mavzu. Qidiruv tizimida so'rovlar operatorining o'zaro ta'siri asoslari.**

Samarali ball va reyting, noaniq yuqori  $K$  hujjatni olish, indeksni yo'q qilish, statik sifat ballari va buyurtma, ta'sir tartibi, klastemi kesish, axborot-qidiruv tizimining tarkibiy qismlari, darajali indekslar, so'rov muddatining yaqinligi, tahlil qilish va baholash funktsiyalarini loyihalash, vektor fazosini baholash va so'rovlar operatorining o'zaro ta'siri.

#### **9-mavzu. Axborot-qidiruv tizimini baholash usullari.**

Axborot-qidiruv tizimini baholash, standart test to'plamlari, reytingsiz qidirish to'plamlarini baholash, reytingli qidiruv natijalarini baholash, muvofiqlikni baholash, kengroq istiqbol: tizim sifati va foydalanuvchi yordami, tizim muammolari, joylashtirilgan tizimni takomillashtirish.

#### **10-mavzu. So'rovlarni kengaytirish asoslari.**

Tegishli fikr-mulohazalar va psevd-mulohazalar, tegishli fikr-mulohazalar uchun Rocchio algoritmi, ehtimoliy ahamiyatga ega bo'lgan fikr-mulohazalar, tegishli aloqa strategiyalarini baholash, psevd-mulohaza bildirish, bilvosita aloqadorlik haqida fikr-mulohazalar. So'rovlarni qayta shakllantirishning global usullari, so'rovlarni qayta shakllantirish uchun lug'at vositalari, so'rovni kengaytirish.

#### **11-mavzu. XML qidiruv asoslari.**

Asosiy XML tushunchalari, XML-ni qidirishdagi qiyinchiliklar, XML qidirish uchun vektor fazo modeli, XML qidirishni baholash, matnga asoslangan va ma'lumotlarga asoslangan XML qidiruvi.

#### **12-mavzu. Ehtimoliy ma'lumot olish asoslari.**

Asosiy ehtimollar nazariyasini ko'rib chiqish, ehtimollik reytingi printsiplari, 1/0 yo'qotish holati, so'rov shartlari uchun reyting funktsiyasi, nazariy jihatdan ehtimollik taxminlari, tegishli fikr-mulohazalarga ehtimoliy yondashuvlar, baholash va ba'zi kengaytmalar, ehtimoliy modellarni baholash, terminlar orasidagi daraxt tuzilmali bog'liqliklar.

#### **13-mavzu. Ma'lumot olish uchun til modellari.**

Til modellari, cheklangan avtomatlar vatil modellari, til modellarining turlari, so'zlar bo'yicha multinomial taqsimotlar, so'rovlar ehtimoli modeli, IRda so'rovlar ehtimolliги tili modellaridan foydalanish, so'rovlarni yaratish ehtimolini baholash, Ponte va Kroft tajribalari, tilni modellashtirish, kengaytirilgan tilni modellashtirish yondashuvlari

#### **14-mavzu. Matn tasnifi va Naive Bayes.**

Matn tasniflash muammosi, Naive Bayes matn tasnifi, Bernoulli modeli, Naive Bayesning xususiyatlari, Multinomial modelning variant, xususiyatlarni tanlash, o'zaro ma'lumot xususiyatlarini tanlash usuli sifatida baholash, X-kvadratni xususiyatni tanlash usuli, chastotaga asoslangan xususiyatni tanlash, bir nechta tasniflagichlar uchun xususiyat tanlash, xususiyatlarni tanlash usullarini taqqoslash, matn tasnifini baholash.

#### **15-mavzu. Hujjat ko'rinishlari va vektor fazolardagi bog'liqlik o'chovlari.**

Hujjat ko'rinishlari va vektor fazolardagi bog'liqlik o'chovlari, Rocchio tasnifi, chiziqli va chiziqli bog'lanmagan tasniflagichlar, ikkitadan ortiq sinflar bilan tasniflash, tarafsizlik-variantlik almashinuvi.

**16-mavzu. Hujjatlarni indekslashda vektor mashinalaridan foydalanish.**

Vektorli mashinalarni qo'llab-quvvatlash: chiziqli ajratiladigan holat, SVM modeliga kengaytmalar, ko'p sinfli SVMlar, nochiziqli SVMlar, eksperimental natijalar, matnli hujjatlarni tasniflash masalalari, klassifikator ish faoliyatini yaxshilash, katta va qiyin toifali taksonomiyalar, matn uchun xususiyatlar, matn tasnifidagi hujjat zonalari, hujjatlarni yuklash zonalari, hujjat zonalari uchun alohida bo'shliqlar, matnni umumlashtirishga ulanishlar. Ad hoc ma'lumot olishda mashinani o'rganish usullari, mashinada o'rganilgan ballarni hisoblash.

### **17-mavzu. Axborot izlashda klasterlash.**

Axborot izlashda klasterlash, muammo bayonoti, terminologiya tushunchalari, Kardinallik - klasterlar soni, klasterlashtirishni baholash. Modelga asoslangan klasterlash.

### **18-mavzu. Ierarxik klasterlash.**

Ierarxik aglomerativ klasterlash, yagona va to'liq bo'g'inli klasterlash, HACning vaqt murakkabligi, Guruh-o'rtacha aglomerativ klasterlash, Centroid klasterlash, BoMinuvchi klasterlash.

### **19-mavzu. Yashirin semantik indekslash.**

Chiziqli algebra sharhi, matritsalar bo'inishi. Termin-hujjat matritsalar va singulyar qiymatlarning bo'inishi. Yashirin semantik indeksatsiya.

### **20-mavzu. Veb-qidiruv asoslari. Veb skanerlash va indekslar.**

Orqa fon, veb xususiyatlari, veb grafika, spam, reklama iqtisodiy model sifatida. Foydalanuvchi so'rovlari. Indeks hajmi. Brauzer xususiyatlari. Crawler arxitekturasi. Brauzerni tarqatish. DNS ruxsati. URL chegarasi. Tarqatish indeksleri. Ulanish serverlari.

## **III. Amaliy mashg'ulotlar bo'yicha ko'rsatma va tavsiyalar**

Amaliy mashg'ulotlar uchun quyidagi mavzular tavsiya etiladi:

1. Axborotlarni izlash va ajratib olish nazariy mezonlari, izlash strategiyasini loyihalash bosqichlari
2. Mantiqiy qidiruvda invert indeksni yaratish, mantiqiy so'rovlarni qayta ishlash
3. Lug'at va e'lonlar ro'yxatini tashkilda chegarasi va belgilar ketma-ketligini dekodlash, ketma-ketligini olish va terminlarning so'z boyligini aniqlash
4. Lug'atlarni qayta tiklashda qidirish, joker belgili so'rovlar, vaqtinchalik indekslar, joker belgilar
5. Axborot izlash tizimlarida bloklangan tartib, xotirada bir martalik indekslash asosida indekslash

6. Axborot izlash tizimlarida indekslarni siqishda Zipf qonuni asosida atamalarning taqsimlanishini modellashtirish, lug'atni siqish.
7. Muddatli vazn va vektor makon modeli
8. Qidiruv tizimida so'rovlar operatorining o'zaro ta'siri
9. Axborot-qidiruv tizimini standart test to'plamlari, reytingsiz qidirish to'plamlarini baholash, reytingli qidiruv natijalarini baholash
10. So'rovlarni kengaytirishda fikr-mulohazalar va psevd-mulohazalar, Rocchio algoritmi
11. XML-ni qidirishdagi qiyinchiliklar, XML qidirish uchun vektor fazo modeli.
12. Ehtimoliy ma'lumot olishda so'rov shartlari uchun reyting funktsiyasi, ehtimoliy modellarni baholash
13. Ma'lumot olish uchun til modellari. Cheklangan avtomatlar vatil modellari, til modellarining turlari, so'zlar bo'yicha multinomial taqsimotlar, so'rovlar ehtimoli modeli
14. Matn tasniflashda Naive Bayes klassifikatori
15. Hujjat ko'rinishlari va vektor fazolardagi bog'qlik o'chovlari
16. Hujjatlarni indekslashda vektor mashinalaridan foydalanish.
17. Axborot izlashda klasterlash algoritmlari
18. Ierarxik klasterlashda ikki asosiy turi yuqoridan pastga va pastdan yuqoriga
19. Yashirin semantik indekslash
20. Veb-dasturiy vositalarida qidiruv. Veb skanerlash va indekslar

### **Mustaqil ta'lim va mustaqil ishlar**

Mustaqil ta'limni baholash – bu talabalarning jamoaviy tartibda va yakka tartibda berilgan amaliy loyihalarni bajarishlari orqali amalga oshiriladi. Bunda har bir talabaga bitta jamoaviy loyiha va ikkita yakka tartibda bajariladigan loyiha beriladi. Talaba berilgan loyihaning maqsad va vazifalarini, mohiyatini tushungan holda qo'yilgan masalani o'rganib, izlanishlar olib boradi. Olingan natijalarni tahlil qilib, hulosalari bilan taqdimotlar tayyorlab himoya qiladi. Ishchi fan dasturida loyihalarning soni, mavzusi, mazmuni bajarish usullari va topshirish muddatlari to'liq ochib beriladi.

### **Mustaqil ta'lim uchun tavsiya etiladigan mavzular:**

1. Axborot qidirish: asosiy tushunchalar, tashkil etish turlari va shakllari
2. Windows operatsion tizimida ma'lumot qidirish
3. Dasturiy ta'minotni qidirish xizmatlari
4. Simli va simsiz aloqa
5. So'rovlar va qidiruv natijalarini rus tilidan ingliz tiliga va aksincha tarjima qilish;
6. So'rovlarning imloni avtomatik tekshirish;
7. Topilgan saytlar uchun qidiruv natijalarini yanada ma'lumotli ko'rsatish (nafaqat topilgan hujjatning birinchi jumlasini ko'rsatiladi);
8. Har qanday grammatik shaklda qidirish qobiliyati (bu rus tili uchun ayniqsa muhimdir);

9. Professional foydalanuvchilar uchun kuchli ilg'or so'rovlar tili;
10. Rus tili uchun beshta asosiy kod sahifalarini (turli operatsion tizimlar) qo'llab-quvvatlash;
11. Hujjatlarning URL (manzil) va sanasi bo'yicha cheklovlardan foydalangan holda qidiruv texnologiyasi;
12. Qidiruv faqat matn bo'yicha emas, balki rasmlarga sarlavhalar, sharhlar va sarlavhalar va boshqalar bo'yicha ham amalga oshiriladi;
13. Qidiruv parametrlarini va ma'lum miqdordagi oldingi foydalanuvchi so'rovlarini saqlash;
14. Turli serverlarda joylashgan hujjat nusxalarini birlashtirish.
15. Internetda qidirishning asosiy usullari.
16. Qidiruv algoritmi
17. Ochiq manbalar bilan ishlash
18. Mutaxassisdan ma'lumot olish
19. Ma'lumotlarni qayta ishlash
20. Ma'lumotlarni qayta ishlash usullari
21. Natijani taqdim etish
22. Axborot bilan ishlash uchun uchta layfxik
23. UX tadqiqoti davomida sifatli va miqdoriy ma'lumotlarning tematik tahlili
24. Sifatli ma'lumotlarni tahlil qilish bilan bog'liq muammolar

3.

**VII. Ta'lim natijalari (shakllanadigan kompetensiyalar)**

**Fanni o'zlashtirish natijasida talaba:**

- "Algoritmlarni loyihalashtirish va tahlil qilish" fanining tushunchasi, Predmet sohani tahlil qilish jarayonida axborotni izlash va axborotni ajratishning loyihalash usullarini takomillashtirish, masalani yechish uchun axborotni izlash modullarining tuzilishida axborotni ajratib olish, ekspertizadan o'tkazish, izlash va ulaming fimksional tuzilishi hamda ular asosida zamonaviy avtomatlashtirilgan axborot tizimlarida axborotlarni izlashni loyihalash bo'yicha **tasavvurga ega bo'lishi; (bilim)**

Axborotlarni izlash va shakllantirish prinsiplari, axborotlarni izlashni loyihalashtirish mezonlari va uslublarini o'rganish; axborotlarni izlashda dasturiy vositalarning ishlash modullarining tuzilishi asosida axborotni ajratish va izlashni bilish va **foydalana olishi; (ko'nikma).**

Axborot izlash tizimlarining xarakteristikalarini tahlil qilish, axborotlarni ajratish va izlashning nazariy va amaliy loyihalashtirish uslublarini o'rganishni amalga oshirish muammolari bo'yicha yechimlar qabul qilish **malakalarga ega bo'lishi kerak.**

4.

**VIII. Ta'lim texnologiyalari va metodlari:**

- ma'ruzalar;
- interfaol keys-stadilar;
- seminarlar (mantiqiy fikrlash, tezkor savol-javoblar);
- guruhlarda ishlash;
- taqdimotlarni qilish;
- individual loyihalar;

- jamoa bo'lib ishlash va hioya qilish uchun loyihalar

### IX. Kreditlarni olish uchun talablar:

5.

Fanga oid nazariy va uslubiy tushunchalarni to'la o'zlashtirish, tahlil natijalarini to'g'ri aks ettira olish, o'rganilayotgan jarayonlar va tushunchalar haqida mustaqil mushohada yuritish, joriy va oraliq nazorat shakllarida berilgan vazifa va topshiriqlarni bajarish, yakuniy nazorat bo'yicha variantlar asosida yozma topshiriqlarni bajarishi zarur.

6.

### X. Asosiy adabiyotlar:

- 1- Ф.Д.Бахромов. Ахборот технологиялари Тошкент.SHARQ"O'zFMJ"-2012
- 2- N.A.Ignatev. Berilganlarning intellektual tahlili. "MUMTOZ SO`Z" – 2018
- 3- P.A.Дадабаева. Ахборот-Коммуникатсия технологиялари ва тизимлар. Тошкент."Sano-standart"-2017 ЎзРОЎМТВ

### XI. Qo'shimcha adabiyotlar

1. Г. М. Чечин, Е. В. Положенцев, С. В. Нижникова "Поиск информации в сети Internet", Методические указания для студентов РГУ Ростов-на-Дону 2001 г.
2. В.Н.Романенко, Г.В.Никитина Сетевой информационный поиск СПБ «Профессия», 2005, 288 ст.
3. Аграновский А.В. Теоретические аспекты алгоритмов обработки и классификации речевых сигналов. 2004

### Axborot manbalari

<http://www.edu.uz>–O'zbekiston Respublikasi Oliy ta'lim, fan va innovatsiyalar vazirligi rasmiy veb sayti.

<http://www.gov.uz>– O'zbekiston Respublikasi xukumati portali.

[www.cspi.uz](http://www.cspi.uz) –Chirchiq Davlat Pedagogika Universiteti

[www.natlib.uz](http://www.natlib.uz) (A.Navoiy nomidagi O'z.MK)

[www.ziyonet.uz](http://www.ziyonet.uz) - Ziyonet axborot-ta'lim resurslari portali

7.

**Chirchiq davlat pedagogika universiteti tomonidan ishlab chiqilgan va universitet Kengashining 2024 yil " \_\_\_\_\_ " \_\_\_\_\_dagi qarori bilan tasdiqlangan**

8.

**Fan/modul uchun ma'sul:**

**A.X.Boymurodov** CHDPU, "Informatika va axborot texnologiyalari" kafedrasida v.v.b.dotsenti.

9.

**Taqrizchilar:**

**D.G'.Axmedjanov** - ChDPU Informatika o'qitish metodikasi kafedrasida mudiri, texnika fanlari doktori.

**I.A.Yuldashev**- Qarshi DU dotsenti, texnika fanlari bo'yicha (PhD)