

✓

**O‘ZBEKISTON RESPUBLIKASI**  
**OLIV TA‘LIM, FAN VA INNOVATSIYALAR VAZIRLIGI**  
**TOSHKENT DAVLAT SHARQSHUNOSLIK UNIVERSITETI**

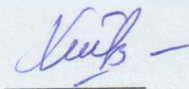


**TILSHUNOSLIKDA MATEMATIK USULLAR**  
**FANINING O‘QUV DASTURI**

**Bilim sohasi:** 200 000 - San‘at va gumanitar fanlar  
**Ta‘lim sohasi:** 230 000 – Tillar  
**Ta‘lim yo‘nalishi:** 60230400 Kompyuter lingvistikasi

**Toshkent – 2025**

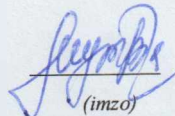
Sharq xalqlari tillari va adabiyoti  
instituti direktori:



X.V. Mirzaxmedova

(imzo)

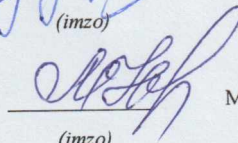
Tarjimashunoslik, tilshunoslik va  
xalqaro jurnalistika oliy maktabi boshlig'i:



S.T. Mustafayeva

(imzo)

Arm boshlig'i:



M. Yuldasheva

(imzo)

<b>Fan/modul kodi</b> TMU 1504		<b>O'quv yili</b> 2027/2028	<b>Semestr</b> 5	<b>YeCTS - Kreditlar</b> 4		
<b>Fan/modul turi</b> Majburiy		<b>Ta'lim tili</b> O'zbek		<b>Haftadagi dars soatlari</b> 4		
<b>1.</b>	<b>Fanning nomi</b>	<b>Auditoriya mashg'ulotlari (soat)</b>	<b>Ma'ruza</b>	<b>Amaliy</b>	<b>Mustaqil ta'lim (soat)</b>	<b>Jami yuklama (soat)</b>
	Tilshunoslikda matematik usullar	48	24	24	72	120
<b>2.</b>	<b>I. Fanning mazmuni:</b>					
<p><b>Fanning maqsadi:</b>  Tilshunoslikda matematik usullar fanining asosiy maqsadi – tilshunoslik jarayonlarini matematik modellar, statistik tahlil va algoritmik yondashuvlar yordamida chuqur o'rganish va tahlil qilish ko'nikmalarini shakllantirishdir. Bu fan tilshunoslikda sodir bo'ladigan murakkab jarayonlarni aniq matematik yondashuvlar orqali modellashtirishga qaratilgan bo'lib, talabalarni lingvistik ma'lumotlarni samarali tahlil qilish, tizimlashtirish va bashorat qilishga o'rgatadi.</p>						
<p><b>Fan vazifasi:</b>  Tilshunoslikdagi jarayonlarni tushinishda matematik va statistik usullarni qo'llash orqali yangi yondashuvlarni shakllantirish. Lingvistik jarayonlarni modellashtirish uchun matematik va algoritmik usullarni chuqur o'zlashtirish. Til birliklarini (so'zlar, gaplar va matnlar) statistik tahlil qilish, so'z chastotasi va ehtimolliklarini aniqlash ko'nikmalarini rivojlantirish. Lingvistik ma'lumotlarni yig'ish, qayta ishlash va ulardan samarali foydalanish uchun amaliy ko'nikmalarni shakllantirish. Lingvistik tahlil va tadqiqotlarda N-gramma modellar, Markov zanjirlari va boshqa algoritmik usullardan foydalanishni o'rgatish. Sintaktik va semantik jarayonlarni grafiklar va daraxtlar yordamida tahlil qilishni o'rgatish.</p>						
<p>Mazkur fan dasturi xalqaro tan olingan reytinglarda birinchi top 300 talik ro'yxatga kiruvchi Berlin Erkin Universiteti (Freie Universität Berlin 104 THE) xorijiy tajribasini inobatga olgan holda takomillashtirildi.</p>						
<p><a href="https://www.fu-berlin.de/studium/studienangebot/grundstaendige/informatik-kombi-lo/index.html">https://www.fu-berlin.de/studium/studienangebot/grundstaendige/informatik-kombi-lo/index.html</a></p>						

## II. Asosiy nazariy qism (ma'ruza mashg'ulotlari)

### II.1. Fan tarkibiga quyidagi mavzular kiradi:

#### **1-mavzu. Tilshunoslikda matematik usullarga kirish**

Ushbu mavzu tilshunoslikdagi murakkab jarayonlarni matematik usullar yordamida tahlil qilishning asosiy tamoyillarini tushuntiradi. Matematik va statistik yondashuvlarning tilshunoslikda tutgan o'рни, ahamiyati, shuningdek, lingvistik jarayonlarni modellashtirish, tahlil qilish va bashoratlash imkoniyatlari haqida umumiy ma'lumot beriladi. Mavzuda talabalar matematik lingvistikaga oid muhim tushunchalar bilan tanishadi va kelgusi mavzularga tayyorgarlik ko'radi.

#### **2- mavzu. Statistika va ehtimollikning asosiy tushunchalari**

Ushbu mavzu statistik va ehtimollik nazariyasining asosiy tushunchalarini o'rgatadi. Talabalar ma'lumotlarni yig'ish, tahlil qilish va tasniflashning statistik usullarini o'zlashtiradilar. Ehtimollik nazariyasi orqali tilshunoslikdagi tasodifiy jarayonlarni model qilish va ehtimoliy yondashuvlarni qo'llash imkoniyatlari haqida tushuncha hosil qilinadi. Bu mavzu, ayniqsa, til ma'lumotlarining tasniflanishi va tahlili uchun statistik va ehtimollik asoslarini o'rganishga yordam beradi.

#### **3- mavzu. Zipf qonuni va lingvistik birliklarning chastotasi**

Ushbu mavzu Zipf qonunini va uning lingvistik tahlildagi ahamiyatini o'rganadi. Zipf qonuni, odatda, so'zlar yoki til birliklarining chastotasi va ularning o'рни o'rtasidagi teskari bog'liqlikni tasvirlaydi: eng ko'p ishlatiladigan so'zlar kamdan-kam ishlatiladiganlarga nisbatan sezilarli darajada ko'proq uchraydi. Talabalar bu qonunni tilshunoslikda, xususan, so'zning chastotasini o'lchashda va matnlarni tahlil qilishda qanday qo'llash mumkinligini o'rganadilar. Mavzu, lingvistik birliklarning taqsimoti va uning statistik modellari bilan tanishishga yordam beradi.

#### **4- mavzu. Matematik modellashtirish va til birliklarini statistik tahlil qilish**

Ushbu mavzu til birliklarini matematik modellar yordamida tasvirlashni o'rgatadi. Matematik modellashtirish, tilshunoslikdagi murakkab jarayonlarni aniq va o'lchovli tarzda ifodalash imkonini beradi. Talabalar til birliklarini (so'zlar, gaplar, frazalar) matematik yondashuvlar, masalan, vektorlar, matritsalar yoki grafiklar yordamida qanday modellashtirish mumkinligini o'rganadilar. Mavzu, tilshunoslikdagi sintaktik va semantik strukturalarni matematik ravishda ifodalash va tahlil qilish imkoniyatlarini tushuntiradi.

#### **5- mavzu. N-gramma modellar, ularning qo'llanilishi, korpus lingvistikasi va matematik tahlil hamda ehtimollik va statistik usulla**

Ushbu mavzu N-gramma modellarini va ularning tilshunoslikda qanday qo'llanilishini o'rgatadi. N-gramma modeli — bu so'zlar yoki harflar ketma-ketligini o'rganish va tahlil qilish uchun ishlatiladigan statistik usul. Bu modelda "N" — ketma-ketlikdagi so'zlar yoki belgilar sonini bildiradi. Talabalar N-gramma modellarini matnlarni tahlil qilish, tilni avtomatik ravishda qayta ishlash, masalan, so'z taxminlari, tilni tarjima qilish yoki matnni tasniflash kabi sohalarda qanday

ishlatilishini o'rganadilar. Mavzu, tilshunoslikda statistik yondashuvlarning samarali qo'llanishini tushunishga yordam beradi. Korpus lingvistikasi — bu til ma'lumotlarini to'plash, saqlash va tahlil qilish uchun ishlatiladigan yondashuv bo'lib, katta hajmdagi til ma'lumotlaridan (korpuslardan) foydalanishni o'z ichiga oladi. Matematik tahlil esa bu ma'lumotlarni statistik va ehtimollik usullari orqali qayta ishlashni anglatadi. Talabalar lingvistik korpuslarni yaratish, ularni matematik usullar yordamida tahlil qilish va tilshunoslikda zamonaviy texnologiyalarni qo'llashni o'rganadilar. Mavzu, tilshunoslikda katta ma'lumotlar bilan ishlash va avtomatik tahlil qilish ko'nikmalarini rivojlantirishga yordam beradi.

#### **6- mavzu. Graf va daraxt modellarini lingvistikada qo'llash va Markov zanjirlari va lingvistik tadqiqotlar**

Ushbu mavzu graf va daraxt modellarining lingvistik tahlilda qanday qo'llanilishini o'rgatadi. Graf va daraxt modellar, tilshunoslikda sintaksis va semantikaning strukturalarini matematik ravishda ifodalash uchun ishlatiladi. Daraxt modellari, ayniqsa, sintaktik tahlilda (masalan, gap tuzilishini aniqlash) keng qo'llaniladi, chunki ular tilning ierarxiy tuzilishini ko'rsatadi. Graf modellar esa til birikmalarining o'zaro bog'liqligini tasvirlashda ishlatiladi. Talabalar bu modellarni tilshunoslikda matn tahlili, sintaktik tahlil va semantik tahlil uchun qanday qo'llashni o'rganadilar. Mavzu, matematik modellarni lingvistik jarayonlarni tushunish va modellashtirishda qo'llashni o'rgatadi. Markov zanjiri — bu ehtimollik nazariyasiga asoslangan model bo'lib, tizimning kelajakdagi holati faqat hozirgi holatga bog'liq bo'ladi. Tilshunoslikda Markov zanjirlari matnlarni yoki gaplarni tahlil qilishda, masalan, til modellari, so'z ketma-ketliklarini prognoz qilish yoki matnlarni tasniflashda ishlatiladi. Talabalar bu modelni tilshunoslikda qanday qo'llashni, ayniqsa tilni qayta ishlash va avtomatik tahlil qilishda qanday yordam berishini o'rganadilar. Mavzu, ehtimollik asosidagi yondashuvlarni tilshunoslikda samarali qo'llashni tushunishga yordam beradi.

#### **7- mavzu. Lingvistik ma'lumotlarni tasniflash va klasterlash**

Ushbu mavzu lingvistik ma'lumotlarni tasniflash va klasterlash usullarini o'rgatadi. Tasniflash — bu ma'lumotlarni oldindan belgilangan toifalar yoki kategoriyalar bo'yicha ajratish jarayonidir. Klasterlash esa, ma'lumotlarni tabiiy guruhlar yoki klasterlarga ajratish, ya'ni ma'lumotlarni o'zaro o'xshashliklari asosida guruhlash jarayonidir. Lingvistikada bu usullar matnlarni, so'zlarni yoki gaplarni ma'lum xususiyatlarga ko'ra tasniflash yoki o'xshash guruhlariga ajratishda qo'llaniladi. Talabalar bu yondashuvlarni tilshunoslikdagi amaliy muammolarni hal qilishda, masalan, matnlarni tasniflash, sentiment tahlili yoki tilni avtomatik tahlil qilishda qanday qo'llashni o'rganadilar. Mavzu, lingvistik ma'lumotlar bilan ishlashda statistik va algoritmik yondashuvlarni o'rganishga yordam beradi.

#### **8- mavzu. Mashina o'qitishning asoslari, lingvistikada qo'llanilishi va sun'iy intellekt (NLP) asoslari**

Ushbu mavzu mashina o'qitishning asosiy tamoyillari va uning lingvistikada qo'llanilishini o'rgatadi. Mashina o'qitish — bu algoritmlar va statistik usullar

yordamida kompyuterlarga ma'lumotlardan o'rganish va qarorlar qabul qilishni o'rgatish jarayonidir. Lingvistikada mashina o'qitish usullari, tilni avtomatik tahlil qilish, matnlarni tasniflash, tarjima qilish, sentiment tahlili va boshqa ko'plab sohalarda qo'llaniladi. Talabalar, mashina o'qitishning asosiy algoritmlari, masalan, qaror daraxtlari, neyron tarmoqlar va klasterlash usullarini o'rganadilar, shuningdek, bu texnologiyalarni tilshunoslikda qanday samarali ishlatish mumkinligini tushunadilar. Mavzu, tilshunoslik va sun'iy intellektning kesishgan sohasidagi zamonaviy texnologiyalarni o'rganishga yordam beradi.

#### **9- mavzu. Semantik tahlil va yashirin semantik modellar (LSA)**

Ushbu mavzu semantik tahlil va yashirin semantik modellar (Latent Semantic Analysis, LSA) ning asosiy prinsiplari va ularning tilshunoslikda qo'llanilishini o'rgatadi. Semantik tahlil — bu matn yoki til birliklarining ma'nolarini aniqlash va ularni tizimli ravishda tahlil qilish jarayonidir. LSA, matnlardagi yashirin semantik tuzilmalarni aniqlash uchun ishlatiladigan statistik usul bo'lib, bu metod matnlar orasidagi semantik bog'liqliklarni aniqlash va tushunishga yordam beradi. Talabalar LSA ni matnlarni tahlil qilish, ma'lumotlarni yig'ish va til ma'lumotlarining semantik strukturalarini aniqlashda qanday qo'llashni o'rganadilar. Mavzu, tilshunoslikda matnni semantik jihatdan tushunish va avtomatik tahlil qilish ko'nikmalarini rivojlantirishga yordam beradi.

#### **10- mavzu. Tilshunoslikda ehtimollik modellaridan foydalanish**

Ushbu mavzu tilshunoslikda ehtimollik modellarining qanday qo'llanilishini o'rgatadi. Ehtimollik modellar tilning murakkab strukturalarini tasvirlash va tahlil qilishda statistik yondashuvlarni qo'llaydi. Bu modellar til birliklarining (so'zlar, gaplar) ehtimoliy taqsimotini o'rganishga asoslanadi va tilni avtomatik qayta ishlashda, masalan, so'z taxminlari, nutqni aniqlash yoki tarjima qilishda keng qo'llaniladi. Talabalar ehtimollik modellarining asosiy prinsiplari, shuningdek, N-gram, Markov zanjirlari va boshqa ehtimollik asosidagi metodlarni tilshunoslikda qanday qo'llashni o'rganadilar. Mavzu, tilshunoslikdagi ma'lumotlarni statistik usullar yordamida tahlil qilish va undan samarali foydalanish ko'nikmalarini rivojlantiradi.

#### **11- mavzu. Ma'lumotlarni vizualizatsiya qilish**

Ushbu mavzu ma'lumotlarni vizualizatsiya qilishning asosiy usullari va vositalarini o'rgatadi. Ma'lumotlarni vizualizatsiya qilish — bu murakkab ma'lumotlar to'plamlarini grafik ko'rinishda ifodalash jarayonidir, bu esa ma'lumotlarni yaxshiroq tushunishga va tahlil qilishga yordam beradi. Vizualizatsiya yordamida ma'lumotlar o'qilishi oson bo'lgan diagrammalar, grafiklar, xaritalar va boshqa tasvirlar shaklida taqdim etiladi. Talabalar, ma'lumotlarni vizualizatsiya qilishda ishlatiladigan vositalar (masalan, Matplotlib, Seaborn, Plotly) va ularning tilshunoslikdagi qo'llanilishi, masalan, til ma'lumotlarini tasvirlash va tahlil qilishda qanday ishlatilishini o'rganadilar. Mavzu, ma'lumotlarni samarali vizual tarzda taqdim etish va ulardan tahlil qilish ko'nikmalarini rivojlantirishga yordam beradi.

## **12- mavzu. Matematik usullarni lingvistik amaliyotda qo'llash**

Ushbu mavzu matematik usullarni lingvistik amaliyotda qanday qo'llashni o'rgatadi. Matematik usullar tilshunoslikda til ma'lumotlarini tahlil qilish, model yaratish va turli lingvistik jarayonlarni avtomatlashtirish uchun ishlatiladi. Ushbu metodlar, masalan, statistik tahlil, ehtimollik nazariyasi, algoritmik yondashuvlar, graf nazariyasi va matematik modellashtirish, tilni avtomatik tahlil qilish, ma'lumotlarni chiqarish, matnni tasniflash, so'zlarni aniqlash kabi ko'plab sohalarda qo'llaniladi. Talabalar, matematik usullar yordamida lingvistik jarayonlarni qanday modellashtirish, til ma'lumotlarini qayta ishlash va tahlil qilishni o'rganadilar. Mavzu, tilshunoslik va matematik modellashtirishning kesishgan sohasida zamonaviy texnologiyalarni qo'llashni tushunishga yordam beradi.

### **II.2. Amaliy mashg'ulotlari bo'yicha ko'rsatma va tavsiyalar**

*"Tilshunoslikda matematik usullar" fani bo'yicha amaliy mashg'ulotlar uchun quyidagi mavzular tavsiya etiladi:*

#### **1-mavzu. Tilshunoslikda matematik usullarga kirish**

Ushbu mavzu tilshunoslikdagi murakkab jarayonlarni matematik usullar yordamida tahlil qilishning asosiy tamoyillarini tushuntiradi. Matematik va statistik yondashuvlarning tilshunoslikda tutgan o'rni, ahamiyati, shuningdek, lingvistik jarayonlarni modellashtirish, tahlil qilish va bashoratlash imkoniyatlari haqida umumiy ma'lumot beriladi. Mavzuda talabalar matematik lingvistikaga oid muhim tushunchalar bilan tanishadi va kelgusi mavzularga tayyorgarlik ko'radi.

#### **2- mavzu. Statistika va ehtimollikning asosiy tushunchalari**

Ushbu mavzu statistik va ehtimollik nazariyasining asosiy tushunchalarini o'rgatadi. Talabalar ma'lumotlarni yig'ish, tahlil qilish va tasniflashning statistik usullarini o'zlashtiradilar. Ehtimollik nazariyasi orqali tilshunoslikdagi tasodifiy jarayonlarni model qilish va ehtimoliy yondashuvlarni qo'llash imkoniyatlari haqida tushuncha hosil qilinadi. Bu mavzu, ayniqsa, til ma'lumotlarining tasniflanishi va tahlili uchun statistik va ehtimollik asoslarini o'rganishga yordam beradi.

#### **3- mavzu. Zipf qonuni va lingvistik birliklarning chastotasi**

Ushbu mavzu Zipf qonunini va uning lingvistik tahlildagi ahamiyatini o'rganadi. Zipf qonuni, odatda, so'zlar yoki til birliklarining chastotasi va ularning o'rni o'rtasidagi teskari bog'liqlikni tasvirlaydi: eng ko'p ishlatiladigan so'zlar kamdan-kam ishlatiladiganlarga nisbatan sezilarli darajada ko'proq uchraydi. Talabalar bu qonunni tilshunoslikda, xususan, so'zning chastotasini o'lchashda va matnlarni tahlil qilishda qanday qo'llash mumkinligini o'rganadilar. Mavzu, lingvistik birliklarning taqsimoti va uning statistik modellari bilan tanishishga yordam beradi.

#### **4- mavzu. Matematik modellashtirish va til birliklarini statistik tahlil qilish**

Ushbu mavzu til birliklarini matematik modellar yordamida tasvirlashni o'rgatadi. Matematik modellashtirish, tilshunoslikdagi murakkab jarayonlarni aniq va o'lchovli tarzda ifodalash imkonini beradi. Talabalar til birliklarini (so'zlar, gaplar, frazalar) matematik yondashuvlar, masalan, vektorlar, matritsalar yoki grafiklar yordamida qanday modellashtirish mumkinligini o'rganadilar. Mavzu, tilshunoslikdagi sintaktik va semantik strukturalarni matematik ravishda ifodalash va tahlil qilish imkoniyatlarini tushuntiradi.

#### **5- mavzu. N-gramma modellar, ularning qo'llanilishi, korpus lingvistikasi va matematik tahlil hamda ehtimollik va statistik usulla**

Ushbu mavzu N-gramma modellarini va ularning tilshunoslikda qanday qo'llanilishini o'rgatadi. N-gramma modeli — bu so'zlar yoki harflar ketma-ketligini o'rganish va tahlil qilish uchun ishlatiladigan statistik usul. Bu modelda "N" – ketma-ketlikdagi so'zlar yoki belgilar sonini bildiradi. Talabalar N-gramma modellarini matnlarni tahlil qilish, tilni avtomatik ravishda qayta ishlash, masalan, so'z taxminlari, tilni tarjima qilish yoki matnni tasniflash kabi sohalarda qanday ishlatilishini o'rganadilar. Mavzu, tilshunoslikda statistik yondashuvlarning samarali qo'llanilishini tushunishga yordam beradi. Korpus lingvistikasi — bu til ma'lumotlarini to'plash, saqlash va tahlil qilish uchun ishlatiladigan yondashuv bo'lib, katta hajmdagi til ma'lumotlaridan (korpuslardan) foydalanishni o'z ichiga oladi. Matematik tahlil esa bu ma'lumotlarni statistik va ehtimollik usullari orqali qayta ishlashni anglatadi. Talabalar lingvistik korpuslarni yaratish, ularni matematik usullar yordamida tahlil qilish va tilshunoslikda zamonaviy texnologiyalarni qo'llashni o'rganadilar. Mavzu, tilshunoslikda katta ma'lumotlar bilan ishlash va avtomatik tahlil qilish ko'nikmalarini rivojlantirishga yordam beradi.

#### **6- mavzu. Graf va daraxt modellarini lingvistikada qo'llash va Markov zanjirlari va lingvistik tadqiqotlar**

Ushbu mavzu graf va daraxt modellarining lingvistik tahlilda qanday qo'llanilishini o'rgatadi. Graf va daraxt modellar, tilshunoslikda sintaksis va semantikaning strukturalarini matematik ravishda ifodalash uchun ishlatiladi. Daraxt modellari, ayniqsa, sintaktik tahlilda (masalan, gap tuzilishini aniqlash) keng qo'llaniladi, chunki ular tilning ierarxik tuzilishini ko'rsatadi. Graf modellar esa til birikmalarining o'zaro bog'liqligini tasvirlashda ishlatiladi. Talabalar bu modellarni tilshunoslikda matn tahlili, sintaktik tahlil va semantik tahlil uchun qanday qo'llashni o'rganadilar. Mavzu, matematik modellarni lingvistik jarayonlarni tushunish va modellashtirishda qo'llashni o'rgatadi. Markov zanjiri — bu ehtimollik nazariyasiga asoslangan model bo'lib, tizimning kelajakdagi holati faqat hozirgi holatga bog'liq bo'ladi. Tilshunoslikda Markov zanjirlari matnlarni yoki gaplarni tahlil qilishda, masalan, til modellari, so'z ketma-ketliklarini prognoz qilish yoki matnlarni tasniflashda ishlatiladi. Talabalar bu modelni tilshunoslikda qanday qo'llashni, ayniqsa tilni qayta ishlash va avtomatik tahlil qilishda qanday yordam berishini o'rganadilar. Mavzu, ehtimollik asosidagi yondashuvlarni tilshunoslikda samarali qo'llashni tushunishga yordam beradi.

#### **7- mavzu. Lingvistik ma'lumotlarni tasniflash va klasterlash**

Ushbu mavzu lingvistik ma'lumotlarni tasniflash va klasterlash usullarini o'rgatadi. Tasniflash — bu ma'lumotlarni oldindan belgilangan toifalar yoki kategoriyalar bo'yicha ajratish jarayonidir. Klasterlash esa, ma'lumotlarni tabiiy guruhlar yoki klasterlarga ajratish, ya'ni ma'lumotlarni o'zaro o'xshashliklari asosida guruhlash jarayonidir. Lingvistikada bu usullar matnlarni, so'zlarni yoki gaplarni ma'lum xususiyatlarga ko'ra tasniflash yoki o'xshash guruhlarga ajratishda qo'llaniladi. Talabalar bu yondashuvlarni tilshunoslikdagi amaliy muammolarni hal qilishda, masalan, matnlarni tasniflash, sentiment tahlili yoki tilni avtomatik tahlil qilishda qanday qo'llashni o'rganadilar. Mavzu, lingvistik ma'lumotlar bilan ishlashda statistik va algoritmik yondashuvlarni o'rganishga yordam beradi.

#### **8- mavzu. Mashina o'qitishning asoslari, lingvistikada qo'llanilishi va sun'iy intellekt (NLP) asoslari**

Ushbu mavzu mashina o'qitishning asosiy tamoyillari va uning lingvistikada qo'llanilishini o'rgatadi. Mashina o'qitish — bu algoritmlar va statistik usullar yordamida kompyuterlarga ma'lumotlardan o'rganish va qarorlar qabul qilishni o'rgatish jarayonidir. Lingvistikada mashina o'qitish usullari, tilni avtomatik tahlil qilish, matnlarni tasniflash, tarjima qilish, sentiment tahlili va boshqa ko'plab sohalarida qo'llaniladi. Talabalar, mashina o'qitishning asosiy algoritmlari, masalan, qaror daraxtlari, neyron tarmoqlar va klasterlash usullarini o'rganadilar, shuningdek, bu texnologiyalarni tilshunoslikda qanday samarali ishlatish mumkinligini tushunadilar. Mavzu, tilshunoslik va sun'iy intellektning kesishgan sohalaridagi zamonaviy texnologiyalarni o'rganishga yordam beradi.

#### **9- mavzu. Semantik tahlil va yashirin semantik modellar (LSA)**

Ushbu mavzu semantik tahlil va yashirin semantik modellar (Latent Semantic Analysis, LSA) ning asosiy prinsiplari va ularning tilshunoslikda qo'llanilishini o'rgatadi. Semantik tahlil — bu matn yoki til birliklarining ma'nolarini aniqlash va ularni tizimli ravishda tahlil qilish jarayonidir. LSA, matnlardagi yashirin semantik tuzilmalarni aniqlash uchun ishlatiladigan statistik usul bo'lib, bu metod matnlar orasidagi semantik bog'liqliklarni aniqlash va tushunishga yordam beradi. Talabalar LSA ni matnlarni tahlil qilish, ma'lumotlarni yig'ish va til ma'lumotlarining semantik strukturalarini aniqlashda qanday qo'llashni o'rganadilar. Mavzu, tilshunoslikda matnni semantik jihatdan tushunish va avtomatik tahlil qilish ko'nikmalarini rivojlantirishga yordam beradi.

#### **10- mavzu. Tilshunoslikda ehtimollik modellaridan foydalanish**

Ushbu mavzu tilshunoslikda ehtimollik modellarining qanday qo'llanilishini o'rgatadi. Ehtimollik modellar tilning murakkab strukturalarini tasvirlash va tahlil qilishda statistik yondashuvlarni qo'llaydi. Bu modellar til birliklarining (so'zlar, gaplar) ehtimoliy taqsimotini o'rganishga asoslanadi va tilni avtomatik qayta ishlashda, masalan, so'z taxminlari, nutqni aniqlash yoki tarjima qilishda keng

qo'llaniladi. Talabalar ehtimollik modellarining asosiy prinsiplari, shuningdek, N-gram, Markov zanjirlari va boshqa ehtimollik asosidagi metodlarni tilshunoslikda qanday qo'llashni o'rganadilar. Mavzu, tilshunoslikdagi ma'lumotlarni statistik usullar yordamida tahlil qilish va undan samarali foydalanish ko'nikmalarini rivojlantiradi.

#### **11- mavzu. Ma'lumotlarni vizualizatsiya qilish**

Ushbu mavzu ma'lumotlarni vizualizatsiya qilishning asosiy usullari va vositalarini o'rgatadi. Ma'lumotlarni vizualizatsiya qilish — bu murakkab ma'lumotlar to'plamlarini grafik ko'rinishda ifodalash jarayonidir, bu esa ma'lumotlarni yaxshiroq tushunishga va tahlil qilishga yordam beradi. Vizualizatsiya yordamida ma'lumotlar o'qilishi oson bo'lgan diagrammalar, grafiklar, xaritalar va boshqa tasvirlar shaklida taqdim etiladi. Talabalar, ma'lumotlarni vizualizatsiya qilishda ishlatiladigan vositalar (masalan, Matplotlib, Seaborn, Plotly) va ularning tilshunoslikdagi qo'llanilishi, masalan, til ma'lumotlarini tasvirlash va tahlil qilishda qanday ishlatilishini o'rganadilar. Mavzu, ma'lumotlarni samarali vizual tarzda taqdim etish va ulardan tahlil qilish ko'nikmalarini rivojlantirishga yordam beradi.

#### **12- mavzu. Matematik usullarni lingvistik amaliyotda qo'llash**

Ushbu mavzu matematik usullarni lingvistik amaliyotda qanday qo'llashni o'rgatadi. Matematik usullar tilshunoslikda til ma'lumotlarini tahlil qilish, model yaratish va turli lingvistik jarayonlarni avtomatlashtirish uchun ishlatiladi. Ushbu metodlar, masalan, statistik tahlil, ehtimollik nazariyasi, algoritmik yondashuvlar, graf nazariyasi va matematik modellashtirish, tilni avtomatik tahlil qilish, ma'lumotlarni chiqarish, matnni tasniflash, so'zlarni aniqlash kabi ko'plab sohalarda qo'llaniladi. Talabalar, matematik usullar yordamida lingvistik jarayonlarni qanday modellashtirish, til ma'lumotlarini qayta ishlash va tahlil qilishni o'rganadilar. Mavzu, tilshunoslik va matematik modellashtirishning kesishgan sohasida zamonaviy texnologiyalarni qo'llashni tushunishga yordam beradi.

### **III. Mustaqil ta'lim\***

#### **III.1. Mustaqil ta'lim va mustaqil ishlar**

Talabalarni mustaqil ta'lim shaklini tashkil etishga qo'yilgan talablar O'zbekiston Respublikasi oliy ta'lim, fan va innovatsiyalar vazirligining 2024-yil 29-apreldagi 136-sonli "Oliy ta'lim muassasalari talabalari mustaqil ta'limini tashkil etish bo'yicha namunaviy tartibni tasdiqlash to'g'risida"gi buyrug'i asosida ishlab chiqilgan.

Mustaqil ta'limni baholash semestr davomida berilgan topshiriq asosida bajarilgan ishlarni HEMISda ilova qilish, shuningdek, oraliq va yakuniy test va savollarga javob berish asosida oshiriladi.

“Tilshunoslikda matematik usullar” fanidan mustaqil ta’lim uchun tavsiya etiladigan mavzular:

Tilshunoslik va matematik usullarga kirish  
Statistika va ehtimollikning asosiy tushunchalari  
Zipf qonuni va lingvistik birliklarning chastotasi  
Matematik modellashtirish va til birliklari  
N-gramma modellar va ularning qo’llanilishi  
Korpus lingvistikasi va matematik tahlil  
Graf va daraxt modellarini lingvistikada qo’llash  
Markov zanjirlari va lingvistik tadqiqotlar  
Lingvistik ma’lumotlarni tasniflash va klasterlash  
Mashina o’qitishning asoslari va lingvistikada qo’llanilishi  
Semantik tahlil va yashirin semantik modellar (LSA)  
Tilshunoslikda ehtimollik modellaridan foydalanish  
Ma’lumotlarni vizualizatsiya qilish  
Matematik usullarni lingvistik amaliyotda qo’llash

Talabalarni mustaqil ta’lim shaklini tashkil etishga qo’yilgan talablar O‘zbekiston Respublikasi oliy ta’lim, fan va innovatsiyalar vazirligining 2024-yil 29-apreldagi 136-sonli “Oliy ta’lim muassasalari talabalari mustaqil ta’limini tashkil etish bo’yicha namunaviy tartibni tasdiqlash to’g’risida”gi buyrug’i asosida ishlab chiqilgan.

Mustaqil ta’limni baholash semestr davomida berilgan topshiriq asosida bajarilgan ishlarni HEMISda ilova qilish, shuningdek, oraliq va yakuniy test va savollarga javob berish asosida oshiriladi.

\* *Izoh: Mustaqil ta’lim ishi mavzulari fan doirasida o’zgarishi mumkin.*

**4. V. Fan o’qitilishining natijalari (shakllanadigan kompetensiyalar)**

**Matematik yondashuvlar asosida tilni tahlil qilish**

Talabalar matematik usullarni tilshunoslik amaliyotiga qo’llay olishadi, til birliklarini modellashtirish, tahlil qilish va avtomatik qayta ishlashda zamonaviy matematik yondashuvlarni ishlatadilar.

**Matematika va tilshunoslikni integratsiya qilish**

Talabalar tilshunoslikdagi matematik modellashtirish va statistik tahlil usullarini birlashtirgan holda, lingvistik muammolarni yechish uchun matematik metodlarni qo’llay olishadi.

**Ehtimollik va statistik modellarni qo’llash**

Talabalar ehtimollik modellari, statistik tahlil va algoritmlar yordamida tilshunoslikdagi ma'lumotlarni tahlil qilish va model yaratishda samarali ishlatish ko'nikmalarini egallaydilar.

#### **Mashina o'qitish va sun'iy intellekti qo'llash**

Talabalar lingvistik ma'lumotlarni tahlil qilishda mashina o'qitish, neyron tarmoqlar va sun'iy intellekt texnologiyalaridan foydalanish ko'nikmalariga ega bo'ladilar.

#### **Matnlarni avtomatik qayta ishlash va semantik tahlil**

Talabalar matnlarni avtomatik ravishda tahlil qilish, semantik tahlil qilish va ma'lumotlarni samarali vizualizatsiya qilishda ilg'or metodlardan foydalanishni o'rganadilar.

#### **Lingvistik ma'lumotlarni klasterlash va tasniflash**

Talabalar lingvistik ma'lumotlarni klasterlash va tasniflash metodlarini qo'llay olishadi, shuningdek, matnlarni yoki so'zlarni avtomatik tasniflash va guruhlashni amalga oshirish ko'nikmalarini rivojlantiradilar.

#### **Lingvistik tadqiqotlar uchun matematik modellarni yaratish**

Talabalar tilshunoslikdagi tadqiqotlarni o'tkazishda matematik modellarni yaratish, tahlil qilish va ular yordamida tilni chuqurroq o'rganish ko'nikmalarini egallaydilar.

#### **Kompyuter dasturlash va statistik tahlil vositalarini qo'llash**

Talabalar Python, R, MATLAB kabi dasturlash tillari va statistik tahlil vositalarini tilshunoslikdagi amaliy masalalarni hal qilishda qo'llay olishadi.

5.

### **VI. TA'LIM TEXNOLOGIYALARI VA METODLARI:**

#### **Interaktiv ta'lim metodlari:**

Talabalar o'z bilimlarini amaliy mashqlar orqali rivojlantiradilar. Bu metodlar orqali talabalar mavzularni o'rganishda faollik ko'rsatadilar, muammolarni mustaqil ravishda hal qilishga undaydilar. Masalan, guruhli ishlash, munozaralar va interaktiv faoliyatlar orqali ta'lim olib boriladi.

#### **Masalalar va misollar yechish:**

O'qitish jarayonida onlayn resurslar, ta'lim platformalari (masalan, Moodle, Google Classroom), ma'lumotlarni vizualizatsiya qilish vositalari (Matplotlib, Seaborn, Plotly) va tahlil vositalaridan foydalaniladi. Bu texnologiyalar talabalarni zamonaviy ilmiy tadqiqotlar va amaliy masalalar bilan tanishtiradi. O'qitish jarayonida talabalar mashina o'qitish va sun'iy intellekt metodlarini o'rganadilar, masalan, lingvistik ma'lumotlarni tasniflash va klasterlash, semantik tahlil va lingvistik model yaratishda. Talabalarni matematik modellashtirish va sun'iy intellekt asosidagi tizimlar bilan tanishtiradi. O'quv jarayonida nazariy bilimlar amaliy masalalar orqali mustahkamlanadi. Talabalar matematik modellarni haqiqiy lingvistik ma'lumotlar ustida qo'llab, amaliy loyihalarni amalga oshiradilar. Bu usul talabalar uchun kengaytirilgan o'rganish imkoniyatini yaratadi.

6. **VII. Kreditlarni olish uchun talablar:**

Topshiriqni muvaffaqiyatli bajarish uchun asosiy ko'nikmalar:

**Nazariy bilimlar:** Talabalar matematik va statistik metodlarni qo'llash orqali tilshunoslikdagi ma'lumotlarni tahlil qilish, matematik modellar yaratish va bu modellarni amaliyotda qo'llashda ko'nikmalarni rivojlantiradilar. Python, R, MATLAB va boshqa dasturlash tillarini bilish, statistik tahlil va matematik modellar vositalaridan samarali foydalanish ko'nikmalarini egallash talab etiladi. Talabalar kompyuter dasturlari va matematik vositalar yordamida tilshunoslik masalalarini yechishlari kerak.

**Amaliy ko'nikmalar:** Ehtimollik nazariyasi va statistik metodlarni qo'llash, ma'lumotlar ustida ehtimollik modellari qurish va statistik tahlilni amalga oshirish ko'nikmalari zarur. Tilshunoslikdagi turli ma'lumotlarni, jumladan, korpuslar, tekstlar va lingvistik birliklarni tahlil qilish, ularni klassifikatsiya qilish va klasterlash, hamda matematik tahlil metodlarini qo'llash ko'nikmalarini rivojlantirish. Mashina o'qitish va sun'iy intellekt usullaridan foydalanish, masalan, tilshunoslikdagi klasterlash, tasniflash va semantik tahlilni bajarish. Semantik tahlil qilish, yashirin semantik modellarni (LSA) qo'llash, ma'lumotlarni yashirin tuzilmalar asosida tahlil qilish. Matematik modellarni va lingvistik tahlil natijalarini vizual tarzda taqdim etish, ma'lumotlarni grafiklar, diagrammalar va boshqa vizual vositalar orqali tushunarli qilish ko'nikmalari.

Muloqot ko'nikmalari: O'z qarorlarini aniq va ravshan tushuntirish qobiliyati.

7. **VIII. ASOSIY VA QO'SHIMCHA O'QUV ADABIYOTLAR HAMDA AXBOROT MANBALARI**

**Asosiy adabiyotlar**

1. Zoran Gacovski Natural Language Processing Arcler Press 224 Shoreacres Road Burlington, ON L7L 2H2 Canada www.arclerpress.com Email: orders@arclereducation.com e-book Edition 2021 ISBN: 978-1-77407-979-9 (e-book)
2. Pushpak Bhattacharyya and Aditya Joshi Natural Language Processing Published by Wiley India Pvt. Ltd., 4436/7, Ansari Road, Daryaganj, New Delhi-110002. Printed at: Yash Printographics Edition: 2023 ISBN: 978-93-5746-238-9 ISBN: 978-93-5746-239-6 (ebk)

**Qo'shimcha adabiyotlar**

3. Mirziyoyev Sh.M. Hozirgi zamon va Yangi O'zbekiston. - Toshkent: O'zbekiston, 2024.
4. Mirziyoyev Sh.M. Yangi O'zbekistonda taraqqiyot strategiyasi asosida demokratik islohotlar yo'lini qat'iy davom ettiramiz. 6-jild. - Toshkent: O'zbekiston, 2023

5. Saidov D. Python dasturlash tili. O'quv - uslubiy qo'llanma, 2019
6. Barbara H. Partee Mathematical methods in linguistics Department of Linguistics University of Washington 1987 ISBN 90-277-2244-7
7. Григорьев Ю.Д., Мартыненко Г.Я. Последовательности типа Фибоначчи. Теория и прикладные аспекты Год издания: 2024 ISBN 978-5-507-50228-8 516 с.
8. Николаев И. С., Митренина О. В., Ландо Т. М. (Ред.). Прикладная и компьютерная лингвистика. М.: URSS, 2016. 320 с. ISBN 978-5-9710-3472-8
9. Батура, Т. В. Математическая лингвистика и автоматическая обработка текстов : учеб. пособие / Т. В. Батура ; Новосиб. гос. ун-т. – Новосибирск : РИЦ НГУ, 2016. – 166 с.
10. УДК 81'32 ББК 81.1 Т415 Тимофеева М. К. Введение в математическую лингвистику: Практикум / Новосиб. гос. ун-т. Новосибирск, 2018. С. 56

#### Scopusdan maqolalar

11. A Closer Look at Classification Evaluation Metrics and a Critical Reflection of Common Evaluation Practice Juri Opitz Transactions of the Association for Computational Linguistics (2024) 12: 820–836.  
<https://doi.org/10.1162/tacl a 00675>  
<https://direct.mit.edu/tacl/article/doi/10.1162/tacl a 00675/122720/A-Closer-Look-at-Classification-Evaluation-Metrics>
12. Improving Probability-based Prompt Selection Through Unified Evaluation and Analysis Open Access Sohee Yang, Jonghyeon Kim, Joel Jang, Seonghyeon Ye, Hyunji Lee, Minjoon Seo Transactions of the Association for Computational Linguistics (2024) 12: 664–680. <https://doi.org/10.1162/tacl a 00666>

<https://direct.mit.edu/tacl/article/doi/10.1162/tacl a 00666/121197/Improving-Probability-based-Prompt-Selection>

#### Axborot manbalar:

- <https://www.w3schools.com/>  
<https://www.tutorialsteacher.com/>  
<https://uzbekdevs.uz/>  
<https://www.studytonight.com/>

8.	<p>Fanning o'quv dasturi Tarjimashunoslik, tilshunoslik va xalqaro jurnalistika oliy maktabida ishlab chiqilgan va 2025 yil "18" 06" dagi 24-sonli bayonnomasi bilan ma'qullangan.</p> <p>O'quv dasturi Toshkent davlat sharqshunoslik universiteti Kengashining 2025 yil "28" 06" sonli bayoni bilan tasdiqlangan.</p>
9.	<p><b>Fan/modul uchun ma'sullar:</b></p> <p>1. Allanyazov R.B.– TDShU Tarjimashunoslik, tilshunoslik va xalqaro jurnalistika oliy maktabi o'qituvchisi</p>
10.	<p><b>Taqrizchilar:</b></p> <p>1. Usmanova Shoira Rustamovna- Tarjimashunoslik, tilshunoslik va xalqaro jurnalistika oliy maktabi professori, f.f.d.</p> <p>2. N.Abdurahmonova - O'zMU, Kompyuter lingvistikasi va amaliy tilshunoslik" kafedrası professori, f.f.d</p>

dep

Faint, illegible text in the upper section of the page, possibly a header or introductory paragraph.

Large block of faint, illegible text in the middle section of the page, appearing to be the main body of the document.

Faint, illegible text in the lower section of the page, possibly a conclusion or footer.