

✓

O'ZBEKISTON RESPUBLIKASI
OLIY TA'LIM, FAN VA INNOVATSIYALAR VAZIRLIGI
TOSHKENT DAVLAT SHARQSHUNOSLIK UNIVERSITETI



KATTA MA'LUMOTLAR TEXNOLOGIYALARI (BIG DATA)
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

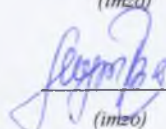
Sharq xalqlari tillari va adabiyoti
instituti direktori:



X.V. Mirzaxmedova

(imzo)

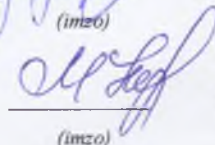
Tarjimashunoslik, tilshunoslik va
xalqaro jumalistika oliy maktabi boshlig'i:



S.T. Mustafayeva

(imzo)

Arm boshlig'i:



M. Yuldasheva

(imzo)

Fan/modul kodi KMT 16-08		O'quv yili 2027/2028	Semestr 6	YeCTS - Kreditlar 8		
Fan/modul turi Majburiy		Ta'lim tili O'zbek		Haftadagi dars soatlari 8		
1.	Fanning nomi	Auditoriya mashg'ulotlari (soat)	Ma'ruza	Amaliy	Mustaqil ta'lim (soat)	Jami yuklama (soat)
	Katta ma'lumotlar texnologiyalari (Big Data)	96	48	48	144	240
2.	<p align="center">I. Fanning mazmuni:</p> <p>Fanning maqsadi: Katta ma'lumotlar texnologiyalari fanining mazmunida katta hajmdagi ma'lumotlarni saqlash, qayta ishlash va tahlil qilish usullari o'rganiladi. Fan katta ma'lumotlar (Big Data) konsepsiyasi, uning hayot sikli, ma'lumotlarni yig'ish, saqlash va tahlil qilish texnologiyalari, shuningdek, ma'lumotlarni parallel qayta ishlash va taqsimlangan hisoblashlar kabi mavzularni qamrab oladi. Ushbu fan shuningdek, Hadoop, Spark va NoSQL ma'lumotlar bazalari kabi zamonaviy texnologiyalar bilan tanishtirishni maqsad qiladi. Katta ma'lumotlar asosidagi tahlil va prognozlashning real hayotdagi qo'llanishi ham o'rganiladi.</p> <p>Fan vazifasi - Katta ma'lumotlar texnologiyalari fani talabalarga katta hajmdagi ma'lumotlar bilan ishlashga oid asosiy tushunchalar va texnologiyalarni o'rgatishga yo'naltirilgan. Ushbu fanning asosiy vazifasi talabalarga ma'lumotlarni samarali qayta ishlash, tahlil qilish va ulardan qiymat chiqarib olish yo'llarini o'rgatishdan iborat. Shuningdek, talabalarni zamonaviy katta ma'lumotlar texnologiyalari bilan tanishtirish, ularning imkoniyatlari va cheklovlarini tushuntirish, hamda ulardan foydalanish bo'yicha amaliy ko'nikmalar berishdir.</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>https://www.fu-berlin.de/studium/studienangebot/grundstaendige/informatik_kombi_lo/index.html</p>					

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

II.1. Fan tarkibiga quyidagi mavzular kiradi:

1-mavzu. Katta ma'lumotlar va Data Science-ga kirish

Ushbu mavzu katta ma'lumotlar texnologiyalari va Data Science asoslarini o'rganishni boshlaydi. Ma'ruza davomida talabalar katta ma'lumotlar tushunchasi, ularning turlari, va ahamiyati bilan tanishadilar. Katta ma'lumotlar nafaqat hajmi katta bo'lgan ma'lumotlar to'plami, balki ularning tezligi, xilma-xilligi va o'zgaruvchanligi bilan ajralib turishini tushunadilar. Shuningdek, Data Science tushunchasi, uning jarayonlari va asosiy elementlari: ma'lumotlarni yig'ish, tahlil qilish, modellashtirish va natijalarni vizualizatsiya qilish kabilar o'rganiladi. Talabalar katta ma'lumotlarning zamonaviy biznes va texnologiyalar sohasidagi muhim rolini tushunib olishlari uchun amaliy misollar va ilovalar ko'rib chiqiladi. Bu mavzu Data Science va katta ma'lumotlar bo'yicha chuqurroq bilimlarni o'zlashtirish uchun mustahkam asos bo'ladi.

2- mavzu. Katta ma'lumotlar texnologiyalarining qo'llanilishi

Ushbu mavzu katta ma'lumotlar texnologiyalarining zamonaviy sanoat va ilmiy sohalarda qanday qo'llanilishini o'rganishga qaratilgan. Ma'ruza davomida talabalar katta ma'lumotlardan foydalanish orqali turli xil sohalarda qanday muammolarni hal qilish, mumkinligini tushunadilar. Ular moliya, sog'liqni saqlash, marketing, ishlab chiqarish, ta'lim, logistika va hukumat tashkilotlaridagi amaliyotlardan misollar ko'radilar.

3- mavzu. Data Science jarayoni va hayotiy tsikli

Ushbu mavzu Data Science jarayonining asosiy bosqichlari va hayotiy tsiklini o'rganishga qaratilgan. Ma'ruza davomida talabalar Data Science loyihasining turli bosqichlari bilan tanishadilar: ma'lumotlarni yig'ish, tozalash va tayyorlash, tahlil qilish, modellashtirish, natijalarni baholash va vizualizatsiya qilish. Har bir bosqichning o'ziga xosligi, qiyinchiliklari va muhim jihatlari tushuntiriladi. Talabalar ushbu jarayonlarni amaliy misollar yordamida o'rganib, Data Science hayotiy tsiklida qanday qilib ma'lumotlardan qimmatli bilimlar olish mumkinligini bilib olishadi. Mavzu Data Science loyihalarini muvaffaqiyatli amalga oshirish uchun zarur bo'lgan ko'nikmalarni shakllantirishga yordam beradi.

4- mavzu. Ma'lumotlarni yig'ish va to'plash usullari Ma'lumotlarni tozalash va tayyorlash

Ushbu mavzu ma'lumotlarni yig'ishning asosiy usullari va manbalarini o'rganishga qaratilgan. Talabalar web scraping, API-lar orqali ma'lumot olish, ma'lumotlar bazalaridan foydalanish va boshqa manbalar orqali ma'lumotlarni yig'ish usullarini bilib oladilar. Shuningdek, ma'lumotlarni to'plash jarayonida duch keladigan qiyinchiliklar va ularni hal qilish yo'llari ham tushuntiriladi. Mavzu Data Science loyihalarini amalga oshirishda sifatli va ishonchli ma'lumotlar yig'ishning ahamiyatini tushunishga yordam beradi. Ma'lumotlarni tozalash – bu Big Data tahlilining boshlang'ich bosqichi bo'lib, noto'g'ri, chala yoki keraksiz ma'lumotlarni aniqlash va ularni to'g'ri formatga keltirishni o'z ichiga oladi.

Ma'lumotlarni tayyorlash esa tahlil va modellashtirish uchun zarur bo'lgan jarayon bo'lib: Turli manbalardagi ma'lumotlar birlashtiriladi.

Kategoriyalar raqamli formatga o'tkaziladi (masalan, bir-hot encoding).

Qarashli qiymatlar (outliers) bilan ishlanadi.

Ushbu jarayonlar to'g'ri va aniq tahlil qilish, modellarni sifatli tuzish uchun asos yaratadi. Ma'lumotlarni tozalash va tayyorlashda Python yoki Apache Spark kabi texnologiyalar qo'llaniladi.

5- mavzu. Python yordamida ma'lumotlarni tahlil qilish: pandas va NumPy

Ushbu mavzu Python dasturlash tilida ma'lumotlarni tahlil qilish uchun asosiy kutubxonalar — pandas va NumPy bilan tanishishga qaratilgan. Talabalar pandas yordamida ma'lumotlarni ko'rish, tozalash, manipulyatsiya qilish va tahlil qilish jarayonlarini o'rganadilar. Shuningdek, NumPy kutubxonasi yordamida ko'p o'lchovli ma'lumotlar bilan ishlash va matematik hisob-kitoblarni amalga oshirish usullari ko'rsatiladi. Ushbu mavzu, talabalarga ma'lumotlarni samarali tahlil qilish va ushbu jarayonni avtomatlashtirishga yordam beradigan kuchli vositalarni o'zlashtirish imkonini beradi.

6- mavzu. Ma'lumotlarni tasvirlash: Matplotlib va Seaborn

Ushbu mavzu ma'lumotlarni vizualizatsiya qilishda Matplotlib va Seaborn kutubxonalaridan foydalanishni o'rganishga qaratilgan. Talabalar Matplotlib yordamida grafiklar, diagrammalar va boshqa vizual tasvirlarni qanday yaratish mumkinligini bilib oladilar. Seaborn esa ma'lumotlarni tasvirlash uchun qo'shimcha qulayliklar va estetikani taqdim etadi, bu esa murakkab ma'lumotlarni tushunishni osonlashtiradi. Mavzu davomida talabalar ma'lumotlarni vizualizatsiya qilishda muhim jihatlar, grafikalar yaratish usullari va tasvirlarning ahamiyatini o'rganadilar, shuningdek, amaliy misollar yordamida o'z bilimlarini mustahkamlash imkoniyatiga ega bo'lishadi.

7- mavzu. Data visualization vositalari: Plotly va Dash Amaliy vizualizatsiya loyihalari

Ushbu mavzu ma'lumotlarni vizualizatsiya qilishda Plotly va Dash kabi vositalardan foydalanishni o'rganishga bag'ishlangan. Talabalar Plotly kutubxonasi yordamida grafikalar va diagrammalar yaratish jarayonini bilib olishadi. Ushbu vosita yordamida interaktiv va estetik jihatdan jozibador grafikalar tayyorlash imkoniyati ko'rib chiqiladi. Dash esa Plotly asosida qurilgan, veb-ilovalar yaratish uchun mo'ljallangan freymorkdir. Talabalar Dash yordamida interaktiv vizualizatsiyalarni qanday yaratish va ularga foydalanuvchi interfeyslarini qo'shishni o'rganadilar. Mavzu ma'lumotlarni taqdim etish va tahlil qilishda vizualizatsiya vositalarining ahamiyatini tushunishga yordam beradi. Amaliy vizualizatsiya loyihalarini yaratish va ularni amalga oshirish jarayoniga qaratilgan. Talabalar ma'lumotlarni vizualizatsiya qilish uchun eng yaxshi amaliyotlar, vositalar va texnikalar bilan tanishadilar. Loyihalar davomida grafikalar, diagrammalar va interaktiv vizualizatsiyalarni yaratish orqali ma'lumotlarni taqdim etish ko'nikmalari rivojlantiriladi. Mavzu talabalarga vizualizatsiyaning ma'lumotlarni tushunishni osonlashtirishdagi ahamiyatini ko'rsatadi va ularning o'z loyihalarini muvaffaqiyatli amalga oshirishlari uchun zarur bo'lgan amaliy tajribani beradi.

8- mavzu. Hadoop va uning asoslari HDFS (Hadoop Distributed File System) va ma'lumotlarni taqsimlas

Ushbu mavzu Hadoop texnologiyasining asoslari va uning katta ma'lumotlar boshqarishdagi rolini o'rganishga qaratilgan. Talabalar Hadoop arxitekturasi, asosiy komponentlari, jumladan HDFS (Hadoop Distributed File System) va MapReduce bilan tanishadilar. Shuningdek, Hadoopning tarixi, rivojlanishi va uning katta ma'lumotlar tahlilidagi afzalliklari ham muhokama qilinadi. Mavzu davomida talabalar Hadoopni o'rnatish va asosiy operatsiyalarni bajarish bo'yicha amaliy ko'nikmalarni ham rivojlantiradilar. Bu

mavzu katta ma'lumotlar texnologiyalarida Hadoopni muvaffaqiyatli ishlatish uchun zarur bo'lgan bilimlarni taqdim etadi. HDFS (Hadoop Distributed File System) ning asosiy printsiplari va ma'lumotlarni taqsimlash jarayonini o'rganishga qaratilgan. Talabalar HDFS arxitekturasi, uning komponentlari va ma'lumotlarni qanday qilib tarqatilgan holda saqlash va boshqarish haqida bilib oladilar. Mavzu davomida ma'lumotlarni taqsimlashning ahamiyati, HDFSning yuqori darajadagi mavjudligi va chidamliligi, shuningdek, katta ma'lumotlar tizimlarida HDFSning qanday ishlashi muhokama qilinadi. Talabalar amaliy misollar orqali HDFSni o'rnatish, ma'lumotlarni yuklash va ularni boshqarish bo'yicha ko'nikmalarni rivojlantiradilar. Bu mavzu katta ma'lumotlar infratuzilmasida HDFSning rolini tushunishga yordam beradi.

9- mavzu. Apache Spark va uning imkoniyatlari

Ushbu mavzu Apache Spark texnologiyasining asoslari va uning katta ma'lumotlar tahlilidagi imkoniyatlarini o'rganishga qaratilgan. Talabalar Spark arxitekturasi, uning komponentlari, jumladan Spark Core, Spark SQL, Spark Streaming, MLlib va GraphX bilan tanishadilar. Mavzu davomida Apache Sparkning yuqori tezlikda ma'lumotlarni qayta ishlash, parallellik va tarqatilgan hisoblash imkoniyatlari muhokama qilinadi. Talabalar Spark yordamida katta ma'lumotlar to'plamini tahlil qilish, real vaqt rejimida ma'lumotlarni qayta ishlash va mashinani o'rganish modellari yaratish bo'yicha amaliy ko'nikmalarni rivojlantiradilar. Bu mavzu Apache Sparkni katta ma'lumotlar ekotizimida muvaffaqiyatli ishlatish uchun zarur bo'lgan bilimlarni taqdim etadi.

10- mavzu. SQL va NoSQL ma'lumotlar bazalari bilan ishlash va MongoDB Cassandra bilan tanishish

Ushbu mavzu SQL (Structured Query Language) va NoSQL (Not Only SQL) ma'lumotlar bazalari bilan ishlashni o'rganishga qaratilgan. Talabalar SQL asosidagi ma'lumotlar bazalarini, jumladan, ular qanday tuzilishi, ma'lumotlarni saqlash va qayta ishlash usullarini bilib oladilar. Shuningdek, NoSQL ma'lumotlar bazalarining xilma-xilligi, ularning afzalliklari va qanday holatlarda qo'llanilishi haqida ham ma'lumot beriladi. Mavzu davomida talabalar SQL va NoSQL tizimlari o'rtasidagi farqlarni, ularning ishlash prinsiplari va qaysi holatlarda qaysi turini tanlash kerakligini o'rganadilar. Amaliy mashg'ulotlar orqali talabalar SQL so'roqlari yozish va NoSQL ma'lumotlar bazalari bilan ishlash bo'yicha ko'nikmalarini rivojlantiradilar. Bu mavzu ma'lumotlar bazalari bilan samarali ishlash uchun zarur bo'lgan nazariy va amaliy bilimlarni taqdim etadi. Talabalar har ikkala tizimning asosiy xususiyatlari, arxitekturasi va ishlash printsiplari bilan tanishadilar. MongoDB va Cassandra o'rtasidagi farqlar, ularning ma'lumotlarni saqlash va boshqarish usullari, shuningdek, qaysi holatlarda ushbu tizimlardan foydalanishning afzalliklari va kamchiliklari muhokama qilinadi. Mavzu davomida talabalar MongoDB va Cassandra bilan ishlash bo'yicha amaliy misollar ko'rib, asosiy CRUD (Create, Read, Update, Delete) operatsiyalarini bajarish ko'nikmalarini rivojlantiradilar. Bu mavzu talabalarni NoSQL ma'lumotlar bazalarini tanlash va ishlatish bo'yicha chuqurroq bilimlar bilan ta'minlaydi.

11- mavzu. Cloud Computing va katta ma'lumotlar (AWS, Azure, Google Cloud)

Ushbu mavzu bulutli hisoblash (Cloud Computing) va katta ma'lumotlar (Big Data) texnologiyalarining integratsiyasini o'rganishga qaratilgan. Talabalar Amazon Web Services (AWS), Microsoft Azure va Google Cloud Platform kabi asosiy bulut xizmatlari ta'minotchilari bilan tanishadilar. Mavzu davomida bulutli xizmatlarning turli xil

modellarini (IaaS, PaaS, SaaS) o'rganish va katta ma'lumotlarni saqlash, qayta ishlash va tahlil qilishda bulutli yechimlarning afzalliklari muhokama qilinadi. Talabalar bulutli platformalarda katta ma'lumotlarni qanday qilib samarali boshqarish va ulardan foydalangan holda tahlil qilishni bilib olishadi. Bu mavzu bulutli hisoblash va katta ma'lumotlar sohasida zamonaviy yechimlar va amaliyotlarni tushunishga yordam beradi.

12- mavzu. Machine Learning tushunchalari va amaliyot

Ushbu mavzu Machine Learning (ML) tushunchalarini va ularni amaliyotda qanday qo'llashni o'rganishga qaratilgan. Talabalar Machine Learningning asosiy prinsiplarini, turli algoritmlarni (masalan, nazoratli, nazoratsiz va yarim nazoratli o'rganish) va ulardan qaysi holatlarda foydalanish mumkinligini bilib oladilar. Mavzu davomida talabalar ML jarayonining asosiy bosqichlarini, jumladan ma'lumotlarni tayyorlash, modelni o'qitish, baholash va natijalarni interpretatsiya qilishni o'rganadilar. Amaliy mashg'ulotlar orqali talabalar real hayotdagi muammolarni hal qilish uchun ML modellarini yaratish va qo'llash ko'nikmalarini rivojlantiradilar. Bu mavzu Machine Learningga kirish va uni amaliyotda muvaffaqiyatli qo'llash uchun zarur bo'lgan asosiy bilimlarni taqdim etadi.

13- mavzu. Tasniflash algoritmlari: KNN, Decision Trees

Ushbu mavzu tasniflash algoritmlari bo'yicha KNN (K-Nearest Neighbors) va Decision Trees (Qaror daraxtlari)ning asosiy tushunchalarini o'rganishga qaratilgan. Talabalar KNN algoritmining ishlash prinsiplarini, ya'ni ma'lumotlar nuqtalari o'rtasidagi masofalarni qanday hisoblash va tasniflash uchun eng yaqin qo'shnilarni tanlashni bilib oladilar.

Shuningdek, Qaror daraxtlari algoritmi, uning tuzilishi, qoidalari va qanday qilib qarorlar qabul qilishda foydali ekanligi haqida ma'lumot beriladi. Mavzu davomida talabalar har ikkala algoritmni amaliy misollar yordamida qo'llashni, ularning afzalliklari va kamchiliklarini taqqoslashni o'rganadilar. Bu mavzu tasniflash jarayonlarida algoritmlarni tanlash va ulardan foydalanishni tushunishga yordam beradi.

14- mavzu. Regression tahlili va uning turlari

Ushbu mavzu regression tahlilining asosiy printsiplari va turli turlarini o'rganishga qaratilgan. Talabalar regression tahlilining maqsadi, uning statistik model sifatida qanday ishlashi va ma'lumotlar orasidagi bog'lanishlarni aniqlashdagi ahamiyatini tushunadilar. Mavzu davomida oddiy va ko'p o'zgaruvchan regression modellari, shuningdek, polinom, logistika va boshqa regression turlari bilan tanishish imkoniyatiga ega bo'lishadi. Talabalar amaliy misollar orqali regression tahlilini qanday o'tkazish, natijalarni baholash va bashorat qilish jarayonlarini o'rganadilar. Ushbu mavzu ma'lumotlar tahlilida regression tahlilining o'rnini va qo'llanilishini yaxshiroq tushunishga yordam beradi.

15- mavzu. Deep Learning va sun'iy neyron tarmoqlar

Ushbu mavzu Deep Learning (chuqur o'rganish) va sun'iy neyron tarmoqlarining asosiy konseptlarini o'rganishga qaratilgan. Talabalar sun'iy neyron tarmoqlarining tuzilishi, ishlash prinsiplari va ularning ma'lumotlarni tahlil qilishdagi roli bilan tanishadilar. Mavzu davomida chuqur o'rganish algoritmlari, arxitekturasi (masalan, konvolyutsion va takroriy neyron tarmoqlar), hamda ularning tasvirni tanib olish, tilni qayta ishlash va boshqa sohalardagi ilovalari ko'rib chiqiladi. Shuningdek, talabalar Deep Learningning amaliy qo'llanilishlari va uning zamonaviy texnologiyalar bilan integratsiyasini o'rganadilar. Bu mavzu chuqur o'rganish bo'yicha nazariy va amaliy bilimlarni shakllantirishda muhim ahamiyatga ega.

16- mavzu. TensorFlow va Keras bilan ishlash

Ushbu mavzu TensorFlow va Keras kutubxonalari orqali sun'iy intellekt va chuqur o'qitish modellarini yaratish va rivojlantirish jarayonini o'rganishga qaratilgan. Talabalar TensorFlowning asosiy tushunchalari, Kerasning soddaligi va qulayligi yordamida qanday qilib tez va samarali modellar yaratish mumkinligini bilib oladilar. Mavzu davomida neyron tarmoqlarning arxitekturasi, qatlamlar, yo'qotish funksiyalari, optimizatsiya usullari va modelni baholash jarayonlari tushuntiriladi. Talabalar amaliy misollar orqali real hayotdagi muammolarni hal qilishda TensorFlow va Kerasni qanday ishlatishni o'rganadilar. Bu mavzu sun'iy intellekt sohasida professional ko'nikmalarni rivojlantirish uchun zarur bo'lgan bilimlarni taqdim etadi.

17- mavzu. Deep Learning loyihalari va amaliyot

Ushbu mavzu deep learning texnologiyalaridan foydalangan holda loyihalarni amalga oshirish jarayoniga qaratilgan. Talabalar turli xil deep learning arxitekturallari, masalan, sun'iy neyron tarmoqlari, konvolyutsion neyron tarmoqlari (CNN) va rekurrent neyron tarmoqlari (RNN) bilan tanishadilar. Mavzu davomida amaliy misollar orqali deep learning loyihalarini yaratish, ma'lumotlarni tayyorlash, modelni o'qitish va natijalarni baholash ko'nikmalari rivojlantiriladi. Talabalar deep learning texnologiyalarining real dunyo ilovalarini, jumladan, tasvirni tanib olish, tabiat tilini qayta ishlash va tavsiflash kabi misollarni ko'rib chiqadilar. Ushbu mavzu, shuningdek, deep learning jarayonidagi qiyinchiliklarni hal qilish va muvaffaqiyatli loyihalarni amalga oshirish uchun zarur bo'lgan amaliy tajribani beradi.

18- mavzu. Apache Kafka va Spark Streaming

Ushbu mavzu Apache Kafka va Spark Streaming texnologiyalarini o'rganishga qaratilgan. Talabalar Kafka arxitekturasi, ma'lumotlarni real vaqtda uzatish jarayoni va uning o'ziga xos imkoniyatlari bilan tanishadilar. Kafka yordamida katta hajmdagi ma'lumotlarni uzatish, saqlash va qayta ishlash jarayoni tushuntiriladi. Shuningdek, Spark Streaming yordamida real vaqt rejimida ma'lumotlarni tahlil qilish va olingan natijalarni qanday qilib tez va samarali tarzda boshqarish mumkinligi muhokama qilinadi. Mavzu davomida talabalar ushbu texnologiyalarning birgalikda qanday ishlashini o'rganib, amaliy misollar yordamida real vaqtda ma'lumotlarni uzatish va tahlil qilish bo'yicha ko'nikmalarni rivojlantiradilar. Bu mavzu katta ma'lumotlar va real vaqt tahlili sohasida zamonaviy yondashuvlarni tushunishga yordam beradi.

19- mavzu. Real vaqt ma'lumotlar tahlili va dasturlash

Ushbu mavzu real vaqt ma'lumotlar tahlili va dasturlash jarayonini o'rganishga qaratilgan. Talabalar real vaqt rejimida ma'lumotlarni yig'ish, tahlil qilish va ulardan qimmatli bilimlar olish uchun zarur bo'lgan texnologiyalar va metodologiyalar bilan tanishadilar. Mavzu davomida talabalar real vaqt tahlilining ahamiyati, uning amaliy qo'llanilishi, shuningdek, turli xil dasturlash tillari va vositalari (masalan, Python, Apache Kafka) yordamida ma'lumotlarni qanday tezda qayta ishlash va tahlil qilish mumkinligini o'rganadilar. Amaliy mashg'ulotlar orqali talabalar real vaqt ma'lumotlar tahlilining asosiy usullari va strategiyalarini qo'llash ko'nikmalarini rivojlantiradilar. Bu mavzu, shuningdek, real vaqt ma'lumotlari bilan ishlashda muammolarni hal qilishga yordam beradigan tajribalarni ham taqdim etadi.

20- mavzu. Flink va real vaqt dasturlari

Ushbu mavzu Apache Flink texnologiyasi va real vaqt dasturlarini yaratish jarayonini o'rganishga qaratilgan. Talabalar Flink arxitekturasi, uning asosiy komponentlari va ma'lumotlarni oqim shaklida qayta ishlash usullarini bilib oladilar. Mavzu davomida Flinkning qanday qilib real vaqtli ma'lumotlarni yig'ish, tahlil qilish va ularga javob berish uchun ishlatilishi muhokama qilinadi. Talabalar amaliy misollar orqali Flinkni o'rnatish, dasturlar yaratish va ularni optimallashtirish bo'yicha ko'nikmalarni rivojlantiradilar. Ushbu mavzu real vaqtli dasturlarni yaratish uchun zarur bo'lgan bilimlarni taqdim etadi va Flinkning katta ma'lumotlar analitika sohasidagi ahamiyatini tushunishga yordam beradi.

21- mavzu. ETL jarayonlari: Ma'lumotlarni yig'ish, tozalash va tayyorlash

Ushbu mavzu ETL (Extract, Transform, Load) jarayonlarining asosiy bosqichlarini o'rganishga qaratilgan, ya'ni ma'lumotlarni yig'ish, tozalash va tayyorlash. Talabalar ma'lumotlarni turli manbalardan qanday qilib to'plash, to'plangan ma'lumotlarni tozalash jarayoni (noto'g'ri yoki keraksiz ma'lumotlarni olib tashlash) va ularni tahlil qilish uchun qanday tayyorlash zarurligini bilib oladilar. Mavzu davomida ETL jarayonining ahamiyati, ma'lumotlar integratsiyasi va sifatini oshirishda qanday rol o'ynashi muhokama qilinadi. Talabalar amaliy misollar orqali ETL jarayonlarini amalga oshirish bo'yicha ko'nikmalarini rivojlantiradilar, bu esa ularga ma'lumotlarni samarali boshqarish va tahlil qilish uchun zarur bilimlarni beradi.

22- mavzu. Ma'lumotlar pipelines va Airflow bilan avtomatlashtirish

Ushbu mavzu ma'lumotlar pipelines'larini yaratish va Airflow yordamida avtomatlashtirish jarayonini o'rganishga qaratilgan. Talabalar ma'lumotlarni yig'ish, tozalash va tahlil qilish jarayonlarini qanday qilib avtomatlashtirish mumkinligini bilib oladilar. Airflow dasturiy ta'minoti orqali ma'lumotlar oqimlarini yaratish, rejalashtirish va boshqarish imkoniyatlari ko'rib chiqiladi. Mavzu davomida talabalar Airflow'da DAG (Directed Acyclic Graph) tushunchasi, operatsiyalarni yaratish va ularni boshqarish bo'yicha amaliy ko'nikmalarni rivojlantiradilar. Ushbu mavzu ma'lumotlar pipelines'larini avtomatlashtirish orqali samaradorlikni oshirish va vaqtni tejash uchun zarur bo'lgan bilimlarni taqdim etadi.

23- mavzu. Data Warehousing va Data Lakes tushunchalari

Ushbu mavzu Data Warehousing (ma'lumot omborlari) va Data Lakes (ma'lumot ko'llari) tushunchalarini o'rganishga qaratilgan. Talabalar ma'lumot omborlarining tuzilishi, funksiyalari va ularning analitik maqsadlarda qanday ishlatilishini bilib oladilar. Shuningdek, ma'lumot ko'llarining xususiyatlari, ularning foydalari va qanday qilib katta hajmdagi tuzilmali va tuzilmasiz ma'lumotlarni saqlash imkoniyatlari muhokama qilinadi. Mavzu davomida Data Warehousing va Data Lakes o'rtasidagi farqlar, ularning foydalanish holatlari va qaysi vaziyatlarda qaysi birini tanlash kerakligi ham ko'rib chiqiladi. Bu mavzu katta ma'lumotlar arxitekturasida har ikkala tushuncha va ularning amaliy qo'llanilishiga oid zarur bilimlarni taqdim etadi.

24- mavzu. Data Science va Big Data uchun amaliy case studies

Ushbu mavzu Data Science va Big Data sohasida amaliy case study-larni o'rganishga qaratilgan. Talabalar turli sohalarda (masalan, moliya, sog'liqni saqlash, marketing va

ishlab chiqarish) muvaffaqiyatli qo'llanilgan Data Science va Big Data strategiyalarini tahlil qiladilar. Case study-lar orqali talabalar real hayotdagi muammolarni hal qilish uchun ma'lumotlarni qanday yig'ish, tahlil qilish va natijalarni qanday vizualizatsiya qilish kerakligini bilib olishadi. Ushbu amaliy tadqiqotlar orqali talabalar Data Science jarayonlarining amaliyotdagi ahamiyatini tushunadilar va o'z loyihalarida qanday qilib muvaffaqiyatli yechimlarni qo'llash mumkinligini o'rganadilar. Mavzu talabalarni analitik fikrlash va muammolarni hal qilish ko'nikmalarini rivojlantirishga yordam beradi.

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

“Katta ma'lumotlar texnologiyalari (Big Data)” fani bo'yicha amaliy mashg'ulotlar uchun quyidagi mavzular tavsiya etiladi:

1-mavzu. Katta ma'lumotlar va Data Science-ga kirish

Ushbu mavzu katta ma'lumotlar texnologiyalari va Data Science asoslarini o'rganishni boshlaydi. Ma'ruza davomida talabalar katta ma'lumotlar tushunchasi, ularning turlari, va ahamiyati bilan tanishadilar. Katta ma'lumotlar nafaqat hajmi katta bo'lgan ma'lumotlar to'plami, balki ularning tezligi, xilma-xilligi va o'zgaruvchanligi bilan ajralib turishini tushunadilar. Shuningdek, Data Science tushunchasi, uning jarayonlari va asosiy elementlari: ma'lumotlarni yig'ish, tahlil qilish, modellashtirish va natijalarni vizualizatsiya qilish kabilar o'rganiladi. Talabalar katta ma'lumotlarning zamonaviy biznes va texnologiyalar sohasidagi muhim rolini tushunib olishlari uchun amaliy misollar va ilovalar ko'rib chiqiladi. Bu mavzu Data Science va katta ma'lumotlar bo'yicha chuqurroq bilimlarni o'zlashtirish uchun mustahkam asos bo'ladi.

2- mavzu. Katta ma'lumotlar texnologiyalarining qo'llanilishi

Ushbu mavzu katta ma'lumotlar texnologiyalarining zamonaviy sanoat va ilmiy sohalarda qanday qo'llanilishini o'rganishga qaratilgan. Ma'ruza davomida talabalar katta ma'lumotlardan foydalanish orqali turli xil sohalarda qanday muammolarni hal qilish mumkinligini tushunadilar. Ular moliya, sog'liqni saqlash, marketing, ishlab chiqarish, ta'lim, logistika va hukumat tashkilotlaridagi amaliyotlardan misollar ko'radilar.

3- mavzu. Data Science jarayoni va hayotiy tsikli

Ushbu mavzu Data Science jarayonining asosiy bosqichlari va hayotiy tsiklini o'rganishga qaratilgan. Ma'ruza davomida talabalar Data Science loyihasining turli bosqichlari bilan tanishadilar: ma'lumotlarni yig'ish, tozalash va tayyorlash, tahlil qilish, modellashtirish, natijalarni baholash va vizualizatsiya qilish. Har bir bosqichning o'ziga xosligi, qiyinchiliklari va muhim jihatlari tushuntiriladi. Talabalar ushbu jarayonlarni amaliy misollar yordamida o'rganib, Data Science hayotiy tsiklida qanday qilib ma'lumotlardan qimmatli bilimlar olish mumkinligini bilib olishadi. Mavzu Data Science loyihalarini muvaffaqiyatli amalga oshirish uchun zarur bo'lgan ko'nikmalarni shakllantirishga yordam beradi.

4- mavzu. Ma'lumotlarni yig'ish va to'plash usullari Ma'lumotlarni tozalash va tayyorlash

Ushbu mavzu ma'lumotlarni yig'ishning asosiy usullari va manbalarini o'rganishga qaratilgan. Talabalar web scraping, API-lar orqali ma'lumot olish, ma'lumotlar bazalaridan

foydalanish va boshqa manbalar orqali ma'lumotlarni yig'ish usullarini bilib oladilar. Shuningdek, ma'lumotlarni to'plash jarayonida duch keladigan qiyinchiliklar va ularni hal qilish yo'llari ham tushuntiriladi. Mavzu Data Science loyihalarini amalga oshirishda sifatli va ishonchli ma'lumotlar yig'ishning ahamiyatini tushunishga yordam beradi. Ma'lumotlarni tozalash – bu Big Data tahlilining boshlang'ich bosqichi bo'lib, noto'g'ri, chala yoki keraksiz ma'lumotlarni aniqlash va ularni to'g'ri formatga keltirishni o'z ichiga oladi.

Ma'lumotlarni tayyorlash esa tahlil va modellashtirish uchun zarur bo'lgan jarayon bo'lib: Turli manbalardagi ma'lumotlar birlashtiriladi.

Kategoriyalar raqamli formatga o'tkaziladi (masalan, bir-hot encoding).

Qarashli qiymatlar (outliers) bilan ishlanadi.

Ushbu jarayonlar to'g'ri va aniq tahlil qilish, modellarni sifatli tuzish uchun asos yaratadi. Ma'lumotlarni tozalash va tayyorlashda Python yoki Apache Spark kabi texnologiyalar qo'llaniladi.

5- mavzu. Python yordamida ma'lumotlarni tahlil qilish: pandas va NumPy

Ushbu mavzu Python dasturlash tilida ma'lumotlarni tahlil qilish uchun asosiy kutubxonalar — pandas va NumPy bilan tanishishga qaratilgan. Talabalar pandas yordamida ma'lumotlarni ko'rish, tozalash, manipulyatsiya qilish va tahlil qilish jarayonlarini o'rganadilar. Shuningdek, NumPy kutubxonasi yordamida ko'p o'lchovli ma'lumotlar bilan ishlash va matematik hisob-kitoblarni amalga oshirish usullari ko'rsatiladi. Ushbu mavzu, talabalarga ma'lumotlarni samarali tahlil qilish va ushbu jarayonni avtomatlashtirishga yordam beradigan kuchli vositalarni o'zlashtirish imkonini beradi.

6- mavzu. Ma'lumotlarni tasvirlash: Matplotlib va Seaborn

Ushbu mavzu ma'lumotlarni vizualizatsiya qilishda Matplotlib va Seaborn kutubxonalaridan foydalanishni o'rganishga qaratilgan. Talabalar Matplotlib yordamida grafiklar, diagrammalar va boshqa vizual tasvirlarni qanday yaratish mumkinligini bilib oladilar. Seaborn esa ma'lumotlarni tasvirlash uchun qo'shimcha qulayliklar va estetikani taqdim etadi, bu esa murakkab ma'lumotlarni tushunishni osonlashtiradi. Mavzu davomida talabalar ma'lumotlarni vizualizatsiya qilishda muhim jihatlarni, grafikalar yaratish usullari va tasvirlarning ahamiyatini o'rganadilar, shuningdek, amaliy misollar yordamida o'z bilimlarini mustahkamlash imkoniyatiga ega bo'lishadi.

7- mavzu. Data visualization vositalari: Plotly va Dash Amaliy vizualizatsiya loyihalari

Ushbu mavzu ma'lumotlarni vizualizatsiya qilishda Plotly va Dash kabi vositalardan foydalanishni o'rganishga bag'ishlangan. Talabalar Plotly kutubxonasi yordamida grafikalar va diagrammalar yaratish jarayonini bilib olishadi. Ushbu vosita yordamida interaktiv va estetik jihatdan jozibador grafikalar tayyorlash imkoniyati ko'rib chiqiladi. Dash esa Plotly asosida qurilgan, veb-ilovalar yaratish uchun mo'ljallangan freymorkdir. Talabalar Dash yordamida interaktiv vizualizatsiyalarni qanday yaratish va ularga foydalanuvchi interfeyslarini qo'shishni o'rganadilar. Mavzu ma'lumotlarni taqdim etish va tahlil qilishda vizualizatsiya vositalarining ahamiyatini tushunishga yordam beradi. Amaliy vizualizatsiya loyihalarini yaratish va ularni amalga oshirish jarayoniga qaratilgan. Talabalar ma'lumotlarni vizualizatsiya qilish uchun eng yaxshi amaliyotlar, vositalar va texnikalar bilan tanishadilar. Loyihalar davomida grafikalar, diagrammalar va interaktiv vizualizatsiyalarni yaratish orqali ma'lumotlarni taqdim etish ko'nikmalari rivojlantiriladi. Mavzu talabalarga vizualizatsiyaning ma'lumotlarni tushunishni osonlashtirishdagi

ahamiyatini ko'rsatadi va ularning o'z loyihalarini muvaffaqiyatli amalga oshirishlari uchun zarur bo'lgan amaliy tajribani beradi.

8- mavzu. Hadoop va uning asoslari HDFS (Hadoop Distributed File System) va ma'lumotlarni taqsimlas

Ushbu mavzu Hadoop texnologiyasining asoslari va uning katta ma'lumotlar boshqarishdagi rolini o'rganishga qaratilgan. Talabalar Hadoop arxitekturasi, asosiy komponentlari, jumladan HDFS (Hadoop Distributed File System) va MapReduce bilan tanishadilar. Shuningdek, Hadoopning tarixi, rivojlanishi va uning katta ma'lumotlar tahlilidagi afzalliklari ham muhokama qilinadi. Mavzu davomida talabalar Hadoopni o'rnatish va asosiy operatsiyalarni bajarish bo'yicha amaliy ko'nikmalarni ham rivojlantiradilar. Bu mavzu katta ma'lumotlar texnologiyalarida Hadoopni muvaffaqiyatli ishlatish uchun zarur bo'lgan bilimlarni taqdim etadi. HDFS (Hadoop Distributed File System) ning asosiy printsiplari va ma'lumotlarni taqsimlash jarayonini o'rganishga qaratilgan. Talabalar HDFS arxitekturasi, uning komponentlari va ma'lumotlarni qanday qilib tarqatilgan holda saqlash va boshqarish haqida bilib oladilar. Mavzu davomida ma'lumotlarni taqsimlashning ahamiyati, HDFSning yuqori darajadagi mavjudligi va chidamliligi, shuningdek, katta ma'lumotlar tizimlarida HDFSning qanday ishlashi muhokama qilinadi. Talabalar amaliy misollar orqali HDFSni o'rnatish, ma'lumotlarni yuklash va ularni boshqarish bo'yicha ko'nikmalarni rivojlantiradilar. Bu mavzu katta ma'lumotlar infratuzilmasida HDFSning rolini tushunishga yordam beradi.

9- mavzu. Apache Spark va uning imkoniyatlari

Ushbu mavzu Apache Spark texnologiyasining asoslari va uning katta ma'lumotlar tahlilidagi imkoniyatlarini o'rganishga qaratilgan. Talabalar Spark arxitekturasi, uning komponentlari, jumladan Spark Core, Spark SQL, Spark Streaming, MLlib va GraphX bilan tanishadilar. Mavzu davomida Apache Sparkning yuqori tezlikda ma'lumotlarni qayta ishlash, parallellik va tarqatilgan hisoblash imkoniyatlari muhokama qilinadi. Talabalar Spark yordamida katta ma'lumotlar to'plamini tahlil qilish, real vaqt rejimida ma'lumotlarni qayta ishlash va mashinani o'rganish modellari yaratish bo'yicha amaliy ko'nikmalarni rivojlantiradilar. Bu mavzu Apache Sparkni katta ma'lumotlar ekotizimida muvaffaqiyatli ishlatish uchun zarur bo'lgan bilimlarni taqdim etadi.

10- mavzu. SQL va NoSQL ma'lumotlar bazalari bilan ishlash va MongoDB Cassandra bilan tanishish

Ushbu mavzu SQL (Structured Query Language) va NoSQL (Not Only SQL) ma'lumotlar bazalari bilan ishlashni o'rganishga qaratilgan. Talabalar SQL asosidagi ma'lumotlar bazalarini, jumladan, ular qanday tuzilishi, ma'lumotlarni saqlash va qayta ishlash usullarini bilib oladilar. Shuningdek, NoSQL ma'lumotlar bazalarining xilma-xilligi, ularning afzalliklari va qanday holatlarda qo'llanilishi haqida ham ma'lumot beriladi. Mavzu davomida talabalar SQL va NoSQL tizimlari o'rtasidagi farqlarni, ularning ishlash printsiplari va qaysi holatlarda qaysi turini tanlash kerakligini o'rganadilar. Amaliy mashg'ulotlar orqali talabalar SQL so'roqlari yozish va NoSQL ma'lumotlar bazalari bilan ishlash bo'yicha ko'nikmalarini rivojlantiradilar. Bu mavzu ma'lumotlar bazalari bilan samarali ishlash uchun zarur bo'lgan nazariy va amaliy bilimlarni taqdim etadi. Talabalar har ikkala tizimning asosiy xususiyatlari, arxitekturasi va ishlash printsiplari bilan tanishadilar. MongoDB va Cassandra o'rtasidagi farqlar, ularning ma'lumotlarni saqlash va boshqarish usullari, shuningdek, qaysi holatlarda ushbu tizimlardan foydalanishning

afzalliklari va kamchiliklari muhokama qilinadi. Mavzu davomida talabalar MongoDB va Cassandra bilan ishlash bo'yicha amaliy misollar ko'rib, asosiy CRUD (Create, Read, Update, Delete) operatsiyalarini bajarish ko'nikmalarini rivojlantiradilar. Bu mavzu talabalarni NoSQL ma'lumotlar bazalarini tanlash va ishlatish bo'yicha chuqurroq bilimlar bilan ta'minlaydi.

11- mavzu. Cloud Computing va katta ma'lumotlar (AWS, Azure, Google Cloud)

Ushbu mavzu bulutli hisoblash (Cloud Computing) va katta ma'lumotlar (Big Data) texnologiyalarining integratsiyasini o'rganishga qaratilgan. Talabalar Amazon Web Services (AWS), Microsoft Azure va Google Cloud Platform kabi asosiy bulut xizmatlari ta'minotchilari bilan tanishadilar. Mavzu davomida bulutli xizmatlarning turli xil modellarini (IaaS, PaaS, SaaS) o'rganish va katta ma'lumotlarni saqlash, qayta ishlash va tahlil qilishda bulutli yechimlarning afzalliklari muhokama qilinadi. Talabalar bulutli platformalarda katta ma'lumotlarni qanday qilib samarali boshqarish va ulardan foydalangan holda tahlil qilishni bilib olishadi. Bu mavzu bulutli hisoblash va katta ma'lumotlar sohasida zamonaviy yechimlar va amaliyotlarni tushunishga yordam beradi.

12- mavzu. Machine Learning tushunchalari va amaliyot

Ushbu mavzu Machine Learning (ML) tushunchalarini va ularni amaliyotda qanday qo'llashni o'rganishga qaratilgan. Talabalar Machine Learningning asosiy prinsiplarini, turli algoritmlarni (masalan, nazoratli, nazoratsiz va yarim nazoratli o'rganish) va ulardan qaysi holatlarda foydalanish mumkinligini bilib oladilar. Mavzu davomida talabalar ML jarayonining asosiy bosqichlarini, jumladan ma'lumotlarni tayyorlash, modelni o'qitish, baholash va natijalarni interpretatsiya qilishni o'rganadilar. Amaliy mashg'ulotlar orqali talabalar real hayotdagi muammolarni hal qilish uchun ML modellarini yaratish va qo'llash ko'nikmalarini rivojlantiradilar. Bu mavzu Machine Learningga kirish va uni amaliyotda muvaffaqiyatli qo'llash uchun zarur bo'lgan asosiy bilimlarni taqdim etadi.

13- mavzu. Tasniflash algoritmlari: KNN, Decision Trees

Ushbu mavzu tasniflash algoritmlari bo'yicha KNN (K-Nearest Neighbors) va Decision Trees (Qaror daraxtlari)ning asosiy tushunchalarini o'rganishga qaratilgan. Talabalar KNN algoritmining ishlash prinsiplarini, ya'ni ma'lumotlar nuqtalari o'rtasidagi masofalarni qanday hisoblash va tasniflash uchun eng yaqin qo'shnilarni tanlashni bilib oladilar.

Shuningdek, Qaror daraxtlari algoritmi, uning tuzilishi, qoidalari va qanday qilib qarorlar qabul qilishda foydali ekanligi haqida ma'lumot beriladi. Mavzu davomida talabalar har ikkala algoritmni amaliy misollar yordamida qo'llashni, ularning afzalliklari va kamchiliklarini taqqoslashni o'rganadilar. Bu mavzu tasniflash jarayonlarida algoritmlarni tanlash va ulardan foydalanishni tushunishga yordam beradi.

14- mavzu. Regression tahlili va uning turlari

Ushbu mavzu regression tahlilining asosiy printsiplari va turli turlarini o'rganishga qaratilgan. Talabalar regression tahlilining maqsadi, uning statistik model sifatida qanday ishlashi va ma'lumotlar orasidagi bog'lanishlarni aniqlashdagi ahamiyatini tushunadilar. Mavzu davomida oddiy va ko'p o'zgaruvchan regression modellari, shuningdek, polinom, logistika va boshqa regression turlari bilan tanishish imkoniyatiga ega bo'lishadi. Talabalar amaliy misollar orqali regression tahlilini qanday o'tkazish, natijalarni baholash va bashorat qilish jarayonlarini o'rganadilar. Ushbu mavzu ma'lumotlar tahlilida regression tahlilining o'rini va qo'llanilishini yaxshiroq tushunishga yordam beradi.

15- mavzu. Deep Learning, sun'iy neyron tarmoqlar va ularning turlari

Ushbu mavzu Deep Learning (chuqur o'rganish) va sun'iy neyron tarmoqlarining asosiy konseptlarini o'rganishga qaratilgan. Talabalar sun'iy neyron tarmoqlarining tuzilishi, ishlash prinsiplari va ularning ma'lumotlarni tahlil qilishdagi roli bilan tanishadilar. Mavzu davomida chuqur o'rganish algoritmlari, arxitekturasi (masalan, konvolyutsion va takroriy neyron tarmoqlar), hamda ularning tasvirni tanib olish, tilni qayta ishlash va boshqa sohalaridagi ilovalari ko'rib chiqiladi. Shuningdek, talabalar Deep Learningning amaliy qo'llanilishlari va uning zamonaviy texnologiyalar bilan integratsiyasini o'rganadilar. Bu mavzu chuqur o'rganish bo'yicha nazariy va amaliy bilimlarni shakllantirishda muhim ahamiyatga ega.

16- mavzu. TensorFlow va Keras bilan ishlash hamda PyTorch kutubxonasi asoslari

Ushbu mavzu TensorFlow va Keras kutubxonalari orqali sun'iy intellekt va chuqur o'qitish modellarini yaratish va rivojlantirish jarayonini o'rganishga qaratilgan. Talabalar TensorFlowning asosiy tushunchalari, Kerasning soddaligi va qulayligi yordamida qanday qilib tez va samarali modellar yaratish mumkinligini bilib oladilar. Mavzu davomida neyron tarmoqlarning arxitekturasi, qatlamlar, yo'qotish funksiyalari, optimizatsiya usullari va modelni baholash jarayonlari tushuntiriladi. Talabalar amaliy misollar orqali real hayotdagi muammolarni hal qilishda TensorFlow va Kerasni qanday ishlatishni o'rganadilar. Bu mavzu sun'iy intellekt sohasida professional ko'nikmalarni rivojlantirish uchun zarur bo'lgan bilimlarni taqdim etadi.

17- mavzu. Deep Learning loyihalari va amaliyot

Ushbu mavzu deep learning texnologiyalaridan foydalangan holda loyihalarni amalga oshirish jarayoniga qaratilgan. Talabalar turli xil deep learning arxitekturalari, masalan, sun'iy neyron tarmoqlari, konvolyutsion neyron tarmoqlari (CNN) va rekurrent neyron tarmoqlari (RNN) bilan tanishadilar. Mavzu davomida amaliy misollar orqali deep learning loyihalarini yaratish, ma'lumotlarni tayyorlash, modelni o'qitish va natijalarni baholash ko'nikmalari rivojlantiriladi. Talabalar deep learning texnologiyalarining real dunyo ilovalarini, jumladan, tasvirni tanib olish, tabiat tilini qayta ishlash va tavsiflash kabi misollarni ko'rib chiqadilar. Ushbu mavzu, shuningdek, deep learning jarayonidagi qiyinchiliklarni hal qilish va muvaffaqiyatli loyihalarni amalga oshirish uchun zarur bo'lgan amaliy tajribani beradi.

18- mavzu. Apache Kafka, Spark Streaming va real-time ma'lumotlarni qayta ishlash

Ushbu mavzu Apache Kafka va Spark Streaming texnologiyalarini o'rganishga qaratilgan. Talabalar Kafka arxitekturasi, ma'lumotlarni real vaqtda uzatish jarayoni va uning o'ziga xos imkoniyatlari bilan tanishadilar. Kafka yordamida katta hajmdagi ma'lumotlarni uzatish, saqlash va qayta ishlash jarayoni tushuntiriladi. Shuningdek, Spark Streaming yordamida real vaqt rejimida ma'lumotlarni tahlil qilish va olingan natijalarni qanday qilib tez va samarali tarzda boshqarish mumkinligi muhokama qilinadi. Mavzu davomida talabalar ushbu texnologiyalarning birgalikda qanday ishlashini o'rganib, amaliy misollar yordamida real vaqtda ma'lumotlarni uzatish va tahlil qilish bo'yicha ko'nikmalarni rivojlantiradilar. Bu mavzu katta ma'lumotlar va real vaqt tahlili sohalarida zamonaviy yondashuvlarni tushunishga yordam beradi.

19- mavzu. Real vaqt ma'lumotlar tahlili va dasturlash

Ushbu mavzu real vaqt ma'lumotlar tahlili va dasturlash jarayonini o'rganishga qaratilgan. Talabalar real vaqt rejimida ma'lumotlarni yig'ish, tahlil qilish va ulardan qimmatli bilimlar

olish uchun zarur bo'lgan texnologiyalar va metodologiyalar bilan tanishadilar. Mavzu davomida talabalar real vaqt tahlilining ahamiyati, uning amaliy qo'llanilishi, shuningdek, turli xil dasturlash tillari va vositalari (masalan, Python, Apache Kafka) yordamida ma'lumotlarni qanday tezda qayta ishlash va tahlil qilish mumkinligini o'rganadilar. Amaliy mashg'ulotlar orqali talabalar real vaqt ma'lumotlar tahlilining asosiy usullari va strategiyalarini qo'llash ko'nikmalarini rivojlantiradilar. Bu mavzu, shuningdek, real vaqt ma'lumotlari bilan ishlashda muammolarni hal qilishga yordam beradigan tajribalarni ham taqdim etadi.

20- mavzu. Flink va real vaqt dasturlari

Ushbu mavzu Apache Flink texnologiyasi va real vaqt dasturlarini yaratish jarayonini o'rganishga qaratilgan. Talabalar Flink arxitekturasi, uning asosiy komponentlari va ma'lumotlarni oqim shaklida qayta ishlash usullarini bilib oladilar. Mavzu davomida Flinkning qanday qilib real vaqqli ma'lumotlarni yig'ish, tahlil qilish va ularga javob berish uchun ishlatilishi muhokama qilinadi. Talabalar amaliy misollar orqali Flinkni o'rnatish, dasturlar yaratish va ularni optimallashtirish bo'yicha ko'nikmalarni rivojlantiradilar. Ushbu mavzu real vaqqli dasturlarni yaratish uchun zarur bo'lgan bilimlarni taqdim etadi va Flinkning katta ma'lumotlar analitika sohasidagi ahamiyatini tushunishga yordam beradi.

21- mavzu. ETL jarayonlari: Ma'lumotlarni yig'ish, tozalash va tayyorlash

Ushbu mavzu ETL (Extract, Transform, Load) jarayonlarining asosiy bosqichlarini o'rganishga qaratilgan, ya'ni ma'lumotlarni yig'ish, tozalash va tayyorlash. Talabalar ma'lumotlarni turli manbalardan qanday qilib to'plash, to'plangan ma'lumotlarni tozalash jarayoni (noto'g'ri yoki keraksiz ma'lumotlarni olib tashlash) va ularni tahlil qilish uchun qanday tayyorlash zarurligini bilib oladilar. Mavzu davomida ETL jarayonining ahamiyati, ma'lumotlar integratsiyasi va sifatini oshirishda qanday rol o'ynashi muhokama qilinadi. Talabalar amaliy misollar orqali ETL jarayonlarini amalga oshirish bo'yicha ko'nikmalarini rivojlantiradilar, bu esa ularga ma'lumotlarni samarali boshqarish va tahlil qilish uchun zarur bilimlarni beradi.

22- mavzu. Ma'lumotlar pipelines va Airflow bilan avtomatlashtirish

Ushbu mavzu ma'lumotlar pipelines'larini yaratish va Airflow yordamida avtomatlashtirish jarayonini o'rganishga qaratilgan. Talabalar ma'lumotlarni yig'ish, tozalash va tahlil qilish jarayonlarini qanday qilib avtomatlashtirish mumkinligini bilib oladilar. Airflow dasturiy ta'minoti orqali ma'lumotlar oqimlarini yaratish, rejalashtirish va boshqarish imkoniyatlari ko'rib chiqiladi. Mavzu davomida talabalar Airflow'da DAG (Directed Acyclic Graph) tushunchasi, operatsiyalarni yaratish va ularni boshqarish bo'yicha amaliy ko'nikmalarni rivojlantiradilar. Ushbu mavzu ma'lumotlar pipelines'larini avtomatlashtirish orqali samaradorlikni oshirish va vaqtni tejash uchun zarur bo'lgan bilimlarni taqdim etadi.

23- mavzu. Data Warehousing va Data Lakes tushunchalari

Ushbu mavzu Data Warehousing (ma'lumot omborlari) va Data Lakes (ma'lumot ko'llari) tushunchalarini o'rganishga qaratilgan. Talabalar ma'lumot omborlarining tuzilishi, funksiyalari va ularning analitik maqsadlarda qanday ishlatilishini bilib oladilar. Shuningdek, ma'lumot ko'llarining xususiyatlari, ularning foydalari va qanday qilib katta

hajmdagi tuzilmali va tuzilmasiz ma'lumotlarni saqlash imkoniyatlari muhokama qilinadi. Mavzu davomida Data Warehousing va Data Lakes o'rtasidagi farqlar, ularning foydalanish holatlari va qaysi vaziyatlarda qaysi birini tanlash kerakligi ham ko'rib chiqiladi. Bu mavzu katta ma'lumotlar arxitekturasida har ikkala tushuncha va ularning amaliy qo'llanilishiga oid zarur bilimlarni taqdim etadi.

24- mavzu. Data Science va Big Data uchun amaliy case studies

Ushbu mavzu Data Science va Big Data sohasida amaliy case study-larni o'rganishga qaratilgan. Talabalar turli sohalarda (masalan, moliya, sog'liqni saqlash, marketing va ishlab chiqarish) muvaffaqiyatli qo'llanilgan Data Science va Big Data strategiyalarini tahlil qiladilar. Case study-lar orqali talabalar real hayotdagi muammolarni hal qilish uchun ma'lumotlarni qanday yig'ish, tahlil qilish va natijalarni qanday vizualizatsiya qilish kerakligini bilib olishadi. Ushbu amaliy tadqiqotlar orqali talabalar Data Science jarayonlarining amaliyotdagi ahamiyatini tushunadilar va o'z loyihalarida qanday qilib muvaffaqiyatli yechimlarni qo'llash mumkinligini o'rganadilar. Mavzu talabalarni analitik fikrlash va muammolarni hal qilish ko'nikmalarini rivojlantirishga 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.

"Katta ma'lumotlar texnologiyalari (Big Data)" fanidan mustaqil ta'lim uchun tavsiya etiladigan mavzular:

Ma'lumotlarni yig'ish va to'plash usullari Ma'lumotlarni tozalash va tayyorlash
Python yordamida ma'lumotlarni tahlil qilish: pandas va NumPy Ma'lumotlarni tasvirlash: Matplotlib va Seaborn Data visualization vositalari: Plotly va Dash Amaliy vizualizatsiya loyihalari Hadoop va uning asoslari HDFS (Hadoop Distributed File System) va ma'lumotlarni taqsimlash Apache Spark va uning imkoniyatlari SQL va NoSQL ma'lumotlar bazalari bilan ishlash MongoDB va Cassandra bilan tanishish Cloud Computing va katta ma'lumotlar (AWS, Azure, Google Cloud) Machine Learning tushunchalari va amaliyot Tasniflash algoritmlari: KNN, Decision Trees Regression tahlili va uning turlari Deep Learning va sun'iy neyron tarmoqlar TensorFlow va Keras bilan ishlash Deep Learning loyihalari va amaliyot Apache Kafka va Spark Streaming Real vaqt ma'lumotlar tahlili va dasturlash Flink va real vaqt dasturlari ETL jarayonlari: Ma'lumotlarni yig'ish, tozalash va tayyorlash Ma'lumotlar pipelines va Airflow bilan

	<p>avtomatlashtirish Data Warehousing va Data Lakes tushunchalari Data Science va Big Data uchun amaliy case studies</p> <p>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.</p> <p>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.</p> <p><i>* Izoh: Mustaqil ta'lim ishi mavzulari fan doirasida o'zgarishi mumkin.</i></p>
4.	<p>V. Fan o'qitilishining natijalari (shakllanadigan kompetensiyalar)</p> <p>Katta ma'lumotlar texnologiyalari (Big Data) fanini o'qitish jarayonida talabalar quyidagi kompetensiyalarni shakllantiradilar:</p> <ol style="list-style-type: none"> 1. Ma'lumotlarni tahlil qilish ko'nikmalari: Talabalar katta ma'lumotlar to'plamlarini tahlil qilish, natijalarni chiqarish va qaror qabul qilish jarayonida ma'lumotlardan foydalanish ko'nikmalarini egallaydilar. 2. Texnologiyalar bilan ishlash: HDFS, Hadoop, Spark, va boshqa katta ma'lumotlar texnologiyalarini o'z ichiga olgan amaliy bilimlarga ega bo'lish. 3. Ma'lumotlarni yig'ish va boshqarish: Talabalar turli usullar yordamida ma'lumotlarni yig'ish, saqlash va boshqarish jarayonlarini o'rganadilar, shu jumladan, SQL va NoSQL ma'lumotlar bazalari bilan ishlash. 4. Vizualizatsiya ko'nikmalari: Ma'lumotlarni samarali ravishda vizualizatsiya qilish uchun grafikalar, diagrammalar va boshqa vizual vositalardan foydalanish ko'nikmalarini rivojlantirish. 5. Muammolarni hal qilish: Katta ma'lumotlar yordamida real muammolarni aniqlash va ularni hal etish uchun kreativ fikrlash va muammolarni tahlil qilish qobiliyatini shakllantirish. 6. Tahliliy fikrlash: Talabalar ma'lumotlardan oqilona xulosalar chiqarish va ularni amaliyotda qo'llash qobiliyatini rivojlantiradilar.
5.	<p>VI. TA'LIM TEXNOLOGIYALARI VA METODLARI:</p> <p>Interaktiv ta'lim metodlari:</p> <p>Katta ma'lumotlar texnologiyalari (Big Data) fani bo'yicha interaktiv ta'lim metodlari talabalarni faol ishtirok etishga, muammolarni hal qilishga va analiy ko'nikmalarni rivojlantirishga yordam beradi. Quyidagi interaktiv ta'lim metodlari bu jarayonda qo'llanilishi mumkin:</p> <ol style="list-style-type: none"> 1. Guruh muhokamalari: Talabalar o'z fikrlarini baham ko'rishlari, turli nazariyalar va yondashuvlarni muhokama qilishlari uchun guruhlar tashkil etiladi. Bu usul fikr almashish va jamoaviy qaror qabul qilishni rivojlantiradi.

2. **Amaliy loyihalar:** Talabalar real hayotdagi masalalarni hal qilish uchun amaliy loyiha ustida ishlashadi. Bu ularga katta ma'lumotlar texnologiyalarini amalda qo'llash imkonini beradi.
3. **Simulatsiyalar:** Talabalar katta ma'lumotlar tahlilini amalga oshirish uchun simulatsiya vositalaridan foydalanadilar. Masalan, ma'lumotlar to'plamlarini tahlil qilish yoki vizualizatsiya qilish uchun dasturlar ishlatishlari mumkin.
4. **Viktorinalar va o'yinlar:** Ma'lumotlarni o'rganishni qiziqarli va samarali qilish uchun viktorina yoki o'yin formatida darslar o'tkazish. Bu usul o'zlashtirilgan bilimlarni mustahkamlashda yordam beradi.
5. **Katta ma'lumotlar tahlilidagi ishlar:** Talabalar kichik jamoalar shaklida birgalikda katta ma'lumotlar tahlilini amalga oshiradilar, natijalarni tahlil qilish va xulosalar chiqarish bo'yicha ish olib boradilar.
6. **Webinars va onlayn seminarlar:** Tajribali mutaxassislar tomonidan o'tkaziladigan webinarlar orqali talabalar soha yangiliklari, tendentsiyalari va amaliy uslublari bilan tanishadilar.
7. **Kross-darslar:** Katta ma'lumotlar bilan bog'liq fanlardan (masalan, statistika, dasturlash, ma'lumotlar bazalari) kross-darslar o'tkazish orqali talabalar bir necha sohada bilimlarini kengaytiradilar.
8. **Feedback va o'z-o'zini baholash:** Talabalar o'z ishlarini baholash va bir-birlariga fikr bildirish imkoniyatiga ega bo'lishadi, bu esa ularning o'z-o'zini baholash ko'nikmalarini rivojlantiradi.

Masalalar va misollar yechish:

Ushbu bo'limda talabalar katta ma'lumotlar texnologiyalariga doir muammolar va misollarni yechish orqali nazariy bilimlarini amaliy ko'nikmalarga aylantiradilar. Misollar yordamida real hayotdagi vaziyatlar asosida yechimlar taklif etiladi.

6.

VII. Kreditlarni olish uchun talablar:

Topshiriqni muvaffaqiyatli bajarish uchun asosiy ko'nikmalar:
 Nazariy bilimlar: Katta ma'lumotlar to'plamlarini tahlil qilish va ulardan foydali ma'lumotlar chiqarish qobiliyati. Bu statistik usullar, ma'lumotlarni o'rganish va analitik fikrlash ko'nikmalarini o'z ichiga oladi.
 Amaliy ko'nikmalar: Python, R yoki Scala kabi dasturlash tillarini bilish, bu tillar yordamida ma'lumotlarni tahlil qilish va modellash imkonini beradi.
 Ma'lumotlar bazasi bilan ishlash: SQL va NoSQL tizimlari bilan ishlash ko'nikmalari, ma'lumotlarni yig'ish, saqlash va boshqarish jarayonlarini amalga oshirish uchun zarurdir.
 Hadoop va Spark texnologiyalarini bilish: HDFS, MapReduce, Apache Spark kabi katta ma'lumotlar platformalarini tushunish va ularni qo'llash ko'nikmalari.
 Ma'lumotlarni vizualizatsiya qilish: Ma'lumotlarni samarali ravishda taqdim etish uchun vizualizatsiya vositalaridan (masalan, Tableau, Matplotlib, Seaborn) foydalanish qobiliyati

7.

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

Asosiy adabiyotlar

✓ Грас Д. Data Science. Наука о данных с нуля: Пер. с англ. - 2-е изд., перераб. и доп. - СПб.: БХВ-Петербург, 2021. - 416 с.: ил. ISBN 978-5-9775-6731-2

- 3/ Маккини У. Python и анализ данных / пер. с англ. А. А. Слинкина. – М.: ДМК Пресс, 2020. – 540 с.: ил. ISBN 978-5-94074-590-5
- 3- Дейтел Пол, Дейтел Харви Python: Искусственный интеллект, большие данные и облачные вычисления. — СПб.: Питер, 2020. — 864 с.: ил. — (Серия «Для профессионалов»). ISBN 978-5-4461-1432-0

Qo'shimcha adabiyotlar

4. Mirziyoyev Sh.M. Hozirgi zamon va Yangi O'zbekiston. - Toshkent: O'zbekiston, 2024.
5. Mirziyoyev Sh.M. Yangi O'zbekistonda taraqqiyot strategiyasi asosida demokratik islohotlar yo'lini qat'iy davom ettiramiz. 6-jild. - Toshkent: O'zbekiston, 2023
6. Радченко И.А, Николаев И.Н. Технологии и инфраструктура Big Data. – СПб: Университет ИТМО, 2018. – 52 с.
7. Радченко И.А, Николаев И.Н. Технологии и инфраструктура Big Data. – СПб: Университет ИТМО, 2018. – 52 с.
8. Y.Lakshimi Prasad Big Data Analytics Made Easy First Published by Notion Press 2016 Copyright © Y. Lakshmi Prasad 2016 All Rights Reserved. ISBN 978-1-946390-72-1
9. Силен Дэви, Мейсман Арно, Али Мохамед Основы Data Science и Big Data. Python и наука о данных. — СПб.: Питер, 2017.— 336 с.: ил. — (Серия «Библиотека программиста»). ISBN 978-5-496-02517-1
10. Dr. Rajni Bhalla INTRODUCTION TO BIG DATA Published By : Lovely Professional University Publisher Address: Lovely Professional University, Jalandhar Delhi GT road, Phagwara - 144411 Printer Detail: Lovely Professional University Edition Detail: (I) ISBN: 978-93-94068-26-1

Scopusdan maqolalar

11. Advancing multimodal emotion recognition in big data through prompt engineering and deep adaptive learning Abeer A. Wafal*, Mai M. Eldefrawi2† and Marwa S. Farhan2,3† Journal of Big Data Springer Nature E-ISSN:2196-1115
<https://www.scopus.com/sourceid/21100791292>
<https://journalofbigdata.springeropen.com/counter/pdf/10.1186/s40537-025-01264-w.pdf>
12. A guide to similarity measures and their data science applications Avivit Levy1*, B. Riva Shalom1† and Michal Chalamish1† Levy et al. Journal of Big Data (2025) 12:188 <https://doi.org/10.1186/s40537-025-01227-1>
<https://journalofbigdata.springeropen.com/counter/pdf/10.1186/s40537-025-01227-1.pdf>

duf

	<p style="text-align: center;">Axborot manbalar:</p> <p>https://mudk.tict.uz/</p> <p>https://www.w3schools.com/</p> <p>https://www.tutorialsteacher.com/</p> <p>https://uzbekdevs.uz/</p> <p>https://pythonworld.ru/</p>
8.	<p>Fanning o'quv dasturi Tarjimashunoslik, tilshunoslik va xalqaro jurnalistika oliy maktabida ishlab chiqilgan va 202<u>5</u> yil "<u>18</u>" <u>06</u> dagi <u>24</u>-sonli bayonnomasi bilan ma'qullangan.</p> <p>O'quv dasturi Toshkent davlat sharqshunoslik universiteti Kengashining 202__yil "<u>28</u>" <u>06</u> <u>11</u> sonli bayoni bilan tasdiqlangan.</p>
9.	<p>Fan/modul uchun ma'sullar:</p> <p>1. Allanyazov R.B.– TDSHU Tarjimashunoslik, tilshunoslik va xalqaro jurnalistika oliy maktabi o'qituvchisi</p>
10.	<p>Taqrizchilar:</p> <p>1. Usmanova Shoir Rustamovna - Tarjimashunoslik, tilshunoslik va xalqaro jurnalistika oliy maktabi professori, f.f.d.</p> <p>2. N.Abdurahmonova - O'zMU, Kompyuter lingvistikasi va amaliy tilshunoslik kafedrasi professori, f.f.d</p>