

АХМЕТ БАЙТҰРСЫНҰЛЫ АТЫНДАҒЫ ҚОСТАНАЙ Өңірлік университеті
КОСТАНАЙСКИЙ РЕГИОНАЛЬНЫЙ УНИВЕРСИТЕТ
ИМЕНИ АХМЕТ БАЙТҰРСЫНҰЛЫ
AKHMET BAITURSYNULY KOSTANAY REGIONAL UNIVERSITY



ҚОСЫМША БІЛІМ БЕРУ
БАҒДАРЛАМАСЫ
(MINOR)
ГЕНЕРАТИВТІ ЖАСАНДЫ ИНТЕЛЛЕКТ

ДОПОЛНИТЕЛЬНАЯ
ОБРАЗОВАТЕЛЬНАЯ ПРОГРАММА
(MINOR)
ГЕНЕРАТИВНЫЙ ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

ADDITIONAL
EDUCATIONAL PROGRAM
(MINOR)
GENERATIVE ARTIFICIAL INTELLIGENCE

Қостанай, 2025

ӘЗІРЛЕУШІЛЕР/ РАЗРАБОТЧИКИ/ DEVELOPERS:

Салыкова Ольга Сергеевна, бағдарламалық қамтамасыз ету кафедрасының қауымдастырылған профессоры, т.ғ.к./Салыкова Ольга Сергеевна, к.т.н., ассоциированный профессор кафедры программного обеспечения/ Salykova Olga PhD, Associate Professor of the Department of Software

Сатмаганбетова Жанар Зарлыкановна, бағдарламалық қамтамасыз ету кафедрасының аға оқытушысы / Сатмаганбетова Жанар Зарлыкановна, старший преподаватель кафедры программного обеспечения / Zhanar Satmaganbetova, senior lecturer of the Department of Software

Ахмет Байтұрсынұлы атындағы ҚӨУ-дың оқу-әдістемелік кеңес отырысында бекітілді, 28.01.2025 ж. №1 хаттама

Утвержден на заседании учебно-методического совета КРУ имени Ахмет Байтұрсынұлы, протокол от 28.01.2025 г. №1

Published by decision educational and methodical council of the Akhmet Baitursynuly Kostanay Regional University (Protocol №1 from 28.01.2025)

Курстың қысқаша сипаттамасы / Краткое описание курса / Brief description of the course:

Жасанды интеллект негіздері туралы білім IT мамандықтарының студенттері үшін білім беру бағдарламасының маңызды бөлігі болып табылады, өйткені жасанды интеллект (AI) және машиналық оқыту медицина мен қаржыдан бастап білім мен өнеркәсіпке дейінгі әртүрлі салаларға белсенді түрде енгізіліп жатқан технологияларды дамытуда шешуші рөл атқарады.

Деректер мен есептеу қуаттылығының қарқынды өсуі жағдайында AI-мен жұмыс істеу білімі мен дағдылары инновациялық шешімдерді жасау үшін қажет болады.

Машиналық оқытудың заманауи әдістерін, соның ішінде табиғи тілді өңдеу (NLP) және үлкен тілдік модельдермен (LLM) жұмыс істеуді меңгеру ерекше маңызды, бұл студенттерге осы технологияларды нақты жобалар мен ғылыми зерттеулерде тиімді қолдануға мүмкіндік береді.

Бұл курста өнеркәсіптік инженерияның негізгі аспектілері, генеративті AI және табиғи тілді өңдеу (NLP) және үлкен тілдік модельдермен жұмыс істеу міндеттері үшін қолданылатын машиналық оқытудың заманауи әдістері қарастырылады. Студенттер интеллектуалды жүйелерді құрудың принциптері мен алгоритмдерімен танысады, мәтінді құру, тілді түсіну және AI арқылы пайдаланушымен өзара әрекеттесу сияқты әртүрлі тапсырмалар үшін модельдерді әзірлеуді және оңтайландыруды үйренеді.

NLP-ді дамыту адамның сөйлеуін түсінуге, түсіндіруге және генерациялауға қабілетті тиімді жүйелерді құруға мүмкіндік береді, бұл ақылды көмекшілерді, чатботтарды, аударма жүйелерін және басқа да көптеген қосымшаларды құруға жаңа мүмкіндіктер ашады.

Minor B057 Ақпараттық технологиялар бағдарламалары тобының студенттеріне бағытталған.

Знания в области основы искусственного интеллекта является важной частью образовательной программы для студентов IT-специальностей, так как искусственный интеллект (ИИ) и машинное обучение играют ключевую роль в развитии технологий, которые уже активно внедряются в различные отрасли от медицины и финансов до образования и промышленности.

В условиях стремительного роста данных и вычислительных мощностей, знания и навыки работы с ИИ становятся необходимыми для создания инновационных решений.

Особенно важным является освоение современных методов машинного обучения, в том числе в области обработки естественного языка (NLP) и работы с большими языковыми моделями (LLM), что позволит студентам эффективно применять эти технологии в реальных проектах и научных исследованиях.

В данном курсе рассматриваются основные аспекты промпт-инженерии, генеративного ИИ и современные методы машинного обучения, применяемые для задач обработки естественного языка (NLP) и работы с большими языковыми моделями. Студенты познакомятся с принципами и алгоритмами, лежащими в основе создания интеллектуальных систем, научатся разрабатывать и оптимизировать модели для различных задач, таких как генерация текста, понимание языка и взаимодействие с пользователем через ИИ.

Развитие NLP делает возможным создание эффективных систем, способных понимать, интерпретировать и генерировать человеческую речь, что открывает новые горизонты для создания интеллектуальных ассистентов, чат-ботов, систем перевода и многих других приложений.

Minor ориентирован на обучающихся группы образовательных программ B057 Информационные технологии.

Knowledge of the basics of artificial intelligence is an important part of the educational program for IT students, as artificial intelligence (AI) and machine learning play a key role in the development of technologies that are already being actively implemented in various sectors from medicine and finance to education and industry.

With the rapid growth of data and computing power, knowledge and skills of working with AI are becoming necessary to create innovative solutions.

It is especially important to master modern machine learning methods, including in the field of natural language processing (NLP) and working with large language models (LLM), which will allow students to effectively apply these technologies in real projects and scientific research.

This course covers the main aspects of industrial engineering, generative AI, and modern machine learning methods used for natural language processing (NLP) tasks and working with large language models. Students will get acquainted with the principles and algorithms underlying the creation of intelligent systems, learn how to develop and optimize models for various tasks, such as text generation, language understanding and user interaction through AI.

The development of NLP makes it possible to create effective systems capable of understanding, interpreting and generating human speech, which opens up new horizons for the creation of intelligent assistants, chatbots, translation systems and many other applications.

Minor is aimed at students of the group educational programs B057 Information Technologies.

Мақсаты/Цель/Aim: студенттерді жасанды интеллекттің негізгі тұжырымдамалары мен әдістерімен таныстыру, сондай-ақ NLP мәселелерін шешу үшін машиналық оқытудың қолданылуын терең зерттеу/ ознакомить студентов с основными концепциями и методами искусственного интеллекта, а также углубленное изучение применения машинного обучения для решения задач NLP/ to introduce students to the basic concepts and methods of artificial intelligence, as well as an in-depth study of the application of machine learning to solve NLP problems.

Оқу міндеттері / Учебные задачи / Learning Objectives: Big Data, компьютерлік лингвистикадағы негізгі терминдер мен ұғымдарды білу; осы бағытта жұмыс істеу үшін негізгі бағдарламаларды қолдана білу / знать основные термины и понятия в Big Data, компьютерной лингвистике; уметь использовать основные программы для работы в данном направлении / know the basic terms and concepts in Big Data, computational linguistics; be able to use basic programs to work in this direction.

Қалыптастыру керек негізгі құзыреттер /Базовые компетенции, которые предстоит сформировать/ Basic competencies to be formed: NLP үшін машиналық оқыту құралдарымен жұмыс істеу дағдылары. Тілді өңдеу саласындағы нақты мәселелерді шешу үшін модельдердің тиімділігін бағалау және жасанды интеллект негізінде модельдерді оңтайландыру мүмкіндігі. Табиғи тілді өңдеу үшін жасанды интеллект технологиясын қолданатын қарапайым қосымшаларды әзірлеу/ навыки работы с инструментами машинного обучения для NLP. Умение оценивать эффективность моделей и оптимизировать модели на основе искусственного интеллекта для решения реальных задач в области обработки языка. Разрабатывать простые приложения, использующие технологии искусственного интеллекта для обработки естественного языка/ skills in working with NLP machine learning tools. The ability to evaluate the effectiveness of models and optimize models based on artificial intelligence to solve real-world problems in the field of language processing. Develop simple applications using artificial intelligence technologies for natural language processing

Оқытудың нәтижелері / Результаты обучения / Learning outcomes:

Өнеркәсіптік инженерия негіздерін және генеративті тіл модельдерімен жұмыс істеу принциптерін білу. NLP есептерін шешу үшін машиналық оқыту алгоритмдерін әзірлеу және енгізу. Мәтіндерден ақпаратты алу, жіктеу және генерациялау үшін NLP заманауи әдістері мен модельдерін қолданыңыз. Модельдер мен алгоритмдердің тиімділігін бағалау, сондай-ақ оларды теңшеу және оңтайландыру. Жасанды интеллектке негізделген шешімдерді әзірлеу үшін құралдармен және кітапханалармен жұмыс істеу.

Знать основы промпт-инженерии и принципы работы с генеративными языковыми моделями. Разработать и реализовать алгоритмы машинного обучения для решения задач NLP. Применять современные методы и модели NLP для извлечения, классификации и генерации информации из текстов. Оценивать эффективность моделей и алгоритмов, а также проводить их настройку и оптимизацию. Работать с инструментами и библиотеками для разработки решений на основе искусственного интеллекта.

Know the basics of industrial engineering and the principles of working with generative language models. To develop and implement machine learning algorithms for solving NLP problems. Apply modern NLP methods and models to extract, classify, and generate information from texts. Evaluate the effectiveness of models and algorithms, as well as adjust and optimize them. Work with tools and

libraries to develop solutions based on artificial intelligence.

Пәндер / Дисциплины / Disciplines:

1. Терең оқыту (5 кредит, ауызша емтихан)/ Глубокое обучение (5 кредитов, устный экзамен) / Deep learning (5 credits, oral examination)
2. Мәтінді автоматты өңдеу технологиялары (5 кредит, ауызша емтихан) / Технологии автоматической обработки текста (5 кредитов, УЭ) / Automatic text processing technologies (5 credits, oral examination)
3. Big Data Талдауы (5 кредит, ауызша емтихан) / Аналитика Big Data (5 кредитов, УЭ) / Big Data Analytics (5 credits, oral examination)

Еңбек сыйымдылығы / Трудоемкость / Labor intensity: 15 кредитов.

Пререквизиттер / Пререквизиты / Prerequisite: Ақпараттық-коммуникациялық технологиялар / Информационно-коммуникационные технологии / Information and Communication Technologies

Тындаушылардың ең аз саны / Минимальное количество слушателей / Minimum number of listeners: 16

Тындаушылардың максималды саны / Максимальное количество слушателей / Maximum number of listener: 50

Пән 1 / Дисциплина 1 / Discipline 1

Терең оқыту / Глубокое обучение / Deep learning

Пәннің қысқаша сипаттамасы / Краткое описание дисциплины / Discipline Summary:

Анықтамасы және, оның түрлері мен қолдану салалары. AI технологиялары мен қолданбаларына шолу. Генеративті AI. Генеративті ai модельдері мен платформалары, жетілдірілген генеративті AI әдістеріне шолу. Өнеркәсіптік инженерия. Өнеркәсіптерді құрудың озық әдістері. Мәтіндік және визуалды модельдердегі Өнеркәсіптік инженерия. Генеративті AI-дегі этикалық аспектілер мен шектеулер. Терең және машиналық оқытуға кіріспе. Нейрондық желілердің негіздері мен түсініктері. Нейрондық желілерді оқыту үшін деректерді дайындау. Зейін үлгілері және трансформаторлар. Трансформаторлардың архитектурасымен танысу. трансформаторлар және олардың NLP-дегі рөлі, үлкен тілдік модельдер (LLM), Kaz-LLM/ Определение ИИ, его типы и области применения. Обзор технологий и приложений ИИ. Генеративный ИИ. Модели и платформы генеративного ИИ, Обзор продвинутых методов генеративного ИИ. Промпт инженеринг. Продвинутое техники создания промптов. Промпт-инженерия в текстовых и визуальных моделях. Этические аспекты и ограничения в генеративном ИИ. Введение в машинное и глубокое обучение. Основы и концепции нейронных сетей. Подготовка данных для обучения нейронных сетей. Модели внимания и трансформеры. Знакомство с архитектурами трансформеров. трансформеры и их роль в NLP, большие языковые модели (LLM), Kaz-LLM/ The definition of And, its types and areas of application. Overview of AI technologies and applications. Generative AI. Models and platforms of generative AI, An overview of advanced methods of generative AI. Industrial engineering. Advanced techniques for creating prompta. Industrial engineering in text and visual models. Ethical aspects and limitations in generative AI. An introduction to machine learning and Deep learning. Fundamentals and concepts of neural networks. Data preparation for neural network training. Attention models and transformers. Introduction to transformer architectures. transformers and their role in NLP, large language models (LLM), Kaz-LLM

Сабақ түрлері / Виды занятий / Types of lesson: дәріс/ лекция/ lecture; зертханалық жұмыс/ лабораторная работа/ laboratory work

Оқытушы / Преподаватель / Teacher Сатмаганбетова Жанар Зарлыкановна, бағдарламалық қамтамасыз ету кафедрасының аға оқытушысы / Сатмаганбетова Жанар Зарлыкановна, старший преподаватель кафедры программного обеспечения / Zhanar

Пән 2 / Дисциплина 2 / Discipline 2

Мәтінді автоматты өңдеу технологиялары/ Технологии автоматической обработки текста/
Automatic text processing technologies

Пәннің қысқаша сипаттамасы / Краткое описание дисциплины / Discipline

Summary: NLP-ге кіріспе. NLP негізгі тұжырымдамалары. NLP-де модельдеу. Табиғи тілді модельдеудің математикалық әдістері. Python және мәтінді өңдеуге арналған негізгі құралдар мен кітапханалар. Мәтіндік деректерді жинау және өңдеу. Корпус лингвистикасы. Морфологиялық белгілеу. Талдау. Семантика-Wordnet. Қателерді түзету. Сөздерді санаттау және таңбалау. Сөйлеуді өңдеу. NLP үшін бұлтты қызметтер. Машиналық оқытудың рөлі NL P. сөздердің векторлық көрінісі. Алгоритмдер N-грамм. Жасырын Марков модельдері. Ақпарат алу. Мәтіндердің жіктелуі, тақырыптық модельдеу. NLP заманауи модельдері мен қосымшалары. NLP негізіндегі практикалық жобалар/ Введение в NLP. Основные концепции NLP. Моделирование в NLP. Математические методы моделирования естественного языка. Python и основные инструменты и библиотеки для обработки текста. Сбор и предобработка текстовых данных. Корпусная лингвистика. Морфологическая разметка. Синтаксический анализ. Семантика-Wordnet. Исправление опечаток. Категоризация и маркировка слов. Обработка речи. Облачные сервисы для NLP. Роль машинного обучения в NLP. Векторное представление слов. Алгоритмы N-граммы. Скрытые марковские модели. Извлечение информации. Классификация текстов, Тематическое моделирование. Современные модели и приложения NLP. Практические проекты на основе NLP/ An introduction to NLP. Basic NLP concepts. NLP modeling. Mathematical methods of natural language modeling. Python and basic text processing tools and libraries. Collecting and preprocessing text data. Corpus linguistics. Morphological markup. Syntactic analysis. Semantics-Wordnet. Correction of typos. Categorization and labeling of words. Speech processing. Cloud services for NLP. The role of machine learning in NLP Vector representation of words. N-gram algorithms. Hidden Markov models. Information extraction. Text classification, Thematic modeling. Modern NLP models and applications. NLP-based practical projects

Сабақ түрлері / Виды занятий / Types of lesson: дәріс/ лекция/ lecture; зертханалық жұмыс/ лабораторная работа/ laboratory work

Оқытушы /Преподаватель / Teacher Сатмаганбетова Жанар Зарлыкановна, бағдарламалық қамтамасыз ету кафедрасының аға оқытушысы / Сатмаганбетова Жанар Зарлыкановна, старший преподаватель кафедры программного обеспечения / Zhanar Satmaganbetova, senior lecturer of the Department of Software

Пән 3 / Дисциплина 3 / Discipline 3

Big Data Талдауы / Аналитика Big Data / Big Data
Analytics

Пәннің қысқаша сипаттамасы / Краткое описание дисциплины / Discipline

Summary: Үлкен деректерді талдау мамандары-бұл көзге көрінбейтін нәрсені қарастыруға көмектесетін мамандардың жаңа түрі. Олардың көмегімен деректермен жұмыс жаңа сапалы деңгейге шығады. Деректер талдаушысы өзінің бизнес тарихында талап етілмеген деректердің үлкен көлемін жинақтаған компанияларда қажет. Аналитик статистика және SQL сияқты стандартты құралдарды ғана емес, сонымен қатар үлкен деректермен жұмыс істеудің негізгі принциптерін білуі керек, Hadoop жүйесінің компоненттері мен big Data шешімдерін жүзеге асыруға арналған бұлтты платформалар туралы түсінікке ие болуы керек. Бұл курста әртүрлі құрылымдау дәрежесіндегі мәліметтермен жұмыс қарастырылады. Үлкен деректерді өңдеу әдістері мен технологиялары. Құрылымдық және құрылымданбаған деректерді өндіру. Негізгі бизнес мәселелерін шешу үшін статистикалық әдістерді, машиналық оқытуды және

жетілдірілген болжамды талдау әдістерін қолдану/ Специалисты в области анализа больших данных - это новый тип профессионалов, которые помогают рассмотреть то, что нельзя увидеть невооружённым глазом. С их помощью работа с данными выходит на новый качественный уровень. Аналитик данных нужен именно в тех компаниях, которые накопили за свою бизнес-историю огромное количество не востребуемых данных. Специалисту-аналитику нужно владеть не только стандартными инструментами вроде статистики и SQL, но и знать основные принципы работы с большими данными, иметь представление о компонентах системы Hadoop и облачных платформах для реализации решений по big data. В данном курсе рассматривается работа с данными в различной степени структуризации. Методы и технологии обработки больших данных. Интеллектуальный анализ структурированных и неструктурированных данных. Использование методов статистики, машинное обучение и продвинутые методы предиктивной аналитики для решения ключевых бизнес-задач/ Experts in the field of big data analysis are a new type of professionals who help to consider what cannot be seen with the naked eye. With their help, working with data reaches a new qualitative level. A data analyst is needed precisely in those companies that have accumulated a huge amount of unclaimed data over their business history. An analyst needs to know not only standard tools like statistics and SQL, but also the basic principles of working with big data, have an understanding of the components of the Hadoop system and cloud platforms for implementing big data solutions. This course examines working with data in varying degrees of structuring. Methods and technologies of big data processing. Intelligent analysis of structured and unstructured data. Using statistical methods, machine learning, and advanced predictive analytics techniques to solve key business problems

Сабақ түрлері / Виды занятий / Types of lesson: дәріс/ лекция/ lecture; зертханалық жұмыс/ лабораторная работа/ laboratory work

Оқытушы /Преподаватель / Teacher Салыкова Ольга Сергеевна, бағдарламалық қамтамасыз ету кафедрасының қауымдастырылған профессоры, т.ғ.к/Салыкова Ольга Сергеевна, к.т.н., ассоциированный профессор кафедры программного обеспечения/ Salykova Olga PhD, Associate Professor of the Department of Software