



Уважаемый слушатель!

Мы рады приветствовать тебя и благодарим за выбор Цифровой кафедры Казанского ГМУ!

В рамках программы дополнительного профессионального образования «**Прикладной анализ данных в здравоохранении**» Вы научитесь работать с медицинскими данными на более высоком уровне с помощью инструментов для анализа: SPSS Statistics и Stattech. Узнаете об анализе данных и том, как решать сложные задачи с помощью искусственного интеллекта и машинного обучения на Python.

Обучение проходит с использованием дистанционных образовательных технологий, благодаря чему **учиться Вы сможете в комфортном формате.**

В разработке курса принимали участие ведущие специалисты из области ИТ, MedTech - в ходе обучения у Вас будет **уникальная возможность** задать свои вопросы экспертам.

Не упустите шанс получить дополнительную ИТ-квалификацию, приходите к нам на обучение в Казанский ГМУ!

Важно: допуск к обучению возможен только после успешного прохождения отборочного этапа!

Что нужно сделать?

1. Прослушать курс «Цифровизация здравоохранения. Вводный курс» на образовательном портале Казанского ГМУ – изучить материалы, посвященные основам цифровых технологий в здравоохранении.

Ссылка на курс: <https://clck.ru/3NJ7dC>

2. Пройти тестирование – успешно сдать тест для проверки усвоенных знаний.

Результаты тестирования будут учитываться при **отборе** студентов на обучение в рамках "Цифровой кафедры".

Содержание программы курса

Модуль 1. Введение в прикладной анализ данных в здравоохранении

№ п/п	Наименование тем, виды учебной работы и краткое содержание учебного материала	Объем, часов
1.	<p>Тема 1 Основные понятия анализа данных в здравоохранении.</p> <p>Лекция: Определение анализа данных и его роль в здравоохранении. Типы данных в здравоохранении: структурированные и неструктурированные данные. Основные этапы анализа данных: сбор, обработка, анализ, интерпретация. Ключевые термины: дата-центрированное здравоохранение, электронные медицинские записи (ЭМЗ), большие данные, машинное обучение. Важность данных для принятия обоснованных решений в здравоохранении. Влияние анализа данных на улучшение качества медицинской помощи и оптимизацию процессов.</p> <p>Семинар: Вопросы и ответы по основным понятиям. Обсуждение примеров использования анализа данных в здравоохранении.</p> <p>Практическое занятие: Разбор реальных примеров успешного применения анализа данных в медицинской практике и здравоохранении.</p>	6
2	<p>Тема 2. Обзор области применения анализа данных в здравоохранении.</p> <p>Лекция 1: Роль анализа данных в клинических исследованиях. Примеры использования: контроль качества данных, анализ результатов, прогнозирование исходов. Применение анализа данных в эпидемиологии. Примеры использования: отслеживание распространения заболеваний, прогнозирование эпидемий, оценка эффективности профилактических мер.</p> <p>Лекция 2: Анализ данных в управлении здравоохранением. Примеры использования: оптимизация ресурсов, улучшение качества обслуживания, мониторинг показателей эффективности здравоохранения.</p> <p>Семинар: Обсуждение примеров клинических исследований, где анализ данных сыграл ключевую роль.</p> <p>Практическое занятие: Разбор кейса по анализу данных эпидемиологического исследования.</p>	12

№ п/п	Наименование тем, виды учебной работы и краткое содержание учебного материала	Объем, часов
	Самостоятельная работа: изучение примеров успешного применения анализа данных в управлении здравоохранением.	
3	<p>Тема 3. Применение специализированного программного обеспечения для анализа данных в здравоохранении.</p> <p>Лекция 1: Обзор основного специализированного программного обеспечения для анализа данных в здравоохранении. Примеры использования каждого инструмента.</p> <p>Лекция 2: Возможности ПО «Stattech» в анализе данных.</p> <p>Лекция 3: Возможности ПО «SPSS Statistics»</p> <p>Практическое занятие 1: Анализ данных в ПО «Stattech». Регистрация, формирование базы данных для загрузки. Инструменты анализа. Разбор кейса по анализу клинических данных.</p> <p>Практическое занятие 2: Анализ данных в ПО «Stattech». Разбор кейса по анализу данных эпидемиологического исследования. Анализ динамических рядов.</p> <p>Практическое занятие 3: Анализ данных в ПО «SPSS Statistics» Разбор кейса по анализу клинических данных.</p> <p>Самостоятельная работа: Анализ данных в ПО «Stattech». Разбор кейса по анализу клинических данных. Построение прогностической модели вероятности определенного исхода.</p>	16
4	<p>Тема 4. Анализ данных в Python.</p> <p>Лекция: Краткий обзор основных инструментов для анализа данных в Python. Примеры использования каждого инструмента.</p> <p>Практическое занятие 1: Установка Python и необходимых библиотек (pandas, numpy, matplotlib, scikit-learn). Инструктаж: пошаговые инструкции по установке и настройке программного обеспечения. Решение возможных проблем при установке.</p> <p>Практическое занятие 2: Настройка Jupyter Notebook для Python. Обзор основных возможностей среды разработки. Примеры работы в Jupyter Notebook и RStudio.</p> <p>Практическое занятие 3: Написание простого скрипта для обработки и</p>	8

№ п/п	Наименование тем, виды учебной работы и краткое содержание учебного материала	Объем, часов
	анализа данных. Визуализация данных с помощью matplotlib (Python)	
4	<p>Тема 5. Анализ медико-биологических данных.</p> <p>Лекция: Основные методы и техники анализа данных. Статистические методы в анализе медико-биологических данных</p> <p>Практическое занятие: Применение статистических методов (регрессия, ANOVA) с использованием Python. Обработка и анализ данных пациентов (с использованием библиотеки Pandas/NumPy)</p>	4
5	Промежуточная аттестация в формате тестирования	2

Модуль 2. Сбор и подготовка данных здравоохранения для анализа

№ п/п	Наименование тем, виды учебной работы и краткое содержание учебного материала	Объем, часов
1.	<p>Тема 1 Источники данных в здравоохранении: электронные медицинские записи, биобанки, клинические исследования</p> <p>Лекция 1: Источники данных в здравоохранении: электронные медицинские записи, базы страхования, данные из реальной/рутинной клинической практики, доказательства из реальной/рутинной клинической практики, эволюция восприятия данных и доказательств реальной клинической практики.</p> <p>Лекция 2. Структурированные электронные медицинские документы: виды, принципы построения. Справочники НСИ ЕГИСЗ. ЕГИСЗ (цифровой контур здравоохранения) – инструмент генерации, передачи, обработки и хранения данных.</p> <p>Лекция 3. Регистры и реестры в здравоохранении и медицине, базы данных здравоохранения. История разработки регистров. Виды, классификации медицинских регистров.</p> <p>Семинар: Вопросы и ответы по основным понятиям. Обсуждение примеров использования анализа данных в здравоохранении.</p> <p>Практическое занятие 1: Разбор реальных примеров успешного применения анализа данных в медицинской практике и здравоохранении.</p>	16

№ п/п	Наименование тем, виды учебной работы и краткое содержание учебного материала	Объем, часов
	<p>Практическое занятие 2: Работа в региональном сегменте ЕГИСЗ на примере медицинской информационной системы ГИС «ЭЗ РТ». Формирование СЭМД.</p> <p>Самостоятельная работа: изучение открытых источников данных здравоохранения</p>	
2	<p>Тема 2. Этические аспекты и конфиденциальность данных</p> <p>Лекция: Введение в этику работы с данными. Основные этические принципы в здравоохранении (принцип уважения, благодеяния, справедливости). Исторические примеры и важность этических норм (Таскигское исследование, случаи несанкционированного использования данных)</p> <p>Семинар: Обсуждение реальных примеров этических дилемм в работе с медико-биологическими данными. Обеспечение этичности и конфиденциальности данных в своей работе.</p>	4
3	<p>Тема 3. Конфиденциальность данных пациентов.</p> <p>Лекция 1. Понятие конфиденциальности данных в здравоохранении. Законы и нормативные акты, регулирующие конфиденциальность данных в здравоохранении.</p> <p>Лекция 2. Согласие пациента на использование данных. – Принцип информированного согласия. Процедуры получения согласия на использование данных для научных и медицинских исследований. Инструменты и атрибуты формирования информированных согласий в медицинских информационных системах.</p> <p>Практическое занятие 1: Исследование и сравнение требований различных нормативных актов в области конфиденциальности данных.</p> <p>Практическое занятие 2: Разработка шаблона согласия на использование данных, включающего все необходимые компоненты в медицинской информационной системе (на примере ГИС «ЭЗ РТ»).</p>	8
4	<p>Тема 4. Анонимизация и псевдонимизация данных.</p> <p>Лекция . Методы анонимизации данных (удаление идентифицирующей информации, генерация случайных идентификаторов). Псевдонимизация данных как способ защиты конфиденциальности.</p>	4

№ п/п	Наименование тем, виды учебной работы и краткое содержание учебного материала	Объем, часов
	<p>Практическое занятие: Анонимизация набора данных пациентов с использованием Python.</p> <p>Самостоятельная работа: принципы и инструменты обезличивания данных в здравоохранении.</p>	
5	<p>Тема 5. Управление доступом к данным</p> <p>Лекция 1. Принципы управления доступом к данным. Роли и права доступа в системах управления данными, медицинских информационных системах. Парольная политика.</p> <p>Лекция 2. Разбор реальных кейсов утечек данных и ошибок в обработке медико-биологических данных. Уроки, извлеченные из этих кейсов.</p> <p>Семинар: Обсуждение мер предосторожности и профилактики утечек данных в здравоохранении.</p> <p>Практическое занятие: Настройка ролей и прав доступа в симулированной системе управления данными, медицинских информационных системах, на примере «ГИС ЭЗ РТ».</p> <p>Самостоятельная работа: нормативно-правовое регулирование вопросов защиты данных в здравоохранении</p>	12
6	Промежуточная аттестация в формате тестирования	2

Модуль 3. Анализ больших данных в здравоохранении

№ п/п	Наименование тем, виды учебной работы и краткое содержание учебного материала	Объем, часов
1.	<p>Тема 1 Введение в большие данные в здравоохранении.</p> <p>Лекция 1: Определение больших данных (Big Data). 5V: объем (Volume), скорость (Velocity), разнообразие (Variety), достоверность (Veracity), ценность (Value). Источники больших данных в здравоохранении: электронные медицинские записи, геномные данные, данные из носимых устройств.</p> <p>Лекция 2. Роль больших данных в здравоохранении. Как большие данные трансформируют здравоохранение. Примеры применения:</p>	8

№ п/п	Наименование тем, виды учебной работы и краткое содержание учебного материала	Объем, часов
	<p>персонализированная медицина, мониторинг состояния здоровья, прогнозирование заболеваний.</p> <p>Семинар 1 Обсуждение особенностей больших данных и их примеров в здравоохранении.</p> <p>Практическое занятие 1: Анализ статьи или кейса, демонстрирующего использование больших данных в здравоохранении.</p>	
2	<p>Тема 2. Методы и инструменты для работы с большими данными</p> <p>Лекция 1: Методы машинного обучения: кластеризация, классификация, регрессия. Методы обработки и анализа текстовых данных: обработка естественного языка (NLP). Анализ потоковых данных. Инструменты для работы с большими данными: Apache Hadoop: основы и применение. Apache Spark: основы и применение. NoSQL базы данных: MongoDB, Cassandra.</p> <p>Лекция 2. Основы Apache Spark. Архитектура Apache Spark. Основные компоненты: Spark Core, Spark SQL, Spark Streaming, Mllib, GraphX.</p> <p>Семинар 1: Обсуждение преимуществ и ограничений различных методов анализа больших данных</p> <p>Практическое занятие 1: Установка и настройка Apache Spark. Краткий обзор интерфейса и функционала.</p> <p>Практическое занятие 2. Использование Apache Spark для обработки больших данных. Запуск Spark-сессии. Выполнение простых операций с данными (чтение, обработка, запись).</p>	10
3	<p>Тема 3. Машинное обучение</p> <p>Лекция 1. Основные алгоритмы машинного обучения. Применение машинного обучения для анализа данных в здравоохранении.</p> <p>Лекция 2. Анализ больших данных с использованием Hadoop. Архитектура Hadoop. Основные компоненты: HDFS, MapReduce, YARN.</p> <p>Практическое занятие 1: Создание и обучение модели машинного обучения.</p> <p>Оценка модели и интерпретация результатов.</p>	10

№ п/п	Наименование тем, виды учебной работы и краткое содержание учебного материала	Объем, часов
	Практическое занятие 2: Установка и настройка Hadoop. Запуск простых задач MapReduce.	
4	<p>Тема 4. Реальные кейсы применения анализа больших данных в здравоохранении</p> <p>Лекция. Примеры успешного применения анализа больших данных в здравоохранении. Кейсы и их результаты.</p> <p>Семинар: Обсуждение проблем и решений при анализе больших данных.</p> <p>Самостоятельная работа: изучение примеров успешного применения анализа данных в управлении здравоохранением.</p>	8
5	<p>Тема 5. Разработка комплексного проект</p> <p>Лекция. Постановка задачи для проекта (например, анализ данных пациентов для прогнозирования вероятности возникновения событий: заболеваний, операций, смерти и т.д.).</p> <p>Практическое занятие 1: Сбор и подготовка данных для анализа. Использование Apache Spark и/или Hadoop для обработки данных. Применение методов машинного обучения для анализа данных.</p> <p>Практическое занятие 2. Презентация результатов проектов. Представление результатов анализа данных. Обсуждение выводов и рекомендаций.</p> <p>Самостоятельная работа: подготовка проекта.</p>	12
6	Промежуточная аттестация в формате тестирования	2

Модуль 4. Искусственный интеллект и машинное обучение в здравоохранении

№ п/п	Наименование тем, виды учебной работы и краткое содержание учебного материала	Объем, часов
1.	<p>Тема 1 Основные понятия искусственного интеллекта и машинного обучения</p> <p>Лекция 1: Определение искусственного интеллекта и машинного обучения. Различие между ИИ, МО и глубоким обучением. Типы машинного обучения: обучение с учителем, обучение без учителя, обучение с подкреплением. Основные этапы разработки модели МО:</p>	12

№ п/п	Наименование тем, виды учебной работы и краткое содержание учебного материала	Объем, часов
	<p>сбор данных, подготовка данных, выбор и обучение модели, оценка модели, внедрение.</p> <p>Лекция 2: Применение ИИ и МО в здравоохранении. Примеры использования ИИ и МО: диагностика заболеваний, прогнозирование, персонализированная медицина.</p> <p>Семинар: Восприятие. Когнитивные процессы. Принятие решений. Как это работает в природе и как это можно описать? Философия вопроса. (Хебб и др.)</p> <p>Практическое занятие: Разбор статьи или кейса, демонстрирующего применение ИИ в здравоохранении.</p> <p>Самостоятельная работа: изучение примеров применения ИИ, СППР в здравоохранении и клинической практике</p>	
2	<p>Тема 2. Сбор и подготовка данных для обучения моделей ИИ</p> <p>Лекция 1. Проблемы и вызовы при сборе данных: качество данных, неполные данные, разнородность данных. Основные этапы подготовки данных: очистка данных, обработка пропущенных значений, кодирование категориальных переменных, нормализация и масштабирование.</p> <p>Практическое занятие 1: Сбор данных из открытых источников (например, набор данных UCI Machine Learning Repository).</p> <p>Практическое занятие 2: Очистка и подготовка данных с использованием Python. Обработка пропущенных значений и кодирование категориальных переменных.</p> <p>Самостоятельная работа: изучение открытых источников данных для обучения моделей ИИ</p>	10
3	<p>Тема 3. Выбор и обучение моделей машинного обучения</p> <p>Лекция 1. Основные алгоритмы машинного обучения. Алгоритмы обучения с учителем: линейная регрессия, логистическая регрессия, деревья решений, случайный лес, градиентный бустинг, нейронные сети. Алгоритмы обучения без учителя: кластеризация (k-means), ассоциативные правила. Алгоритмы глубокого обучения: нейронные сети, свёрточные нейронные сети (CNN), рекуррентные нейронные сети (RNN).</p>	8

№ п/п	Наименование тем, виды учебной работы и краткое содержание учебного материала	Объем, часов
	<p>Лекция 2. Обучение моделей. Процесс обучения модели: разделение данных на тренировочные и тестовые, кросс-валидация, гиперпараметры.</p> <p>Практическое занятие 1: Применение алгоритмов машинного обучения с использованием Python.</p> <p>Практическое занятие 2: Обучение моделей на примере набора данных здравоохранения. Кросс-валидация и настройка гиперпараметров.</p>	
4	<p>Тема 4. Оценка моделей машинного обучения.</p> <p>Лекция 1. Метрики оценки моделей. Метрики для задач классификации: точность, полнота, F1-score, ROC-кривая, площадь под кривой (AUC). Метрики для задач регрессии: среднеквадратичная ошибка (MSE), средняя абсолютная ошибка (MAE), коэффициент детерминации (R^2).</p> <p>Лекция 2. Интерпретация результатов оценки моделей. Примеры ошибок и способы их устранения</p> <p>Практическое занятие 1: Оценка модели классификации и регрессии на примере набора данных здравоохранения.</p> <p>Практическое занятие 2. Интерпретация результатов оценки модели и рекомендации по улучшению.</p>	8
5	<p>Тема 5. Применение нейронных сетей для анализа медико-биологических данных и данных здравоохранения.</p> <p>Лекция 1. Структура и принципы работы нейронных сетей. Основные типы нейронных сетей: полносвязные сети, свёрточные нейронные сети (CNN), рекуррентные нейронные сети (RNN).</p> <p>Лекция 2. Применение нейронных сетей в здравоохранении. Примеры использования нейронных сетей для анализа изображений (рентген, МРТ), последовательностей данных (анализ ЭКГ, геномных данных)</p> <p>Практическое занятие 1: Создание простой нейронной сети с использованием Python.</p> <p>Практическое занятие 2: Создание простой нейронной сети с</p>	10

№ п/п	Наименование тем, виды учебной работы и краткое содержание учебного материала	Объем, часов
	<p>использованием Python (продолжение)</p> <p>Практическое занятие 3. Обучение и оценка нейронной сети на примере данных здравоохранения (например, классификация изображений рентгеновских снимков).</p>	
6	Промежуточная аттестация в формате тестирования	2