Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Мухарямова Лайсан Музиповна
Должность: и.о.первого проректора
Дата подписания: 12.03.2026 18:04:43
Уникальный программный ключ:
b57b96507511d4669a7e8b1e807a3d3e7412a55d

Министерство здравоохранения Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования «Казанский государственный медицинский университет»

«УТВЕРЖДАЮ»
Ректор ФГБОУ ВО Казанский ГМУ
Минздрава России, профессор

_____А.С. Созинов

_____2024г

Дополнительная профессиональная программа
(программа профессиональной переподготовки)

Прикладной анализ данных в здравоохранении

2024 г.

Содержание

I. Общие положения 4
1. Нормативная правовая основа Программы: 4
2. Термины и определения, используемые в Программе 5
3. Требования к поступающим 7
II. Планируемые результаты обучения и структура Программы 8
Структура образовательных результатов 9
III. Учебный план Программы 10
IV. Календарный учебный график 11
V. Рабочие программы модулей (курсов, дисциплин) 12
Рабочая программа модуля 1. Введение в прикладной анализ данных в здравоохранении 13
Рабочая программа модуля 2. Сбор и подготовка данных здравоохранения для анализа 22
Рабочая программа модуля 3. Анализ больших данных в здравоохранении 31
Рабочая программа модуля 4. Искусственный интеллект и машинное обучение в здравоохранении 39
Рабочая программа практики 47
VI. Итоговая аттестация по Программе 52
VII. Завершение обучения по Программе 53

Аннотация

Дополнительная профессиональная программа профессиональной переподготовки ИТ-профиля «Прикладной анализ данных в здравоохранении» (далее – Программа) предназначена для обучающихся по специальностям и направлениям подготовки, не отнесённым к ИТ-сфере, здравоохранение и медицинские науки.

Целью профессиональной переподготовки является получение актуальной для здравоохранения дополнительной ИТ-квалификации 06.042 Специалист по большим данным.

Нормативный срок освоения программы 252 часа при очно-заочной (с применением электронного обучения, дистанционных образовательных технологий) форме подготовки.

Авторы и преподаватели:
Галиахметов А.И. – ассистент кафедры общественного здоровья и организации здравоохранения ФГБОУ ВО Казанский государственный медицинский университет» Минздрава России.
Искандаров И.Р. – к.м.н., доцент кафедры общественного здоровья и организации здравоохранения ФГБОУ ВО Казанский государственный медицинский университет» Минздрава России.
Альмухаметов А.А. – руководитель проектов ЗАО «Витакор»
Марапов Д.И. – к.м.н., генеральный директор ООО «Статтех»
Классен В.И. – д.т.н., профессор, генеральный директор АО «Радиокомпания «Вектор»
Просвиркин И.А. - к.т.н. – ИТ-директор ООО «ФтизисБиоМед»
Гогоберидзе Ю.Т. - старший разработчик ООО «ФтизисБиоМед»
Сабитов Р.А. - старший ML-инженер ООО «ФтизисБиоМед»
Фатихов А.Ф. - ML-инженер ООО «ФтизисБиоМед»

Программа рассмотрена и утверждена центральным координационно-методическим советом ФГБОУ ВО Казанский ГМУ Минздрава России на заседании № 7 от 9 июля 2024г.

Председатель ЦКМС,
Профессор Мухарямова Л.М.

I. Общие положения

1. Нормативная правовая основа Программы:

– Федеральный закон от 29 декабря 2012 года №273-ФЗ «Об образовании в Российской Федерации»;

– постановление Правительства Российской Федерации от 13 мая 2021 г. № 729 «О мерах по реализации программы стратегического лидерства «Приоритет-2030»;

– паспорт федерального проекта «Развитие кадрового потенциала ИТ-отрасли» национальной программы «Цифровая экономика Российской Федерации»;

– приказ Минцифры России от 29.12.2023 № 1180 «Об утверждении методик расчета показателей федеральных проектов «Развитие кадрового потенциала ИТ-отрасли» и «Обеспечение доступа в Интернет за счет развития спутниковой связи» национальной программы «Цифровая экономика Российской Федерации», а также внесении изменений в некоторые приказы Министерства цифрового развития, связи и массовых коммуникаций Российской Федерации об утверждении методик расчета показателей федеральных проектов национальной программы «Цифровая экономика Российской Федерации» (далее – приказ Минцифры России № 1180);

– приказ Минобрнауки России от 1 июля2013 г. № 499 «Об утверждении Порядка организации и осуществления образовательной деятельности по дополнительным профессиональным программам» (с изменениями, внесенными приказом Минобрнауки России о 15 ноября 2013 г. № 1244 «О внесении изменений в Порядок организации и осуществления образовательной деятельности по дополнительным профессиональным программам, утвержденный приказом Министерства образования и науки Российской Федерации от 1 июля 2013 г. № 499»);

– приказ Минтруда России от 12 апреля 2013 г. № 148н «Об утверждении уровней квалификации в целях разработки проектов профессиональных стандартов»;

– методические рекомендации по разработке основных профессиональных образовательных программ и дополнительных профессиональных программ с учетом соответствующих профессиональных стандартов (утв. Минобрнауки России 22 января 2015 г. № ДЛ-1/05вн);

– постановление Правительства Российской Федерации от 11 октября 2023 г. № 1678 «Об утверждении Правил применения организациями, осуществляющими образовательную деятельность, электронного обучения, дистанционных образовательных технологий при реализации образовательных программ»;

– федеральный государственный образовательный стандарт федерального государственного образовательного стандарта высшего образования по направлению подготовки 09.03.01 Информатика и вычислительная техника (уровень бакалавриата), утвержденного приказом Минобрнауки России от 12 января 2016 г. № 5 (далее вместе – ФГОС ВО));

– профессиональный стандарт 06.042 «Специалист по большим данным», утвержденный приказом Министерства труда и социальной защиты РФ от 6 июля 2020 года N 405н (далее – профессиональный стандарт)

2. Термины и определения, используемые в Программе

Дополнительная ИТ-квалификация – квалификация, приобретаемая в ходе освоения Программы обучающимися:

1) специальностей и направлений подготовки, отнесённых к ИТ-сфере, – в части формирования навыков использования и формирования цифровых компетенций, необходимых для выполнения нового вида профессиональной деятельности в соответствии с перечнем областей цифровых компетенций согласно приложению 1 к Методике расчета показателя «Количество обученных, получивших дополнительную ИТ-квалификацию на «цифровых кафедрах», утверждённой приказом Минцифры России № 1180 (далее – Методика расчета Показателя);

2) специальностей и направлений подготовки, не отнесённых к ИТ-сфере, – в области создания алгоритмов и компьютерных программ, пригодных для практического применения.

Специальности и направления подготовки, отнесённые к ИТ-сфере, – специальности и направления подготовки, перечисленные в перечне направлений подготовки (бакалавриат) и специальностей (специалитет) высшего образования в приложении 2 к Методике расчета Показателя.

Специальности и направления подготовки, не отнесённые к ИТ-сфере, – специальности и направления подготовки (бакалавриат, специалитет, магистратура, ординатура), не указанные в перечне направлений подготовки (бакалавриат) и специальностей (специалитет) высшего образования в приложении 2 к Методике расчета Показателя.

Цифровая компетенция (компетенция) – образовательный результат, формируемый при освоении Программы, необходимый для приобретения дополнительной ИТ-квалификации и выражающийся в осуществлении деятельности в области создания алгоритмов и компьютерных программ, пригодных для практического применения, выполнении нового вида профессиональной деятельности.

Целевой уровень сформированности компетенций – установленный Программой уровень сформированности компетенций в соответствии с Матрицей компетенций, актуальных для цифровой экономики, с приоритетом компетенций в ИТ-сфере.

Матрица цифровых компетенций – матрица компетенций, актуальных для цифровой экономики, с приоритетом компетенций в ИТ-сфере, разработанная Университетом Иннополис при участии ИТ-компаний и университетов-участников программы «Приоритет-2030», представляющая собой перечень компетенций, структурированный по сферам применения, типу компетенций, уровням их сформированности и характеристикам.

Знание (З) – информация о свойствах объектов, закономерностях процессов и явлений, правилах использования этой информации для принятия решений, присвоенная обучающимся на одном из уровней, позволяющих выполнять над ней мыслительные операции.

Умение (У) – освоенный субъектом способ выполнения действия, обеспечиваемый совокупностью приобретенных знаний и навыков; операция (действие), выполняемая определенным способом и с определенным качеством.

Опыт практической деятельности (ОПД) – образовательный результат, включающий выполнение обучающимся деятельности, завершающейся получением результата / продукта (элемента продукта), значимого при выполнении трудовой функции, в условиях реального производства или в модельной ситуации.

Дополнительная профессиональная программа профессиональной переподготовки (Программа) – комплекс основных характеристик образования (объем, содержание, планируемые результаты) и организационно-педагогических условий, который представлен в виде учебного плана, календарного учебного графика, рабочих программ учебных курсов, дисциплин (модулей), оценочных и методических материалов, а также программ учебной и производственной практик, стажировок и форм аттестации, иных компонентов и обеспечивает приобретение дополнительной квалификации. Программа может разрабатываться с учетом положений профессиональных стандартов, федеральных государственных образовательных стандартов, требований рынка труда (индустрии).

Рабочая программа – нормативный документ в составе Программы, регламентирующий взаимодействие преподавателя и обучающихся в ходе учебного процесса при реализации структурных элементов Программы (модуль, дисциплина, курс).

Профессиональный модуль (ПМ) – структурный элемент Программы, предназначенный для формирования определенных компетенций.

Учебная дисциплина (УД) – структурный элемент Программы, предназначенный для формирования знаний и умений в соответствующей сфере профессиональной деятельности.

Междисциплинарный курс (МДК) – структурный элемент Программы или программы профессионального модуля, предназначенный для формирования знаний и умений, объединенных по прагматическим основаниям с нарушением академических границ отраслей знаний.

Практика (практическая подготовка) – форма организации образовательной деятельности при освоении образовательной программы в условиях выполнения обучающимися определенных видов работ, связанных с будущей профессиональной деятельностью и направленных на формирование, закрепление, развитие практических навыков и компетенции по профилю соответствующей образовательной программы.

Стажировка – формирование и закрепление полученных в результате теоретической подготовки профессиональных знаний и умений в рамках выполнения практических заданий (функций) на базе профильной компании (организации). Допускается заключение срочных трудовых договоров, предусматривающих прохождение обучающимся оплачиваемой стажировки. Время прохождения стажировки целесообразно учитывать в качестве учебной или производственной практики.

Электронное обучение – организация образовательной деятельности с применением содержащейся в базах данных и используемой при реализации образовательных программ информации и обеспечивающих ее обработку информационных технологий, технических средств, а также информационно-телекоммуникационных сетей, обеспечивающих передачу по линиям связи указанной информации, взаимодействие обучающихся и педагогических работников.

Дистанционные образовательные технологии – это образовательные технологии, реализуемые в основном с применением информационно- телекоммуникационных сетей при опосредованном (на расстоянии) взаимодействии обучающихся и педагогических работников.

Фонды оценочных средств (ФОС) – совокупность оценочных средств, используемых на различных этапах педагогической диагностики.

Оценочные средства (ОС) – дидактические средства для оценки качества подготовленности обучающихся.

Оценка цифровых компетенций (ассесмент) – проводимая на платформе Минцифры России оценка уровня сформированности цифровых компетенций, состоящая из трёх этапов:

1) входная оценка – оценка входного уровня цифровых компетенций обучающихся, которая проводится на этапе зачисления и начала обучения по Программе.

2) промежуточная оценка – это оценка уровня сформированности цифровых компетенций обучающихся, которая проводится в процессе обучения по Программе.

3) итоговая оценка – оценка достижения обучающимися целевого уровня сформированности цифровых компетенций, которая проводится на этапе завершения обучения по Программе.

3. Требования к поступающим

К обучению по Программе допускаются обучающиеся по очной или по очно-заочной форме за счет бюджетных средств или по договорам об оказании платных образовательных услуг, освоившие программы специалитета в объеме не менее 4 курса (студенты 5 курса), и/или программы ординатуры (ординаторы) и магистры по специальностям и направлениям подготовки здравоохранение и медицинские науки.

4. Квалификационная характеристика выпускника

Выпускникам Программы присваивается дополнительная ИТ-квалификация в области создания алгоритмов и компьютерных программ, пригодных для практического применения.

Выпускник Программы будет готов к выполнению профессиональной деятельности, трудовой функции профессионального стандарта 06.042 «Специалист по большим данным»: анализ больших данных с использованием существующей в организации методологической и технологической инфраструктуры в качестве аналитика, исследователя данных, руководителя отдела по информационным технологиям, заместителя руководителя медицинской организации по цифровым технологиям, цифровой трансформации.

Квалификационный уровень по национальной рамке квалификаций: 6.

II. Планируемые результаты обучения и структура Программы

Получение дополнительной ИТ-квалификации 06.042 «Специалист по большим данным», обеспечивается формированием приведённых в таблице цифровых компетенций:

Наименование сферы	ID и наименование компетенции	Инструменты профессиональной деятельности	Целевой уровень формирования компетенций в Программе			
			Минимальный (исходный)	Базовый	Продвинутый	Экспертный
Анализ данных	11. Использует специализированное программное обеспечение для анализа медико-биологических данных	Stattech SPSS Statistics Python	-	-	Самостоятельно применяет оптимальные алгоритмы анализа медико-биологических данных, используя специализированное программное обеспечение. Выбирает и применяет оптимальные критерии и методы для описания данных и аналитики	-
Средства программной разработки	28. Применяет языки программирования для решения профессиональных задач	Python	-	Применяет языки программирования для решения профессиональных задач под контролем более опытных специалистов	-	-
Искусственный интеллект и машинное обучение	170. Осуществляет сбор и подготовку данных для обучения моделей искусственного интеллекта	Python	-	Осуществляет критический отбор данных, проверяя их на целостность и непротиворечивость. Использует методы поиска данных и достоверные источники данных	-	-

Структура образовательных результатов

Формирование цифровых компетенций, необходимых для получения обучающимися дополнительной ИТ-квалификации, обеспечивается последовательным формированием промежуточных образовательных результатов, начиная со знаний.

ID и формулировка целевого уровня формирования компетенций	Промежуточные образовательные результаты		
	Опыт практической деятельности (ОПД)	Умения (У)	Знания (З)
11. Самостоятельно применяет оптимальные алгоритмы анализа медико-биологических данных, используя специализированное программное обеспечение. Выбирает и применяет оптимальные критерии и методы для описания данных и аналитики	ОПД 1 Применяет для анализа данных здравоохранения программное обеспечение Stattech ОПД 2 Применяет для анализа данных здравоохранения программное обеспечение SPSS Statistics	У 1 Умеет выбирать оптимальные критерии и методы для описания данных и аналитики в здравоохранении	З 1. Знает основные понятия и терминологию анализа данных в здравоохранении. З 2. Знает основные методы анализа данных в здравоохранении
28. Применяет языки программирования для решения профессиональных задач	ОПД 1. Применяет язык программирования Python под контролем опытных наставников	У 1. Умеет выбирать типовые алгоритмы разработки	З 1. Знает принципы и основы программирования.
170. Осуществляет критический отбор данных, проверяя их на целостность и непротиворечивость. Использует методы поиска данных и достоверные источники данных	ОПД 1. Применяет Python на практике для решения задач в здравоохранении.	У 1. Умеет собирать и подготавливать данные, для обучения и оценки моделей машинного обучения.	З 1. Знает основные концепции и методы искусственного интеллекта (ИИ) и машинного обучения (МО).

III. Учебный план Программы

Объем Программы составляет 252 часа

Учебный план Программы определяет перечень, последовательность, общую трудоемкость разделов и формы контроля знаний.

Структурные элементы (разделы Программы)	Общая трудоемкость. часов	Обязательная аудиторная учебная нагрузка		Самостоя-тельная работа, часов	Практики, стажировки, часов	Промежуточная аттестация,
		всего, часов	в т.ч. практические занятия, часов			
Модуль 1: Введение в прикладной анализ данных в здравоохранении	48	38	22	8		2
Модуль 2: Сбор и подготовка данных здравоохранения для анализа	50	36	12	12		2
Модуль 3: Анализ больших данных в здравоохранении	50	38	14	10		2
Модуль 4: Искусственный интеллект и машинное обучение в здравоохранении	50	38	20	10		2
Практика / стажировка	50				50	
Итоговая аттестация в формате демонстрационного экзамена (включая подготовку к аттестации)	4			4		
Итого:	**252**	**150**	**70**	**44**	**50**	**8**

IV. Календарный учебный график

Календарный учебный график представляет собой график учебного процесса, устанавливающий последовательность и продолжительность обучения, включая практику /стажировку, и итоговой аттестации по месяцам, а также этапы ассесмента. При этом время, выделяемое на прохождение оценки сформированности цифровых компетенций, в общей трудоёмкости Программы, отражённой в Учебном плане, не учитывается.

Структурные элементы (разделы Программы)[1] и этапы ассесмента	**месяцы**								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Входная оценка цифровых компетенций	+								
Модуль 1: Введение в прикладной анализ данных в здравоохранении	+	+							
Модуль 2: Сбор и подготовка данных здравоохранения для анализа			+	+					
Промежуточная оценка цифровых компетенций				+					
Модуль 3: Анализ больших данных в здравоохранении					+	+			
Модуль 4: Искусственный интеллект и машинное обучение в здравоохранении							+	+	
Практика / стажировка									+
Итоговая оценка цифровых компетенций									+
Итоговая аттестация									+

V. Рабочие программы модулей (курсов, дисциплин)

Рабочие программы разрабатываются для структурных элементов (разделов) Программы, указанных в Структуре Программы и Учебном плане, и содержат:

– перечень тем, включающих лекции, семинары, мастер-классы, практические занятия, самостоятельную работу, консультации и иные виды учебной работы с указанием краткого содержания и трудоёмкости,

– образцы оценочных средств,

– методические материалы для преподавателей и обучающихся,

– сведения о кадровом обеспечении образовательного процесса.

Рабочая программа практики / стажировки предусматривает определение цели и задач практической деятельности обучающихся, площадку (площадки) прохождения практики, задания (индивидуальные или групповые), критерии оценки результатов практической деятельности обучающихся.

Министерство здравоохранения Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования «Казанский государственный медицинский университет»

««УТВЕРЖДАЮ»
Ректор ФГБОУ ВО Казанский ГМУ
Минздрава России, профессор

_____А.С. Созинов

«___»_____2024г

Рабочая программа модуля 1
ВВЕДЕНИЕ В ПРИКЛАДНОЙ АНАЛИЗ ДАННЫХ В ЗДРАВООХРАНЕНИИ

2024 г.

1. Область применения рабочей программы

Рабочая программа **модуля «Введение в прикладной анализ данных в здравоохранении»** (далее – рабочая программа) является частью дополнительной профессиональной программы профессиональной переподготовки **ИТ «Прикладной анализ данных в здравоохранении»** (далее – программа) и направлена на формирование цифровых компетенций:

- «Использует специализированное программное обеспечение для анализа медико- биологических данных» (ID 11 – продвинутый уровень):

Формулировка целевого уровня формирования компетенций	**Промежуточные образовательные результаты**		
	Опыт практической деятельности (ОПД)	**Умения (У)**	**Знания (З)**
Самостоятельно применяет оптимальные алгоритмы анализа медико-биологических данных, используя специализированное программное обеспечение. Выбирает и применяет оптимальные критерии и методы для описания данных и аналитики	ОПД 1 Применяет для анализа данных здравоохранения программное обеспечение Stattech ОПД 2 Применяет для анализа данных здравоохранения программное обеспечение SPSS Statistics ОПД 3. Применяет для анализа данных Python	У 1 Умеет выбирать оптимальные критерии и методы для описания данных и аналитики в здравоохранении	З 1. Знает основные понятия и терминологию анализа данных в здравоохранении. З 2. Знает основные методы анализа данных в здравоохранении

- «Осуществляет сбор и подготовку данных для обучения моделей искусственного интеллекта» (ID 170 – базовый уровень):

Формулировка целевого уровня формирования компетенций	**Промежуточные образовательные результаты**		
	Опыт практической деятельности (ОПД)	**Умения (У)**	**Знания (З)**
Осуществляет критический отбор данных, проверяя их на целостность и непротиворечивость. Использует методы поиска данных и достоверные источники данных	ОПД 1. Применяет Python на практике для решения задач в здравоохранении.	У 1. Умеет собирать и подготавливать данные, для обучения и оценки моделей машинного обучения.	З 1. Знает основные концепции и методы искусственного интеллекта (ИИ) и машинного обучения (МО).

Освоение рабочей программы является обязательным для всех обучающихся по Программе.

2. Структура и краткое содержание рабочей программы

№ п/п	Наименование тем, виды учебной работы и краткое содержание учебного материала	Объем, часов
1.	**Тема 1 Основные понятия анализа данных в здравоохранении.** 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. Семинар: Вопросы и ответы по основным понятиям. Обсуждение примеров использования анализа данных в здравоохранении. Практическое занятие: Разбор реальных примеров успешного применения анализа данных в медицинской практике и здравоохранении.	6
2	**Тема 2. Обзор области применения анализа данных в здравоохранении.** 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. Лекция 2: Анализ данных в управлении здравоохранением. Примеры использования: оптимизация ресурсов, улучшение качества обслуживания, мониторинг показателей эффективности здравоохранения. Семинар: Обсуждение примеров клинических исследований, где анализ данных сыграл ключевую роль. Практическое занятие: Разбор кейса по анализу данных эпидемиологического исследования. Самостоятельная работа: изучение примеров успешного применения анализа данных в управлении здравоохранением.	12
3	**Тема 3. Применение специализированного программного обеспечения для анализа данных в здравоохранении.**	16

№ п/п	Наименование тем, виды учебной работы и краткое содержание учебного материала	Объем, часов
	Лекция 1**:** Обзор основного специализированного программного обеспечения для анализа данных в здравоохранении. Примеры использования каждого инструмента. Лекция 2: Возможности ПО «Stattech» в анализе данных. Лекция 3: Возможности ПО «SPSS Statistics» Практическое занятие 1: Анализ данных в ПО «Stattech». Регистрация, формирование базы данных для загрузки. Инструменты анализа. Разбор кейса по анализу клинических данных. Практическое занятие 2: Анализ данных в ПО «Stattech». Разбор кейса по анализу данных эпидемиологического исследования. Анализ динамических рядов. Практическое занятие 3: Анализ данных в ПО «SPSS Statistics» Разбор кейса по анализу клинических данных. Самостоятельная работа: Анализ данных в ПО «Stattech». Разбор кейса по анализу клинических данных. Построение прогностической модели вероятности определенного исхода.	
4	**Тема 4. Анализ данных в Python.** Лекция: Краткий обзор основных инструментов для анализа данных в Python. Примеры использования каждого инструмента. 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. Практическое занятие 2: Настройка Jupyter Notebook для Python. Обзор основных возможностей среды разработки. Примеры работы в Jupyter Notebook и RStudio. Практическое занятие 3: Написание простого скрипта для обработки и анализа данных. Визуализация данных с помощью matplotlib (Python)	8
4	**Тема 5. Анализ медико-биологических данных.** Лекция: Основные методы и техники анализа данных. Статистические методы в анализе медико-биологических данных Практическое занятие: Применение статистических методов (регрессия,	4

№ п/п	Наименование тем, виды учебной работы и краткое содержание учебного материала	Объем, часов
	ANOVA) с использованием Python. Обработка и анализ данных пациентов (с использованием библиотеки Pandas/NumPy)	
5	Промежуточная аттестация в формате тестирования	2

3. Учебно-тематический план рабочей программы

№ п/п	Наименование и краткое содержание структурного элемента (раздела) Программы	Количество часов		
		аудиторных		самостоятель-ной работы
		лекции, семинары	практические занятия	
1.	Тема 1 Основные понятия анализа данных в здравоохранении	2	2	-
2	Тема 2. Обзор области применения анализа данных в здравоохранении.	6	2	4
3	Тема 3. Применение специализированного программного обеспечения для анализа данных в здравоохранении.	6	6	4
4	Тема 4. Анализ данных в Python.	2	6	-
5	Тема 5. Анализ медико-биологических данных.	2	2	-
	Промежуточная аттестация	2		
	Итого	**48**		

4. Контроль и оценка результатов освоения рабочей программы

Образовательная организация высшего образования, реализующая рабочую программу, обеспечивает организацию и проведение текущий, промежуточный контроли демонстрируемых обучающимися образовательных результатов.

Текущий контроль проводится преподавателем на основе оценивания результатов практических работ и самостоятельной работы обучающихся. Промежуточный контроль проводится в форме тестирования. Формы и методы текущего и промежуточного контроля, критерии оценивания доводятся до сведения обучающихся в начале обучения.

Для текущего и промежуточного контроля создаются фонды оценочных средств (ФОС). ФОС включают в себя педагогические контрольно-измерительные материалы, предназначенные для определения соответствия (или несоответствия) индивидуальных образовательных достижений запланированным образовательным результатам.

4.1. Примеры оценочных средств.

Тестовые задания:

1. Что такое прикладной анализ данных в здравоохранении и какие задачи он решает?
- A) Описание демографических характеристик населения
- B) Сбор и анализ данных для улучшения клинических решений и результатов лечения
- C) Оценка финансовой эффективности медицинских учреждений
- D) Все выше перечисленное

2. Какие виды данных наиболее часто используются в прикладном анализе данных в

здравоохранении?

- A) Клинические данные (медицинские карты, результаты анализов)
- B) Административные данные (записи о госпитализации, расходы)
- C) Пациент-генерированные данные (опросы, данные носимых устройств)
- D) Все выше перечисленное

3. Какой метод анализа данных наиболее часто применяется для выявления скрытых закономерностей в больших наборах медицинских данных?

- A) Регрессионный анализ
- B) Анализ временных рядов
- C) Кластерный анализ
- D) Дискриминантный анализ

4. Какие виды «выбросов» данных существуют?

- А) Мягкие и жёсткие.
- Б) Близкие и далёкие.
- В) Сильные и слабые.
- Г) Правильного ответа нет.

5. Что из перечисленного является примером использования анализа данных в клинических исследованиях?

- А) Контроль качества данных.
- Б) Анализ результатов.
- В) Прогнозирование исходов.
- Г) Всё вышеперечисленное.

5. Образцы учебно-методических материалов для обучающихся и преподавателей

Примеры заданий на самостоятельную и практическую работу

5.1. Вопросы для собеседования

1. Что такое прикладной анализ данных в здравоохранении и какие задачи он решает?

- Описание: Этот вопрос проверяет базовые знания о предмете, включая определение прикладного анализа данных в контексте здравоохранения и его основные цели и задачи.
- Ответ: Прикладной анализ данных в здравоохранении — это использование методов анализа данных для улучшения качества медицинской помощи, прогнозирования исходов лечения, оптимизации ресурсов и снижения затрат. Основные задачи включают анализ клинических данных, оценку эффективности лечения, прогнозирование эпидемий и улучшение персонализированного ухода за пациентами.

2. Какие методы анализа данных наиболее часто применяются в здравоохранении? Приведите примеры.

- Описание: Этот вопрос проверяет знание основных методов анализа данных и их применения в здравоохранении.
- Ответ: Наиболее часто применяемые методы включают статистический анализ, машинное обучение, искусственный интеллект, обработку больших данных и визуализацию данных. Примеры: регрессионный анализ для предсказания рисков заболеваний, кластеризация для сегментации пациентов, нейронные сети для диагностики по изображениям.

3. Какую роль играют электронные медицинские записи (ЭМЗ) в анализе данных в здравоохранении?

- Описание: Этот вопрос оценивает понимание роли и значимости ЭМЗ в анализе данных.
- Ответ: Электронные медицинские записи играют ключевую роль в анализе данных в здравоохранении, так как они содержат большой объем структурированных и неструктурированных данных о пациентах, включая историю болезни, результаты анализов, диагностику и лечение. Эти данные используются для анализа и улучшения качества медицинской помощи, персонализации лечения и принятия обоснованных клинических решений.

5.2. Решение кейсов.

1. Анализ заболеваемости и прогнозирование эпидемий.
Описание задания: Используйте набор данных о заболеваемости гриппом за последние 10 лет в конкретном регионе. Проанализируйте тренды, определите сезонные колебания и разработайте модель для прогнозирования будущих вспышек эпидемий.
Цели:
- Провести временной анализ данных.
- Применить методы сезонной декомпозиции.
- Разработать и оценить модель прогнозирования (например, ARIMA, SARIMA, LSTM).
- Представить результаты в виде отчета и графиков.

2. Разработка модели для предсказания риска сердечно-сосудистых заболеваний.
Описание задания: Используйте данные пациентов, включая демографические данные, медицинскую историю, результаты анализов и образ жизни, чтобы разработать модель машинного обучения для предсказания риска сердечно-сосудистых заболеваний.
Цели:
- Подготовить данные для анализа (чистка, нормализация, работа с пропущенными значениями).
- Применить методы классификации (например, логистическая регрессия, случайный лес, градиентный бустинг).
- Оценить производительность моделей (метрики точности, ROC-кривая).
- Создать отчет с рекомендациями по использованию модели в клинической практике.

3. Оценка эффективности нового медицинского препарата.
Описание задания: Используйте клинические данные об эффективности нового препарата против диабета, собранные в ходе испытаний. Проведите анализ данных, чтобы определить, насколько эффективен новый препарат по сравнению с традиционными методами лечения.
Цели:
- Провести описательную статистику данных.
- Применить методы статистического анализа для оценки эффективности (например, t-тест, ANOVA).
- Выполнить регрессионный анализ для выявления факторов, влияющих на результат лечения.
- Сформировать отчет с выводами и рекомендациями.

4. Анализ базы данных пациентов с гипертонической болезнью.
Описание задания: Проведите анализ базы данных (файл в формате .xlsx), собранной в рамках исследования эффективности лечения пациентов с артериальной гипертонией. Дата проведения опроса исследуемых – 01.06.2024.

Цели:
- Подготовить данные для последующего анализа: исправить ошибки, привести содержимое столбцов в соответствие с типом переменных (категориальные или количественные).
- Определить инструмент для статистического анализа данных: вручную с применением калькулятора, электронные таблицы, статистическое программное обеспечение (Stattech, SPPS и пр.).
- Описать выборку по показателям: пол, возраст, профессиональный статус, наличие курения, наличие сопутствующих заболеваний. Категориальные показатели описать с помощью абсолютных значений и процентных долей, количественные показатели с помощью средних значений и стандартных отклонений.

1. Оценить зависимость АД до лечения от различных факторов.
2. Оценить связь курения и сопутствующих заболеваний.
3. Сформировать вывод по результатам анализа.

6. Кадровое обеспечение образовательного процесса

Галиахметов А.И. – ассистент кафедры общественного здоровья и организации здравоохранения ФГБОУ ВО Казанский государственный медицинский университет» МЗ РФ.

Искандаров И.Р. – к.м.н., доцент кафедры общественного здоровья и организации здравоохранения ФГБОУ ВО Казанский государственный медицинский университет» МЗ РФ.

Альмухаметов А.А. – руководитель проектов ЗАО «Витакор»

Марапов Д.И. – к.м.н., генеральный директор ООО «Статтех»

Васильева А.Г. – аналитик ЗАО «Витакор»

7. Материально-техническое обеспечение реализации рабочей программы

Необходимый для реализации программы модуля перечень материально-технического обеспечения включает в себя специально оборудованные помещения для проведения учебных занятий, в том числе:

аудитории, оборудованные мультимедийными и иными средствами обучения, позволяющими использовать информационные технологии, в количестве, позволяющем обучающимся осваивать умения и навыки, предусмотренные профессиональной деятельностью, индивидуально;

дисплейный класс, с мультимедийным и компьютерным оборудованием.

помещения для самостоятельной работы обучающихся оснащены компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду Университета.

Перечень информационных технологий, необходимых для освоения программы, включая перечень программного обеспечения и информационных справочных систем

Для достижения целей педагогического образования применяются следующие информационные технологии:

1. Образовательный портал дистанционного обучения Казанского ГМУ. Дистанционный курс в составе образовательного портала создан в системе MOODLE и содержит в себе лекции, презентации, задания, гиперссылки на первоисточники учебного материала, тесты / задания для самоконтроля, контрольные и итоговые тесты по курсу.
2. Пакет прикладных программ МойОфис.

3. Программа для ЭВМ StatTech (Статтех).
4. Программа для ЭВМ «SPSS Statistics»

Используемое программное обеспечение имеет лицензию и ежегодно и / или своевременно обновляется.

8. Информационное обеспечение реализации рабочей программы

8.1. Основная литература.

Статистика в биомедицине, фармации и фармацевтике : учебное пособие / Н. Н. Зубов, В. И. Кувакин, С. З. Умаров [и др.] ; под общей редакцией И. А. Наркевича ; Санкт-Петербургский государственный химико-фармацевтический университет. - Москва : КноРус, 2021. – 298.

8.2. Дополнительная литература.

Бююль А., Цефель П. SPSS: искусство обработки информации. Анализ статистических данных и восстановление скрытых закономерностей. — СПб.: ДиаСофтЮП, 2005.

Введение в биостатистику для медиков / С. Л. Плавинский ; Открытый институт здоровья. - М. : [б. и.], 2011. - 582, [2] с. : рис., табл. ; 21 см. - Библиогр.: с. 579-582.

Принципы и практика клинических исследований / Д. И. Галлин, Ф. П. Огнибене ; пер. с англ. под общ. ред. Г. Т. Сухих. - Москва : Практическая медицина, 2013. - XVIII, 451 с. : ил. ; 29 см. - Библиогр. в конце глав. - Пер. изд. : Principles and practice of clinical research / J. I. Gallin, F. P. Ognibene

Программирование на Python в примерах и задачах / Алексей Васильев. - Москва : Эксмо, 2021. - 616 с.

8.3. Интернет-ресурсы.

Единая образовательная платформа Университета Иннополис https://oneid.unionepro.ru
Официальный сайт Stattech https://stattech.ru/
Официальный сайт ФГБУ «Центральный научно-исследовательский институт организации и информатизации здравоохранения» МЗ РФ https://mednet.ru/
Сайт Федеральной электронной медицинской библиотеки − http://www.femb.ru
Сайт ФГБОУ ВО Казанского ГМУ Минздрава России – https://kazangmu.ru/
Сайт Научной электронной библиотеки − http://elibrary.ru/

Министерство здравоохранения Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования «Казанский государственный медицинский университет»

«УТВЕРЖДАЮ»
Ректор ФГБОУ ВО Казанский ГМУ
Минздрава России, профессор
_____А.С. Созинов

«__» _____ 2024 г.

Рабочая программа модуля 2

СБОР И ПОДГОТОВКА ДАННЫХ ЗДРАВООХРАНЕНИЯ ДЛЯ АНАЛИЗА

2024 г.

1. Область применения рабочей программы

Рабочая программа **модуля «Сбор и подготовка данных здравоохранения для анализа»** (далее – рабочая программа) является частью дополнительной профессиональной программы профессиональной переподготовки **ИТ «Прикладной анализ данных в здравоохранении»** (далее – программа) и направлена на формирование цифровых компетенций:

- «Применяет языки программирования для решения профессиональных задач» (ID 28 – базовый уровень):

Формулировка целевого уровня формирования компетенций	Промежуточные образовательные результаты		
	Опыт практической деятельности (ОПД)	Умения (У)	Знания (З)
Применяет языки программирования для решения профессиональных задач под контролем более опытных специалистов	У 1. Умеет выбирать типовые алгоритмы разработки	З 1. Знает принципы и основы программирования.	ОПД 1. Применяет язык программирования Python под контролем опытных наставников

- «Осуществляет сбор и подготовку данных для обучения моделей искусственного интеллекта» (ID 170 – базовый уровень):

Формулировка целевого уровня формирования компетенций	Промежуточные образовательные результаты		
	Опыт практической деятельности (ОПД)	Умения (У)	Знания (З)
Осуществляет критический отбор данных, проверяя их на целостность и непротиворечивость. Использует методы поиска данных и достоверные источники данных	ОПД 1. Применяет Python на практике для решения задач в здравоохранении.	У 1. Умеет собирать и подготавливать данные, для обучения и оценки моделей машинного обучения.	З 1. Знает основные концепции и методы искусственного интеллекта (ИИ) и машинного обучения (МО).

Освоение рабочей программы является обязательным для всех обучающихся по Программе.

2. Структура и краткое содержание рабочей программы

№ п/п	Наименование тем, виды учебной работы и краткое содержание учебного материала	Объем, часов
1.	**Тема 1 Источники данных в здравоохранении: электронные медицинские записи, биобанки, клинические исследования** 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.	16

№ п/п	Наименование тем, виды учебной работы и краткое содержание учебного материала	Объем, часов
	Лекция 2. Структурированные электронные медицинские документы: виды, принципы построения. Справочники НСИ ЕГИСЗ. ЕГИСЗ (цифровой контур здравоохранения) – инструмент генерации, передачи, обработки и хранения данных. Лекция 3. Регистры и реестры в здравоохранении и медицине, базы данных здравоохранения. История разработки регистров. Виды, классификации медицинских регистров. Семинар: Вопросы и ответы по основным понятиям. Обсуждение примеров использования анализа данных в здравоохранении. Практическое занятие 1: Разбор реальных примеров успешного применения анализа данных в медицинской практике и здравоохранении. Практическое занятие 2: Работа в региональном сегменте ЕГИСЗ на примере медицинской информационной системы ГИС «ЭЗ РТ». Формирование СЭМД. Самостоятельная работа: изучение открытых источников данных здравоохранения	
2	**Тема 2. Этические аспекты и конфиденциальность данных** Лекция: Введение в этику работы с данными. Основные этические принципы в здравоохранении (принцип уважения, благодеяния, справедливости). Исторические примеры и важность этических норм (Таскигийское исследование, случаи несанкционированного использования данных) Семинар: Обсуждение реальных примеров этических дилемм в работе с медико-биологическими данными. Обеспечение этичности и конфиденциальности данных в своей работе.	4
3	**Тема 3. Конфиденциальность данных пациентов.** Лекция 1. Понятие конфиденциальности данных в здравоохранении. Законы и нормативные акты, регулирующие конфиденциальность данных в здравоохранении. 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. Практическое занятие 1: Исследование и сравнение требований различных нормативных актов в области конфиденциальности данных. Практическое занятие 2: Разработка шаблона согласия на использование данных, включающего все необходимые компоненты в медицинской информациорнной системе (на примере ГИС «ЭЗ РТ»).	8

№ п/п	Наименование тем, виды учебной работы и краткое содержание учебного материала	Объем, часов
4	**Тема 4. Анонимизация и псевдонимизация данных.** Лекция . Методы анонимизации данных (удаление идентифицирующей информации, генерация случайных идентификаторов). Псевдонимизация данных как способ защиты конфиденциальности. Практическое занятие: Анонимизация набора данных пациентов с использованием Python. Самостоятельная работа: принципы и инструменты обезличивания данных в здравоохранении.	4
5	**Тема 5. Управление доступом к данным** Лекция 1. Принципы управления доступом к данным. Роли и права доступа в системах управления данными, медицинских информационных системах. Парольная политика. Лекция 2. Разбор реальных кейсов утечек данных и ошибок в обработке медико-биологических данных. Уроки, извлеченные из этих кейсов. Семинар: Обсуждение мер предосторожности и профилактики утечек данных в здравоохранении. Практическое занятие: Настройка ролей и прав доступа в симулированной системе управления данными, медицинских информационных системах, на примере «ГИС ЭЗ РТ». Самостоятельная работа: нормативно-правовое регулирование вопросов защиты данных в здравоохранении	12
6	Промежуточная аттестация в формате тестирования	2

3. Учебно-тематический план рабочей программы

№ п/п	Наименование и краткое содержание структурного элемента (раздела) Программы	Количество часов		
		аудиторных		самостоятель-ной работы
		лекции, семинары	практические занятия	
1.	Тема 1 Источники данных в здравоохранении: электронные медицинские записи, биобанки, клинические исследования	8	4	4
2	Тема 2. Этические аспекты и конфиденциальность данных.	4	-	-
3	Тема 3. Конфиденциальность данных пациентов.	4	4	-
4	Тема 4. Анонимизация и псевдонимизация данных.	2	2	4
5	Тема 5. Управление доступом к данным	6	2	4
	Промежуточная аттестация	2		
	Итого	**50**		

4. Контроль и оценка результатов освоения рабочей программы

Образовательная организация высшего образования, реализующая рабочую программу, обеспечивает организацию и проведение текущий, промежуточный контроли демонстрируемых обучающимися образовательных результатов.

Текущий контроль проводится преподавателем на основе оценивания результатов практических работ и самостоятельной работы обучающихся. Промежуточный контроль проводится в форме тестирования. Формы и методы текущего и промежуточного контроля, критерии оценивания доводятся до сведения обучающихся в начале обучения.

Для текущего и промежуточного контроля создаются фонды оценочных средств (ФОС). ФОС включают в себя педагогические контрольно-измерительные материалы, предназначенные для определения соответствия (или несоответствия) индивидуальных образовательных достижений запланированным образовательным результатам.

4.1. Примеры оценочных средств.

Тестовые задания:

1. Какие типы данных наиболее часто используются в здравоохранении и каковы их основные источники?

- А) Клинические данные, электронные медицинские записи (ЭМЗ)
- Б) Финансовые данные, страховые отчеты
- В) Административные данные, отчеты о пациентах
- Г) Все вышеперечисленные

2. Какие шаги включены в процесс очистки данных здравоохранения?

- А) Удаление дубликатов, корректировка отсутствующих значений
- Б) Изменение формата данных, нормализация данных
- В) Кодирование категориальных данных, проверка на аномалии
- Г) Все вышеперечисленные

3. Какие методы используются для обработки отсутствующих данных в медицинских наборах данных?

- А) Удаление строк с отсутствующими значениями
- Б) Импутация средним значением или медианой
- В) Использование алгоритмов машинного обучения для предсказания

4. Что из перечисленного является примером структурированного электронного медицинского документа?

- А) Медицинская карта пациента.
- Б) Результаты лабораторных исследований.
- В) Рецепт на лекарственный препарат.
- Г) Всё вышеперечисленное.

5. Что из перечисленного является методом анонимизации данных?

- А) Удаление идентифицирующей информации.

- Б) Генерация случайных идентификаторов.
- В) Псевдонимизация данных.
- Г) Всё вышеперечисленное.

5. Образцы учебно-методических материалов для обучающихся и преподавателей

Примеры заданий на самостоятельную и практическую работу

5.1. Вопросы для собеседования

1. Какие этические и правовые аспекты необходимо учитывать при анализе данных в здравоохранении?

Описание: Этот вопрос проверяет знание этических и правовых аспектов работы с данными в здравоохранении.

Ответ: При анализе данных в здравоохранении необходимо учитывать конфиденциальность и защиту персональных данных пациентов, соблюдение нормативных актов и законов о защите данных, а также этические вопросы, такие как обеспечение информированного согласия и недопущение дискриминации.

2. Какие этические дилеммы могут возникнуть при работе с медико-биологическими данными?

Описание: Этот вопрос проверяет знание этических и правовых аспектов работы с данными в здравоохранении.

Ответ: Дилемма конфиденциальности и открытости данных. Этичность использования данных в коммерческих целях. Вопросы информированного согласия пациентов.

3. Что такое конфиденциальность данных в здравоохранении?

Описание: Этот вопрос проверяет понимание сущности конфиденциальности данных.

Ответ: Конфиденциальность данных в здравоохранении — это защита личной информации пациентов от несанкционированного доступа, использования или раскрытия. Это включает в себя медицинские записи, результаты анализов, историю болезни и другую личную информацию.

4. Какие законы и нормативные акты регулируют конфиденциальность данных в здравоохранении в России?

Описание: этот вопрос проверяет знание законов и нормативных актов, регулирующих защиту данных в здравоохранении.

Ответ: В России конфиденциальность данных в сфере здравоохранения регулируется Федеральным законом № 152-ФЗ «О персональных данных», а также другими нормативными актами, такими как Федеральный закон № 323-ФЗ «Об основах охраны здоровья граждан в Российской Федерации».

5. Что такое парольная политика и зачем она нужна? требования предъявляются к паролям.

Описание: этот вопрос проверяет знание сущности и принципов парольной политики, вопросов разграничения доступа к данным.

Ответ: Парольная политика — это набор правил и рекомендаций по использованию

паролей в информационной системе. Она нужна для обеспечения безопасности данных и защиты от несанкционированного доступа. Основные требования к паролям: сложность, уникальность, срок действия, частота смены.

5.2. Решение кейсов

Анонимизация и псевдонимизация данных здравоохранения для анализа.

Цель:

Научиться собирать, очищать и подготавливать данные здравоохранения для последующего анализа. Задание позволяет определить степень развития компетенций после освоения модулей 1 и 2 программы

Описание задания:

Вам предоставлен набор данных, содержащий информацию о пациентах, посещениях медицинских учреждений и проведенных процедурах. Ваша задача - провести предварительную обработку данных, очистку, проверку на корректность и подготовку к дальнейшему анализу.

Шаги выполнения:

1. Импорт данных:

- Загрузите предоставленный CSV-файл с данными.

- Изучите структуру данных, определите типы данных в каждом столбце (числовые, категориальные, текстовые и т.д.).

2. Очистка данных:

- Проверьте данные на наличие пропущенных значений.

- Обработайте пропущенные значения:

- Если пропущенные значения составляют менее 5% данных, заполните их средними или медианными значениями для числовых данных, а наиболее частыми значениями для категориальных данных.

- Если пропущенные значения составляют более 5% данных, решите, стоит ли удалить соответствующие строки или столбцы.

3. Обработка дубликатов:

- Найдите и удалите дубликаты строк в наборе данных.

4. Нормализация данных:

- Приведите данные к единому формату, если это необходимо (например, даты, единицы измерения).

- Нормализуйте числовые данные, если это требуется для дальнейшего анализа (например, стандартное масштабирование).

5. Кодирование категориальных данных:

- Преобразуйте категориальные данные в числовой формат с помощью метода One-Hot Encoding или Label Encoding.

6. Выявление и обработка выбросов:

- Используйте методы статистического анализа (например, метод межквартильного размаха) для выявления выбросов.

- Решите, как обработать найденные выбросы: удаление, замена, исправление.

7. Анализ и визуализация данных:

- Постройте несколько графиков (гистограммы, диаграммы рассеяния, ящики с усами и т.д.) для визуализации основных характеристик данных.

- Проанализируйте взаимосвязи между различными переменными.

8. Документирование процесса:

- Подготовьте отчет, описывающий каждый шаг обработки данных, включая методы и инструменты, использованные на каждом этапе.

- Включите основные выводы и визуализации в отчет.

Требования к сдаче:

- Jupyter Notebook или Python-скрипт с кодом выполнения всех шагов.

- Отчет в формате PDF, описывающий процесс и выводы.

Рекомендуемые инструменты:

- Язык программирования Python

- Библиотеки: pandas, numpy, matplotlib, seaborn, scikit-learn.

6. Кадровое обеспечение образовательного процесса

Галиахметов А.И. – ассистент кафедры общественного здоровья и организации здравоохранения ФГБОУ ВО Казанский государственный медицинский университет» МЗ РФ.

Искандаров И.Р. – к.м.н., доцент кафедры общественного здоровья и организации здравоохранения ФГБОУ ВО Казанский государственный медицинский университет» МЗ РФ.

Альмухаметов А.А. – руководитель проектов ЗАО «Витакор»

Марапов Д.И. – к.м.н., генеральный директор ООО «Статтех»

Васильева А.Г. – аналитик ЗАО «Витакор»

7. Материально-техническое обеспечение реализации рабочей программы

Необходимый для реализации программы модуля перечень материально-технического обеспечения включает в себя специально оборудованные помещения для проведения учебных занятий, в том числе:

аудитории, оборудованные мультимедийными и иными средствами обучения, позволяющими использовать информационные технологии, в количестве, позволяющем обучающимся осваивать умения и навыки, предусмотренные профессиональной деятельностью, индивидуально;

дисплейный класс, с мультимедийным и компьютерным оборудованием.

помещения для самостоятельной работы обучающихся оснащены компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду Университета.

Перечень информационных технологий, необходимых для освоения программы, включая перечень программного обеспечения и информационных справочных систем

Для достижения целей педагогического образования применяются следующие информационные технологии:

1.Образовательный портал дистанционного обучения Казанского ГМУ. Дистанционный курс в составе образовательного портала создан в системе MOODLE и

содержит в себе лекции, презентации, задания, гиперссылки на первоисточники учебного материала, тесты / задания для самоконтроля, контрольные и итоговые тесты по курсу.

2. Пакет прикладных программ МойОфис.

3. Демо-версия ГИС ЭЗ РТ.

Используемое программное обеспечение имеет лицензию и ежегодно и / или своевременно обновляется.

8. Информационное обеспечение реализации рабочей программы

8.1. Основная литература.

Статистика в биомедицине, фармации и фармацевтике : учебное пособие / Н. Н. Зубов, В. И. Кувакин, С. З. Умаров [и др.] ; под общей редакцией И. А. Наркевича ; Санкт-Петербургский государственный химико-фармацевтический университет. - Москва : КноРус, 2021. – 298.

8.2. Дополнительная литература.

Федеральный закон "О персональных данных" от 27.07.2006 N 152-ФЗ. Доступно по: https://www.consultant.ru/document/cons_doc_LAW_61801/ Ссылка активна на 11.07.2024.

Концепция создания единой государственной информационной системы в сфере здравоохранения. Приказ Минздравсоцразвития России от 28.04.2011 г. № 364. Доступно по: http://portal.egisz.rosminzdrav.ru/materials/99. Ссылка активна на 11.07.2024.

Программирование на Python в примерах и задачах / Алексей Васильев. - Москва : Эксмо, 2021. - 616 с.

Столбов А.П., Кузнецов П.П. Автоматизированная обработка и защита персональных данных в медицинских учреждениях. - М.: ИД «Менеджер здравоохранения», 2010. - 176 с.

8.3. Интернет-ресурсы.

Единая образовательная платформа Университета Иннополис https://oneid.unionepro.ru
Официальный сайт ФГБУ «Центральный научно-исследовательский институт организации и информатизации здравоохранения» МЗ РФ https://mednet.ru/
Сайт Федеральной электронной медицинской библиотеки − http://www.femb.ru
Сайт ФГБОУ ВО Казанского ГМУ Минздрава России – https://kazangmu.ru/
Сайт Научной электронной библиотеки − http://elibrary.ru/
Паспорт национального проекта «Здравоохранение» http://government.ru/info/35561/
Официальный сайт ФГБУ «Центральный научно-исследовательский институт организации и информатизации здравоохранения» МЗ РФ https://mednet.ru/
Портал информационной поддержки пользователей ГИС ЭЗ РТ https://wiki.vitacore.ru/display/rmis
Портал нормативно-справочной информации ЕГИСЗ https://nsi.rosminzdrav.ru/

Министерство здравоохранения Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования «Казанский государственный медицинский университет»

«УТВЕРЖДАЮ»
Ректор ФГБОУ ВО Казанский ГМУ
Минздрава России, профессор
_____ А.С. Созинов

«___» _____ 2024 г.

Рабочая программа модуля 3

АНАЛИЗ БОЛЬШИХ ДАННЫХ В ЗДРАВООХРАНЕНИИ

2024 г.

1. Область применения рабочей программы

Рабочая программа **модуля «Анализ больших данных в здравоохранении»** (далее – рабочая программа) является частью дополнительной профессиональной программы профессиональной переподготовки **ИТ «Прикладной анализ данных в здравоохранении»** (далее – программа) и направлена на формирование цифровых компетенций:

- «Применяет языки программирования для решения профессиональных задач» (ID 28 - базовый уровень):

Формулировка целевого уровня формирования компетенций	Промежуточные образовательные результаты		
	Опыт практической деятельности (ОПД)	Умения (У)	Знания (З)
Применяет языки программирования для решения профессиональных задач под контролем более опытных специалистов	У 1. Умеет выбирать типовые алгоритмы разработки	З 1. Знает принципы и основы программирования.	ОПД 1. Применяет язык программирования Python под контролем опытных наставников

- «Осуществляет сбор и подготовку данных для обучения моделей искусственного интеллекта» (ID 170 - базовый уровень):

Формулировка целевого уровня формирования компетенций	Промежуточные образовательные результаты		
	Опыт практической деятельности (ОПД)	Умения (У)	Знания (З)
Осуществляет критический отбор данных, проверяя их на целостность и непротиворечивость. Использует методы поиска данных и достоверные источники данных	ОПД 1. Применяет Python на практике для решения задач в здравоохранении.	У 1. Умеет собирать и подготавливать данные, для обучения и оценки моделей машинного обучения.	З 1. Знает основные концепции и методы искусственного интеллекта (ИИ) и машинного обучения (МО).

Освоение рабочей программы является обязательным для всех обучающихся по Программе.

2. Структура и краткое содержание рабочей программы

№ п/п	Наименование тем, виды учебной работы и краткое содержание учебного материала	Объем, часов
1.	**Тема 1 Введение в большие данные в здравоохранении.** 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. Лекция 2. Роль больших данных в здравоохранении. Как большие данные трансформируют здравоохранение. Примеры применения: персонализированная медицина, мониторинг состояния здоровья, прогнозирование заболеваний.	8

№ п/п	Наименование тем, виды учебной работы и краткое содержание учебного материала	Объем, часов
	Семинар 1 Обсуждение особенностей больших данных и их примеров в здравоохранении. Практическое занятие 1: Анализ статьи или кейса, демонстрирующего использование больших данных в здравоохранении.	
2	**Тема 2. Методы и инструменты для работы с большими данными** 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. Лекция 2. Основы Apache Spark. Архитектура Apache Spark. Основные компоненты: Spark Core, Spark SQL, Spark Streaming, MLlib, GraphX. Семинар 1: Обсуждение преимуществ и ограничений различных методов анализа больших данных Практическое занятие 1: Установка и настройка Apache Spark. Краткий обзор интерфейса и функционала. Практическое занятие 2. Использование Apache Spark для обработки больших данных. Запуск Spark-сессии. Выполнение простых операций с данными (чтение, обработка, запись).	10
3	**Тема 3. Машинное обучение** Лекция 1. Основные алгоритмы машинного обучения. Применение машинного обучения для анализа данных в здравоохранении. Лекция 2. Анализ больших данных с использованием Hadoop. Архитектура Hadoop. Основные компоненты: HDFS, MapReduce, YARN. Практическое занятие 1: Создание и обучение модели машинного обучения. Оценка модели и интерпретация результатов. Практическое занятие 2: Установка и настройка Hadoop. Запуск простых задач MapReduce.	10
4	**Тема 4. Реальные кейсы применения анализа больших данных в здравоохранении** Лекция. Примеры успешного применения анализа больших данных в здравоохранении. Кейсы и их результаты. Семинар: Обсуждение проблем и решений при анализе больших данных. Самостоятельная работа: изучение примеров успешного применения анализа данных в управлении здравоохранением.	8
5	**Тема 5. Разработка комплексного проект**	12

№ п/п	Наименование тем, виды учебной работы и краткое содержание учебного материала	Объем, часов
	Лекция. Постановка задачи для проекта (например, анализ данных пациентов для прогнозирования вероятности возникновения событий: заболеваний, операций, смерти и т.д.). 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. Практическое занятие 2. Презентация результатов проектов. Представление результатов анализа данных. Обсуждение выводов и рекомендаций. Самостоятельная работа: подготовка проекта.	
6	Промежуточная аттестация в формате тестирования	2

3. Учебно-тематический план рабочей программы

№ п/п	Наименование и краткое содержание структурного элемента (раздела) Программы	Количество часов		
		аудиторных		самостоятель-ной работы
		лекции, семинары	практические занятия	
1.	Тема 1 Введение в большие данные в здравоохранении.	6	2	-
2	Тема 2. Методы и инструменты для работы с большими данными	6	4	
3	Тема 3. Машинное обучение	4	4	
4	Тема 4. Реальные кейсы применения анализа больших данных в здравоохранении.	4		4
5	Тема 5. Разработка комплексного проект	2	4	6
	Промежуточная аттестация	2		
	Итого	**50**		

4. Контроль и оценка результатов освоения рабочей программы

Образовательная организация высшего образования, реализующая рабочую программу, обеспечивает организацию и проведение текущий, промежуточный контроли демонстрируемых обучающимися образовательных результатов.

Текущий контроль проводится преподавателем на основе оценивания результатов практических работ и самостоятельной работы обучающихся. Промежуточный контроль проводится в форме тестирования. Формы и методы текущего и промежуточного контроля, критерии оценивания доводятся до сведения обучающихся в начале обучения.

Для текущего и промежуточного контроля создаются фонды оценочных средств (ФОС). ФОС включают в себя педагогические контрольно-измерительные материалы, предназначенные для определения соответствия (или несоответствия) индивидуальных образовательных достижений запланированным образовательным результатам.

4.1. Примеры оценочных средств.

Тестовые задания:

1. Каковы основные источники больших данных в здравоохранении и какие типы данных они содержат?

- A) Электронные медицинские карты (ЭМК), данные геномики, данные носимых устройств
- B) Данные социальных сетей, данные биржевых торгов, данные о погоде
- C) Данные телевизионных просмотров, данные банковских операций, данные о дорожном трафике
- D) Данные интернет-поисковиков, данные о посещаемости музеев, данные о продаже билетов

2. Какой из следующих методов машинного обучения наиболее часто используется для предсказания заболеваний у пациентов на основе анализа больших данных?

- A) Кластеризация (Clustering)
- B) Линейная регрессия (Linear Regression)
- C) Решающие деревья (Decision Trees)
- D) Понижение размерности (Dimensionality Reduction)

3. Какие основные преимущества применения больших данных в здравоохранении?

- A) Улучшение персонализированного лечения, снижение затрат, предсказание вспышек заболеваний
- B) Увеличение времени ожидания, повышение сложности систем, увеличение числа ошибок
- C) Повышение производительности труда, улучшение качества продукта, снижение налогов
- D) Снижение количества медицинских кадров, увеличение бумажной работы, увеличение расходов на обучение

4. Какие меры безопасности необходимо предпринимать при работе с большими данными в здравоохранении для защиты конфиденциальности пациентов?

- A) Шифрование данных, управление доступом, анонимизация данных
- B) Открытый доступ к данным, отсутствие аутентификации, совместное использование паролей
- C) Хранение данных на личных устройствах сотрудников, отключение антивирусных программ, использование общих учетных записей
- D) Публикация данных в открытых источниках, упрощение процедур безопасности, использование несертифицированного ПО

5. Как анализ больших данных может помочь в управлении здравоохранением на государственном уровне?

- A) Путем предоставления точных прогнозов эпидемий, оптимизации ресурсов и

улучшения планирования здравоохранения

- B) Путем увеличения бюрократических процедур, повышения сложности управления, снижения качества обслуживания

- C) Путем закрытия малых медицинских учреждений, увеличения зависимости от внешних консультантов, сокращения финансирования здравоохранения

- D) Путем уменьшения числа доступных данных, усложнения доступа к медицинской помощи, увеличения времени обработки запросов

5. Образцы учебно-методических материалов для обучающихся и преподавателей

Примеры заданий на самостоятельную и практическую работу.

Решение кейсов:

Анализ больших данных для улучшения медицинских услуг

Цель задания:

Научиться применять методы анализа больших данных для решения реальных задач в области здравоохранения.

Описание:

Вам предстоит провести анализ набора данных, содержащих информацию о пациентах и их медицинских историях, чтобы выявить ключевые факторы, влияющие на развитие хронических заболеваний. Результаты анализа помогут в разработке стратегии для улучшения профилактики и лечения этих заболеваний.

Шаги выполнения задания:

1. Подготовка данных:

- Получите набор данных, содержащий информацию о пациентах (например, демографические данные, результаты медицинских тестов, истории болезней, образ жизни).

- Оцените качество данных: проверьте на наличие пропусков, выбросов и некорректных значений. При необходимости очистите и предобработайте данные.

2. Анализ данных:

- Проведите исследовательский анализ данных (Exploratory Data Analysis, EDA), чтобы выявить основные тенденции и закономерности.

- Визуализируйте данные с помощью графиков и диаграмм (гистограммы, коробчатые диаграммы, тепловые карты и т.д.).

3. Моделирование:

- Выберите целевую переменную (например, наличие или отсутствие хронического заболевания).

- Разделите данные на обучающую и тестовую выборки.

- Постройте несколько моделей машинного обучения для предсказания целевой переменной (например, логистическая регрессия, решающие деревья, случайный лес).

- Оцените качество моделей с помощью метрик (точность, полнота, F-мера, ROC-AUC).

4. Интерпретация результатов:

- Определите наиболее значимые переменные, влияющие на развитие хронических заболеваний.

- Проанализируйте результаты моделей и сделайте выводы о факторах риска.

5. Рекомендации:

- На основе полученных результатов разработайте рекомендации для улучшения профилактики и лечения хронических заболеваний.

- Подготовьте отчет с выводами и предложениями, подкрепленный визуализациями и результатами анализа.

Требования к выполнению задания:

- Использование языков программирования Python

- Применение библиотек для анализа данных и машинного обучения (например, pandas, numpy, scikit-learn, matplotlib, seaborn для Python).

- Описание всех этапов работы, включая подготовку данных, анализ, моделирование и интерпретацию результатов.

- Визуализация ключевых находок и выводов.

Результат:

- Код выполнения задания.

- Отчет в формате PDF, содержащий описание этапов работы, визуализации и выводы.

- Презентация (Мой офис или аналог), в которой кратко представлены основные результаты и рекомендации.

6. Кадровое обеспечение образовательного процесса

Галиахметов А. И., ассистент кафедры общественного здоровья и организации здравоохранения ФГБОУ ВО Казанский государственный медицинский университет» МЗ РФ.

Искандаров И. Р., к.м.н., доцент кафедры общественного здоровья и организации здравоохранения ФГБОУ ВО Казанский государственный медицинский университет» МЗ РФ.

Альмухаметов А. А., руководитель проектов ЗАО «Витакор»

Марапов Д. И., к.м.н., генеральный директор ООО «Статтех»

Васильева А. Г., аналитик ЗАО «Витакор»

7. Материально-техническое обеспечение реализации рабочей программы

Необходимый для реализации программы модуля перечень материально-технического обеспечения включает в себя специально оборудованные помещения для проведения учебных занятий, в том числе:

аудитории, оборудованные мультимедийными и иными средствами обучения, позволяющими использовать информационные технологии, в количестве, позволяющем обучающимся осваивать умения и навыки, предусмотренные профессиональной деятельностью, индивидуально;

дисплейный класс, с мультимедийным и компьютерным оборудованием.

помещения для самостоятельной работы обучающихся оснащены компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду Университета.

Перечень информационных технологий, необходимых для освоения программы, включая перечень программного обеспечения и информационных справочных систем

Для достижения целей педагогического образования применяются следующие информационные технологии:

1.Образовательный портал дистанционного обучения Казанского ГМУ. Дистанционный курс в составе образовательного портала создан в системе MOODLE и содержит в себе лекции, презентации, задания, гиперссылки на первоисточники учебного материала, тесты / задания для самоконтроля, контрольные и итоговые тесты по курсу.

2.Пакет прикладных программ МойОфис.

Используемое программное обеспечение имеет лицензию и ежегодно и / или своевременно обновляется.

8. Информационное обеспечение реализации рабочей программы

8.1. Основная литература.

Статистика в биомедицине, фармации и фармацевтике : учебное пособие / Н. Н. Зубов, В. И. Кувакин, С. З. Умаров [и др.] ; под общей редакцией И. А. Наркевича ; Санкт-Петербургский государственный химико-фармацевтический университет. - Москва : КноРус, 2021. – 298.

8.2. Дополнительная литература.

Силен Д., Мейсман А., Али М. «Основы Data Science и Big Data. Python и наука о данных». — СПб.: Питер, 2023.

Программирование на Python в примерах и задачах / Алексей Васильев. - Москва : Эксмо, 2021. - 616 с.

Иванова А. А. Применение Big Data в сфере здравоохранения: российский и зарубежный опыт. Научные записки молодых исследователей. 2020;8(5):42-53.

8.3. Интернет-ресурсы.

Единая образовательная платформа Университета Иннополис https://oneid.unionepro.ru
Официальный сайт ФГБУ «Центральный научно-исследовательский институт организации и информатизации здравоохранения» МЗ РФ https://mednet.ru/
Сайт Федеральной электронной медицинской библиотеки − http://www.femb.ru
Сайт ФГБОУ ВО Казанского ГМУ Минздрава России – https://kazangmu.ru/
Сайт Научной электронной библиотеки − http://elibrary.ru/

Министерство здравоохранения Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования «Казанский государственный медицинский университет»

«УТВЕРЖДАЮ»
Ректор ФГБОУ ВО Казанский ГМУ
Минздрава России, профессор
_____ А.С. Созинов

«__» _____ 2024 г.

Рабочая программа модуля 4

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И МАШИННОЕ ОБУЧЕНИЕ В ЗДРАВООХРАНЕНИИ

2024 г.

1. Область применения рабочей программы

Рабочая программа **модуля «Искусственный интеллект и машинное обучение в здравоохранении»** (далее – рабочая программа) является частью дополнительной профессиональной программы профессиональной переподготовки **ИТ «Прикладной анализ данных в здравоохранении»** (далее – программа) и направлена на формирование цифровых компетенций:

- «Применяет языки программирования для решения профессиональных задач» (ID 28 - базовый уровень):

Формулировка целевого уровня формирования компетенций	Промежуточные образовательные результаты		
	Опыт практической деятельности (ОПД)	Умения (У)	Знания (З)
Применяет языки программирования для решения профессиональных задач под контролем более опытных специалистов	У 1. Умеет выбирать типовые алгоритмы разработки	З 1. Знает принципы и основы программирования.	ОПД 1. Применяет язык программирования Python под контролем опытных наставников

- «Осуществляет сбор и подготовку данных для обучения моделей искусственного интеллекта» (ID 170 - базовый уровень):

Формулировка целевого уровня формирования компетенций	Промежуточные образовательные результаты		
	Опыт практической деятельности (ОПД)	Умения (У)	Знания (З)
Осуществляет критический отбор данных, проверяя их на целостность и непротиворечивость. Использует методы поиска данных и достоверные источники данных	ОПД 1. Применяет Python на практике для решения задач в здравоохранении.	У 1. Умеет собирать и подготавливать данные, для обучения и оценки моделей машинного обучения.	З 1. Знает основные концепции и методы искусственного интеллекта (ИИ) и машинного обучения (МО).

-

Освоение рабочей программы является обязательным для всех обучающихся по Программе.

2. Структура и краткое содержание рабочей программы

№ п/п	Наименование тем, виды учебной работы и краткое содержание учебного материала	Объем, часов
1.	**Тема 1 Основные понятия искусственного интеллекта и машинного обучения** 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: сбор данных, подготовка данных, выбор и обучение модели, оценка модели, внедрение. Лекция 2: Применение ИИ и МО в здравоохранении. Примеры использования ИИ и МО: диагностика заболеваний, прогнозирование, персонализированная медицина.	12

№ п/п	Наименование тем, виды учебной работы и краткое содержание учебного материала	Объем, часов
	Семинар: Восприятие. Когнитивные процессы. Принятие решений. Как это работает в природе и как это можно описать? Философия вопроса. (Хебб и др.) Практическое занятие: Разбор статьи или кейса, демонстрирующего применение ИИ в здравоохранении. Самостоятельная работа: изучение примеров применения ИИ, СППР в здравоохранении и клинической практике	
2	**Тема 2. Сбор и подготовка данных для обучения моделей ИИ** Лекция1. 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. Практическое занятие 1: Сбор данных из открытых источников (например, набор данных UCI Machine Learning Repository). Практическое занятие 2: Очистка и подготовка данных с использованием Python. Обработка пропущенных значений и кодирование категориальных переменных. Самостоятельная работа: изучение открытых источников данных для обучения моделей ИИ	10
3	**Тема 3. Выбор и обучение моделей машинного обучения** 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). Лекция 2. Обучение моделей. Процесс обучения модели: разделение данных на тренировочные и тестовые, кросс-валидация, гиперпараметры. Практическое занятие 1: Применение алгоритмов машинного обучения с использованием Python. Практическое занятие 2: Обучение моделей на примере набора данных здравоохранения. Кросс-валидация и настройка гиперпараметров.	8
4	**Тема 4. Оценка моделей машинного обучения.** 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).	8

№ п/п	Наименование тем, виды учебной работы и краткое содержание учебного материала	Объем, часов
	Лекция 2. Интерпретация результатов оценки моделей. Примеры ошибок и способы их устранения Практическое занятие 1: Оценка модели классификации и регрессии на примере набора данных здравоохранения. Практическое занятие 2. Интерпретация результатов оценки модели и рекомендации по улучшению.	
5	**Тема 5. Применение нейронных сетей для анализа медико-биологических данных и данных здравоохранения.** Лекция 1. Структура и принципы работы нейронных сетей. Основные типы нейронных сетей: полносвязные сети, свёрточные нейронные сети (CNN), рекуррентные нейронные сети (RNN). Лекция 2. Применение нейронных сетей в здравоохранении. Примеры использования нейронных сетей для анализа изображений (рентген, МРТ), последовательностей данных (анализ ЭКГ, геномных данных) Практическое занятие 1: Создание простой нейронной сети с использованием Python. Практическое занятие 2: Создание простой нейронной сети с использованием Python (продолжение) Практическое занятие 3. Обучение и оценка нейронной сети на примере данных здравоохранения (например, классификация изображений рентгеновских снимков).	10
6	Промежуточная аттестация в формате тестирования	2

3. Учебно-тематический план рабочей программы

№ п/п	Наименование и краткое содержание структурного элемента (раздела) Программы	Количество часов		
		аудиторных		самостоятель-ной работы
		лекции, семинары	практические занятия	
1.	Тема 1 Основные понятия искусственного интеллекта и машинного обучения	6	2	4
2	Тема 2. Сбор и подготовка данных для обучения моделей ИИ	2	4	2
3	Тема 3. Выбор и обучение моделей машинного обучения	2	4	4
4	Тема 4. Оценка моделей машинного обучения.	4	4	-
5	Тема 5. Применение нейронных сетей для анализа	4	6	-

	медико-биологических данных.			
6	Промежуточная аттестация	2		
	Итого	**50**		

4. Контроль и оценка результатов освоения рабочей программы

Образовательная организация высшего образования, реализующая рабочую программу, обеспечивает организацию и проведение текущий, промежуточный контроли демонстрируемых обучающимися образовательных результатов.

Текущий контроль проводится преподавателем на основе оценивания результатов практических работ и самостоятельной работы обучающихся. Промежуточный контроль проводится в форме тестирования. Формы и методы текущего и промежуточного контроля, критерии оценивания доводятся до сведения обучающихся в начале обучения.

Для текущего и промежуточного контроля создаются фонды оценочных средств (ФОС). ФОС включают в себя педагогические контрольно-измерительные материалы, предназначенные для определения соответствия (или несоответствия) индивидуальных образовательных достижений запланированным образовательным результатам.

4.1. Примеры оценочных средств.

Тестовые задания:

1. Каковы основные преимущества использования искусственного интеллекта в диагностике заболеваний?

- A. Повышенная точность диагностики
- B. Уменьшение времени на постановку диагноза
- C. Автоматизация рутинных задач
- D. Все вышеперечисленное

2. Какой метод машинного обучения чаще всего используется для анализа медицинских изображений (например, рентгеновских снимков или МРТ)?

- A. Регрессия
- B. Кластеризация
- C. Глубокое обучение
- D. Байесовские сети

3. Какие данные могут быть использованы для обучения моделей машинного обучения в здравоохранении?

- A. Электронные медицинские записи
- B. Генетическая информация
- C. Данные с носимых устройств
- D. Все вышеперечисленное

4. Каковы основные вызовы при внедрении искусственного интеллекта в здравоохранении?

- A. Проблемы с конфиденциальностью данных
- B. Нехватка обученных специалистов
- C. Высокая стоимость разработки и внедрения
- D. Все вышеперечисленное

5. Какую роль играет естественный язык в применении искусственного интеллекта в здравоохранении?
- A. Помогает в обработке и анализе медицинских записей
- B. Позволяет создавать интеллектуальных ассистентов для пациентов
- C. Обеспечивает перевод медицинских документов
- D. Все вышеперечисленное

5. Образцы учебно-методических материалов для обучающихся и преподавателей

Примеры заданий на самостоятельную и практическую работу

Решение кейса:

Задание: Разработка модели машинного обучения для прогнозирования риска развития диабета

Цель:

Разработать и обучить модель машинного обучения для прогнозирования риска развития диабета у пациентов на основе их медицинских данных.

Шаги выполнения:

1. Сбор и подготовка данных:

- Найдите и скачайте набор данных о пациентах, содержащий информацию о возрасте, поле, индексе массы тела (ИМТ), артериальном давлении, уровнях глюкозы, инсулина, и другие релевантные медицинские параметры.

- Примером такого набора данных может быть Pima Indians Diabetes Dataset, доступный в библиотеке sklearn.datasets или на платформе Kaggle.

2. Предварительная обработка данных:

- Очистите данные от пропущенных значений и аномалий.

- Нормализуйте или стандартизируйте числовые признаки.

- Закодируйте категориальные признаки (если есть) с использованием one-hot encoding или других методов.

3. Разделение данных:

- Разделите данные на тренировочную и тестовую выборки в соотношении 80:20.

4. Обучение модели:

- Выберите алгоритм машинного обучения для классификации (например, логистическую регрессию, случайный лес, градиентный бустинг или нейронную сеть).

- Обучите модель на тренировочной выборке.

5. Оценка модели:

- Оцените точность модели на тестовой выборке с использованием метрик, таких как точность (accuracy), полнота (recall), точность (precision) и F1-меру.

- Постройте ROC-кривую и вычислите площадь под кривой (AUC-ROC).

6. Интерпретация результатов:

- Определите, какие признаки имеют наибольшее влияние на предсказания модели.

- Обсудите возможные шаги по улучшению модели, такие как подбор гиперпараметров, использование более сложных моделей или добавление новых признаков.

7. Документирование:

- Подготовьте отчет с описанием выполнения каждого шага, результатов оценки модели и выводами.

- Включите графики, диаграммы и таблицы для визуализации результатов.

Дополнительно:

- Постройте и визуализируйте важность признаков (feature importance) для вашей модели.

- Попробуйте использовать методы уменьшения размерности, такие как PCA (Principal Component Analysis), и оцените их влияние на производительность модели.

Ожидаемые результаты:

- Разработанная и обученная модель, способная предсказывать риск развития диабета с удовлетворительной точностью.

- Подробный отчет, описывающий процесс разработки, результаты и выводы.

Программные инструменты:

- Язык программирования: Python

- Библиотеки: pandas, numpy, scikit-learn, matplotlib, seaborn, tensorflow, пр.

6. Кадровое обеспечение образовательного процесса

Просвиркин И.А., к.т.н., ИТ-директор ООО «ФтизисБиоМед»
Гогоберидзе Ю.Т., старший разработчик ООО «ФтизисБиоМед»
Сабитов Р.А., старший ML-инженер ООО «ФтизисБиоМед»
Фатихов А.Ф., ML-инженер ООО «ФтизисБиоМед»

7. Материально-техническое обеспечение реализации рабочей программы

Необходимый для реализации программы модуля перечень материально-технического обеспечения включает в себя специально оборудованные помещения для проведения учебных занятий, в том числе:

аудитории, оборудованные мультимедийными и иными средствами обучения, позволяющими использовать информационные технологии, в количестве, позволяющем обучающимся осваивать умения и навыки, предусмотренные профессиональной деятельностью, индивидуально;

дисплейный класс, с мультимедийным и компьютерным оборудованием.

помещения для самостоятельной работы обучающихся оснащены компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду Университета.

Перечень информационных технологий, необходимых для освоения программы, включая перечень программного обеспечения и информационных справочных систем

Для достижения целей педагогического образования применяются следующие информационные технологии:

1.Образовательный портал дистанционного обучения Казанского ГМУ. Дистанционный курс в составе образовательного портала создан в системе MOODLE и содержит в себе лекции, презентации, задания, гиперссылки на первоисточники учебного материала, тесты / задания для самоконтроля, контрольные и итоговые тесты по курсу.

2.Пакет прикладных программ МойОфис.

Используемое программное обеспечение имеет лицензию и ежегодно и / или своевременно обновляется.

8. Информационное обеспечение реализации рабочей программы

8.1. Основная литература.

Основы цифровой трансформации общества : учебник для вузов / Н. А. Горелов, О. Н. Кораблева. - 2-е изд., перераб. и доп. - Москва : Юрайт, 2024. – 336.

Цифровая трансформация здравоохранения и ведомственной медицины / Столяр В. П., Крайнюков П. Е., Калачёв О. В. - Москва : Планета, 2020. - 199

8.2. Дополнительная литература.

Гусев А. В. «Искусственный интеллект в медицине. Как умные технологии меняют подход к лечению». — М.: Эксмо, 2023.

Емельянов Д. О., Копаница Г. Д., Кумсков М. И. «Применение технологий искусственного интеллекта для диагностики и прогнозирования заболеваний». — Томск: Издательство Томского политехнического университета, 2022.

Камаев В. А., Щербаков Д. С., Федотова Г. В., Леванов В. М. «Интеллектуальные системы поддержки принятия решений в здравоохранении». — Нижний Новгород: Издательство Приволжского исследовательского медицинского университета, 2021.

Корнилов В. В., Шишкин А. Г. «Использование искусственного интеллекта при принятии управленческих решений в сфере здравоохранения». — СПб: Санкт-Петербургский государственный экономический университет, 2020.

Остроушко Н. И., Остроушко А. В., Сафронова К. В. «Современные аспекты применения искусственного интеллекта в медицинской диагностике». — Курск: Университетская книга, 2022.

Силен Д., Мейсман А., Али М. «Основы Data Science и Big Data. Python и наука о данных». — СПб.: Питер, 2023.

Программирование на Python в примерах и задачах / Алексей Васильев. - Москва : Эксмо, 2021. - 616 с.

8.3. Интернет-ресурсы.

Единая образовательная платформа Университета Иннополис https://oneid.unionepro.ru
Официальный сайт ФГБУ «Центральный научно-исследовательский институт организации и информатизации здравоохранения» МЗ РФ https://mednet.ru/
Сайт Федеральной электронной медицинской библиотеки − http://www.femb.ru
Сайт ФГБОУ ВО Казанского ГМУ Минздрава России – https://kazangmu.ru/
Сайт Научной электронной библиотеки − http://elibrary.ru/

Министерство здравоохранения Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования «Казанский государственный медицинский университет»

«УТВЕРЖДАЮ»
Ректор ФГБОУ ВО Казанский ГМУ
Минздрава России, профессор
_____ А.С. Созинов

«__» _____ 2024 г.

Рабочая программа практики

Дополнительная профессиональная программа
Прикладной анализ данных в здравоохранении

2024 г.

1. Область применения рабочей программы

Рабочая программа **практики дополнительной профессиональной программы «Прикладной анализ данных в здравоохранении»** (далее – рабочая программа) является частью дополнительной профессиональной программы профессиональной переподготовки **ИТ «Прикладной анализ данных в здравоохранении»** (далее – программа) и направлена на формирование цифровых компетенций:

ID и наименование компетенции	формулировка целевого уровня формирования компетенций	Промежуточные образовательные результаты		
		Опыт практической деятельности (ОПД)	Умения (У)	Знания (З)
11. Использует специализированное программное обеспечение для анализа медико-биологических данных (Продвинутый уровень)	Самостоятельно применяет оптимальные алгоритмы анализа медико-биологических данных, используя специализированное программное обеспечение. Выбирает и применяет оптимальные критерии и методы для описания данных и аналитики	ОПД 1 Применяет для анализа данных здравоохранения программное обеспечение Stattech ОПД 2 Применяет для анализа данных здравоохранения программное обеспечение SPSS Statistics	У 1 Умеет выбирать оптимальные критерии и методы для описания данных и аналитики в здравоохранении	З 1. Знает основные понятия и терминологию анализа данных в здравоохранении. З 2. Знает основные методы анализа данных в здравоохранении
28. Применяет языки программирования для решения профессиональных задач (Базовый уровень)	Применяет языки программирования для решения профессиональных задач	ОПД 1. Применяет язык программирования Python под контролем опытных наставников	У 1. Умеет выбирать типовые алгоритмы разработки	З 1. Знает принципы и основы программирования.
170. Осуществляет сбор и подготовку данных для обучения моделей искусственного интеллекта (Базовый уровень)	Осуществляет критический отбор данных, проверяя их на целостность и непротиворечивость. Использует методы поиска данных и достоверные источники данных	ОПД 1. Применяет Python на практике для решения задач в здравоохранении.	У 1. Умеет собирать и подготавливать данные, для обучения и оценки моделей машинного обучения.	З 1. Знает основные концепции и методы искусственного интеллекта (ИИ) и машинного обучения (МО).

Освоение рабочей программы является обязательным для всех обучающихся по Программе.

2. Структура и краткое содержание рабочей программы

№ п/п	Наименование тем, виды учебной работы и краткое содержание учебного материала	Объем, часов
1.	Выполнение проектной работы на базе индустриального партнера	50

3. Учебно-тематический план рабочей программы

№ п/п	Наименование и краткое содержание структурного элемента (раздела) Программы	Количество часов			
		аудиторных		самостоятельной работы	Практика / стажировка
		лекции, семинары	практические занятия		
1.	Выполнение проектной работы на базе индустриального партнера	-	-	-	50
6	Промежуточная аттестация	Не предусмотрена			
	Итого	**50**			

3. Контроль и оценка результатов освоения рабочей программы

Образовательная организация высшего образования, реализующая рабочую программу, совместно с индустриальном партнером обеспечивает организацию и проведение текущий, промежуточный контроли демонстрируемых обучающимися образовательных результатов.

Текущий контроль проводится преподавателем на основе оценивания результатов практических работ и самостоятельной работы обучающихся. Промежуточный контроль проводится в форме тестирования. Формы и методы текущего и промежуточного контроля, критерии оценивания доводятся до сведения обучающихся в начале обучения.

Для текущего и промежуточного контроля создаются фонды оценочных средств (ФОС). ФОС включают в себя педагогические контрольно-измерительные материалы, предназначенные для определения соответствия (или несоответствия) индивидуальных образовательных достижений запланированным образовательным результатам.

4. Примеры оценочных средств

Примеры тематик проектов для выполнения:

1. Формирование обучающей базы для разработки ИИ прогнозирования риска сосудистых катастроф.
2. Формирование обучающей базы для разработки ИИ прогнозирования онкологического заболевания
3. Разработка инструмента, анализа истории болезни пациента для прогнозирования будущих событий и их исходов.
4. Прогноз событий с помощью нейронного классификатора и рекурентной нейронной сети.

5. Образцы учебно-методических материалов для обучающихся и преподавателей

Выполнение проекта должно сопровождаться отчетом по следующей структуре: титульный лист;содержание работы; основная часть с описанием проделанной работы; заключение; список использованных источников; приложения.

В отчете необходимо осветить перечень поставленных и решенных задач, использованные инструменты и технологии.

Критерии оценивания: Соответствует требованиям задания; В ходе работы программы не возникает ошибок в процессе выполнения; Программа в результате своей работы возвращает корректный результат, требуемый по условиямзадания.

6. Кадровое обеспечение образовательного процесса

Галиахметов А. И., ассистент кафедры общественного здоровья и организации здравоохранения ФГБОУ ВО Казанский государственный медицинский университет» МЗ РФ.
Искандаров И. Р., к.м.н., доцент кафедры общественного здоровья и организации здравоохранения ФГБОУ ВО Казанский государственный медицинский университет» МЗ РФ.
Альмухаметов А. А., руководитель проектов ЗАО «Витакор»
Марапов Д. И., к.м.н., генеральный директор ООО «Статтех»
Васильева А. Г., аналитик ЗАО «Витакор»
Просвиркин И.А., к.т.н., ИТ-директор ООО «ФтизисБиоМед»
Гогоберидзе Ю.Т., старший разработчик ООО «ФтизисБиоМед»
Сабитов Р.А., старший ML-инженер ООО «ФтизисБиоМед»
Фатихов А.Ф., ML-инженер ООО «ФтизисБиоМед»

7. Материально-техническое обеспечение реализации рабочей программы

Необходимый для реализации программы модуля перечень материально-технического обеспечения включает в себя специально оборудованные помещения для проведения учебных занятий, в том числе:

аудитории, оборудованные мультимедийными и иными средствами обучения, позволяющими использовать информационные технологии, в количестве, позволяющем обучающимся осваивать умения и навыки, предусмотренные профессиональной деятельностью, индивидуально;

дисплейный класс, с мультимедийным и компьютерным оборудованием.

помещения для самостоятельной работы обучающихся оснащены компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду Университета.

Перечень информационных технологий, необходимых для освоения программы, включая перечень программного обеспечения и информационных справочных систем

Для достижения целей педагогического образования применяются следующие информационные технологии:

1.Образовательный портал дистанционного обучения Казанского ГМУ. Дистанционный курс в составе образовательного портала создан в системе MOODLE и содержит в себе лекции, презентации, задания, гиперссылки на первоисточники учебного материала, тесты / задания для самоконтроля, контрольные и итоговые тесты по курсу.

2.Пакет прикладных программ МойОфис.

3. Программа для ЭВМ StatTech (Статтех).

4. Демо-версия ГИС ЭЗ РТ

Используемое программное обеспечение имеет лицензию и ежегодно и / или своевременно обновляется.

8. Информационное обеспечение реализации рабочей программы

8.1. Основная литература.

Основы цифровой трансформации общества : учебник для вузов / Н. А. Горелов, О. Н. Кораблева. - 2-е изд., перераб. и доп. - Москва : Юрайт, 2024. – 336.

Цифровая трансформация здравоохранения и ведомственной медицины / Столяр В. П., Крайнюков П. Е., Калачёв О. В. - Москва : Планета, 2020. - 199

Статистика в биомедицине, фармации и фармацевтике : учебное пособие / Н. Н. Зубов, В. И. Кувакин, С. З. Умаров [и др.] ; под общей редакцией И. А. Наркевича ; Санкт-Петербургский государственный химико-фармацевтический университет. - Москва : КноРус, 2021. – 298.

8.2. Дополнительная литература.

Бююль А., Цефель П. SPSS: искусство обработки информации. Анализ статистических данных и восстановление скрытых закономерностей. — СПб.: ДиаСофтЮП, 2005.

Введение в биостатистику для медиков / С. Л. Плавинский ; Открытый институт здоровья. - М. : [б. и.], 2011. - 582, [2] с. : рис., табл. ; 21 см. - Библиогр.: с. 579-582.

Принципы и практика клинических исследований / Д. И. Галлин, Ф. П. Огнибене ; пер. с англ. под общ. ред. Г. Т. Сухих. - Москва : Практическая медицина, 2013. - XVIII, 451 с. : ил. ; 29 см. - Библиогр. в конце глав. - Пер. изд. : Principles and practice of clinical research / J. I. Gallin, F. P. Ognibene

Программирование на Python в примерах и задачах / Алексей Васильев. - Москва : Эксмо, 2021. - 616 с.

Федеральный закон "О персональных данных" от 27.07.2006 N 152-ФЗ. Доступно по: https://www.consultant.ru/document/cons_doc_LAW_61801/ Ссылка активна на 11.07.2024.

Концепция создания единой государственной информационной системы в сфере здравоохранения. Приказ Минздравсоцразвития России от 28.04.2011 г. № 364. Доступно по: http://portal.egisz.rosminzdrav.ru/materials/99. Ссылка активна на 11.07.2024.

Программирование на Python в примерах и задачах / Алексей Васильев. - Москва : Эксмо, 2021. - 616 с.

Столбов А.П., Кузнецов П.П. Автоматизированная обработка и защита персональных данных в медицинских учреждениях. - М.: ИД «Менеджер здравоохранения», 2010. - 176 с.

Силен Д., Мейсман А., Али М. «Основы Data Science и Big Data. Python и наука о данных». — СПб.: Питер, 2023.

8.3. Интернет-ресурсы.

Единая образовательная платформа Университета Иннополис https://oneid.unionepro.ru
Официальный сайт Stattech https://stattech.ru/
Официальный сайт ФГБУ «Центральный научно-исследовательский институт организации и информатизации здравоохранения» МЗ РФ https://mednet.ru/
Сайт Федеральной электронной медицинской библиотеки − http://www.femb.ru
Сайт ФГБОУ ВО Казанского ГМУ Минздрава России – https://kazangmu.ru/
Сайт Научной электронной библиотеки − http://elibrary.ru/
Паспорт национального проекта «Здравоохранение» http://government.ru/info/35561/
Официальный сайт ФГБУ «Центральный научно-исследовательский институт организации и информатизации здравоохранения» МЗ РФ https://mednet.ru/
Портал информационной поддержки пользователей ГИС ЭЗ РТ https://wiki.vitacore.ru/display/rmis

Портал нормативно-справочной информации ЕГИСЗ https://nsi.rosminzdrav.ru/

VI. Итоговая аттестация по Программе

После завершения обучения по Программе и прохождения итоговой оценки сформированности цифровых компетенций обучающиеся допускаются к итоговой аттестации.

Итоговая аттестация проводится с участием представителей профильных индустриальных партнёров в форме демонстрационного экзамена и предусматривает выполнение обучающимся профессиональных задач и оценку результатов и/или процесса выполнения – проверку сформированности в рамках Программы цифровых компетенций.

Задания демонстрационного экзамена разрабатываются с участием организаций-работодателей, отраслевых партнёров и профессиональных сообществ. Демонстрационный экзамен должен предусматривать выполнение (демонстрацию) обучающимся деятельности, завершающейся получением результата (продукта или его элемента), значимого при выполнении трудовой функции или трудовых действий.

Для обеспечения организации и проведения итоговой аттестации разрабатывается положение об итоговой аттестации, регулирующее требования к выполнению, оформлению и оцениванию работ, заданий, условия проведения итоговой аттестации, требования к составу аттестационной комиссии. Состав комиссии, перечень тем итоговых аттестационных работ, портфолио, практических заданий и требований к выполнению разрабатывается и актуализируется при участии индустриальных партнёров.

Примеры тем и заданий для демонстрационного экзамена

Перечень тем, образцы заданий для каждой из цифровых компетенций, включая предмет и показатели оценки

Примеры тем для компетенции ID 11 «Использует специализированное программное обеспечение для анализа медико- биологических данных»:

1. Методы анализа данных наиболее часто применяются в здравоохранении
2. Модели для предсказания риска сердечно-сосудистых заболеваний
3. Оценка эффективности нового медицинского препарата
4. Анализ данных о пациентах для улучшения персонализированного лечения

Примеры тем для компетенции ID 28 «Применяет языки программирования для решения профессиональных задач»:

1. Обработка отсутствующих данных в медицинских наборах данных
2. Процесс очистки данных здравоохранения
3. Выявление и обработка выбросов данных

Примеры тем для компетенции ID 170 «Осуществляет сбор и подготовку данных для обучения моделей искусственного интеллекта»

1. Оценка качества данных
2. Исследовательский анализ данных (Exploratory Data Analysis, EDA), для выявления основных тенденций и закономерностей.

3. Построение моделей машинного обучения для предсказания целевой переменной

Примеры заданий для оценки компетенций:

ID 11 «Использует специализированное программное обеспечение для анализа медико-биологических данных»:

Задание: Оценка эффективности нового медицинского препарата

Описание задания: Используйте клинические данные об эффективности нового препарата против диабета, собранные в ходе испытаний. Проведите анализ данных, чтобы определить, насколько эффективен новый препарат по сравнению с традиционными методами лечения.

Критерий оценки: самостоятельно применяет оптимальные алгоритмы анализа медико-биологических данных, используя специализированное программное обеспечение. Выбирает и применяет оптимальные критерии и методы для описания данных и аналитики

ID 28 «Применяет языки программирования для решения профессиональных задач»:

Задание: предварительная обработка данных для анализа

Описание задания: обработать предоставленный набор данных, содержащий информацию о пациентах, посещениях медицинских учреждений и проведенных процедурах. Провести предварительную обработку данных, очистку, проверку на корректность и подготовку к дальнейшему анализу.

Критерии оценки: применяет языки программирования (Python) для решения задания. Представлен Python-скрипт с кодом выполнения всех шагов. Отчет в формате PDF, описывающий процесс и выводы.

ID 170 «Осуществляет сбор и подготовку данных для обучения моделей искусственного интеллекта»

Задание: применение метода анализа больших данных для анализа данных пациентов ССЗ

Описание задания: провести анализ набора данных, содержащих информацию о пациентах и их медицинских историях, чтобы выявить ключевые факторы, влияющие на развитие хронических заболеваний. Результаты анализа помогут в разработке стратегии для улучшения профилактики и лечения этих заболеваний.

Критерии оценки: Осуществляет критический отбор данных, проверяя их на целостность и непротиворечивость. Использует методы поиска данных и достоверные источники данных

VII. Завершение обучения по Программе

Лицам, завершившим обучение по Программе и достигших целевого уровня сформированности цифровых компетенций по результатам итоговой оценки и прошедших итоговую аттестацию, присваивается дополнительная ИТ-квалификация, установленная Программой.

При освоении Программы параллельно с получением высшего образования диплом о профессиональной переподготовке выдается не ранее получения соответствующего документа об образовании и о квалификации (за исключением лиц, имеющих среднее профессиональное или высшее образование).

Лицам, не прошедшим итоговую аттестацию или получившим на итоговой аттестации неудовлетворительные результаты, а также лицам, освоившим часть Программы и (или) отчисленным из образовательной организации высшего образования, реализующей Программу, выдается справка об обучении или о периоде обучения по образцу, самостоятельно устанавливаемому образовательной организацией высшего образования.