

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ
ФЕДЕРАЦИИ

Федеральное государственное автономное образовательное учреждение высшего
образования "Казанский (Приволжский) федеральный университет"
Институт математики и механики им. Н.И. Лобачевского



УТВЕРЖДЕНА

Проректор по образовательной
деятельности КФУ

Е.А. Турилова

20__ г.



**Дополнительная профессиональная программа
(программа профессиональной переподготовки)**

Анализ данных в педагогической деятельности

(наименование программы)

Утверждена Ученым советом Института математики и механики им.Н.И.Лобачевского
КФУ (протокол № 8 от «13» апреля 2023 г.)

Председатель Ученого совета Турилова Екатерина Александровна

(подпись)

Руководитель подразделения,
реализующего ДПП ПП

«__» _____ 20__ г.

(подпись)

Е.А. Турилова
(инициалы, фамилия)

Казань, 2023 г.

I. Общие положения

1. Дополнительная профессиональная программа (программа профессиональной переподготовки) ИТ-профиля «Аналитика данных в педагогической деятельности» (далее – Программа) разработана в соответствии с нормами Федерального закона РФ от 29 декабря 2012 года № 273-ФЗ «Об образовании в Российской Федерации», с учетом требований приказа Минобрнауки России от 1 июля 2013 г. № 499 «Об утверждении Порядка организации и осуществления образовательной деятельности по дополнительным профессиональным программам», с изменениями, внесенными приказом Минобрнауки России от 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», приказа Министерства образования и науки РФ от 23 августа 2017 г. N 816 «Об утверждении Порядка применения организациями, осуществляющими образовательную деятельность, электронного обучения, дистанционных образовательных технологий при реализации образовательных программ» (указать при необходимости); паспорта федерального проекта «Развитие кадрового потенциала ИТ-отрасли» национальной программы «Цифровая экономика Российской Федерации»; постановления Правительства Российской Федерации от 13 мая 2021 г. № 729 «О мерах по реализации программы стратегического лидерства «Приоритет-2030» (в редакции постановления Правительства Российской Федерации от 14 марта 2022 г. № 357 «О внесении изменений в постановление Правительства Российской Федерации от 13 мая 2021 г. № 729»); приказа Министерства цифрового развития, связи и массовых коммуникаций Российской Федерации от 28 февраля 2022 г. № 143 «Об утверждении методик расчета показателей федеральных проектов национальной программы «Цифровая экономика Российской Федерации» и признании утратившими силу некоторых приказов Министерства цифрового развития, связи и массовых коммуникаций Российской Федерации

Федерации об утверждении методик расчета показателей федеральных проектов национальной программы «Цифровая экономика Российской Федерации» (далее – приказ Министерства цифрового развития, связи и массовых коммуникаций Российской Федерации № 143); федерального государственного образовательного стандарта 01.03.01 «Математика», (далее вместе – ФГОС ВО), а также профессионального стандарта «Системный аналитик», утвержденного приказом Минтруда России от 28.10.2014 N 809н .

2. Профессиональная переподготовка заинтересованных лиц (далее – Слушатели), осуществляемая в соответствии с Программой (далее – Подготовка), имеющей отраслевую направленность¹ «Образование», проводится в ФГАО ВО «Казанский (Приволжский) федеральный университет» (далее – Университет) в соответствии с учебным планом в очной форме обучения².

3. Разделы, включенные в учебный план Программы, используются для последующей разработки календарного учебного графика, учебно-тематического плана, рабочей программы, оценочных и методических материалов. Перечисленные документы разрабатываются Университетом самостоятельно, с учетом актуальных положений законодательства об образовании, законодательства в области информационных технологий и смежных областей знаний ФГОС ВО и профессионального стандарта «Системный аналитик».

4. Программа регламентирует требования к профессиональной переподготовке в области разработки требований к программному обеспечению, продукту, средству.

Срок освоения Программы составляет 9 месяцев, общая трудоемкость 252 часа.

К освоению Программы в рамках проекта допускаются лица:

- получающие высшее образование по очной (очно-заочной) форме, лица,

¹ Варианты отраслевой направленности: «Городское хозяйство»; «Финансовые услуги»; «Строительство»; «Добывающая промышленность»; «Обрабатывающая промышленность»; «Транспортная инфраструктура»; «Здравоохранение»; «Энергетическая инфраструктура»; «Образование»; «Сельское хозяйство и агропромышленный комплекс»; «Информационно-коммуникационные технологии»; «Искусство и культура»

² При реализации Программы допускается использовать сетевую форму обучения с организациями реального сектора экономики субъекта Российской Федерации

освоившие основную профессиональную образовательную программу (далее – ОПОП ВО) бакалавриата – в объеме не менее первого курса (бакалавры 2-го курса), ОПОП ВО специалитета – не менее первого и второго курсов (специалисты 3-го курса). Также к освоению ДПП ПП допускаются лица, обучающиеся по программам магистратуры, которые не относятся к ИТ-профилю (согласно приложению к Методике расчета показателя граждан, прошедших обучение по дополнительным образовательным программам) и по программам ординатуры.

5. Область профессиональной деятельности: Об «Связь, информационные и коммуникационные технологии».

II. Цель

6. Целью подготовки слушателей по Программе является получение компетенции обучающимися по специальностям и направлениям подготовки, не отнесенным к ИТ-сфере³, необходимой для выполнения нового вида профессиональной деятельности в области связи и информационных и коммуникационных технологий; приобретение новой квалификации «Системный аналитик» обучающимися по специальностям и направлениям подготовки, не отнесенным к ИТ-сфере. Освоение слушателями методов эмпирических исследований и количественного анализа данных с применением современных способов хранения (облачные хранения) и обработки (Python) социально-педагогической информации.

³Указать целевые группы обучающихся, определенные паспортом Федерального проекта: – обучающиеся по специальностям и направлениям подготовки, не отнесенным к ИТ-сфере, – обучающиеся по специальностям и направлениям подготовки ИТ-сферы (выбрать нужное)

III. Характеристика новой квалификации и связанных с ней видов профессиональной деятельности, трудовых функций и (или) уровней квалификации

7. Виды профессиональной деятельности, трудовая функция, указанные в профессиональном стандарте по должности «Младший системный аналитик», представлены в таблице 1:

Таблица 1

Характеристика новой квалификации, связанной с видом профессиональной деятельности и трудовыми функциями в соответствии с профессиональным стандартом «Системный аналитик»

Область профессиональной деятельности	Тип задач профессиональной деятельности	Код и наименование профессиональной компетенции	Трудовые действия	Трудовая функция	Обобщенная трудовая функция	Вид профессиональной деятельности
Связь, информационные и коммуникационные технологии	Проектный	ПК -1 Применяет методы системного анализа для разработки алгоритмов	Выявление существенных явлений проблемной ситуации Установка причинно-следственных связей между явлениями проблемной ситуации Установка причин проблем, которые могут быть устранены за счет автоматизации Определение и описание технико-экономических характеристик вариантов концептуальной архитектуры	Анализ проблемной ситуации заинтересованных лиц Разработка концепции системы	Концептуальное, функциональное и логическое проектирование систем среднего и крупного масштаба и сложности	Проектно-исследовательская деятельность в области информационных технологий

Таблица 2

Характеристика новой и развиваемой цифровой компетенции в ИТ-сфере, связанной с уровнем формирования и развития в результате освоения Программы⁴ «Анализ данных в педагогической деятельности»

Наименование сферы	Код и наименование профессиональной компетенции	Пример инструментов	0 — способность не проявляется/ проявляется в степени, недостаточной для отнесения к 1 уровню сформированности компетенции	1 — способность проявляется под внешним контролем / при внешней постановке задачи/ обучающийся пользуется готовыми, рекомендованными продуктами	2 — способность проявляется, но обучающийся эпизодически прибегает к экспертной консультации/ самостоятельно подбирает и пользуется готовыми продуктами	3 — способность проявляется системно / обучающийся модифицирует способность под определенные задачи / создает новый продукт, обучает других
Средства программной разработки	ПК-2 Применяет языки программирования для решения профессиональных задач (ID-28)	Python	(+)	(+)	(+)	(-)
Информационные ресурсы и продукты	ПК-3 Оценивает результаты обучения с использованием	Moodle	(+)	(+)	(+)	(-)

⁴ На основании Матрицы компетенций, актуальных для цифровой экономики, указанной в Приложении 1 в Требованиях к ДПП ПП.

	цифровых ресурсов и продуктов (ID-247)					
--	---	--	--	--	--	--

IV. Характеристика новых и развиваемых цифровых компетенций, формирующихся в результате освоения программы

8. В ходе освоения Программы Слушателем приобретаются следующие профессиональные компетенции:

- ПК-1 Применяет методы системного анализа для разработки алгоритмов

9. В ходе освоения Программы Слушателем совершенствуются следующие профессиональные компетенции:

- ПК-2 Применяет языки программирования для решения профессиональных задач;

- ПК-3 Оценивает результаты обучения с использованием цифровых ресурсов и продуктов.

V. Планируемые результаты обучения по ДПП ПП

10. Результатами подготовки слушателей по Программе является получение компетенции, необходимой для выполнения нового вида профессиональной деятельности в области связи и информационных и коммуникационных технологий; приобретение новой квалификации «Системный аналитик».

11. В результате освоения Программы слушатель должен:

ПК-1 Применяет методы системного анализа для разработки алгоритмов.

Знать:

- базовые способы проверки работоспособности кода программного обеспечения,
- наиболее простые способы интеграции программных модулей и компонент.

Уметь:

- выполнять базовую проверку работоспособности кода программного обеспечения,
- интегрировать программные модули и компоненты в рамках простых информационных систем.

Иметь навыки:

- проверки работоспособности и простого рефакторинга кода программного обеспечения,
- интеграции программных модулей и компоненты для простых информационных систем.

ПК-2 Применяет языки программирования для решения профессиональных задач.

Знать:

- синтаксис выбранного языка программирования, особенности программирования на этом языке, стандартные библиотеки языка программирования
- методологии и технологии проектирования и использования баз данных
- технологии программирования
- особенности выбранной среды программирования и системы управления базами данных
- методы повышения читаемости программного кода
- средства проверки работоспособности и отладки программного обеспечения.

Уметь:

- выбирать средства реализации требований к программному обеспечению
- вырабатывать варианты реализации программного обеспечения.

Иметь навыки:

- использования методов и приемов формализации и алгоритмизации поставленных задач

- использования выбранной среды программирования и средства системы управления базами данных
- применения современных компиляторов, отладчиков и оптимизаторов программного кода
- применения инструментария для создания и актуализации исходных текстов программ.

ПК-3 Оценивает результаты обучения с использованием цифровых ресурсов и продуктов.

Знать:

- теоретические основы и приемы проектирования баз данных.

Уметь:

- использовать реляционные базы данных при проектировании и разработке программного обеспечения
- проводить оценку и обоснование рекомендуемых решений.

Иметь навыки:

- использования одной из популярных систем управления базами данных SQL.

VI. Организационно-педагогические условия реализации ДПП

12. Реализация Программы должна обеспечить получение компетенции, необходимой для выполнения нового вида профессиональной деятельности в области связи, информационных и коммуникационных технологий; приобретение новой квалификации «Системный аналитик».

13. Учебный процесс организуется с применением электронного обучения и дистанционных образовательных технологий, инновационных технологий и методик обучения, способных обеспечить получение слушателями знаний, умений и навыков в области связи, информационных и коммуникационных технологий.

14. Реализация Программы обеспечивается научно-педагогическими кадрами Университета, допустимо привлечение к образовательному процессу высококвалифицированных специалистов ИТ-сферы и/или дополнительного профессионального образования в части, касающейся профессиональных компетенций в области разработки мобильных приложений различного назначения., с обязательным участием представителей профильных организаций-работодателей. Возможно привлечение региональных руководителей цифровой трансформации (отраслевых ведомственных и/или корпоративных) к проведению итоговой аттестации, привлечение работников организаций реального сектора экономики субъектов Российской Федерации. Не менее 50% общего объема аудиторных часов в рамках ДПП ПП реализуются научно-педагогическими работниками, отвечающими следующим критериям:

– наличие высшего профильного образования в ИТ-сфере и/или дополнительного профессионального образования в части, касающейся профессиональных компетенций в области разработки мобильных приложений различного назначения.;

– наличие стажа педагогической работы в образовательных организациях высшего образования Российской Федерации и/или стажа практической работы в профильной организации ИТ-отрасли не менее 3 лет.

Не менее 20% от общего объема аудиторных часов в рамках ДПП ПП реализуются лицами, имеющими подтвержденный стаж в профессии в ИТ-сфере или в отрасли цифровой экономики не менее двух лет, полученный не более четырех лет назад.

VII. Учебный план ДПП

15. Объем Программы составляет 9 месяцев, трудоемкость – 252 часа.

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

Учебный план программы профессиональной переподготовки «Анализ
данных в педагогической деятельности»

№ п/п	Наименование раздела (модуля)	Общая трудоемкость (252 часов)	Форма контроля
1.	Введение в математическую статистику.	52	тест
2.	Применение Python для анализа данных	52	зачет
3.	Машинное обучение на Python. Современные хранилища данных, анализ больших данных	52	зачет
	Промежуточная аттестация	6	Итоговое тестирование
	Практика	54	зачет
	Итоговая аттестация	36	Итоговый проект
	Итого:	252	

VIII. Календарный учебный график

18. Календарный учебный график представляет собой график учебного процесса, устанавливающий последовательность и продолжительность обучения и итоговой аттестации по учебным дням.

Календарный учебный график программы профессиональной переподготовки «Анализ данных в педагогической деятельности»

№	Наименование раздела (модуля)	Учебные недели																																						
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36			
1.	Введение в математическую статистику.	+	+	+	+	+	+																																	
2.	Применение Python для анализа данных							+	+	+	+	+																												
3.	Машинное обучение на Python. Современные хранилища данных, анализ больших данных.													+	+	+	+	+																						
4.	Промежуточная аттестация																			+																				
5.	Практика на предприятии																					+	+	+	+	+	+	+	+	+	+									
6.	Итоговая аттестация (подготовка и защита итогового проекта)																																			+	+	+	+	+

IX. Рабочая программа учебных предметов, курсов, дисциплин (модулей)

19. Рабочая программа содержит перечень разделов и тем, а также рассматриваемых в них вопросов с учетом их трудоемкости.

Рабочая программа разрабатывается Университетом с учетом профессионального стандарта «Системный аналитик», утвержденного приказом Минтруда России от 28.10.2014 N 809н .

№ п/п	Наименование и краткое содержание раздела(модуля)	Объем, часов
1.	<p>Введение в математическую статистику. <i>Краткое содержание:</i> Основы теории вероятностей для изучения математической статистики. Простой случайный выбор. Генеральная совокупность и выборка. Характеристики выборки. Понятие о статистическом моделировании. Понятие оценок параметров выборки. Свойства оценок, методы нахождения оценок. Интервальные оценки. Оценка параметров выборки из нормального распределения. Критерии значимости. Проверка статистических гипотез. Общие понятия теории Пирсона-Неймана проверки статистических гипотез. Понятие о регрессионном анализе. Линейная регрессия.</p>	52
2.	<p>Применение Python для анализа данных <i>Краткое содержание:</i> Основы программирования на Python. Синтаксис языка Python. Форматирование блоков кода. Числовые, булевы, строковые литералы. Основные операции над простыми типами. Простые управляющие конструкции. Декларация функций. Модули (импортирование). Соглашения по оформлению кода (PEP8). Организация кода в модули и пакеты Встроенные типы данных. ООП. Исключения в Python. Регулярные выражения в Python. Модуль re. Функциональные методы программирования. Декораторы. Дата и время в Python. Основные библиотеки Python для анализа данных и машинного обучения. Обработка и разведочный анализ данных (EDA) с использованием библиотек pandas и numpy Визуализация данных с использованием библиотек matplotlib, seaborn, plotly</p>	52
3.	<p>Машинное обучение на Python. Современные хранилища данных, анализ больших данных <i>Краткое содержание:</i></p>	52

	<p>Машинное обучение для решения педагогических задач. Три типа машинного обучения.</p> <p>Основные фреймворки машинного обучения</p> <p>Обучение с учителем: задача регрессии и классификации. Деревья решений, бэггинг, бустинг, стекинг.</p> <p>Обучение без учителя: кластерный анализ, алгоритм k-means.</p> <p>Обучение с подкреплением</p> <p>Введение в глубокое обучение. Нейронные сети.</p> <p>Введение в Big Data.</p>	
4.	<p>Промежуточная аттестация</p> <p><i>Краткое содержание:</i></p> <p>Методы математической статистики, основные выборочные характеристики.</p> <p>Приёмы предварительного анализа данных.</p> <p>Математические основы методов классификации данных.</p> <p>Машинное обучение и его типы: с учителем, без учителя, с подкреплением.</p> <p>Примеры алгоритмов: decision tree, k-means, nearest neighbor</p>	6
5.	<p>Практика на предприятии</p> <p><i>Краткое содержание:</i></p> <p>Практика проводится на базе организаций различных организационно-правовых форм и форм собственности или их основных структурных подразделений, осуществляющих деятельность, соответствующую виду (видам) деятельности, к которому (которым) готовится обучающийся.</p>	54
6.	<p>Итоговая аттестация (подготовка и защита итогового проекта)</p> <p><i>Краткое содержание:</i></p> <p>В ходе итоговой аттестации обучающиеся обеспечивают презентацию (защиту) разработанного цифрового решения (проекта), а также перечня решаемых им проблем. Оценивается использование технологий, изученных в курсе, самостоятельность выполняемого решения, степень участия каждого члена команды в разработке.</p>	36

20. Учебно-тематический план Программы определяет тематическое содержание, последовательность разделов и (или) тем и их трудоемкость.

Учебный план 1 модуля «Введение в математическую статистику»

№ п/п	Наименование учебных тем, курсов, дисциплин (модулей), вида аттестации	Всего ауд. часов	Аудиторная работа		Сам. работа	Форма промежуточного контроля
			Лекции	Практ. занятия		
1	Основы теории вероятностей для изучения математической статистики	4	2	2	4	Тест №1
2	Простой случайный выбор. Генеральная совокупность и	8	4	4	2	

	выборка. Характеристики выборки. Понятие о статистическом моделировании.					
3	Понятие оценок параметров выборки. Свойства оценок, методы нахождения оценок. Интервальные оценки. Оценка параметров выборки из нормального распределения.	4	2	2	2	
4	Критерии значимости. Проверка статистических гипотез.	8	4	4	2	Практическая работа №1.
5	Общие понятия теории Пирсона- Неймана проверки статистических гипотез.	8	4	4	2	
6	Понятие о регрессионном анализе. Линейная регрессия.	4	2	2	4	
	Тест по модулю 1					2

Учебный план 2 модуля «Применение Python для анализа данных»

№ п/п	Наименование учебных тем, курсов, дисциплин (модулей), вида аттестации	Всего ауд. часов	Аудиторная работа		Сам. работа	Форма промежуточного контроля
			Лекции	Практ. занятия		
1	Основы программирования на Python. Синтаксис языка Python.	10	6	4		
2	Основные библиотеки Python для анализа данных и машинного обучения.	8	4	4		Практическая работа №2
3.	Обработка и разведочный анализ данных (EDA) с использованием библиотек pandas и numpy	12	6	4		

4.	Визуализация данных с использованием библиотек matplotlib, seaborn, plotly	6	2	4		
5	Тест по модулю 2	2				

Учебный план 3 модуля «Машинное обучение на Python. Современные хранилища данных, анализ больших данных».

№ п/п	Наименование учебных тем, курсов, дисциплин (модулей), вида аттестации	Всего ауд. часов	Аудиторная работа		Сам. работа	Форма промежуточного контроля
			Лекции	Практ. занятия		
1	Машинное обучение для решения педагогических задач. Три типа машинного обучения.	10	4	6		
2	Основные фреймворки машинного обучения	4	2	2		
3.	Обучение с учителем: задача регрессии и классификации. Деревья решений, бэггинг, бустинг, стекинг.	8	4	4		
4.	Обучение без учителя: кластерный анализ, алгоритм k-means.	6	4	2		Практическая работа №3
5.	Обучение с подкреплением					
6.	Введение в глубокое обучение. Нейронные сети.	4	2	2		
7.	Введение в Big Data.	4	2	2		
8.	Тест по модулю 3	2				

№ п/п	Наименование раздела(модуля)	Количество часов		
		аудиторных		самостоятельной работы (выполнение _____ * заданий)
		Лекции	Семинары	
1.	Введение в математическую статистику	18	18	16
2.	Применение Python для анализа данных	18	18	16
3.	Машинное обучение на Python. Современные хранилища данных, анализ больших данных.	18	18	16
4.	Практика	54		
	Промежуточная аттестация	6		
	Итоговая аттестация	36		

**указать вид (-ы) запланированной самостоятельной работы*

Х. Формы аттестации

21. Слушатели, успешно выполнившие все элементы учебного плана, допускаются к итоговой аттестации.

Итоговая аттестация по Программе проводится в форме защиты итогового проекта.

22. Лицам, успешно освоившим Программу (в области создания алгоритмов и программ, пригодных для практического применения, или навыков использования и освоения цифровых технологий, необходимых для выполнения нового вида профессиональной деятельности) и прошедшим итоговую аттестацию в рамках проекта «Цифровые кафедры», выдается документ о квалификации: диплом о профессиональной переподготовке.

При освоении ДПП III 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).

23. Лицам, не прошедшим итоговую аттестацию или получившим на итоговой аттестации неудовлетворительные результаты, а также лицам, освоившим часть Программы и (или) отчисленным из Университета, выдается справка об обучении или о периоде обучения по образцу, самостоятельно устанавливаемому Университетом.

XI. Оценочные материалы

24. Контроль знаний, полученных слушателями при освоении разделов (модулей) Программы, осуществляется в следующих формах:

- текущий контроль успеваемости – обеспечивает оценивание хода освоения разделов Программы, проводится в форме решения тестов и выполнения практических заданий;

- промежуточная аттестация – завершает изучение отдельного модуля Программы, проводится в форме итогового тестирования;

- итоговая аттестация – завершает изучение всей программы, проводится в форме итогового проекта.

25. В ходе освоения Программы каждый слушатель выполняет следующие отчетные работы:

№ п/п	Наименование раздела (модуля)	Задание	Критерии оценки
1.	Введение в математическую статистику	Практическое задание	50% процентов выполненных требований к заданию
		Тест	60% правильных ответов на тест
2.	Применение Python для анализа данных	Практическое задание	50% процентов выполненных требований к заданию
		Тест	60% правильных ответов на тест

3.	Машинное обучение на Python. Современные хранилища данных, анализ больших данных.	Практическое задание	50% процентов выполненных требований к заданию
		Тест	60% правильных ответов на тест
4.	Практика	Выполнение задания на практику	Отчет по практике
5.	Промежуточная аттестация	Тест	60% правильных ответов на тест
6.	Итоговая аттестация	Выполнение и защита проекта	Оценка экзаменационной комиссии на основе выполнения требований, указанных в описании проекта

26. Текущий контроль. Перечень примерных тестовых заданий

26.1 Модуль «Введение в математическую статистику»

26.1.1 Тема «Основы теории вероятностей для изучения математической статистики» Образец вопросов теста

Вопрос 1. Размещения – это

- а) возможность переставлять местами набор элементов;
- б) комбинации, составленные выбором из различных элементов различных элементов, отличающиеся либо составом элементов, либо порядком их следования;
- в) комбинации m элементов из n элементов, отличающиеся составом или порядком следования, причем выбранный элемент возвращается на место и может участвовать в дальнейшем выборе;
- г) комбинации, составленные выбором различных элементов из различных элементов, отличающиеся только составом (но не порядком следования);

Вопрос 2. Классическая формула для вычисления вероятности применима

- а) в любом опыте;
- б) если опыт обладает равновозможностью исходов;

в) если исходы опыта образуют исчерпывающий набор его равновозможных и исключающих друг друга исходов;

г) если исходы опыта образуют последовательность зависимых друг от друга событий

Вопрос 3. Вероятность невозможного события

а) неопределенна; б) равна нулю;

в) равна 0,5; г) равна единице.

26.1.2 Тема «Критерии значимости. Проверка статистических гипотез»

Образец практического задания

1. Получены данные о результатах теста в двух группах:

При следующих данных проверить гипотезу о равенстве среднего балла в двух группах, при условии, что ГС не нормально распределены.

X	15.6, 15.8, 16.0, 15.7, 15.4, 18.0, 16.5, 16.3, 16.0, 15.6
Y	16.1, 16.3, 16.0, 17.0, 17.2, 17.3, 15.5, 15.2, 18.1, 14.9

26.2 Модуль «Применение Python для анализа данных»

26.2.1. Тема «Основы программирования на Python. Синтаксис языка Python» . Образец практическая работа №1.

Цель: ознакомиться с основами языка Python, получить умения для выполнения дальнейших лабораторных работ.

Задачи практической работы:

- изучить типизацию данных;
- научиться пользоваться циклами «for» и «while»;
- рассмотреть «ветвление» в Python;
- отработать задачи с использованием конструкции «try-except»;
- разобрать функции и пространства имён.

1. Проверьте на истинность выражения :

$1 > 2$

$32 > 23$

'a' < 'b'

'c' < 'b'

'a' in ['b','a','c']

2. Напишите функцию `f_abs()` нахождения модуля числа, которая получает на вход число и возвращает его, если оно положительно, иначе возвращает число с обратным знаком. Примените ее к числу $170**2 - 150**3$.

26.2.2. Темы: «Обработка и разведочный анализ данных (EDA) с использованием библиотек `pandas` и `numpy`», « Визуализация данных с использованием библиотек `matplotlib`, `seaborn`, `plotly`»

Образец практической работы №2.

Для `house_train.csv`:

- 1.) провести разведочный анализ данных
- 2.) Получить основную информацию о данных: количество наблюдений, количество переменных, типы данных, количество пропущенных значений.
- 3.) Вычислить основные статистические характеристики данных: средние значения, дисперсии, корреляции, минимумы, максимумы, квантили.
- 4.) Нарисовать гистограммы и боксплоты.
- 5.) построить графики зависимостей целевой переменной от остальных признаков.
- 6.) Сделать выводы по результатам анализа полученных графиков.
(обязательно использовать `pandas`, `numpy`, `matplotlib`)

26.3. Модуль «Машинное обучение на Python. Современные хранилища данных, анализ больших данных».

26.3.1. Тема « Обучение с учителем: задача регрессии и классификации. Деревья решений, бэггинг, бустинг, стекинг.»

Образец практической работы №3. Построить модель линейной регрессии энергопотребления здания, используя температуру воздуха (`air_temperature`) и влажность (`dew_temperature`).

Рассчитать качество построенной модели по проверочным данным.

26.3.2. Тема «Обучение без учителя: кластерный анализ, алгоритм k-means.» Образец практической работы №4.

Используйте готовый ноутбук с кластеризацией базы покупок интернет магазина:

1. Удалите из выборки все строки про доставку (автоматически, уже из `carts`)
2. Выведите две гистограммы
 - А.) Распределения размеров классов - сколько корзин
 - Б.) Распределение количества уникальных позиций в классе
3. Напишите продвинутую функцию визуализации одного выбранного класса
 - А.) Размер класса
 - Б.) Количество уникальных позиций
 - В.) Позиции, которые есть минимум в 20% корзин
 - Г.) Уникальные слова в описаниях позиций, которые встречаются минимум в 20% позиций, для отслеживания моно-брендовых и моно-товарных закупок

27. Промежуточная аттестация. Итоговое тестирование по каждому модулю.

Образцы вопросов:

1. Совокупность всех возможных объектов данного вида, над которыми проводятся наблюдения с целью получения конкретных значений определенной случайной величины называется ...

а) выборкой; б) вариантами;

в) генеральной совокупностью; г) выборочной совокупностью.

2. Основное условие, которому должна удовлетворять наилучшая оценка:

а) математическое ожидание квадрата отклонения оценки от оцениваемого параметра должно быть как можно меньшим;

б) оценка должна быть как можно меньшим числом;

в) предел разности между оценкой и оцениваемым параметром должен быть как можно меньшим;

г) такового нет.

3. Какие характеристики можно отнести к языку программирования Python?

а) имеет эффективный компилятор в коды процессора

б) использует раннее связывание

в) объектно-ориентированный

г) универсальный язык программирования

4. Машинное обучение - это ...

а) научное направление, задачей которого является создание интеллектуальных систем, лежит на стыке информатики, статистики и анализа данных, а также занимается вопросами, связанными с философией и этичностью использования интеллектуальных систем

б) математическая область, связанная с построением предсказательных алгоритмов (как правило представленных статистическими моделями) на основе данных

в) объединение ряда научных областей, занимающихся построением систем анализа и обработки данных

г) прикладное направление, развивающее методы построения баз знаний и правил, явным образом описывающих знания экспертов

28. Итоговая аттестация. Перечень примерных заданий

1. Собрать данные о студентах и их успеваемости.

2. Получить основную информацию о данных: количество наблюдений, количество переменных, типы данных, количество пропущенных значений.

3. Вычислить основные статистические характеристики данных: средние значения, дисперсии, корреляции, минимумы, максимумы, квартили.

4. Нарисовать гистограммы, боксплоты, тепловую карту для матрицы корреляции.

5. Указать выбросы, если есть (вручную или любым другим эвристическим методом).

6. На основе выбранных данных провести алгоритм проверки гипотезы.

7. Выбрать два признака номер группы и количество баллов и построить линейную регрессию.

8. Используя алгоритм k-means разбить данные с результатами теста на группы

9. Написать отчет о полученных результатах

ХII. Материально-техническое и учебно-методическое обеспечение Программы

Электронная информационно-образовательная среда КФУ (ЭИОС) представляет собой совокупность электронных информационных ресурсов, электронных образовательных ресурсов, информационных технологий, телекоммуникационных технологий, соответствующих технологических средств, обеспечивающих освоение обучающимися образовательных

программ или их частей, а также взаимодействие между всеми участниками образовательного процесса независимо от места их нахождения;

ЭИОС обеспечивает:

—доступ к учебным планам, рабочим программам дисциплин (модулей), практик, к изданиям электронных библиотечных систем и электронным образовательным ресурсам;

—фиксацию хода образовательного процесса, результатов промежуточной аттестации и результатов освоения основной образовательной программы;

—проведение всех видов занятий, процедур оценки результатов обучения, реализация которых предусмотрена с применением электронного обучения, дистанционных образовательных технологий; <https://edu.kpfu.ru/>

—формирование электронного портфолио обучающегося, в том числе сохранение работ обучающегося, рецензий и оценок на эти работы со стороны любых участников образовательного процесса;

—взаимодействие между участниками образовательного процесса, в том числе синхронное и (или) асинхронное взаимодействие посредством сети «Интернет».

Система «Антиплагиат.ВУЗ» и другие ресурсы позволяющие обеспечивать освоение обучающимися образовательных программ в полном объеме независимо от места нахождения обучающихся.

Основными элементами ЭИОС КФУ являются:

а) электронные информационные ресурсы:

—официальный сайт КФУ (<https://kpfu.ru/>);

—личные кабинеты участников образовательного процесса, обеспечивающие доступ к

—компонентам ЭИОС КФУ;

—корпоративная электронная почта;

—сайт Научной библиотеки им. Н.И. Лобачевского;

—информационно-аналитическая система управления образовательным

процессом;

—система автоматического поиска текстовых заимствований;

—другие базы данных и файловые системы, используемые в образовательном процессе;

б) электронные образовательные ресурсы:

—система управления обучением Moodle;

—сайт дистанционного обучения (<https://edu.kpfu.ru/>), содержащий более 3500 цифровых образовательных ресурсов;

—площадка для создания и тестирования курсов (<https://do.kpfu.ru/>);

в) электронные библиотечные системы:

—внутренняя электронная библиотечная система КФУ, обеспечивающая доступ к информационным ресурсам, включающая печатные и электронные документы на русском и иностранных языках;

—внешние электронные библиотечные системы и электронные библиотеки, доступ к которым осуществляется на договорной основе;

г) средства вычислительной техники:

—серверное оборудование КФУ;

—компьютеры, эксплуатируемые в КФУ;

—ноутбуки, планшеты, смартфоны и другие портативные, мобильные персональные компьютеры;

—средства организационной и множительной техники;

—мультимедийное оборудование.

Система электронного (дистанционного) обучения (далее – СДО) – электронная информационно-образовательная среда в виде системно-организованной совокупности информационно-коммуникационных средств и технологий, процессов программно-аппаратного и организационно-методического обеспечения, деятельности научно-педагогического, педагогического, учебно-вспомогательного и инженерного персонала (работников), ориентированная на реализацию системы сопровождения учебного процесса с целью удовлетворения образовательных потребностей

обучающихся независимо от места их нахождения

Доступ в СДО обеспечивается непрерывно (в круглосуточном режиме с коэффициентом доступности всех компонентов среды не ниже 99,5 %) и из любой точки подключения к сети Интернет с заданными характеристиками канала связи.

Доступ ко всем сервисам СДО является персонализированным (под единой учетной записью).

Освоение ДПП ПП предполагает использование следующего программного обеспечения и информационно-справочных систем:

- Язык Python версии 3, среда выполнения языка
- Блокноты разработки – Jupyter Notebook, Google Colab
- СУБД PostgreSQL
- Библиотеки NumPy, Pandas, Scikit-learn, PyTorch, Tensorflow, Keras

1. XIII. Список литературы

2. Блягоз, З. У. Теория вероятностей и математическая статистика. Курс лекций : учебное пособие / З. У. Блягоз. - 2-е изд., испр. - Санкт-Петербург : Лань, 2018. - 224 с. - ISBN 978-5-8114-2934-9. - Текст : электронный // Лань : электронно-библиотечная система. - URL: <https://e.lanbook.com/book/103061> (дата обращения: 10.03.2020). - Режим доступа: для авториз. пользователей.
3. Бородин, А. Н. Элементарный курс теории вероятностей и математической статистики : учебное пособие / А. Н. Бородин. - 8-е изд., стер. - Санкт-Петербург : Лань, 2011. - 256 с. - ISBN 978-5-8114-0442-1. - Текст : электронный
4. // Лань : электронно-библиотечная система. - URL: <https://e.lanbook.com/book/2026> (дата обращения: 10.03.2020). - Режим доступа: для авториз. пользователей.
5. Гусева Е.Н., Теория вероятностей и математическая статистика / Е.Н. Гусева - Москва : ФЛИНТА, 2016. - 220 с.
6. ISBN 978-5-9765-1192-7 - Текст : электронный // ЭБС 'Консультант студента' : [сайт]. - URL : <http://www.studentlibrary.ru/book/ISBN9785976511927.html> (дата обращения: 10.03.2020). - Режим доступа : по подписке.
7. Емельянов, Г. В. Задачник по теории вероятностей и математической статистике : учебное пособие / Г. В. Емельянов, В. П. Скитович. - 3-е изд., стер. - Санкт-Петербург : Лань, 2019. - 332 с. - ISBN 978-5-8114-3984-3. - Текст : электронный // Лань : электронно-библиотечная система. - URL: <https://e.lanbook.com/book/113941> (дата обращения: 10.03.2020). - Режим доступа: для авториз. пользователей.
8. Хуснутдинов, Р. Ш. Сборник задач по курсу теории вероятностей и математической статистики : учебное пособие / Р. Ш. Хуснутдинов. - 2-е изд., испр. - Санкт-Петербург : Лань, 2014. -

320 с. - ISBN 978-5-8114-1668-4. - Текст : электронный // Лань :
электронно-библиотечная система. - URL:
<https://e.lanbook.com/book/53676> (дата обращения: 10.03.2020). -
Режим доступа: для авториз. пользователей.

9. Буре, В. М. Теория вероятностей и математическая статистика : учебник / В. М. Буре, Е. М. Парилина. - Санкт-Петербург : Лань, 2013. - 416 с. - ISBN 978-5-8114-1508-3. - Текст : электронный // Лань : электронно-библиотечная система. - URL: <https://e.lanbook.com/book/10249> (дата обращения: 10.03.2020). - Режим доступа: для авториз. пользователей.
10. Python и машинное обучение, Себастьян Рашка, Вахид Мирджалили
11. [Ян Лекун](#). Как учится машина. Революция в области нейронных сетей и глубокого обучения. (Библиотека Сбера: Искусственный интеллект). — М.: Альпина нон-фикшн, 2021. — [ISBN 978-5-907394-29-2](#).
12. Практическая статистика для специалистов Data Science | Брюс Эндрю, Брюс Питер
13. Гмурман, В.Е. Теория вероятностей и математическая статистика: Учебник для прикладного бакалавриата / В.Е. Гмурман. - Люберцы: Юрайт, 2016. - 479 с.
14. Кремер, Н.Ш. Теория вероятностей и математическая статистика: Учебник / Н.Ш. Кремер. - М.: Юнити, 2016. - 240 с
15. Python для сложных задач. Наука о данных и машинное обучение | Вандер Плас Дж.

Перечень Интернет-ресурсов, необходимых для освоения дисциплины

17. <https://habr.com/ru/all/> - Статьи о программировании
18. <https://www.kaggle.com/> - платформа для соревнований по машинному обучению
19. Константин Воронцов. [Курс «машинное обучение»](#) школы анализа

данных компании [Яндекс](#).-

https://academy.yandex.ru/dataschool/lectures/machine_learning.xml