

МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«АМУРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

Утверждаю

Ректор

_____ А.Д. Шутенко

(подпись)

« 22 » 06 2024 г.

Согласовано

Проректор по МП

_____ В.В. Ерёмкина

(подпись)

« 22 » 06 2024 г.

Согласовано

Директор ЦЭОиДПО

_____ Н.С. Бодруг

(подпись)

« 22 » 06 2024 г.

**ПРОГРАММА
ПРОФЕССИОНАЛЬНОЙ ПЕРЕПОДГОТОВКИ**

Анализ данных

(наименование программы)

системный аналитик

(наименование присваиваемой квалификации)

Благовещенск, 2024

1. ОБЩАЯ ХАРАКТЕРИСТИКА ПРОГРАММЫ

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»;

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 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 высшего образования по направлениям подготовки: 01.03.02 «Прикладная математика и информатика», 09.03.01 «Информатика и вычислительная техника», 09.03.04 «Программная инженерия», 09.03.02 «Информационные системы и технологии», 10.03.01 «Информационная безопасность», 24.03.01 «Ракетные комплексы и космонавтика», 38.03.01 «Экономика», 38.03.02 «Управление логистикой организации», 38.03.04 «Государственное и муниципальное управление» и специалитета 24.05.01 «Пилотируемые и автоматические космические аппараты и системы» в пределах специальности «Проектирование, производство и эксплуатация ракет и ракетно-космических комплексов», 38.05.01 «Экономическая безопасность», утвержденного приказом Минобрнауки России от РФ от 17 ноября 2020 г. № 1427 (далее вместе – ФГОС ВО)), а также профессионального стандарта 06.022 «Системный аналитик», утвержденного приказом Министерства труда и социальной защиты РФ от 24.11.2014г. No34882.

1.1. Цель реализации программы

Цель: формирование у слушателей профессиональных компетенций, необходимых для приобретения квалификации «системный аналитик».

1.2. Характеристика нового вида профессиональной деятельности, новой квалификации

Область профессиональной деятельности и сфера профессиональной деятельности, в которых слушатель, освоивший программу, осуществляет профессиональную деятельность: 06 Связь, информационные и коммуникационные технологии (в сфере выявления и анализа требований системы с последующим формированием технического задания, планирования, организации и проведения аналитических работ на предприятии).

Объектами профессиональной деятельности являются: информационные и коммуникационные системы и подсистемы, процессы анализа требований к системам, разработка технического задания, разработка методик выполнения аналитических работ, планирование аналитических работ.

В рамках освоения дополнительной квалификации слушатель программы готовится к следующим видам деятельности:

06.019 - Разработка технической документации и методического обеспечения продукции в сфере информационных технологий (далее – ИТ).

06.022 - Проектно-исследовательская деятельность в области информационных технологий.

1.3 Требования к результатам освоения программы

а) Слушатель в результате освоения программы должен обладать следующими профессиональными компетенциями в соответствии с видом деятельности:

ВД 1 - Разработка технической документации и методического обеспечения продукции в сфере информационных технологий:

ПК-1 Выявление и анализ требований системы с последующим формированием технического задания.

ВД 2 - Проектно-исследовательская деятельность в области информационных технологий:

ПК-2 Планирование, организация и проведение аналитических работ на предприятии.

б) Формирование результатов освоения дополнительной профессиональной программы:

В результате освоения программы слушатель должен обладать следующими профессиональными компетенциями, соответствующими видам деятельности:

Формируемые Компетенции	Знания	Умения	Практический опыт
ВД1. Разработка технической документации и методического обеспечения продукции в сфере информационных технологий			
ПК 1. Выявление и анализ требований системы с последующим формированием технического задания	З 1. Нормативной документации по предметной области функции системы	У 1. Выявление требований к системе и подсистеме	ПрО 1. Создание и сопровождение требований и технических заданий на разработку и модернизацию систем и подсистем малого и среднего масштаба и сложности
ВД 2. Проектно-исследовательская деятельность в области информационных технологий			
ПК 2. Планирование, организация и проведение аналитических работ на предприятии	З 2. Методики выполнения аналитических работ	У 2. Организация аналитических работ в ИТ-проекте	ПрО 2. Управление аналитическими работами и подразделением

1.4. Требования к уровню подготовки поступающего на обучение, необходимому для освоения программы

Лица, желающие освоить дополнительную профессиональную программу, должны иметь среднее профессиональное или высшее непрофильное техническое образование, получающие среднее профессиональное и (или) высшее непрофильное техническое

образование. Наличие указанного образования должно подтверждаться документом государственного или установленного образца.

1.5. Трудоемкость обучения

Нормативная трудоемкость обучения по данной программе – 256 часов, включая все виды дистанционной, аудиторной и внеаудиторной (самостоятельной) учебной работы слушателя.

1.6. Форма обучения – очная.

1.7. Режим занятий (график учебных занятий)

Программа реализуется в течение 9-ми месяцев (2 ак. часа в неделю) и предполагает дистанционное обучение с применением электронного обучения и дистанционных образовательных технологий (система дистанционного обучения «Moodle»).

Все виды дистанционной учебной работы (практические занятия, лекции, СРС, текущий контроль, промежуточный контроль) слушателя проводятся с помощью системы дистанционного обучения «Moodle».

2. СОДЕРЖАНИЕ ПРОГРАММЫ

2.1. Учебный план

Наименование дисциплин	Общая трудоемкость в часах	Аудиторные занятия, час.				Дистанционные занятия, час.				СРС (час.)	Текущий контроль (шт.)			Промежуточная аттестация	
		Всего	Из них			Всего	Из них				РК, РГР, Реф.	КР	КП	Зачет	Экзамен
			Лекции	Лабораторные раб.	Практич. занятия		Лекции	Лабораторные раб.	Практич. занятия						
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Модуль 1. Математика для анализа данных															
Линейная алгебра	15						1		4	10				Д	Д
Математический анализ	24						3		5	16				Д	
Приложения линейной алгебры в анализе данных	19						2		4	13				Д	
Теория вероятностей и основы статистики	18						2		4	12				Д	
Статистические методы	20						2		5	13				Д	
Модуль 2. Инструменты анализа данных															
Визуальные инструменты	15						1	4		10				Д	
Основы программирования на Python	30						2	7		21				Д	
Библиотека NumPy	10						1	3		6				Д	
Сбор, организация и	30						3	7		20				Д	

1.5	<i>Основные темы:</i> Статистические методы <i>Краткое содержание:</i> Статистическая оценка параметров: вероятность и правдоподобие; оценка параметров распределения; метод максимального правдоподобия; поиск параметров линейной регрессии с вероятностной точки зрения. Метод главных компонент: матрица ковариации; применение PCA. Статистические эксперименты и проверка гипотез: вероятностное принятие решений; доверительные интервалы; A/B тестирование; параметрические тесты; размер выборки, его связь с ошибкой. Методы статистической проверки гипотез: непараметрические тесты; бутстрэп; нелинейное преобразование данных; множественная проверка гипотез.	23
2	Модуль 2. Инструменты анализа данных	120
2.1	<i>Основные темы:</i> Визуальные инструменты <i>Краткое содержание:</i> Классификация инструментов анализа данных. Обзор инструментов, не требующих навыков программирования. Обзор инструментов анализа данных, в которых применяется программирование. Microsoft Excel. KNIME. Microsoft Power BI. Tableau.	15
2.2	<i>Основные темы:</i> Основы программирования на Python <i>Краткое содержание:</i> Назначение и применение языка Python. Основы работы с интерактивной средой разработки Jupyter. Основные понятия языка. Основные типы данных. Линейные алгоритмы. Условные алгоритмы. Циклические алгоритмы. Функции. Работа с потоковым вводом-выводом.	30
2.3	<i>Основные темы:</i> Библиотека Numpy <i>Краткое содержание:</i> Назначение библиотеки. Основные понятия и особенности. Типы данных Numpy. Основные операции над типами данных Numpy.	10
2.4	<i>Основные темы:</i> Сбор, организация и хранение данных <i>Краткое содержание:</i> Понятие «большие данные». Основные возможности по нахождению и сбору данных. Обзор основных источников данных. Организация и выполнение сбора данных. Основные инструменты сбора и организации данных. Основные виды представления данных. Понятие SQL. Работа с базами данных в SQLiteStudio. Основы SQL. Типы данных SQL. Основные виды запросов SQL. Основы устройства веб-сайтов. Извлечение данных с веб-сайтов.	30
2.5	<i>Основные темы:</i> Работа с данными в Pandas <i>Краткое содержание:</i> Понятие модульности в Python. Краткий обзор основных модулей, применяемых в анализе данных. Назначение и применение библиотеки Pandas. Основные типы Pandas. Извлечение данных из DataFrame. Описательные статистики Pandas. Основные операции, проводимые над типами данных в Pandas. Назначение предобработки данных. Методы предобработки данных. Возможности библиотеки Pandas при построении графиков.	11
2.6	<i>Основные темы:</i> Визуализация данных средствами языка Python <i>Краткое содержание:</i> Визуализация как эффективная форма представления данных. Назначение библиотеки Matplotlib. Интеграция Matplotlib с другими библиотеками. Основные функции библиотеки. Построение графиков по данным. Трактовка графиков. Настройка визуального представления графиков в соответствии с целями и задачами исследования. Назначение библиотеки Seaborn. Особенности и отличия от других модулей построения графиков. Интеграция с другими библиотеками. Основные типы графиков Seaborn. Построение графиков по данным. Настройка представления графиков.	24
2.8	Практика	30
2.9	Итоговая аттестация: выпускная аттестационная работа	10

3. УСЛОВИЯ РЕАЛИЗАЦИИ ПРОГРАММЫ

3.1. Материально-технические условия реализации программы

Наименование	Вид занятий	Наименование оборудования,
--------------	-------------	----------------------------

специализированных аудиторий, кабинетов, лабораторий		программного обеспечения
1	2	3
Компьютерный класс	Лекции	Система дистанционного обучения «Moodle»
Компьютерный класс	Лабораторные и практические занятия, тестирование, итоговая аттестация	Система дистанционного обучения «Moodle» Программное обеспечение: SQLite Studio, Pycharm, Jupyter Notebook, Open Office

3.2. Учебно-методическое обеспечение программы

1. Айзек М.П. Графика, формулы, анализ данных в Excel. Пошаговые примеры / М.П. Айзек. — СПб.: Наука и техника, 2019. — 384 с.
2. Анализ данных Учебник / Мхитарян В.С. - Отв. ред. – М.: Издательство Юрайт, 2017 .– 490 .– ISBN 978-5-534-00616-2: 190.90, 4 .– URL:<http://www.biblio-online.ru/book/CC38E97ACCE5-4470-90F1-3B6D35ACC0B4>.
3. Балдин, К.В. Математический анализ: Учебник / К.В. Балдин, В.Н. Башлыков, А.В. Рукосуев. - М.: Флинта, МПСУ, 2015. - 368 с.
4. Громов, Е. И. Статистические методы прогнозирования : учебное пособие / Е. И. Громов, О. П. Григорьева, Ю. С. Скрипниченко. — Ставрополь : СтГАУ, 2020. — 168 с. — ISBN 978-5-9596-1732-5. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/169742>
5. Документация по работе с реляционной СУБД <https://www.postgresql.org/docs/>
6. Замятин, А. В. Интеллектуальный анализ данных: учебное пособие / А. В. Замятин. - Томск: Издательский Дом Томского государственного университета, 2020. - 196 с. - ISBN 978-5- 94621-898-6. - Текст : электронный // ЭБС «Консультант студента»: [сайт]. - URL: <https://www.studentlibrary.ru/book/ISBN9785946218986.html> ЭБС «Консультант студента»).
7. Кацко, И. А. Теория вероятностей и математическая статистика / И. А. Кацко, П. С. Бондаренко, Г. В. Горелова. — 3-е изд., испр. и доп. — Санкт-Петербург : Лань, 2023. — 436 с. — ISBN 978-5-507-45492-1. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/302663>.
8. Кострикин А.И. Введение в алгебру: В 3-х частях. Часть 1. Основы алгебры: учеб. пособие / А.И. Кострикин. – М.: МЦНМО, 2022. – 272 с.
9. Кулаичев А.П. Методы и средства комплексного анализа данных: Учебное пособие / А.П. Кулаичев. — М.: Форум, 2018. — 160 с.
10. Маккинли У. Python и анализ данных / Уэс Маккинли - Москва: ДМК Пресс, 2015. - 482 с. - ISBN 978-5-97060-315-4. - Текст: электронный // ЭБС «Консультант студента»: [сайт]. - URL: <https://www.studentlibrary.ru/book/ISBN9785970603154.html>
11. Макшанов А.В. Технологии интеллектуального анализа данных. — М.: Лань. 2019. 212 с.
12. Миркин Б.Г. Введение в анализ данных: Учебник и практикум. М.: Издательство Юрайт, 2017.– 174с. Режим доступа:<http://www.biblio-online.ru/book/46A41F93-BC46-401C-A30E27C0FB60B9DE>
13. Мхитарян В. С. Теория планирования эксперимента и анализ статистических данных. — М.: Юрайт. 2020. 491 с.
14. Ниворожкина Л.И. Статистические методы анализа данных: Учебник / Л.И. Ниворожкина, С.В. Арженовский, А.А. Рудяга. — М.: Риор, 2018. — 320 с.
15. Организационное и правовое обеспечение информационной безопасности: учебник и практикум для вузов / Т.А. Полякова, А.А. Стрельцов, С.Г. Чубукова, В.А.

Ниесов; под редакцией Т.А. Поляковой, А.А. Стрельцова. – Москва: Издательство Юрайт, 2020. – 325 с. – (Высшее образование). – ISBN 978-5-534-03600-8. – URL: <https://urait.ru/bcode/450371>.

16. Организационно-правовое обеспечение информационной безопасности: учебное пособие / А.А. Стрельцов, В.С. Горбатов, Т.А.Полякова и др. / Под ред. А.А. Стрельцова. – М.: Издательский центр «Академия», 2008. – 256 с.

17. Рыбина Г. В. Основы построения интеллектуальных систем : учеб. пособ. / Рыбина Г. В. - Москва: Финансы и статистика, 2021. - 432 с. - ISBN 978-5-00184-030-5. - Текст: электронный // ЭБС «Консультант студента»: [сайт]. - URL: <https://www.studentlibrary.ru/book/ISBN9785001840305.html> (ЭБС «Консультант студента»).

18. Сборник задач по теории вероятностей и математической статистике / И. А. Кацко, П. С. Бондаренко, Г. В. Горелова [и др.]. — Санкт-Петербург : Лань, 2023. — 204 с. — ISBN 978-5-507-45493-8. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/302672>

19. Северенс Ч., Введение в программирование на Python, М.: Национальный Открытый Университет «ИНТУИТ», 2016. [Электронный ресурс]. URL: <http://biblioclub.ru/index.php?page=book&id=429184>

20. Теория вероятностей: учеб. для вузов. / А.В. Печинкин, О.И. Тескин, Г.М. Цветкова и др.; Под ред. В.С. Зарубина, А.П. Крищенко. – М.: Изд-во МГТУ им. Н.Э. Баумана. 2004. - 456 с. ISBN 5-7038-2485-0 https://vk.com/wall-186208863_1914

21. Халафян А. А., Боровиков В. П., Калайдина Г. В. Теория вероятностей, математическая статистика и анализ данных. Основы теории и практика на компьютере. Statistica. Excel. Более 150 примеров решения задач. Учебное пособие. — М.: Ленанд. 2017. 320 с.

22. Целых А. Н. Современные методы прикладной информатики в задачах анализа данных: учебное пособие по курсу «Методы интеллектуального анализа данных» / А. Н. Целых, А. А. Целых, Э. М. Котов. - Ростов-на-Дону: ЮФУ, 2021. - 130 с. - ISBN 978-5-9275-3783-9. - Текст: электронный // ЭБС «Консультант студента»: [сайт]. - URL: <https://www.studentlibrary.ru/book/ISBN9785927537839.html> (ЭБС «Консультант студента»).

23. Чашкин, Ю.Р. Математическая статистика. Анализ и обработка данных: Учебное пособие / Ю.Р. Чашкин; Под ред. С.Н. Смоленский. — Рн/Д: Феникс, 2017. — 236 с.

24. Ясницкий Л. Н. Интеллектуальные системы: учебник / Ясницкий Л. Н. - 2-е изд. - Москва: Лаборатория знаний, 2020. - 224 с. Систем. требования: Adobe Reader XI; экран 10». (Учебник для высшей школы) - ISBN 978-5-00101-897-1. - Текст: электронный // ЭБС «Консультант студента»: [сайт]. - URL: <https://www.studentlibrary.ru/book/ISBN9785001018971.html> (ЭБС «Консультант студента»).

Программное обеспечение и Интернет-ресурсы:

№	Наименование ресурса	Краткая характеристика
1.	http://www.edu.ru/	Российское образование. Федеральный портал
2.	ЭБС IPRbooks http://www.iprbookshop.ru	Электронно-библиотечная система IPRbooks — научно-образовательный ресурс для решения задач обучения в России и за рубежом. Контент ЭБС IPRbooks отвечает требованиям стандартов высшей школы, СПО, дополнительного и дистанционного образования. ЭБС IPRbooks в полном объеме соответствует требованиям законодательства РФ в сфере образования
3.	ЭБС Издательства «Лань». http://e.lanbook.com	Ресурс, включающий в себя как электронные версии книг издательства «Лань» и других ведущих издательств учебной литературы, так и электронные версии периодических изданий по естественным, техническим и гуманитарным наукам.

4.	ЭБС ЮРАЙТ https://urait.ru	Фонд электронной библиотеки составляет более 4000 наименований. Учебники и учебные пособия для всех уровней профессионального образования от ведущих научных школ с соблюдением требований новых ФГОСов.
----	--------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

4. ОЦЕНКА КАЧЕСТВА ОСВОЕНИЯ ПРОГРАММЫ

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

- практика – обеспечивает комплексное освоение всех видов профессиональной деятельности системного аналитика;

- итоговая аттестация – завершает изучение всей программы.

В ходе освоения Программы каждый слушатель выполняет следующие отчетные работы:

№ п/п	Наименование раздела (модуля)	Задание	Критерии оценки
1.	Модуль 1. Математика для анализа данных		
1.1	Линейная алгебра	Тест	90-100 баллов – зачтено; 70-89 баллов – зачтено; 50-69 баллов – зачтено; 0-49 баллов – не зачтено.
1.2	Математический анализ	Тест	
1.3	Приложения линейной алгебры в анализе данных	Тест	
1.4	Теория вероятностей и основы статистики	Тест	
1.5	Статистические методы	Тест	
2.	Модуль 2. Инструменты анализа данных		
2.1	Визуальные инструменты	Выполнение практического кейса	90-100 баллов – зачтено; 70-89 баллов – зачтено; 60-69 баллов – зачтено; 0-59 баллов – не зачтено.
2.2	Основы программирования на Python	Выполнение практического кейса	
	Библиотека Numpy	Выполнение практического кейса	
	Сбор, организация и хранение данных	Выполнение практического кейса	
	Работа с данными в Pandas	Выполнение практического кейса	
	Визуализация данных средствами языка	Выполнение практического кейса	

	Python	кейса	
3	Практика	Проектная работа	По результатам публичной защиты проекта.
4	Итоговая аттестация	Выпускная аттестационная работа	<p>Оценка «Отлично» выставляется за ВАР, которая носит исследовательский характер, имеет грамотно изложенную теоретическую главу, глубокий анализ, критический разбор практической деятельности, логичное, последовательное изложение материала с соответствующими выводами и обоснованными предложениями. При ее защите выпускник показывает глубокое знание вопросов темы, свободно оперирует данными исследования, вносит обоснованные предложения, а во время доклада использует раздаточный материал, легко отвечает на поставленные вопросы.</p> <p>Оценка «Хорошо» выставляется за ВАР, которая носит исследовательский характер, имеет грамотно изложенную теоретическую главу, в ней представлены достаточно подробный анализ и критический разбор практической деятельности, последовательное изложение материала с соответствующими выводами, однако с не вполне обоснованными предложениями. При ее защите выпускник показывает знание вопросов темы, оперирует данными исследования, вносит предложения по теме исследования, во время доклада использует раздаточный материал, без особых затруднений отвечает</p>

		<p>на поставленные вопросы.</p> <p>Оценка «Удовлетворительно» выставляется за ВАР, которая носит исследовательский характер, имеет теоретическую главу, базируется на практическом материале, но имеет поверхностный анализ и недостаточно критический разбор, в ней просматривается непоследовательность изложения материала, представлены необоснованные предложения. При ее защите выпускник проявляет неуверенность, показывает слабое знание вопросов темы, не всегда дает исчерпывающие аргументированные ответы на заданные вопросы.</p> <p>Оценка «Неудовлетворительно» выставляется за ВАР, которая не носит исследовательского характера, не имеет анализа, не отвечает требованиям, изложенным в методических указаниях кафедры. В работе нет выводов либо они носят декларативный характер. При защите ВКР выпускник затрудняется отвечать на поставленные вопросы по ее теме, не знает теории вопроса, при ответе допускает существенные ошибки. К защите не подготовлены наглядные пособия и раздаточный материал.</p>
--	--	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Тестирование

Модуль 1. Математика для анализа данных

Тема «Линейная алгебра»

Пример теста

1. Найдите матрицу X, если известно, что:

- $X = \begin{pmatrix} 3 & 2 \\ 4 & 9 \end{pmatrix}$
- $X = \begin{pmatrix} 8 & 5 \\ 3 & 7 \end{pmatrix}$
- $X = \begin{pmatrix} 3 & 6 \\ 4 & 1 \end{pmatrix}$
- $X = \begin{pmatrix} 5 & 1 \\ 3 & 4 \end{pmatrix}$

2. Найдите x, y и z, если известно, что определитель матрицы $\begin{pmatrix} x-1 & 2 & 4 \\ y-2 & 3 & 1 \\ z+6 & -2 & 3 \end{pmatrix}$ равен -

25, определитель матрицы $\begin{pmatrix} -2 & x+1 & 5 \\ 1 & y & 3 \\ 4 & z+6 & 2 \end{pmatrix}$ равен -71, и определитель

матрицы $\begin{pmatrix} 3 & 1 & x \\ -5 & -1 & y-2 \\ 4 & 5 & z+2 \end{pmatrix}$ равен -45:

- $x = 2, y = 1, z = -1$
- $x = -1, y = 2, z = 1$
- $x = 1, y = -1, z = 2$
- $x = 2, y = -1, z = 1$

3. Раскрыть скобки и вычислить вектор $((i+k) \times (i-k)) \times j$:

- i
- 0
- j
- k

Тема «Математический анализ»

Пример теста

1. Найти градиент функции $f(x,y) = \frac{xy}{x^2+y^2+1}$ в точке M(0, 3):

- $\text{grad}f = \{0; -0,3\}$
- $\text{grad}f = \{0,3; 0\}$
- $\text{grad}f = \{-0,3; 0\}$
- $\text{grad}f = \{0; 0,3\}$

2. Заменяя приращение функции дифференциалом, вычислить приближено $1,05^{0,97}$:

- 0,96
- 0,97
- 1
- -0,98

3. Функция $y = f(x)$ имеет первообразную $F(x) = 1/2 * x^2 - \cos x + C$. Зная это, найдите

производную функции $y = f(x)$.

- $2\sin^2(x/2)$
- $2\cos^2(x/2)$
- $1 + 2\cos x$
- $2\sin^2(\pi/4 - x/2)$

Тема «Приложение линейной алгебры в анализе данных»

Пример теста

1. Как изменятся стандартные ошибки коэффициентов линейной регрессии, если значения случайного члена во всех наблюдениях вырастут в n раз (при постоянстве остальных величин):

- не изменятся
- вырастут в n раз
- вырастут в n^2 раз
- уменьшатся в n раз

2. Под полной проблемой собственных значений для матрицы A понимается:

- нахождение всех собственных значений
- нахождение нескольких собственных значений
- нахождение нескольких собственных векторов
- нахождение всех собственных векторов матрицы

Тема «Теория вероятностей и основа статистики»

Пример теста

1. При увеличении объема выборки n и одном и том же уровне значимости α , ширина доверительного интервала:

- может как уменьшиться, так и увеличиться
- уменьшается
- не изменяется
- увеличивается

2. Как называется число m_0 (наступления события в n независимых испытаниях, в каждом из которых вероятность появления события равна p), определяемое из неравенства $np - q \leq m_0 \leq np + r$:

- наибольшее
- оптимальное
- наименее вероятное
- невозможное

3. Если имеется группа из n несовместных событий H_i , образующих полную группу событий, и известны вероятности $P(H_i)$, а событие A может наступить после реализации одного из H_i , известны вероятности $P(A/H_i)$, то $P(A)$ вычисляется по формуле

- Бернулли
- Байеса
- полной вероятности

Тема «Статистические методы»

Пример теста

1. Перечислите основные свойства точечных оценок:

- несмещенность и эффективность
- эффективность и состоятельность
- несмещенность, эффективность и состоятельность
- несмещенность и состоятельность

2. Если эмпирическая статистика K_3 находится в «зоне значимости», то принимается

статистическое решение:

- не отклонять нулевую гипотезу H_0
- отклонить нулевую гипотезу H_0 ;
- не отклонять нулевую гипотезу H_0 , хотя в действительности она будет не верна

3. Согласно F-критерию Фишера, правильным является следующий статистический

вывод:

- статистически значимых различий между дисперсиями в первой и второй группах нет. Различия в степени однородности показателей уровня вербального мышления между группами отсутствуют

- имеется статистическое различие между дисперсиями в первой и второй группах. По степени однородности такого показателя, как уровень вербального мышления, имеются различия между группами

- эмпирическое значение F-критерия Фишера находится в «зоне неопределенности». Необходимо увеличить объемы выборок и продолжить исследование

Модуль 2. Инструменты анализа данных

Тема «Визуальные инструменты»

Пример кейса

Анализ данных на базе Microsoft Excel:

1. Скачайте набор данных «Данные о продуктах супермаркета». Это может быть набор данных из любого открытого источника, например, Kaggle, UCI, Data.gov и т.д.

2. Импортируйте набор данных в Microsoft Excel.

3. Оцените качество и подготовку данных. Определите, есть ли в данных отсутствующие значения, выбросы или дубликаты. Исправьте любые ошибки или недостатки в данных по необходимости.

4. Создайте сводную таблицу, чтобы посмотреть среднее, медиану и стандартное отклонение для каждого продукта по количеству проданных единиц и общей выручке.

5. Создайте диаграммы для визуализации данных. Например, можно построить гистограммы для поиска распределения для каждого продукта в категориях проданных единиц и общей выручки.

6. Проверьте гипотезы относительно данных, используя функцию анализа данных в Microsoft Excel. Например, можно проверить, зависят ли проданные единицы от типа продукта или от сезона (например, лето, зима, весна, осень).

7. Представьте результаты вашего анализа с помощью графиков и диаграмм. Составьте отчет о результате работы в Microsoft Word или PowerPoint.

8. Сделайте выводы и рекомендации, основанные на анализе. Например, можно рекомендовать увеличить производство и продажу тех продуктов, которые имеют большой спрос.

Тема «Основы программирования на Python»

Пример кейса

Разработка приложения «планировщик задач» на языке Python

1. Определите подходящую версию языка Python, обосновав сделанный выбор.

2. Самостоятельно осуществите выбор подходящей интегрированной среды разработки (IDE) с обоснованием выбора.

3. Разработайте техническое задание (ТЗ) к приложению. Проведите проектирование приложения с учетом функционала, который должна включать в себя реализуемая программа, а именно:

- создание задач со сроками выполнения;
- вывод списка всех созданных задач;
- редактирование и удаление задач;
- фильтрация задач по названию или срокам выполнения;
- уведомления о крайних сроках выполнения задач.

Учтите, что между запусками программа должна хранить свои данные на накопителе данных. Используйте для этого любой подходящий формат хранения данных на ваш выбор.

4. Разработайте прототип приложения на основе созданного проекта.

5. Проанализируйте работу прототипа. ЗадOCUMENTИРУЙТЕ все найденные в программе ошибки и неточности. Оцените удобство пользования созданной программой, а также эффективность ее работы с точки зрения скорости достижения результата. Например, за какое минимально возможное количество действий можно создать задачу в программе.

6. На основе выявленных недостатков исправьте и улучшите созданную программу. По необходимости проведите повторное проектирование программы и создайте новый прототип.

7. Подготовьте отчет, включающий в себя экранные формы работы приложения. Сделайте выводы об особенностях разработки приложений типа «планировщик задач», с которыми вам пришлось столкнуться по ходу разработки. Отметьте наиболее удачные решения, которые вам удалось применить при разработке программы. Опишите возникшие в ходе разработки проблемы. Расскажите, каким функционалом можно дополнить вашу текущую версию программы.

Тема «Библиотека NumPy»

Пример кейса

Анализ данных в NumPy на основе набора данных «Данные о футбольных матчах»

1. Скачайте набор данных «Данные о футбольных матчах». Это может быть набор данных из любого открытого источника, например, Kaggle, UCI, Data.gov и т.д.

2. Импортируйте набор данных в Python с помощью библиотеки NumPy.

3. Оцените качество и подготовку данных. Определите, есть ли в данных отсутствующие значения, выбросы или дубликаты. Исправьте любые ошибки или недостатки в данных по необходимости.

4. Проведите статистический анализ данных. Включите в ваши исследования следующие показатели и таблицы:

- расчет среднего количества голов, забитых каждой из команд;
- сводная таблица с результатами каждой из команд в виде «победы / ничьи / поражения»;
- список из 5 лучших команд по результатам проведенных матчей.

В зависимости от имеющихся данных дополните или измените вышеприведенные пункты. Например, можно провести расчет количества полученных красных карточек каждой командой за определенный промежуток времени, или посчитать среднее количество получаемых травм, характерное для конкретной команды на протяжении нескольких игровых сезонов.

5. Подготовьте и проверьте несколько гипотез относительно имеющихся данных. Например, докажите или опровергните гипотезу о том, что спортсмены имеют лучшие показатели в первой половине матча, чем во второй.

6. Представьте полученные результаты в виде отчета. Отчет можно создать в программах Microsoft Word и Microsoft PowerPoint. По желанию полученные с помощью NumPy сводные таблицы и показатели можно импортировать в Microsoft Excel, чтобы создать диаграммы и графики для визуализации полученных данных.

7. Дополните отчет выводами и прогнозами касательно результатов команд или их участников. Например, сделайте прогноз на результативность команды в следующем игровом сезоне, учитывая предыдущие показатели команды, состояние спортсменов и другие показатели.

Тема «Сбор, организация и хранение данных»

Пример кейса

Создание набора данных на основе интернет-каталога книг

1. Найдите на просторах сети Интернет каталоги книг. Обычно их можно найти на сайтах, занимающихся дистрибуцией книжной продукции.

2. Внимательно изучите содержимое каталогов, а также данные, которыми описывают книги в каталоге.

3. Отберите каталог, наиболее близкий по содержанию к следующим характеристикам:

- у книг есть заголовок и описание;
- указано авторство;
- отмечено количество страниц;
- присутствует год издания;
- указано возрастное ограничение.

4. Продумайте формат хранения будущих данных. Учтите, что вам понадобится информация, перечисленная в пунктах выше. Помимо этой информации необходимо подумать, какая еще информация может понадобиться для выполнения следующих задач:

- отбор наиболее популярных книг с группировкой по жанрам;
- вывод 5 самых обсуждаемых книг;
- отображение самой дешевой и самой дорогой книг в каталоге.

Информацию для решения обозначенных задач также необходимо включить в набор данных.

5. Проведите сбор данных с помощью средств языка Python. Например, посредством библиотек Beautiful Soup и Scrapy.

6. Сохраните данные в выбранном формате хранения данных. Проверьте их целостность и достоверность. В случае ошибок установите источник искажений. Например, ошибка в программе или разрыв интернет-соединения.

7. Подготовьте отчет, включающий в себя выдержки из собранных данных и описание к ним. Объясните, как собранные данные могут помочь в решении задач из пункта 4 настоящего кейса, а также опишите, какие еще задачи можно решить с помощью собранных данных.

Тема «Работа с данными в Pandas»

Пример кейса

Исследование показателей продажи видеоигр с помощью Pandas

1. Загрузите набор данных о продажах видеоигр.

2. Откройте его в Python с помощью библиотеки Pandas.

3. Используйте основные методы библиотеки Pandas, такие как head(), info() и describe(), чтобы ознакомиться с данными. Обратите внимание на получаемые статистические показатели.

4. Ознакомившись с данными, проверьте набор данных на ошибки, пропуски, прочие искажения. По возможности исправьте их.

5. Проведите расчет и вывод следующих показателей:

- среднее количество проданных копий видеоигр для каждого издателя;
- наиболее «успешный» издатель видеоигр по показателю количества проданных игр по странам;

- издатель с наибольшим количеством проданных копий видеоигр;

- игровая платформа с наибольшим количеством реализованных копий видеоигр.

Расчет дополнительных показателей приветствуется.

6. Загрузите данные в Microsoft Excel и постройте графики. Например, круговая диаграмма по данным о среднем количестве проданных видеоигр каждым издателем.

7. Включите полученные графики в отчет по выполненной работе. Сделайте соответствующие выводы. Сопроводите все материалы в отчете комментариями. Дополнительно отметьте сильные и слабые стороны Pandas при работе с табличными данными относительно библиотеки NumPy.

Тема «Визуализация данных средствами языка Python»

Пример кейса

Визуализация набора данных «Бриллианты»

1. Загрузите данные на компьютер. Откройте набор данных в Python. Например, с помощью средств библиотек Matplotlib или Seaborn.

2. Внимательно ознакомьтесь с содержимым наборов данных. При необходимости воспользуйтесь уже изученными библиотеками, например, Pandas.

3. Убедитесь в целостности и корректности данных. По необходимости исправьте и очистите данные.

4. Проведите построение различных графиков на основе имеющихся данных:

- scatter plot;
- box plot;
- bar plot;
- pair plot;
- line plot;
- heat map;
- violin plot.

По желанию постройте другие типы графиков. Данные для графиков отберите самостоятельно, руководствуясь целесообразностью. Например, имеет смысл проверить гипотезу о зависимости цены бриллианта от количества карат, но не очень разумно проверять зависимость цвета от размера.

5. Обобщите полученные сведения, включив графики в отчет по проделанной работе. Сопроводите каждый график комментариями.

6. Сделайте выводы о полученных результатах. Подумайте, как можно применить полученные сведения на практике. Например, автоматизировать процесс назначения предварительной цены каждому бриллианту в зависимости от его характеристик, что может быть актуально в системах с применением компьютерного зрения.

7. Представьте результаты работы в удобном для вас виде. Предпочтительно в формате презентации Microsoft PowerPoint или файла pdf.

Практика

Цель производственной практики: разработать модель анализа данных клиентов и предоставить рекомендации по улучшению качества сервиса компании.

Задачи производственной практики:

1. Сбор данных о клиентах компании из различных источников (CRM, сайт, социальные сети и т.д.).

2. Оценка качества данных и их предобработка (удаление дубликатов, заполнение пропусков и т.д.).

3. Проведение исследовательского анализа данных, включающего:

- Визуальный анализ данных (графики, диаграммы, распределения и т.д.).
- Оценку корреляции между признаками.
- Определение наиболее значимых признаков для моделирования.

4. Выбор модели машинного обучения для предсказания уровня удовлетворенности клиентов.

5. Обучение выбранной модели на обучающих данных и оценка ее качества на тестовых данных.

6. Проведение анализа важности признаков для моделирования и составление рекомендаций для улучшения качества сервиса компании.

Ожидаемые результаты производственной практики:

1. Отчет о сборе и предобработке данных.
2. Отчет об исследовательском анализе данных.
3. Отчет о выбранной модели машинного обучения и результаты ее оценки.
4. Отчет о важности признаков для моделирования и рекомендации по улучшению качества сервиса компании.

Итоговая аттестация. Перечень примерных заданий

Общими требованиями к ВАР являются:

- целевая направленность;
- логическая последовательность изложения материала;
- глубина исследования и полнота освещения вопросов;
- убедительность аргументаций;
- краткость и четкость формулировок;
- конкретность изложения результатов работы;
- доказательность выводов и обоснованность рекомендаций;
- грамотное оформление.

Структура пояснительной записки и графической части ВАР

Структура пояснительной записки:

- титульный лист;
- реферат;
- список использованных сокращений, условных обозначений символов, единиц и терминов;
- содержание;
- введение;
- основной раздел, в котором приводятся материалы по исследованию предметной области и самого предмета, по анализу вариантов решения поставленной задачи и выбору конкретного варианта;
- заключение;
- библиографический список использованных источников;
- приложения.

Объем пояснительной записки должен быть не более 50 страниц (без приложений); количество использованных источников – не менее 25.

Графические материалы должны содержать информацию, позволяющую оценить:

- а) постановку и формализацию задачи, используемые математические методы;
- б) общий принцип функционирования аппаратного либо программного продукта;
- в) конкретные проектные решения, выполненные студентом;
- г) результаты экспериментов (если предусмотрено их проведение);
- д) вопросы технологии и эксплуатации.

Примерные темы ВКР

1. "Анализ и оптимизация процессов целевого взаимодействия с клиентами на основе данных CRM-системы"
2. "Разработка алгоритма автоматической классификации клиентов для повышения эффективности маркетинга"
3. "Прогнозирование изменения уровня удовлетворенности клиентов в зависимости от различных факторов на примере компании X"
4. "Анализ пользовательского поведения в мобильном приложении для улучшения функциональности и удобства использования"
5. "Системный анализ эффективности работы интернет-магазина на основе данных о покупках и поведении клиентов"
6. "Разработка методики оценки качества продукта на основе анализа обратной связи и отзывов клиентов"
7. "Анализ и оптимизация процессов логистики на основе данных о поставках и времени доставки"
8. "Исследование влияния изменений цен на продукцию на динамику продаж с использованием методов машинного обучения"
9. "Применение системного анализа для определения оптимальных параметров производства на основе спроса и прогнозов рынка"
10. "Анализ эффективности программы лояльности клиентов на основе данных о частоте и объеме покупок".

При условии успешного прохождения установленного вида аттестационного испытания, слушателю присваивается соответствующая квалификация «Системный аналитик» и выдается документ установленного образца о профессиональной переподготовке.

5. СОСТАВИТЕЛИ ПРОГРАММЫ

Ерёмина В.В., доцент, к.ф.-м.н., проректор по МП

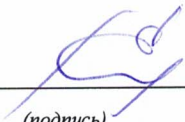
Павельчук А.В., доцент, к.ф.-м.н., директор ЦРСКД «АмурТехноЦентр» (ДНК им. М.Т. Луценко)

Мишаченко К.Г., старший преподаватель ЦК, преподаватель ЦРСКД «АмурТехноЦентр» (ДНК им. М.Т. Луценко)

Ким А.В., старший преподаватель ЦК, преподаватель ЦРСКД «АмурТехноЦентр» (ДНК им. М.Т. Луценко)

Руководитель УКЦ

«Современный университет»



(подпись)

В.В. Ерёмина