

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
федеральное государственное бюджетное образовательное учреждение высшего
образования «Вятский государственный университет»
(ВятГУ)
г. Киров

Утверждаю
Директор/Декан Репкин Д. А.



Номер регистрации
РПД_3-10.05.02.01_2021_122310
Актуализировано: 11.05.2021

Рабочая программа дисциплины
Основы анализа данных

	наименование дисциплины
Квалификация выпускника	Специалист по защите информации
Специальность	10.05.02
	шифр
	Информационная безопасность телекоммуникационных систем
	наименование
Специализация	Системы подвижной цифровой защищенной связи
	наименование
Формы обучения	Очная
	наименование
Кафедра-разработчик	Кафедра радиоэлектронных средств
	наименование
Выпускающая кафедра	Кафедра радиоэлектронных средств
	наименование

Киров, 2021 г.

Сведения о разработчиках рабочей программы дисциплины

Прозоров Дмитрий Евгеньевич

ФИО

Метелев Александр Петрович

ФИО

Цели и задачи дисциплины

Цель дисциплины	Цель дисциплины – сформировать у студентов системное представление о технологиях анализа данных, интеллектуального анализа данных (Data Mining), их применении и инструментах, изучить основные методы прикладного анализа данных, развить навыки практического применения методов многомерного анализа и Data Mining для решения различных научных и технических задач в инфокоммуникационных системах.
Задачи дисциплины	<p>Задачи дисциплины:</p> <ul style="list-style-type: none"> - формирование представлений о целях, способах реализации и инструментах анализа данных; - изучение сфер применения, методов и средств Data Mining; - формирование практических навыков анализа данных; - получение теоретических знаний и практических навыков при решении типовых задач анализа данных; - рассмотрение структуры и функций хранилищ данных.

Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы

Компетенция ОПК-2

Способен применять информационно-коммуникационные технологии, программные средства системного и прикладного назначений, в том числе отечественного производства, для решения задач профессиональной деятельности		
Знает	Умеет	Владеет
базовые возможности языка Python; основные понятия машинного обучения и анализа данных; основные типы задач машинного обучения; математический аппарат, применяемый для анализа данных; типы признаков данных; возможные приложения машинного обучения	выполнять сбор данных для последующей обработки; осуществлять предобработку данных; проводить визуализацию в задачах машинного обучения; использовать базовые алгоритмы в задачах классификации; применять основные методы в задачах кластеризации; строить композиции алгоритмов машинного обучения	навыками работы в одной из сред разработки языка программирования Python; способностью интерпретировать полученные результаты в задачах машинного обучения; представлением о метриках качества в задачах анализа данных; навыками кросс-валидации обучающей выборки; способностью выполнять регуляризацию линейных моделей; навыками масштабирования информативных признаков

Структура дисциплины
Тематический план

№ п/п	Наименование разделов дисциплины	Шифр формируемых компетенций
1	Основные задачи анализа данных	ОПК-2
2	Регрессия и классификация	ОПК-2
3	Кластеризация	ОПК-2
4	Подготовка и прохождение промежуточной аттестации	ОПК-2

Формы промежуточной аттестации

Зачет	9 семестр (Очная форма обучения)
Экзамен	Не предусмотрен (Очная форма обучения)
Курсовая работа	Не предусмотрена (Очная форма обучения)
Курсовой проект	Не предусмотрена (Очная форма обучения)

Трудоемкость дисциплины

Форма обучения	Курсы	Семестры	Общий объем (трудоемкость)		Контактная работа, час	в том числе аудиторная контактная работа обучающихся с преподавателем, час				Самостоятельная работа, час	Курсовая работа (проект), семестр	Зачет, семестр	Экзамен, семестр
			Часов	ЗЕТ		Всего	Лекции	Семинарские, практические занятия	Лабораторные занятия				
Очная форма обучения	5	9	144	4	92.5	54	18	0	36	51.5		9	

Содержание дисциплины

Очная форма обучения

Код занятия	Наименование тем занятий	Трудоемкость, академических часов
Раздел 1 «Основные задачи анализа данных»		34.00
Лекции		
Л1.1	Типы задачи анализа данных	4.00
Лабораторные занятия		
Р1.1	Основы Python	8.00
Самостоятельная работа		
С1.1	Подготовка к лекции Л1.1	10.00
Контактная внеаудиторная работа		
КВР1.1	Контактная внеаудиторная работа	12.00
Раздел 2 «Регрессия и классификация»		82.00
Лекции		
Л2.1	Обучение с учителем. Регрессия и классификация.	4.00
Л2.2	Градиентный спуск	2.00
Л2.3	Метод опорных векторов	2.00
Л2.4	Метрики качества	2.00
Лабораторные занятия		
Р2.1	Python: работа с библиотеками numpy, pandas, matplotlib	8.00
Р2.2	Python: предобработка данных	12.00
Р2.3	Python: решающие деревья	8.00
Самостоятельная работа		
С2.1	Подготовка к лекциям Л2.1-Л2.4	26.00
Контактная внеаудиторная работа		
КВР2.1	Контактная внеаудиторная работа	18.00
Раздел 3 «Кластеризация»		24.00
Лекции		
Л3.1	Обучение без учителя. Кластеризация.	4.00
Самостоятельная работа		
С3.1	Подготовка к лекции Л3.1	12.00
Контактная внеаудиторная работа		
КВР3.1	Контактная внеаудиторная работа	8.00
Раздел 4 «Подготовка и прохождение промежуточной аттестации»		4.00
З4.1	Подготовка к сдаче зачета	3.50
КВР4.1	Сдача зачета	0.50
ИТОГО		144.00

Содержание дисциплины данной рабочей программы используется при обучении по индивидуальному учебному плану, при ускоренном обучении, при применении дистанционных образовательных технологий и электронном обучении (при наличии).

Методические указания для обучающихся по освоению дисциплины

Успешное освоение дисциплины предполагает активное, творческое участие обучающегося на всех этапах ее освоения путем планомерной, повседневной работы. Обучающийся обязан посещать лекции, семинарские, практические и лабораторные занятия (при их наличии), получать консультации преподавателя и выполнять самостоятельную работу.

Изучение дисциплины следует начинать с проработки настоящей рабочей программы, методических указаний и разработок, указанных в программе, особое внимание уделить целям, задачам, структуре и содержанию дисциплины.

Главной задачей каждой лекции является раскрытие сущности темы и анализ ее основных положений. Тематика лекций определяется настоящей рабочей программой дисциплины.

Лекции – это систематическое устное изложение учебного материала. На них обучающийся получает основной объем информации по каждой конкретной теме. Лекции обычно носят проблемный характер и нацелены на освещение наиболее трудных и дискуссионных вопросов.

Предполагается, что обучающиеся приходят на лекции, предварительно проработав соответствующий учебный материал по источникам, рекомендованным программой. Часто обучающимся трудно разобраться с дискуссионными вопросами, дать однозначный ответ. Преподаватель, сравнивая различные точки зрения, излагает свой взгляд и нацеливает их на дальнейшие исследования и поиск научных решений. После лекции желательно вечером перечитать и закрепить полученную информацию, тогда эффективность ее усвоения значительно возрастает. При работе с конспектом лекции необходимо отметить материал, который вызывает затруднения для понимания, попытаться найти ответы на затруднительные вопросы, используя предлагаемую литературу. Если самостоятельно не удалось разобраться в материале, сформулируйте вопросы и обратитесь за помощью к преподавателю.

Целью семинарских занятий является проверка уровня понимания обучающимися вопросов, рассмотренных на лекциях и в учебной литературе.

Целью практических и лабораторных занятий является формирование у обучающихся умений и навыков применения теоретических знаний в реальной практике решения задач; восполнение пробелов в пройденной теоретической части курса.

Семинарские, практические и лабораторные занятия в равной мере направлены на совершенствование индивидуальных навыков решения теоретических и прикладных задач, выработку навыков интеллектуальной работы, а также ведения дискуссий. Для успешного участия в семинарских, практических и лабораторных занятиях обучающемуся следует тщательно подготовиться.

Основной формой подготовки обучающихся к практическим (лабораторным) занятиям является самостоятельная работа с учебно-методическими материалами, научной литературой, статистическими данными и т.п.

Изучив конкретную тему, обучающийся может определить, насколько хорошо он в ней разобрался. Если какие-то моменты остались непонятными, целесообразно составить список вопросов и на занятии задать их преподавателю. Практические (лабораторные) занятия предоставляют обучающемуся возможность творчески раскрыться, проявить инициативу и развить навыки публичного ведения дискуссий и общения.

Самостоятельная работа обучающихся включает в себя выполнение различного рода заданий (изучение учебной и научной литературы, материалов лекций, систематизацию прочитанного материала, подготовку контрольной работы, решение

задач, подготовка докладов, написание рефератов, публикация тезисов, научных статей, подготовка и защита курсовой работы / проекта и другие), которые ориентированы на глубокое усвоение материала изучаемой дисциплины.

Обучающимся рекомендуется систематически отводить время для повторения пройденного материала, проверяя свои знания, умения и навыки.

Внутренняя система оценки качества освоения дисциплины включает входной контроль уровня подготовленности обучающихся, текущий контроль успеваемости, промежуточную аттестацию, направленную на оценивание промежуточных и окончательных результатов обучения по дисциплине (в том числе результатов курсового проектирования (выполнения курсовых работ) при наличии).

При проведении промежуточной аттестации обучающегося учитываются результаты текущего контроля, проводимого в течение освоения дисциплины.

Процедура оценивания результатов освоения дисциплины осуществляется на основе действующих локальных нормативных актов ФГБОУ ВО «Вятский государственный университет», с которыми обучающиеся знакомятся на официальном сайте университета www.vyatsu.ru.

Учебно-методическое обеспечение дисциплины, в том числе учебно-методическое обеспечение самостоятельной работы обучающегося по дисциплине

Учебная литература (основная)

1) Сидняев, Николай Иванович. Теория планирования эксперимента и анализ статистических данных : Учебник и практикум для вузов / Н. И. Сидняев. - 2-е изд., пер. и доп. - Москва : Юрайт, 2020. - 495 с. - (Высшее образование). - ISBN 978-5-534-05070-7 : 909.00 р. - URL: <https://urait.ru/bcode/449686> (дата обращения: 08.05.2020). - Режим доступа: Образовательная платформа Юрайт. - Текст : электронный.

2) Осипенко, С. А. Статистические методы обработки и планирования эксперимента : учебное пособие / С.А. Осипенко. - Москва|Берлин : Директ-Медиа, 2020. - 62 с. : ил., табл. - Библиогр.: с. 54. - ISBN 978-5-4499-1629-7 : Б. ц. - URL: <http://biblioclub.ru/index.php?page=book&id=598682/> (дата обращения: 24.03.2020). - Режим доступа: ЭБС Университетская библиотека ONLINE. - Текст : электронный.

3) Крутиков, В. Н. Методы оптимизации : учебное пособие / В.Н. Крутиков, В.В. Мешечкин. - 2-е изд., исправ. и доп. - Кемерово : Кемеровский государственный университет, 2019. - 106 с. : ил. - Библиогр.: с. 101 - 102. - ISBN 978-5-8353-2437-8 : Б. ц. - URL: <http://biblioclub.ru/index.php?page=book&id=600281/> (дата обращения: 24.03.2020). - Режим доступа: ЭБС Университетская библиотека ONLINE. - Текст : электронный.

4) Шкаберина, Г. Ш. Программирование. Основы языка Python : учебное пособие / Г. Ш. Шкаберина, Н. Л. Резова. - Красноярск : СибГУ им. академика М. Ф. Решетнёва, 2018. - 92 с. - Б. ц. - URL: <https://e.lanbook.com/book/147450> (дата обращения: 15.05.2020). - Режим доступа: ЭБС Лань. - Текст : электронный.

5) Шелудько, В. М. Язык программирования высокого уровня Python: функции, структуры данных, дополнительные модули : учебное пособие / В.М. Шелудько. - Ростов-на-Дону|Таганрог : Южный федеральный университет, 2017. - 108 с. : ил. - Библиогр. в кн. - ISBN 978-5-9275-2648-2 : Б. ц. - URL: <http://biblioclub.ru/index.php?page=book&id=500060/> (дата обращения: 03.03.2021). - Режим доступа: ЭБС Университетская библиотека ONLINE. - Текст : электронный.

Учебная литература (дополнительная)

1) Кибзун, Андрей Иванович. Теория вероятностей и математическая статистика : базовый курс с примерами и задачами : [учебник для вузов по техническим и экономическим специальностям] / А. И. Кибзун, Е. Р. Горяинова, А. В. Наумов ; ред. А. И. Кибзун. - 4-е изд., перераб. и доп. - Москва : Физматлит, 2014. - 231 с. : ил. - Библиогр.: с. 228. - Предм. указ.: с. 229-231. - ISBN 978-5-9221-1484-4 : 615.00 р. - Текст : непосредственный.

2) Гмурман, Владимир Ефимович. Руководство к решению задач по теории вероятностей и математической статистике : учеб. пособие для бакалавриата и специалитета / В. Е. Гмурман. - 11-е изд. - Москва : Юрайт, 2019. - 406 с. - (Бакалавр. Специалист). - ISBN 978-5-534-08389-7 : 849.00 р. - Текст : непосредственный.

3) Глебов, В. И. Практикум по математической статистике: проверка гипотез с использованием Excel, MatCalc, R и Python : учебное пособие / В.И. Глебов, С.Я. Криволапов. - Москва : Прометей, 2019. - 87 с. : ил. - Библиогр. в кн. - ISBN 978-5-907100-66-4 : Б. ц. - URL: <http://biblioclub.ru/index.php?page=book&id=576035/> (дата обращения: 03.03.2021). - Режим доступа: ЭБС Университетская библиотека ONLINE. - Текст : электронный.

4) Программные системы статистического анализа: обнаружение закономерностей в данных с использованием системы R и языка Python : учебное пособие / В.М. Волкова, М.А. Семенова, Е.С. Четвертакова, С.С. Вожов. - Новосибирск : Новосибирский государственный технический университет, 2017. - 74 с. : ил., табл. - Библиогр.: с. 48. - ISBN 978-5-7782-3183-2 : Б. ц. - URL: <http://biblioclub.ru/index.php?page=book&id=576496/> (дата обращения: 03.03.2021). - Режим доступа: ЭБС Университетская библиотека ONLINE. - Текст : электронный.

Учебно-методические издания

1) Математика. Основы теории вероятностей : методические указания для практических занятий. - Новочеркасск : Новочерк. инж.-мелиор. ин-т Донской ГАУ, 2020. - 58 с. - Б. ц. - URL: <https://e.lanbook.com/book/148541> (дата обращения: 15.05.2020). - Режим доступа: ЭБС Лань. - Текст : электронный.

2) Технологии анализа данных: Data Mining, Visual Mining, Text Mining, OLAP : учеб. пособие / А. А. Барсегян, М. С. Куприянов, В. В. Степаненко, И. И. Холод. - 2-е изд. - СПб. : БХВ-Петербург, 2008. - 375 с. + 1 эл. опт. диск (CD-ROM). - Библиогр.: с.368-372. - Предм. указ.: с. 372-375. - ISBN 978-5-94157-991-4 : 368.00 р. - Текст : непосредственный.

3) Петрунин, Юрий Юрьевич. Информационные технологии анализа данных. Data analysis : учеб. пособие по дисциплине "Информатика" для студентов высших учебных заведений, обучающихся по управленческим и экономическим специальностям и направлениям / Ю. Ю. Петрунин ; МГУ. - М. : Ун-т книжный дом (КДУ), 2008. - 291 с. : ил., табл. ; 20 см. - Библиогр.: с. 290-291. - ISBN 978-5-98227-416-8 : 264.00 р. - Текст : непосредственный.

Учебно-наглядное пособие

1) Хранилища данных. Лекция 1. Понятия о хранилищах. Презентация. - Москва : Национальный Открытый Университет «ИНТУИТ», 2014. - 46 с. - Б. ц. - URL: <http://biblioclub.ru/index.php?page=book&id=237105/> (дата обращения: 24.03.2020). - Режим доступа: ЭБС Университетская библиотека ONLINE. - Текст : электронный.

3) Хранилища данных. Лекция 4. Создание многомерного хранилища данных на основе MS SQL Server 2005. Презентация. - Москва : Национальный Открытый Университет «ИНТУИТ», 2014. - 67 с. - Б. ц. - URL: <http://biblioclub.ru/index.php?page=book&id=237114/> (дата обращения: 24.03.2020). - Режим доступа: ЭБС Университетская библиотека ONLINE. - Текст : электронный.

4) Хранилища данных. Лекция 6. Работа с OLAP срезами. Презентация. - Москва : Национальный Открытый Университет «ИНТУИТ», 2014. - 103 с. - Б. ц. - URL: <http://biblioclub.ru/index.php?page=book&id=237115/> (дата обращения: 24.03.2020). - Режим доступа: ЭБС Университетская библиотека ONLINE. - Текст : электронный.

5) Хранилища данных. Лекция 7. SQL Server – ProClarity. Презентация. - Москва : Национальный Открытый Университет «ИНТУИТ», 2014. - 75 с. - Б. ц. - URL: <http://biblioclub.ru/index.php?page=book&id=237116/> (дата обращения: 24.03.2020). - Режим доступа: ЭБС Университетская библиотека ONLINE. - Текст : электронный.

6) Хранилища данных. Лекц. 8. - Москва : Национальный Открытый Университет «ИНТУИТ», 2014. - 35 с. - Б. ц. - URL: <http://biblioclub.ru/index.php?page=book&id=237117/> (дата обращения: 24.03.2020). - Режим доступа: ЭБС Университетская библиотека ONLINE. - Текст : электронный.

2) Хранилища данных. Лекция 3. Создание куба в SQL Server 2005. Презентация. - Москва : Национальный Открытый Университет «ИНТУИТ», 2014. - 79 с. - Б. ц. - URL: <http://biblioclub.ru/index.php?page=book&id=237113/> (дата обращения: 24.03.2020). - Режим доступа: ЭБС Университетская библиотека ONLINE. - Текст : электронный.

Электронные образовательные ресурсы

1) Портал дистанционного обучения ВятГУ [электронный ресурс] / - Режим доступа: <http://mooc.do-kirov.ru/>

2) Раздел официального сайта ВятГУ, содержащий описание образовательной программы [электронный ресурс] / - Режим доступа: https://www.vyatsu.ru/php/programms/eduPrograms.php?Program_ID=3-10.05.02.01

3) Личный кабинет студента на официальном сайте ВятГУ [электронный ресурс] / - Режим доступа: <https://new.vyatsu.ru/account/>

4) Единое окно доступа к образовательным ресурсам <http://window.edu.ru/>

Электронные библиотечные системы (ЭБС)

• ЭБС «Научная электронная библиотека eLIBRARY» (<http://elibrary.ru/defaultx.asp>)

- ЭБС «Издательства Лань» (<http://e.lanbook.com/>)
- ЭБС «Университетская библиотека online» (www.biblioclub.ru)
- Внутренняя электронно-библиотечная система ВятГУ (<http://lib.vyatsu.ru/>)
- ЭБС «ЮРАЙТ» (<https://urait.ru>)

Современные профессиональные базы данных и информационные справочные системы

- ГАРАНТ
- КонсультантПлюс
- Техэксперт: Нормы, правила, стандарты
- Роспатент (<https://www1.fips.ru/elektronnye-servisy/informatsionno-poiskovaya-sistema>)
- Web of Science® (<http://webofscience.com>)

Материально-техническое обеспечение дисциплины

Демонстрационное оборудование

Перечень используемого оборудования
МУЛЬТИМЕДИА ПРОЕКТОР CASIO XJ-A141V С ЭКРАНОМ НАСТЕННЫМ 180*180СМ, ШТАТИВОМ PROFFIX 63-100СМ И КАБЕЛЕМ VGA 15.2М

Специализированное оборудование

Перечень используемого оборудования
ПЕРСОНАЛЬНЫЙ КОМПЬЮТЕР ICL S253.MI (МОНОБЛОК)
ПЕРСОНАЛЬНЫЙ КОМПЬЮТЕР ICL SafeRay S251.Mi (МОНОБЛОК)

Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, в том числе лицензионное и свободно распространяемое ПО (включая ПО отечественного производства)

№ п.п	Наименование ПО	Краткая характеристика назначения ПО
1	Программная система с модулями для обнаружения текстовых заимствований в учебных и научных работах «Антиплагиат.ВУЗ»	Программный комплекс для проверки текстов на предмет заимствования из Интернет-источников, в коллекции диссертация и авторефератов Российской государственной библиотеки (РГБ) и коллекции нормативно-правовой документации LEXPRO
2	Microsoft Office 365 ProPlusEdu ALNG SubsVL MVL AddOn toOPP	Набор веб-сервисов, предоставляющий доступ к различным программам и услугам на основе платформы Microsoft Office, электронной почте бизнес-класса, функционалу для общения и управления документами
3	Office Professional Plus 2016	Пакет приложений для работы с различными типами документов: текстами, электронными таблицами, базами данных, презентациями
4	Windows Professional	Операционная система
5	Kaspersky Endpoint Security для бизнеса	Антивирусное программное обеспечение
6	Справочная правовая система «Консультант Плюс»	Справочно-правовая система по законодательству Российской Федерации
7	Электронный периодический справочник ГАРАНТ Аналитик	Справочно-правовая система по законодательству Российской Федерации
8	Security Essentials (Защитник Windows)	Защита в режиме реального времени от шпионского программного обеспечения, вирусов.
9	МойОфис Стандартный	Набор приложений для работы с документами, почтой, календарями и контактами на компьютерах и веб браузерах
10	2012 Внеб. Бессроч. лиценз. MATLAB Academic new Product From 25 fo 24 Concurrent Licenses	Специализированное лицензионное ПО
11	Python	Язык программирования
12	Anaconda	дистрибутив языков программирования Python и R с набором приложений. По умолчанию в Anaconda Navigator доступны следующие приложения: JupyterLab Jupyter Notebook QtConsole Spyder Glue Orange RStudio Visual Studio Code

Обновленный список программного обеспечения данной рабочей программы находится по адресу:
https://www.vyatsu.ru/php/list_it/index.php?op_id=122310