

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Вятский государственный университет»

УТВЕРЖДАЮ
Начальник Управления
дополнительного образования
Курагина Курагина К.А.

«29» июня 2022 г.
03-04-2022-0580-1009

РАБОЧАЯ ПРОГРАММА

учебной дисциплины
«Анализ данных»

дополнительной профессиональной программы –
программы профессиональной переподготовки
«Разработка прикладного программного обеспечения»

Киров, 2022

Рабочая программа составлена в соответствии с требованиями дополнительной профессиональной программы «Разработка прикладного программного обеспечения»

Рабочая программа разработана:

Чистяков Геннадий Андреевич, кандидат технических наук, доцент, заместитель директора института математики и информационных систем ФГБОУ ВО «ВятГУ»

© Вятский государственный университет, 2022

© Чистяков Геннадий Андреевич, 2022

1. РАБОЧАЯ УЧЕБНАЯ ПРОГРАММА

1.1 Пояснительная записка

Актуальность и значение учебной дисциплины «Анализ данных» определяются необходимостью понимания принципов разработки программных продуктов, включающих в себя функционал по анализу массивов структурированных и/или слабоструктурированных данных, знание которых является ключевым требованием для осуществления деятельности в ИТ-сфере при решении широкого профиля задач.

Цели и задачи учебной дисциплины

Цель учебной дисциплины	Формирование базовых навыков в области анализа данных, выявления скрытых закономерностей и построения соответствующих программных моделей, пригодных для практического применения.
Задачи учебной дисциплины	Формирование представления об анализе данных как части раздела теории интеллектуальных систем. Знакомство с типовым циклом решения задачи анализа данных. Получения навыков решения задач регрессии и классификации. Формирование навыков работы с программными библиотеками анализа данных.

Компетенции слушателя, формируемые в результате освоения учебной дисциплины / модуля

В результате освоения учебной дисциплины (модуля) слушатель должен демонстрировать следующие результаты обучения:

Виды деятельности	Профессиональные компетенции	Практический опыт	Умения	Знания
Производственно-технологическая Организационно-управленческая Научно-исследовательская Проектная Информационно-аналитическая	ПК 1: Применяет языки программирования для решения профессиональных задач	Применяет языки программирования (в т.ч. скрипты) для решения профессиональных задач под контролем более опытных специалистов	Способен создавать программное обеспечение и/или его элементы (модули, подсистемы и т.д.) на языках программирования Python и/или JavaScript	Базовые принципы программирования, принципы работы с языками динамической типизации, стандартные библиотеки Python и/или JavaScript

1.2 Содержание учебной дисциплины (модуля)
Объем учебной дисциплины и виды учебной работы

Форма обучения	Общий объем (трудоемкость) часов	В том числе аудиторная контактная работа обучающихся с преподавателем, час					Самостоятельная работа, час	Контроль, час	Форма промежуточной аттестации
		Всего	Лекции	Практические (семинарские) занятия	Лабораторные занятия	Консультации			
Очная с применением ДОТ	32	16	10	-	6	-	14	2	Зачет

Тематический план

№ п/п	Основные темы курса	Часы		Самостоятельная работа	Контроль
		Лекции	Лабораторные занятия		
1.	Основы анализа данных	4	2	2	
2.	Подходы к решению задач регрессии и классификации	4	2	2	
3.	Библиотеки анализа данных	2	2	10	
4.	Промежуточная аттестация	-	-	-	2
	Итого:	10	6	14	2

Матрица соотнесения разделов / тем учебной дисциплины / модуля и формируемых в них компетенций

Темы учебной дисциплины	Количество часов	Компетенции	
		ПК 1	Общее количество компетенций
Тема 1. Основы анализа данных	8	+	1
Тема 2. Подходы к решению задач регрессии и классификации	8	+	1
Тема 3. Библиотеки анализа данных	14	+	1
Промежуточная аттестация	2	+	1
Итого:	32		

Краткое содержание учебной дисциплины «Анализ данных»

Тема 1. Основы анализа данных

Классический цикл решения задачи анализа данных. Сбор данных. Подготовка данных. Выбор моделей. Параметризация моделей. Обучение. Оценка.

Тема 2. Подходы к решению задач регрессии и классификации

Классические модели. Линейная и логистическая регрессия. Регуляризация. Дерево решений. Построение ансамблей. Нейросетевые модели.

Тема 3. Библиотеки анализа данных

Особенности работы с библиотекой `sklearn`. Особенности работы с библиотекой `xgboost`. Фреймворки `tensorflow` и `keras`.

2. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

2.1. Методические рекомендации для преподавателя

Выбор методов и средств обучения, образовательных технологий осуществляется преподавателем исходя из необходимости достижения обучающимися планируемых результатов освоения дисциплины, а также с учетом индивидуальных возможностей обучающихся из числа инвалидов и лиц с ограниченными возможностями здоровья.

Организация учебного процесса предусматривает применение инновационных форм учебных занятий, развивающих у обучающихся навыки командной работы, межличностной коммуникации, принятия решений, лидерские качества (включая, при необходимости, проведение интерактивных лекций, тренингов, анализ ситуаций и имитационных моделей, в том числе с учетом региональных особенностей профессиональной деятельности выпускников и потребностей работодателей).

2.2. Методические указания для слушателей

Успешное освоение учебной дисциплины предполагает активное, творческое участие слушателей на всех этапах ее освоения путем планомерной, повседневной работы. Слушатели обязаны посещать лекции и лабораторные занятия, 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. Содержание лекций определяется настоящей рабочей программой дисциплины.

Лекции – это систематическое устное изложение учебного материала. На них слушатели получают основной объем информации по каждой конкретной теме. Лекции обычно носят проблемный характер и нацелены на освещение наиболее трудных и дискуссионных вопросов, кроме того, они способствуют формированию у слушателей навыков самостоятельной работы с научной литературой.

Предполагается, что слушатели приходят на лекции, предварительно проработав соответствующий учебный материал по источникам, рекомендуемым программой. Часто слушателям 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 лабораторных занятий является проверка уровня понимания слушателями вопросов, рассмотренных на лекциях и в учебной литературе, степени и качества усвоения материала; применение теоретических знаний в реальной практике решения задач; восполнение пробелов в пройденной теоретической части курса и оказания помощи в его освоении.

Лабораторные занятия в равной мере направлены на совершенствование индивидуальных навыков решения теоретических и прикладных задач, выработку навыков интеллектуальной работы, а также ведения дискуссий. Конкретные пропорции разных видов работы в группе, а также способы их оценки определяются преподавателем, ведущим занятия.

На лабораторных занятиях под руководством преподавателя слушатели обсуждают дискуссионные вопросы, отвечают на вопросы тестов, закрепляя приобретенные знания, выполняют лабораторные задания и т.п. Для успешного проведения лабораторного занятия слушателям следует тщательно подготовиться.

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

Изучив конкретную тему, слушатель может определить, насколько хорошо он в ней разобрался. Если какие-то моменты остались непонятными, целесообразно составить список вопросов и на занятии задать их

преподавателю. Лабораторные занятия предоставляют слушателю возможность творчески раскрыться, проявить инициативу и развить навыки публичного ведения дискуссий и общения, сформировать определенные навыки и умения и т.п.

Самостоятельная работа слушателей включает в себя освоение материала на внешних ресурсах, рекомендованных преподавателем, и выполнение различного рода заданий (изучение учебной и научной литературы, материалов лекций, систематизацию прочитанного материала, подготовку контрольной работы, решение задач и т.п.), которые ориентированы на самопроверку и более глубокое усвоение материала изучаемой дисциплины.

Регулярно рекомендуется отводить время для повторения пройденного материала, проверяя свои знания, умения и навыки по контрольным вопросам.

Результатом самостоятельной работы должно стать формирование у слушателей определенных компетенций.

Процедура оценивания результатов освоения учебной дисциплины осуществляется на основе действующего Положения об организации текущего контроля успеваемости и промежуточной аттестации обучающихся ВятГУ.

Для приобретения требуемых компетенций, хороших знаний и пороговой оценки по дисциплине слушателям необходимо выполнять все виды работ своевременно в течение всего периода обучения.

3. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ УЧЕБНОЙ ДИСЦИПЛИНЫ

Основная литература

1. Технологии интеллектуального анализа данных : учебное пособие / Н. Г. Талипов. - Казань : КНИТУ-КАИ, 2020. - 308 с. - URL: <https://e.lanbook.com/book/193530>. - Режим доступа: ЭБС Лань. - ISBN 978-5-7579-2488-5.

2. Интеллектуальный анализ данных / О. Н. Сапрыкин. - Самара : СамГУ, 2020. - 80 с. - URL: <https://e.lanbook.com/book/188906>. - Режим доступа: ЭБС Лань. - ISBN 978-5-7883-1563-8.

3. Технологии интеллектуального анализа данных : учебное пособие / Д. С. Алексеев. - Кострома : КГУ им. Н.А. Некрасова, 2020. - 141 с. - URL: <https://e.lanbook.com/book/160082>.

4. Анализ данных [Электронный ресурс]. - Москва : РТУ МИРЭА, 2020 - Ч. 1 : учебное пособие / О. А. Новикова, Е. Г. Андрианова. - Москва : РТУ МИРЭА, 2020. - 162 с. - URL: <https://e.lanbook.com/book/167597>.

5. Анализ данных : учебно-методическое пособие / П. Е. Котиков. - Санкт-Петербург : СПбГПУ, 2019. - 48 с. - URL:

<https://e.lanbook.com/book/174498>. - Режим доступа: ЭБС Лань. - ISBN 978-5-907184-46-6.

Дополнительная литература

1. Информационное обеспечение систем анализа данных : учебное пособие / Б. А. Лаговский. - Москва : РТУ МИРЭА, 2021. - 66 с. - URL: <https://e.lanbook.com/book/171549>.

2. Современные методы прикладной информатики в задачах анализа данных: учебное пособие по курсу «Методы интеллектуального анализа данных» : учебное пособие / А. Н. Целых, А. А. Целых, Э. М. Котов. - Ростов-на-Дону, Таганрог : Южный федеральный университет, 2021. - 130 с. : ил., табл., схем.. - Библиогр. в кн - URL: <https://biblioclub.ru/index.php?page=book&id=683920/>. - Режим доступа: ЭБС Университетская библиотека ONLINE. - ISBN 978-5-9275-3783-9.

3. Статистический анализ данных в информационных системах : учебно-методическое пособие / И. Е. Тарасов. - Москва : РТУ МИРЭА, 2020. - 96 с. - URL: <https://e.lanbook.com/book/163854>.

4. Методы анализа данных : учебное пособие / В. И. Кузьмин, А. Ф. Гадзаов. - 2-е изд., перераб. и доп.. - Москва : РТУ МИРЭА, 2020. - 155 с. - URL: <https://e.lanbook.com/book/171433>.

5. Анализ данных в среде R : практикум / С. А. Агалаков. - Омск : Омский государственный университет им. Ф.М. Достоевского (ОмГУ), 2020. - 52 с.. - Библиогр.: с. 52. - URL: <https://biblioclub.ru/index.php?page=book&id=614033/>. - Режим доступа: ЭБС Университетская библиотека ONLINE. - ISBN 978-5-7779-2438-4.

Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины

1. Портал дистанционного обучения ВятГУ.
2. Раздел официального сайта ВятГУ, содержащий описание образовательной программы

Описание материально-технической базы, необходимой для осуществления образовательного процесса

Перечень специализированных аудиторий (лабораторий)

Вид занятий	Назначение аудитории
Лекция	Учебная аудитория, оснащенная мультимедийным оборудованием
Лабораторная работа	Компьютерный класс
Самостоятельная работа	Читальные залы библиотеки

Перечень специализированного оборудования

Перечень используемого оборудования	
Мультимедиа-проектор с экраном настенным	
Ноутбук (персональный компьютер)	

Перечень информационных технологий, используемых при осуществлении образовательного процесса по учебной дисциплине

№ п.п	Наименование ПО	Краткая характеристика назначения ПО	Производитель ПО и/или поставщик 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	ЗАО "Анти-Плагиат"
2	Microsoft Office 365 Student Advantage	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	ООО "Рубикон"
3	Office Professional Plus 2013 Russian OLP NL Academic.	Пакет приложений для работы с различными типами документов: текстами, электронными таблицами, базами данных, презентациями	ООО "СофтЛайн" (Москва)
4	Windows 7 Professional and Professional K	Операционная система	ООО "Рубикон"
5	Kaspersky Endpoint Security для бизнеса	Антивирусное программное обеспечение	ООО «Рубикон»
6	Информационная система КонсультантПлюс	Справочно-правовая система по законодательству Российской Федерации	ООО «КонсультантКиров»
7	Электронный периодический справочник «Система ГАРАНТ»	Справочно-правовая система по законодательству Российской Федерации	ООО «Гарант-Сервис»
8	Security Essentials (Защитник Windows)	Защита в режиме реального времени от шпионского программного обеспечения, вирусов.	Microsoft

4. МАТЕРИАЛЫ, УСТАНОВЛИВАЮЩИЕ СОДЕРЖАНИЕ ТЕКУЩЕГО КОНТРОЛЯ УСПЕВАЕМОСТИ (ТКУ) И САМОСТОЯТЕЛЬНОЙ РАБОТЫ СЛУШАТЕЛЕЙ

Формы ТКУ:

- собеседование;
- тест;
- контрольная работа;
- лабораторная, и т.п. работа.

Формы самостоятельной работы:

- конспектирование;
- реферирование литературы;
- аннотирование книг, статей;
- выполнение заданий поисково-исследовательского характера;
- углубленный анализ научно-методической литературы;
- работа с лекционным материалом: проработка конспекта лекций, работа на полях конспекта с терминами, дополнение конспекта материалами из рекомендованной литературы;
- лабораторно-практические занятия: выполнение задания в соответствии с инструкциями и методическими указаниями преподавателя, получение результата.

5. МАТЕРИАЛЫ, УСТАНОВЛИВАЮЩИЕ СОДЕРЖАНИЕ И ПОРЯДОК ПРОВЕДЕНИЯ ПРОМЕЖУТОЧНЫХ АТТЕСТАЦИЙ

К сдаче зачета допускаются все слушатели, проходящие обучение на данной ДПП, вне зависимости от результатов текущего контроля успеваемости и посещаемости занятий, при этом, результаты текущего контроля успеваемости могут быть использованы преподавателем при оценке уровня усвоения обучающимися знаний, приобретения умений, навыков и сформированности компетенций в результате изучения учебной дисциплины.

Промежуточная аттестация проводится в форме зачета (тестовых заданий).

Зачет принимается преподавателями, проводившими лекции по данной учебной дисциплине. Допускается использование средств электронной информационно-образовательной среды.

Методические рекомендации по подготовке и проведению промежуточной аттестации

Промежуточная аттестация проводится в целях повышения эффективности обучения, определения уровня профессиональной подготовки

обучающихся и контролем за обеспечением выполнения стандартов обучения.

Примерный перечень вопросов

1. Стандартный цикл решения задачи анализа данных.
2. Оценочные метрики.
3. Особенности работы с тестовой и обучающей выборками. Кросс-валидация.
4. Типовые действия этапа подготовки данных.
5. Классификация задач анализа данных.
6. Классические модели для решения задачи регрессии.
7. Классические модели для решения задачи классификации.
8. Построение ансамблей моделей.
9. Особенности работы с библиотекой sklearn.
10. Особенности работы с библиотекой xgboost.

Критерии оценивания

Результаты промежуточной аттестации определяются по системе: «зачтено», «не зачтено».

Оценки «зачтено» заслуживает ответ слушателя, в котором полностью раскрыто теоретическое содержание заявленных в экзаменационном билете вопросов. Представлен анализ практической составляющей вопроса, слушатель приводит примеры, аргументирует и соотносит теоретические знания с профессиональной сферой; использует творческий подход к решению проблемных вопросов; владеет навыками обобщения, систематизации и обоснования выводов, предложений по конкретному вопросу.

Оценки «не зачтено» заслуживает слушатель, который обнаруживает существенные пробелы в знании основного учебного материала, допустивший принципиальные ошибки; если слушатель не дал правильных ответов на большинство вопросов. Слушатель демонстрирует несформированность компетенций.