

Факультет электроники и вычислительной техники

УТВЕРЖДЕНО

Факультет электроники и вычислительной
техники

Декан Авдеюк О.А.
Г.

Технологии анализа данных

рабочая программа дисциплины (модуля, практики)

Закреплена за кафедрой **Системы автоматизированного проектирования и поискового конструирования**

Учебный план	Направление 09.04.01 Информатика и вычислительная техника Программа "Анализ данных и интеллектуальные технологии"
--------------	--

Профиль

Квалификация	Магистр
--------------	---------

Срок обучения **2 года**

Форма обучения	очная	Общая трудоемкость	3 ЗЕТ
----------------	--------------	--------------------	--------------

Виды контроля в экзамены 1 семестрах:

Семестр(Курс.Номер семестра на курсе)	1(1.1)		Итого	
	УП	ПП	УП	ПП
Лекции	16	16	16	16
Лабораторные	16	16	16	16
Итого ауд.	32	32	32	32
Контактная работа	32.35	32.35	32.35	32.35
Сам. работа	40	55.65	40	55.65
Часы на контроль	35.65	20	35.65	20
Практическая подготовка	0	0	0	0
Итого трудоемкость в часах	108	108	0	0

ЛИСТ ОДОБРЕНИЯ, СОГЛАСОВАНИЯ И АКТУАЛИЗАЦИИ РАБОЧЕЙ ПРОГРАММЫ

Разработчик(и) программы:

Профессор Садовникова Н.П. д.т.н.

Рецензент(ы):

(при наличии)

к.т.н., директор по развитию, ООО «Эдванс Софт», Набока М.В.

Рабочая программа дисциплины (модуля, практики)

Технологии анализа данных

разработана в соответствии с ФГОС ВО:

Федеральный государственный образовательный стандарт высшего образования - магистратура по направлению подготовки 09.04.01 Информатика и вычислительная техника (приказ Минобрнауки России от 19.09.2017 г. № 918)

составлена на основании учебного плана:

Направление 09.04.01 Информатика и вычислительная техника

Программа "Анализ данных и интеллектуальные технологии"

Профиль:

утвержденного учёным советом вуза от 05.06.2019 протокол № 12.

Рабочая программа одобрена на заседании кафедры

Системы автоматизированного проектирования и поискового конструирования

номер протокола 2019 г.

Зав. кафедрой Щербаков Максим Владимирович

Рабочая программа дисциплины (модуля, практики) актуализирована 30.08.2024

СОГЛАСОВАНО:

Факультет электроники и вычислительной техники

Председатель НМС факультета: Авдеюк О.А.

Протокол заседания НМС от

г. №

1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ, ПРАКТИКИ). ВИД, ТИП ПРАКТИКИ, СПОСОБ И ФОРМА (ФОРМЫ) ЕЕ ПРОВЕДЕНИЯ.
Цели и задачи
изучение современных технологий сбора, хранения и подготовки данных к анализу;
изучение основных методов поиска закономерностей, связей, правил в массивах данных; подходов к визуализации данных на различных этапах анализа;
овладение практическими умениями и навыками реализации технологий интеллектуального анализа данных, формирования и проверки гипотез о их природе и структуре; формирование умений и навыков применения специализированного программного обеспечения к задачам анализа данных.

2. МЕСТО ДИСЦИПЛИНЫ (МОДУЛЯ, ПРАКТИКИ) В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ				
Цикл (раздел) ОП:	Б1.В			
2.1	Требования к предварительной подготовке обучающегося:			
2.1.1	Междисциплинарный курсовой проект			
2.1.2	Системы обработки больших данных			
2.1.3	Теория вероятностей, математическая статистика и численные методы			
2.2	Дисциплины (модули) и практики, для которых освоение данной дисциплины (модуля) необходимо как предшествующее:			
2.2.1	Машинное обучение			
2.2.2	Онтологический инжиниринг и семантические технологии			
2.2.3	Системы поддержки принятия решений			
2.2.4	Учебная практика: Технологическая (проектно-технологическая) практика			
2.2.5	Выполнение и защита выпускной квалификационной работы			
2.2.6	Производственная практика: Преддипломная практика			
3. КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ, ПРАКТИКИ)				
ПК-3: Администрирование систем управления базами данных и системного программного обеспечения инфокоммуникационной системы организации				
ПК-3.1: Знает: основные принципы систем управления базами данных и системного программного обеспечения инфокоммуникационной системы организации.				
Результаты обучения:				
ПК-3.2: Умеет: администрировать СУБД и управлять системным программным обеспечением инфокоммуникационной системы организации.				
Результаты обучения:				
ПК-3.3: Владеет навыками: применения современных инструментов управления базами данных и управления системным программным обеспечением инфокоммуникационной системы организации				
Результаты обучения:				
ПК-11: Управление сервисами ИТ				
ПК-11.1: Знает: основы управления сервисами ИТ.				
Результаты обучения:				
ПК-11.2: Умеет: управлять сервисами ИТ.				
Результаты обучения:				
ПК-11.3: Владеет навыками: применения современных инструментов управления сервисами ИТ.				
Результаты обучения:				
ПК-16: Управление аналитическими работами и подразделением				
ПК-16.1: Знает: методы управления аналитическими работами в составе подразделений.				
Результаты обучения:				
ПК-16.2: Умеет: организовывать и управлять аналитическими работами и подразделением.				
Результаты обучения:				
ПК-16.3: Владеет навыками: использования современных средств управления аналитическими работами в составе междисциплинарных подразделений.				
Результаты обучения:				
4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ, ПРАКТИКИ)				
Код занятия	Наименование разделов и тем /вид занятия/	Семестр / Курс	Часов	Форма контроля
1	Раздел 1.			

1.1	Задачи и методы анализа данных /Тема/	1	0	
1.1.1	Задачи анализа данных. Типы данных. Источники данных. Шкалы измерений. Примеры прикладных задач. Задачи анализа данных. Типы данных. Источники данных. Шкалы измерений. Примеры прикладных задач. /Лек/	1	2	Эк
1.2	Современные технологии сбора данных. Концепция СППР на основе анализа данных /Тема/	1	0	
1.2.1	Понятие «Big data». Data Mining. СППР на основе данных /Лек/	1	2	Эк
1.2.2	Основы работы в R. /Лаб/	1	4	К
1.3	Технологии хранения и обработки данных. /Тема/	1	0	
1.3.1	Процесс ETL. Извлечение данных. Преобразования данных. моделирование ETL -процессов. Программные средства ETL. Разработка ETL-систем /Лек/	1	2	Эк
1.4	Методы анализа качества данных. /Тема/	1	0	
1.4.1	Подготовка данных к анализу. Технологии и методы оценки качества данных. Профайлинг. Очистка и предобработка данных /Лек/	1	2	Эк
1.4.2	Классические методы статистики и визуализация в R /Лаб/	1	4	К
1.5	Методологии ведения проектов анализа данных. /Тема/	1	0	
1.5.1	Концепция проекта анализа данных. CRISP-DM — межотраслевой стандарт ведения проектов анализа данных. Этапы CRISP-DM. Методология SEMMA /Лек/	1	2	Эк
1.6	Классические методы анализа данных. /Тема/	1	0	
1.6.1	Числовые характеристики распределений, проверка статистических гипотез. Дисперсионный анализ. Описательная статистика. Визуализация данных /Лек/	1	2	Эк
1.6.2	Регрессионный анализ в R /Лаб/	1	4	К
1.7	Методы анализа данных на основе машинного обучения. /Тема/	1	0	
1.7.1	Кластерный анализ. Регрессионный анализ. Методы классификации данных /Лек/	1	2	Эк
1.8	Специализированные языки и инструменты анализа данных. /Тема/	1	0	
1.8.1	Пакеты STATISTICA. SAS. Языки R, Julia, Scala, Python /Лек/	1	2	Эк
1.8.2	Кластерный анализ и классификация в R /Лаб/	1	4	К
1.8.3	/Ср/	1	40	К
1.8.4	/КоПа/	1	0.35	Ко
1.8.5	/Экзамен/	1	20	Эк
1.8.6	контрольная работа /Контр.раб./	1	15.65	К

Примечание. Формы контроля: Эк – экзамен, К- контрольная работа, Ко- контрольный опрос, Сз- семестровое задание, З-зачет, ОП- отчет по практике, Зд-задание, Р-реферат.

5. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

Оценочные средства планируемых результатов обучения представлены в виде фондов оценочных средств (ФОС), разработанных в соответствии с локальным нормативным актом университета. ФОС может быть представлен в Приложении к рабочей программе.

1. Формы представления, типы и виды анализируемых данных.
2. Общая схема анализа данных. Требования к алгоритмам анализа данных.
3. Способы получения и предобработки данных.
4. Основные виды измерительных шкал
5. Примеры задач анализа данных.
6. Числовые характеристики выборки
7. Понятие распределения случайной величины.
8. Статистические гипотезы.
9. Дисперсионный анализ.
10. Критерий Фишера.
11. Цели, задачи и основное содержание кластерного анализа.
12. Классификация методов кластеризации.
13. Способы определения меры расстояния между кластерами.
14. Алгоритм кластеризации k-means.
15. Цели, задачи классификации.
16. Корреляционный анализ
17. Цели, задачи регрессии.

18. Простая линейная регрессия.
19. Оценка соответствия простой линейной регрессии реальным данным.
20. Оценка значимости простой регрессионной модели
21. Множественная линейная регрессия.
22. Оценка значимости множественной регрессионной модели.
23. Этапы регрессионного анализа.
24. Метод наименьших квадратов.
25. Ограничения применимости регрессионных моделей.
26. Логистическая регрессия.
27. Общая характеристика алгоритмов построения деревьев решений.
28. Наивный байесовский классификатор.

6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ, ПРАКТИКИ)

6.1. Рекомендуемая литература

	Авторы, составители	Заглавие	Издательство, год.	Электронный адрес
Л1.1	Бонцанини М., Логунова А. В.	Анализ социальных медиа на Python. Извлекайте и анализируйте данные из всех уголков социальной паутины на Python	Москва: ДМК Пресс, 2018	https://e.lanbook.com/reader/book/108129/#6
Л1.2	Мартин О., Снастина В. С.	Байесовский анализ на Python	Москва: ДМК Пресс, 2020	https://e.lanbook.com/reader/book/140585/#9
Л1.3	Рочев К. В.	Информационные технологии. Анализ и проектирование информационных систем: учебное пособие	Санкт-Петербург: Лань, 2019	https://e.lanbook.com/reader/book/122181/#128
Л1.4	Маккинни У.	Python и анализ данных	Москва: ДМК Пресс, 2020	https://e.lanbook.com/book/131721
Л1.5	Груздев А. В., Хейдт М.	Изучаем Pandas	Москва: ДМК Пресс, 2019	https://e.lanbook.com/book/131693

6.2. Перечень ресурсов информационно-телекоммуникационной сети "Интернет"

Э1	Онлайн-курс "Технологии анализа данных" в СДО «Moodle»
Э2	Анализ данных с использованием R. Use R to clean, analyze, and visualize data. Learn how to ask the right questions, obtain data, and perform reproducible research. Use GitHub to manage data science projects. https://ru.coursera.org/learn/data-scientists-tools#about

6.3 Перечень программного обеспечения

6.3.1.1	Adobe Acrobat Reader DC — бесплатное решение для просмотра файлов PDF
6.3.1.2	LibreOffice — офисный пакет

6.4 Перечень информационных справочных систем и электронных библиотечных систем (ЭБС)

6.3.2.1	Библиотека (НТБ), http://library.vstu.ru/sci-nci
6.3.2.2	Электронная информационно-образовательная среда университета, http://eos.vstu.ru
6.3.2.3	ЭБС "Лань", https://e.lanbook.com/
6.3.2.4	ЭБС "Book.ru", https://www.book.ru/
6.3.2.5	Электронная библиотека "Grebennikon", https://grebennikon.ru/

7. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ, ПРАКТИКИ) /ОБОРУДОВАНИЕ

7.1	Мультимедийная учебная аудитория для проведения занятий лекционного и семинарского типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации. /Учебная доска, учебная мебель, интерактивная трибуна, видеопроектор.
7.2	Лаборатория информационных технологий. /Учебная мебель, компьютерная техника, оснащенная программным обеспечением, доступом в Интернет и в электронную информационно-образовательную среду университета.
7.3	Аудитория для самостоятельной работы обучающихся./Учебная мебель, компьютерная техника с возможностью подключения к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду университета (читальный зал информационно-библиотечного центра)

8. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ (МОДУЛЯ, ПРАКТИКИ)

Организация образовательного процесса по данной дисциплине регламентируется учебным планом и расписанием учебных занятий. При формировании своей индивидуальной образовательной траектории обучающийся имеет право на перезачет дисциплины (переаттестации ее части), если она была освоена в процессе предшествующего обучения. Перезачёт (переаттестации ее части) освобождает обучающегося от необходимости повторного освоения дисциплины (полностью или частично).

Учебный процесс при преподавании курса основывается на использовании традиционных, инновационных и информационных образовательных технологий. Традиционные образовательные технологии представлены лекциями и

практическими занятиями. Инновационные образовательные технологии используются в виде широкого применения активных и интерактивных форм проведения занятий. Информационные образовательные технологии реализуются путем активизации самостоятельной работы студентов в электронной информационной образовательной среде.

Лекционный курс предполагает систематизированное изложение основных вопросов учебного плана. На первой лекции лектор информирует студентов о рекомендуемой литературе и электронных источниках информации по дисциплине, с указанием, какой учебник (учебное пособие) является базовым.

Практические занятия представляют собой детализацию лекционного теоретического материала, проводятся в целях закрепления курса и охватывают основные разделы дисциплины.

Основной формой проведения практических занятий является решение конкретных задач, аналогичные которым, будут выполнять студенты на лабораторных работах.

Лабораторные работы предполагают выполнение и отчет заданий по темам, рассмотренным на лекционных и закрепленных на практических занятиях. Каждому лабораторному занятию предшествует самостоятельная подготовка студента, включающая: ознакомление с содержанием лабораторной работы по методическим указаниям; проработку теоретической части по лекционному материалу и учебникам, рекомендованным в методических указаниях;

Самостоятельная работа студентов включает изучение законспектированного на лекционных занятиях материала, дополнение его с учетом рекомендованной по данной теме литературы, самостоятельную подготовку к лабораторным работам, самостоятельное выполнение и оформление заданий контрольной работы, аналогичных выполненным на занятиях.

Перечень методических указаний для освоения дисциплины представлен в таблице 6.1.3

В течении семестра для студентов проводятся групповые текущие консультации по учебной дисциплине, а также консультация перед экзаменом.

Методические рекомендации по обучению лиц с ограниченными возможностями здоровья и инвалидов

Профессорско-педагогический состав знакомится с психолого-физиологическими особенностями обучающихся инвалидов и лиц с ограниченными возможностями здоровья (ОВЗ), индивидуальными программами реабилитации инвалидов (при наличии). При необходимости осуществляется дополнительная поддержка преподавания тьюторами, психологами, социальными работниками, прошедшими подготовку ассистентами.

В соответствии с методическими рекомендациями Минобрнауки РФ (утв. 8 апреля 2014 г. N АК-44/05вн), в курсе предполагается использовать социально-активные и рефлексивные методы обучения, технологии социокультурной реабилитации с целью оказания помощи в установлении полноценных межличностных отношений с другими студентами, создании комфортного психологического климата в студенческой группе. Подбор и разработка учебных материалов производится с учетом предоставления материала в различных формах: аудиальной, визуальной, с использованием специальных технических средств и информационных систем.

Освоение дисциплины лицами с ОВЗ осуществляется с использованием средств обучения общего и специального назначения (персонального и коллективного использования). Материально-техническое обеспечение предусматривает приспособление аудиторий к нуждам лиц с ОВЗ (при необходимости).

Форма проведения аттестации для студентов-инвалидов устанавливается с учетом индивидуальных психофизических особенностей. Для студентов с ОВЗ предусматривается доступная форма предоставления заданий оценочных средств. Студентам с инвалидностью увеличивается время на подготовку ответов на контрольные вопросы. Для таких студентов предусматривается доступная форма предоставления ответов на задания.

При необходимости для обучающихся с инвалидностью процедура оценивания результатов обучения может проводиться в несколько этапов.