

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Максимов Алексей Борисович

Должность: директор департамента по образовательной политике

Дата подписания: 22.11.2025 10:46:06

Уникальный программный код:

8db180d1a3f02ac9e60521a5672742735c18b1d6

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«МОСКОВСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»
(МОСКОВСКИЙ ПОЛИТЕХ)

Факультет информационных технологий

УТВЕРЖДЕНО

Декан факультета

Информационных технологий

/ Д.Г. Демидов /



«16» 02 2023 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ «Анализ и автоматическая обработка данных»

Направление подготовки

09.04.01 «Информатика и вычислительная техника»

Профиль

«Интеллектуальные системы»

Квалификация

Магистр

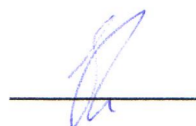
Форма обучения

Очная

Москва, 2023 г.

Разработчики:

преподаватель кафедры
«Инфокогнитивные технологии»



Кружалов А. С.

Согласовано:

Заведующий кафедрой
«Инфокогнитивные технологии»
к. т. н., доцент



Пухова Е. А.

Содержание

1	Цель, задачи и планируемые результаты обучения по дисциплине	4
2	Место дисциплины в структуре образовательной программы	5
3	Структура и содержание дисциплины	6
3.1	Виды учебной работы и трудоемкость	6
3.2	Тематический план изучения дисциплины	6
3.3	Содержание дисциплины	7
3.4	Тематика практических и лабораторных занятий	7
4	Учебно-методическое и информационное обеспечение	7
4.1	Основная литература	7
4.2	Дополнительная литература	8
5	Материально-техническое обеспечение	8
6	Методические рекомендации	8
7	Фонд оценочных средств	9
7.1	Методы контроля и оценивания результатов обучения	9
7.2	Шкала и критерии оценивания результатов обучения	10
7.3	Оценочные средства	13
	Приложение 1. Образец оформления титульного листа к отчётам	15

1 Цель, задачи и планируемые результаты обучения по дисциплине

Целью освоения дисциплины является приобретение студентами теоретических знаний и практических навыков работы с методами анализа данных и построения математических моделей на основе экспериментальных данных.

Задачи освоения дисциплины:

- изучение основ кластерного анализа и линейного регрессионного анализа;
- изучение логических алгоритмов классификации;
- приобретение практических навыков построения ансамблей обучающихся алгоритмов.

Обучение по дисциплине «Анализ и автоматическая обработка данных» направлено на формирование у обучающихся следующих компетенций:

Код и наименование компетенций	Индикаторы достижения компетенции
ПК-2. Управление программно-техническими, технологическими и человеческими ресурсами.	ИПК 2.1. Знает: методологии разработки ПО и управления; методы и средства организации проектных данных, управления рисками; нормативно-технические документы (стандарты и регламенты), описывающие процессы управления инфраструктурой коллективной среды разработки, управления рисками, оценки сложности, трудоемкости, сроков выполнения работ; основные принципы и методы управления персоналом; методы и программные средства для оценки сложности, трудоемкости и сроков выполнения работ. ИПК 2.2. Умеет: применять методологии разработки программного обеспечения, управления проектами разработки программного обеспечения; методы и средства организации проектных данных, управления рисками, оценки сложности, трудоемкости и сроков выполнения работ; основные принципы и методы управления персоналом; нормативно-технические документы (стандарты и регламенты), применять основные принципы и методы управления персоналом; определять потребность в персонале; применять методы планирования развития и обучения персонала, методы оценки квалификации персонала. ИПК 2.3. Владеет: методом выбора инструментальных средств разработки; определением набора библиотек повторно используемых модулей, областей применения процесса управления рисками, методами выявления и отслеживания рисков в процессе разработки ПО в медицинской сфере; определения критериев (показателей) оценки сложности, трудоемкости, сроков выполнения работ; формирование запросов на поиск персонала; планирования и организации обучения и развития персонала, оценок квалификации персонала.

<p>УК-2. Способен управлять проектом на всех этапах его жизненного цикла.</p>	<p>ИУК-2.1. Разрабатывает концепцию управления проектом на всех этапах его жизненного цикла в рамках обозначенной проблемы: формулирует цель и пути достижения, задачи и способы их решения, обосновывает актуальность, значимость, ожидаемые результаты и возможные сферы их применения.</p> <p>ИУК-2.2. Разрабатывает план реализации проекта в соответствии с существующими условиями, необходимыми ресурсами, возможными рисками и распределением зон ответственности участников проекта.</p> <p>ИУК-2.3. Осуществляет мониторинг реализации проекта на всех этапах его жизненного цикла, вносит необходимые изменения в план реализации проекта с учетом количественных и качественных параметров достигнутых промежуточных результатов.</p>
---	---

2 Место дисциплины в структуре образовательной программы

Дисциплина «Анализ и автоматическая обработка данных» относится к дисциплинам вариативной части учебного плана по направлению подготовки 09.04.01 «Информатика и вычислительная техника».

Дисциплина рассчитана на слушателей, имеющих базовые знания и компетенции в области проектирования интеллектуальных систем, теории вероятности и математической статистики, полученные в рамках обучения в бакалавриате, а также при изучении дисциплины «Проектирование интеллектуальных систем».

3 Структура и содержание дисциплины

Общая трудоемкость дисциплины составляет 2 зачетные единицы (72 часа).

3.1 Виды учебной работы и трудоемкость

3.1.1 Очная форма обучения

№ п/п	Вид учебной работы	Количество часов	Семестры	
			Семестр	Неделя семестра
1	Аудиторные занятия	28	4	1-11
	В том числе:			
1.1	Лекции	8	4	1-4
1.2	Практические занятия	10	4	5-7
1.3	Лабораторные занятия	10	4	8-11
2	Самостоятельная работа	44	4	1-11
3	Промежуточная аттестация		4	12
	Зачёт			
	Итого:	72		

3.2 Тематический план изучения дисциплины

3.2.1 Очная форма обучения

№ п/п	Наименование тем, разделов	Трудоемкость, час				
		Всего	Аудиторная работа			Самостоятельная работа
			Лекции	Практические занятия	Лабораторные занятия	
1	Тема 1. Кластерный анализ и визуализация многомерных данных	8	2	6	0	11
2	Тема 2. Линейный регрессионный анализ	6	2	4	0	11
3	Тема 3. Деревья решений и методы ансамблирования обучающихся алгоритмов	8	2	0	6	11
4	Тема 4. Градиентный бустинг	6	2	0	4	11
	Итого	28	8	10	10	44

3.3 Содержание дисциплины

Тема 1. Кластерный анализ и визуализация многомерных данных

Постановка задачи кластеризации. Методы определения схожести объектов. Методы нормализации. Определение наличия кластерной структуры. Статистика Хопкинса. Алгоритм К средних. Нейронная сеть Кохонена. Коэффициент силуэта. Самоорганизующиеся карты Кохонена. Примеры использования самоорганизующихся карт.

Тема 2. Линейный регрессионный анализ

Основные понятия регрессионного анализа: целевая переменная и объясняющая переменная (признак, предиктор, регрессор, фактор). Диаграмма рассеяния (поле корреляции). Парная линейная регрессия. Метод наименьших квадратов (МНК). Интерпретация параметров регрессионной модели. Коэффициент детерминации. Ковариация и коэффициент корреляции Пирсона. Множественная регрессия. Мультиколлинеарность (полная и частичная). Отбор факторов на основе анализа матрицы корреляции. Стандартизированные коэффициенты. Работа с категориальными переменными. Полиномиальная регрессия.

Тема 3. Деревья решений и методы ансамблирования обучающихся алгоритмов

Дерево решений: определение, алгоритм построения дерева решений по обучающей выборке, области применения, достоинства и недостатки деревьев решений. Методы учёта количественных и категориальных признаков при построении тестирующих процедур во внутренних узлах дерева решений. Функционал качества разбиения для построения тестирующих процедур во внутренних узлах дерева решений. Прирост информации. Относительный прирост информации. Критерий Джини (англ. Gini Impurity). Бэггинг. Метод случайных подпространств. Случайный лес: алгоритм построения, достоинства и недостатки.

Тема 4. Градиентный бустинг

Идея последовательного построения ансамбля методом бустинга. Экспоненциальная функция потерь для решения задачи классификации. Математическая постановка задачи построения ансамбля алгоритмов методом градиентного бустинга.

3.4 Тематика практических и лабораторных занятий

3.4.1 Практические занятия

1. Кластеризация текстовых или фактографических данных с помощью сети Кохонена.
2. Линейная регрессия.

3.4.2 Лабораторные занятия

1. Реализация дерева решений.

4 Учебно-методическое и информационное обеспечение

4.1 Основная литература

1. *Миркин Б. Г.* Введение в анализ данных : учебник и практикум / Б. Г. Миркин. — Москва : Издательство Юрайт, 2023. — 174 с. — (Высшее образование). — ISBN 978-5-9916-5009-0. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/511121> (дата обращения: 10.07.2023).

2. *Платонов А. В.* Машинное обучение : учебное пособие для вузов / А. В. Платонов. — Москва : Издательство Юрайт, 2023. — 85 с. — (Высшее образование). — ISBN 978-5-534-15561-7. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/520544> (дата обращения: 10.07.2023).

4.2 Дополнительная литература

1. *Хайкин С.* Нейронные сети: полный курс, 2-е издание.: Пер. с англ. М.: Издательский дом «Вильямс», 2006. 1104 с.
2. *Мандель И. Д.* Кластерный анализ. М.: «Финансы и статистика», 1988. 176 с.
3. *Фёрстер Э.* Методы корреляционного и регрессионного анализа. Руководство для экономистов. / Э. Фёрстер, Б. М. Рёнц – М.: «Финансы и статистика», 1983 г. – 304 с.
4. *Воронцов К. В.* Лекции по искусственным нейронным сетям, 2007. URL: <http://www.ccas.ru/voron/download/NeuralNets.pdf> (дата обращения: 17.09.2020)
5. *Bishop C.* *Pattern Recognition and Machine Learning.* Springer, 2006.
6. *Goodfellow I. J., Bengio Y., Courville A.* Deep Learning. MIT Press, 2016.

4.3 Электронные образовательные ресурсы

1. Электронный образовательный ресурс: Анализ и автоматическая обработка данных <https://online.mospolytech.ru/enrol/index.php?id=7322>
2. ЭБС Лань (lanbook.com)
3. Образовательная платформа Юрайт. Для вузов и ссузов. (urait.ru)

4.4 Современные профессиональные базы данных и информационные справочные системы

- <http://www.grandars.ru/student/statistika/metodologiya-statistiki.html>
- <http://statlab.kubsu.ru/node/4>
- http://studopedia.ru/5_60066_statisticheskie-metodi.html
- ЭБС Лань (lanbook.com)
- Образовательная платформа Юрайт. Для вузов и ссузов. (urait.ru)

5 Материально-техническое обеспечение

Для проведения всех видов занятий необходимо презентационное оборудование (мультимедийный проектор, экран) — 1 комплект.

Для проведения практических и лабораторных занятий требуется компьютерный класс, оборудованный современной вычислительной техникой из расчета одно рабочее место на одного обучаемого с возможностью доступа в Интернет. На компьютерах должен быть установлен современный веб-браузер (Google Chrome версии не ниже 96 или Mozilla Firefox версии не ниже 97), интерпретатор Python (версии не ниже 3.10), среда Jupyter Notebook, Microsoft Office (Libre Office). При изучении дисциплины могут использоваться только лицензионное программное обеспечение и свободно распространяемые интернет-ресурсы.

Компьютеры должны быть оборудованы не менее чем 8 Гб оперативной памяти и иметь не менее 50 Гб свободного дискового пространства.

6 Методические рекомендации

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

Методика преподавания дисциплины и реализация компетентного подхода в изложении и восприятии материала предусматривает использование следующих активных и интерактивных форм проведения групповых, индивидуальных, аудиторных занятий в сочетании с внеаудиторной работой с целью формирования и развития профессиональных навыков у обучающихся:

- выполнение лабораторных работ в лабораториях вуза;
- выполнение практических заданий под руководством преподавателя;
- индивидуальные и групповые консультации студентов преподавателем, в том числе в виде защиты выполненных заданий в рамках самостоятельной работы.

Самостоятельная внеаудиторная работа студентов составляет 61% от общего объема дисциплины и состоит из:

- подготовки к выполнению и защите практических и лабораторных работ;
- чтения литературы и освоения дополнительного материала в рамках тематики дисциплины;
- подготовки к промежуточной аттестации.

7 Фонд оценочных средств

7.1 Методы контроля и оценивания результатов обучения

В процессе обучения используются следующие оценочные формы самостоятельной работы студентов, оценочные средства текущего контроля успеваемости и промежуточных аттестаций:

- выполнение лабораторных работ;
- выполнение практических работ;
- зачёт.

Приём практических и лабораторных работ осуществляется путём предоставления обучающимися отчета о выполненной работе с выводами, содержащими анализ полученных результатов, а также защиты полученных результатов на устном собеседовании. Выполнение работы оценивается от 0 до 35 баллов в зависимости от полноты выполнения и сроков сдачи.

Зачёт проводится в формате устного собеседования на основе предоставленных студентом отчетных материалов (отчётах о выполнении практических и лабораторных работ). Если обучающийся защитил все выполненные работы в течение семестра и набранное им количество баллов удовлетворяет критериям для получения положительной оценки, зачёт может быть проставлен без опроса.

7.2 Шкала и критерии оценивания результатов обучения

Показателем оценивания компетенций на различных этапах их формирования является достижение обучающимися планируемых результатов обучения по дисциплине.

Показатель	Критерии оценивания	
	не зачтено	зачтено
УК-2. Способен управлять проектом на всех этапах его жизненного цикла.		
ИУК-2.1. Разрабатывает концепцию управления проектом на всех этапах его жизненного цикла в рамках обозначенной проблемы: формулирует цель и пути достижения, задачи и способы их решения, обосновывает актуальность, значимость, ожидаемые результаты и возможные сферы их применения	Обучающийся демонстрирует полное отсутствие или недостаточное соответствие знаний, относящихся к компетенции.	Обучающийся демонстрирует полное соответствие знаний, относящихся к компетенции.
ИУК-2.2. Разрабатывает план реализации проекта в соответствии с существующими условиями, необходимыми ресурсами, возможными рисками и распределением зон ответственности участников проекта.	Обучающийся не умеет или в недостаточной степени умеет разрабатывать план реализации проекта в соответствии с существующими условиями, необходимыми ресурсами, возможными рисками и распределением зон ответственности участников проекта.	Обучающийся демонстрирует полное владение умением разрабатывать план реализации проекта в соответствии с существующими условиями, необходимыми ресурсами, возможными рисками и распределением зон ответственности участников проекта.
ИУК-2.3. Осуществляет мониторинг реализации проекта на всех этапах его жизненного цикла, вносит необходимые изменения в план реализации проекта с учётом количественных и качественных параметров достигнутых промежуточных результатов.	Обучающийся не владеет или в недостаточной степени владеет методами осуществления мониторинга реализации проекта на всех этапах его жизненного цикла, внесения необходимых изменений в план реализации проекта с учётом количественных и качественных параметров достигнутых промежуточных результатов.	Обучающийся в полной мере владеет методами осуществления мониторинга реализации проекта на всех этапах его жизненного цикла, внесения необходимых изменений в план реализации проекта с учётом количественных и качественных параметров достигнутых промежуточных результатов.

ПК-2. Управление программно-техническими, технологическими и человеческими ресурсами.		
<p>ИПК 2.1. Знает: методологии разработки ПО и управления; методы и средства организации проектных данных, управления рисками; нормативно-технические документы (стандарты и регламенты), описывающие процессы управления инфраструктурой коллективной среды разработки, управления рисками, оценки сложности, трудоемкости, сроков выполнения работ; основные принципы и методы управления персоналом; методы и программные средства для оценки сложности, трудоемкости и сроков выполнения работ.</p>	<p>Обучающийся демонстрирует полное отсутствие или недостаточное соответствие знаний, относящихся к компетенции.</p>	<p>Обучающийся демонстрирует полное соответствие знаний, относящихся к компетенции.</p>
<p>ИПК 2.2. Умеет: применять методологии разработки программного обеспечения, управления проектами разработки программного обеспечения; методы и средства организации проектных данных, управления рисками, оценки сложности, трудоемкости и сроков выполнения работ; основные принципы и методы управления персоналом; нормативно-технические документы (стандарты и регламенты), применять основные принципы и методы управления персоналом; определять потребность в персонале; применять методы планирования развития и обучения персонала, методы оценки квалификации персонала.</p>	<p>Обучающийся не умеет или в недостаточной степени умеет применять методологии разработки программного обеспечения, управления проектами разработки программного обеспечения; методы и средства организации проектных данных, управления рисками, оценки сложности, трудоемкости и сроков выполнения работ; основные принципы и методы управления персоналом; нормативно-технические документы (стандарты и регламенты), применять основные принципы и методы управления персоналом; определять потребность в персонале; применять методы планирования развития и обучения персонала, методы оценки квалификации персонала.</p>	<p>Обучающийся демонстрирует полное владение умением применять методологии разработки программного обеспечения, управления проектами разработки программного обеспечения; методы и средства организации проектных данных, управления рисками, оценки сложности, трудоемкости и сроков выполнения работ; основные принципы и методы управления персоналом; нормативно-технические документы (стандарты и регламенты), применять основные принципы и методы управления персоналом; определять потребность в персонале; применять методы планирования развития и обучения персонала, методы оценки квалификации персонала.</p>
<p>ИУК-2.3. Осуществляет мониторинг реализации проекта на всех этапах его жизненного цикла, вносит необходимые изменения в план реализации проекта с учётом количественных и качественных параметров достигнутых промежуточных результатов.</p>	<p>Обучающийся не владеет или в недостаточной степени владеет методами осуществления мониторинга реализации проекта на всех этапах его жизненного цикла, внесения необходимых изменений в план реализации проекта с учётом количественных и качественных параметров достигнутых промежуточных результатов.</p>	<p>Обучающийся в полной мере владеет методами осуществления мониторинга реализации проекта на всех этапах его жизненного цикла, внесения необходимых изменений в план реализации проекта с учётом количественных и качественных параметров достигнутых промежуточных результатов.</p>

7.2.1 Критерии оценки ответа на зачёте

«Зачтено»

Студент глубоко и прочно усвоил весь программный материал, исчерпывающе, последовательно, грамотно и логически, стройно его излагает, тесно увязывает с практикой в соответствующей предметной области, не затрудняется с ответом при видоизменении задания, правильно обосновывает принятые решения, умеет самостоятельно обобщать и излагать материал, не допуская ошибок. Допускается наличие неточностей, не искажающих суть излагаемого материала. За выполнение и защиту практических и лабораторных работ студент набрал не менее 60 баллов.

«Не зачтено»

Студент не знает значительной части программного материала, допускает существенные ошибки, за выполнение и защиту практических и лабораторных работ студент набрал менее 60 баллов.

7.2.2 Критерии оценки работы обучающегося на практических и лабораторных занятиях

0 баллов

- Студент не выполнил задание и не предоставил отчёт.
- Студент выполнил задание, но допустил существенные ошибки при выполнении, которые не были исправлены после замечания преподавателя.
- При защите выполненной работы на устном собеседовании студент не смог ответить на вопросы преподавателя, либо допустил существенные ошибки.

15-25 баллов

Студент выполнил работу в полном объёме, предоставил отчёт и защитил работу после установленного преподавателем срока выполнения. Допускаются неточности в ходе выполнения работы, которые были частично исправлены обучающимся после замечания преподавателя.

25-35 баллов

Студент выполнил работу в полном объёме, предоставил отчёт и защитил работу в установленный преподавателем срок.

Оразец оформления титульного листа отчёта по практическим и лабораторным работам представлен в приложении 1. Требования к содержанию отчёта зависят от конкретной работы и приводятся в задании к работе.

7.3 Оценочные средства

7.3.1 Перечень оценочных средств

№ ОС	Наименование ОС	Краткая характеристика ОС	Представление ОС в ФОС
1	Собеседование	Средство контроля, организованное как специальная беседа преподавателя с обучающимся на темы, связанные с изучаемой дисциплиной, и рассчитанное на выяснение объема знаний обучающегося по определенному разделу, теме, проблеме и т. п.	Контрольные вопросы
2	Проект	Конечный продукт, получаемый в результате планирования и выполнения комплекса учебных и исследовательских заданий. Позволяет оценить умения обучающихся самостоятельно конструировать свои знания в процессе решения практических задач и проблем, ориентироваться в информационном пространстве и уровень сформированности аналитических, исследовательских навыков, навыков практического и творческого мышления. Может выполняться в индивидуальном порядке или группой обучающихся.	Темы групповых и/или индивидуальных проектов

7.3.2 Вопросы для зачёта

1. Обучение без учителя. Постановка задачи кластеризации.
2. Метрика. Аксиомы метрики. Евклидово расстояние и линейное расстояние. Методы нормализации.
3. Архитектура нейронной сети Кохонена. Методика конкурентного обучения winner takes all (WTA). Правило Кохонена для обновления весов.
4. Проблема «мёртвых» нейронов и способы её решения.
5. Задача визуализации многомерных данных. Самоорганизующиеся карты Кохонена. Архитектура и сферы применения.
6. Методика конкурентного обучения winner takes most (WTM). Функция топологической окрестности. Алгоритм обучения самоорганизующихся карт Кохонена.
7. Постановка задачи регрессионного анализа. Парная линейная регрессия. Метод наименьших квадратов (МНК). Интерпретация параметров регрессионной модели. Коэффициент детерминации.
8. Множественная линейная регрессия. Мультиколлинеарность. Отбор факторов на основе анализа матрицы корреляции. Работа с категориальными переменными. Полиномиальная регрессия.

9. Дерево решений: определение, алгоритм построения дерева решений по обучающей выборке, области применения, достоинства и недостатки деревьев решений.
10. Методы учёта количественных и категориальных признаков при построении тестирующих процедур во внутренних узлах дерева решений.
11. Функционал качества разбиения для построения тестирующих процедур во внутренних узлах дерева решений. Прирост информации. Относительный прирост информации. Критерий Джини (англ. Gini Impurity).
12. Методы ансамблирования обучающихся алгоритмов. Бэггинг. Метод случайных подпространств. Случайный лес: алгоритм построения, достоинства и недостатки.
13. Градиентный бустинг: алгоритм построения ансамбля, достоинства и недостатки.

Образец оформления титульного листа к отчётам

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«МОСКОВСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»
(МОСКОВСКИЙ ПОЛИТЕХ)

Кафедра «Инфокогнитивные технологии»

Практические и лабораторные занятия по дисциплине
«Анализ и автоматическая обработка данных»

Лабораторная работа № 1
«Реализация дерева решений»

Группа	234-321
Студент	Иванов Иван Иванович
Преподаватель	Кружалов Алексей Сергеевич