

МИНОБРНАУКИ РОССИИ

**Федеральное государственное автономное образовательное
учреждение высшего образования "Пермский
государственный национальный исследовательский
университет"**

Авторы-составители: **Радионова Марина Владимировна
Ильин Иван Вадимович**

Рабочая программа дисциплины

**МАТЕМАТИЧЕСКИЕ МЕТОДЫ И ТЕХНОЛОГИИ ОБРАБОТКИ БОЛЬШИХ
ДАННЫХ**

Код УМК 99862

Утверждено
Протокол №8
от «17» апреля 2024 г.

Пермь, 2024

1. Наименование дисциплины

Математические методы и технологии обработки больших данных

2. Место дисциплины в структуре образовательной программы

Дисциплина входит в обязательную часть Блока « Б.1 » образовательной программы по направлениям подготовки (специальностям):

Направление подготовки: **01.03.02** Прикладная математика и информатика

направленность Анализ данных и искусственный интеллект в цифровой экономике

3. Планируемые результаты обучения по дисциплине

В результате освоения дисциплины **Математические методы и технологии обработки больших данных** у обучающегося должны быть сформированы следующие компетенции:

01.03.02 Прикладная математика и информатика (направленность : Анализ данных и искусственный интеллект в цифровой экономике)

ПК.10 Способен к планированию и организации аналитических работ, в том числе, с использованием технологий больших данных

Индикаторы

ПК.10.1 Разрабатывает, обсуждает и утверждает содержание аналитических работ, в том числе, с использованием технологий больших данных

ПК.10.3 Определяет необходимые ресурсы для проведения аналитических работ

4. Объем и содержание дисциплины

Направление подготовки	01.03.02 Прикладная математика и информатика (направленность: Анализ данных и искусственный интеллект в цифровой экономике)
форма обучения	очная
№№ триместров, выделенных для изучения дисциплины	7
Объем дисциплины (з.е.)	4
Объем дисциплины (ак.час.)	144
Контактная работа с преподавателем (ак.час.), в том числе:	56
Проведение лекционных занятий	28
Проведение практических занятий, семинаров	14
Проведение лабораторных работ, занятий по иностранному языку	14
Самостоятельная работа (ак.час.)	88
Формы текущего контроля	Итоговое контрольное мероприятие (1) Письменное контрольное мероприятие (2)
Формы промежуточной аттестации	Экзамен (7 триместр)

5. Аннотированное описание содержания разделов и тем дисциплины

Интеллектуальный анализ данных

Введение в datamining. Данные и метаданные.

Методы и стадии Data Mining. Задачи Data Mining. Извлечение, данных, информации и знаний. Сферы применения Data Mining.

Сбор, хранение, обработка, извлечение больших потоков данных

Основные характеристики и типы больших данных. Что такое статистическое мышление? Обработка данных для контроля качества процессов и технологии six sigma. Статистический подход к анализу данных. Точечное и интервальное оценивание.

Методы анализа больших потоков данных

Анализ последовательностей. Отбор признаков и снижение размерности. Работа с выбросами и пропущенными значениями.

Методы классификации

Классификация и кластеризация. Поиск признаковых зависимостей и частых множеств. Мультимодальная кластеризация.

Методы кластеризации

Предиктивные алгоритмы. Основные методы кластеризации. Прогнозирование и визуализация данных

Введение в машинное обучение

Итоговое контрольное мероприятие по темам "Сбор, хранение, обработка, извлечение и анализ больших потоков данных" и "Методы классификации и кластеризация"

Введение в разработку данных

Основные аналитические методы обработки данных. Машинное обучение и майнинг больших данных (Big Data). Нейронные сети как реализация алгоритмов машинного обучения.

Способы визуализации, интерпретации, проектная организация и бизнес рекомендации

Применение технологий больших данных для задач управления в банковской, страховой, финансовой индустриях, в масс медиа,

FinTech как область применения бизнес анализа. Способы визуального представления данных. Методы визуализации. Когнитивная информатика, экономика знаний, требования к специалистам в области интеллектуальной обработки данных для бизнеса

Контрольная точка 1

Контрольная работа включает проверку знаний по темам Интеллектуальный анализ данных, Сбор, хранение, обработка, извлечение и анализ больших потоков данных

Контрольная точка 2

Контрольная работа по темам Методы классификации и прогнозирования, Введение в машинное обучение и разработку данных

Итоговое контрольное мероприятие

Итоговое контрольное мероприятие по всем темам дисциплины

6. Методические указания для обучающихся по освоению дисциплины

Освоение дисциплины требует систематического изучения всех тем в той последовательности, в какой они указаны в рабочей программе.

Основными видами учебной работы являются аудиторные занятия. Их цель - расширить базовые знания обучающихся по осваиваемой дисциплине и систему теоретических ориентиров для последующего более глубокого освоения программного материала в ходе самостоятельной работы. Обучающемуся важно помнить, что контактная работа с преподавателем эффективно помогает ему овладеть программным материалом благодаря расстановке необходимых акцентов и удержанию внимания интонационными модуляциями голоса, а также подключением аудио-визуального механизма восприятия информации.

Самостоятельная работа преследует следующие цели:

- закрепление и совершенствование теоретических знаний, полученных на лекционных занятиях;
- формирование навыков подготовки текстовой составляющей информации учебного и научного назначения для размещения в различных информационных системах;
- совершенствование навыков поиска научных публикаций и образовательных ресурсов, размещенных в сети Интернет;
- самоконтроль освоения программного материала.

Обучающемуся необходимо помнить, что результаты самостоятельной работы контролируются преподавателем во время проведения мероприятий текущего контроля и учитываются при промежуточной аттестации.

Обучающимся с ОВЗ и инвалидов предоставляется возможность выбора форм проведения мероприятий текущего контроля, альтернативных формам, предусмотренным рабочей программой дисциплины. Предусматривается возможность увеличения в пределах 1 академического часа времени, отводимого на выполнение контрольных мероприятий.

Процедура оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья по дисциплине предусматривает предоставление информации в формах, адаптированных к ограничениям их здоровья и восприятия информации.

При проведении текущего контроля применяются оценочные средства, обеспечивающие передачу информации, от обучающегося к преподавателю, с учетом психофизиологических особенностей здоровья обучающихся.

7. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине

При самостоятельной работе обучающимся следует использовать:

- конспекты лекций;
- литературу из перечня основной и дополнительной учебной литературы, необходимой для освоения дисциплины (модуля);
- текст лекций на электронных носителях;
- ресурсы информационно-телекоммуникационной сети "Интернет", необходимые для освоения дисциплины;
- лицензионное и свободно распространяемое программное обеспечение из перечня информационных технологий, используемых при осуществлении образовательного процесса по дисциплине;
- методические указания для обучающихся по освоению дисциплины.

8. Перечень основной и дополнительной учебной литературы

Основная:

1. Анализ данных : учебник для академического бакалавриата / В. С. Мхитарян [и др.] ; под редакцией В. С. Мхитаряна. — Москва : Издательство Юрайт, 2019. — 490 с. — (Бакалавр. Академический курс). — ISBN 978-5-534-00616-2. — Текст : электронный // ЭБС Юрайт [сайт]. <https://www.urait.ru/bcode/432178>
2. Орешков, В. И. Интеллектуальный анализ данных : учебное пособие / В. И. Орешков. — Рязань : Рязанский государственный радиотехнический университет, 2017. — 161 с. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. <https://www.iprbookshop.ru/121842>
3. Неделько, В. М. Основы статистических методов машинного обучения : учебное пособие / В. М. Неделько. — Новосибирск : Новосибирский государственный технический университет, 2010. — 72 с. — ISBN 978-5-7782-1385-2. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. <http://www.iprbookshop.ru/45418>
4. Просто о больших данных:перевод с английского/Д. Гурвиц, А. Ньюджент, Ф. Халпер, М. Кауфман.- Москва:Эксмо,2015, ISBN 978-5-699-85807-1.-400.-Глоссарий: с. 354-368. - Указ.: с. 369-391

Дополнительная:

1. Пальмов, С. В. Интеллектуальный анализ данных : учебное пособие / С. В. Пальмов. — Самара : Поволжский государственный университет телекоммуникаций и информатики, 2017. — 127 с. — ISBN 2227-8397. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. <http://www.iprbookshop.ru/75376.html>
2. Целых, А. Н. Применение временных рядов для анализа больших данных : учебное пособие по курсу «Математические методы анализа больших данных» / А. Н. Целых, В. С. Васильев, Э. М. Котов. — Ростов-на-Дону, Таганрог : Издательство Южного федерального университета, 2021. — 84 с. — ISBN 978-5-9275-3983-3. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. <https://www.iprbookshop.ru/121929>
3. Маккинли, Уэс Python и анализ данных / Уэс Маккинли ; перевод А. Слинкина. — 2-е изд. — Саратов : Профобразование, 2019. — 482 с. — ISBN 978-5-4488-0046-7. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. <http://www.iprbookshop.ru/88752.html>
4. Адлер, Ю. П. Статистическое управление процессами. «Большие данные» : учебное пособие / Ю. П. Адлер, Е. А. Черных. — Москва : Издательский Дом МИСиС, 2016. — 52 с. — ISBN 978-5-87623-969-3. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. <http://www.iprbookshop.ru/64199>

9. Перечень ресурсов сети Интернет, необходимых для освоения дисциплины

<http://repo.ssau.ru/handle/Uchebnye-izdaniya/Laboratornyi-praktikum-po-kursu-%C2%ABIntellektualnyi-analiz-dannyh%C2%BB-Elektronnyi-resurs-praktikum-78475>

Интеллектуальный анализ данных

[http://wiki.cs.hse.ru/%D0%92%D0%B2%D0%B5%D0%B4%D0%B5%D0%BD%D0%B8%D0%B5_%D0%B2_%D0%B0%D0%BD%D0%B0%D0%BB%D0%B8%D0%B7_%D0%B4%D0%B0%D0%BD%D0%BD%D1%8B%D1%85_\(%D0%BC%D0%B0%D0%B9%D0%BD%D0%BE%D1%80_%D0%98%D0%90%D0%94\)](http://wiki.cs.hse.ru/%D0%92%D0%B2%D0%B5%D0%B4%D0%B5%D0%BD%D0%B8%D0%B5_%D0%B2_%D0%B0%D0%BD%D0%B0%D0%BB%D0%B8%D0%B7_%D0%B4%D0%B0%D0%BD%D0%BD%D1%8B%D1%85_(%D0%BC%D0%B0%D0%B9%D0%BD%D0%BE%D1%80_%D0%98%D0%90%D0%94)) Анализ данных

10. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине

Образовательный процесс по дисциплине **Математические методы и технологии обработки больших данных** предполагает использование следующего программного обеспечения и информационных справочных систем:

- доступ в режиме on-line в Электронную библиотечную систему (ЭБС);
- доступ в электронную информационно-образовательную среду университета
- Интернет-сервисы и электронные ресурсы

Необходимое лицензионное и свободно распространяемое программное обеспечение:

- офисный пакет приложений

При освоении материала и выполнения заданий по дисциплине рекомендуется использование материалов, размещенных в Личных кабинетах обучающихся ЕТИС ПГНИУ (**student.psu.ru**).

При организации дистанционной работы и проведении занятий в режиме онлайн могут использоваться:

система видеоконференцсвязи на основе платформы BigBlueButton (<https://bigbluebutton.org/>).

система LMS Moodle (<http://e-learn.psu.ru/>), которая поддерживает возможность использования текстовых материалов и презентаций, аудио- и видеоконтент, а так же тесты, проверяемые задания, задания для совместной работы.

система тестирования Indigo (<https://indigotech.ru/>).

11. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине

1. Лекционные занятия - аудитория, оснащенная презентационной техникой (проектор, экран, компьютер/ноутбук) с соответствующим программным обеспечением, меловой (и) или маркерной доской.
2. Занятий семинарского типа (практические занятия) - аудитория, оснащенная презентационной техникой (проектор, экран, компьютер/ноутбук) с соответствующим программным обеспечением, меловой (и) или маркерной доской.
3. Лабораторные занятия - Компьютерный класс, оснащенный персональными ЭВМ и соответствующим программным обеспечением. Состав оборудования определен в Паспорте компьютерного класса.
4. Самостоятельная работа - аудитория для самостоятельной работы, оснащенная компьютерной техникой с возможностью подключения к сети «Интернет», обеспеченная доступом в электронную информационно-образовательную среду университета. Помещения Научной библиотеки ПГНИУ.
5. Текущий контроль и промежуточная аттестация - аудитория, оснащенная меловой (и) или маркерной доской.
6. Индивидуальные и групповые консультации - аудитория, оснащенная презентационной техникой

(проектор, экран, компьютер/ноутбук) с соответствующим программным обеспечением, меловой (и) или маркерной доской или аудитория, оснащенная меловой (и) или маркерной доской.

Помещения научной библиотеки ПГНИУ для обеспечения самостоятельной работы обучающихся:

1. Научно-библиографический отдел, корп.1, ауд. 142. Оборудован 3 персональными компьютера с доступом к локальной и глобальной компьютерным сетям.

2. Читальный зал гуманитарной литературы, корп. 2, ауд. 418. Оборудован 7 персональными компьютерами с доступом к локальной и глобальной компьютерным сетям.

3. Читальный зал естественной литературы, корп.6, ауд. 107а. Оборудован 5 персональными компьютерами с доступом к локальной и глобальной компьютерным сетям.

4. Отдел иностранной литературы, корп.2 ауд. 207. Оборудован 1 персональным компьютером с доступом к локальной и глобальной компьютерным сетям.

5. Библиотека юридического факультета, корп.9, ауд. 4. Оборудована 11 персональными компьютерами с доступом к локальной и глобальной компьютерным сетям.

6. Читальный зал географического факультета, корп.8, ауд. 419. Оборудован 6 персональными компьютерами с доступом к локальной и глобальной компьютерным сетям.

Все компьютеры, установленные в помещениях научной библиотеки, оснащены следующим программным обеспечением:

Операционная система ALT Linux;

Офисный пакет Libreoffice.

Справочно-правовая система «КонсультантПлюс»

**Фонды оценочных средств для аттестации по дисциплине
Математические методы и технологии обработки больших данных**

**Планируемые результаты обучения по дисциплине для формирования компетенции.
Индикаторы и критерии их оценивания**

ПК.10

Способен к планированию и организации аналитических работ, в том числе, с использованием технологий больших данных

Компетенция (индикатор)	Планируемые результаты обучения	Критерии оценивания результатов обучения
ПК.10.3 Определяет необходимые ресурсы для проведения аналитических работ	Знает программные компоненты для проведения аналитических работ	Неудовлетворител НЕУДОВЛЕТВОРИТЕЛЬНО выставляется студенту, который не знает программные компоненты для проведения аналитических работ. Заявленная часть компетенции в полной мере не сформирована. Имеющихся знаний, умений, навыков недостаточно для решения практических задач. Требуется повторное обучение. Удовлетворительн УДОВЛЕТВОРИТЕЛЬНО выставляется студенту, который не уверенно знает программные компоненты для проведения аналитических работ. Сформированы знания и умения в области анализа данных, необходимые для дальнейшего обучения. Сформированность заявленной части компетенции соответствует минимальным требованиям. Хорошо ХОРОШО выставляется студенту, который в целом хорошо знает программные компоненты для проведения аналитических работ. Сформировано умение использовать соответствующие программные инструменты в стандартных ситуациях. Компетенция в целом соответствует требованиям. Отлично ОТЛИЧНО выставляется студенту, который отлично знает программные компоненты для проведения аналитических работ. Сформированы систематические знания и умения в области ИИ. Обучающийся свободно справляется с практическими

Компетенция (индикатор)	Планируемые результаты обучения	Критерии оценивания результатов обучения
		Отлично задачами, владеет разносторонними приемами выполнения соответствующий видов работ. Сформированность заявленной части компетенции превышает стандартные требования. Имеющихся знаний, умений, навыков и мотивации в полной мере достаточно для применения творческого подхода к решению сложных практических задач.
ПК.10.1 Разрабатывает, обсуждает и утверждает содержание аналитических работ, в том числе, с использованием технологий больших данных	Знает методику проведения анализа данных средствами технологий больших данных	Неудовлетворител Выставляется студенту, который не знает методику проведения анализа данных (Big Data). Заявленная часть компетенции в полной мере не сформирована. Имеющихся знаний, умений, навыков недостаточно для решения практических задач. Требуется повторное обучение. Удовлетворительн Выставляется студенту, который не уверенно знает методику проведения анализа данных. Сформированы знания и умения в области Big Data, необходимые для дальнейшего обучения. Сформированность заявленной части компетенции соответствует минимальным требованиям. Хорошо Выставляется студенту, который в целом хорошо знает методику проведения анализа данных. Сформировано умение использовать соответствующие программные инструменты в стандартных ситуациях. Компетенция в целом соответствует требованиям. Отлично Выставляется студенту, который отлично знает методику проведения анализа данных. Сформированы систематические знания и умения в области Big Data. Обучающийся свободно справляется с практическими задачами, владеет разносторонними приемами выполнения соответствующий видов работ. Сформированность заявленной части компетенции превышает стандартные

Компетенция (индикатор)	Планируемые результаты обучения	Критерии оценивания результатов обучения
		Отлично требования. Имеющихся знаний, умений, навыков и мотивации в полной мере достаточно для применения творческого подхода к решению сложных практических задач.

Оценочные средства текущего контроля и промежуточной аттестации

Схема доставки : набор 2024

Вид мероприятия промежуточной аттестации : Экзамен

Способ проведения мероприятия промежуточной аттестации : Оценка по дисциплине в рамках промежуточной аттестации определяется на основе баллов, набранных обучающимся на контрольных мероприятиях, проводимых в течение учебного периода.

Максимальное количество баллов : 100

Конвертация баллов в отметки

«отлично» - от 81 до 100

«хорошо» - от 61 до 80

«удовлетворительно» - от 43 до 60

«неудовлетворительно» / «незачтено» менее 43 балла

Компетенция (индикатор)	Мероприятие текущего контроля	Контролируемые элементы результатов обучения
ПК.10.3 Определяет необходимые ресурсы для проведения аналитических работ ПК.10.1 Разрабатывает, обсуждает и утверждает содержание аналитических работ, в том числе, с использованием технологий больших данных	Контрольная точка 1 Письменное контрольное мероприятие	Знает и умеет применять термины и задачи анализа больших данных; современные программные средства анализа больших данных
ПК.10.3 Определяет необходимые ресурсы для проведения аналитических работ ПК.10.1 Разрабатывает, обсуждает и утверждает содержание аналитических работ, в том числе, с использованием технологий больших данных	Контрольная точка 2 Письменное контрольное мероприятие	Знает и умеет выбирать для последующего анализа применять основные аналитические методы обработки данных, машинное обучение и майнинг больших данных (Big Data), нейронные сети.

Компетенция (индикатор)	Мероприятие текущего контроля	Контролируемые элементы результатов обучения
ПК.10.3 Определяет необходимые ресурсы для проведения аналитических работ ПК.10.1 Разрабатывает, обсуждает и утверждает содержание аналитических работ, в том числе, с использованием технологий больших данных	Итоговое контрольное мероприятие Итоговое контрольное мероприятие	Знает и умеет применять основные понятия анализа больших данных: «данные», «информация», «знания», понятие машинного обучения и его компоненты; основные определения и термины машинного обучения; виды задач в машинном обучении и процесс их решения; обучение без учителя: кластерный анализ, интеллектуальный анализ больших данных, современные программные средства для анализа больших данных.

Спецификация мероприятий текущего контроля

Контрольная точка 1

Продолжительность проведения мероприятия промежуточной аттестации: **2 часа**

Условия проведения мероприятия: **в часы самостоятельной работы**

Максимальный балл, выставляемый за мероприятие промежуточной аттестации: **30**

Проходной балл: **12.5**

Показатели оценивания	Баллы
Верно решенное задание	5

Контрольная точка 2

Продолжительность проведения мероприятия промежуточной аттестации: **2 часа**

Условия проведения мероприятия: **в часы самостоятельной работы**

Максимальный балл, выставляемый за мероприятие промежуточной аттестации: **30**

Проходной балл: **12.5**

Показатели оценивания	Баллы
Верно решенное задание	5

Итоговое контрольное мероприятие

Продолжительность проведения мероприятия промежуточной аттестации: **2 часа**

Условия проведения мероприятия: **в часы самостоятельной работы**

Максимальный балл, выставляемый за мероприятие промежуточной аттестации: **40**

Проходной балл: **17.5**

Показатели оценивания	Баллы
Верно решенное задание	5