

МИНОБРНАУКИ РОССИИ

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

Институт компьютерных наук и технологий

Авторы-составители: Кнутова Наталия Сергеевна

Рабочая программа дисциплины

**ТРЕК "ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И БОЛЬШИЕ ДАННЫЕ
(РЕГРЕССИОННЫЙ АНАЛИЗ)"**

Код УМК 100570

Утверждено
Протокол №1
от «28» июня 2024 г.

Пермь, 2024

1. Наименование дисциплины

Трек "Искусственный интеллект и большие данные (Регрессионный анализ)"

2. Место дисциплины в структуре образовательной программы

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

Направление подготовки: **01.03.02** Прикладная математика и информатика
направленность Искусственный интеллект и большие данные

3. Планируемые результаты обучения по дисциплине

В результате освоения дисциплины Трек "**Искусственный интеллект и большие данные (Регрессионный анализ)**" у обучающегося должны быть сформированы следующие компетенции:

01.03.02 Прикладная математика и информатика (направленность : Искусственный интеллект и большие данные)

ОПК.1 Способен применять фундаментальные знания, полученные в области математических и (или) естественных наук, и использовать их в профессиональной деятельности

Индикаторы

ОПК.1.2 Осуществляет первичный сбор и анализ материала, интерпретирует различные математические и физические объекты

ПК.1 Способен проводить работы по сбору, обработке и анализу информации и результатов исследований в предметной области

Индикаторы

ПК.1.1 Применяет актуальную нормативную документацию в соответствующей области знаний (обладает базовыми знаниями, полученными в области математических и (или) естественных наук)

ПК.1.2 Применяет методы анализа научных данных, методы и средства планирования и организации исследований и разработок, в том числе с использованием пакетов прикладных программ

ПК.1.3 Осуществляет теоретическое обобщение научных данных, результатов экспериментов и наблюдений

4. Объем и содержание дисциплины

Направление подготовки	01.03.02 Прикладная математика и информатика (направленность: Искусственный интеллект и большие данные)
форма обучения	очная
№.№ семестров, выделенных для изучения дисциплины	7
Объем дисциплины (з.е.)	3
Объем дисциплины (ак.час.)	108
Контактная работа с преподавателем (ак.час.), в том числе:	52
Проведение лекционных занятий	18
Проведение практических занятий, семинаров	34
Самостоятельная работа (ак.час.)	56
Формы текущего контроля	Защищаемое контрольное мероприятие (3) Итоговое контрольное мероприятие (1)
Формы промежуточной аттестации	Экзамен (7 семестр)

5. Аннотированное описание содержания разделов и тем дисциплины

Регрессионный анализ

Раздел 1 Анализ однофакторной регрессионной модели.

В данной теме рассматриваются основы регрессионного анализа с одной независимой переменной. Изучаются методы построения и интерпретации простой линейной регрессии, проверка значимости коэффициентов, анализ остатков и оценка качества модели. Особое внимание уделяется графическим методам визуализации зависимости между переменными и диагностике нарушений предположений регрессионного анализа (линейность, гомоскедастичность, нормальность остатков).

Раздел 2 Анализ множественной линейной регрессионной модели.

Тема посвящена расширению регрессионного анализа на случай нескольких независимых переменных. Рассматриваются методы оценки параметров модели (МНК), интерпретация коэффициентов регрессии, мультиколлинеарность и способы её выявления (VIF, корреляционная матрица). Изучаются критерии качества модели, а также методы отбора значимых предикторов (пошаговая регрессия, методы регуляризации).

Раздел 3 Различные аспекты множественной регрессии.

В рамках темы исследуются расширенные возможности множественной регрессии, включая работу с категориальными переменными (фиктивные переменные), взаимодействия между предикторами и нелинейные эффекты. Рассматриваются методы диагностики модели (анализ выбросов, влияние наблюдений, графики остатков), а также способы улучшения модели (преобразования переменных, устранение гетероскедастичности).

Раздел 4 Некоторые обобщения множественной регрессии.

Тема охватывает расширенные модели регрессии, выходящие за рамки классической линейной регрессии. Изучаются обобщённые линейные модели (GLM), логистическая регрессия для бинарных данных, пуассоновская регрессия для счётных данных, а также устойчивые (робастные) методы регрессии. Рассматриваются случаи, когда нарушаются стандартные предположения (например, автокорреляция в временных рядах).

Раздел 5 Прогнозирование в регрессионных моделях.

В данной теме рассматриваются методы использования регрессионных моделей для прогнозирования. Изучаются оценка точности прогнозов (MSE, MAE, MAPE), разбиение данных на обучающую и тестовую выборки, кросс-валидация. Особое внимание уделяется проблемам переобучения, методам регуляризации (Ridge, Lasso) и сравнению альтернативных моделей прогнозирования.

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. Афанасьев В. Н., Юзбашев М. М., Гуляева Т. И. Эконометрика: учебник для вузов / В. Н. Афанасьев, М. М. Юзбашев, Т. И. Гуляева ; ред. В. Н. Афанасьев. - Москва: Финансы и статистика, 2006, ISBN 5-279-02738-3. - 256. - Библиогр.: с. 254-255

Дополнительная:

1. Дрейпер Норман, Смит Г. Прикладной регрессионный анализ / Пер. с англ., науч. ред. и предисл. Ю. П. Адлера, В. Г. Горского. - М.: Статистика, 1973. - 391.

2. Вуколов Э. А. Основы статистического анализа. Практикум по статистическим методам и исследованию операций с использованием пакетов STATISTICA и EXCEL: учебное пособие по специальности "Менеджмент организации" / Э. А. Вуколов. - Москва: ФОРУМ, 2004, ISBN 5-8199-0141-X. - 464. - Библиогр.: с. 454-455

3. Бабешко Л. О. Основы эконометрического моделирования: учебное пособие для студентов, обучающихся по специальностям "Бухгалтерский учет, анализ и аудит", "Финансы и кредит", "Мировая экономика", "Налоги и налогообложение" / Л. О. Бабешко. - Москва: КомКнига, 2006, ISBN 5-484-00757-7. - 432. - Библиогр.: с. 428

9. Перечень ресурсов сети Интернет, необходимых для освоения дисциплины

При освоении дисциплины использование ресурсов сети Интернет не предусмотрено.

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

Образовательный процесс по дисциплине **Трек "Искусственный интеллект и большие данные (Регрессионный анализ)"** предполагает использование следующего программного обеспечения и информационных справочных систем:

Образовательный процесс по дисциплине 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.

Необходимое лицензионное и (или) свободно распространяемое программное обеспечение:

- приложение позволяющее просматривать и воспроизводить медиаконтент PDF-файлов «Adobe Acrobat Reader DC»;
- офисный пакет приложений «LibreOffice»;

Специализированное программное обеспечение не требуется

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

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

Для лекционных занятий требуется аудитория, оснащенная презентационной техникой (проектор, экран, компьютер/ноутбук) с соответствующим программным обеспечением, меловой (и) или маркерной доской.

Для проведения практических занятий - аудитория, оснащенная презентационной техникой (проектор, экран, компьютер/ноутбук) с соответствующим программным обеспечением, меловой (и) или маркерной доской.

Для проведения лабораторных занятий - компьютерный класс. Состав оборудования определен в Паспорте компьютерного класса.

Для групповых (индивидуальных) консультаций - аудитория, оснащенная презентационной техникой (проектор, экран, компьютер/ноутбук) с соответствующим программным обеспечением, меловой (и) или маркерной доской.

Для проведения текущего контроля - аудитория, оснащенная меловой (и) или маркерной доской.

Для самостоятельной работы студентов - аудитория, оснащенная компьютерной техникой с возможностью подключения к сети «Интернет», с обеспеченным доступом в электронную информационно-образовательную среду университета, помещения Научной библиотеки ПГНИУ.

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
Трек "Искусственный интеллект и большие данные (Регрессионный анализ)"**

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

ОПК.1

Способен применять фундаментальные знания, полученные в области математических и (или) естественных наук, и использовать их в профессиональной деятельности

Компетенция (индикатор)	Планируемые результаты обучения	Критерии оценивания результатов обучения
ОПК.1.2 Осуществляет первичный сбор и анализ материала, интерпретирует различные математические и физические объекты	Навыки работы с различными источниками данных	<p style="text-align: center;">Неудовлетворительно</p> <p>Не выполнены условия на "удовлетворительно"</p> <p style="text-align: center;">Удовлетворительно</p> <p>Не выполнены условия на "хорошо". Студент допускает большое количество ошибок, которое может исправить самостоятельно или с помощью преподавателя</p> <p style="text-align: center;">Хорошо</p> <p>Не выполнены условия на "отлично". Студент допускает небольшое количество ошибок, которые может исправить самостоятельно</p> <p style="text-align: center;">Отлично</p> <p>Студент владеет методами очистки и предобработки данных для регрессионного анализа, самостоятельно выявляет и корректирует аномалии в исходных данных</p>

ПК.1

Способен проводить работы по сбору, обработке и анализу информации и результатов исследований в предметной области

Компетенция (индикатор)	Планируемые результаты обучения	Критерии оценивания результатов обучения
ПК.1.1 Применяет актуальную нормативную документацию в соответствующей области знаний (обладает базовыми знаниями, полученными в области математических и (или) естественных наук)	Применение международных стандартов статистического анализа (ISO, ГОСТ) при обработке данных, соблюдение принципов воспроизводимости исследований и документации методов анализа	<p style="text-align: center;">Неудовлетворительно</p> <p>Не выполнены условия на "удовлетворительно"</p> <p style="text-align: center;">Удовлетворительно</p> <p>Не выполнены условия на "хорошо". Студент допускает большое количество ошибок, которое может исправить самостоятельно или с помощью преподавателя</p> <p style="text-align: center;">Хорошо</p> <p>Не выполнены условия на "отлично". Студент допускает небольшое количество ошибок, которые может исправить</p>

Компетенция (индикатор)	Планируемые результаты обучения	Критерии оценивания результатов обучения
		<p>Хорошо самостоятельно</p> <p>Отлично Студент демонстрирует понимание взаимосвязи теоретических основ и практического применения регрессионного анализа, способен работать в рамках нормативных требований к статистическому анализу, владеет навыками корректного оформления и представления результатов исследования</p>
<p>ПК.1.2 Применяет методы анализа научных данных, методы и средства планирования и организации исследований и разработок, в том числе с использованием пакетов прикладных программ</p>	<p>Практическое владение специализированными пакетами (R, Python, SPSS, STATISTICA)</p>	<p>Неудовлетворительно Не выполнены условия на "удовлетворительно"</p> <p>Удовлетворительно Не выполнены условия на "хорошо". Студент допускает большое количество ошибок, которое может исправить самостоятельно или с помощью преподавателя</p> <p>Хорошо Не выполнены условия на "отлично". Студент допускает небольшое количество ошибок, которые может исправить самостоятельно</p> <p>Отлично Студент самостоятельно проводит полный цикл статистического исследования, применяя современные вычислительные средства для анализа данных</p>
<p>ПК.1.3 Осуществляет теоретическое обобщение научных данных, результатов экспериментов и наблюдений</p>	<p>Способность формулировать теоретические положения на основе эмпирических данных</p>	<p>Неудовлетворительно Не выполнены условия на "удовлетворительно"</p> <p>Удовлетворительно Не выполнены условия на "хорошо". Студент допускает большое количество ошибок, которое может исправить самостоятельно или с помощью преподавателя</p> <p>Хорошо Не выполнены условия на "отлично". Студент допускает небольшое количество ошибок, которые может исправить самостоятельно</p> <p>Отлично Студент самостоятельно формулирует практические рекомендации на основе анализа моделей, прогнозирует поведение систем на основе регрессионных</p>

Компетенция (индикатор)	Планируемые результаты обучения	Критерии оценивания результатов обучения
		Отлично зависимостей

Оценочные средства текущего контроля и промежуточной аттестации

Схема доставки : Базовая

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 44 до 60

«неудовлетворительно» / «незачтено» менее 44 балла

Компетенция (индикатор)	Мероприятие текущего контроля	Контролируемые элементы результатов обучения
ОПК.1.2 Осуществляет первичный сбор и анализ материала, интерпретирует различные математические и физические объекты ПК.1.3 Осуществляет теоретическое обобщение научных данных, результатов экспериментов и наблюдений	Раздел 2 Анализ множественной линейной регрессионной модели. Защищаемое контрольное мероприятие	анализ множественной линейной регрессионной модели с использованием статистических пакетов
ПК.1.2 Применяет методы анализа научных данных, методы и средства планирования и организации исследований и разработок, в том числе с использованием пакетов прикладных программ	Раздел 3 Различные аспекты множественной регрессии. Защищаемое контрольное мероприятие	понимание теоретических основ множественной регрессии, навыки работы с критериями выбора моделей
ПК.1.2 Применяет методы анализа научных данных, методы и средства планирования и организации исследований и разработок, в том числе с использованием пакетов прикладных программ ПК.1.3 Осуществляет теоретическое обобщение научных данных, результатов экспериментов и наблюдений	Раздел 4 Некоторые обобщения множественной регрессии. Защищаемое контрольное мероприятие	проверка гипотез

Компетенция (индикатор)	Мероприятие текущего контроля	Контролируемые элементы результатов обучения
ОПК.1.2 Осуществляет первичный сбор и анализ материала, интерпретирует различные математические и физические объекты ПК.1.1 Применяет актуальную нормативную документацию в соответствующей области знаний (обладает базовыми знаниями, полученными в области математических и (или) естественных наук) ПК.1.2 Применяет методы анализа научных данных, методы и средства планирования и организации исследований и разработок, в том числе с использованием пакетов прикладных программ ПК.1.3 Осуществляет теоретическое обобщение научных данных, результатов экспериментов и наблюдений	Раздел 5 Прогнозирование в регрессионных моделях. Итоговое контрольное мероприятие	тест по всем разделам дисциплины

Спецификация мероприятий текущего контроля

Раздел 2 Анализ множественной линейной регрессионной модели.

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

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

Максимальный балл, выставляемый за мероприятие промежуточной аттестации: **20**

Проходной балл: **9**

Показатели оценивания	Баллы
Устная защита результатов	8
Отчет с анализом и выводами	7
Практическое задание в статистических пакетах	5

Раздел 3 Различные аспекты множественной регрессии.

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

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

Максимальный балл, выставляемый за мероприятие промежуточной аттестации: **20**

Проходной балл: **9**

Показатели оценивания	Баллы
-----------------------	-------

Самостоятельная работа "Модели множественной регрессии"	20

Раздел 4 Некоторые обобщения множественной регрессии.

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

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

Максимальный балл, выставляемый за мероприятие промежуточной аттестации: **20**

Проходной балл: **9**

Показатели оценивания	Баллы
Самостоятельная работа "Проверка гипотезы о гетероскедастичности ошибок"	10
Самостоятельная работа "Проверка гипотезы об ограничении"	5
Самостоятельная работа "Проверка гипотезы о случайности и нормальности остатков с помощью критерия Жарка-Бэра"	5

Раздел 5 Прогнозирование в регрессионных моделях.

Продолжительность проведения мероприятия промежуточной аттестации: **1 часа**

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

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

Проходной балл: **17**

Показатели оценивания	Баллы
Тест "Регрессионный анализ"	40