

МИНОБРНАУКИ РОССИИ

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

Кафедра информационных систем и математических методов в экономике

Авторы-составители: **Ильин Иван Вадимович**
Радионова Марина Владимировна
Гарафутдинов Роберт Викторович
Васёва Галина Сергеевна

Рабочая программа дисциплины
**СОВРЕМЕННЫЕ ТЕХНОЛОГИИ МАШИННОГО ОБУЧЕНИЯ И
ИСКУССТВЕННОГО ИНТЕЛЛЕКТА**

Код УМК 101465

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

Пермь, 2024

1. Наименование дисциплины

Современные технологии машинного обучения и искусственного интеллекта

2. Место дисциплины в структуре образовательной программы

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

Специальность: **10.05.04** Информационно-аналитические системы безопасности
специализация Информационная безопасность финансовых и экономических структур

3. Планируемые результаты обучения по дисциплине

В результате освоения дисциплины **Современные технологии машинного обучения и искусственного интеллекта** у обучающегося должны быть сформированы следующие компетенции:

10.05.04 Информационно-аналитические системы безопасности (специализация : Информационная безопасность финансовых и экономических структур)

ОПК.2.4 Способен разрабатывать и применять автоматизированные технологии обработки больших информационных потоков (массивов) финансовой и/или экономической информации в режиме реального времени

Индикаторы

ОПК.2.4.1 Формирует разделы технических заданий на создание автоматизированных технологий обработки больших массивов экономических и финансовых данных

ОПК.2.4.2 Применяет автоматизированные технологии обработки больших информационных потоков (массивов) финансовой и/или экономической информации в режиме реального времени

4. Объем и содержание дисциплины

Специальность	10.05.04 Информационно-аналитические системы безопасности (специализация: Информационная безопасность финансовых и экономических структур)
форма обучения	очная
№№ семестров, выделенных для изучения дисциплины	6
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 лабораторных работ, занятий по иностранному языку	28
Самостоятельная работа (ак.час.)	88
Формы текущего контроля	Входное тестирование (1) Защищаемое контрольное мероприятие (2) Итоговое контрольное мероприятие (1)
Формы промежуточной аттестации	Экзамен (6 семестр)

5. Аннотированное описание содержания разделов и тем дисциплины

Тема 1. Введение в искусственный интеллект и машинное обучение

Понятие ИИ. Этапы развития. Области ИИ. Кривая Гартнера. Интеграция ИИ в бизнес-процессы. Автоматизация и прогнозирование. Кейсы применения ИИ в бизнесе. Условия применения ИИ: данные, ресурсы, специалисты, метрики. Методология CRISP-DM. Понятие машинного обучения и его классификация. Основные задачи машинного обучения

Тема 2. Разведочный анализ данных

Предобработка данных. Качество данных: пропуски и выбросы. Нормализация данных. Выявление закономерностей в данных: доверительные интервалы, статистические гипотезы, корреляционный анализ.

Тема 3. Задача регрессии

Обучение с учителем: задача предсказания. Общая схема решения задачи предсказания. Деление набора данных на обучающую и тестовую выборки. Задача регрессии. Парная и множественная линейная регрессия. Полиномиальная регрессия. Регуляризация. Метрики качества регрессии.

Тема 4. Задача классификации

Бинарная и множественная классификация. Алгоритмы классификации. Гиперпараметры. Метрики качества классификации. Кросс-валидация и подбор гиперпараметров.

Тема 5. Задача кластеризации

Обучение без учителя: кластеризация. Расстояние между объектами. Алгоритмы кластеризации. Метрики качества кластеризации.

Тема 6. Ансамбли алгоритмов

Методы ансамблирования: комитет, стекинг, бустинг. Бустрэп и бэггинг. Лес решений. Блендинг. Адаптивный и градиентный бустинг.

Итоговое КМ

Итоговый тест

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 аудиторские 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. Целых, А. Н. Извлечение знаний методами машинного обучения : учебное пособие по курсам «Модели и методы инженерии знаний», «Методы машинного обучения» / А. Н. Целых, Э. М. Котов. — Ростов-на-Дону, Таганрог : Издательство Южного федерального университета, 2022. — 105 с. — ISBN 978-5-9275-4215-4. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт].
<https://www.iprbookshop.ru/131448>

Дополнительная:

1. Теория и практика машинного обучения : учебное пособие / В. В. Воронина, А. В. Михеев, Н. Г. Ярушкина, К. В. Святков. — Ульяновск : Ульяновский государственный технический университет, 2017. — 291 с. — ISBN 978-5-9795-1712-4. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. <https://www.iprbookshop.ru/106120.html>

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 **Современные технологии машинного обучения и искусственного интеллекта** предполагает использование следующего программного обеспечения и информационных справочных систем:

- 1) презентационные материалы (слайды по темам лекционных занятий);
- 2) доступ в режиме on-line в Электронную библиотечную систему (ЭБС);
- 3) доступ в электронную информационно-образовательную среду университета;
- 4) интернет-сервисы и электронные ресурсы.

Перечень необходимого лицензионного и (или) свободно распространяемого программного обеспечения:

- приложение позволяющее просматривать и воспроизводить медиаконтент PDF-файлов «AdobeAcrobatReader DC»;
- ОС "Альт Образование"
- Linux Mint

Специализированное программное обеспечение компьютерного класса.

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 специализированной мебелью, презентационной техникой (проектор, экран, компьютер/ноутбук) с соответствующим программным обеспечением, меловой (и) или маркерной доской.

Для проведения лабораторных занятий - Компьютерный класс со специализированным программным обеспечением.

Для групповых (индивидуальных) консультаций, текущего контроля и промежуточной аттестации - аудитория, оснащенная специализированной мебелью, презентационной техникой (проектор, экран, компьютер/ноутбук) с соответствующим программным обеспечением, 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
Современные технологии машинного обучения и искусственного интеллекта**

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

ОПК.2.4

Способен разрабатывать и применять автоматизированные технологии обработки больших информационных потоков (массивов) финансовой и/или экономической информации в режиме реального времени

Индикатор	Планируемые результаты обучения	Критерии оценивания результатов обучения
<p>ОПК.2.4.2 Применяет автоматизированные технологии обработки больших информационных потоков (массивов) финансовой и/или экономической информации в режиме реального времени</p>	<p>умение применять автоматизированные технологии обработки больших информационных потоков (массивов) финансовой и/или экономической информации в режиме реального времени, используя методы машинного обучения и искусственного интеллекта для анализа, интерпретации и визуализации данных.</p>	<p align="center">Неудовлетворител Студент не может эффективно применять технологии, неадекватно интерпретирует данные, не создает приемлемые визуализации и не демонстрирует умений работать с потоковыми данными.</p> <p align="center">Удовлетворительн Студент использует технологии, но результаты анализа имеют значительные недостатки, визуализации недостаточно информативны, а работа с потоковыми данными вызывает трудности.</p> <p align="center">Хорошо Студент применяет технологии и интерпретирует данные с незначительными ошибками, создаёт визуализации и работает в режиме реального времени.</p> <p align="center">Отлично Студент точно применяет технологии, глубоко анализирует и интерпретирует данные, создает информативные визуализации и успешно работает с потоковыми данными в режиме реального времени.</p>
<p>ОПК.2.4.1 Формирует разделы технических заданий на создание автоматизированных технологий обработки больших массивов экономических и финансовых данных</p>	<p>знает разделы технических заданий на создание автоматизированных технологий обработки больших массивов экономических и финансовых данных, понимает требования к данным, умеет делать выбор алгоритмов и технологий, анализ и обоснование архитектуры системы</p>	<p align="center">Неудовлетворител Студент не определяет требования к данным, не предлагает алгоритмов и технологий, не структурирует информацию и не предлагает базовые идеи архитектуры системы.</p> <p align="center">Удовлетворительн Студент демонстрирует базовое понимание требований к данным, выбор алгоритмов и технологий производит с ошибками, проводит поверхностный анализ архитектуры системы.</p>

Индикатор	Планируемые результаты обучения	Критерии оценивания результатов обучения
		<p style="text-align: center;">Хорошо</p> <p>Студент в целом понимает требования к данным, выбирает подходящие алгоритмы и технологии с незначительными недостатками в обосновании, предлагает архитектуру системы с неполным обоснованием</p> <p style="text-align: center;">Отлично</p> <p>Студент полностью и глубоко понимает требования к данным, безошибочно выбирает и обосновывает алгоритмы и технологии, излагает материал логично и последовательно, а также предлагает обоснованную архитектуру системы с анализом альтернатив.</p>

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 50 до 60

«неудовлетворительно» / «незачтено» менее 50 балла

Компетенция (индикатор)	Мероприятие текущего контроля	Контролируемые элементы результатов обучения
Входной контроль	Тема 1. Введение в искусственный интеллект и машинное обучение Входное тестирование	Остаточные знания по анализу данных и Питону
ОПК.2.4.1 Формирует разделы технических заданий на создание автоматизированных технологий обработки больших массивов экономических и финансовых данных	Тема 4. Задача классификации Защищаемое контрольное мероприятие	Знание основных метрик качества регрессии и классификации. Умение использовать распространенные библиотеки языка Python для решения задач машинного обучения (scikit-learn, statsmodels). Умение строить модели регрессии и классификации и оценивать их качество с помощью Python.
ОПК.2.4.2 Применяет автоматизированные технологии обработки больших информационных потоков (массивов) финансовой и/или экономической информации в режиме реального времени	Тема 6. Ансамбли алгоритмов Защищаемое контрольное мероприятие	Умение использовать распространенные библиотеки языка Python для решения задач машинного обучения (scikit-learn, statsmodels, XGBoost, CatBoost). Умение строить модели регрессии, классификации и кластеризации и оценивать их качество с помощью Python. Знание подходов к построению композиций моделей и алгоритмов ансамблирования. Уметь применять ансамблирование на практике в Python.

Компетенция (индикатор)	Мероприятие текущего контроля	Контролируемые элементы результатов обучения
ОПК.2.4.2 Применяет автоматизированные технологии обработки больших информационных потоков (массивов) финансовой и/или экономической информации в режиме реального времени ОПК.2.4.1 Формирует разделы технических заданий на создание автоматизированных технологий обработки больших массивов экономических и финансовых данных	Итоговое КМ Итоговое контрольное мероприятие	Итоговый тест

Спецификация мероприятий текущего контроля

Тема 1. Введение в искусственный интеллект и машинное обучение

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

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

Максимальный балл, выставляемый за мероприятие промежуточной аттестации: **0**

Проходной балл: **0**

Показатели оценивания	Баллы
«10 баллов» оценивается успешно написанный тест входного контроля (10 б.)	10
«7 баллов» оценивается написанный тест входного контроля (7 б.)	7
«5 баллов» оценивается удовлетворительно написанный тест входного контроля (5 б.)	5

Тема 4. Задача классификации

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

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

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

Проходной балл: **15**

Показатели оценивания	Баллы
«30 баллов» выставляется за сданные все практические и контрольные работы (тема методы классификации). Возможно добавление / снижение баллов за качество выполненных заданий	30
«20 баллов» выставляется за частично сданные практические и контрольные работы (темы методы классификации). Возможно добавление / снижение баллов за качество выполненных заданий	20
«15 баллов» выставляется за минимальный набор сданных практических и контрольных работ (темы методы классификации). Возможно добавление / снижение баллов за качество выполненных заданий	15

Тема 6. Ансамбли алгоритмов

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

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

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

Проходной балл: **15**

Показатели оценивания	Баллы
«30 баллов» выставляется за сданные все практические и контрольные работы (темы методы ансамблей). Возможно добавление / снижение баллов за качество выполненных заданий	30
«20 баллов» выставляется за частично сданные практические и контрольные работы (темы методы ансамблей). Возможно добавление / снижение баллов за качество выполненных заданий	20
«15 баллов» выставляется за минимальный набор сданных практических и контрольных работ (темы методы ансамблей). Возможно добавление / снижение баллов за качество выполненных заданий	15

Итоговое КМ

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

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

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

Проходной балл: **20**

Показатели оценивания	Баллы
«40 баллов» выставляется за сданные все практические и контрольные работы (темы включают все содержательные линии дисциплины). Возможно добавление / снижение баллов за качество выполненных заданий. А также за 80 % выполненных тестовых заданий	40
«30 баллов» выставляется за 60 % выполненных тестовых заданий. Также за частично сданные практические и контрольные работы (темы включают все содержательные линии дисциплины). Возможно добавление / снижение баллов за качество выполненных заданий	30
«20 баллов» выставляется за минимальный набор сданных практических и контрольных работ (темы включают все содержательные линии дисциплины). Возможно добавление / снижение баллов за качество выполненных заданий. А также за 50 % выполненных тестовых заданий	20