

МИНОБРНАУКИ РОССИИ

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

Физико-математический институт

**Авторы-составители: Радионова Марина Владимировна
Русакова Ольга Леонидовна**

Рабочая программа дисциплины
МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ
Код УМК 100226

Утверждено
Протокол №9
от «30» мая 2023 г.

Пермь, 2023

1. Наименование дисциплины

Методы машинного обучения

2. Место дисциплины в структуре образовательной программы

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

Направление подготовки: **05.03.03** Картография и геоинформатика
направленность Геоинформатика

3. Планируемые результаты обучения по дисциплине

В результате освоения дисциплины **Методы машинного обучения** у обучающегося должны быть сформированы следующие компетенции:

05.03.03 Картография и геоинформатика (направленность : Геоинформатика)

ОПК.2 Способен понимать принципы работы современных информационно-коммуникационных технологий и использовать их для решения профессиональных задач с учетом требований информационной безопасности

Индикаторы

ОПК.2.1 Демонстрирует базовые знания в области информационно-коммуникационных технологий

ОПК.2.2 Ориентируясь на задачи профессиональной деятельности, обоснованно выбирает информационно-коммуникационные технологии и использует их в профессиональной деятельности с учетом требований информационной безопасности

4. Объем и содержание дисциплины

Направление подготовки	05.03.03 Картография и геоинформатика (направленность: Геоинформатика)
форма обучения	очная
№№ триместров, выделенных для изучения дисциплины	8
Объем дисциплины (з.е.)	3
Объем дисциплины (ак.час.)	108
Контактная работа с преподавателем (ак.час.), в том числе:	42
Проведение лекционных занятий	14
Проведение лабораторных работ, занятий по иностранному языку	28
Самостоятельная работа (ак.час.)	66
Формы текущего контроля	Защищаемое контрольное мероприятие (2) Итоговое контрольное мероприятие (1)
Формы промежуточной аттестации	Зачет (8 триместр)

5. Аннотированное описание содержания разделов и тем дисциплины

Методы машинного обучения

Тема 1. Проверка статистических гипотез

Тестирование статистических гипотез. Основные понятия. Виды статистических гипотез. Ошибки I и II рода. Ошибка I рода. Уровень значимости и p-value. Основные принципы проверки статистических гипотез. Параметрические гипотезы. Критерии согласия. Критерии однородности.

Тема 2. Задачи классификации и кластеризации

Общая постановка задачи классификации. Метрики расстояния и близости между объектами, расстояния между кластерами. Метод k-средних. Визуализация. Дерево решений, случайный лес и ансамбли деревьев. Классификации без обучения, понятия кластерного анализа. Основные типы задач и алгоритмов кластерного анализа. Иерархические, параллельные и последовательные процедуры кластерного анализа. Визуализация

Тема 3. Методы корреляционного и регрессионного анализа

Основные показатели исследования взаимосвязей. Коэффициенты корреляции Пирсона и Спирмена. Таблицы сопряжённости и их интерпретация. Предпосылки линейной регрессии. Построение уравнения простой и множественной линейной регрессии. Оценка качества построенной модели.

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 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. 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>
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. Бессмертный, И. А. Интеллектуальные системы : учебник и практикум для академического бакалавриата / И. А. Бессмертный, А. Б. Нугуманова, А. В. Платонов. — Москва : Издательство Юрайт, 2019. — 243 с. — (Бакалавр. Академический курс). — ISBN 978-5-534-01042-8. — Текст : электронный // ЭБС Юрайт [сайт]. <https://www.urait.ru/bcode/433716>

9. Перечень ресурсов сети Интернет, необходимых для освоения дисциплины

<https://academy.yandex.ru/handbook/ml> Учебник по машинному обучению

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

Образовательный процесс по дисциплине **Методы машинного обучения** 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.
- IDE для Python версии 3 и выше.

Самостоятельная работа студентов организуется и поддерживается учебными материалами онлайн-курса, размещенного на платформе online.psu.ru

Доступ к курсу по ссылке

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 и маркерной доской.

Для проведения лабораторных занятий - компьютерный класс. Состав оборудования определен в Паспорте компьютерного класса (требуется доступ в Интернет и должна быть установлена IDE для Python).

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

Для проведения текущего контроля - компьютерный класс. Состав оборудования определен в Паспорте компьютерного класса (требуется доступ в Интернет и должна быть установлена IDE для Python).

Самостоятельная работа 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 компьютерами 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
Методы машинного обучения**

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

ОПК.2

Способен понимать принципы работы современных информационно-коммуникационных технологий и использовать их для решения профессиональных задач с учетом требований информационной безопасности

Компетенция (индикатор)	Планируемые результаты обучения	Критерии оценивания результатов обучения
<p>ОПК.2.1 Демонстрирует базовые знания в области информационно-коммуникационных технологий</p>	<p>Знание: - основных элементов языка программирования Python; - специальных структур данных и библиотек Python для обработки, анализа данных и построения моделей машинного обучения. Умение демонстрировать приобретённые знания для решения задач машинного обучения в конкретной прикладной области.</p>	<p align="center">Неудовлетворител</p> <p>Не знает: - современные методы машинного обучения; - библиотек Python для обработки, визуализации, анализа данных и построения моделей машинного обучения.. Не владеет: - основными приёмами первичной обработки исходных данных; - навыками использования современного инструментария Python для построения моделей машинного обучения.</p> <p align="center">Удовлетворительн</p> <p>Знает: - современные методы машинного обучения; - библиотек Python для обработки, визуализации, анализа данных и построения моделей машинного обучения.. Владеет: - основными приёмами подготовки исходных данных для построения моделей машинного обучения; - навыками использования современного инструментария Python для построения моделей машинного обучения, определения качества построенных моделей. Но может допускать существенные ошибки.</p> <p align="center">Хорошо</p> <p>Знает: - современные методы машинного обучения; - библиотек Python для обработки, визуализации, анализа данных и построения моделей машинного обучения.. Владеет:</p>

Компетенция (индикатор)	Планируемые результаты обучения	Критерии оценивания результатов обучения
		<p style="text-align: center;">Хорошо</p> <ul style="list-style-type: none"> - основными приёмами подготовки исходных данных для построения моделей машинного обучения; - навыками использования современного инструментария Python для построения моделей машинного обучения, определения качества построенных моделей и их оптимизации. <p>Но может допускать незначительные ошибки.</p> <p style="text-align: center;">Отлично</p> <p>Знает:</p> <ul style="list-style-type: none"> - современные методы машинного обучения; - библиотек Python для обработки, визуализации, анализа данных и построения моделей машинного обучения.. <p>Владеет:</p> <ul style="list-style-type: none"> - основными приёмами подготовки исходных данных для построения моделей машинного обучения; - навыками использования современного инструментария Python для построения моделей машинного обучения, определения качества построенных моделей и их оптимизации.
<p>ОПК.2.2 Ориентируясь на задачи профессиональной деятельности, обоснованно выбирает информационно-коммуникационные технологии и использует их в профессиональной деятельности с учетом требований информационной безопасности</p>	<p>Умение обоснованно выбирать информационно-коммуникационные технологии для решения задачи построения машинного обучения в рамках конкретной прикладной задачи</p>	<p style="text-align: center;">Неудовлетворител</p> <p>Не знает и не может, ориентируясь на конкретную задачу, выбрать информационно-коммуникационные технологии для решения задачи машинного обучения.</p> <p style="text-align: center;">Удовлетворительн</p> <p>Самостоятельно, ориентируясь на конкретную задачу, выбирает информационно-коммуникационные технологии для решения задачи машинного обучения. Но может допустить существенные ошибки.</p> <p style="text-align: center;">Хорошо</p> <p>Самостоятельно, ориентируясь на конкретную задачу, выбирает информационно-коммуникационные технологии для решения задачи машинного обучения. Но может допустить</p>

Компетенция (индикатор)	Планируемые результаты обучения	Критерии оценивания результатов обучения
		<p style="text-align: center;">Хорошо несущественные ошибки.</p> <p style="text-align: center;">Отлично Самостоятельно, ориентируясь на конкретную задачу, выбирает информационно-коммуникационные технологии для решения задачи машинного обучения.</p>

Оценочные средства текущего контроля и промежуточной аттестации

Схема доставки : Базовая

Вид мероприятия промежуточной аттестации : Зачет

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
ОПК.2.2 Ориентируясь на задачи профессиональной деятельности, обоснованно выбирает информационно-коммуникационные технологии и использует их в профессиональной деятельности с учетом требований информационной безопасности ОПК.2.1 Демонстрирует базовые знания в области информационно-коммуникационных технологий	Тема 2. Задачи классификации и кластеризации Защищаемое контрольное мероприятие	Знание основных методов классификации и кластеризации объектов. Умение:- реализовать решение задач классификации и кластеризации с помощью методов языка Python; - анализировать результаты работы выбранного метода;- подобрать оптимальные параметры метода.
ОПК.2.2 Ориентируясь на задачи профессиональной деятельности, обоснованно выбирает информационно-коммуникационные технологии и использует их в профессиональной деятельности с учетом требований информационной безопасности ОПК.2.1 Демонстрирует базовые знания в области информационно-коммуникационных технологий	Тема 3. Методы корреляционного и регрессионного анализа Защищаемое контрольное мероприятие	Знание основных понятий и методов корреляционного и регрессионного анализа. Умение:- применять методы языка Python для решения задачи корреляционного и регрессионного анализа;- оценить качество построенной модели машинного обучения;- интерпретировать результаты. Обязательным является прохождение теста на платформе https://testing.assessment.unionepro.ru/courses/65421ed3f501deb120e07101

Компетенция (индикатор)	Мероприятие текущего контроля	Контролируемые элементы результатов обучения
ОПК.2.2 Ориентируясь на задачи профессиональной деятельности, обоснованно выбирает информационно-коммуникационные технологии и использует их в профессиональной деятельности с учетом требований информационной безопасности ОПК.2.1 Демонстрирует базовые знания в области информационно-коммуникационных технологий	Итоговое контрольное мероприятие Итоговое контрольное мероприятие	Знание основных методов машинного обучения. Умение практически решать задачи машинного обучения.

Спецификация мероприятий текущего контроля

Тема 2. Задачи классификации и кластеризации

Продолжительность проведения мероприятия промежуточной аттестации: **4 часа**

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

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

Проходной балл: **13**

Показатели оценивания	Баллы
Задача классификации с помощью решающих деревьев	20
Задача кластеризации	10

Тема 3. Методы корреляционного и регрессионного анализа

Продолжительность проведения мероприятия промежуточной аттестации: **4 часа**

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

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

Проходной балл: **13**

Показатели оценивания	Баллы
Построение модели не менее трёх моделей. Для каждой модели подбор оптимальных параметров, оценка качества модели.	15
Отчёт	5
Подготовка данных для анализа (оценка полноты данных)	5
Анализ диаграммы рассеивания, коэффициента корреляции, коэффициента детерминации и их интерпретация.	5

Итоговое контрольное мероприятие

Продолжительность проведения мероприятия промежуточной аттестации: **4 часа**

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

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

Проходной балл: **17**

Показатели оценивания	Баллы
Алгоритм решения предложен. Возможные проблемы при решении задачи указаны.	10
Даны оценка качества построенной модели и интерпретация результатов.	10
Для решения задачи выбран подходящий инструментарий. Преимущества и недостатки инструментария перечислены.	10
Требования к исходным данным перечислены.	5
Прикладная задача классифицирована корректно.	5