

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

Авторы-составители: **Селетков Илья Павлович**

Рабочая программа дисциплины
МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ
Код УМК 98476

Утверждено
Протокол №9
от «18» мая 2022 г.

Пермь, 2022

1. Наименование дисциплины

Методы машинного обучения

2. Место дисциплины в структуре образовательной программы

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

Направление: **01.04.02** Прикладная математика и информатика
направленность Перспективные методы искусственного интеллекта в сетях передачи и обработки данных

3. Планируемые результаты обучения по дисциплине

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

01.04.02 Прикладная математика и информатика (направленность : Перспективные методы искусственного интеллекта в сетях передачи и обработки данных)

ПК.7 Способен руководить проектами по созданию, внедрению и использованию одной или нескольких сквозных цифровых технологий искусственного интеллекта в прикладных областях

Индикаторы

ПК.7.1 Руководит проектами в области сквозной цифровой технологии «Рекомендательные системы и системы поддержки принятия решений»

ПК.7.2 Исследует и анализирует развитие новых направлений и перспективных методов и технологий в области искусственного интеллекта, участвует в исследовательских проектах по развитию новых направлений в области искусственного интеллекта (алгоритмическая имитация биологических систем принятия решений, автономное самообучение и развитие адаптивности алгоритмов к новым задачам, автономная декомпозиция сложных задач, поиск и синтез решений)

4. Объем и содержание дисциплины

Направления подготовки	01.04.02 Прикладная математика и информатика (направленность: Перспективные методы искусственного интеллекта в сетях передачи и обработки данных)
форма обучения	очная
№№ триместров, выделенных для изучения дисциплины	2
Объем дисциплины (з.е.)	4
Объем дисциплины (ак.час.)	144
Контактная работа с преподавателем (ак.час.), в том числе:	72
Проведение лекционных занятий	36
Проведение практических занятий, семинаров	36
Самостоятельная работа (ак.час.)	72
Формы промежуточной аттестации	Экзамен (2 триместр)

5. Аннотированное описание содержания разделов и тем дисциплины

<i>Наименование разделов (тем) дисциплины</i>	<i>Содержание разделов (тем) дисциплин</i>
Тема 1. Систематизация задач и методов машинного обучения	Основания систематизации: по целевому признаку, по выходному признаку, по доступности распределений, по структуре данных, по разметке обучения виды моделей: дискриминативная функция, дискриминативная модель, генеративная модель
Тема 2. Базовые технологии	Библиотека numpy, модель данных ndarray, механизм broadcasting, внутреннее представление моделей данных Библиотека pandas, индексы на строках и столбцах, модель данных dataframe Библиотека matplotlib, программный интерфейс pyplot Библиотека scikit-learn: раздел classification, раздел regression Библиотека scikit-learn: раздел model selection Библиотека scikit-learn: раздел clustering, раздел dimensionality reduction Библиотека scikit-learn: раздел preprocessing, раздел feature extraction
Тема 3. Современные технологии	Библиотека tensorflow Библиотека и программный интерфейс keras
Тема 4. Задачи с целевым признаком	Бинарная классификация многоклассовая классификация функция потерь: 0-1-loss, mean absolute error, mean squared error, cross-entropy средний риск байесовский классификатор эмпирический риск одноклассовая классификация, P-classification классификация с пересекающимися классами (multi-label, multi-output) детекция идентификация верификация локализация SVM как дискриминативная функция kNN как дискриминативная модель линейные модели, их вероятностная интерпретация проблема смещения-дисперсии ансамбли классификаторов
Тема 5. Задачи без целевого признака	Восстановление плотности KDE как генеративная модель метод главных компонент автоэнкодеры кластеризация сегментация
Тема 6. Слабоструктурированная информация	Представление изображений Сверточные нейронные сети Представление последовательностей. Текст как последовательность. Представление текстов в виде мешка слов. tf-idf Модель языка. N-граммы. Рекуррентные нейронные сети. Представление элементов описания как точек в линейном пространстве (embeddings)

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 литература

1. Флах П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных. – М: ДМК Пресс. – 2015. – 400 с. ISBN 978-5-97060-273-7 (Flach P. Machine learning: the art and science of algorithms that makes sense of data. – Cambridge University Press, 2012)

Дополнительная литература

1. Bishop C. M. Pattern recognition and machine learning. – Springer, 2006
2. Коэльо Л. П., Ричарт В. Построение систем машинного обучения на языке Python. – М: ДМК Пресс. – 2016. (Coelho L. P., Richert W. Building machine learning systems with Python. — 2nd ed. — Packt Publishing Ltd, 2015.)
3. Max Kuhn, Kjell Johnson. Applied Predictive Modeling. — Springer, 2013.
4. Hastie, T., Tibshirani R., Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. — 2nd ed. — Springer-Verlag, 2009. — 746 p. — ISBN 978-0-387-84857-0.
5. Журавлев Ю. И., Рязанов В. В., Сенько О. В. «Распознавание». Математические методы. Программная система. Практические применения. — М.: Фазис, 2006. ISBN 5-7036-0108-8.
6. I.H. Witten, E. Frank Data Mining: Practical Machine Learning Tools and Techniques. — 2nd ed. — Morgan Kaufmann, 2005 ISBN 0-12-088407-0
7. Шлезингер М., Главач В. Десять лекций по статистическому и структурному распознаванию. — Киев: Наукова думка, 2004. ISBN 966-00-0341-2.

9. Перечень ресурсов сети Интернет, необходимых для освоения дисциплины

<http://www.openet.ru> - Российский портал открытого образования

<http://www.mon.gov.ru> - Министерство образования и науки Российской Федерации

mmp.cs.msu.ru

www.machinelearning.ru

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:

1. Презентационные материалы (слайды по темам лекционных и практических 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 (поисковые системы, электронная почта, мессенджеры).
5. Электронные ресурсы <https://scikit-learn.org>, <https://pandas.pydata.org/>, <https://www.tensorflow.org/learn>, <https://pytorch.org/>, <https://code.google.com/archive/p/neurolab/>

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

1. Пакеты офисных программ (тестовые процессоры, табличные редакторы, программы для создания презентаций и др.). Например, «LibreOffice», Notepad++.
2. Программный продукт Python 3+ (64-bit).
3. Среда выполнения алгоритмов машинного обучения на языке Python. Например, anaconda или google colab.
4. Операционная система Windows 10 (академическая лицензия), или Ubuntu, или ALT Linux.
5. Статистический пакет MATLAB (или свободный аналог Octave)

При освоении материала и выполнении заданий по дисциплине рекомендуется использование материалов, размещенных в Личных кабинетах обучающихся ЕГИС 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 - меловая и (или) маркерная доска, компьютерный класс (аппаратное и программное обеспечение определено в Паспортах компьютерных классов)

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.

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 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
Методы машинного обучения**

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

ПК-7. Способен руководить проектами по созданию, внедрению и использованию одной или нескольких сквозных цифровых технологий искусственного интеллекта в прикладных областях

Индикатор	Планируемые результаты обучения	Критерии оценивания результатов обучения
<p>ПК.7.1 Руководит проектами в области сквозной цифровой технологии «Рекомендательные системы и системы поддержки принятия решений»</p>	<p>умение руководить проектами по созданию, внедрению и поддержке систем искусственного интеллекта на основе сквозной цифровой технологии «Рекомендательные системы и системы поддержки принятия решений»</p>	<p align="center">Неудовлетворител Не удовлетворяет требованиям на "удовлетворительно".</p> <p align="center">Удовлетворительн Только с помощью может назвать принципы построения рекомендательных систем и систем поддержки принятия решений, методы и подходы к планированию и реализации проектов по созданию систем искусственного интеллекта, допускает существенные ошибки.</p> <p align="center">Хорошо Не всегда самостоятельно может назвать принципы построения рекомендательных систем и систем поддержки принятия решений, методы и подходы к планированию и реализации проектов по созданию систем искусственного интеллекта на основе сквозной цифровой технологии «Рекомендательные системы и системы поддержки принятия решений», допускает несущественные ошибки</p> <p align="center">Отлично Самостоятельно может назвать принципы построения рекомендательных систем и систем поддержки принятия решений, методы и подходы к планированию и реализации проектов по созданию систем искусственного интеллекта на основе сквозной цифровой технологии «Рекомендательные системы и системы поддержки принятия решений»</p>
<p>ПК-7.2. Исследует и анализирует развитие новых направлений и перспективных методов и технологий в области</p>	<p>умение проводить анализ новых направлений, методов технологий в области искусственного интеллекта и определять наиболее перспективные</p>	<p align="center">Неудовлетворител Не удовлетворяет требованиям на "удовлетворительно".</p> <p align="center">Удовлетворительн Только с помощью может рассказать о современном состоянии и перспективах</p>

искусственного интеллекта, участвует в исследовательских проектах по развитию новых направлений в области искусственного интеллекта (алгоритмическая имитация биологических систем принятия решений, автономное самообучение и развитие адаптивности алгоритмов к новым задачам, автономная декомпозиция сложных задач, поиск и синтез решений)	для различных областей применения	<p>развития новых направлений, методов и технологий в области искусственного интеллекта, допускает ошибки.</p> <p style="text-align: center;">Хорошо</p> <p>Знает современное состояние и перспективы развития новых направлений, методов и технологий в области искусственного интеллекта, но допускает неточности или ошибки.</p> <p style="text-align: center;">Отлично</p> <p>Знает современное состояние и перспективы развития новых направлений, методов и технологий в области искусственного интеллекта</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

41 до 60

«неудовлетворительно» / «незачтено» менее 41 балла

Компетенция (индикатор)	Мероприятие текущего контроля	Контролируемые элементы результатов обучения
<p>ПК.7.1 Руководит проектами в области сквозной цифровой технологии «Рекомендательные системы и системы поддержки принятия решений»</p> <p>ПК.7.2. Исследует и анализирует развитие новых направлений и перспективных методов и технологий в области искусственного интеллекта, участвует в исследовательских проектах по развитию новых направлений в области</p>	Контрольная точка 1	<p>1. Виды оснований систематизаций</p> <p>2. Виды моделей</p> <p>3. Библиотеки numpy, pandas, matplotlib, cikit-learn, scikit-learn, scikit-learn, scikit-learn</p>

<p>искусственного интеллекта (алгоритмическая имитация биологических систем принятия решений, автономное самообучение и развитие адаптивности алгоритмов к новым задачам, автономная декомпозиция сложных задач, поиск и синтез решений)</p>		
<p>ПК.7.1 Руководит проектами в области сквозной цифровой технологии «Рекомендательные системы и системы поддержки принятия решений» ПК.7.2. Исследует и анализирует развитие новых направлений и перспективных методов и технологий в области искусственного интеллекта, участвует в исследовательских проектах по развитию новых направлений в области искусственного интеллекта (алгоритмическая имитация биологических систем принятия решений, автономное самообучение и развитие адаптивности алгоритмов к новым задачам, автономная декомпозиция сложных задач, поиск и синтез решений)</p>	<p>Контрольная точка 2</p>	<ol style="list-style-type: none"> 1. Библиотеки tensorflow и keras 2. Классификации, функция потерь, риски 3. SVM как дискриминативная функция, kNN как дискриминативная модель линейные модели
<p>ПК.7.1 Руководит проектами в области сквозной цифровой технологии «Рекомендательные системы и системы поддержки принятия решений» ПК.7.2. Исследует и анализирует развитие новых направлений и перспективных методов и технологий в области искусственного интеллекта, участвует в исследовательских проектах по развитию новых направлений в области искусственного интеллекта (алгоритмическая имитация биологических систем принятия решений, автономное самообучение и развитие адаптивности алгоритмов к новым задачам, автономная</p>	<p>Контрольная точка 3</p>	<ol style="list-style-type: none"> 1. Восстановление плотности KDE как генеративная модель метод главных компонент автоэнкодеры кластеризация сегментация 2. Представление изображений, последовательностей, текстов в виде мешка слов. 3. Рекуррентные нейронные сети. Представление элементов описания как точек в линейном пространстве

декомпозиция сложных задач, поиск и синтез решений)		
---	--	--

Спецификация мероприятий текущего контроля

Контрольная точка 1.

Продолжительность проведения мероприятия промежуточной аттестации: 2 часа

Условия проведения мероприятия: в часы самостоятельной работы

Максимальный балл, выставяемый за мероприятие промежуточной аттестации: 30

Проходной балл: 13

Показатели оценивания	Баллы
1. Виды оснований систематизаций	10
2. Виды моделей	10
3. Библиотеки numpy, pandas, matplotlib, cikit-learn, scikit-learn, scikit-learn	10

Контрольная точка 2.

Продолжительность проведения мероприятия промежуточной аттестации: 2 часа

Условия проведения мероприятия: в часы самостоятельной работы

Максимальный балл, выставяемый за мероприятие промежуточной аттестации: 30

Проходной балл: 13

Показатели оценивания	Баллы
1. Библиотеки tensorflow и keras	10
2. Классификации, функция потерь, риски	10
3. SVM как дискриминативная функция, kNN как дискриминативная модель линейные модели	10

Контрольная точка 3.

Продолжительность проведения мероприятия промежуточной аттестации: 2 часа

Условия проведения мероприятия: в часы самостоятельной работы

Максимальный балл, выставяемый за мероприятие промежуточной аттестации: 40

Проходной балл: 17

Показатели оценивания	Баллы
1. Восстановление плотности KDE как генеративная модель метод главных компонент автоэнкодеры кластеризация сегментация	13
2. Представление изображений, последовательностей, текстов в виде мешка слов.	13
3. Рекуррентные нейронные сети. Представление элементов описания как точек в линейном пространстве	14