

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Ястребов Олег Александрович

Должность: Ректор

Дата подписания: 29.05.2024 11:28:12

Уникальный программный ключ:

ca953a01204891083f939673078ef1a989dae18a

**Федеральное государственное автономное образовательное учреждение высшего образования**

**«Российский университет дружбы народов имени Патриса Лумумбы»**

**Факультет искусственного интеллекта**

(наименование основного учебного подразделения (ОУП)-разработчика ОП ВО)

## **РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**

### **МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ**

(наименование дисциплины/модуля)

**Рекомендована МССН для направления подготовки/специальности:**

### **02.04.02 ФУНДАМЕНТАЛЬНАЯ ИНФОРМАТИКА И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ**

(код и наименование направления подготовки/специальности)

**Освоение дисциплины ведется в рамках реализации основной профессиональной образовательной программы высшего образования (ОП ВО):**

### **УПРАВЛЕНИЕ ДАННЫМИ И ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ**

(наименование (профиль/специализация) ОП ВО)

**2024 г.**

## 1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Дисциплина «Методы машинного обучения» входит в программу магистратуры «Управление данными и искусственный интеллект» по направлению 02.04.02 «Фундаментальная информатика и информационные технологии» и изучается в 1 семестре 1 курса. Дисциплину реализует NOT DEFINED. Дисциплина состоит из 6 разделов и 46 тем и направлена на изучение

Целью освоения дисциплины является Целью освоения дисциплины «Методы машинного обучения» является изучение современных математических методов машинного обучения, предназначенных для анализа данных и построения предсказательных моделей, изучение математических основ методов машинного обучения и соответствующих алгоритмов; изучение современных программных сред и библиотек, позволяющих проводить анализ, визуализацию данных, применять современные математические методы машинного обучения; развитие практических навыков использования методов машинного обучения в прикладных задачах.

## 2. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Освоение дисциплины «Методы машинного обучения» направлено на формирование у обучающихся следующих компетенций (части компетенций):

*Таблица 2.1. Перечень компетенций, формируемых у обучающихся при освоении дисциплины (результаты освоения дисциплины)*

Шифр	Компетенция	Индикаторы достижения компетенции (в рамках данной дисциплины)
УК-6	Способен определить и реализовать приоритеты собственной деятельности и способы ее совершенствования на основе самооценки	УК-6.2 Умеет планировать свое рабочее время и время для саморазвития. Формулировать цели личного и профессионального развития и условия их достижения, исходя из тенденций развития области профессиональной деятельности, индивидуально-личностных особенностей; УК-6.3 Имеет практический опыт получения дополнительного образования, изучения дополнительных образовательных программ;
ОПК-2	Способен применять компьютерные / суперкомпьютерные методы, современное программное обеспечение (в том числе отечественного производства) для решения задач профессиональной деятельности	ОПК-2.1 Знает основные положения и концепции в области программирования, архитектуру языков программирования, теории коммуникации, знает основную терминологию, знаком с перечнем ПО, включенного в Единый Реестр Российских программ;
ОПК-5	Способен устанавливать и сопровождать программное обеспечение информационных систем, осуществлять эффективное управление разработкой программных средств и проектов	ОПК-5.1 Знает методику установки и администрирования информационных систем и баз данных. Знаком с перечнем ПО, входящим в Единый реестр российских программ; ОПК-5.3 Имеет практические навыки установки и инсталляции программных комплексов;
ПК-2	Способен к овладению методами математического моделирования объектов и процессов при разработке требований и проектировании программного обеспечения систем искусственного интеллекта	ПК-2.1 Знает методы математического моделирования, используемые при разработке требований и проектировании программного обеспечения систем искусственного интеллекта; ПК-2.2 Умеет выбирать методы математического моделирования объектов и процессов при разработке требований и проектировании программного обеспечения систем искусственного интеллекта;

### 3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОП ВО

Дисциплина «Методы машинного обучения» относится к обязательной части блока 1 «Дисциплины (модули)» образовательной программы высшего образования.

В рамках образовательной программы высшего образования обучающиеся также осваивают другие дисциплины и/или практики, способствующие достижению запланированных результатов освоения дисциплины «Методы машинного обучения».

*Таблица 3.1. Перечень компонентов ОП ВО, способствующих достижению запланированных результатов освоения дисциплины*

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики*	Последующие дисциплины/модули, практики*
УК-6	Способен определить и реализовать приоритеты собственной деятельности и способы ее совершенствования на основе самооценки		Управление проектами и Agile; Машинный перевод; Технологическая (проектно-технологическая) практика (учебная); Преддипломная практика; Технологическая (проектно-технологическая) практика (производственная);
ОПК-2	Способен применять компьютерные / суперкомпьютерные методы, современное программное обеспечение (в том числе отечественного производства) для решения задач профессиональной деятельности		Технологическая (проектно-технологическая) практика (производственная); Вопросно-ответные системы; Генеративные методы в NLP; Методы глубокого обучения для обработки естественного языка; Компьютерное зрение; Методы оптимизации; Практикум по программированию;
ОПК-5	Способен устанавливать и сопровождать программное обеспечение информационных систем, осуществлять эффективное управление разработкой программных средств и проектов		Машинный перевод; Практикум по программированию; Генеративные методы в NLP; Обучение с подкреплением;
ПК-2	Способен к овладению методами математического моделирования объектов и процессов при разработке требований и проектировании программного обеспечения систем искусственного		Технологическая (проектно-технологическая) практика (производственная); Преддипломная практика; Компьютерное зрение; Практикум по программированию; Обучение с подкреплением; Генеративный искусственный интеллект;

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики*	Последующие дисциплины/модули, практики*
	интеллекта		<i>Искусственный интеллект в компьютерных играх**;</i> <i>Искусственный интеллект в финансовых технологиях**;</i> Генеративные методы в NLP;

\* - заполняется в соответствии с матрицей компетенций и СУП ОП ВО

\*\* - элективные дисциплины /практики

#### 4. ОБЪЕМ ДИСЦИПЛИНЫ И ВИДЫ УЧЕБНОЙ РАБОТЫ

Общая трудоемкость дисциплины «Методы машинного обучения» составляет «4» зачетные единицы.

Таблица 4.1. Виды учебной работы по периодам освоения образовательной программы высшего образования для очной формы обучения.

Вид учебной работы	ВСЕГО, ак.ч.		Семестр(-ы)
			1
<i>Контактная работа, ак.ч.</i>	36		36
Лекции (ЛК)	18		18
Лабораторные работы (ЛР)	0		0
Практические/семинарские занятия (СЗ)	18		18
<i>Самостоятельная работа обучающихся, ак.ч.</i>	81		81
<i>Контроль (экзамен/зачет с оценкой), ак.ч.</i>	27		27
<b>Общая трудоемкость дисциплины</b>	<b>ак.ч.</b>	<b>144</b>	<b>144</b>
	<b>зач.ед.</b>	<b>4</b>	<b>4</b>

## 5. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

Таблица 5.1. Содержание дисциплины (модуля) по видам учебной работы

Номер раздела	Наименование раздела дисциплины	Содержание раздела (темы)		Вид учебной работы*
Раздел 1	Математические основы машинного обучения	1.1	Постановка задач обучения по прецедентам. Объекты и признаки. Типы шкал: бинарные, номинальные, порядковые, количественные	ЛК, СЗ
		1.2	Типы задач: классификация, регрессия, прогнозирование, ранжирование	ЛК, СЗ
		1.3	Основные понятия: модель алгоритмов, метод обучения, функция потерь и функционал качества, принцип минимизации эмпирического риска, обобщающая способность, скользящий контроль	ЛК, СЗ
		1.4	Линейные модели регрессии и классификации. Метод наименьших квадратов. Полиномиальная регрессия	ЛК, СЗ
		1.5	Примеры прикладных задач.	ЛК, СЗ
Раздел 2	Линейный классификатор и стохастический градиент	2.1	Линейный классификатор, модель МакКаллока-Питтса, непрерывные аппроксимации пороговой функции потерь	ЛК, СЗ
		2.2	Метод стохастического градиента SG и стохастического среднего градиента SAG	ЛК, СЗ
		2.3	Эвристики: инициализация весов, порядок предъявления объектов, выбор величины градиентного шага, «выбивание» из локальных минимумов	ЛК, СЗ
		2.4	Проблема мультиколлинеарности и переобучения, регуляризация или редукция весов (weight decay). Гауссовский и лапласовский регуляризаторы	ЛК, СЗ
		2.5	Вероятностная постановка задачи классификации. Принцип максимума правдоподобия	ЛК, СЗ
		2.6	Логистическая регрессия. Принцип максимума правдоподобия и логарифмическая функция потерь. Метод стохастического градиента для логарифмической функции потерь. Многоклассовая логистическая регрессия. Регуляризованная логистическая регрессия. Калибровка Платта.	ЛК, СЗ
Раздел 3	Метрические методы классификации и регрессии	3.1	Гипотезы компактности и непрерывности	ЛК, СЗ
		3.2	Обобщённый метрический классификатор	ЛК, СЗ
		3.3	Метод ближайших соседей kNN и его обобщения. Подбор числа k по критерию скользящего контроля	ЛК, СЗ
		3.4	Метод окна Парзена с постоянной и переменной шириной окна.	ЛК, СЗ
		3.5	Метод потенциальных функций и его связь с линейной моделью классификации	ЛК, СЗ
		3.6	Оценка Надарая-Ватсона с постоянной и переменной шириной окна. Выбор функции ядра и ширины окна сглаживания	ЛК, СЗ
		3.7	Задача отсева выбросов. Робастная непараметрическая регрессия. Алгоритм LOWESS.	ЛК, СЗ
Раздел 4	Метод опорных векторов	4.1	Оптимальная разделяющая гиперплоскость. Понятие зазора между классами (margin).	ЛК, СЗ
		4.2	Случаи линейной разделимости и отсутствия	ЛК, СЗ

Номер раздела	Наименование раздела дисциплины	Содержание раздела (темы)		Вид учебной работы*
			линейной разделимости. Связь с минимизацией регуляризованного эмпирического риска. Кусочно-линейная функция потерь	
		4.3	Задача квадратичного программирования и двойственная задача. Понятие опорных векторов.	ЛК, СЗ
		4.4	Рекомендации по выбору константы C	ЛК, СЗ
		4.5	Функция ядра (kernel functions), спрямляющее пространство, теорема Мерсера	ЛК, СЗ
		4.6	Способы конструктивного построения ядер. Примеры ядер	ЛК, СЗ
		4.7	SVM-регрессия	ЛК, СЗ
		4.8	Регуляризации для отбора признаков: LASSO SVM, Elastic Net SVM, SFM, RFM.	ЛК, СЗ
		4.9	Метод релевантных векторов RVM	ЛК, СЗ
Раздел 5	Многомерная линейная регрессия	5.1	Задача регрессии, многомерная линейная регрессия.	ЛК, СЗ
		5.2	Метод наименьших квадратов, его вероятностный смысл и геометрический смысл	ЛК, СЗ
		5.3	Сингулярное разложение.	ЛК, СЗ
		5.4	Проблемы мультиколлинеарности и переобучения	ЛК, СЗ
		5.5	Регуляризация. Гребневая регрессия через сингулярное разложение	ЛК, СЗ
		5.6	Методы отбора признаков: Лассо Тибширани, Elastic Net, сравнение с гребневой регрессией	ЛК, СЗ
		5.7	Метод главных компонент и декоррелирующее преобразование Карунена-Лоэва, его связь с сингулярным разложением	ЛК, СЗ
		5.8	Спектральный подход к решению задачи наименьших квадратов	ЛК, СЗ
		5.9	Задачи и методы низкоранговых матричных разложений	ЛК, СЗ
Раздел 6	Критерии выбора моделей и методы отбора признаков	6.1	Критерии качества классификации: чувствительность и специфичность, ROC-кривая и AUC, точность и полнота, AUC-PR	ЛК, СЗ
		6.2	Внутренние и внешние критерии. Эмпирические и аналитические критерии	ЛК, СЗ
		6.3	Скольльзящий контроль, разновидности эмпирических оценок скользящего контроля. Критерий непротиворечивост	ЛК, СЗ
		6.4	Разновидности аналитических оценок. Регуляризация. Критерий Акаике (AIC). Байесовский информационный критерий (BIC). Оценка Вапника-Червоненкиса	ЛК, СЗ
		6.5	Сложность задачи отбора признаков. Полный перебор	ЛК, СЗ
		6.6	Метод добавления и удаления, шаговая регрессия	ЛК, СЗ
		6.7	Поиск в глубину, метод ветвей и границ	ЛК, СЗ
		6.8	Усечённый поиск в ширину, многорядный итерационный алгоритм МГУА	ЛК, СЗ
		6.9	Генетический алгоритм, его сходство с МГУА	ЛК, СЗ
		6.10	Случайный поиск и случайный поиск с адаптацией (СПА).	ЛК

\* - заполняется только по **ОЧНОЙ** форме обучения: ЛК – лекции; ЛР – лабораторные работы; СЗ – практические/семинарские занятия.

## 6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Таблица 6.1. Материально-техническое обеспечение дисциплины

Тип аудитории	Оснащение аудитории	Специализированное учебное/лабораторное оборудование, ПО и материалы для освоения дисциплины (при необходимости)
Лекционная	Аудитория для проведения занятий лекционного типа, оснащенная комплектом специализированной мебели; доской (экраном) и техническими средствами мультимедиа презентаций.	
Семинарская	Аудитория для проведения занятий семинарского типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, оснащенная комплектом специализированной мебели и техническими средствами мультимедиа презентаций.	
Для самостоятельной работы	Аудитория для самостоятельной работы обучающихся (может использоваться для проведения семинарских занятий и консультаций), оснащенная комплектом специализированной мебели и компьютерами с доступом в ЭИОС.	

\* - аудитория для самостоятельной работы обучающихся указывается **ОБЯЗАТЕЛЬНО!**

## 7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

### Основная литература:

1. Айвазян С. А., Енюков И. С., Мешалкин Л. Д. Прикладная статистика: основы моделирования и первичная обработка данных. — М.: Финансы и статистика, 1983.
2. Загоруйко Н. Г. Прикладные методы анализа данных и знаний. — Новосибирск: ИМ СО РАН, 1999. ISBN 5-86134-060-9.
3. Witten I.H., Frank E. Data Mining: Practical Machine Learning Tools and Techniques (Second Edition). — Morgan Kaufmann, 2005 ISBN 0-12-088407-0

### Дополнительная литература:

- 1.
- 2.

### Ресурсы информационно-телекоммуникационной сети «Интернет»:

1. ЭБС РУДН и сторонние ЭБС, к которым студенты университета имеют доступ на основании заключенных договоров
  - Электронно-библиотечная система РУДН – ЭБС РУДН <http://lib.rudn.ru/MegaPro/Web>
  - ЭБС «Университетская библиотека онлайн» <http://www.biblioclub.ru>
  - ЭБС Юрайт <http://www.biblio-online.ru>
  - ЭБС «Консультант студента» [www.studentlibrary.ru](http://www.studentlibrary.ru)
  - ЭБС «Троицкий мост»
2. Базы данных и поисковые системы



- электронный фонд правовой и нормативно-технической документации  
<http://docs.cntd.ru/>

- поисковая система Яндекс <https://www.yandex.ru/>

- поисковая система Google <https://www.google.ru/>

- реферативная база данных SCOPUS

<http://www.elsevierscience.ru/products/scopus/>

*Учебно-методические материалы для самостоятельной работы обучающихся при освоении дисциплины/модуля\*:*

1. Курс лекций по дисциплине «Методы машинного обучения».

\* - все учебно-методические материалы для самостоятельной работы обучающихся размещаются в соответствии с действующим порядком на странице дисциплины **в ТУИС!**

## **8. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ И БАЛЛЬНО-РЕЙТИНГОВАЯ СИСТЕМА ОЦЕНИВАНИЯ УРОВНЯ СФОРМИРОВАННОСТИ КОМПЕТЕНЦИЙ ПО ДИСЦИПЛИНЕ**

Оценочные материалы и балльно-рейтинговая система\* оценивания уровня сформированности компетенций (части компетенций) по итогам освоения дисциплины «Методы машинного обучения» представлены в Приложении к настоящей Рабочей программе дисциплины.

\* - ОМ и БРС формируются на основании требований соответствующего локального нормативного акта РУДН.

**РАЗРАБОТЧИК:**

<hr/>	<hr/>	<hr/>
<i>Должность, БУП</i>	<i>Подпись</i>	Баум Валентина Владимировна <i>Фамилия И.О.</i>

**РУКОВОДИТЕЛЬ БУП:**

<hr/>	<hr/>	<hr/>
Заведующий кафедрой <i>Должность БУП</i>	<i>Подпись</i>	<i>Фамилия И.О.</i>

**РУКОВОДИТЕЛЬ ОП ВО:**

<hr/>	<hr/>	<hr/>
<i>Должность, БУП</i>	<i>Подпись</i>	Воробьева Александра Андреевна <i>Фамилия И.О.</i>