

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Ястребов Олег Александрович
Должность: Ректор
Дата подписания: 23.05.2025 13:10:44
Уникальный программный ключ:
sa953a0120d891083f939673078ef1a989dae18a

**Федеральное государственное автономное образовательное учреждение высшего образования
«Российский университет дружбы народов имени Патриса Лумумбы»**

Инженерная академия

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

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

МАШИННОЕ ОБУЧЕНИЕ И АНАЛИЗ БОЛЬШИХ ДАННЫХ

(наименование дисциплины/модуля)

Рекомендована МССН для направления подготовки/специальности:

01.04.02 ПРИКЛАДНАЯ МАТЕМАТИКА И ИНФОРМАТИКА

(код и наименование направления подготовки/специальности)

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

ПРОЕКТИРОВАНИЕ КОСМИЧЕСКИХ МИССИЙ И СИСТЕМ

(наименование (профиль/специализация) ОП ВО)

2025 г.

1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Дисциплина «Machine Learning and Big Data Mining» входит в программу магистратуры «Проектирование космических миссий и систем» по направлению 01.04.02 «Прикладная математика и информатика» и изучается в 1 семестре 1 курса. Дисциплину реализует Кафедра механики и процессов управления. Дисциплина состоит из 9 разделов и 27 тем и направлена на изучение fundamentals of clustering and its basic techniques, classification and its basic techniques, frequent itemset mining and association rules, feature selection and dimensionality reduction, outlier detection, recommender systems and algorithms, ensemble clustering and classification, multimodal relational clustering, artificial neural methods and stochastic optimization, elements of statistical learning; collection of basic methods for solving typical problems and familiarization with the area of their application in professional activities.

Целью освоения дисциплины является to develop fundamental knowledge and skills in applying problem solving methods necessary for professional activities, to improve the general level of students' literacy in the discipline Machine Learning and Big Data Mining, to familiarize them with a new rapidly evolving field and provide practical knowledge experience in the analysis of real world data.

2. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Освоение дисциплины «Машинное обучение и анализ больших данных» 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)
ПК-1	Способен формулировать цели, задачи научных исследований в области прикладной математики и информатики, вычислительной техники и современных технологий программирования, выбирать методы и средства решения задач	ПК-1.1 Обладает фундаментальными знаниями, полученными в области математических и (или) естественных наук, программирования и информационных технологий;; ПК-1.2 Умеет находить, формулировать и решать стандартные задачи в собственной научно-исследовательской деятельности в области прикладной математики и информатики, вычислительной техники и современных технологий программирования;; ПК-1.3 Имеет практический опыт научно-исследовательской деятельности в области прикладной математики и информатики, вычислительной техники и современных технологий программирования.;
ПК-2	Способен применять современные теоретические и экспериментальные методы разработки математических моделей исследуемых объектов и процессов, относящихся к профессиональной деятельности по направлению подготовки и участвовать в их реализации в виде программных продуктов	ПК-2.1 Знает современные теоретические и экспериментальные методы разработки математических моделей, инновационные инструментальные средства проектирования и элементы архитектурных решений информационных систем;; ПК-2.2 Умеет разрабатывать и реализовывать алгоритмы математических моделей на базе языков и пакетов прикладных программ моделирования;; ПК-2.3 Имеет практический опыт разработки вариантов реализации информационных систем с использованием инновационных инструментальных средств.;

3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОП ВО

Дисциплина «Машинное обучение и анализ больших данных» относится к части, формируемой участниками образовательных отношений 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 «Машинное обучение и анализ больших данных».

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

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики*	Последующие дисциплины/модули, практики*
ПК-1	Способен формулировать цели, задачи научных исследований в области прикладной математики и информатики, вычислительной техники и современных технологий программирования, выбирать методы и средства решения задач		Современные методы механики космического полета; <i>Искусственные нейронные сети (Глубокое обучение)**;</i> <i>Искусственные нейронные сети (Обучение с подкреплением)**;</i> <i>Artificial Neural Networks (Deep Learning)**;</i> <i>Artificial Neural Networks (Reinforcement Learning)**;</i> Dynamics and Control of Space Systems; Геоинформационные системы и их применение; Разработка и безопасность веб-приложений; Современные методы дистанционного зондирования Земли; Научно-исследовательская работа; Технологическая практика; Преддипломная практика;
ПК-2	Способен применять современные теоретические и экспериментальные методы разработки математических моделей исследуемых объектов и процессов, относящихся к профессиональной деятельности по направлению подготовки и участвовать в их реализации в виде программных продуктов		<i>Искусственные нейронные сети (Глубокое обучение)**;</i> <i>Искусственные нейронные сети (Обучение с подкреплением)**;</i> <i>Artificial Neural Networks (Deep Learning)**;</i> <i>Artificial Neural Networks (Reinforcement Learning)**;</i> Dynamics and Control of Space Systems; Геоинформационные системы и их применение; Научно-исследовательская работа; Технологическая практика; Преддипломная практика;

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

** - элективные дисциплины /практики

4. ОБЪЕМ ДИСЦИПЛИНЫ И ВИДЫ УЧЕБНОЙ РАБОТЫ

Общая трудоемкость дисциплины «Machine Learning and Big Data Mining» составляет «5» зачетных единиц.

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
Контактная работа, ак.ч.	34		34
Лекции (ЛК)	17		17
Лабораторные работы (ЛР)	17		17
Практические/семинарские занятия (СЗ)	0		0
Самостоятельная работа обучающихся, ак.ч.	110		110
Контроль (экзамен/зачет с оценкой), ак.ч.	36		36
Общая трудоемкость дисциплины	ак.ч.	180	180
	зач.ед.	5	5

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	Introduction to Machine Learning and Data Mining	1.1	Introduction to modern data analysis	ЛК, ЛР
		1.2	Machine Learning. Data Mining and Knowledge Discovery in Data Bases	ЛК, ЛР
Раздел 2	Clustering and its basic techniques	2.1	The task of clusterization	ЛК, ЛР
		2.2	K-means and its modifications (k-medoids and fuzzy cmeans clustering)	ЛК, ЛР
		2.3	Density-based methods: DB-scan and Mean Shift	ЛК, ЛР
		2.4	Hierarchical clustering	ЛК, ЛР
		2.5	Criteria of quality	ЛК, ЛР
Раздел 3	Classification and its basic techniques	3.1	The task of classification	ЛК, ЛР
		3.2	1-Rules. K-Nearest Neighbours approach	ЛК, ЛР
		3.3	Naïve Bayes. Decision Trees. Logistic Regression	ЛК, ЛР
		3.4	Quality assessment: precision, recall, F - measure, loss-function, confusion-matrix, cross- validation and learning curves (ROC, lift etc.)	ЛК, ЛР
		3.5	Multi-class and multi-label classification	ЛК, ЛР
Раздел 4	Frequent Itemset Mining and Association Rules	4.1	Frequent itemsets. Apriori and FP-growth algorithms	ЛК, ЛР
		4.2	Association rules. Interestingness measures: support and confidence. Closed itemsets	ЛК, ЛР
		4.3	Connection with Lattice Theory and Formal Concept Analysis. Applications	ЛК, ЛР
Раздел 5	Feature Selection and Dimensionality Reduction. Outlier detection	5.1	Feature selection versus feature extraction and generation	ЛК, ЛР
		5.2	Singular Value Decomposition, Latent Semantic Analysis and Principal Component Analysis. Boolean Matrix Factorization	ЛК, ЛР
		5.3	Outlier and novelty detection techniques	ЛК, ЛР
Раздел 6	Recommender Systems and Algorithms	6.1	Collaborative filtering. User-based and item-based methods. Slope one	ЛК, ЛР
		6.2	Association rules based and bicluster-based techniques. Quality assessment: MAE, precision and recall	ЛК, ЛР
		6.3	SVD-based approaches: pureSVD, SVD++ and time-SVD. Factorization machines	ЛК, ЛР
Раздел 7	Ensemble Clustering and Classification	7.1	Ensemble methods of clusterization for k-means partitions' aggregation	ЛК, ЛР
		7.2	Ensemble methods of classification: Bagging, Boosting, and Random Forest	ЛК, ЛР
Раздел 8	Multimodal relational clustering	8.1	Biclustering. Spectral co-clustering. Triclustering	ЛК, ЛР
		8.2	Two-mode networks. Folksonomies and resource-sharing systems. Multimodal approaches	ЛК, ЛР
		8.3	Applications: Community detection in Socail Network Analysis and gene expression analysis	ЛК, ЛР
Раздел 9	Artificial Neural Methods and Stochastic Optimization. Elements of Statistical Learning	9.1	Artificial Neural Networks. Basic ideas of Deep Learning. (Stochastic) gradient descent. Statistical (Bayesian) view on Machine learning	ЛК, ЛР

* - 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		
Компьютерный класс		
Для самостоятельной работы		

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

7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Основная литература:

1. Han, J., Kamber, M., Pei, J. Data Mining: Concepts and Techniques, Third Edition. – Morgan Kaufmann Publishers, 2011. – 740 pp.

Дополнительная литература:

1. Hall, M., Witten, Ian H., Frank, E. Data Mining: practical machine learning tools and techniques. – 2011. – 664 pp

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

- 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 «Машинное обучение и анализ больших данных».

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

РАЗРАБОТЧИК:

Доцент

Должность, БУП

Подпись

Салтыкова Ольга
Александровна

Фамилия И.О.

РУКОВОДИТЕЛЬ БУП:

Заведующий кафедрой

Должность БУП

Подпись

Разумный Юрий
Николаевич

Фамилия И.О.

РУКОВОДИТЕЛЬ ОП ВО:

Профессор

Должность, БУП

Подпись

Разумный Юрий
Николаевич

Фамилия И.О.