МИНИСТЕРСТВО ТРАНСПОРТА РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

«РОССИЙСКИЙ УНИВЕРСИТЕТ ТРАНСПОРТА»
(РУТ (МИИТ)

Рабочая программа дисциплины (модуля),
как компонент образовательной программы
высшего образования - программы бакалавриата
по направлению подготовки
09.03.01 Информатика и вычислительная техника,
утвержденной первым проректором РУТ (МИИТ)
Тимониным В.С.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Машинное обучение и анализ данных

Направление подготовки:	09.03.01 Информатика и вычислительная техника
Направленность (профиль):	IT-сервисы и технологии обработки данных на транспорте
Форма обучения:	Очная

Рабочая программа дисциплины (модуля) в виде электронного документа выгружена из единой корпоративной информационной системы управления университетом и соответствует оригиналу

Простая электронная подпись, выданная РУТ (МИИТ)
ID подписи: 170737
Подписал: заместитель директора академии Паринов Денис Владимирович
Дата: 29.05.2023

1. Общие сведения о дисциплине (модуле).

Целями освоения учебной дисциплины являются формирование у обучающихся профессиональных знаний и навыков в области использования моделей и методов систем искусственного интеллекта (ИИ); усвоение основ разработки и применения методов систем ИИ при решении прикладных задач обработки информации и оптимального управления, а также формирование у обучающихся компетенций в области построения и функционирования систем ИИ, методов оптимального представления и обработки знаний в этих системах, методов и алгоритмов обработки информации и оптимального управления.

Дисциплина предназначена для формирования и закрепления у обучающихся компетенций, необходимых для решения следующих профессиональных задач (в рамках проектной деятельности):

- Проверка реализации запросов на изменение (верификация) в соответствии с полученным планом;
- Обеспечение качества в проектах в области ИТ в соответствии с установленными регламентами;
- Организация приемо-сдаточных испытаний (валидация) в проектах малого и среднего уровня сложности в области ИТ в соответствии с установленными регламентами.

2. Планируемые результаты обучения по дисциплине (модулю).

Перечень формируемых результатов освоения образовательной программы (компетенций) в результате обучения по дисциплине (модулю):

ОПК-8 - Способен разрабатывать алгоритмы и программы, пригодные для практического применения;

ПК-1 - Способен анализировать большие данные с использованием существующей в организации методологической и технологической инфраструктуры.

Обучение по дисциплине (модулю) предполагает, что по его результатам обучающийся будет:

Знать:

требования к персональным помпьютерам и серверным ресурсам, реализующим алгоритмы машинного обучения

основные методы оптимального представления и обработки больших данных в системах ИИ

особенности машинного обучения с/без учителя

программы для практической реализации машинного обучения

Уметь:

разрабатывать и проводить сравнительный анализ алгоритмов распознавания и обработки больших данных

проектировать вычислительную инфраструктуру для хранения больших объемов данных

осуществлять отбор данных для машинного обучения

Владеть:

навыками применения методов машинного обучения при решении профессиональных задач

навыками разработки алгоритмов машинного обучения

навыками подбора компонентов для инфраструктуры задач глубокого обучения

навыками формирования технического задания на машинное обучение

3. Объем дисциплины (модуля).

3.1. Общая трудоемкость дисциплины (модуля).

Общая трудоемкость дисциплины (модуля) составляет 3 з.е. (108 академических часа(ов).

3.2. Объем дисциплины (модуля) в форме контактной работы обучающихся с педагогическими работниками и (или) лицами, привлекаемыми к реализации образовательной программы на иных условиях, при проведении учебных занятий:

Тип учебных занятий	Количество часов	
	Всего	Сем. №4
Контактная работа при проведении учебных занятий (всего):	64	64
В том числе:		
Занятия лекционного типа	32	32
Занятия семинарского типа	32	32

3.3. Объем дисциплины (модуля) в форме самостоятельной работы обучающихся, а также в форме контактной работы обучающихся с педагогическими работниками и (или) лицами, привлекаемыми к реализации

образовательной программы на иных условиях, при проведении промежуточной аттестации составляет 44 академических часа (ов).

3.4. При обучении по индивидуальному учебному плану, в том числе при ускоренном обучении, объем дисциплины (модуля) может быть реализован полностью в форме самостоятельной работы обучающихся, а также в форме контактной работы обучающихся с педагогическими работниками и (или) лицами, привлекаемыми к реализации образовательной программы на иных условиях, при проведении промежуточной аттестации.

4. Содержание дисциплины (модуля).

4.1. Занятия лекционного типа.

№ п/п	Тематика лекционных занятий / краткое содержание
1	Этапы машинного обучения Рассматриваемые вопросы: Классификация алгоритмов машинного обучения Этапы машинного обучения
2	Классификация Рассматриваемые вопросы: Постановка задач классификации, регрессии, прогнозирования и ранжирования Современные методы решения задач распознавания образов. Рекомендательные системы
3	Методы кластерного анализа Рассматриваемые вопросы: Постановка задач кластеризации, поиска ассоциативных правил, фильтрации выбросов, построение доверительной области, сокращение размерности, заполнение пропущенных значений Этапы и методы проведения кластерного анализа Алгоритмы генерации ассоциативных правил. Фильтрация данных
4	Нейронные сети Рассматриваемые вопросы: Классификация архитектур нейронных сетей Нейронные сети глубокого обучения Рекуррентные нейронные сети. Сверточные нейронные сети
5	Алгоритмы Рассматриваемые вопросы: Основные понятия и этапы работы генетических алгоритмов. Основные понятия агентного подхода. Многоагентные системы. Алгоритмы SARSA и Q-learning
6	Интерпретация моделей машинного обучения Рассматриваемые вопросы: Особенности обработки естественного языка. Библиотеки и фреймворки для распознавания речи, анализа и генерирования текста, синтеза речи

4.2. Занятия семинарского типа.

Практические занятия

№ п/п	Тематика практических занятий/краткое содержание
1	Этапы машинного обучения Рассматриваемые вопросы: Разведочный анализ данных. Заполнение пропусков. Генерация признаков
2	Методы кластерного анализа Рассматриваемые вопросы: Машинное обучение без учителя. Кластеризация и поиск аномалий.
3	Регрессия Рассматриваемые вопросы: Регрессия (линейная регрессия, деревья решений, простая нейронная сеть).
4	Классификация Рассматриваемые вопросы: Машинное обучение с учителем. Классификация (логистическая регрессия, дерево решений, случайный лес, градиентный бустинг)
5	Интерпретация моделей машинного обучения Рассматриваемые вопросы: Интерпретация результатов моделей машинного обучения с учителем. Частичные зависимости (PD) и локальная интерпретация (LIME).
6	Подготовка математических моделей данных для формирования кейсов по машинному обучению Рассмотрение основных подходов к построению математических моделей данных и математические модели данных для кейсов по машинному обучению

4.3. Самостоятельная работа обучающихся.

№ п/п	Вид самостоятельной работы
1	Изучение литературы
2	Работа с документацией библиотек pandas, matplotlib, seaborn, pandas_profiling
3	Выполнение курсовой работы.
4	Подготовка к промежуточной аттестации.
5	Подготовка к текущему контролю.

4.4. Примерный перечень тем курсовых работ

1. Использование метода опорных векторов в задачах машинного обучения.

2. Использование метода наименьших квадратов в задачах машинного обучения.

3. Использование нейронных сетей долгой краткосрочной памяти в задачах анализа временных рядов.

4. Использование библиотеки TensorFlow в задачах машинного обучения.

5. Использование IBM Watson Studio для создания когнитивных приложений.

6. Использование IBM Watson Studio в задачах распознавания образов.

7. Обработка естественного языка в когнитивных приложениях.

8. Алгоритмы поиска ассоциативных правил.

9. Рекомендательные системы.

10. Распознавание образов с использованием искусственных нейронных сетей глубокого обучения.

11. Обучение с подкреплением в системах искусственного интеллекта.

12. Сегментация изображений с использованием методов кластерного анализа.

5. Перечень изданий, которые рекомендуется использовать при освоении дисциплины (модуля).

№ п/п	Библиографическое описание	Место доступа
1	1. Габдуллин, Н. М. Развитие человеческого капитала и цифровой экономики в регионах России: факторный и кластерный анализ : монография / Н. М. Габдуллин. — Казань : КФУ, 2019. — 268 с. — ISBN 978-5-00130-291-9. — Текст : электронный // Лань : электронно-библиотечная система	https://e.lanbook.com/book/173018
2	2. Гитис, Л. Х. Статистическая классификация и кластерный анализ / Л. Х. Гитис. — Москва : Горная книга, 2003. — 157 с. — ISBN 5-7418-0010-6. — Текст : электронный // Лань : электронно-библиотечная система.	https://e.lanbook.com/book/3493
3	3. Гласснер, Э. Глубокое обучение без математики. Том 2. Практика : руководство / Э. Гласснер ; перевод с английского В. А. Яроцкого. — Москва : ДМК Пресс, 2020. — 610 с. — ISBN 978-5-97060-767-1. — Текст : электронный // Лань : электронно-библиотечная система	https://e.lanbook.com/book/131710
4	4. Гультяева, Т. А. Методы статистического обучения в задачах регрессии и классификации : монография / Т. А. Гультяева, А. А. Попов, А. С. Саутин. — Новосибирск : НГТУ, 2016. — 323 с.	URL: https://e.lanbook.com/book/118291

	— ISBN 978-5-7782-2817-7. — Текст : электронный // Лань : электронно-библиотечная систе	
5	5. Кук, Д. Машинное обучение с использованием библиотеки H2O / Д. Кук ; перевод с английского А. Б. Огурцова. — Москва : ДМК Пресс, 2018. — 250 с. — ISBN 978-5-97060-508-0. — Текст : электронный // Лань : электронно-библиотечная система.	https://e.lanbook.com/book/97353
6	6. Флах, П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных / П. Флах. — Москва : ДМК Пресс, 2015. — 400 с. — ISBN 978-5-97060-273-7. — Текст : электронный // Лань : электронно-библиотечная система.	https://e.lanbook.com/book/69955
7	7. Целых, А. Н. Современные технологии противодействия финансовым преступлениям : учебное пособие / А. Н. Целых. — Ростов-на-Дону : ЮФУ, 2019. — 119 с. — ISBN 978-5-9275-3286-5. — Текст : электронный // Лань : электронно-библиотечная система	https://e.lanbook.com/book/141063
8	8. Шалев-Шварц, Ш. Идеи машинного обучения : учебное пособие / Ш. Шалев-Шварц, Бен-ДавидШ. ; перевод с английского А. А. Слинкина. — Москва : ДМК Пресс, 2019. — 436 с. — ISBN 978-5-97060-673-5. — Текст : электронный // Лань : электронно-библиотечная система.	https://e.lanbook.com/book/131686

6. Перечень современных профессиональных баз данных и информационных справочных систем, которые могут использоваться при освоении дисциплины (модуля).

Документация библиотеки pandas. https://pandas.pydata.org/docs/

Документация библиотеки matplotlib. https://matplotlib.org/stable/users/index

Документация библиотеки seaborn. https://seaborn.pydata.org/tutorial.html

Документация библиотеки pandas_profiling. https://pandas-profiling.github.io/pandas-profiling/docs/master/index.html#pandas-profiling

Документация scikit-learn https://scikit-learn.org/stable/user_guide.html

Документация библиотеки lime https://github.com/marcotcr/lime

http://aiportal.ru

http://library.miit.ru

https://e.lanbook.com

7. Перечень лицензионного и свободно распространяемого программного обеспечения, в том числе отечественного производства, необходимого для освоения дисциплины (модуля).

MS Office (Excel, Word)
Notepad++
Браузер Chrome
Anaconda

8. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю).

2 учебных класса (столы, стулья - по 25 ед)
Компьютер преподавателя
Intel Core i7-9700 / Asus PRIME H310M-R R2.0 / 2x8GB / SSD 250Gb / DVDRW
Компьютеры студентов (24 ед)
Intel Core i9-9900 / B365M Pro4 / 2x16GB / SSD 512Gb
Монитор (25 ед)
Клавиатура (25 ед)
Мышка (25 ед)
Лазерный принтер (2 ед)
Проектор Optoma W340UST
Экран для проектора
Маркерная доска

9. Форма промежуточной аттестации:

Зачет в 4 семестре.
Курсовая работа в 4 семестре.

10. Оценочные материалы.

Оценочные материалы, применяемые при проведении промежуточной аттестации, разрабатываются в соответствии с локальным нормативным актом РУТ (МИИТ).

Авторы:

доцент, к.н. Академии "Высшая инженерная школа"	Б.В. Игольников

Согласовано:

Заместитель директора академии	Д.В. Паринов
Председатель учебно-методической комиссии	Д.В. Паринов