

МИНИСТЕРСТВО ТРАНСПОРТА РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«РОССИЙСКИЙ УНИВЕРСИТЕТ ТРАНСПОРТА»
(РУТ (МИИТ))



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

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Практикум по валидации и интерпретации моделей

Направление подготовки: 09.03.01 Информатика и вычислительная техника

Направленность (профиль): IT-сервисы и технологии обработки данных на транспорте (Российско-Китайская программа)

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 ФГАОУ ВО РУТ
(МИИТ)

ID подписи: 35225

Подписал: заместитель директора Игольников Борис
Владимирович

Дата: 16.10.2026

1. Общие сведения о дисциплине (модуле).

Целью освоения учебной дисциплины является формирование у обучающихся системных знаний и практических навыков в области валидации, тестирования и интерпретации моделей машинного обучения, включая понимание методов оценки качества, подходов к обеспечению надёжности моделей, техник объяснимого искусственного интеллекта и автоматизации работы с признаками.

Задачами освоения дисциплины являются:

Формирование теоретических знаний о методах валидации моделей машинного обучения, включая кросс-валидацию, бутстрэп и методы оценки обобщающей способности.

Освоение практических навыков тестирования моделей на различных типах данных и оценки их устойчивости к изменениям.

Приобретение умений применять методы интерпретации моделей, включая SHAP, LIME, анализ важности признаков и визуализацию принятия решений.

Формирование компетенций по автоматизации работы с признаками, включая отбор, генерацию и трансформацию признаков для повышения качества моделей.

Освоение алгоритмов автоматического машинного обучения и их адаптации к специфике прикладных задач.

Приобретение навыков обоснования выбора моделей и методов их валидации для решения конкретных бизнес-задач.

Дисциплина предназначена для формирования компетенций, необходимых для решения профессиональных задач, связанных с проектированием, разработкой и улучшением алгоритмов и моделей машинного обучения, включая обеспечение их надёжности, интерпретируемости и эффективности.

2. Планируемые результаты обучения по дисциплине (модулю).

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

ОПК-3 - Способен использовать современные информационные технологии и программно-аппаратные средства, в том числе отечественного производства, при решении задач профессиональной деятельности;

ПК-1 - Способен проектировать, разрабатывает и улучшать алгоритмы и модели машинного обучения.

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

Знать:

Методы валидации и оценки качества моделей машинного обучения, включая кросс-валидацию, стратифицированную выборку, временную кросс-валидацию и методы оценки обобщающей способности.

Алгоритмы автоматизации работы с признаками, включая методы отбора, генерации, трансформации и кодирования признаков, а также подходы к автоматическому поиску оптимальных признаков пространств.

Методы интерпретации и объяснения моделей машинного обучения, включая локальные и глобальные методы объяснимости, анализ важности признаков и визуализацию принятия решений.

Уметь:

Применять методы валидации для объективной оценки качества моделей, выявлять проблемы переобучения и недообучения, обосновывать выбор метрик и процедур валидации для конкретных задач.

Обосновывать способы и варианты применения алгоритмов автоматического машинного обучения в задачах ИИ, включая их преобразование и адаптацию к специфике задачи.

Применять методы интерпретации моделей для объяснения их решений, выявления факторов, влияющих на предсказания, и формулировки рекомендаций на основе анализа моделей.

Владеть:

Навыками применения методов валидации и тестирования моделей с использованием современных библиотек и фреймворков.

Навыками автоматизации работы с признаками, включая использование инструментов автоматического отбора и генерации признаков.

Навыками интерпретации моделей с использованием методов объяснимого ИИ и визуализации принятия решений.

3. Объем дисциплины (модуля).

3.1. Общая трудоемкость дисциплины (модуля).

Общая трудоемкость дисциплины (модуля) составляет 5 з.е. (180 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	Семестр №5
Контактная работа при проведении учебных занятий (всего):	80	80
В том числе:		
Занятия лекционного типа	16	16
Занятия семинарского типа	64	64

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 100 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	Тема 1. Введение в валидацию и интерпретацию моделей. Рассматриваемые вопросы: - Роль валидации и интерпретации в жизненном цикле ML-моделей. - Основные понятия: обобщающая способность, переобучение, недообучение. - Этические и регуляторные аспекты интерпретируемости моделей. - Обзор методов валидации и интерпретации.
2	Тема 2. Методы валидации моделей. Рассматриваемые вопросы: - Hold-out метод: разделение выборки на обучающую, валидационную и тестовую. - К-блочная кросс-валидация: алгоритм, выбор числа блоков. - Стратифицированная кросс-валидация для несбалансированных данных. - Временная кросс-валидация для временных рядов. - Бутстрэп-методы оценки качества.
3	Тема 3. Метрики оценки качества моделей. Рассматриваемые вопросы: - Метрики для задач классификации: Accuracy, Precision, Recall, F1, ROC-AUC, PR-AUC.

№ п/п	Тематика лекционных занятий / краткое содержание
	- Метрики для задач регрессии: MAE, MSE, RMSE, MAPE, R². - Метрики для задач ранжирования: MAP, NDCG. - Выбор метрик в зависимости от бизнес-задачи. - Интерпретация метрик для стейкхолдеров.
4	Тема 4. Автоматизация работы с признаками (Feature Engineering Automation). Рассматриваемые вопросы: - Роль признаков в качестве моделей машинного обучения. - Автоматический отбор признаков: фильтрующие, обёрточные и встроенные методы. - Генерация новых признаков: трансформации, взаимодействия, полиномиальные признаки. - Автоматическое кодирование категориальных признаков. - Инструменты автоматизации: Featuretools, Tsfresh, AutoFeat.
5	Тема 5. Автоматическое машинное обучение (AutoML). Рассматриваемые вопросы: - Концепция AutoML и её место в MLOps. - Автоматический поиск архитектуры моделей. - Автоматический подбор гиперпараметров: Grid Search, Random Search, Bayesian Optimization. - Современные библиотеки AutoML: AutoGluon, H2O.ai, TPOT. - Ограничения и вызовы AutoML.
6	Тема 6. Методы интерпретации моделей. Рассматриваемые вопросы: - Уровни интерпретации: глобальная и локальная. - Анализ важности признаков: Permutation Importance, Feature Importance из встроенных моделей. - LIME: локальные интерпретации, не зависящие от модели. - SHAP: унифицированный подход к интерпретации. - Визуализация принятия решений.
7	Тема 7. Интерпретация чёрных ящиков. Рассматриваемые вопросы: - Интерпретация сложных моделей: ансамблей, нейросетей. - Градиентные методы интерпретации: Grad-CAM, Integrated Gradients. - Surogate-модели для интерпретации. - Интерпретация в задачах NLP и CV. - Компромисс между точностью и интерпретируемостью.
8	Тема 8. Практические аспекты валидации и интерпретации. Рассматриваемые вопросы: - Интеграция валидации в пайплайны ML. - Автоматизация отчётов о качестве моделей. - Обоснование выбора модели для стейкхолдеров. - Кейсы внедрения интерпретации в бизнес-процессы. - Этические аспекты использования интерпретируемых моделей. - Ответственный искусственный интеллект .

4.2. Занятия семинарского типа.

Практические занятия

№ п/п	Тематика практических занятий/краткое содержание
1	Тема 1. Установка и настройка среды для практикума. Рассматриваемые вопросы: - Установка и настройка библиотек для валидации и интерпретации.

№ п/п	Тематика практических занятий/краткое содержание
	- Настройка окружения для воспроизводимых экспериментов. - Загрузка и предобработка датасетов. - Проверка работоспособности.
2	Тема 2. Простое разделение выборки и базовое обучение. Рассматриваемые вопросы: - Разделение данных на обучающую и тестовую выборки. - Обучение базовой модели. - Оценка качества на тестовой выборке. - Анализ результатов.
3	Тема 3. Реализация кросс-валидации. Рассматриваемые вопросы: - Реализация K-блочной кросс-валидации. - Сравнение с простым разделением. - Стратифицированная кросс-валидация. - Анализ стабильности оценок качества.
4	Тема 4. Временная кросс-валидация. Рассматриваемые вопросы: - Применение временной кросс-валидации для временных рядов. - Сравнение с обычной кросс-валидацией. - Прогнозирование и оценка качества. - Анализ результатов.
5	Тема 5. Расчёт метрик классификации. Рассматриваемые вопросы: - Расчёт Accuracy, Precision, Recall, F1. - Построение матрицы ошибок. - Расчёт ROC-AUC и PR-AUC. - Интерпретация метрик для бизнес-задач.
6	Тема 6. Расчёт метрик регрессии. Рассматриваемые вопросы: - Расчёт MAE, MSE, RMSE, MAPE. - Расчёт R². - Визуализация остатков. - Интерпретация метрик для бизнес-задач.
7	Тема 7. Выбор метрик в зависимости от задачи. Рассматриваемые вопросы: - Анализ бизнес-задачи и выбор метрик. - Сравнение различных метрик на одном датасете. - Обоснование выбора метрик. - Подготовка отчёта по качеству модели.
8	Тема 8. Автоматический отбор признаков. Рассматриваемые вопросы: - Применение фильтрующих методов отбора признаков. - Применение обёрточных методов. - Использование встроенных методов. - Сравнение результатов.
9	Тема 9. Генерация новых признаков. Рассматриваемые вопросы: - Применение полиномиальных признаков. - Создание признаков взаимодействий.

№ п/п	Тематика практических занятий/краткое содержание
	- Использование Featuretools для автоматической генерации. - Оценка влияния новых признаков на качество модели.
10	Тема 10. Автоматическое кодирование категориальных признаков. Рассматриваемые вопросы: - Применение One-Hot Encoding. - Применение Target Encoding. - Сравнение различных методов кодирования. - Выбор оптимального метода.
11	Тема 11. Автоматический подбор гиперпараметров. Рассматриваемые вопросы: - Применение Grid Search. - Применение Random Search. - Применение Bayesian Optimization. - Сравнение эффективности методов.
12	Тема 12. Использование библиотек AutoML. Рассматриваемые вопросы: - Применение AutoGluon. - Применение H2O.ai. - Сравнение качества моделей AutoML с ручной настройкой. - Анализ преимуществ и ограничений.
13	Тема 13. Анализ важности признаков. Рассматриваемые вопросы: - Использование встроенной важности признаков из моделей. - Применение Permutation Importance. - Сравнение различных подходов. - Визуализация важности признаков.
14	Тема 14. Интерпретация с LIME. Рассматриваемые вопросы: - Применение LIME для объяснения предсказаний. - Визуализация локальных интерпретаций. - Сравнение интерпретаций для различных моделей. - Анализ результатов.
15	Тема 15. Интерпретация с SHAP. Рассматриваемые вопросы: - Применение SHAP для глобальной интерпретации. - Применение SHAP для локальной интерпретации. - Визуализация SHAP-значений. - Интерпретация результатов для бизнес-задач.
16	Тема 16. Сравнение методов интерпретации. Рассматриваемые вопросы: - Сравнение LIME и SHAP на одном датасете. - Анализ согласованности интерпретаций. - Выбор метода интерпретации в зависимости от задачи. - Обоснование выбора.
17	Тема 17. Интерпретация ансамблевых моделей. Рассматриваемые вопросы: - Интерпретация случайного леса. - Интерпретация градиентного бустинга. - Сравнение важности признаков в разных ансамблях. - Визуализация принятия решений.

№ п/п	Тематика практических занятий/краткое содержание
18	Тема 18. Интерпретация нейронных сетей. Рассматриваемые вопросы: - Применение градиентных методов интерпретации. - Визуализация активаций. - Интерпретация сверточных нейросетей. - Анализ результатов.
19	Тема 19. Интерпретация NLP-моделей. Рассматриваемые вопросы: - Интерпретация предсказаний текстовых моделей. - Визуализация внимания в трансформерах. - Применение SHAP и LIME для текстов. - Анализ влияния слов на предсказания.
20	Тема 20. Интерпретация рекомендательных систем. Рассматриваемые вопросы: - Интерпретация рекомендаций. - Объяснение факторов, влияющих на рекомендации. - Визуализация рекомендаций. - Анализ результатов для бизнес-задач.
21	Тема 21. Валидация моделей на несбалансированных данных. Рассматриваемые вопросы: - Применение стратифицированной кросс-валидации. - Использование специальных метрик . - Анализ влияния дисбаланса на качество. - Методы борьбы с дисбалансом.
22	Тема 22. Валидация моделей на временных рядах. Рассматриваемые вопросы: - Применение временной кросс-валидации. - Прогнозирование и оценка качества. - Сравнение с простым разделением. - Анализ результатов.
23	Тема 23. Тестирование устойчивости моделей. Рассматриваемые вопросы: - Тестирование на шумовых данных. - Тестирование на выбросах. - Анализ устойчивости предсказаний. - Оценка робастности моделей.
24	Тема 24. Анализ ошибок моделей. Рассматриваемые вопросы: - Анализ ошибок на тестовой выборке. - Выявление систематических ошибок. - Визуализация ошибок. - Формулировка рекомендаций по улучшению.
25	Тема 25. Сравнительный анализ моделей. Рассматриваемые вопросы: - Обучение нескольких моделей на одном датасете. - Сравнение качества моделей. - Визуализация сравнения. - Обоснование выбора лучшей модели.
26	Тема 26. Обоснование выбора модели для бизнес-задачи. Рассматриваемые вопросы:

№ п/п	Тематика практических занятий/краткое содержание
	- Формулировка бизнес-задачи. - Выбор метрик и методов валидации. - Сравнение моделей с учётом бизнес-критериев. - Подготовка обоснования для стейкхолдеров.
27	Тема 27. Автоматизация отчётов о качестве моделей. Рассматриваемые вопросы: - Создание шаблонов отчётов. - Автоматическая генерация отчётов. - Визуализация результатов. - Настройка автоматической отправки отчётов.
28	Тема 28. Практика интерпретации для бизнес-задач. Рассматриваемые вопросы: - Формулировка бизнес-задачи. - Применение методов интерпретации. - Формулировка выводов для бизнеса. - Подготовка презентации для стейкхолдеров.
29	Тема 29. Комплексный проект: аналитика и проектирование. Рассматриваемые вопросы: - Формулировка прикладной задачи. - Выбор датасета и предобработка данных. - Проектирование эксперимента по валидации и интерпретации. - Планирование этапов работы.
30	Тема 30. Комплексный проект: реализация. Рассматриваемые вопросы: - Обучение моделей и оценка качества. - Применение методов валидации. - Применение методов интерпретации. - Сбор и анализ результатов.
31	Тема 31. Комплексный проект: формулировка выводов. Рассматриваемые вопросы: - Анализ полученных результатов. - Обоснование выбора модели. - Формулировка бизнес-рекомендаций. - Подготовка итогового отчёта.
32	Тема 32. Комплексный проект: защита. Рассматриваемые вопросы: - Подготовка презентации. - Защита проекта. - Демонстрация результатов валидации и интерпретации. - Обсуждение результатов с преподавателем.

4.3. Самостоятельная работа обучающихся.

№ п/п	Вид самостоятельной работы
1	Работа с лекционным материалом и дополнительной литературой.
2	Изучение документации по библиотекам для валидации и интерпретации.
3	Подготовка к практическим занятиям и выполнение домашних заданий.

№ п/п	Вид самостоятельной работы
4	Самостоятельное исследование методов валидации и интерпретации на различных датасетах.
5	Работа над комплексным проектом.
6	Подготовка к промежуточной аттестации и текущему контролю.
7	Подготовка к промежуточной аттестации.
8	Подготовка к текущему контролю.

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

№ п/п	Библиографическое описание	Место доступа
1	Попов, И. Ю. Валидация систем искусственного интеллекта : учебно-методическое пособие / И. Ю. Попов, А. Я. Бучаев, Д. А. Есипов. — Санкт-Петербург : НИУ ИТМО, 2024. — 29 с.	https://e.lanbook.com/book/460214
2	Ложников, П. С. Обеспечение безопасности моделей искусственного интеллекта : учебное пособие / П. С. Ложников. — Омск : ОмГТУ, 2025. — 132 с. — ISBN 978-5-8149-3941-8.	https://e.lanbook.com/book/526459
3	Чесалин, А. Н. Основы искусственного интеллекта с приложениями в информационной безопасности : учебное пособие / А. Н. Чесалин. — Москва : РТУ МИРЭА, 2021. — 155 с. — ISBN 978-5-7339-1589-0	https://e.lanbook.com/book/182429

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

Профессиональная база знаний. Документация библиотеки pandas.
<https://pandas.pydata.org/docs/>

Профессиональная база знаний. Документация библиотеки matplotlib.
<https://matplotlib.org/stable/users/index>

Профессиональная база знаний. Документация библиотеки seaborn.
<https://seaborn.pydata.org/tutorial.html>

Профессиональная база знаний. Документация библиотеки pandas_profiling.
<https://pandas-profiling.github.io/pandas-profiling/docs/master/index.html#pandas-profiling>

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

Профессиональная база знаний. Документация библиотеки lime <https://github.com/marcotcr/lime>

Электронно-библиотечная система Научно-технической библиотеки МИИТ - <http://library.miit.ru/>

Электронно-библиотечная система Лань - <https://e.lanbook.com/>

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

Python 3.x

Библиотеки: scikit-learn, pandas, numpy, matplotlib, seaborn, shap, lime

Jupyter Notebook

Система контроля версий Git

Офисный пакет приложений

Браузер

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

Компьютер преподавателя

Компьютеры студентов

Клавиатура

Мышка

Проектор

Экран для проектора

Маркерная доска

9. Форма промежуточной аттестации:

Зачет в 5 семестре.

10. Оценочные материалы.

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

Авторы:

заместитель директора

Б.В. Игольников

Согласовано:

Директор

Д.В. Паринов

Заместитель директора

Б.В. Игольников

Председатель учебно-методической
комиссии

Д.В. Паринов