СРМ №1

ИССЛЕДОВАНИЕ И СРАВНИТЕЛЬНЫЙ АНАЛИЗ СОВРЕМЕННЫХ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

Срок выполнения по календарно-тематическому плану: 3-я неделя.

Цель работы

Самостоятельно исследовать современные методы машинного обучения, провести их содержательное и экспериментальное сравнение и обосновать выбор методов, потенциально применимых в индивидуальном семестровом проекте.

Связь с дисциплиной и индивидуальным семестровым проектом

Работа выполняется после ПЗ №1–2 и параллельно с освоением Python в ПЗ №3. Результаты используются при выборе методов для последующих ПЗ №5–8 и индивидуального семестрового проекта.

Краткие теоретические сведения

Современное машинное обучение включает методы обучения с учителем, без учителя и другие подходы. На уровне магистратуры важно не перечислять алгоритмы, а сопоставлять их предпосылки, требования к данным, вычислительную сложность, интерпретируемость, устойчивость и качество на конкретной задаче.

Методические указания

Работа выполняется преимущественно самостоятельно. Магистрант сам определяет конкретные методы, данные, параметры эксперимента и форму представления результатов в рамках заданных требований. Необходимо отделять обзор литературы от собственных экспериментальных результатов. Все сравниваемые варианты должны оцениваться в сопоставимых условиях. Каждый вывод должен опираться либо на источник, либо на полученный экспериментальный результат.

При использовании AI-сервисов запрещается передавать конфиденциальные, персональные или служебные данные без разрешения. Код и вычислительные эксперименты должны быть воспроизводимыми; для стохастических методов фиксируется seed/random_state, а данные, метрики и параметры описываются в отчёте.

Порядок выполнения

1. Уточнить исследовательский вопрос текущей работы.
2. Определить исходные данные и критерии оценки.
3. Изучить и сравнить подходящие методы.
4. Выполнить практический эксперимент или разработать прототип.
5. Проанализировать ошибки и ограничения.
6. Связать результат с индивидуальным семестровым проектом.
7. Оформить выводы и подготовиться к защите.

Задания

Задание 1. Выберите одну прикладную задачу, связанную с темой индивидуального семестрового проекта.

Задание 2. Исследуйте не менее 5 методов машинного обучения, из которых минимум 3 должны быть пригодны для выбранной задачи.

Задание 3. Составьте сравнительную таблицу: принцип работы, тип данных, преимущества, ограничения, вычислительные требования, интерпретируемость и подходящие метрики.

Задание 4. Выберите минимум 2 метода и проведите небольшой воспроизводимый эксперимент на одном наборе данных.

Задание 5. Сравните результаты по одинаковым метрикам и объясните различия.

Задание 6. Сформулируйте обоснованный выбор baseline и альтернативного метода для дальнейшего развития проекта.

Пример выполнения

Для проекта обнаружения фишинговых URL можно исследовать Logistic Regression, Decision Tree, Random Forest, SVM и Gradient Boosting. Затем на одном train/test-разбиении сравнить Logistic Regression и Random Forest по precision, recall, F1 и ROC-AUC. Если Random Forest показывает выше F1, но требует больше ресурсов, это необходимо отразить в выводе; Logistic Regression может остаться интерпретируемым baseline.

Индивидуальное задание

Подготовьте аналитико-экспериментальный обзор методов именно для собственной темы. В конце укажите, какие 2–3 метода вы будете проверять в ПЗ №5–6 и почему.

Рекомендуемая форма представления результатов

Раздел	**Что должно быть представлено**
Постановка	Цель, исследовательский вопрос, связь с индивидуальным проектом
Данные/материалы	Источник, структура, ограничения и допустимость использования
Методика	Алгоритмы, параметры, критерии и схема эксперимента
Результаты	Таблицы, метрики, схемы, примеры или журналы работы
Анализ	Сравнение, ошибки, ограничения, риски
Решение по проекту	Что включается в итоговую систему и почему

Требования к отчёту

6–8 страниц: постановка задачи; обзор минимум 5 методов; сравнительная таблица; описание данных; эксперимент минимум с 2 методами; метрики; анализ ошибок; вывод для индивидуального проекта.

Критерии оценивания

Работа оценивается в соответствии с критериями силлабуса: аккуратность и точность — 30 %, творчество и креативность — 30 %, полнота и зрелость — 40 %. Оригинальность учитывается коэффициентом 1,0; 0,5 или 0. Работа должна быть самостоятельной, а использование готовых решений без собственного анализа снижает коэффициент оригинальности.

Контрольные вопросы

1. Какой исследовательский вопрос решается в данной работе?

2. Почему выбранные методы соответствуют поставленной задаче?

3. Какие альтернативные подходы были рассмотрены?

4. Какие данные использованы и какие у них ограничения?

5. По каким критериям сравнивались варианты решения?

6. Какие ошибки или нестабильные случаи выявлены?

7. Какие угрозы достоверности эксперимента необходимо учитывать?

8. Какие вопросы конфиденциальности и безопасности возникают?

9. Как результаты работы изменяют индивидуальный семестровый проект?

10. Какой следующий этап исследования необходимо выполнить?

Список источников

1. Бессмертный И. А., Авдюшина А. Е., Кугаевских А. В., Королева Ю. А., Рущенко Н. Г. Системы искусственного интеллекта : методическое пособие. – Санкт-Петербург : Университет ИТМО, 2024. – 281 с.

2. Соснило А. И. Атлас искусственного интеллекта для бизнеса и власти : учебное пособие. – Санкт-Петербург : Университет ИТМО, 2022. – 98 с.

3. Сидоркина И. Г. Системы искусственного интеллекта : учебное пособие. – Москва : КНОРУС, 2018. – 248 с.

4. Баланов А. Н. Искусственный интеллект. Понимание, применение и перспективы : учебник для вузов. – Санкт-Петербург : Лань, 2024. – 312 с.

5. Калинин М. О., Крундышев В. М. Основы искусственного интеллекта. Безопасность искусственного интеллекта : учебное пособие. – Санкт-Петербург, 2024. – 220 с.

6. Рассел С., Норвиг П. Искусственный интеллект: современный подход. – 2-е изд. – Москва : Вильямс, 2018. – 230 с.

7. Андреянов Н. В., Евдокимова Т. С., Павлов А. Д. [и др.]. Методы искусственного интеллекта : учебное пособие. – Казань : КНИТУ-КАИ, 2024. – 392 с.