

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Писарев Сергей Станиславович

Должность: Ректор

Негосударственное образовательное учреждение высшего образования

Дата подписания: 25.01.2026 18:15:46

Уникальный программный ключ:

b9d7463b91f434da3d4dc1afa9a0cf32d3c58650

«Школа управления СКОЛКОВО»



РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ Основы машинного обучения

Направление подготовки	38.03.02 Менеджмент
Квалификация выпускника	Бакалавр
Образовательная программа	Управление и предпринимательство
Форма обучения	Очная
Рабочая программа дисциплины разработана	

Трудоемкость	Контактная работа			Самостоятельная работа	Форма контроля	Семестр
	з.е.	часы	лекции	семинарские занятия		
5	180	24	24	132	Экзамен	3

Москва

2026

1. АННОТАЦИЯ ДИСЦИПЛИНЫ

Цель дисциплины “Основы машинного обучения” - сформировать у студентов систематизированные знания, практические умения и навыки в области машинного обучения, включая понимание основных методов, алгоритмов и принципов их применения для решения практических задач анализа данных.

2. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОСВОЕНИЯ ДИСЦИПЛИНЫ

В случае успешного освоения курса студенты будут:

знать

- Основные понятия, цели и типы задач машинного обучения;
- Принципы работы, достоинства и недостатки ключевых алгоритмов;
- Методы оценки и сравнения моделей, борьбы с переобучением;
- Основные библиотеки Python для машинного обучения;

уметь

- Формулировать задачу в терминах машинного обучения;
- Проводить разведочный анализ и предобработку данных;
- Реализовывать, обучать и тестировать модели с использованием готовых библиотек;
- Интерпретировать результаты работы алгоритмов;
- Визуализировать данные и результаты моделирования;

владеть

- Навыками применения классических алгоритмов ML для решения типовых задач;
- Технологиями кросс-валидации и настройки гиперпараметров моделей;
- Инструментами экосистемы Python для Data Science.

Дисциплина направлена на развитие следующих компетенций и их индикаторов:

Код компетенции	Формулировка компетенции и/или ее индикатора (ов)
ОПК-2.	Способен осуществлять сбор, обработку и анализ данных, необходимых для решения поставленных управлеченческих задач, с использованием современного инструментария и интеллектуальных информационно-аналитических систем
ОПК-2-1.	Определяет источники данных и выбирает методы и инструменты поиска, корректно осуществляет анализ литературы и документов
ОПК-2-2.	Применяет методы сбора, обработки и анализа данных, необходимых для решения управлеченческих задач, с использованием современных цифровых технологий, воспринимает, анализирует, запоминает и передает информацию с использованием цифровых средств, а также с помощью алгоритмов при работе с полученными из различных источниками данными
ОПК-5.	Способен использовать при решении профессиональных задач современные информационные технологии и программные средства, включая управление крупными массивами данных и их интеллектуальный анализ
ОПК-5-1.	Использует цифровые технологии, включая информационные системы и

	базы данных, системы искусственного интеллекта и системы интеллектуального анализа и обработки данных для решения профессиональных задач
ОПК-5-2.	Оценивает возможности и целесообразность использования цифровых технологий в деятельности организации, использует современные цифровые технологии и программные продукты для решения профессиональных задач
ОПК-6.	Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности
ОПК-6-1.	Знает процессы, методы поиска, сбора, хранения, обработки, представления, распространения информации и способы осуществления таких процессов и методов; современные инструментальные среды, программно-технические платформы и программные средства
ОПК-6.2.	Рационально выбирает современные информационные технологии для решения задач профессиональной деятельности

3. СОДЕРЖАНИЕ И СТРУКТУРА ДИСЦИПЛИНЫ

Название раздела/темы	Всего часов	Трудоемкость (час.) по видам учебных занятий			
		Контактная работа			Самостоятельная работа
		Всего	Лекции	семинары	
Тема 1. Введение в машинное обучение	30	8	4	4	22
Тема 2. Предобработка данных	30	8	4	4	22
Тема 3. Линейные модели и основы обучения	30	8	4	4	22
Тема 4. Ансамбли на основе деревьев решений	30	8	4	4	22
Тема 5. Оценка, валидация и настройка моделей	30	8	4	4	22
Тема 6. Обучение без учителя: кластеризация и поиск структуры	30	8	4	4	22
Итого	180	48	24	24	132

Тема 1. Введение в машинное обучение

Предмет и история машинного обучения. Типы обучения. Жизненный цикл ML-проекта.

Тема 2. Предобработка данных

Построение пайплайнов. Кодирование категориальных переменных (One-Hot, Label Encoding). Масштабирование числовых признаков (StandardScaler, MinMaxScaler). Разделение данных на обучающую и тестовую выборки.

Тема 3. Линейные модели и основы обучения

Линейная и логистическая регрессия. Основы оптимизации. Интерпретация весов признаков. Концепция регуляризации.

Тема 4. Ансамбли на основе деревьев решений

Недостатки простых моделей. Ансамблевых методов: бэггинга (Random Forest) и бустинга (Gradient Boosting).

Тема 5. Оценка, валидация и настройка моделей

Выбор метрик (Accuracy, Precision-Recall, ROC-AUC). Методы кросс-валидации (k-Fold). Автоматизированный подбор гиперпараметров (GridSearchCV).

Тема 6. Обучение без учителя: кластеризация и поиск структуры

Алгоритмы без учителя K-means. Снижение размерности (PCA) для визуализации и упрощения.

4. ОЦЕНОЧНЫЕ СРЕДСТВА И ПРИМЕРЫ ЗАДАНИЙ ДЛЯ ОЦЕНКИ РЕЗУЛЬТАТОВ ОСВОЕНИЯ ДИСЦИПЛИНЫ

4.1 Текущий контроль

Оценка за курс складывается из следующих видов заданий текущего контроля, каждый из которых обладает своим весом в общей системе:

Компоненты	Процент в итоговой оценке
Выполнение тестов по теоретической части	20%
Выполнение домашних заданий	30%
Презентация проекта	50%

На курсе используется 10 балльная система оценивания. За каждое задание студент получает от 1 до 10 баллов. Итоговый балл за каждый вид заданий рассчитывается как среднее арифметическое всех полученных баллов за все задания в рамках одного вида (О1, О2, О3). Невыполненное в срок задание оценивается в 0 баллов.

Общая оценка за курс (О) рассчитывается как:

$$O = O1 \times 0,2 + O2 \times 0,3 + O3 \times 0,5.$$

Если по результатам текущего контроля студент получил положительную оценку (не ниже “удовлетворительно”), оценка за промежуточную аттестацию выставляется автоматически.

Выполнение тестов по теоретической части

Данный компонент оценивается по следующим параметрам:

- количество верных ответов на поставленный вопрос;
- полнота ответа при наличии открытых вопросов.

Домашние задания

Выполнение домашних заданий оценивается по следующим параметрам:

- понимание описываемых процессов и явлений;

- полнота материала, свидетельствующая об освоении материала курса, грамотность использования терминов, определений, фактов;
- выполнение сроков сдачи.

Групповая работа над проектом

Оцениваются

- общий результат работы группы: задача выполнена в установленные сроки, приняли участие все члены группы в соответствии со своими ролями, результаты представлены в соответствии с заданными условиями.
- индивидуальный результат: студент действовал в соответствии со своей ролью, вклад в работу группы существенен, студент полностью владеет материалом, с которым работала группа, выражает готовность дополнить/исправить других студентов, четко отвечает на вопросы преподавателя.

4.2 Промежуточная аттестация

Студентам, набравшим достаточные для удовлетворительной оценки баллы за текущий контроль, оценка за дисциплину выставляется равной оценке за текущий контроль (См. п. 4.1)

Студентам, получившим неудовлетворительную оценку по результатам текущего контроля, необходимо по согласованию с преподавателем сдать один или несколько компонентов текущего контроля. Преподаватель вправе предложить студентам выполнить задание, не повторяющее задание текущего контроля, но проверяющее аналогичные знания, умения и навыки.

4.3 Примеры заданий

Примерные тестовые задания

1. Задача, в которой целевая переменная является категориальной (например, "спам"/"не спам"),

- A) Регрессия
 B) Классификация
 C) Кластеризация

D) Ранжирование

2. Какова основная цель разбиения данных на обучающую и тестовую выборки?

- A) Ускорить процесс обучения
 B) Оценить способность модели обобщать на новых, невиданных данных
 C) Уменьшить объем данных для обучения
 D) Сбалансировать классы

3. Почему важно масштабировать признаки перед использованием моделей, основанных на градиентном спуске или расстояниях (KNN, SVM, K-means)?

A) Чтобы ускорить сходимость градиентного спуска и избежать доминирования признаков с большим размахом

- B) Чтобы сделать данные красивыми для графиков
 C) Чтобы удалить выбросы
 D) Это необязательный шаг

Примерные домашние задания

Используйте библиотеки `numpy`, `pandas`, `matplotlib`. Не используйте `sklearn.linear_model` для реализации самой модели (только для сравнения).

1. Подготовка данных:

- Загрузите датасет `boston.csv` (или `california_housing.csv` для более современного аналога).
- Выберите один признак (например, `RM` – среднее количество комнат или `LSTAT` – процент населения низкого статуса) и целевую переменную `MEDV` (стоимость дома).
- Визуализируйте зависимость целевой переменной от выбранного признака с помощью `scatter plot`.
- Разбейте данные на обучающую и тестовую выборки (например, 80/20).

2. Реализация линейной регрессии:

- Реализуйте класс `LinearRegressionGD` с методами:

`__init__(self, learning_rate=0.01, n_iters=1000)` – конструктор.
`fit(self, X, y)` – метод обучения с помощью градиентного спуска.

- Инициализируйте веса `w` и `b` нулями.

Реализуйте цикл на `n_iters` итераций.

На каждой итерации вычисляйте предсказания, ошибку и градиенты .

Обновляйте веса `w` и `b` по правилу градиентного спуска.

- Обучите модель на тренировочных данных. Постройте график сходимости функции потерь (loss по итерациям).

3. Анализ и сравнение:

На одном графике изобразите:

- Исходные точки разброса (тестовые данные).
- Линию регрессии, полученную вашей моделью.

Вычислите метрики качества на тестовой выборке: MSE и R^2 (коэффициент детерминации). Реализуйте вычисление метрик самостоятельно.

- Обучите модель `LinearRegression` из библиотеки `sklearn` на тех же данных.

- Сравните полученные веса (`w`, `b`) и метрики (MSE , R^2) вашей реализации и реализации из `sklearn`. Результаты должны быть практически идентичными.

Сделайте вывод.

5. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

5.1 Литература

Платонов, А. В. Машинальное обучение : учебное пособие для вузов / А. В. Платонов. — 2-е изд. — Москва : Издательство Юрайт, 2025. — 89 с. — (Высшее образование). — ISBN 978-5-534-20732-3. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/558662> (дата обращения: 29.12.2025).

5.2 Электронные образовательные ресурсы

Материалы дисциплины размещены в LMS: <https://l.skolkovo.ru/login/index.php>

5.3 Профессиональные базы данных и информационные справочные системы (при наличии)

Справочная правовая система КонсультантПлюс <https://consultantkhv.ru/o-sisteme/>

6. ЛИЦЕНЗИОННОЕ И СВОБОДНО РАСПРОСТРАНЯЕМОЕ ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ

Операционная система Simple Linux, браузер Yandex браузер, антивирусное ПО Calmantivirus;

Свободно распространяемое ПО, в том числе отечественного производства:

Офисный пакет Libre Office, Okular PDF Reader, 7-Zip Архиватор, GIMP Редактирования фотографий, Inkscape Векторная графика, Blender 3D графика, Kdenlive Видеоредактор, Audacity Аудиоредактор, VLC Медиаплеер, Thunderbird Почтовый клиент, Flameshot Создание скриншотов

7. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Учебная аудитория для проведения занятий лекционного типа, оснащенная мультимедийным оборудованием, учебной мебелью, доской или со стенами с маркерным покрытием.

Учебная аудитория для проведения занятий семинарского типа, оснащенная мультимедийным оборудованием, учебной мебелью, доской или со стенами с маркерным покрытием.

Аудитория (коворкинг) для самостоятельной работы, оснащенная учебной мебелью, ноутбуками.

Материально-техническое обеспечение аудиторий представлено на официальном сайте <https://bbask.ru/sveden/objects/>.