МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ,
МИНИСТЕРСТВО НАУКИ, ВЫСШЕГО ОБРАЗОВАНИЯ И ИННОВАЦИЙ
КЫРГЫЗСКОЙ РЕСПУБЛИКИ

МОО ВО Кыргызско-Российский Славянский университет
имени первого Президента Российской Федерации Б.Н. Ельцина

УТВЕРЖДАЮ
директор Политехнического института
_____ Г.В. Попов
_3 сентября 2026 г.

Машинное обучение и анализ данных

аннотация дисциплины (модуля)

Разработано	**Высшая школа Цифровых технологий**
Учебный план	g090404_25_2пи_рпис.plx Направление подготовки 09.04.04 - РФ, 710400 - КР Программная инженерия Магистерская программа "Разработка программно-информационных систем"
Квалификация	**магистр**
Форма обучения	**очная**

Распределение часов дисциплины по семестрам

Семестр (<Курс>.<Семестр на курсе>)	**2 (1.2)**		Итого	
Недель	18			
Вид занятий	УП	РП	УП	РП
Лекции	16	16	16	16
Практические	22	22	22	22
Контактная работа в период теоретического обучения	1	1	1	1
В том числе инт.	8	8	8	8
В том числе в форме практ.подготовки	22	22	22	22
Итого ауд.	38	38	38	38
Контактная работа	39	39	39	39
Сам. работа	105	105	105	105
Итого	144	144	144	144

1. ЦЕЛИ ОСВОЕНИЯ ДИСЦИПЛИНЫ	
1.1	Цели освоения дисциплины (модуля)
1.2	Целью освоения дисциплины «Введение в машинное обучение» является формирование у студентов фундаментальных знаний, навыков и компетенций в области машинного обучения, позволяющих разрабатывать, анализировать и применять алгоритмы обработки данных для решения практических задач в программной инженерии.
1.3	
1.4	Задачи курса
1.5	Ознакомить студентов с ключевыми концепциями машинного обучения, его методами и областями применения.
1.6	Научить студентов использовать популярные библиотеки Python (NumPy, Pandas, Matplotlib, Scikit-Learn) для предобработки, анализа и визуализации данных.
1.7	Изучить основные алгоритмы машинного обучения, включая линейную и логистическую регрессию, деревья решений, метод опорных векторов, случайные леса и алгоритмы кластеризации.
1.8	Дать представление о методах повышения качества моделей, таких как регуляризация, кросс-валидация, подбор гиперпараметров и ансамблирование (бустинг, случайные леса).
1.9	Рассмотреть принципы обработки естественного языка (NLP) и их применение в анализе текстовых данных.
1.10	Научить студентов применять полученные знания для решения реальных задач, в том числе предсказательного моделирования, анализа данных и создания рекомендательных систем.
1.11	Развить практические навыки работы с инструментами машинного обучения и подготовки данных для моделирования.
1.12	В результате освоения дисциплины студенты получат базовые знания и практические навыки, необходимые для работы в сфере Data Science, искусственного интеллекта и анализа данных.

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ООП	
Цикл (раздел) ООП:	Б1.В
2.1	**Требования к предварительной подготовке обучающегося:**
2.1.1	Для успешного освоения дисциплины «Введение в машинное обучение» студент должен обладать следующими знаниями и навыками:
2.1.2	1. Математическая подготовка:
2.1.3	Основы линейной алгебры (матрицы, векторы, операции с ними).
2.1.4	Основы математического анализа (производные, интегралы, оптимизация).
2.1.5	Основы теории вероятностей и статистики (среднее, дисперсия, распределения, корреляция).
2.1.6	2. Программирование и алгоритмы:
2.1.7	Базовые знания языка Python (переменные, функции, циклы, условия).
2.1.8	Опыт работы с основными структурами данных (списки, словари, множества, кортежи).
2.1.9	Понимание принципов алгоритмизации и написания программ.
2.1.10	3. Работа с инструментами и библиотеками:
2.1.11	Опыт работы в Jupyter Notebook или других средах разработки Python.
2.1.12	Знание основ работы с библиотеками NumPy, Pandas, Matplotlib.
2.1.13	4. Базовые знания в области обработки данных:
2.1.14	Чтение, запись и обработка данных в различных форматах (CSV, JSON, Excel).
2.1.15	Основы работы с базами данных и SQL (желательно).
2.1.16	Навыки визуализации данных.
2.1.17	Желательно, но не обязательно:
2.1.18	Опыт работы с базами данных и SQL.
2.1.19	Знание основ объектно-ориентированного программирования (ООП).
2.1.20	Понимание принципов оптимизации и многопоточного программирования.
2.1.21	Эти знания и навыки позволят студентам эффективно осваивать курс, понимать основные концепции машинного обучения и применять их на практике.
2.1.22	Теория вероятностей и математическая статистика
2.1.23	Математические основы программной инженерии
2.1.24	Структурное программирование
2.1.25	Вычислительная математика
2.1.26	Алгоритмы и структуры данных
2.1.27	Методы оптимизации
2.2	**Дисциплины и практики, для которых освоение данной дисциплины (модуля) необходимо как предшествующее:**
2.2.1	Освоение дисциплины «Введение в машинное обучение» является важным этапом подготовки студентов и служит основой для дальнейшего изучения ряда дисциплин и практик в области программной инженерии, анализа данных и искусственного интеллекта.

2.2.2	1. Дисциплины (модули), опирающиеся на данный курс:
2.2.3	Машинное обучение и глубокое обучение
2.2.4	(Углубленное изучение методов машинного обучения, нейросетей, алгоритмов глубокого обучения, работа с фреймворками TensorFlow и PyTorch.)
2.2.5	Анализ данных и Data Science
2.2.6	(Развитие компетенций по сбору, обработке и анализу данных, работа с большими данными.)
2.2.7	Компьютерное зрение
2.2.8	(Использование алгоритмов машинного обучения для обработки и анализа изображений, нейросетевые методы компьютерного зрения.)
2.2.9	Обработка естественного языка (NLP)
2.2.10	(Продвинутые методы обработки текстовых данных, работа с моделями на основе глубоких нейросетей.)
2.2.11	Разработка интеллектуальных систем
2.2.12	(Создание рекомендательных систем, чат-ботов, экспертных систем.)
2.2.13	Методы искусственного интеллекта
2.2.14	(Исследование различных подходов к созданию интеллектуальных систем, включая символические методы, алгоритмы обучения и адаптивные системы.)
2.2.15	2. Практики, требующие знаний по данному курсу:
2.2.16	Практика по анализу данных и моделированию
2.2.17	(Работа с реальными датасетами, применение методов машинного обучения, написание аналитических отчетов.)
2.2.18	Научно-исследовательская практика
2.2.19	(Разработка и тестирование алгоритмов машинного обучения, работа над научными статьями и исследовательскими проектами.)
2.2.20	Преддипломная практика
2.2.21	(Применение методов машинного обучения и анализа данных в реальных проектах.)
2.2.22	Проектная деятельность по программной инженерии
2.2.23	(Разработка программных решений с элементами машинного обучения, анализ данных для бизнес-приложений.)
2.2.24	Таким образом, данная дисциплина закладывает фундаментальные знания и навыки, необходимые для дальнейшего изучения искусственного интеллекта, анализа данных, компьютерного зрения, обработки текстов и разработки интеллектуальных систем.
2.2.25	Выполнение и защита выпускной квалификационной работы
2.2.26	Технологическая (проектно-технологическая) практика
2.2.27	Технологическая (проектно-технологическая) практика
2.2.28	Преддипломная практика 2
2.2.29	Преддипломная практика 1
2.2.30	Нейронные сети

3. КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)

ПК-4: Владение навыками создания программного обеспечения для анализа, распознавания и обработки информации, систем цифровой обработки сигналов

Знать:	
Уровень 1	- методы создания программного обеспечения для анализа, распознавания и обработки информации, систем цифровой обработки сигналов.
Уметь:	
Уровень 1	- использовать методы создания программного обеспечения для анализа, распознавания и обработки информации, систем цифровой обработки сигналов.
Владеть:	
Уровень 1	- методами и средствами создания программного обеспечения для анализа, распознавания и обработки информации, систем цифровой обработки сигналов с помощью современных компьютерных

В результате освоения дисциплины обучающийся должен

3.1	**Знать:**
Основные принципы моделирования данных и методы предобработки данных. Теоретические основы машинного обучения, включая методы обучения с учителем и без учителя. Методы регрессии, классификации, кластеризации и принципы их работы. Основы статистического анализа и способы оценки качества моделей. Базовые алгоритмы и принципы работы нейросетевых моделей. Методы визуализации данных для интерпретации результатов моделирования. Принципы обработки естественного языка (NLP) и их применение. Основы программирования на Python и работы с библиотеками NumPy, Pandas, Matplotlib, Scikit-Learn. Подходы к построению экспериментальных исследований в области машинного обучения.	

3.2	Уметь:

Применять методы предобработки данных (очистка, нормализация, обработка пропущенных значений, кодирование категориальных признаков).
Разрабатывать и реализовывать алгоритмы машинного обучения для решения задач регрессии, классификации, кластеризации.
Использовать Python и специализированные библиотеки (NumPy, Pandas, Matplotlib, Seaborn, Scikit-Learn) для анализа, обработки и визуализации данных.
Выполнять оценку качества моделей машинного обучения с использованием метрик (MSE, RMSE, R^2, Accuracy, Precision, Recall, F1-score, ROC-AUC).
Применять методы кросс-валидации и подбора гиперпараметров для повышения точности моделей.
Реализовывать и интерпретировать методы регуляризации (Lasso, Ridge, Elastic Net) для предотвращения переобучения.
Работать с различными алгоритмами классификации (логистическая регрессия, SVM, деревья решений, случайные леса, градиентный бустинг).
Использовать методы кластеризации (K-Means, DBSCAN, иерархическая кластеризация) и анализа главных компонент (PCA) для снижения размерности данных.
Применять алгоритмы обработки естественного языка (NLP), включая наивный Байес, TF-IDF, Bag of Words.
Создавать и анализировать машинные модели на реальных данных, а также представлять результаты исследований в виде отчетов и визуализаций.
Разрабатывать прототипы интеллектуальных систем для автоматизации аналитических задач.

3.3	Владеть:

Работа с Python и его основными библиотеками (NumPy, Pandas, Matplotlib, Seaborn, Scikit-Learn).
Использование Jupyter Notebook и Anaconda для выполнения вычислений и визуализации данных.
Обработка и анализ разнородных наборов данных (CSV, Excel, SQL, JSON, API-данные).
Методами машинного обучения:

Применение регрессионных алгоритмов (линейная, полиномиальная регрессия, регуляризация).
Использование методов классификации (логистическая регрессия, KNN, SVM, деревья решений, случайные леса).
Разработка моделей кластеризации (K-Means, DBSCAN, иерархическая кластеризация).
Применение алгоритмов ансамблирования (бустинг, случайные леса, градиентный бустинг).
Владение методами подбора гиперпараметров (кросс-валидация, GridSearch, RandomSearch).
Навыками обнаружения выбросов и работы с пропущенными значениями.
Методами предобработки данных и визуализации:

Владение методами нормализации и стандартизации данных.
Использование методов Feature Engineering (создание новых признаков, их отбор и преобразование).
Визуализация данных с помощью Seaborn и Matplotlib.
Основами обработки естественного языка (NLP):

Работа с текстовыми данными (Tokenization, Stop Words, Lemmatization, Stemming).
Извлечение признаков с помощью TF-IDF, Bag of Words.
Классификация текстов с использованием наивного Байесовского классификатора.
Навыками работы с моделями машинного обучения в продакшене:
Деплой моделей с использованием Flask/FastAPI, Docker.
Владение базовыми принципами DevOps для работы с моделями ML.
Оптимизация производительности моделей для работы с большими данными.