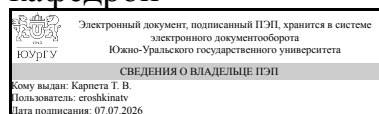


ЮЖНО-УРАЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

УТВЕРЖДАЮ:
Заведующий выпускающей
кафедрой



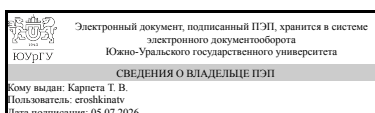
Т. В. Карпета

РАБОЧАЯ ПРОГРАММА

дисциплины 1.Ф.П0.05 Алгоритмы машинного обучения
для направления 01.03.02 Прикладная математика и информатика
уровень Бакалавриат
профиль подготовки Прикладная математика и искусственный интеллект
форма обучения очная
кафедра-разработчик Прикладная математика и программирование

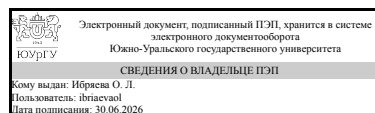
Рабочая программа составлена в соответствии с ФГОС ВО по направлению подготовки 01.03.02 Прикладная математика и информатика, утверждённым приказом Минобрнауки от 10.01.2018 № 9

Зав.кафедрой разработчика,
к.физ.-мат.н.



Т. В. Карпета

Разработчик программы,
д.техн.н., доцент



О. Л. Ибряева

1. Цели и задачи дисциплины

Целью дисциплины является формирование у студентов фундаментальных теоретических знаний и практических навыков в области машинного обучения, понимания математических основ и принципов работы основных алгоритмов. В ходе изучения курса студенты осваивают методы построения, обучения и оценки моделей для задач классификации, регрессии, кластеризации и ранжирования, учатся применять метрики качества и интерпретировать результаты. Особое внимание уделяется глубокому пониманию работы алгоритмов «под капотом» через ручные вычисления и геометрическую интерпретацию, что позволяет осознанно выбирать и настраивать модели для решения прикладных задач.

Краткое содержание дисциплины

Дисциплина охватывает ключевые разделы машинного обучения: введение в ML и метод ближайших соседей, метрики качества классификации и регрессии, линейную и логистическую регрессию с методами оптимизации, метод опорных векторов с жёстким и мягким зазором. Изучаются деревья решений и ансамбли моделей, методы отбора признаков и интерпретируемости, автоматизированное машинное обучение (AutoML). Рассматриваются алгоритмы понижения размерности (PCA, t-SNE, PAMAP), методы кластеризации (K-Means, DBSCAN, иерархическая кластеризация), задача ранжирования и метрики ROC-AUC, Precision-Recall.

2. Компетенции обучающегося, формируемые в результате освоения дисциплины

Планируемые результаты освоения ОП ВО (компетенции)	Планируемые результаты обучения по дисциплине
ПК-5 (ПК-4 модели) Способен разрабатывать и применять методы машинного обучения для решения задач	Знает: [ПК-4.1. 3-1.] принципы и методы машинного обучения, типы и классы задач машинного обучения, методологию ML Ops; [ПК-4.1. 3-2.] статистические методы анализа данных; [ПК-4.3. 3-1.] классические методы и алгоритмы машинного обучения: предиктивные - обучение с учителем, дескриптивные - обучение без учителя Умеет: [ПК-4.1. У-1.] сопоставить задачам предметной области классы задач машинного обучения; [ПК-4.1. У-2.] использовать статистические методы анализа данных при решении задач машинного обучения; [ПК-4.3. У-1.] проводить сравнительный анализ и осуществлять выбор, настройку при необходимости, разработку методов и алгоритмов для решения задач машинного обучения Имеет практический опыт: участия в разработке алгоритмов для решения задач машинного обучения
ПК-8 (ПК-5 модели) Способен использовать инструментальные средства для решения задач машинного обучения	Умеет: [ПК-5.1. У-1.] проводить сравнительный анализ и осуществлять выбор инструментальных средств для решения задач машинного обучения

3. Место дисциплины в структуре ОП ВО

Перечень предшествующих дисциплин, видов работ учебного плана	Перечень последующих дисциплин, видов работ
Программирование для анализа данных	Современные архитектуры глубоких искусственных нейронных сетей, Анализ требований и проектирование систем искусственного интеллекта, Высокопроизводительные параллельные вычисления, Производственная практика (преддипломная) (8 семестр), Производственная практика (практика по получению профессиональных умений и опыта профессиональной деятельности) (8 семестр)

Требования к «входным» знаниям, умениям, навыкам студента, необходимым при освоении данной дисциплины и приобретенным в результате освоения предшествующих дисциплин:

Дисциплина	Требования
Программирование для анализа данных	Знает: [ПК-5.1. 3-1.] возможности современных инструментальных средств и систем программирования для решения задач анализа данных и машинного обучения Умеет: [ПК-2.2. У-1.] разрабатывать программные приложения систем искусственного интеллекта, с использованием современных языков программирования, библиотек и программных платформ функционального, логического, объектно-ориентированного программирования (Python, R, C++, C#) Имеет практический опыт: участия в разработке программных приложений систем искусственного интеллекта, с использованием современных языков программирования

4. Объём и виды учебной работы

Общая трудоемкость дисциплины составляет 3 з.е., 108 ч., 72,5 ч. контактной работы

Вид учебной работы	Всего часов	Распределение по семестрам в часах
		Номер семестра
		5
Общая трудоёмкость дисциплины	108	108
Аудиторные занятия:	64	64
Лекции (Л)	32	32
Практические занятия, семинары и (или) другие виды аудиторных занятий (ПЗ)	0	0

Лабораторные работы (ЛР)	32	32
Самостоятельная работа (СРС)	35,5	35,5
Изучение материала в пособии "Отбор признаков"	2	2
Изучение материала в пособии "Линейная и логистическая регрессия"	2	2
Решение индивидуальных типовых заданий	31,5	31.5
Консультации и промежуточная аттестация	8,5	8,5
Вид контроля (зачет, диф.зачет, экзамен)	-	экзамен

5. Содержание дисциплины

№ раздела	Наименование разделов дисциплины	Объем аудиторных занятий по видам в часах			
		Всего	Л	ПЗ	ЛР
1	Основы машинного обучения и линейные модели	22	10	0	12
2	Методы обучения с учителем и автоматизация построения моделей	26	12	0	14
3	Методы обучения без учителя и специализированные задачи машинного обучения	16	10	0	6

5.1. Лекции

№ лекции	№ раздела	Наименование или краткое содержание лекционного занятия	Кол-во часов
1	1	Введение в ML. Метод ближайших соседей. Проклятие размерности. Наивный Байес.	2
2	1	Метрики оценки качества моделей классификации и регрессии.	2
3	1	Линейная регрессии. Целевая функция. Метод градиентного спуска.	2
4	1	Нормальное уравнение. Линейная регрессия для нелинейных моделей.	2
5	1	Логистическая регрессия. Бинарная и многоклассовая классификация. Регуляризация.	2
6	2	Метод опорных векторов. Жесткий зазор. Hinge loss.	2
7	2	SVM с мягким зазором. Ядерный трюк. Кросс-валидация. Решетчатый поиск.	2
8	2	Деревья решений. Критерии разбиения, борьба с переобучением.	2
9	2	Ансамбли деревьев решений. Важность признаков.	2
10	2	Отбор признаков, интерпретируемость.	2
11	2	Автоматизированное машинное обучение (AutoML)	2
12	3	Понижение размерности. PCA, t-SNE, UMAP, t-SNE	2
13	3	Кластеризация. K-Means, DBSCAN.	2
14	3	Иерархическая кластеризация. Метрики оценки качества кластеризации.	2
15	3	Задача ранжирования. Метрики оценки качества ранжирования.	2
16	3	Метрика ROC-AUC, кривая точности- полноты. Методы decision_function и predict_proba, настройка порогов для редких классов.	2

5.2. Практические занятия, семинары

Не предусмотрены

5.3. Лабораторные работы

№ занятия	№ раздела	Наименование или краткое содержание лабораторной работы	Кол-во часов
1	1	Метод ближайших соседей и наивный байесовский классификатор.	2
2	1	Лабораторная работа №1 «Метод ближайших соседей и наивный байесовский классификатор».	2
3	1	Контрольная работа №1.	2
4	1	Лабораторная работа №2 «Линейная регрессия». Сдача лабораторной работы №1.	2
5	1	Лабораторная работа №3 «Логистическая регрессия».	2
6	1	Сдача лабораторной работы №2.	2
7	2	Лабораторная работа №4 «Метод опорных векторов». Сдача лабораторной работы №3.	2
8	2	Контрольная работа №2.	2
9	2	Лабораторная работа №5 «Ансамблевые методы машинного обучения». Сдача лабораторной работы №4.	2
10	2	Постановка задачи мини-проекта и выбор набора данных.	2
11	2	Разработка и анализ моделей машинного обучения для мини-проекта. Использование средств AutoML. Сдача лабораторной работы №5.	2
12	2	Сдача Jupyter Notebook для мини-проекта и промежуточный контроль результатов.	2
13	2	Защита мини-проектов.	2
14	3	Лабораторная работа №6 «Кластеризация и понижение размерности данных».	2
15	3	Лабораторная работа №7 «Методы ранжирования объектов». Сдача лабораторной работы №6.	2
16	3	Контрольная работа №3. Сдача лабораторной работы №7.	2

5.4. Самостоятельная работа студента

Выполнение СРС			
Подвид СРС	Список литературы (с указанием разделов, глав, страниц) / ссылка на ресурс	Семестр	Кол-во часов
Изучение материала в пособии "Отбор признаков"	приложенное методическое пособие	5	2
Изучение материала в пособии "Линейная и логистическая регрессия"	приложенное методическое пособие	5	2
Решение индивидуальных типовых заданий	приложенные файлы с индивидуальными заданиями	5	31,5

6. Фонд оценочных средств для проведения текущего контроля успеваемости, промежуточной аттестации

Контроль качества освоения образовательной программы осуществляется в соответствии с Положением о балльно-рейтинговой системе оценивания результатов учебной деятельности обучающихся.

6.1. Контрольные мероприятия (КМ)

№	Се-	Вид	Название	Вес	Макс.	Порядок начисления баллов	Учи-
---	-----	-----	----------	-----	-------	---------------------------	------

КМ	местр	контроля	контрольного мероприятия		балл		тыва- ется в ПА
1	5	Текущий контроль	Лабораторная работа 1	1	7	Максимальная оценка за выполнение лабораторной работы составляет 6 баллов + 1 балл за сдачу в установленный срок. Один балл начисляется за корректную загрузку данных, проведение разведочного анализа и стратифицированное разделение выборок. Два балла начисляются за обучение моделей (kNN, GaussianNB, MultinomialNB), исследование влияния числа соседей и построение границ решений. Один балл начисляется за визуализацию матриц ошибок и попарных зависимостей признаков. Один балл начисляется за расчет метрик классификации и регрессии, а также анализ выбросов. Один балл начисляется за формулировку обоснованных выводов и ответы на контрольные вопросы при защите.	экзамен
2	5	Текущий контроль	Лабораторная работа 2	1	7	Максимальная оценка за выполнение лабораторной работы составляет 6 баллов + 1 балл за сдачу в установленный срок. Один балл начисляется за загрузку данных, визуализацию и реализацию функции вычисления среднеквадратичной ошибки. Два балла начисляются за самостоятельную реализацию алгоритма градиентного спуска, проведение экспериментов с шагом обучения и построение множественной регрессии с нормализацией признаков. Один балл начисляется за визуализацию прямой регрессии и графиков сходимости функции потерь. Один балл начисляется за сравнение ручной реализации с библиотечной и построение полиномиальной регрессии. Один балл начисляется за наличие содержательных выводов и ответы на контрольные вопросы при защите.	экзамен
3	5	Текущий контроль	Лабораторная работа 3	1	7	Максимальная оценка за выполнение лабораторной работы составляет 6 баллов + 1 балл за сдачу в установленный срок. Один балл начисляется за подготовку данных и самостоятельную реализацию сигмоиды и функции перекрестной энтропии. Два балла начисляются за оптимизацию параметров, построение линейных и нелинейных границ решений, а также	экзамен

						исследование влияния коэффициента регуляризации на переобучение. Один балл начисляется за визуализацию границ решений для различных значений параметра регуляризации. Один балл начисляется за реализацию многоклассовой классификации или решение задачи средствами sklearn с генерацией полиномиальных признаков. Один балл начисляется за формулировку выводов и ответы на контрольные вопросы при защите.	
4	5	Текущий контроль	Лабораторная работа 4	1	7	Максимальная оценка за выполнение лабораторной работы составляет 6 баллов + 1 балл за сдачу в установленный срок. Один балл начисляется за загрузку линейно и нелинейно разделимых выборок и их первичную визуализацию. Два балла начисляются за обучение линейного и нелинейного SVM, исследование влияния параметров C и γ , а также подбор гиперпараметров с помощью кросс-валидации. Один балл начисляется за визуализацию разделяющих гиперплоскостей и границ решений для различных ядер. Один балл начисляется за анализ влияния предварительной нормализации данных на метрики качества и решение прикладной задачи классификации. Один балл начисляется за наличие обоснованных выводов и ответы на контрольные вопросы при защите.	экзамен
5	5	Текущий контроль	Лабораторная работа 5	1	7	Максимальная оценка за выполнение лабораторной работы составляет 6 баллов + 1 балл за сдачу в установленный срок. Один балл начисляется за подготовку данных и базовое обучение моделей случайного леса и градиентного бустинга. Два балла начисляются за настройку гиперпараметров ансамблей с помощью решетчатого поиска и кросс-валидации, а также построение кривых обучения. Один балл начисляется за визуализацию и сравнение важности признаков методами MDI и Permutation Importance. Один балл начисляется за итоговое сравнение метрик качества и времени обучения одиночного дерева, случайного леса и бустинга. Один балл начисляется за формулировку выводов и ответы на контрольные вопросы при защите.	экзамен
6	5	Текущий	Лабораторная	1	7	Максимальная оценка за выполнение	экзамен

		контроль	работа 6			лабораторной работы составляет 6 баллов + 1 балл за сдачу в установленный срок. Один балл начисляется за нормализацию данных и применение методов понижения размерности (PCA, t-SNE, UMAP) с визуализацией результатов. Два балла начисляются за применение алгоритмов кластеризации (K-Means, DBSCAN, иерархическая), определение оптимального числа кластеров методом локтя и построение дендрограмм. Один балл начисляется за визуализацию кластеров в пространстве главных компонент и анализ обнаруженных выбросов. Один балл начисляется за расчет и сравнение метрик качества кластеризации. Один балл начисляется за наличие выводов и ответы на контрольные вопросы при защите.	
7	5	Текущий контроль	Лабораторная работа 7	1	7	Максимальная оценка за выполнение лабораторной работы составляет 6 баллов + 1 балл за сдачу в установленный срок. Один балл начисляется за корректную загрузку данных, разведочный анализ и специфическое разделение выборок по идентификаторам запросов. Два балла начисляются за обучение модели ранжирования, подбор гиперпараметров и анализ предсказанных скоров. Один балл начисляется за расчет специализированных метрик ранжирования (NDCG, MAP) и построение графиков сходимости модели. Один балл начисляется за визуализацию важности признаков и сравнение качества модели до и после настройки. Один балл начисляется за формулировку обоснованных выводов и ответы на контрольные вопросы при защите.	экзамен
8	5	Текущий контроль	Мини-проект	1	15	Мини-проект оценивается в 15 баллов, которые распределяются между тремя этапами сдачи следующим образом. За первый этап, утверждение набора данных и проведение разведочного анализа, начисляется 3 балла. Критериями оценки на данном этапе являются обоснованность выбора набора данных, соответствие формальным требованиям по объёму и размерности, содержательность поставленной задачи и её практическая значимость, а также глубина и качество выполненного разведочного анализа	экзамен

					данных с визуализацией. За второй этап, сдачу промежуточного Colab-ноутбука с первыми моделями, начисляется 4 балла. Критериями оценки выступают корректность предобработки данных, наличие работающей baseline-модели, применение хотя бы одного метода машинного обучения, воспроизводимость результатов и аккуратность оформления кода. За третий этап, включающий финальный Colab-ноутбук, презентацию и устную защиту, начисляется 8 баллов, из которых 4 балла относятся к качеству финального ноутбука и 4 балла — к качеству презентации и защиты. Критериями оценки финального ноутбука являются полнота применения заявленных методов, наличие и содержательность результатов понижения размерности, корректное применение AutoML, обоснованность сравнения моделей, качество кода и оформления. Критериями оценки презентации и защиты являются качество визуального оформления слайдов, ясность и структурированность устного выступления, глубина понимания применённых методов, способность обосновать выбор моделей и гиперпараметров, качество интерпретации результатов, а также индивидуальный вклад каждого участника команды и ответы на вопросы преподавателя и других студентов. За нарушение сроков сдачи на каждом из трёх этапов предусмотрено снижение итоговой оценки на 1 балл. Таким образом, максимальное суммарное снижение оценки за нарушение дедлайнов составляет 3 балла. Штраф начисляется независимо от качества выполненной работы и применяется при факте просрочки сдачи этапа без уважительной причины.	
9	5	Текущий контроль	Индивидуальные задания 1	1	6	экзамен

						Максимальный балл составляет шесть баллов.	
10	5	Текущий контроль	Контрольная работа 1	1	6	Контрольная работа оценивается по шестибалльной системе, где за каждую из трёх задач начисляется два балла при условии полного и корректного выполнения всех вычислений. Задания включают применение методов классификации, мультиномиальный байесовский классификатор для текста и расчёт метрик качества. Максимальный балл за контрольную работу составляет шесть баллов.	экзамен
11	5	Текущий контроль	Индивидуальные задания 2	1	6	Задание оценивается по шестибалльной системе, где за каждую из шести задач начисляется один балл при условии полного и верного выполнения всех требуемых вычислений и получения корректного итогового результата. Задания охватывают линейную регрессию (решение через нормальное уравнение и метод градиентного спуска), логистическую регрессию (анализ границы решений и обучение модели методом градиентного спуска), метод опорных векторов с жёстким зазором, а также функцию потерь Hinge Loss. Максимальный балл составляет шесть баллов.	экзамен
12	5	Текущий контроль	Контрольная работа 2	1	6	Контрольная работа оценивается по шестибалльной системе, где за каждую из трёх задач начисляется два балла при условии полного и корректного выполнения всех вычислений. Задания включают решение задачи линейной регрессии (нормальным уравнением или градиентным спуском), анализ или обучение модели логистической регрессии, а также построение разделяющей гиперплоскости в методе опорных векторов или вычисление функции потерь Hinge Loss. Максимальный балл за контрольную работу составляет шесть баллов.	экзамен
13	5	Текущий контроль	Индивидуальные задания 3	1	6	Задание оценивается по шестибалльной системе, где за каждую из шести задач начисляется один балл при условии полного и верного выполнения всех требуемых вычислений и получения корректного итогового результата. Задания охватывают деревья решений для задач классификации и регрессии, агломеративную иерархическую кластеризацию, алгоритм K-Means и индекс силуэта, метод главных компонент, а также метрики качества	экзамен

						ранжирования. Максимальный балл составляет шесть баллов.	
14	5	Текущий контроль	Контрольная работа 3	1	6	Контрольная работа оценивается по шестибалльной системе, где за каждую из трёх задач начисляется два балла при условии полного и корректного выполнения всех вычислений и построения требуемых графических иллюстраций. Задания включают построение деревьев решений, задачи кластеризации, метод главных компонент и расчёт метрик качества ранжирования. Максимальный балл за контрольную работу составляет шесть баллов.	экзамен
15	5	Промежуточная аттестация	Экзамен	-	100	Каждый теоретический вопрос оценивается максимум в 25 баллов, каждая практическая задача — максимум в 25 баллов, итого 100 баллов за билет. Сто баллов выставляется, если на оба теоретических вопроса даны полные ответы с корректными определениями, ключевыми формулами и примерами, а обе практические задачи решены верно с подробными пошаговыми выкладками и обоснованными ответами.	экзамен

6.2. Процедура проведения, критерии оценивания

Вид промежуточной аттестации	Процедура проведения	Критерии оценивания
экзамен	Студент получает экзаменационный билет, содержащий два теоретических вопроса и две практические задачи, и в течение 60–70 минут подготавливает письменные решения в черновике. После этого в течение 20–30 минут происходит устная защита билета перед преподавателем, в ходе которой студент отвечает на теоретические вопросы, объясняет ход решения практических задач и отвечает на уточняющие вопросы. Разрешается использовать непрограммируемый калькулятор, справочные материалы не допускаются.	В соответствии с пп. 2.5, 2.6 Положения

6.3. Паспорт фонда оценочных средств

Компетенции	Результаты обучения	№ КМ														
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
ПК-5	Знает: [ПК-4.1. 3-1.] принципы и методы машинного обучения, типы и классы задач машинного обучения, методологию ML Ops; [ПК-4.1. 3-2.] статистические методы анализа данных; [ПК-4.3. 3-1.] классические методы и алгоритмы машинного обучения: предиктивные - обучение с учителем, дескриптивные - обучение без учителя	+	+	+	+	+		+		+		+			+	+

			22.06.2026). — Режим доступа: для авториз. пользователей.
2	Основная литература	ЭБС издательства Лань	Мэрфи, К. П. Вероятностное машинное обучение. Введение / К. П. Мэрфи ; перевод с английского А. А. Слинкина. — Москва : ДМК Пресс, 2022. — 940 с. — ISBN 978-5-93700-119-1. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/314891 (дата обращения: 22.06.2026). — Режим доступа: для авториз. пользователей.
3	Основная литература	ЭБС издательства Лань	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. — URL: https://e.lanbook.com/book/69955 (дата обращения: 22.06.2026). — Режим доступа: для авториз. пользователей.
4	Дополнительная литература	ЭБС издательства Лань	Хуттер, Ф. Введение в автоматизированное машинное обучение (AutoML) : справочник / Ф. Хуттер, Л. Коттхофф, Х. Ваншорен ; перевод с английского В. С. Яценкова.. — Москва : ДМК Пресс, 2023. — 256 с. — ISBN 978-5-93700-196-2. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/348104 (дата обращения: 22.06.2026). — Режим доступа: для авториз. пользователей.

Перечень используемого программного обеспечения:

1. -Python(бессрочно)

Перечень используемых профессиональных баз данных и информационных справочных систем:

Нет

8. Материально-техническое обеспечение дисциплины

Вид занятий	№ ауд.	Основное оборудование, стенды, макеты, компьютерная техника, предустановленное программное обеспечение, используемое для различных видов занятий
Лекции	336 (36)	Доска, компьютер, проектор
Дифференцированный зачет	333 (36)	Доска, компьютеры, проектор
Лабораторные занятия	333 (36)	Доска, компьютеры, проектор