ЮЖНО-УРАЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

УТВЕРЖДАЮ:
Заведующий выпускающей кафедрой

Электронный документ, подписанный ПЭП, хранится в системе электронного документооборота Южно-Уральского государственного университета
СВЕДЕНИЯ О ВЛАДЕЛЬЦЕ ПЭП
Кому выдан: Замышляева А. А.
Пользователь: zamyshliaevaaa
Дата подписания: 08.09.2022

А. А. Замышляева

РАБОЧАЯ ПРОГРАММА

дисциплины 1.Ф.П1.12 Машинное обучение
для направления 01.03.02 Прикладная математика и информатика
уровень Бакалавриат
профиль подготовки Математическое и программное обеспечение вычислительных машин и систем
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,
д.физ.-мат.н., проф.

Электронный документ, подписанный ПЭП, хранится в системе электронного документооборота Южно-Уральского государственного университета
СВЕДЕНИЯ О ВЛАДЕЛЬЦЕ ПЭП
Кому выдан: Замышляева А. А.
Пользователь: zamyshliaevaaa
Дата подписания: 08.09.2022

А. А. Замышляева

Разработчик программы,
старший преподаватель

Электронный документ, подписанный ПЭП, хранится в системе электронного документооборота Южно-Уральского государственного университета
СВЕДЕНИЯ О ВЛАДЕЛЬЦЕ ПЭП
Кому выдан: Ножкина Т. Г.
Пользователь: nozhkinatg
Дата подписания: 06.09.2022

Т. Г. Ножкина

Челябинск

1. Цели и задачи дисциплины

Формирование теоретических знаний по основам машинного обучения для построения формальных математических моделей и интерпретации результатов моделирования; выработка умения по практическому применению методов машинного обучения для построения формальных математических моделей и интерпретации результатов моделирования при решении прикладных задач в различных прикладных областях и умения и навыка использования различных программных инструментов анализа баз данных и систем машинного обучения.

Краткое содержание дисциплины

В процессе обучения изучаются математические основы машинного обучения, особенности и область применения методов машинного обучения. Данная учебная дисциплина реализуется как цикл лекционных и лабораторных занятий, которые знакомят студентов с теоретическими основами и алгоритмами машинного обучения, их возможными практическими реализациями и применением при решении реальных задач. В рамках данного курса студенты должны получить представление о задачах, решаемых с помощью рассматриваемой теории, и принципах построения некоторых основных классификаторов.

2. Компетенции обучающегося, формируемые в результате освоения дисциплины

Планируемые результаты освоения ОП ВО (компетенции)	Планируемые результаты обучения по дисциплине
ПК-4 Способен активно участвовать в исследовании новых математических моделей в естественных науках	Знает: классы методов и алгоритмов машинного обучения Умеет: формализовать прикладные задачи и разрабатывать новые методы и алгоритмы машинного обучения для их решения Имеет практический опыт: программной реализации алгоритмов машинного обучения

3. Место дисциплины в структуре ОП ВО

Перечень предшествующих дисциплин, видов работ учебного плана	Перечень последующих дисциплин, видов работ
Дополнительные главы математического анализа	Проектно-исследовательский семинар, Методики планирования эксперимента

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

Дисциплина	Требования
Дополнительные главы математического анализа	Знает: общенаучную и практическую значимость математического анализа, его роль в математическом моделировании реальных процессов Умеет: Имеет практический опыт: исследования и анализа свойств математических

	и физических объектов средствами математического анализа

4. Объём и виды учебной работы

Общая трудоемкость дисциплины составляет 4 з.е., 144 ч., 74,5 ч. контактной работы

Вид учебной работы	Всего часов	Распределение по семестрам в часах
		Номер семестра
		5
Общая трудоёмкость дисциплины	144	144
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)	69,5	69,5
Подготовка отчётов по лабораторным работам	20	20
Подготовка к коллективной проектной деятельности	10	10
Подготовка к экзамену	19,5	19.5
Изучение теоретического материала	20	20
Консультации и промежуточная аттестация	10,5	10,5
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	Введение в машинное обучение	8	4	0	4
2	Основные методы машинного обучения	8	4	0	4
3	Регрессия. Кластеризация. Классификация. Графовые методы	16	8	0	8
4	Ансамблирование моделей. Бустинг. Переобучение.	16	8	0	8
5	Нейронные сети	16	8	0	8

5.1. Лекции

№ лекции	№ раздела	Наименование или краткое содержание лекционного занятия	Кол-во часов
1 - 2	1	Введение в машинное обучение. Основные понятия. Определение предмета машинного обучения. Примеры задач и областей приложения. Образы и признаки. Типы задач предсказания. Регрессия. Таксономия. Классификация. Типы ошибок классификации. Обобщающая способность классификатора. Принцип минимизации эмпирического риска. Недообучение. Переобучение. Статистический, нейросетевой и структурно - лигвистический подходы к распознаванию образов. Структура типичной системы распознавания	4

		образов. Цикл построения системы распознавания образов.	
3 - 4	2	Классификация. Общие принципы. Этапы классификации. Алгоритмы обучения классификаторов с учителем и без учителя. Дискриминантный анализ. Геометрическая интерпретация задачи классификации. Проективный подход. Метрики в пространстве признаков. Евклидово расстояние. Расстояние Махалонобиса. Ошибки первого и второго рода. Чувствительность и избирательность. Кривая мощности критерия классификации. ROC-кривые. Проверка классификатора. Проверка тестовой выборкой. Перекрестна проверка. Оценка информативности признаков.	4
5 - 6	3	Основные методы машинного обучения. Байесовская классификация. Условная вероятность. Формула полной вероятности. Формула Байеса. Статистическое распознавание образов. Наивный байесовский классификатор. Задача классификации спама. Критерий отношения правдоподобия. Байесовский уровень ошибки. Байесовский риск. Критерий Байеса. Максимальный апостериорный критерий. Критерий максимального правдоподобия. Многоклассовые байесовские классификаторы. Байесовские классификаторы для нормально распределенных классов при различной структуре матрицы ковариации.	4
7 - 8	3	Оценивание функций распределения. Параметрическое оценивание. Метод максимума правдоподобия. Байесовское оценивание. Непараметрическое оценивание. Оценивание ядерным сглаживанием. Окна Парзена. гладкие ядра. Оценка многомерной плотности. Оценивание по K ближайшим соседям. Классификация по K ближайшим соседям. Взвешивание признаков. Повышение скорости поиска ближайших соседей. Метод k-D-дерева. Деревья решений. Основные понятия. Классы решаемых задач: описание данных, классификация, регрессия. Общий алгоритм построения дерева решений. Критерии выбора наилучшего атрибута: прирост информации, относительный прирост информации, индекс Гини. Правила остановки разбиения дерева. Обрезание дерева. Алгоритм ID3. Переобучение деревьев решений. Обработка непрерывных атрибутов. Обучение на данных с пропусками. Программное обеспечение для построения деревьев решений.	4
9 - 10	4	Регрессия. Метод наименьших квадратов. Теорема Гаусса -Маркова. Обобщенный метод наименьших квадратов. Рекурсивный метод наименьших квадратов. Анализ регрессионных остатков. Графическая проверка линейности, гомоскедастичности. Объясненная и необъясненная вариация. Коэффициент детерминации. Неустойчивость МНК к выбросам. Робастная регрессия. Теоретико-множественный подход к регрессии. Ошибки регрессии: нормальность и ограниченность. Теоретико-множественное оценивание параметров регрессии и прогноза при интервальной ошибке. Статус образцов. ПИО-размах и ПИО-отклонение. Выбросы, инсайдеры, аутсайдеры и абсолютные выбросы. Эмпирическое оценивание интервальной ошибки. Планирование эксперимента при построении регрессии с интервальной ошибкой Многомерная регрессия. Особенности построения регрессии по многомерным данным. Множественная линейная регрессия, ее преимущества и недостатки. Мультиколлинеарность данных. Регрессия на главные компоненты. Интерпретация РГК-моделей. Проверка регрессионных моделей. Ошибка обучения и ошибка прогноза. Критерий выбора количества главных компонент при построении РГК.	4
11 - 12	4	Анализ многомерных данных. Корреляционные и причинно-следственные связи. Корреляция признаков и структура данных. Латентные структуры в данных. Формальная и эффективная размерность данных. Структура и ум в данных. Понижение размерности данных. Поиск латентных структур. Отделение структуры от шума. Метод главных компонент как декомпозиция матрицы данных. Матрица счетов. Матрица нагрузок. Матрица ошибок. Объясненная и остаточная вариация в данных. Предобработка данных.	4

		Графическая интерпретация метода главных компонент. Критерии выбора количества главных компонент. Transfer learning, адаптация моделей под задачу.	
13 - 14	5	Кластеризация. Кластеризация как классификация без учителя. Меры сходства и меры различия образов. Критерии качества кластеризации. Итеративна оптимизация разбиения на кластеры. Плоские методы кластеризации. Метод K средних. Метод ISODATA. Метод FOREL. Графовые методы. Иерархическая кластеризация. Агломеративные и разделяющие алгоритмы кластеризации. Дендрограммы. Комитетные методы распознавания образов. Теоретические предпосылки комитетных методов. Одиночные модели и ансамбли моделей. Последовательные методы комитетов: бустинг, AdaBoost. Ошибки классификации комитетными методами. Бустинг и переобучение. Параллельные методы комитетов: бутстреп, бэггинг.	4
15 - 16	5	Нейронные сети. Предпосылки возникновения нейросетей. Перцептрон Розенблатта. Многослойный перцептрон. Карты Кохонена. Сети Хопфилда. Методы обучения нейросетей. Метод опорных векторов. Машинное обучение и теория. Вапника-Червоненкиса. Принцип структурной минимизации риска. Метод опорных векторов. Политика назначения штрафов. Ядерные преобразования. Регрессия опорных векторов.	4

5.2. Практические занятия, семинары

Не предусмотрены

5.3. Лабораторные работы

№ занятия	№ раздела	Наименование или краткое содержание лабораторной работы	Кол-во часов
1	1	Базовые инструменты машинного обучения.	2
2	1	Язык программирования Python. Работа с библиотеками Numpy, Pandas, Matplotlib. Работа с Google Colaboratory.	2
3 - 4	2	Предсказание рака груди на данных Breast Cancer Prediction.	4
5 - 6	3	Решение задачи о типе лесного покрытия из репозитория UCI методом ближайших соседей.	4
7 - 8	3	Решение задачи о типе лесного покрытия из репозитория UCI методом ближайших соседей.	4
9 - 10	4	Построение рекомендательной системы	4
11 - 12	4	Построение рекомендательной системы	4
13 - 14	5	Создание сверточной нейронной сети на фреймоврке глубокого обучения PyTorch для решения задачи распознавания 10 классов объектов на датасете CIFAR10.	4
15 - 16	5	Создание сверточной нейронной сети на фреймоврке глубокого обучения PyTorch для решения задачи распознавания 10 классов объектов на датасете CIFAR10.	4

5.4. Самостоятельная работа студента

Выполнение СРС			
Подвид СРС	Список литературы (с указанием разделов, глав, страниц) / ссылка на ресурс	Семестр	Кол-во часов

Подготовка отчётов по лабораторным работам	ГОСТ Оформления отчёта; ЭУМД, осн. лит. п. 2: https://e.lanbook.com/book/82818.	5	20
Подготовка к коллективной проектной деятельности	ЭУМД, осн. лит. п. 1: https://e.lanbook.com/book/131686; п. 2: https://e.lanbook.com/book/82818; доп. лит. п. 3: https://e.lanbook.com/book100905.	5	10
Подготовка к экзамену	ЭУМД, осн. лит. п. 1: https://e.lanbook.com/book/131686; п. 2: https://e.lanbook.com/book/82818; доп. лит. п. 3: https://e.lanbook.com/book100905.	5	19,5
Изучение теоретического материала	ЭУМД, осн. лит. п. 1: https://e.lanbook.com/book/131686; п. 2: https://e.lanbook.com/book/82818; доп. лит. п. 3: https://e.lanbook.com/book100905.	5	20

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)

№ КМ	Се-местр	Вид контроля	Название контрольного мероприятия	Вес	Макс. балл	Порядок начисления баллов	Учи-тыва-ется в ПА
1	5	Текущий контроль	КМ-1 Лабораторная работа 1	5	4	Верно выполнен пункт 1 лабораторной работы -1 балл. Верно выполнен пункт 2 лабораторной работы -1 балл. Правильно оформлен отчёт по лабораторной работе -1 балл. Защитил отчёт -1 балл. В противном случае -0 баллов.	экзамен
2	5	Текущий контроль	КМ-2 Лабораторная работа 2	10	4	Верно выполнен пункт 1 лабораторной работы -1 балл. Верно выполнен пункт 2 лабораторной работы -1 балл. Правильно оформлен отчёт по лабораторной работе -1 балл. Защитил отчёт -1 балл. В противном случае -0 баллов.	экзамен
3	5	Текущий контроль	КМ-3 Лабораторная работа 3	15	4	Верно выполнен пункт 1 лабораторной работы -1 балл. Верно выполнен пункт 2 лабораторной работы -1 балл. Правильно оформлен отчёт по лабораторной работе -1 балл. Защитил отчёт -1 балл. В противном случае -0 баллов.	экзамен
4	5	Текущий контроль	КМ-4 Лабораторная	15	4	Верно выполнен пункт 1 лабораторной работы -1 балл.	экзамен

			работа 4			Верно выполнен пункт 2 лабораторной работы -1 балл. Правильно оформлен отчёт по лабораторной работе -1 балл. Защитил отчёт -1 балл. В противном случае -0 баллов.	
5	5	Текущий контроль	КМ-5 Лабораторная работа 5	15	4	Верно выполнен пункт 1 лабораторной работы -1 балл. Верно выполнен пункт 2 лабораторной работы -1 балл. Правильно оформлен отчёт по лабораторной работе -1 балл. Защитил отчёт -1 балл. В противном случае -0 баллов.	экзамен
6	5	Текущий контроль	КМ-6 Коллективная проектная деятельность	40	3	Студенты разделяются на команды по 2-4 человека. Каждая команда получает задание. Задание содержит как теоретические, так и практические вопросы. Вопросы связаны с алгоритмами и методами машинного обучения, системами искусственного интеллекта, задачами, решаемыми при создании, поддержке и использовании систем искусственного интеллекта. Каждый член команды получает 3 балла при полном и структурированном ответе. Команда логично и доказательно раскрывает проблему, предложенную в задании, ответ не содержит фактических ошибок и характеризуется глубиной, полнотой; 2 балла -ответ имеет достаточный содержательный уровень, однако отличается слабой структурированностью, имеются неточности при ответе; 1 балл -ответ имеет фрагментарный характер, отличается поверхностностью и малой содержательностью, имеются неточности при ответе, материал в основном излагается, но допущены фактические ошибки; 0 баллов -допускаются существенные фактические ошибки при ответе, на большую часть дополнительных вопросов по содержанию экзамена студент затрудняется дать ответ или не дает верных ответов. Команда может получить максимальный балл (3 балла) при представлении своего проекта, связанное с содержанием дисциплины.	экзамен
7	5	Проме-жуточная аттестация	КМ-7 Экзамен	-	6	Вопрос 1. 2 балла -студент дал полный верный ответ на вопрос; 1 балл -студент дал не полный ответ на	экзамен

						вопрос; 0 баллов -студент не ответил на вопрос или ответ был не верный. Вопрос 2. 2 балла -студент дал полный верный ответ на вопрос; 1 балл -студент дал не полный ответ на вопрос; 0 баллов -студент не ответил на вопрос или ответ был не верный. Вопрос 3. 2 балла -студент дал полный верный ответ на вопрос; 1 балл -студент дал не полный ответ на вопрос; 0 баллов -студент не ответил на вопрос или ответ был не верный.	

6.2. Процедура проведения, критерии оценивания

Вид промежуточной аттестации	Процедура проведения	Критерии оценивания
экзамен	На экзамене происходит оценивание учебной деятельности обучающихся по дисциплине на основе полученных оценок за контрольно-рейтинговые мероприятия текущего контроля. Студент может улучшить свой рейтинг, пройдя контрольное мероприятие промежуточной аттестации, которое не является обязательным. Контрольное мероприятие промежуточной аттестации проводится во время экзамена в виде устного опроса. Студенту выдаётся экзаменационный билет, содержащий 3 вопроса из разных тем курса. Студенту дается 60 минут на подготовку ответов. Затем студент озвучивает свои ответы.	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
ПК-4	Знает: классы методов и алгоритмов машинного обучения	+	+	+	+	+	+	+
ПК-4	Умеет: формализовать прикладные задачи и разрабатывать новые методы и алгоритмы машинного обучения для их решения			+	+	+	+	+
ПК-4	Имеет практический опыт: программной реализации алгоритмов машинного обучения	+	+	+	+	+	+	+

Типовые контрольные задания по каждому мероприятию находятся в приложениях.

7. Учебно-методическое и информационное обеспечение дисциплины

Печатная учебно-методическая документация

а) основная литература:

Не предусмотрена

б) дополнительная литература:

Не предусмотрена

в) отечественные и зарубежные журналы по дисциплине, имеющиеся в библиотеке:

Не предусмотрены

г) методические указания для студентов по освоению дисциплины:

1. ГОСТ Оформления отчёта
2. Методические указания по организации самостоятельной работы студента

из них: учебно-методическое обеспечение самостоятельной работы студента:

1. ГОСТ Оформления отчёта
2. Методические указания по организации самостоятельной работы студента

Электронная учебно-методическая документация

№	Вид литературы	Наименование ресурса в электронной форме	Библиографическое описание
1	Основная литература	Электронно-библиотечная система издательства Лань	Шалев-Шварц, Ш. Идеи машинного обучения : учебное пособие / Ш. Шалев-Шварц, Бен-ДавидШ. ; перевод с английского А. А. Слинкина. — Москва : ДМК Пресс, 2019. — 436 с. — ISBN 978-5-97060-673-5. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/131686 (дата обращения: 06.09.2022). — Режим доступа: для авториз. пользователей.
2	Основная литература	Электронно-библиотечная система издательства Лань	Коэльо, Л. П. Построение систем машинного обучения на языке Python / Л. П. Коэльо, В. Ричарт ; перевод с английского А. А. Слинкин. —2-е изд. —Москва : ДМК Пресс, 2016. —302 с. —ISBN 978-5-97060-330-7. —Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/82818
3	Дополнительная литература	Электронно-библиотечная система издательства Лань	Рашка, С. Python и машинное обучение: крайне необходимое пособие по новейшей предсказательной аналитике, обязательное для более глубокого понимания методологии машинного обучения : руководство / С. Рашка ; перевод с английского А. В. Логунова. —Москва : ДМК Пресс, 2017. —418 с. —ISBN 978-5-97060-409-0. —Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book100905

Перечень используемого программного обеспечения:

1. Microsoft-Office(бессрочно)
2. -Python(бессрочно)
3. Microsoft-Microsoft Imagine Premium (Windows Client, Windows Server, Visual Studio Professional, Visual Studio Premium, Windows Embedded, Visio, Project, OneNote, SQL Server, BizTalk Server, SharePoint Server)(04.08.2019)

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
Лабораторные занятия	333 (3б)	ПК, интернет.
Лекции	333 (3б)	ПК, проектор, экран.
Самостоятельная работа студента	332 (3б)	ПК, интернет.