

Министерство науки и высшего образования Российской Федерации
федеральное государственное бюджетное образовательное учреждение
высшего образования
«Санкт-Петербургский государственный университет промышленных технологий и дизайна»
(СПбГУПТД)

УТВЕРЖДАЮ
Первый проректор, проректор
по УР

_____ А.Е. Рудин

Рабочая программа дисциплины

Б1.В.ДВ.01.02 Машинное обучение

Учебный план: 2026-2027 09.03.03 ИИТА ПИИ ОО №1-1-181.plx

Кафедра: **33** Цифровых и аддитивных технологий

Направление подготовки:
(специальность) 09.03.03 Прикладная информатика

Профиль подготовки: Прикладной искусственный интеллект
(специализация)

Уровень образования: бакалавриат

Форма обучения: очная

План учебного процесса

Семестр (курс для ЗАО)		Контактная работа обучающихся		Сам. работа	Контроль, час.	Трудоёмкость, ЗЕТ	Форма промежуточной аттестации
		Лекции	Практ. занятия				
5	УП	16	32	93,75	2,25	4	Курсовая работа, Зачет
	РПД	16	32	93,75	2,25	4	
Итого	УП	16	32	93,75	2,25	4	
	РПД	16	32	93,75	2,25	4	

Рабочая программа дисциплины составлена в соответствии с федеральным государственным образовательным стандартом высшего образования по направлению подготовки 09.03.03 Прикладная информатика, утвержденным приказом Минобрнауки России от 19.09.2017 г. № 922

Составитель (и):		
кандидат технических наук, доцент	_____	Суханов Михаил Борисович
ассистент	_____	Муккель Артем Алексеевич
От кафедры составителя:		
Заведующий кафедрой цифровых и аддитивных технологий	_____	Сошников Антон Владимирович
От выпускающей кафедры:		
Заведующий кафедрой	_____	Сошников Антон Владимирович

Методический отдел:

1 ВВЕДЕНИЕ К РАБОЧЕЙ ПРОГРАММЕ ДИСЦИПЛИНЫ

1.1 Цель дисциплины: Сформировать компетенции обучающегося в области машинного обучения, включая понимание основных алгоритмов, методов анализа данных и их практического применения в различных предметных областях.

1.2 Задачи дисциплины:

Изучить фундаментальные принципы и математические основы машинного обучения;

Освоить классические алгоритмы обучения с учителем и без учителя;

Получить навыки предварительной обработки и анализа данных;

Изучить методы оценки качества моделей и валидации;

Выработать практические навыки применения алгоритмов машинного обучения для решения реальных задач;

Освоить современные инструментальные средства и библиотеки для машинного обучения.

1.3 Требования к предварительной подготовке обучающегося:

Предварительная подготовка предполагает создание основы для формирования компетенций, указанных в п. 2, при изучении дисциплин:

Математика

Прикладная статистика

Алгоритмизация и программирование

Методы прикладной математики

Исследование операций и методы оптимизации

2 КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ ДИСЦИПЛИНЫ

ПК-3: Способен проектировать компьютерное программное обеспечение					
Знать: Методологические основы применения алгоритмов машинного обучения.					
Уметь: Визуализировать результаты работы алгоритмов машинного обучения, выбирать метод машинного обучения, соответствующий исследовательской задаче, интерпретировать полученные результаты.					
Владеть: Навыками применения методов машинного обучения при разработке программного обеспечения в сфере искусственного интеллекта.					
ПК-4: Способен применять классические алгоритмы машинного обучения					
Знать: Методы машинного обучения; принципы и алгоритмы обучения искусственного интеллекта.					
Уметь: Формировать алгоритм обучения моделей искусственного интеллекта.					
Владеть: Навыками обучения моделей искусственного интеллекта на основе подготовленных данных.					

3 РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ

Наименование и содержание разделов, тем и учебных занятий	Семестр (курс для ЗАО)	Контактная работа		СР (часы)	Инновац. формы занятий	Форма текущего контроля
		Лек. (часы)	Пр. (часы)			
Раздел 1. Введение в машинное обучение	5					Пр
Тема 1. Основы машинного обучения. Практические занятия: Настройка среды Python.		2	4	14		
Раздел 2. Обучение с учителем						Пр
Тема 2. Линейная и логистическая регрессия. Практические занятия: решение индивидуальных заданий.		3	6	14		
Тема 3. Алгоритмы классификации: kNN, SVM, деревья. Практические занятия: решение индивидуальных заданий.		3	6	14		
Тема 4. Ансамбли: случайный лес, бустинг. Практические занятия: решение индивидуальных заданий.		3	6	14		
Раздел 3. Обучение без учителя и оценка						Пр
Тема 5. Кластеризация и снижение размерности. Практические занятия: решение индивидуальных заданий.		3	4	20		

Тема 6. Метрики качества и валидация. Практические занятия: решение индивидуальных заданий.		2	6	17,75		
Итого в семестре (на курсе для ЗАО)		16	32	93,75		
Консультации и промежуточная аттестация (Курсовая работа, Зачет)		2,25				
Всего контактная работа и СР по дисциплине		50,25		93,75		

4 КУРСОВОЕ ПРОЕКТИРОВАНИЕ

4.1 Цели и задачи курсовой работы (проекта): Продемонстрировать владение основными навыками работы с методами машинного обучения с учителем и без учителя, владение основными инструментальными средствами библиотек языка программирования Python, методами и приемами подготовительного и описательного анализа данных, средствами визуализации данных, использования и усовершенствования обучаемых моделей, умение делать выводы из проведенного анализа.

4.2 Тематика курсовой работы (проекта): 1. Дескриптивный анализ численного набора данных с использованием технологий визуализации.

2. Дескриптивный анализ категориального набора данных с использованием технологий визуализации.

3. Дескриптивный анализ смешанного набора данных с использованием технологий визуализации.

4. Сравнение методов регрессии на реальных наборах данных.

5. Сравнение методов классификации на реальных наборах данных.

6. Предварительный анализ данных и построение признаков в задачах обработки финансовой и экономической информации.

7. Предварительный анализ данных и построение признаков в задачах распознавания голоса.

8. Предварительный анализ данных и построение признаков в задачах распознавания текста.

4.3 Требования к выполнению и представлению результатов курсовой работы (проекта):

Работа выполняется в рамках представленного технического задания, с использованием современных технологий проектирования и визуализации.

Результаты представляются в виде:

1. Курсовой работы, объемом не менее 20 страниц, содержащей следующие обязательные элементы:

Титульный лист

Задание

Реферат

Содержание

Введение

Основная часть

В данном разделе пояснительной записки должны быть отражены результаты проведенной работы:

1. Выбрать набор данных для анализа в соответствии с выбранной темой курсовой работы. Описать этот набор и решаемую задачу.

2. Провести предварительный анализ и очистку данных. Этот этап включает в себя вывод информации о количественных характеристиках датасета, информацию об отсутствующих значениях, характеристиках и физическом смысле каждого атрибута данных, его значимости для предсказания целевой переменной, вывод нескольких точек данных для иллюстрации структуры данных.

3. При необходимости, преобразовать атрибуты исходного датасета в числовые признаки. Этот этап сильно зависит от типа исследуемых данных и может включать в себя векторизацию текста, извлечение признаков из аудио и видео данных, преобразование изображений в плоский численный массив и другие преобразования.

4. Провести описательный анализ данных. Сделать выводы. Этот этап включает в себя определение шкалы измерения каждого признака, выявление аномальных значений, визуализацию распределения каждого признака, при необходимости - проверка на нормальность,

построение кореллограмм и совместных распределений каждого признака с целевой переменной, выявление коррелированных признаков и признаков, не несущих информации для данной задачи.

5. Применить при необходимости к данным методы обучения без учителя: кластеризацию, понижение размерности и поиск аномалий. Сделать выводы.

6. Разделить набор данных на обучающую и тестовую выборки. Обосновать количественные характеристики и метод разделения (временной, случайный, последовательный).

7. Обучить несколько моделей для решения выбранной задачи (для задач классификации - не менее 7 различных алгоритмов). Проанализировать результаты, сделать выводы.

8. Выбрать наиболее перспективную модель для решения поставленной задачи. Изменить гиперпараметры модели. Предпочтительно, провести Grid Search. Найти оптимальные гиперпараметры.

9. С учетом сделанных выводов провести усовершенствование моделей. Это можно осуществить с помощью введения регуляризации, изменение параметров модели (для параметрических моделей), введением суррогатных признаков, отбором признаков, нормализацией данных, ансамблированием моделей, изменением алгоритма предварительной обработки данных. Сравнить результаты.

10. Попробовать изменить порядок предобработки данных для повышения эффективности модели. Попробовать применить понижение размерности для создания суррогатных признаков. Сравнить результаты, сделать выводы.

11. Представить результаты моделирования в наглядном виде (графики, линии обучения, таблицы сравнения моделей, таблицы классификации, и другие). Сделать выводы, сравнить с существующими аналогичными решениями, порассуждать о перспективах решения проблемы.

Заключение

Список использованной литературы

Приложение

5. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

5.1 Описание показателей, критериев и системы оценивания результатов обучения

5.1.1 Показатели оценивания

Код компетенции	Показатели оценивания результатов обучения	Наименование оценочного средства
ПК-3	Формулирует суть применения алгоритмов машинного обучения для решения профессиональных задач. Формировать аналитическое заключение по результатам использования алгоритмов машинного обучения для решения профессиональных задач. Демонстрирует результаты применения методов машинного обучения при разработке программного обеспечения в сфере искусственного интеллекта	
ПК-4	Формулирует основные принципы и алгоритмы обучения искусственного интеллекта на основе методов машинного обучения. Строит алгоритм обучения моделей искусственного интеллекта на основе использования методов машинного обучения.	
	Демонстрирует результаты обучения моделей искусственного интеллекта на основе подготовленных данных.	

5.1.2 Система и критерии оценивания

Шкала оценивания	Критерии оценивания сформированности компетенций	
	Устное собеседование	Письменная работа
5 (отлично)	Полный, исчерпывающий ответ, явно демонстрирующий глубокое понимание предмета и широкую эрудицию в оцениваемой области. Критический, оригинальный подход к материалу. Учитываются баллы, накопленные в течение семестра.	Качество исполнения всех элементов задания полностью соответствует всем требованиям. Учитываются баллы, накопленные в течение семестра.
4 (хорошо)	Ответ полный, основанный на проработке всех обязательных источников информации. Подход к материалу ответственный, но стандартный. Учитываются баллы, накопленные в течение семестра.	Задание выполнено в достаточном объеме, но ограничивается только основными подходами. Учитываются баллы, накопленные в течение семестра.
3 (удовлетворительно)	Ответ стандартный, в целом качественный, основан на всех обязательных источниках информации. Присутствуют небольшие пробелы в знаниях или несущественные ошибки. Учитываются баллы, накопленные в течение семестра.	Задание выполнено в соответствии с заданием. Имеются отдельные несущественные ошибки или отступления от правил оформления работы. Учитываются баллы, накопленные в течение семестра.
2 (неудовлетворительно)	Неспособность ответить на вопрос без помощи экзаменатора. Незнание значительной части принципиально важных элементов дисциплины. Многочисленные грубые ошибки. Не учитываются баллы, накопленные в течение семестра.	Отсутствие одного или нескольких обязательных элементов задания, либо многочисленные грубые ошибки в работе, либо грубое нарушение правил оформления или сроков представления работы. Не учитываются баллы, накопленные в течение семестра.
Зачтено	Ответ на теоретический вопрос по материалам лекций полный, с возможными несущественными ошибками. Учитываются баллы, накопленные в течение семестра. Качество исполнения всех элементов практико-ориентированного задания полностью соответствует всем требованиям. Учитываются баллы, накопленные в течение семестра.	Обучающийся своевременно выполнил практические задания. Учитываются баллы, накопленные в течение семестра.
Не зачтено	Ответ на теоретический вопрос не полный, с существенными ошибками. Не учитываются баллы, накопленные в течение семестра. Своевременно не выполняет (выполнил частично) практические задания. Не учитываются баллы, накопленные в течение семестра.	Своевременно не выполняет (выполнил частично) практические задания. Не учитываются баллы, накопленные в течение семестра.

5.2 Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности

5.2.1 Перечень контрольных вопросов

№ п/п	Формулировки вопросов
Семестр 5	
1	Основные типы задач машинного обучения.
2	Обучение с учителем, без учителя и с подкреплением.
3	Переобучение и недообучение.
4	Валидация и кросс-валидация.
5	Линейная регрессия: математические основы.
6	Логистическая регрессия и функция сигмоиды.
7	Метрики качества классификации.

8	Метрики качества регрессии.
9	Алгоритм k-ближайших соседей.
10	Наивный байесовский классификатор.
11	Метод опорных векторов.
12	Решающие деревья.
13	Решающие деревья.
14	Случайный лес.
15	Градиентный бустинг.
16	Алгоритм k-means.
17	Иерархическая кластеризация.
18	Метод главных компонент (PCA).
19	Предварительная обработка данных.
20	Работа с категориальными признаками.
21	Проклятие размерности.

5.2.2 Типовые тестовые задания

Не предусмотрено.

5.2.3 Типовые практико-ориентированные задания (задачи, кейсы)

Учитываются индивидуальные задания, выполненные в течение семестра.

5.3 Методические материалы, определяющие процедуры оценивания знаний, умений, владений (навыков и (или) практического опыта деятельности)

5.3.1 Условия допуска обучающегося к промежуточной аттестации и порядок ликвидации академической задолженности

Проведение промежуточной аттестации регламентировано локальным нормативным актом СПбГУПТД «Положение о проведении текущего контроля успеваемости и промежуточной аттестации обучающихся»

5.3.2 Форма проведения промежуточной аттестации по дисциплине

Устная ☐ + Письменная ☐ + Компьютерное тестирование ☐ Иная ☐

5.3.3 Особенности проведения промежуточной аттестации по дисциплине

зачет:

- время на подготовку к устному собеседованию составляет 15 минут.

защита курсовой работы:

- доклад с использованием презентации по основным элементам курсовой работы – 10 мин;

- ответ на вопросы по материалам курсовой работы – 10 мин.

6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

6.1 Учебная литература

Автор	Заглавие	Издательство	Год издания	Ссылка
6.1.1 Основная учебная литература				
Плас Дж. Вандер	Python для сложных задач: наука о данных и машинное обучение	Санкт-Петербург: Питер	2018	https://ibooks.ru/read/ing.php?short=1&productid=356721
Элбон Крис	Машинное обучение с использованием Python. Сборник рецептов: Пер. с англ.	Санкт-Петербург: БХВ-Петербург	2019	https://ibooks.ru/read/ing.php?short=1&productid=366635
Сузи, Р. А.	Язык программирования Python	Москва: Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа	2020	http://www.iprbookshop.ru/97589.html
6.1.2 Дополнительная учебная литература				
Зурахов В. С.	Интеллектуальный анализ данных и машинное обучение	СПб.: СПбГУПТД	2014	http://publish.sutd.ru/tp_ext_inf_publish.php?id=2025
Уэс Маккинли, Слинкин А. А.	Python и анализ данных	Саратов: Профобразование	2017	http://www.iprbookshop.ru/64058.html

6.2 Перечень профессиональных баз данных и информационно-справочных систем

Информационная система «Единое окно доступа к образовательным ресурсам. Раздел. Информатика и информационные технологии» [Электронный ресурс]. URL: http://window.edu.ru/catalog/?p_rubr=2.2.75.6
Электронно-библиотечная система IPRbooks [Электронный ресурс]. URL: <http://www.iprbookshop.ru>

6.3 Перечень лицензионного и свободно распространяемого программного обеспечения

MicrosoftOfficeProfessional
Microsoft Windows
Figma
Blender
Adobe Audition

6.4 Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине

Аудитория	Оснащение
Лекционная аудитория	Мультимедийное оборудование, специализированная мебель, доска
Компьютерный класс	Мультимедийное оборудование, компьютерная техника с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду