

Машинное обучение и анализ данных рабочая программа дисциплины (модуля)

Закреплена за кафедрой **Промышленной кибербезопасности и защиты геоданных**

Учебный план b090302_25_BIS25.plx
09.03.02 Информационные системы и технологии

Квалификация **Бакалавр**

Форма обучения **очная**

Общая трудоемкость **3 ЗЕТ**

Часов по учебному плану	0	Виды контроля в семестрах:
в том числе:		
аудиторные занятия	0	
самостоятельная работа	0	

Распределение часов дисциплины по семестрам

Семестр (<Курс>.<Семестр на курсе>)	4 (2.2)		Итого	
Неделя	14			
Вид занятий	УП	РП	УП	РП
Лекции	28	28	28	28
Практические	28	28	28	28
Иные виды контактной работы	0,25	0,25	0,25	0,25
Итого ауд.	56,25	56,25	56,25	56,25
Контактная работа	56,25	56,25	56,25	56,25
Сам. работа	51,75	51,75	51,75	51,75
Итого	108	108	108	108

1. ЦЕЛИ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)	
1.1	Цель освоения дисциплины заключается в ознакомлении с базовыми понятиями машинного обучения, с основными алгоритмами машинного обучения, особенностями их применения.

2. МЕСТО ДИСЦИПЛИНЫ (МОДУЛЯ) В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ	
Цикл (раздел) ОП:	Б1.О
2.1	Требования к предварительной подготовке обучающегося:
2.1.1	Системы искусственного интеллекта
2.2	Дисциплины (модули) и практики, для которых освоение данной дисциплины (модуля) необходимо как предшествующее:

3. КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)	
ОПК-6: Способен разрабатывать алгоритмы и программы, пригодные для практического применения в области информационных систем и технологий	
Знать:	
Уметь:	
Владеть:	

В результате освоения дисциплины (модуля) обучающийся должен

3.1	Знать:
3.1.1	основные возможности современных интегрированных сред разработки программного обеспечения на объектно-ориентированных языках программирования;
3.1.2	возможности
3.1.3	компиляторов программных проектов под различные операционные системы;
3.1.4	наборы инструкций для системных утилит автоматической сборки программного обеспечения и установки программных пакетов объектно-ориентированных библиотек и фреймворков, методы разработки алгоритмов и программ в рамках объектно-ориентированной парадигмы программирования на современном языке высокого уровня; принципы объектно-ориентированной парадигмы: абстрагирование, инкапсуляция, наследование, полиморфизм;
3.1.5	основные синтаксические конструкции объектно-ориентированного языка программирования: классы, поля, свойства, методы, выражения, события; методы обобщенного программирования;
3.1.6	методы оценки сложности алгоритмов;
3.1.7	функциональные возможности стандартной библиотеки языка и фреймворка;
3.1.8	язык программирования высокого уровня (основы объектно-ориентированного программирования);
3.1.9	стандартные алгоритмы и методы организации и обработки данных, общие
3.1.10	принципы построения, области и особенности применения языков программирования высокого уровня;
3.1.11	ключевые понятия и особенности разработки программных средств для решения задач профессиональной деятельности с использованием объектно-ориентированного подхода, знает методы и средства самостоятельного решения задач в сфере профессиональной деятельности;
3.2	Уметь:
3.2.1	использовать функциональные возможности современных интегрированных сред разработки программного обеспечения на объектно-ориентированных языках программирования для разработки прикладных программ; использовать утилиты автоматической сборки и развертывания программ в операционных системах, разрабатывать алгоритмы и программы в рамках
3.2.2	объектно-ориентированной парадигмы на современном языке программирования высокого уровня с применением основных синтаксических конструкций и функциональных возможностей стандартной библиотеки языка и фреймворка;
3.2.3	разрабатывать и реализовывать на языке высокого уровня алгоритмы решения типовых профессиональных задач, работать с интегрированной средой разработки программного обеспечения;
3.2.4	использовать технологии разработки программных средств для решения задач профессиональной деятельности, умеет планировать самостоятельную деятельность при решении профессиональных задач;
3.3	Владеть:

3.3.1	работы с основными современными интегрированными средами разработки программного обеспечения на объектно-ориентированных языках; разработки, отладки и развёртывания программного обеспечения в операционных системах семейства Windows и Linux; поиска и анализа возможностей современных интегрированных программных средств разработки прикладного программного обеспечения, разработки алгоритмов и программ; отладки, поиска и устранения ошибок программного кода; оценки сложности алгоритмов; использования
3.3.2	возможностей стандартной библиотеки, сторонних библиотек программного кода и фреймворков;
3.3.3	программирования элементов информационных систем, требующие объектно-ориентированного подхода;

4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Код занятия	Наименование разделов и тем /вид занятия/	Семестр / Курс	Часов	Компетен-ции	Литература	Инте ракт.	Примечание
	Раздел 1. Введение в машинное обучение						
1.1	Разбор возможных приложений алгоритмов машинного обучения. /Лек/	4	2		Л1.1 Л1.2	0	
1.2	Исследование влияния параметров алгоритма на точность алгоритма на обучающей и тестовой выборке. /Лек/	4	2		Л1.1 Л1.2	0	
1.3	Приложений алгоритмов машинного обучения в различных областях. Контрольная точка №1 /Пр/	4	8		Л1.1 Л1.2	0	
	Раздел 2. Модели регрессии						
2.1	Проведение экспериментов с программной реализацией линейной регрессии. /Лек/	4	4		Л1.1 Л1.2	0	
2.2	Проведение экспериментов с программной реализацией логистической регрессии. /Лек/	4	4		Л1.1 Л1.2	0	
2.3	Исследование влияния параметров алгоритма регрессии на точность на обучающей и тестовой выборке. Контрольная точка №2 /Пр/	4	8		Л1.1 Л1.2	0	
2.4	Индивидуальные задания /Ср/	4	25,75		Л1.1 Л1.2	0	
	Раздел 3. Деревья решений						
3.1	Проведение экспериментов с программной реализацией дерева решений. /Лек/	4	4		Л1.1 Л1.2	0	
3.2	Проведение экспериментов с программной реализацией алгоритма AdaBoost. /Лек/	4	4		Л1.1 Л1.2	0	
3.3	Исследование влияния параметров алгоритма дерева решений на точность на обучающей и тестовой выборке. Контрольная точка №3 /Пр/	4	6		Л1.1 Л1.2	0	
	Раздел 4. Кластеризация						
4.1	Программная реализация каскада с деревьями решений в качестве слабых классификаторов. /Лек/	4	4		Л1.1 Л1.2	0	
4.2	Кластеризация. Алгоритм k-means. /Лек/	4	4		Л1.1 Л1.2	0	
4.3	Исследование влияния параметров слабых классификаторов на точность алгоритма кластеризации на обучающей и тестовой выборке. Контрольная точка №4 /Пр/	4	6		Л1.1 Л1.2	0	
4.4	Подготовка к зачету /Ср/	4	26		Л1.1 Л1.2	0	
4.5	Зачет /ИБКР/	4	0,25		Л1.1 Л1.2	0	

5. ОЦЕНОЧНЫЕ СРЕДСТВА

5.1. Контрольные вопросы и задания

Тема 1: Возможные приложения алгоритмов машинного обучения

1. Какие задачи могут быть решены с помощью методов машинного обучения?
2. В каких областях применяется машинное обучение (медицина, финансы, ИБ и т.д.)?
3. Чем отличаются задачи классификации, регрессии и кластеризации?
4. Что такое обучение с учителем и без учителя? Примеры.
5. Какие этапы включает жизненный цикл проекта машинного обучения?

Тема 2: Влияние параметров алгоритма на точность

6. Как выбор гиперпараметров влияет на качество модели?
7. Что такое переобучение и недообучение? Как их избежать?
8. Как проводится подбор оптимальных параметров модели?
9. Что такое кросс-валидация и зачем она нужна?
10. Как оценить обобщающую способность модели?

Тема 3: Линейная регрессия — реализация и анализ

11. Как устроена модель линейной регрессии?
12. Как выглядит функция стоимости в линейной регрессии?
13. Как работает метод градиентного спуска в обучении модели?
14. Как оценить качество модели на обучающей и тестовой выборках?
15. Как влияет нормализация данных на обучение линейной модели?

Тема 4: Логистическая регрессия — реализация и анализ

16. Чем логистическая регрессия отличается от линейной?
17. Как строится функция активации в логистической регрессии?
18. Как интерпретировать выход модели как вероятность?
19. Как оценивается точность модели (accuracy, precision, recall)?
20. Как происходит обучение логистической регрессии?

Тема 5: Деревья решений — программная реализация

21. Что представляет собой дерево решений?
22. Как происходит процесс разбиения узлов дерева?
23. Какие критерии используются для выбора разбиения (прирост информации, Gini)?
24. Как предотвратить переобучение дерева решений?
25. Как оценить важность признаков с помощью дерева?

Тема 6: Алгоритм AdaBoost — программная реализация

26. Что такое ансамблевые методы и где они применяются?
27. Как работает алгоритм AdaBoost?
28. Как формируется финальная модель в AdaBoost?
29. Как влияет количество слабых классификаторов на точность?
30. Какие метрики используются для оценки качества ансамбля?

Тема 7: Каскады с деревьями решений

31. Что такое бустинг с использованием деревьев решений?
32. Как строится композиция деревьев в каскаде?
33. Как происходит коррекция ошибок в последующих деревьях?
34. Какие параметры наиболее существенно влияют на производительность?
35. Как реализуется и настраивается градиентный бустинг?

Тема 8: Кластеризация. Алгоритм k-means

36. Что такое кластеризация и чем она отличается от классификации?
37. Как работает алгоритм k-means?
38. Как выбрать начальные центроиды и как они влияют на результат?
39. Как определить оптимальное число кластеров?
40. Какие метрики используются для оценки качества кластеризации?

5.2. Темы письменных работ

не предусмотрены

5.3. Оценочные средства

Рабочая программа "Машинное обучение и анализ данных" обеспечена оценочными средствами для проведения текущего контроля и промежуточной аттестации, включающими контрольные вопросы для проведения промежуточной аттестации, критерии оценивания учебной деятельности обучающихся по балльно-рейтинговой системе, примеры заданий для практических и лабораторных занятий, билеты для проведения промежуточной аттестации.

Все оценочные средства представлены в Приложении 1.

5.4. Перечень видов оценочных средств

Оценочные средства разработаны для всех видов учебной деятельности студента - лекций, практических занятий, самостоятельной работы и промежуточной аттестации. Оценочные средства представлены в виде:

- средства текущего контроля: проверочных работ по решению задач, дискуссии по теме;
- средств итогового контроля - промежуточной аттестации: экзамена в 4 семестре.

6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

6.1. Рекомендуемая литература

6.1.1. Основная литература			
	Авторы, составители	Заглавие	Издательство, год
Л1.1	Романов П. С., Романова И. П.	Системы искусственного интеллекта. Моделирование нейронных сетей в системе MATLAB. Лабораторный практикум: учебное пособие для вузов	Санкт-Петербург: Лань, 2024
Л1.2	Остроух А. В., Суркова Н. Е.	Системы искусственного интеллекта: монография	Санкт-Петербург: Лань, 2024
6.3.1 Перечень программного обеспечения			
6.3.1.1	Office Professional Plus 2019		
6.3.1.2	Windows 10		
6.3.1.3	МТС-Линк	Комплексная платформа для коммуникаций, обучения и совместной работы, разработанная с использованием современных технологий. Доступны десктопные и мобильные приложения для удобной работы с системой.	
6.3.2 Перечень информационных справочных систем			
6.3.2.1	Электронно-библиотечная система "Лань" Доступ к коллекциям электронных изданий ЭБС "Издательство "Лань"		
6.3.2.2	Электронно-библиотечная система «Книжный Дом Университета» ("БиблиоТех")		
6.3.2.3	База данных научных электронных журналов "eLibrary"		

7. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)			
Аудитория	Назначение	Оснащение	Вид
3-33	Компьютерный класс, Аудитория для лекционных, практических и семинарных занятий, самостоятельной работы, текущего контроля и промежуточной аттестации	Специализированная мебель: набор учебной мебели на 11 посадочных мест; стул преподавательский – 1 шт.; стол преподавательский – 2 шт.; доска маркерная – 1 шт. 12 моноблоков Enigma Venus, 1 моноблок Acer B223w, 1 моноблок iRu M22, 1 принтер HP LJ1020, 1 телевизор bbk 65LEX-8274/UTS2C, в аудитории развернута проводная сеть и подключен доступ к интернет.	
3	Специализированная многофункциональная учебная аудитория № 3 для проведения учебных занятий семинарского типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной/ итоговой аттестации	Компьютерные столы обучающихся; Стулья обучающихся; Письменный стол педагогического работника; Стул педагогического работника; Стеллаж для учебно-методических материалов, в том числе учебно-наглядных пособий; Многофункциональное устройство (принтер, сканер, ксерокс); Интерактивная доска; Мультимедийный проектор; Ноутбуки с возможностью подключения к сети «Интернет» и обеспечением доступа к электронной информационно-образовательной среде	

5	Помещение № 5 для самостоятельной работы обучающихся	Письменный стол обучающегося; Стул обучающегося; Письменный стол обучающегося с ограниченными возможностями здоровья; Стул обучающегося с ограниченными возможностями здоровья; Ноутбук с возможностью подключения к сети «Интернет» и обеспечением доступа к электронной информационно-образовательной среде лицензиата; Моноблок (в том числе, клавиатура, мышь, наушники) с возможностью подключения к сети «Интернет» и обеспечением доступа к электронной информационно-образовательной среде	
Ауд. 8	Аудитория для научно-исследовательской работы обучающихся, курсового и дипломного проектирования № 8	Рабочие места на базе вычислительной техники с набором необходимых для проведения и оформления результатов исследований дополнительных аппаратных и/или программных средств; Письменный стол обучающегося; Стул обучающегося; Письменный стол обучающегося с ограниченными возможностями здоровья; Стул обучающегося с ограниченными возможностями здоровья; Ноутбук с возможностью подключения к сети «Интернет» и обеспечением доступа к электронной информационно-образовательной среде лицензиата; Моноблок (в том числе, клавиатура, мышь, наушники) с возможностью подключения к сети «Интернет» и обеспечением доступа к электронной информационно-образовательной среде; Многофункциональное устройство (принтер, сканер, ксерокс).	

8. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ (МОДУЛЯ)

Методические указания по изучению дисциплины "Машинное обучение и анализ данных" представлены в Приложении 2 и включают в себя:

1. Методические указания для обучающихся по организации учебной деятельности.
2. Методические указания по организации самостоятельной работы обучающихся.
3. Методические указания по организации процедуры оценивания знания, умений, навыков и (или) опыта деятельности,

характеризующих этапы формирования компетенций.