

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Косенок Сергей Михайлович
Должность: ректор
Дата подписания: 20.06.2025 07:53:12
Уникальный программный ключ:
e3a68f3eaa1e62674b54f4998099d3d6bfdcf836

Оценочные материалы для текущего контроля и промежуточной аттестации
по дисциплине

Анализ данных

Код, направление подготовки	09.03.02 ИНФОРМАЦИОННЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ
Направленность (профиль)	Безопасность информационных систем и технологий
Форма обучения	Очная
Кафедра-разработчик	Информатики и вычислительной техники
Выпускающая кафедра	Информатики и вычислительной техники

Задания на контрольную работу

1. Основы синтаксиса языка Python
2. Загрузка веб-страницы: модуль requests.
3. Пакет BeautifulSoup 4
4. Поиск данных на HTML-страницах и работа с API с помощью XML¶.
5. Классы и поиск по дереву. API и XML
6. Работа с API с помощью JSON. Эмуляция действий с браузером
7. Selenium: дистанционное управление для браузера
8. Библиотеки numpy и matplotlib¶.
9. Библиотека numpy: эффективные массивы
10. Библиотека pandas. Датафреймы. Загрузка датафрейма из CSV-файла
11. Постановка задач обучения по прецедентам.
12. Объекты и признаки. Типы шкал: бинарные, номинальные, порядковые, количественные.
13. Байесовский классификатор
14. Понятие формального нейрона. Виды функций активации
15. Принципы обучения нейросетей. Обратное распространение ошибки. Правило Хебба. Конкурентное обучение.
16. Обучающая, тестовая и валидационная выборки. Выбор признаков.
17. Метрики работы нейросетей. Accuracy, precision, recall. F-мера. ROC и AUC.
18. Архитектура многослойного персептрона
19. Сверточные нейросети
20. Сети глубокого обучения
21. LSTM
22. Автокодировщики
23. Нейронечеткие модели. Нейробайесовские модели
24. Загрузка веб-страницы: модуль requests.
25. Пакет BeautifulSoup 4
26. Поиск данных на HTML-страницах и работа с API с помощью XML¶.
27. Классы и поиск по дереву. API и XML
28. Работа с API с помощью JSON. Эмуляция действий с браузером
29. Selenium: дистанционное управление для браузера
30. Библиотеки numpy и matplotlib¶.
31. Библиотека numpy: эффективные массивы
32. Библиотека pandas. Датафреймы. Загрузка датафрейма из CSV-файла
33. Постановка задач обучения по прецедентам.
34. Объекты и признаки. Типы шкал: бинарные, номинальные, порядковые,
35. количественные.
36. Типы задач: классификация, регрессия, прогнозирование, кластеризация.
37. Примеры прикладных задач.
38. Основные понятия: модель алгоритмов, метод обучения, функция потерь и функционал качества, принцип минимизации эмпирического риска, обобщающая способность, скользящий контроль.
39. Метод ближайших соседей и его обобщения
40. Метод потенциальных функций, градиентный алгоритм.
41. Задача регрессии.
42. Многомерная линейная регрессия.
43. Метод наименьших квадратов, его вероятностный смысл и геометрический смысл.
44. Типы задач: классификация, регрессия, прогнозирование, кластеризация.
45. Примеры прикладных задач.

46. Основные понятия: модель алгоритмов, метод обучения, функция потерь и функционал качества, принцип минимизации эмпирического риска, обобщающая способность, скользящий контроль.
47. Метод ближайших соседей и его обобщения
48. Метод потенциальных функций, градиентный алгоритм.
49. Задача регрессии.
50. Многомерная линейная регрессия.
51. Метод наименьших квадратов, его вероятностный смысл и геометрический смысл.

Этап: проведение промежуточного контроля успеваемости по дисциплине

Вид промежуточной аттестации по дисциплине – **зачет**.

Задание для проведения промежуточной аттестации содержит один теоретический и два практических вопроса. Примерные вопросы и задания:

1. Постановка задачи машинного обучения. Обучение с учителем и без учителя. Типы признаков
2. Типы задач. Линейные модели
3. Примеры задач обучения без учителя (3-5 примеров)
4. Примеры задач обучения с учителем (3-5 примеров)
5. Функционал качества. Минимизация эмпирического риска
6. Переобучение. Методы валидации моделей. Кросс-валидация
7. Метрические методы классификации. Виды расстояний
8. Метод ближайшего соседа. Метод k ближайших соседей. Взвешенный метод k ближайших соседей
9. Отбор эталонных объектов. Понятие отступа. Классификация объектов

10. Линейная регрессия. Постановка задачи. Матричная формулировка. Точное решение
11. Регуляризация в линейной регрессии. Виды регуляризаторов
12. Гребневая регрессия. Лассо-регрессия
13. Итеративная версия решения задачи линейной регрессии без регуляризации и с регуляризацией. Метод градиентного спуска
14. Вероятностная модель данных. Максимум правдоподобия
15. Линейный классификатор. Отступы. Функционал качества
16. Метод стохастического градиента (SGD). Преимущества и недостатки
17. Методы улучшения сходимости SGD (метод моментов, метод Нестерова)
18. Методы улучшения сходимости SGD (AdaGrad, RMSProp, Adam)
19. Логистическая регрессия. Простой байесовский классификатор. Сигмоидальная функция
20. Метод опорных векторов (SVM). Постановка задачи для линейно разделимой выборки без выбросов
21. Отступы в SVM. Учет выбросов. Формулировка теоремы Куна-Таккера. Двойственная задача
22. Классификация объектов в SVM. Постановка задачи через двойственные переменные λ
23. Нелинейные ядра в методе SVM. Примеры ядер. Способы их построения
24. Основные метрики качества алгоритмов. ROC-кривые
25. Методы кластеризации. Типы кластерных структур
26. Функционал качества кластеризации.
27. EM-алгоритм в кластерном анализе
28. Метод k-средних
29. Иерархическая кластеризация. Формула Ланса-Уильямса
30. Быстрая агломеративная кластеризация.
31. Регрессия по соседним объектам. Окно Парзена