Федеральное государственное бюджетное образовательное

учреждение высшего профессионального образования

«ОМСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

«УТВЕРЖДАЮ»

Проректор по УМР ОмГТУ

\_\_\_\_\_\_\_\_\_\_\_\_Л. О. Штриплинг

«\_\_\_\_» \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ **«date»** г.

# Р А Б О Ч А Я П Р О Г Р А М М А

по дисциплине

**«name\_of\_discipline»**

**«direction»**

Разработана в соответствии с ООП и ФГОС ВО по направлениям подготовки магистратуры **«kod» «direction».**

Программу составил:

доцент, канд. техн. наук \_\_\_\_\_\_\_\_\_\_/ Е. Ю. Андиева /

«\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_\_**«date»** г.

Обсуждена на заседании кафедры «Прикладная математика и фундаментальная информатика». Протокол № \_\_ от «\_\_\_» \_\_\_\_\_\_\_\_\_ 2017 г.

Зав. кафедрой «Прикладная математика и фундаментальная информатика»

\_\_\_\_\_\_\_\_\_/ А. В. Зыкина /

«\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_\_**«date»** г.

Руководитель ООП:

профессор, д. ф.-м. н.

\_\_\_\_\_\_\_\_\_\_/ А. В. Зыкина /

«\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_\_**«date»** г.

**1. Цели и задачи дисциплины**

~~Целью изучения дисциплины «Анализ данных» является введение в область data science, освоение методов машинного обучения и знакомство с технологиями big data.~~

**target**

Основные задачи дисциплины:

**needs**

**2. Место дисциплины в структуре ООП**

Дисциплина **«name\_of\_discipline»** изучается в 6 и 7 семестрах и относится к дисциплинам по выбору по направлению подготовки бакалавров **«kod» «direction».** Входные знания и умения (компетенции), необходимые для изучения дисциплины **«kod» «direction».** формируются в процессе изучения математических дисциплин (Математика, Алгебра, теория чисел, Математическая логика, Дискретная математика, Оптимизация и принятие решений в технических системах, Алгоритмы и анализ сложности и др.) и дисциплин IT-направленности (Программирование, Объектно-ориентированное программирование, Базы данных и др.) в более ранних семестрах программы бакалавриата.

**3. Требования к результатам освоения дисциплины**

**3.1. Освоение дисциплины «Анализ данных» вносит существенный вклад в формирование следующих компетенций:**

|  |  |
| --- | --- |
| **Шифр направления** | **Формируемая компетенция** |
| **«kod»** | kod\_kom |
| **«kod»** | kod\_kom |

**3.2.** В результате освоения дисциплины студент должен демонстрировать освоение указанными компетенциями по дескрипторам «знания, умения, владения», соответствующие тематическим модулям дисциплины, и применимые в их последующем обучении и профессиональной деятельности:

**Знать:**

**– З.1** основные понятия, подходы и методы, проблемные точки машинного обучения;

**– З.2** основные этапы процесса data science**;**

**– З.3** основные принципы и алгоритмы, заложенные в технологии Big Data.

**Уметь:**

– **У.1** пользоваться возможностями Python при решении задач, связанных с предобработкой данных;

– **У.2** создавать конвейеры на Python под разрабатываемые для решения задач анализа данных процессы data science;

**– У.3** работать с большими данными на одном компьютере.

**Владеть:**

**– В.1** методиками предварительной обработки данных;

**– В.2** библиотеками Python для быстрого сбора конвейера для решения задач анализа данных.

**3.3 Проектируемые результаты и признаки формирования компетенций**

***Компетентностная модель дисциплины***

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Индекс компетенции | Проектируемые результаты освоения дисциплины **«name\_of\_discipline»** и индикаторы формирования компетенций | | | Средства и технологии оценки | Технологии формирования компетенции |
| Знания (З) | Умения (У) | Владения (В) |
| ПК-2 | З.1, З.3 | У.2, У.3 | В.2 | зачет, экзамен,  РГР,  ДЗ | 1) информационно-развивающие технологии;  2) деятельностные практико-ориентированные технологии;  3) развивающие проблемно-ориентированные технологии. |
| ПК-3 | З.1, З.2 | У.1, У.2 | В.1, В.2 | зачет, экзамен,  РГР,  ДЗ | 1) информационно-развивающие технологии;  2) деятельностные практико-ориентированные технологии;  3) развивающие проблемно-ориентированные технологии. |

**4. Объем дисциплины и виды учебной работы в часах и в зачетных единицах**

**Очная форма обучения**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Вид занятий | Всего (час./  зач.ед. ) | С е м е с т р ы | | | | | | | |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| **Всего аудиторных занятий:** | **Z1** | **X4** | C7 | **B1** | N4 | **M7** | **S1** | **D4** | F7 |
| Лекции | **Z2** | **X5** | C8 | B2 | N5 | M8 | S2 | D5 | F8 |
| Практические занятия | Z3 | **X6** | C9 | B3 | N6 | M9 | S3 | D6 | F9 |
| Лабораторные работы | Z4 | X7 | V1 | B4 | N7 | A1 | S4 | D7 | G1 |
| Самостоятельная работа: | Z5 | X8 | V2 | B5 | N8 | A2 | S5 | D8 | G2 |
| Самостоятельное изучение материала дисциплины и подготовка к зачетам | Z6 | X9 | V3 | B6 | N9 | A3 | S6 | D9 | G3 |
| Курсовой проект (работа) | Z7 | C1 | V4 | B7 | M1 | A4 | S7 | F1 | G4 |
| Расчетно-графическая работа | Z8 | C2 | V5 | B8 | M2 | A5 | S8 | F2 | G5 |
| Домашнее задание | Z9 | C3 | V6 | B9 | M3 | A6 | S9 | F3 | G6 |
| Количество часов на экзамен | X1 | C4 | V7 | N1 | M4 | A7 | D1 | F4 | G7 |
| Всего по дисциплине | X2 | C5 | V8 | N2 | M5 | A8 | D2 | F5 | G8 |
| Вид аттестации за семестр (зачет, дифференцированный зачет, экзамен) | X3 | C6 | V9 | N3 | M6 | A9 | D3 | F6 | G9 |

**5. Содержание дисциплины по модулям и видам учебных занятий**

**5.1. Содержание дисциплины по модулям**

1. Введение в Data Science и Big Data.

2. Машинное обучение.

3. Построение выводов по данным.

4. Язык программирования Python.

5. Решение задач анализа данных на языке Python.

|  |  |
| --- | --- |
| **Содержание модулей** | Количество часов (очная форма обучения) |
| **Модуль 1. Введение в Data Science и Big Data** |  |
| **Тема 1. Data Science в мире больших данных.** Определение Big Data и Data Science. Взаимосвязь Big Data и Data Science. Основные категории данных. | 2/4 |
| **Тема 2. Data Science в мире больших данных.** Обзор экосистемы больших данных и Data Science. Пакета Anaconda. Этапы процесса Data Science и их содержательное наполнение. | 2/4 |
| **Тема 3. Data Science в мире больших данных.** Сбор и обработка данных.Масштабирование признаков. Работа с категориальными признаками. | 2/4 |
| **Модуль 2. Машинное обучение** |  |
| **Тема 1.** **Введение в машинное обучение.** Основные понятия машинного обучения. Обозначения и формализация задачи машинного обучения. Классификация задач машинного обучения. | 2/4 |
| **Тема 2.** **Обучение на размеченных данных: линейные модели (регрессия).** Линейная регрессия. Обучение линейной регрессии. Градиентный спуск для линейной регрессии. Стохастический градиентный спуск (SGD). | 2/4 |
| **Тема 3.** **Обучение на размеченных данных: линейные модели (классификация).** Линейная классификация. Задача оценивания вероятностей и логистическая регрессия.  **Тема 4.** **Обучение на размеченных данных: проблема переобучения.** Проблема переобучения. Регуляризация. | 2/4 |
| **Тема 5.** **Обучение на размеченных данных: метрики качества.** Оценивание качества алгоритмов. Выбор гиперпараметров. Метрики качества в задачах регрессии. Метрики качества классификации. Точность и полнота. Качество оценок принадлежности классу. | 2/4 |
| **Тема 6.** **Обучение на размеченных данных: практические рекомендации по линейным моделям.** Спрямляющие пространства  **Тема 7.** **Обучение на размеченных данных: решающие деревья.** Обучение решающих деревьев. Критерии останова и стрижка деревьев. Композиции деревьев. Смещение и разброс. | 2/4 |
| **Тема 8.** **Обучение на размеченных данных: градиентный бустинг.** Композиции простых алгоритмов. Градиентный бустинг. Обучение композиций. Борьба с переобучением. Градиентный бустинг над решающими деревьями. | 2/4 |
| **Тема 9.** **Обучение на размеченных данных: метрические алгоритмы.** Метод k ближайших соседей. Настройка параметров в kNN. Метрики в kNN. Проклятие размерности. Метод опорных векторов (SVM). Ядра в методе опорных векторов. | 2/4 |
| **Тема 10.** **Обучение на размеченных данных: быстрообучающиеся реализации SVM.** Поиск нелинейных SVM с привлечением подвыборки. SVM в крупном масштабе .Добавление нелинейности в алгоритм SGD | 2/4 |
| **Тема 11.** **Обучение на размеченных данных: введение в нейронные сети.** Однослойная нейронная сеть. Многослойная нейронная сеть. Функция ошибки. Оптимизация параметров нейронной сети. Регуляризация и прореживание нейронной сети. | 2/4 |
| **Тема 12.** **Обучение без учителя.** **Введение в кластеризацию.** Задача кластеризации. Методы кластеризации: EM-алгоритм, агломеративная иерархическая кластеризация, графовые методы кластеризации. Оценка качества и рекомендации по решению задачи кластеризации. | 2/5 |
| **Тема 13.** **Обучение без учителя. Понижение размерности и отбор признаков.** Одномерный отбор признаков. Жадные методы отбора признаков. Отбор признаков на основе моделей. Понижение размерности. Метод главных компонент. | 2/5 |
| **Тема 14.** **Матричные разложения.** Прогнозирование неизвестных значений в матрице. Вероятностный взгляд на матричные разложения. Неотрицательные матричные разложения: постановка, решение, функционалы и инициализация. | 2/5 |
| **Тема 15.** **Обучение без учителя. Поиск аномалий и визуализация данных.** Задача обнаружения аномалий. Параметрическое и непараметрическое восстановление плотности. Одноклассовый SVM. Задача визуализации. Многомерное шкалирование. Метод t-SNE. | 2/5 |
| **Модуль 3. Построение выводов по данным.** |  |
| **Тема 1. АБ-тестирование.** Что такое и где используется АБ-тестирование. Метрики. Дизайн эксперимента. Устойчивость. Размер выборки. Параметрические и непараметрические критерии.. | 2/5 |
| **Тема 2.** **Теорема Байеса в машинном обучении.** Теорема Байеса. Байесовский подход к теории вероятностей. Байесовские модели в задачах машинного обучения. | 2/5 |
| **ИТОГО ЧАСОВ** | **36/78** |

Примечание: кол-во часов (х/у, где х – лекционные занятия, у – самостоятельная проработ-ка материала).

**5.2. Содержание практических и лабораторных занятий**

**5.2.1. Содержание практических занятий**

Практические занятия не предусмотрены учебным планом.

**5.2.2. Содержание лабораторных работ**

Цель лабораторного практикума – закрепление теоретического материала дисциплины, овладение методами решения задач.

|  |  |
| --- | --- |
| **Содержание лабораторного практикума** | Количество часов (очная форма обучения) |
| **Модуль 4. Язык программирования Python** |  |
| **Лабораторная работа 1.** **Python: основы.** Использование интерпретатора Python в качестве калькулятора. Строки, операторы, индексация, слайсы, функции строк и форматирование. | 2 |
| **Лабораторная работа 2. Python: основы.** Типы данных: список, кортеж (tuple), словарь (dict), множество (set). Булевый тип данных (bool). Встроенные функции и типы. | 2 |
| **Лабораторная работа 3. Python: основы.** Shell и основные команды. Запуск python скрипта из консоли и передача в него параметров. Python Types | 2 |
| **Лабораторная работа 4**. **Python: основы.** Инструкция if. Инструкции while и for. | 2 |
| **Лабораторная работа 5**. **Python: основы.** Python Sugar. Распаковки, enumerate. Перехват ошибок. |  |
| **Лабораторная работа 6**. **Python: функции.** Документирование функции. Передача параметров в функцию. Функции, возвращающие результат. Анонимные функции lambda. | 2 |
| **Лабораторная работа 7**. **Python: функции.** Реализация стохастического градиентного спуска. | 2 |
| **Лабораторная работа** **8.** **Python: модули.** Создание и использование модулей | 2 |
| **Лабораторная работа** **9.** **Python: библиотеки.** Python Standard Library (STDLIB). Сторонние библиотеки. Установка сторонних модулей, каталог пакетов PyPI. Менеджер пакетов pip. Работаем в виртуальном окружении ( virtualenv). | 2 |
| **Лабораторная работа 10.** **Python: работа с файлами.** Запись и чтение файлов. Режимы работы с файлами. Методы работы с файлами. Запись данных в json и CSV форматах. | 2 |
| **Лабораторная работа 11.** **Python: работа с файлами.** Извлечение данных из веб-ресурсов | 2 |
| **Лабораторная работа 12.** **Python: ООП.** Создание классов. Инстанцирование. Методы экземпляров классов. | 2 |
| **Лабораторная работа 13.** **Python: ООП.** Реализация стохастического градиентного спуска II. |  |
| **Лабораторная работа 14.** **Python: ООП.** Инкапсуляция. Наследование. | 2 |
| **Лабораторная работа 15.** **Python: ООП.** Полиморфизм. Обращение к функциям родителя, метод super. Абстрактные классы. Статические методы. Методы классов. | 2 |
| **Лабораторная работа 16.** **Python: построение графиков и визуализация.** Введение в API библиотеки matplotlib. | 2 |
| **Лабораторная работа 17.** **Python: построение графиков и визуализация.** Функции построения графиков в pandas. Инструментальная экосистема визуализации для python. | 2 |
| **Лабораторная работа 18.** **Практические рекомендации.** Особенности реализации процесса Data Science средствами Python. | 2 |
| **Модуль 5. Решение задач анализа данных на языке Python** |  |
| **Лабораторная работа 1**. Реализация SVM в крупном масштабе на основе SGD. | 4 |
| **Лабораторная работа 2** Реализация перцептрона с онлайновым обучением. | 4 |
| **Лабораторная работа 3** Классификация на основе метода опорных векторов | 4 |
| **Лабораторная работа 4** Решение нелинейных задач ядерным методом SVM | 4 |
| **Лабораторная работа 5** Обучение на основе деревьев решений. Случайный лес. Объединение моделей для методов ансамблевого обучения. | 4 |
| **Лабораторная работа 6** Прогнозирование стоимости домов с помощью регрессии. | 4 |
| **Лабораторная работа 7** Регрессия со штрафом (регуляризованная регрессия) | 4 |
| **Лабораторная работа 8** Кластеризация: поиск взаимосвязанных сообщений | 4 |
| **Лабораторная работа 9** Кластеризация: поиск взаимосвязанных сообщений | 4 |
| **ИТОГО ЧАСОВ** | **72** |

**6. Образовательные технологии**

**6.1.** Для достижения планируемых результатов освоения дисциплины «Анализ данных»используются следующие образовательные технологии:

|  |  |  |  |
| --- | --- | --- | --- |
| **Методы** | **Лекция** | **Лабораторные работы** | **СРС** |
| Проблемное обучение | + | + | + |
| Работа в команде |  | + |  |
| Индивидуальное обучение |  | + | + |
| Опережающая самостоятельная работа |  | + | + |

**6.2. Интерактивные формы обучения**

Проведение занятий в интерактивной форме не предусмотрено.

**7. Самостоятельная работа студентов**

Самостоятельная работа направлена на закрепление и углубление полученных теоретических и практических знаний, развитие навыков практической работы.

**7.1. Объем самостоятельной работы и распределение по видам учебных работ в часах**

**Очная форма обучения**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Вид самостоятельной работы** | **Количество часов** | | | | | | | |
| **Семестры** | | | | | | | |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| **Подготовка к лекционным занятиям, проработка материала** |  |  |  |  |  | S6 | D9 |  |
| **Выполнение РГР/ДЗ** |  |  |  |  |  | S8 | F4 |  |
| **ИТОГО** |  |  |  |  |  | S5 | D8 |  |
| **ИТОГО по дисциплине** |  |  |  |  |  |  |  | **Z1** |

Распределение часов на выполнение СРС проведено на основе личного опыта преподавателя и рекомендаций учебника «Управление факультетом». – С. 236 – 237. (Под. Ред. С. Д. Резника. – М.: ИНФРА-М, 2008. – 696 с.).

**7.2. Темы курсовых работ**

Курсовые работы не предусмотрены учебным планом.

**7.3. Расчетно-графическая работа**

В рамках самостоятельной работ по дисциплине в 6-м семестре предусмотрена расчетно-графическая работа по теме «Построение системы машинного обучения» для решения задач анализа данных из различных прикладных областей.

**7.4. Домашнее задание**

В рамках самостоятельной работы по дисциплине в 7-м семестре предусмотрено домашнее задание по теме «Крупномасштабное машинное обучение».

**8. Методическое обеспечение системы оценки качества освоения программы модуля**

К промежуточной аттестации бакалавров по дисциплине «Анализ данных» могут привлекаться в качестве внешних экспертов: представители работодателей, представители выпускающей кафедры*.*

**8.1. Фонды оценочных средств (в соответствии с ПОмГТУ73.05-2012 «О фонде оценочных средств по дисциплине»)**

Фонд оценочных средств по дисциплине «Анализ данных» включает:

– вопросы к экзамену;

– вопросы к зачету;

– темы РГР;

– темы ДЗ.

Оценка качества освоения программы дисциплины «Анализ данных» включает текущий контроль успеваемости, промежуточную аттестацию (по модулям), итоговую аттестацию.

**8.2 Контрольные вопросы по дисциплине**

**Модуль 1**

1) Постановка задач обучения по прецедентам.

2) Объекты и признаки.

3) Типы шкал: бинарные, номинальные, порядковые, количественные.

4) Типы задач: [классификация](http://www.machinelearning.ru/wiki/index.php?title=%D0%9A%D0%BB%D0%B0%D1%81%D1%81%D0%B8%D1%84%D0%B8%D0%BA%D0%B0%D1%86%D0%B8%D1%8F), [регрессия](http://www.machinelearning.ru/wiki/index.php?title=%D0%A0%D0%B5%D0%B3%D1%80%D0%B5%D1%81%D1%81%D0%B8%D1%8F), [прогнозирование](http://www.machinelearning.ru/wiki/index.php?title=%D0%9F%D1%80%D0%BE%D0%B3%D0%BD%D0%BE%D0%B7%D0%B8%D1%80%D0%BE%D0%B2%D0%B0%D0%BD%D0%B8%D0%B5), [ранжирование](http://www.machinelearning.ru/wiki/index.php?title=%D0%A0%D0%B0%D0%BD%D0%B6%D0%B8%D1%80%D0%BE%D0%B2%D0%B0%D0%BD%D0%B8%D0%B5&action=edit).

5) Основные понятия: [модель алгоритмов](http://www.machinelearning.ru/wiki/index.php?title=%D0%9C%D0%BE%D0%B4%D0%B5%D0%BB%D1%8C_%D0%B0%D0%BB%D0%B3%D0%BE%D1%80%D0%B8%D1%82%D0%BC%D0%BE%D0%B2), [метод обучения](http://www.machinelearning.ru/wiki/index.php?title=%D0%9C%D0%B5%D1%82%D0%BE%D0%B4_%D0%BE%D0%B1%D1%83%D1%87%D0%B5%D0%BD%D0%B8%D1%8F), [функция потерь](http://www.machinelearning.ru/wiki/index.php?title=%D0%A4%D1%83%D0%BD%D0%BA%D1%86%D0%B8%D1%8F_%D0%BF%D0%BE%D1%82%D0%B5%D1%80%D1%8C) и функционал качества, [принцип минимизации эмпирического риска](http://www.machinelearning.ru/wiki/index.php?title=%D0%9F%D1%80%D0%B8%D0%BD%D1%86%D0%B8%D0%BF_%D0%BC%D0%B8%D0%BD%D0%B8%D0%BC%D0%B8%D0%B7%D0%B0%D1%86%D0%B8%D0%B8_%D1%8D%D0%BC%D0%BF%D0%B8%D1%80%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%BE%D0%B3%D0%BE_%D1%80%D0%B8%D1%81%D0%BA%D0%B0&action=edit), [обобщающая способность](http://www.machinelearning.ru/wiki/index.php?title=%D0%9E%D0%B1%D0%BE%D0%B1%D1%89%D0%B0%D1%8E%D1%89%D0%B0%D1%8F_%D1%81%D0%BF%D0%BE%D1%81%D0%BE%D0%B1%D0%BD%D0%BE%D1%81%D1%82%D1%8C), [скользящий контроль](http://www.machinelearning.ru/wiki/index.php?title=%D0%A1%D0%BA%D0%BE%D0%BB%D1%8C%D0%B7%D1%8F%D1%89%D0%B8%D0%B9_%D0%BA%D0%BE%D0%BD%D1%82%D1%80%D0%BE%D0%BB%D1%8C).

6) Линейные модели регрессии и классификации.

7) Метод наименьших квадратов. Полиномиальная регрессия.

8) Методика экспериментального исследования и сравнения алгоритмов на модельных и реальных данных.

9) Линейные классификаторы: персептрон.

10) Метод опорных векторов. Оптимальная гиперплоскость.

11) Алгоритм построения оптимальной гиперплоскости.

12) SVM-метод в пространстве признаков.

13) Задача многомерной регрессии. Простая линейная регрессия. Гребневая регрессия.

14). Регрессия с опорными векторами. Решение задачи регрессии с помощью SVM.

15) Универсальное прогнозирование в режиме онлайн.

16) Алгоритм оптимального распределения\\ потерь в режиме онлайн.

17). Усиление простых классификаторов - бустинг.

18) Записать общую формулу байесовского классификатора.

19) Какие вы знаете три подхода к восстановлению плотности распределения по выборке?

20) Что такое наивный байесовский классификатор?

21) Что такое оценка плотности Парзена-Розенблатта.

22) Выписать формулу алгоритма классификации в методе парзеновского окна.

23) На что влияет ширина окна, а на что вид ядра в методе парзеновского окна?

24) Что такое «проблема мультиколлинеарности», в каких задачах и при использовании каких алгоритмов она возникает? Какие есть подходы к её решению?

25) Что такое «смесь распределений» ?

26) Что такое ЕМ-алгоритм, какова его основная идея?

27) Какая задача решается на Е-шаге, на М-шаге? Каков вероятностный смысл скрытых переменных?

28) Последовательное добавление компонент в ЕМ-алгоритме, основная идея алгоритма.

29) Что такое стохастический ЕМ-алгоритм, какова основная идея? В чём его преимущество (какой недостаток стандартного ЕМ-алгоритма он устраняет)?

30) Что такое «выбросы»? Как осуществляется фильтрация выбросов?

**Модуль 2**

1. В чём отличия внутренних и внешних критериев?
2. Разновидности внешних критериев.
3. Разновидности критерия скользящего контроля.
4. Что такое критерий непротиворечивости? В чём его недостатки?
5. Что такое многоступенчатый выбор модели по совокупности критериев?
6. Основная идея отбора признаков методом полного перебора. Действительно ли это полный перебор?
7. Основная идея отбора признаков методом добавлений и исключений.
8. Основная идея отбора признаков методом поиска в глубину.
9. Основная идея отбора признаков методом поиска в ширину.
10. Основная идея отбора признаков с помощью случайного поиска.
11. В чём отличия случайного поиска от случайного поиска с адаптацией?
12. Метод стохастического градиента.
13. Расписать градиентный шаг для квадратичной функции потерь.
14. Расписать градиентный шаг для сигмоидной функции активации.
15. Недостатки метода стохастического градиента и как с ними бороться?
16. Определите сущность понятия «большие данные»
17. Опишите методики анализа больших данных.
18. Процесс аналитики анализа больших данных.
19. Дайте характеристику Big Data на мировом рынке.
20. Охарактеризуйте Big Data в России.
21. Определите понятие Data Mining.
22. Вопросы безопасности больших данных.

**Модуль 3**

1. Python - основная идея, подход к построению ПО, история
2. Python - структуры данных и организация памяти
3. Python - блоки и специфика синтаксиса
4. Python - выражения и функции, типы данных
5. Python – библиотека Numpy
6. Python – библиотека Scipy
7. Python – библиотека Pandas
8. Python – библиотека Scikit-lern

**9. Ресурсное обеспечение дисциплины**

**9.1.1** Лекционная аудитория, оснащенная техническими средствами (ПК, проектор). Современные специализированные компьютерные классы ОмГТУ. Всемирная информационная сеть Интернет.

**9.1.2.Технические средства обучения и контроля**

Использование проектора для презентаций докладов на лекционных занятиях.

**9.1.3. Вычислительная техника**

resourse

**9.2 Учебно-методическое и информационное обеспечение**

**9.2.1 Основная литература**

osnovnya

**9.2.2 Дополнительная**

dopolnit

**9.2.3. Периодические издания**

period

period1

**9.2.4. Информационные ресурсы**

inform