

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ  
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ  
«ОРЕНБУРГСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ»**

**Методические рекомендации для  
самостоятельной работы обучающихся по дисциплине**

Б1.В.ДВ.06.02 Основы технологии Big Data

**Направление подготовки (специальность)** 27.03.04 Управление в технических системах

**Профиль подготовки (специализация)** Интеллектуальные системы обработки информации и управления

**Квалификация выпускника** бакалавр

**Форма обучения** очная

## СОДЕРЖАНИЕ

<b>1. Организация самостоятельной работы .....</b>	
<b>2. Методические рекомендации по выполнению курсовой работы (проекта) ...</b>	
<b>3. Методические рекомендации по подготовке реферата/эссе .....</b>	
<b>4. Методические рекомендации по выполнению индивидуальных домашних заданий .....</b>	
<b>5. Методические рекомендации по самостоятельному изучению вопросов .....</b>	
<b>6. Методические рекомендации по подготовке к занятиям .....</b>	
<b>6.1. Практическое занятие № ПЗ-1, 2, 3, 4 Определение больших данных. Технологии хранения больших данных .....</b>	
<b>6.2. Практическое занятие № ПЗ-5, 6, 7, 8 Технологии обработки больших данных. Технологии анализа больших данных. Научные проблемы в области больших данных .....</b>	
<b>6.3. Практическое занятие № ПЗ-9, 10, 11, 12 Статистические методы анализа больших данных .....</b>	
<b>6.4. Практическое занятие № ПЗ-13, 14, 15, 16 Современные программные средства анализа больших данных .....</b>	

## 1. ОРГАНИЗАЦИЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ

### 1.1. Организационно-методические данные дисциплины

№ п.п.	Наименование темы	Общий объем часов по видам самостоятельной работы (из табл. 5.1 РПД)				
		подготовка курсового проекта (работы)	подготовка реферата / эссе	индивидуальные домашние задания (ИДЗ)	самостоятельное изучение вопросов (СИБ)	подготовка к занятиям (ПкЗ)
1	2	3	4	5	6	7
1	Раздел 1. Основные понятия теории Big Data	12	-	-	16	17
2	Раздел 2. Принципы анализа Big Data	11	-	-	12	18
5	<b>Итого: 86</b>	23	-	-	28	35

## 2. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПО ВЫПОЛНЕНИЮ КУРСОВОЙ РАБОТЫ (ПРОЕКТА)

### 2.1 Темы курсовых работ (проектов)

1. Языки Python и R.
2. Алгоритмы кластеризации.
3. Нечеткие методы представления данных.
4. Роль аналитика по данным (Data Scientist).
5. Программное обеспечение SAS Enterprise Miner.
6. Распознавание образов.
7. Роль больших данных в здравоохранении.
8. Нейронные сети как реализация алгоритмов машинного обучения.
9. Применение технологий больших данных для задач управления в реальном времени.

### 2.2 Порядок и сроки выполнения курсовой работы (проекта)

1. Предварительный этап:
  - выбор темы курсовой работы и оценка возможности раскрытия данной темы;
  - подача заявления на закрепление темы.
2. Основной этап:
  - составление плана;
  - сбор материала;
  - оформление курсовой работы;
3. Заключительный этап:
  - рецензирование работы руководителем и допуск к защите;
  - защита работы.

### 2.3 Структура курсовой работы (проекта)

- титульный лист;
- содержание;
- введение;
- основная часть;
- заключение;
- список использованной литературы;
- приложения.

### 2.4 Требования к оформлению курсовой работы (проекта)

Страницы текста работы должны соответствовать формату А4 ГОСТ 9327. Текст должен быть выполнен с использованием компьютера и принтера на одной стороне листа белой бумаги через полтора интервала. Цвет шрифта должен быть черным, высота букв, цифр и других знаков должна составлять 14 пунктов (кегель 14), форматирование текста по ширине, заголовков — по центру. При использовании текстового редактора Microsoft Word рекомендуется использовать стандартную гарнитуру шрифта Times New Roman. Текст следует печатать, соблюдая следующие размеры полей: левое – 2,5 см, правое – 1,5 см, верхнее – 2 см, нижнее – 2 см. Абзацный отступ – 1,25 см.

В конце работы указываются использованные учебники и учебные пособия использованные для выполнения курсовой работы.  
Работа должна быть подписана с указанием даты ее окончания.

## 2.5 Критерии оценки

- сроки сдачи;
- правильность и аккуратность оформления;
- соответствие оформление курсовой работы (проекта) установленным требованиям;
- умение работать с документальными и литературными источниками;
- умение формулировать основные выводы по результатам анализа конкретного анализа;

## 2.6 Рекомендованная литература

1. Чекмарев Ю.В. Вычислительные системы, сети и телекоммуникации [Электронный ресурс]/ Чекмарев Ю.В.— Электрон. текстовые данные.— М.: ДМК Пресс, 2013.— 184 с.— Режим доступа: <http://www.iprbookshop.ru/5083>.— ЭБС «IPRbooks», по паролю
2. Управление данными. Цехановский В.В. Чертовской В.Д. Издательство: Лань. 2015г.-432 с.
3. Борисова И.В. Цифровые методы обработки информации [Электронный ресурс]: учебное пособие/ Борисова И.В.— Электрон. текстовые данные.— Новосибирск: Новосибирский государственный технический университет, 2014.— 139 с.— Режим доступа: <http://www.iprbookshop.ru/45061>.— ЭБС «IPRbooks», по паролю

**3 МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ  
ПО ПОДГОТОВКЕ РЕФЕРАТА/ЭССЕ**

*Не предусмотрен рабочим учебным планом.*

**4. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПО ВЫПОЛНЕНИЮ  
ИНДИВИДУАЛЬНЫХ ДОМАШНИХ ЗАДАНИЙ**

*Не предусмотрены рабочим учебным планом.*

## **5. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПО САМОСТОЯТЕЛЬНОМУ ИЗУЧЕНИЮ ВОПРОСОВ**

**5.1 Наименование вопроса.** Технологии Business Intelligence и реляционные системы управления базами данных

При изучении вопроса необходимо обратить внимание на следующие особенности.

- Понятие базы данных и история развития систем управления базами данных

**5.2 Наименование вопроса.** Методы анализа социально-политических и медиа процессов

При изучении вопроса необходимо обратить внимание на следующие особенности.

- Использование модели множественной линейной регрессии для прогнозирования экономических показателей

**5.3 Наименование вопроса.** Предсказание и прогнозирование социально-экономических прогнозов

При изучении вопроса необходимо обратить внимание на следующие особенности.

- Прогнозирование и предвидение: общее и особенное

**5.4 Наименование вопроса.** Специализированные статистические пакеты

При изучении вопроса необходимо обратить внимание на следующие особенности.

- Специальные методы анализа социально-политических и медиа процессов.

## **6. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПО ПОДГОТОВКЕ К ЗАНЯТИЯМ**

**6.1 Практическое занятие № ПЗ-1, 2, 3, 4** Определение больших данных. Технологии хранения больших данных

При подготовке к занятию необходимо обратить внимание на следующие моменты.

- Совокупность подходов, инструментов и методов обработки структурированных и неструктурированных данных огромных объёмов, и значительного многообразия для получения воспринимаемых человеком результатов, эффективных в условиях непрерывного прироста, распределения по многочисленным узлам вычислительной сети

**6.2 Практическое занятие № ПЗ-5, 6, 7, 8** Технологии обработки больших данных. Технологии анализа больших данных. Научные проблемы в области больших данных

При подготовке к занятию необходимо обратить внимание на следующие моменты.

- Методы и техники анализа, применимые к большим данным: методы класса Data Mining: обучение ассоциативным правилам (англ. association rule learning), классификация (методы категоризации новых данных на основе принципов, ранее применённых к уже наличествующим данным), кластерный анализ, регрессионный анализ

**6.3 Практическое занятие № ПЗ-9, 10, 11, 12** Статистические методы анализа больших данных

При подготовке к занятию необходимо обратить внимание на следующие моменты.

- Понятие прогноза и предвидения. Отличие прогнозирования от предвидения. Закон распределения случайной величины. Статистические оценки параметров. Доверительные области. Корреляционный анализ. Использование модели множественной линейной регрессии для прогнозирования экономических показателей

**6.4 Практическое занятие № ПЗ-13, 14, 15, 16** Современные программные средства анализа больших данных

При подготовке к занятию необходимо обратить внимание на следующие моменты.

- SPSS Statistics (аббревиатура англ. "Statistical Package for the Social Sciences", "статистический пакет для социальных наук") - компьютерная программа для статистической обработки данных, один из лидеров рынка в области коммерческих статистических продуктов, предназначенных для проведения прикладных исследований в социальных науках