

**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ  
РОССИЙСКОЙ ФЕДЕРАЦИИ**

Филиал федерального государственного бюджетного образовательного  
учреждения высшего образования  
«Воронежский государственный технический университет»  
в городе Борисоглебске



УТВЕРЖДАЮ

Директор филиала

 /В.В. Григораш/

31 августа 2021 г.

**РАБОЧАЯ ПРОГРАММА  
дисциплины**

**«Технологии обработки больших данных»**

**Направление подготовки 09.03.02 Информационные системы и технологии**

**Профиль Информационные системы и технологии цифровизации**

**Квалификация выпускника бакалавр**

**Нормативный период обучения 4 г**

**Форма обучения очная**

**Год начала подготовки 2020**

Автор программы



О.Г. Яскевич

Заведующий кафедрой  
естественнонаучных дисциплин



Л.И. Матвеева

Руководитель ОПОП



Е.А. Позднова

**Борисоглебск 2021**

# 1. ЦЕЛИ И ЗАДАЧИ ДИСЦИПЛИНЫ

## 1.1. Цель дисциплины

Целью дисциплины является освоение принципов, методов, технологий и инструментов обработки и использования больших данных в информационных системах

## 1.2. Задачи освоения дисциплины

- ознакомление студентов с основными классами задач представления и обработки больших данных;
- изучение методов и технологий подготовки, хранения, обработки и анализа больших данных;
- освоение методов аналитической обработки больших объёмов данных в информационных системах;
- приобретение навыков использования технологий и инструментов Big Data при решении практических задач.

## 2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП

Дисциплина «Технологии обработки больших данных» относится к дисциплинам части, формируемой участниками образовательных отношений блока Б.1 учебного плана.

## 3. ПЕРЕЧЕНЬ ПЛАНИРУЕМЫХ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ

Процесс изучения дисциплины «Технологии обработки больших данных» направлен на формирование следующих компетенций:

ПК-2 - Способен выполнять проектирование информационных систем и ресурсов для различных прикладных областей

ПК-1 - Способен выполнять работы по созданию (модификации) и сопровождению информационных систем и ресурсов для различных прикладных областей

<b>Компетенция</b>	<b>Результаты обучения, характеризующие сформированность компетенции</b>
ПК-1	Знать методы обработки и анализа больших данных
	Уметь использовать технологии обработки больших данных при проектировании и эксплуатации информационных систем
	Владеть инструментами Big Data
ПК-2	Знать технологии распределённой обработки данных
	Уметь разрабатывать компоненты программных систем для обработки и анализа больших данных
	Владеть навыками практического применения технологий Big Data в различных прикладных областях

#### 4. ОБЪЕМ ДИСЦИПЛИНЫ

Общая трудоемкость дисциплины «Технологии обработки больших данных» составляет 5 зачетных единиц.

Распределение трудоемкости дисциплины по видам занятий

##### Очная форма обучения

Вид учебной работы	Всего часов	Семестры
		7
<b>Аудиторные занятия (всего)</b>	72	14
В том числе:		
Лекции	18	18
Практические занятия (ПЗ)	-	-
Лабораторные работы (ЛР)	54	54
<b>Самостоятельная работа</b>	72	72
Курсовой проект(работа)	-	-
Контрольная работа	-	-
Вид промежуточной аттестации (Экзамен)	36	36
Общая трудоемкость	час	180
	зач. ед.	5

#### 5. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

5.1 Содержание разделов дисциплины и распределение трудоемкости по видам занятий

##### очная форма обучения

№ п/п	Наименование темы	Содержание раздела	Лекц	Лаб. зан.	СРС	Всего, час
1	Введение в Big Data	Понятие больших данных. Особенности сбора, хранения, обработки и анализа больших массивов данных. Источники больших данных. Использование больших данных в информационных системах	2	2	10	14
2	Модели и методы анализа больших данных	Основные задачи анализа больших данных в информационных системах. Классификация современных методов анализа данных. Особенности технологии Data Mining. Использование методов Data Mining для анализа больших данных. Решение задач классификации, кластеризации, прогнозирования на основе больших данных. Нейросетевые технологии обработки информации. Технологии визуализации и трансформации данных. Многомерный анализ данных. OLAP-технологии.	6	16	20	42

3	Базы данных NoSQL	Варианты построения распределенных баз данных, репликация, фрагментация. Согласованность. CAP-теорема. Классы NoSQL баз данных. Семейства столбцов. Графовые СУБД. Документо-ориентированные распределённые СУБД. Понятие агрегата. Современные документо-ориентированные СУБД. Запросы к СУБД на языке JSON. Примеры СУБД NoSQL. Организация и принципы функционирования MongoDB.	4	16	20	40
4	Технологии распределённой обработки данных	Базовые архитектуры распределённой обработки данных. Параллельные и облачные вычисления. Использование фреймворка MapReduce в распределенной среде. Реализации MapReduce. Состав и возможности платформы распределённых вычислений Hadoop. Языки поисковых запросов для Hadoop. Распределённые файловые системы. Архитектура и реализация распределённых файловых систем. HDFS. Распределенные базы данных. HBASE. Обзор современных инструментов BigData.	6	20	22	48
<b>Итого</b>			<b>18</b>	<b>54</b>	<b>72</b>	<b>144</b>

## 5.2 Перечень лабораторных работ

1. Основные конструкции языка Python
2. Обработка и анализ данных с использованием библиотек Python
3. Платформа Hadoop и ее компоненты
4. Распределенные файловая система HDFS
5. Модель вычислений MapReduce
6. Распределенные вычисления на платформе Apache Spark
7. Apache Kafka. Поточковая обработка данных
8. Изучение возможностей и работа с документо-ориентированной БД MongoDB

## 6. ПРИМЕРНАЯ ТЕМАТИКА КУРСОВЫХ ПРОЕКТОВ (РАБОТ) И КОНТРОЛЬНЫХ РАБОТ

Учебным планом по дисциплине «Технологии обработки больших данных» не предусмотрено выполнение курсовых проектов (работ) и контрольной работы (контрольных работ) в 7 семестре.

## 7. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ДЛЯ ПРОВЕДЕНИЯ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ

### 7.1. Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания

#### 7.1.1 Этап текущего контроля

Результаты текущего контроля знаний и межсессионной аттестации оцениваются по следующей системе:

«аттестован»;

«не аттестован».

Компетенция	Результаты обучения,	Критерии оценивания	Аттестован
-------------	----------------------	---------------------	------------

	характеризующие сформированность компетенции		
ПК-2	<b>Знать</b> методы обработки и анализа больших данных	Знание особенностей технологии Big Data. Знание моделей и методов Data Mining, классов решаемых задач. Ответы на теоретические вопросы при защите лабораторных работ	Выполнение работ в срок, предусмотренный в рабочих программах
	<b>Уметь</b> использовать технологии обработки больших данных при проектировании и эксплуатации информационных систем	Умение решать задачи классификации, распознавания и прогнозирования с использованием интеллектуальных методов анализа и обработки больших данных. Выполнение лабораторных работ	Выполнение работ в срок, предусмотренный в рабочих программах
	<b>Владеть</b> инструментами Big Data	Владение навыками использования инструментальных средств Big Data при решении практических задач. Выполнение лабораторных работ	Выполнение работ в срок, предусмотренный в рабочих программах
ПК-1	<b>Знать</b> технологии распределённой обработки данных	Знание методов и средств распределённой обработки информации. Ответы на теоретические вопросы при защите лабораторных работ	Выполнение работ в срок, предусмотренный в рабочих программах
	<b>Уметь</b> разрабатывать компоненты программных систем для обработки и анализа больших данных	Умение разрабатывать программное обеспечение для решения задач обработки больших данных. Выполнение лабораторных работ	Выполнение работ в срок, предусмотренный в рабочих программах
	<b>Владеть</b> навыками практического применения технологий Big Data в различных прикладных областях	Решение прикладных задач в конкретной предметной области, выполнение лабораторных работ	Выполнение работ в срок, предусмотренный в рабочих программах

### 7.1.2 Этап промежуточного контроля знаний

Результаты промежуточного контроля знаний оцениваются в 7 семестре для очной формы обучения по системе:

- «отлично»;
- «хорошо»;
- «удовлетворительно»;
- «неудовлетворительно».

Компетенция	Результаты обучения, характеризующие сформированность компетенции	Критерии оценивания	Отлично	Хорошо	Удовл.	Неудовл.
ПК-2	Знать методы обработки и анализа больших данных	Тест	Выполнение теста на 90- 100%	Выполнение теста на 80- 90%	Выполнение теста на 70- 80%	В тесте менее 70% правильных ответов

	Уметь использовать технологии обработки больших данных при проектировании и эксплуатации информационных систем	Решение стандартных практически задач	Задачи решены в полном объеме и получены верные ответы	Продемонстрирован верный ход решения всех, но не получен верный ответ во всех задачах	Продемонстрирован верный ход решения в большинстве задач	Задачи не решены
	Владеть инструментами Big Data	Решение прикладных задач в конкретной предметной области	Задачи решены в полном объеме и получены верные ответы	Продемонстрирован верный ход решения всех, но не получен верный ответ во всех задачах	Продемонстрирован верный ход решения в большинстве задач	Задачи не решены
ПК-1	Знать технологии распределённой обработки данных	Тест	Выполнение теста на 90- 100%	Выполнение теста на 80- 90%	Выполнение теста на 70- 80%	В тесте менее 70% правильных ответов
	Уметь разрабатывать компоненты программных систем для обработки и анализа больших данных	Решение стандартных практически задач	Задачи решены в полном объеме и получены верные ответы	Продемонстрирован верный ход решения всех, но не получен верный ответ во всех задачах	Продемонстрирован верный ход решения в большинстве задач	Задачи не решены
	Владеть навыками практического применения технологий Big Data в различных прикладных областях	Решение прикладных задач в конкретной предметной области	Задачи решены в полном объеме и получены верные ответы	Продемонстрирован верный ход решения всех, но не получен верный ответ во всех задачах	Продемонстрирован верный ход решения в большинстве задач	Задачи не решены

## 7.2 Примерный перечень оценочных средств ( типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности)

### 7.2.1 Примерный перечень заданий для подготовки к тестированию

1 Какими свойствами, как правило, обладают большие данные? (возможно несколько ответов)

- а) структурированность;
- б) неструктурированность;**
- в) объем в терабайтах и петабайтах;**
- г) объем в мегабайтах и гигабайтах.

2 Какая технология используется для оперативного анализа данных?

- а) Data Mining;
- б) OLTP;
- в) DSS;
- г) OLAP.**

3 Каким требованиям должна удовлетворять система анализа данных? (возможно несколько ответов)

а) Доступ к данным пользователя осуществляется по заранее составленным запросам

**б) Запросы к данным могут быть произвольными и заранее не определены;**

**в) Должна быть возможность периодически обновлять данные;**

г) Время отклика системы измеряется в секундах

4 Какие модели данных могут использоваться при реализации OLAP-систем? (возможно несколько ответов)

а) фактографическая;

**б) реляционная;**

в) семантическая;

г) **многомерная.**

5 Какой метод применяется для обнаружения скрытых знаний в больших объемах данных?

а) **Data Mining;**

б) дискриминантный анализ;

в) OLAP;

г) дисперсионный анализ.

6 Какие методы применяются для анализа неструктурированных текстовых данных? (возможно несколько ответов)

а) DSS Mining;

**б) Web Mining;**

в) Machine Mining;

г) **Text Mining.**

7 Обучающей парой называется

**а) входной и выходной векторы**

б) входной вектор и активационная функция

в) входной вектор и вектор весов

г) выходной вектор и активационная функция

8 Какой из методов Data Mining может использоваться для определения групп потребителей со схожими стереотипами поведения с целью предложения интересных именно им товаров?

**а) кластерный анализ;**

б) генетический алгоритм;

в) поиск ассоциативных правил;

г) нечеткая логика.

9 Какой из алгоритмов относится к алгоритмам кластеризации?

- 1 Алгоритм Дейкстры
- 2 k-means – алгоритм**
- 3 Генетический алгоритм
- 4 Алгоритм Форда

10 В OLAP-системах наборы данных, необходимые для описания событий (товары, поставщики и т.д.) являются...

- а) фактами;
- б) измерениями;**
- в) ресурсами;
- г) названиями.

### 7.2.2 Примерный перечень заданий для решения стандартных задач

1 Какая из библиотек Python предназначена для работы с нейронными сетями?

- Pandas
- Keras**
- NumPy
- Statsmodels

2 Какая из библиотек Python предназначена для машинного обучения?

- Scikit Learn**
- Keras
- NumPy
- Pandas

3 Какая из перечисленных программных систем не может быть использована для анализа данных?

- Matlab
- EXCEL
- 3dMax**
- Statistica

4 Как запустить задание MapReduce из java кода?

- Вызвать метод submit() экземпляра класса JobClient
- Вызвать метод submit() и затем start() экземпляра класса JobClient
- Вызвать метод submit() экземпляра класса Job**
- Вызвать метод submit() и затем start() экземпляра класса Job



5 В чём отличие системы хранения и анализа данных Hadoop от реляционной СУБД? Выберите все подходящие варианты ответа (несколько вариантов).

1) Реляционные СУБД используют как интерактивный, так и пакетный режим доступа, в то время как Hadoop - только пакетный.

2) Реляционные СУБД позволяют выполнять многократное чтение и запись данных, в то время как Hadoop - однократную запись и многократное чтение.

3) Реляционные СУБД работают только со структурированными наборами данных, в то время как Hadoop использует слабо и неструктурированные данные.

4) Реляционные СУБД масштабируются линейно, в то время как Hadoop - нелинейно

6 Сколько попыток запустить задачу на трекере задач (TaskTracker) выполняет по умолчанию трекер заданий (JobTracker), прежде чем считает задачу сбойной?

2

3

1

4

7. К OLAP-системе сформулированы следующие требования:

- время ответа системы на запросы является наиболее критичным параметром;

- таблицы измерений изменяются редко.

Какую модель данных лучше подходит для реализации системы:

**а) многомерная;**

б) реляционная, схема «звезда»;

в) реляционная, схема «снежинка»;

г) логическая.

8 Какие программные продукты имеют встроенные компоненты, реализующие методы Data Mining? (возможно несколько ответов)

**а) STATISTICA;**

**б) Microsoft SQL Server;**

в) Microsoft Excel;

**г) Oracle.**

9 Какое свойство алгоритма анализа данных состоит в том, что алгоритм решения задачи должен быть применим для некоторого класса задач, различающихся лишь значениями входных данных?

- 1 **Массовость**
- 2 Дискретность
- 3 Конечность
- 4 Детерминированность

10 С чем может быть связано получение неверного результата при программной реализации вычислительного метода? Укажите неправильный ответ

- 1 **Неверно выбран язык программирования**
- 2 Ошибка в программном коде
- 3 Неверно выбраны параметры метода
- 4 Неправильно построен алгоритм решения задачи

### 7.2.3 Примерный перечень заданий для решения прикладных задач

1 Какое количество фоновых процессов менеджера ресурсов и менеджера узлов работает при стандартной конфигурации кластера с YARN архитектурой?

- 1) **Один процесс менеджера ресурсов и один или более процессов менеджера узлов (по одному на каждом узле кластера)**
- 2) Один процесс менеджера ресурсов и один процесс менеджера узлов
- 3) Один или более процессов менеджера ресурсов (по одному на каждом узле кластера) и один процесс менеджера узлов
- 4) Один или более процессов менеджера ресурсов (по одному на каждом узле кластера) и один или более процессов менеджера узлов (по одному на каждом узле кластера)

2 Для чего в HDFS применяется инструмент архивации файлов HAR?

- 1) Для уменьшения места, которое занимают файлы в HDFS
- 2) Для уменьшения трафика между узлами данных
- 3) HAR не применяется в HDFS
- 4) **Ни один из указанных вариантов**

3 Каков размер блока HDFS по умолчанию (для релиза 2.6.0 и выше)?

- 1) 512 байт
- 2) **128 Мбайт**
- 3) 256 Мбайт

4 Что делает команда `hadoop fs -move <localsrc> <dst>?`

- 1) **Такой команды не существует**
- 2) Перемещает файл из локальной файловой системы в HDFS
- 3) Перемещает (переименовывает) файл из одной директории в HDFS в другую директорию в HDFS
- 4) Ни один из указанных вариантов

5 Для чего предназначена утилита `distcp`?

- 1) Для параллельного копирования больших объёмов данных из локальной файловой системы в HDFS
- 2) Для параллельного копирования образа файловой системы между основным узлом имён и резервным
- 3) Для параллельного копирования больших объёмов данных между кластерами Hadoop**
- 4) Ни один из указанных вариантов

6 Каким образом настроить трекер задач (TaskTracker) так, чтоб он долго выполняемую задачу не помечал как сбойную и не завершал её работу?

- 1) Задать свойство `mapred.task.timeout=0`**
- 2) Задать свойство `mapred.task.kill.hangup.task=false`
- 3) Задать свойство `mapred.task.timeout=NO_TIMEOUT`
- 4) Ни один из указанных вариантов

7 Перечислите правильную последовательность этапов Knowledge Discovery in Databases –процесса обнаружения знаний в базах данных:

- 1 трансформация, интерпретация результатов, выборка, очистка, построение моделей.
- 2 построение моделей, выборка, очистка, трансформация, интерпретация результатов.
- 3 построение моделей, выборка, очистка, трансформация, интерпретация результатов,
- 4 выборка, очистка, трансформация, построение моделей, интерпретация результатов.**

8 Для чего используется разрешение исполнения (x) для файлов в HDFS?

- 1)Разрешение исполнения для файлов в HDFS игнорируется**
- 2)Разрешение исполнения присваивается файлам с расширением '.jar для того, что задания (Job) могли использовать эти файлы
- 3)Разрешение исполнения присваивается файлам с расширением '.jar, '.zip и другим архивным файлам для их успешной распаковки в HDFS

9 Поставьте в соответствие:

- 1) Пространственные данные
- 2) Выборочные данные
- 3) Временные данные

А) совокупность информации, относящейся к разным объектам, полученной за один и тот же период или момент времени

Б) совокупность информации, относящейся к разным объектам, полученной за один и тот же период или момент времени

**1 1-А, 3-Б**

- 2 1-А, 2-Б
- 3 1-Б, 3-А
- 4 2-А, 3-Б

10 Для чего предназначена команда `hdfs dfsadmin -refreshNodes`?

- 1) Проверяет доступность всех узлов данных кластера и выдаёт список существующих, но недоступных по каким-либо причинам узлов
- 2) Вызывает повторное чтение конфигурационных файлов на всех узлах кластера (NameNode и DataNode)
- 3) **Обновляет узел имён (NameNode) новым набором разрешённых узлов данных**
- 4) Повторно запускает фоновые процессы на всех узлах кластера

#### **7.2.4 Примерный перечень вопросов для подготовки к зачету**

Не предусмотрено учебным планом

#### **7.2.5 Примерный перечень заданий для подготовки к экзамену**

- 1. Основные задачи обработки больших данных в информационных системах.
- 2. Классификация современных методов анализа данных. Понятие о технологии Data Mining.
- 3. Понятие искусственной нейронной сети. Задачи, решаемые с использованием искусственных нейронных сетей.
- 4. Классификация искусственных нейронных сетей. Однослойные и многослойные сети.
- 5. Подходы к обучению искусственных нейронных сетей. Обучение с учителем. Обучение без учителя.
- 6. Решение задач прогнозирования на основе нейросетевых технологий.
- 7. Решение задач кластеризации на основе нейросетевых технологий.
- 8. Решение задач классификации и распознавания на основе нейросетевых технологий.
- 9. OLAP-технологии.
- 10. Требования к распределенным информационным системам
- 11. Средства построения распределенных информационных систем
- 12. Технология Map-Reduce
- 13 Система Apache Hadoop
- 14. Базы данных NoSQL. Особенности, классификация
- 15 Возможности NoSQL-баз данных по обеспечению целостности, доступности скорости обработки информации. CAP-теорема.
- 16 Способы репликации и кластеризации баз данных
- 17 Документо-ориентированные базы данных
- 18 Возможности СУБД MongoDB
- 19 Работа с документо-ориентированными БД на языке JSON.
- 20. Современные инструменты BigData

#### **7.2.6. Методика выставления оценки при проведении промежуточной аттестации**

Экзамен проводится по билетам, каждый из которых содержит 2 теоретических вопроса и задачу. Каждый правильный ответ на вопрос в билете оценивается 5 баллами, задача оценивается в 10 баллов (5 баллов за верное решение и 5 баллов за верный ответ). Максимальное количество набранных баллов – 20. В процессе ответа преподаватель задаёт студенту дополнительные вопросы по тематике излагаемого материала.

Критерии оценивания ответов на теоретические вопросы:

5 баллов – студент даёт полный развёрнутый ответ на вопрос, демонстрирует уверенное владение материалом, излагает ответ логично и последовательно, обосновывает все выводы и положения, теоретический материал иллюстрирует примерами, отвечает на дополнительные вопросы преподавателя, свободно владеет терминологией дисциплины.

4 балла – студент даёт достаточно полный ответ на вопрос, показывает умение пользоваться основными понятиями и определениями, отвечает на дополнительные вопросы, однако имеются незначительные неточности в ответах и затруднения в иллюстрации теоретического материала практическими примерами;

3 балла – студент демонстрирует понимание сути вопроса, отвечает на вопрос по существу, правильно применяет теоретические положения, однако ответ представлен недостаточно полно, выводы и положения недостаточно обоснованы, имеются затруднения в ответах на дополнительные вопросы;

2 балла – студент имеет обобщённые знания основных разделов изучаемой дисциплины и содержания лекционных курсов, но не усвоил его деталей. Ответ представлен на обобщённом уровне, содержит существенные неточности, недостаточно правильные формулировки, нарушения логичности и последовательности изложения материала. Отсутствуют ответы более чем на 50% дополнительных вопросов.

1 балл – студент демонстрирует поверхностные знания материала, ответ представлен неполно (на уровне основных понятий) и содержит существенные ошибки, отсутствует логичность и последовательность в изложении, студент не отвечает на большинство дополнительных вопросов.

0 баллов – ответ на вопрос отсутствует или является полностью неверным.

Критерии оценивания результатов экзамена:

1. Оценка «Неудовлетворительно» ставится в случае, если студент набрал менее 6 баллов.

2. Оценка «Удовлетворительно» ставится в случае, если студент набрал от 6 до 10 баллов

3. Оценка «Хорошо» ставится в случае, если студент набрал от 11 до 15 баллов.

**4. Оценка «Отлично» ставится, если студент набрал от 16 до 20 баллов**

#### **7.2.7 Паспорт оценочных материалов**

№ п/п	Контролируемые разделы (темы) дисциплины	Код контролируемой компетенции	Наименование оценочного средства
1	Введение в Big Data	ПК-2, ПК-1	Тест, защита лабораторных работ

2	Модели и методы анализа больших данных	ПК-2, ПК-1	Тест, защита лабораторных работ
3	Базы данных NoSQL	ПК-2, ПК-1	Тест, защита лабораторных работ
4	Технологии распределённой обработки данных	ПК-2, ПК-1	Тест, защита лабораторных работ

### **7.3. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности**

Тестирование осуществляется, либо при помощи компьютерной системы тестирования, либо с использованием выданных тест-заданий на бумажном носителе. Время тестирования 30 мин. Затем осуществляется проверка теста экзаменатором и выставляется оценка согласно методике выставления оценки при проведении промежуточной аттестации.

Решение стандартных задач осуществляется, либо при помощи компьютерной системы тестирования, либо с использованием выданных задач на бумажном носителе. Время решения задач 30 мин. Затем осуществляется проверка решения задач экзаменатором и выставляется оценка, согласно методике выставления оценки при проведении промежуточной аттестации.

Решение прикладных задач осуществляется, либо при помощи компьютерной системы тестирования, либо с использованием выданных задач на бумажном носителе. Время решения задач 30 мин. Затем осуществляется проверка решения задач экзаменатором и выставляется оценка, согласно методике выставления оценки при проведении промежуточной аттестации.

## **8 УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ**

### **8.1 Перечень учебной литературы, необходимой для освоения дисциплины**

1. Питолин, А.В. Искусственные нейронные сети. Теория и практика : Учеб. пособие. - Воронеж : ВГТУ, 2007. – 129 с.
2. Пальмов, С. В. Интеллектуальный анализ данных [Электронный ресурс] : Учебное пособие / С. В. Пальмов. - Самара : Поволжский государственный университет телекоммуникаций и информатики, 2017. - 127 с. - ISBN 2227-8397.URL: <http://www.iprbookshop.ru/75376.html>
3. Кандаурова, Н.В. Технологии обработки информации [Электронный ресурс] : учебное пособие / В.С. Чеканов; Н.В. Кандаурова. - Ставрополь: Северо-Кавказский федеральный университет, 2014. - 175 с. URL: <http://www.iprbookshop.ru/63145.html>
4. Минаева Ю.В. Методы статистического и интеллектуального анализа данных: учеб. пособие. – Воронеж: ВГТУ, 2017. 90 с.

### **8.2 Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень лицензионного программного обеспечения, ресурсов информационно-телекоммуникационной сети «Интернет», современных профес-**

### **сиональных баз данных и информационных справочных систем:**

Перечень ПО, включая перечень лицензионного программного обеспечения:

Microsoft Office 64-bit;  
ОС Windows 7 Pro;  
Mozilla Firefox 81.0 (x64 ru);  
Google

Chrome

Ресурсы информационно-телекоммуникационной сети «Интернет»:

<http://window.edu.ru>

<http://www.edu.ru/>

Образовательный портал ВГТУ

Современных профессиональных баз данных и информационных справочных систем:

<https://habr.com/ru/>

<https://sources.ru>

<http://bigor.bmstu.ru>

## **9 МАТЕРИАЛЬНО-ТЕХНИЧЕСКАЯ БАЗА, НЕОБХОДИМАЯ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА**

Для проведения обучения по дисциплине используется компьютерный класс.

Компьютерный класс оснащен:

- персональные компьютеры с установленным ПО, подключенные к сети Интернет (11 шт.);
- принтер цветной лазерный;
- доска магнитно-маркерная поворотная.

Помещение для самостоятельной работы. Читальный зал с возможностью подключения к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду

## **10. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ**

По дисциплине «Технологии обработки больших данных» читаются лекции, проводятся лабораторные работы.

Основой изучения дисциплины являются лекции, на которых излагаются наиболее существенные и трудные вопросы, а также вопросы, не нашедшие отражения в учебной литературе.

Лабораторные работы выполняются на лабораторном оборудовании в соответствии с методиками, приведенными в указаниях к выполнению работ.

Вид учебных занятий	Деятельность студента
---------------------	-----------------------

Лекция	Написание конспекта лекций: кратко, схематично, последовательно фиксировать основные положения, выводы, формулировки, обобщения; пометать важные мысли, выделять ключевые слова, термины. Проверка терминов, понятий с помощью энциклопедий, словарей, справочников с выписыванием толкований в тетрадь. Обозначение вопросов, терминов, материала, которые вызывают трудности, поиск ответов в рекомендуемой литературе. Если самостоятельно не удастся разобраться в материале, необходимо сформулировать вопрос и задать преподавателю на лекции или на практическом занятии.
Лабораторная работа	Лабораторные работы позволяют научиться применять теоретические знания, полученные на лекции при решении конкретных задач. Чтобы наиболее рационально и полно использовать все возможности лабораторных для подготовки к ним необходимо: следует разобрать лекцию по соответствующей теме, ознакомиться с соответствующим разделом учебника, проработать дополнительную литературу и источники, решить задачи и выполнить другие письменные задания.
Самостоятельная работа	Самостоятельная работа студентов способствует глубокому усвоению учебного материала и развитию навыков самообразования. Самостоятельная работа предполагает следующие составляющие: - работа с текстами: учебниками, справочниками, дополнительной литературой, а также проработка конспектов лекций; - выполнение домашних заданий и расчетов; - работа над темами для самостоятельного изучения; - участие в работе студенческих научных конференций, олимпиад; подготовка к промежуточной аттестации.
Подготовка к промежуточной аттестации	Готовиться к промежуточной аттестации следует систематически, в течение всего семестра. Интенсивная подготовка должна начаться не позднее, чем за месяц-полтора до промежуточной аттестации. Данные перед экзаменом три дня эффективнее всего использовать для повторения и систематизации материала.





## ЛИСТ РЕГИСТРАЦИИ ИЗМЕНЕНИЙ

№ п/п	Перечень вносимых изменений	Дата вне- сения из- менений	Подпись заведую- щего кафедрой, от- ветственной за реал- лизацию ОПОП
1	Актуализирован раздел 8.2 в части состава используемого лицензионного программного обеспечения, современных профессиональных баз данных и справочных информационных систем	31.08.2021	