



**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ  
РОССИЙСКОЙ ФЕДЕРАЦИИ**  
ФГБОУ ВО «Брянский государственный технический  
университет» (БГТУ)

**Факультет информационных технологий**  
*(наименование факультета/института)*

**Кафедра «Компьютерные технологии и системы»**  
*(наименование кафедры, ответственной за реализацию дисциплины)*

**УТВЕРЖДАЮ**  
Первый проректор по учебной  
работе и цифровизации  
\_\_\_\_\_ В.А. Шкаберин  
« 26 » апреля 2024 г.

**РАБОЧАЯ ПРОГРАММА**  
**учебной дисциплины**

**«Методы анализа данных и естественно-языковых текстов»**  
*(наименование дисциплины)*

**09.03.02 Информационные системы и технологии**  
*(код и наименование специальности или направления подготовки)*

**Искусственный интеллект и обработка больших данных**  
*(направленность (профиль)/ специализация образовательной программы)*

**высшее образование – бакалавриат**  
*(уровень образования)*

**бакалавр**  
*(квалификация, присваиваемая по специальности или направлению подготовки)*

**очная**  
*(форма обучения)*

**2024**  
*(год набора)*

**Брянск 2024**

Рабочая программа учебной дисциплины  
«Методы анализа данных и естественно-языковых текстов»

(наименование дисциплины)

09.03.02 Информационные системы и технологии

(код и наименование специальности или направления подготовки)

Искусственный интеллект и обработка больших данных

(направленность (профиль)/специализация образовательной программы)

**Разработал(и):**

Доцент, к.б.н., доц

(должность, ученая степень, ученое звание)

(подпись)

А.А. Кузьменко

(И.О. Фамилия)

(должность, ученая степень, ученое звание)

(подпись)

(И.О. Фамилия)

Рассмотрена и одобрена на заседании кафедры  
«Компьютерные технологии и системы»

(наименование кафедры, ответственной за реализацию дисциплины)

« 28 » марта 2024 г., протокол № 7

Врио заведующего кафедрой

к.т.н., доцент

(ученая степень, ученое звание)

(подпись)

М.В. Терехов

(И.О. Фамилия)

**Согласовано:**

Врио заведующего выпускающей кафедрой

«Компьютерные технологии и системы»

(наименование выпускающей кафедры)

к.т.н., доцент

(ученая степень, ученое звание)

(подпись)

Терехов М.В.

(И.О. Фамилия)

© Кузьменко А.А., 2024

© ФГБОУ ВО «Брянский государственный  
технический университет», 2024

## СОДЕРЖАНИЕ

ПРЕДИСЛОВИЕ.....	5
1. ЦЕЛЬ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ .....	5
2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ ФГОС .....	5
3. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОСВОЕНИЯ ДИСЦИПЛИНЫ .....	5
4. ОБЪЕМ ДИСЦИПЛИНЫ И ВИДЫ УЧЕБНОЙ РАБОТЫ .....	6
5. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ .....	7
5.1. Структура дисциплины.....	7
5.2. Распределение формируемых компетенций по разделам (темам) дисциплины.....	8
5.3. Лекции .....	9
5.4. Лабораторные работы .....	12
5.5. Практические занятия .....	13
5.6. Самостоятельная работа обучающихся .....	15
5.7. Организация текущего контроля успеваемости и промежуточной аттестации обучающихся .....	18
6. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ .....	18
7. РЕАЛИЗАЦИЯ ДИСЦИПЛИНЫ ПРИ ИСПОЛЬЗОВАНИИ ТЕХНОЛОГИЙ ЭЛЕКТРОННОГО ОБУЧЕНИЯ И (ИЛИ) ДИСТАНЦИОННЫХ ОБРАЗОВАТЕЛЬНЫХ ТЕХНОЛОГИЙ.....	19
8. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ .....	20
8.1. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся .....	20
8.2. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины .....	20
8.3. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», используемых при изучении дисциплины .....	21
8.4. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень программного обеспечения и (или) информационных справочных систем .....	21
9. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ .....	21
10. ОСОБЕННОСТИ ОРГАНИЗАЦИИ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ДЛЯ ИНВАЛИДОВ И ЛИЦ С ОГРАНИЧЕННЫМИ ВОЗМОЖНОСТЯМИ ЗДОРОВЬЯ.....	22

11. МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЕ .....	23
11.1. Методические материалы для педагогических работников .....	23
11.2. Методические материалы для обучающихся .....	26
12. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЕ .....	27
12.1. Виды и средства оценивания результатов освоения дисциплины .....	27
12.2. Шкала оценивания при текущем контроле успеваемости .....	27
12.3. Шкала оценивания при промежуточной аттестации обучающихся .....	29
12.4. Оценивание окончательных результатов обучения по дисциплине .....	32
12.5. Характеристика результатов обучения .....	32
12.6. Контрольно-измерительные материалы для текущего контроля успеваемости и промежуточной аттестации обучающихся .....	33
13. ВОСПИТАТЕЛЬНАЯ РАБОТА .....	33

## ПРЕДИСЛОВИЕ

Учебная дисциплина «Методы анализа данных и естественно-языковых текстов» (далее – дисциплина) ориентирована на формирование у обучающихся компетенций в рамках основной профессиональной образовательной программы высшего образования (ОПОП ВО) по направлению подготовки 09.03.02 Информационные системы и технологии, профиль «Искусственный интеллект и обработка больших данных».

### 1. ЦЕЛЬ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

**Цель** освоения дисциплины – получение навыков, методов и подходов, позволяющих разрабатывать и применять автоматизированные технологии для анализа данных и обработки естественно-языковых текстов.

**Задачи** дисциплины:

- Быть ознакомлен с технологиями и алгоритмами обработки естественно-языковых текстов
- Разрабатывать и применять автоматизированные технологии обработки естественно-языковых текстов.
- Использовать и применять навыки создания программного обеспечения, предназначенного для обработки текстов, написанных на естественном языке

### 2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ ФГОС

Дисциплина входит в вариативная часть, формируемую участниками образовательных отношений учебного плана образовательной программы и реализуется на 4 курсе(-ах) в 7 семестре(-ах).

Предварительно изучаются дисциплины: *«Объектно-ориентированное программирование»*, *«Методы и средства проектирования информационных систем и технологий»*, *«Системы искусственного интеллекта»*.

Параллельно изучаются дисциплины: *«Интеллектуальные системы и технологии»*, *«Моделирование систем искусственного интеллекта»*, *«Методы и средства проектирования информационных систем и технологий»*.

Базируются на изучении дисциплины: *«Методы интеллектуального анализа больших данных»*.

### 3. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Изучение дисциплины направлено на формирование у обучающихся компетенций ПК-1, представленных в таблице 1.

Таблица 1 – Требования к результатам освоения учебной дисциплины

Код и наименование компетенции	Индикаторы компетенций	В результате изучения учебной дисциплины обучающиеся должны:		
		знать	уметь	владеть

ПК-1. Способен выполнять работы и управлять работами по созданию (модификации) и сопровождению ИС, автоматизирующих задачи организационного управления и бизнес- процессы	ПК 1.1. Осуществлять инженерно-технологическую поддержку планирования управления требованиями	Возможности планирования в системах обработки естественно-языковых текстов	Анализировать входные данные	Источниками информации, необходимой для анализа в системах обработки естественно-языковых текстов
	ПК.1.2. Проектировать ИС и осуществлять разработку дизайна ИС	Современный отечественный и зарубежный опыт эксплуатации систем обработки естественно-языковых текстов	Применять инструменты и методы проектирования систем обработки естественно-языковых текстов	Устранение обнаруженных несоответствий при проектировании систем обработки естественно-языковых текстов
	ПК 1.3. Разрабатывать технологии интеграции ИС с существующими ИС заказчика	Форматы обмена данными	Осуществлять коммуникации	Знаниями в области современных стандартов информационного взаимодействия систем

#### 4. ОБЪЕМ ДИСЦИПЛИНЫ И ВИДЫ УЧЕБНОЙ РАБОТЫ

Общая трудоемкость дисциплины составляет 5 зачетных единиц(ы) (180 академических часа(-ов)). Распределение трудоемкости дисциплины по видам учебной работы и семестрам представлено в таблице 2.

Таблица 2 – Распределение трудоемкости дисциплины по видам учебной работы и семестрам

[illegible]

Виды учебной работы в соответствии с учебным планом образовательной программы	Трудоемкость, час.												
	Всего	Семестр											
		1	2	3	4	5	6	7	8	9	А	В	С
в том числе:													
3.1. Экзамен, семестр								7					
3.2. Зачет, семестр								-					
3.3. Зачет с оценкой, семестр								-					
3.4. Курсовой проект (контроль), семестр								-					
3.5. Курсовая работа (контроль), семестр								-					
3.6. Контрольная работа (контроль), семестр								-					
<b>Общая трудоемкость (5 з.е.)</b>								180					

## 5. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

### 5.1. Структура дисциплины

Структура дисциплины представлена в виде тематического плана в таблице 3.

Таблица 3 – Тематический план дисциплины

Наименование раздела (темы) дисциплины	Трудоемкость, час.				
	Всего	Лекции	Лабораторные работы	Практические занятия	Самостоятельная работа
<b>Раздел 1. Введение в обработку ЕЯ</b>	<b>16</b>	<b>2</b>		<b>1</b>	<b>13</b>
Тема 1. Роль и использование ЕЯ в компьютерных системах.	7	1	-	-	6
Тема 2. Обзор задач и классификация методов ОЕЯ.	10	1	-	1	7
<b>Раздел 2. Лингвистические ресурсы для ЕЯ систем</b>	<b>25</b>	<b>3</b>	<b>4</b>	<b>3</b>	<b>15</b>
Тема 3. Корпусы текстов, словари, тезаурусы, лингвистические онтологии, WordNet.	13	2	2	2	8
Тема 4. Википедия, как источник для построения лингвистического обеспечения.	11	1	2	1	7
<b>Раздел 3. Технологический подходы к обработки ЕЯ</b>	<b>25</b>	<b>4</b>	<b>2</b>	<b>3</b>	<b>16</b>

Наименование раздела (темы) дисциплины	Трудоемкость, час.				
	Всего	Лекции	Лабораторные работы	Практические занятия	Самостоятельная работа
Тема 5. Графематический анализ (токенизация), морфологический анализ, стемминг, лемматизация, частеречная разметка, снятие морфологической многозначности, синтаксический анализ: грамматики ЕЯ и формальные грамматики, фрагментация предложений, алгоритмы синтаксического разбора.	13	2	2	2	8
Тема 6. Латентно-семантический анализ (LSA), латентное размещение Дирхле (LDA), семантическая близость понятий, разрешение лексической многозначности терминов, обработка ЕЯ с использованием нейронных сетей.	11	2	-	1	8
<b>Раздел 4. Обзор программных инструментов для обработки ЕЯ</b>	<b>19</b>	<b>2</b>	<b>2</b>	<b>2</b>	<b>13</b>
Тема 7. Решения корпоративного уровня и свободно-доступные для ОЕЯ: законченные решения, Фреймворки, библиотеки, словари.	8	1	-	-	7
Тема 8. Системы полнотекстового поиска.	11	1	2	2	6
<b>Раздел 5. Информационный поиск</b>	<b>21</b>	<b>2</b>	<b>2</b>	<b>3</b>	<b>14</b>
Тема 9. Введение в информационный поиск.	9	1	1	1	6
Тема 10. Архитектура поисковых систем.	11	1	1	2	8
<b>Раздел 6. Прочие задачи, связанные с обработкой ЕЯ</b>	<b>29</b>	<b>3</b>	<b>6</b>	<b>4</b>	<b>16</b>
Тема 11. Анализ тональности, классификация, кластеризация текстов, выявление дубликатов.	13	1	2	1	8
Тема 12. Генерация текста, машинный перевод, чат-боты.	18	2	4	3	8
<b>Итого</b>	<b>135</b>	<b>16</b>	<b>16</b>	<b>16</b>	<b>87</b>

## 5.2. Распределение формируемых компетенций по разделам (темам) дисциплины

Распределение формируемых компетенций по разделам дисциплины представлено в таблице 4.



Таблица 4 – Формирование компетенций по разделам дисциплины

Наименование раздела (темы) дисциплины	Код компетенции
	ПК-3
Тема 1. Роль и использование ЕЯ в компьютерных системах.	+
Тема 2. Обзор задач и классификация методов ОЕЯ.	+
Тема 3. Корпусы текстов, словари, тезаурусы, лингвистические онтологии, WordNet.	+
Тема 4. Википедия, как источник для построения лингвистического обеспечения.	+
Тема 5. Графематический анализ (токенизация), морфологический анализ, стемминг, лемматизация, частеречная разметка, снятие морфологической многозначности, синтаксический анализ: грамматики ЕЯ и формальные грамматики, фрагментация предложений, алгоритмы синтаксического разбора.	+
Тема 6. Латентно-семантический анализ (LSA), латентное размещение Дихле (LDA), семантическая близость понятий, разрешение лексической многозначности терминов, обработка ЕЯ с использованием нейронных сетей.	+
Тема 7. Решения корпоративного уровня и свободно-доступные для ОЕЯ: законченные решения, Фреймворки, библиотеки, словари.	+
Тема 8. Системы полнотекстового поиска.	+
Тема 9. Введение в информационный поиск.	+
Тема 10. Архитектура поисковых систем.	+
Тема 11. Анализ тональности, классификация, кластеризация текстов, выявление дубликатов.	+
Тема 12. Генерация текста, машинный перевод, чат-боты.	+

### 5.3. Лекции

Перечень занятий лекционного типа, их содержание и трудоемкость представлены в таблице 5.

Таблица 5 – Тематика и содержание лекций

Наименование темы дисциплины	Тема лекции	Содержание лекции	Трудоемкость, час.
------------------------------	-------------	-------------------	--------------------

Наименование темы дисциплины	Тема лекции	Содержание лекции	Трудоемкость, час.
Тема 1. Роль и использование ЕЯ в компьютерных системах.	Роль и использование ЕЯ в компьютерных системах.	1. Введение в обработку ЕЯ. 2. Роль и использование ЕЯ в компьютерных системах.	1
Тема 2. Обзор задач и классификация методов ОЕЯ.	Обзор задач и классификация методов ОЕЯ.	1. Обзор задач и классификация методов ОЕЯ. 2. Проблемы обработки ЕЯ. 3. Методы оценки качества решения задач обработки ЕЯ. 4. Языковые модели, N-граммы, цепи Маркова.	1
Тема 3. Корпусы текстов, словари, тезаурусы, лингвистические онтологии, WordNet.	Корпусы текстов, словари, тезаурусы, лингвистические онтологии, WordNet.	1. Корпусы и коллекции текстов. 2. Морфологические словари и тезаурусы. 3. Лингвистические онтологии. 4. Лексическая база данных WordNet	2
Тема 4. Википедия, как источник для построения лингвистического обеспечения.	Википедия, как источник для построения лингвистического обеспечения.	Википедия, как источник для построения лингвистического обеспечения.	1
Тема 5. Графематический анализ (токенизация), морфологический анализ, стемминг, лемматизация, частеречная разметка, снятие морфологической многозначности, синтаксический анализ: грамматики ЕЯ и формальные грамматики, фрагментация предложений, алгоритмы синтаксического разбора.	Графематический анализ (токенизация), морфологический анализ, стемминг, лемматизация, частеречная разметка, снятие морфологической многозначности, синтаксический анализ: грамматики ЕЯ и формальные грамматики, фрагментация предложений, алгоритмы синтаксического разбора.	1. Графематический анализ (токенизация). 2. Морфологический анализ: особенности морфологии русского и английского языка, алгоритмы стемминга. 3. Морфологический анализ: лемматизация, частеречная разметка, снятие морфологической многозначности. 4. Синтаксический анализ: грамматики ЕЯ и формальные грамматики, фрагментация предложений, алгоритмы синтаксического разбора. 5. Модели представления семантики. 6. Методы семантического анализа.	2

Наименование темы дисциплины	Тема лекции	Содержание лекции	Трудоемкость, час.
Тема 6. Латентно-семантический анализ (LSA), латентное размещение Дирхле (LDA), семантическая близость понятий, разрешение лексической многозначности терминов, обработка ЕЯ с использованием нейронных сетей.	Латентно-семантический анализ (LSA), латентное размещение Дирхле (LDA), семантическая близость понятий, разрешение лексической многозначности терминов, обработка ЕЯ с использованием нейронных сетей.	1. Латентно-семантический анализ (LSA), латентное размещение Дирхле (LDA) и их модификации. 2. Семантическая близость понятий. 3. Методы разрешения лексической многозначности терминов. 4. Обработка ЕЯ с использованием нейронных сетей.	2
Тема 7. Решения корпоративного уровня и свободно-доступные для ОЕЯ: законченные решения, Фреймворки, библиотеки, словари.	Решения корпоративного уровня и свободно-доступные для ОЕЯ: законченные решения, Фреймворки, библиотеки, словари.	1. Решения корпоративного уровня для ОЕЯ. 2. Свободно-доступные решения для ОЕЯ: фреймворки, библиотеки, словари.	1
Тема 8. Системы полнотекстового поиска.	Системы полнотекстового поиска.	Обзор открытых систем полнотекстового поиска.	1
Тема 9. Введение в информационный поиск.	Введение в информационный поиск.	1. Введение в информационный поиск. 2. Булев и ранжированный поиск. 3. Построение и сжатие индекса. 4. Модель векторного пространства. 5. Ранжирование документов.	1
Тема 10. Архитектура поисковых систем.	Архитектура поисковых систем.	1. Архитектура поисковых систем. 2. Алгоритмы PageRank и HITS. 3. Подходы к извлечению отношений из текстов. 4. Автоматическое реферирование и аннотирование текстов.	1
Тема 11. Анализ тональности, классификация, кластеризация текстов, выявление дубликатов.	Анализ тональности, классификация, кластеризация текстов, выявление дубликатов.	1. Анализ тональности текстов. 2. Классификация и кластеризация текстов, выявление дубликатов. 3. Подходы и особенности анализа современ-	1

Наименование темы дисциплины	Тема лекции	Содержание лекции	Трудоемкость, час.
		ных социальных медиа в интернете.	
Тема 12. Генерация текста, машинный перевод, чат-боты.	Генерация текста, машинный перевод, чат-боты.	1. Генерация текста. 2. Машинный перевод. 3. Разработка чат-ботов и разговорных интерфейсов.	2
<b>Итого</b>	—	—	16

#### 5.4. Лабораторные работы

Лабораторные работы по дисциплине предусмотрены учебным планом образовательной программы (таблица 6).

Наименование темы дисциплины	Тема лабораторной работы	Трудоемкость, час.
Тема 3. Корпусы текстов, словари, тезаурусы, лингвистические онтологии, Word-Net.	Построение статистической языковой модели на основе N- грамм по коллекции текстов.	2
Тема 4. Википедия, как источник для построения лингвистического обеспечения.	Разрешение лексической многозначности терминов.	2
Тема 5. Графематический анализ (токенизация), морфологический анализ, стемминг, лемматизация, частеречная разметка, снятие морфологической многозначности, синтаксический анализ: грамматики ЕЯ и формальные грамматики, фрагментация предложений, алгоритмы синтаксического разбора.	Обработка естественного языка на базе набора библиотек NLTK.	2
Тема 8. Системы полнотекстового поиска.	Реализация поиска по коллекции документов с использованием библиотеки Apache Lucene.	2
Тема 9. Введение в информационный поиск.	Распознавание именованных сущностей в текстах.	1
Тема 10. Архитектура поисковых систем.	Извлечение отношений из текста с помощью шаблонов.	1

Наименование темы дисциплины	Тема лабораторной работы	Трудоемкость, час.
Тема 11. Анализ тональности, классификация, кластеризация текстов, выявление дубликатов.	Реализация тематической классификации текстов.	1
	Создание собственного корпуса.	
	Предварительная обработка и преобразование корпуса. Конвейеры векторизации и преобразования.	1
Тема 12. Генерация текста, машинный перевод, чат-боты.	Кластеризация для выявления сходств в тексте.	1
	Контекстно-зависимый анализ текста.	1
	Визуализация текста.	1
	Графовые методы анализа текста.	1
<b>Итого</b>	–	<b>16</b>

### 5.5. Практические занятия

Практические занятия по дисциплине предусмотрены / не предусмотрены учебным планом образовательной программы.

Перечень практических занятий, их содержание и трудоемкость представлены в таблице 7.

Таблица 7 – Тематика и содержание практических занятий

Наименование темы дисциплины	Тема практического занятия	Содержание практического занятия	Трудоемкость, час.
Тема 1. Роль и использование ЕЯ в компьютерных системах.	-	-	-
Тема 2. Обзор задач и классификация методов ОЕЯ.	Изучение возможностей по обработке ЕЯ продуктов от Microsoft	Изучение возможностей по обработке ЕЯ продуктов от Microsoft	1
Тема 3. Корпусы текстов, словари, тезаурусы, лингвистические онтологии, WordNet.	N-граммы.	N-граммы.	1
	Работа с корпусами текстов русского языка.	Работа с корпусами текстов русского языка.	1
Тема 4. Википедия, как источник для построения лингвистического обеспечения.	Работа с пакетом Apache OpenNLP.	Работа с пакетом Apache OpenNLP.	1
Тема 5. Графематический анализ (токенизация), морфологический анализ, стемминг, лемматизация, частеречная разметка, снятие морфологической многозначности,	Разработка графематического анализатора текста.	Разработка графематического анализатора текста.	1
	Разработка морфологического анализатора на основе стемминга.	Разработка морфологического анализатора на основе стемминга.	1

Наименование темы дисциплины	Тема практического занятия	Содержание практического занятия	Трудоемкость, час.
синтаксический анализ: грамматики ЕЯ и формальные грамматики, фрагментация предложений, алгоритмы синтаксического разбора.			
Тема 6. Латентно-семантический анализ (LSA), латентное размещение Дирхле (LDA), семантическая близость понятий, разрешение лексической многозначности терминов, обработка ЕЯ с использованием нейронных сетей.	Разработка синтаксического анализатора для русского языка.	Разработка синтаксического анализатора для русского языка.	1
Тема 7. Решения корпоративного уровня и свободно-доступные для ОЕЯ: законченные решения, Фреймворки, библиотеки, словари.	-	-	-
Тема 8. Системы полнотекстового поиска.	Изучение подсистем полнотекстового поиска в базах данных.	Изучение подсистем полнотекстового поиска в базах данных.	1
	Работа с поисковыми индексами.	Работа с поисковыми индексами.	1
Тема 9. Введение в информационный поиск.	Сравнение различных функций взвешивания терминов.	Сравнение различных функций взвешивания терминов.	2
Тема 10. Архитектура поисковых систем.	Сравнение методов извлечения терминов из текстов.	Сравнение методов извлечения терминов из текстов.	1
	Извлечение отношении методом бутстрепинга.	Извлечение отношении методом бутстрепинга.	1

Наименование темы дисциплины	Тема практического занятия	Содержание практического занятия	Трудоемкость, час.
Тема 11. Анализ тональности, классификация, кластеризация текстов, выявление дубликатов.	Классификация текстов методом SVM.	Классификация текстов методом SVM.	1
Тема 12. Генерация текста, машинный перевод, чат-боты.	Построение системы машинного перевода на основе параллельного корпуса.	Построение системы машинного перевода на основе параллельного корпуса.	1
	Знакомство с инструментами разработки чат-ботов	Знакомство с инструментами разработки чат-ботов	1
	Знакомство с хранилищем чат-ботов	Знакомство с хранилищем чат-ботов	1
<b>Итого</b>	—	-	<b>16</b>

### 5.6. Самостоятельная работа обучающихся

Вопросы, выносимые на самостоятельное изучение, представлены в таблице 8.

Таблица 8 – Вопросы для самостоятельного изучения дисциплины

Наименование темы дисциплины	Вопросы для самостоятельного изучения темы
Тема 1. Роль и использование ЕЯ в компьютерных системах.	Использование и актуальность обработки ЕЯ в компьютерных системах.
Тема 2. Обзор задач и классификация методов ОЕЯ.	Скрытые модели в обработке ЕЯ.
Тема 3. Корпусы текстов, словари, тезаурусы, лингвистические онтологии, WordNet.	1. Поиск и изучение свободно доступных лингвистических ресурсов. 2. Цепи Маркова.
Тема 4. Википедия, как источник для построения лингвистического обеспечения.	Лингвистические онтологии.
Тема 5. Графематический анализ (токенизация), морфологический анализ, стемминг, лемматизация, частеречная разметка, снятие морфологической многозначности, синтаксический анализ: грамматики ЕЯ и формальные грамматики, фрагментация предложений, алгоритмы синтаксического разбора.	1. Алгоритмы синтаксического разбора (парсинга). 2. Модели представления семантики текста.
Тема 6. Латентно-семантический анализ (LSA), латентное размещение Дирхле (LDA), семантическая близость понятий, разрешение лексической многозначности терминов,	Использование средств ОЕЯ в продуктах от Microsoft.

Наименование темы дисциплины	Вопросы для самостоятельного изучения темы
обработка ЕЯ с использованием нейронных сетей.	
Тема 7. Решения корпоративного уровня и свободно-доступные для ОЕЯ: законченные решения, Фреймворки, библиотеки, словари.	Изучение различных библиотек, фреймворков и ресурсов для ОЕЯ
Тема 8. Системы полнотекстового поиска.	Библиотека Apache Lucene.
Тема 9. Введение в информационный поиск.	1. Параметрические и зонные индексы. 2. Компоненты поисковых систем. 3. Вероятностная модель информационного поиска. 4. Обход и индексирование веба
Тема 10. Архитектура поисковых систем.	1. Разрешение анафоры с помощью машинного обучения. 2. Постобработка реферата при автоматическом реферировании.
Тема 11. Анализ тональности, классификация, кластеризация текстов, выявление дубликатов.	Метод опорных векторов.
Тема 12. Генерация текста, машинный перевод, чат-боты.	1. Платформы и мессенджеры для создания чат-ботов. 2. Вопросно-ответные системы.

В процессе самостоятельной работы обучающиеся должны принимать решение по рассматриваемой проблеме с минимальным участием педагогического работника. Для решения поставленных задач может использоваться дополнительная литература и источники в информационно-коммуникационной сети «Интернет». Для закрепления пройденного материала педагогическим работником могут выдаваться домашние задания.

В таблице 9 указаны виды самостоятельной работы, выполняемые обучающимися при изучении соответствующих тем дисциплины.

Таблица 9 – Виды самостоятельной работы

Наименование темы дисциплины	Виды самостоятельной работы
Тема 1. Роль и использование ЕЯ в компьютерных системах.	Проработка лекционного материала Изучение рекомендуемой литературы Подготовка докладов
Тема 2. Обзор задач и классификация методов ОЕЯ.	Проработка лекционного материала Изучение рекомендуемой литературы Выполнение практического задания Подготовка докладов
Тема 3. Корпусы текстов, словари, тезаурусы, лингвистические онтологии, WordNet.	Проработка лекционного материала Изучение рекомендуемой литературы Выполнение практического задания Выполнение лабораторной работы Подготовка докладов
Тема 4. Википедия, как источник для построения лингвистического обеспечения.	Проработка лекционного материала Изучение рекомендуемой литературы Выполнение практического задания Выполнение лабораторной работы



Наименование темы дисциплины	Виды самостоятельной работы
	Подготовка докладов
Тема 5. Графематический анализ (токенизация), морфологический анализ, стемминг, лемматизация, частеречная разметка, снятие морфологической многозначности, синтаксический анализ: грамматики ЕЯ и формальные грамматики, фрагментация предложений, алгоритмы синтаксического разбора.	Проработка лекционного материала Изучение рекомендуемой литературы Выполнение практического задания Выполнение лабораторной работы Подготовка докладов
Тема 6. Латентно-семантический анализ (LSA), латентное размещение Дихле (LDA), семантическая близость понятий, разрешение лексической многозначности терминов, обработка ЕЯ с использованием нейронных сетей.	Проработка лекционного материала Изучение рекомендуемой литературы Выполнение практического задания Подготовка докладов
Тема 7. Решения корпоративного уровня и свободно-доступные для ОЕЯ: законченные решения, Фреймворки, библиотеки, словари.	Проработка лекционного материала Изучение рекомендуемой литературы Подготовка докладов
Тема 8. Системы полнотекстового поиска.	Проработка лекционного материала Изучение рекомендуемой литературы Выполнение практического задания Выполнение лабораторной работы Подготовка докладов
Тема 9. Введение в информационный поиск.	Проработка лекционного материала Изучение рекомендуемой литературы Выполнение практического задания Выполнение лабораторной работы Подготовка докладов
Тема 10. Архитектура поисковых систем.	Проработка лекционного материала Изучение рекомендуемой литературы Выполнение практического задания Выполнение лабораторной работы Подготовка докладов
Тема 11. Анализ тональности, классификация, кластеризация текстов, выявление дубликатов.	Проработка лекционного материала Изучение рекомендуемой литературы Выполнение практического задания Выполнение лабораторной работы Подготовка докладов
Тема 12. Генерация текста, машинный перевод, чат-боты.	Проработка лекционного материала Изучение рекомендуемой литературы Выполнение практического задания Выполнение лабораторной работы Подготовка докладов

Учебным планом в рамках дисциплины не предусмотрено курсовое проектирование.

Выполнение курсовое проектирование осуществляется в соответствии с методическими указаниями, содержащимися в соответствующем разделе электронного курса «Методы анализа данных и естественно-языковых текстов» информационно-образовательной среды БГТУ (<http://edu.tu-bryansk.ru>).

### 5.7. Организация текущего контроля успеваемости и промежуточной аттестации обучающихся

Текущий контроль успеваемости обеспечивает оценивание хода освоения дисциплины. Формы контрольно-оценочных мероприятий, проводимых в рамках текущего контроля успеваемости, представлены в таблице 10.

Таблица 10 – Формы и периодичность текущего контроля успеваемости

Вид учебной работы	Форма текущего контроля успеваемости	Периодичность осуществления
Практические занятия / Лабораторные работы	Устный экспресс-опрос, экспресс-тестирование.	На каждом занятии
Самостоятельная работа обучающихся	- устная (устный опрос, защита письменной работы, доклада по результатам самостоятельной работы, рефератов и т.д.); - письменная (письменный опрос, выполнение конспектов, глоссариев, курсового проекта / курсовой работы и т.д.); - тестовая (бланочное или компьютерное тестирование)	В течение семестра

Оценивание промежуточных и окончательных результатов обучения по дисциплине (промежуточная аттестация обучающихся) осуществляется в форме зачета и экзамена, проводимого в устной / письменной форме. Аттестационное испытание может включать в себя прохождение теста с использованием технологии компьютерного тестирования. Для уточнения оценки экзаменатор может проводить короткий опрос-собеседование с обучающимся и (или) выдавать ему дополнительные задания.

## 6. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ

В ходе освоения дисциплины применяются следующие образовательные технологии: личностно-ориентированные, активизации деятельности обучающихся, интеллектуальной направленности, проблемного обучения, диалоговые и профессионально-ориентированные (таблица 11).

Таблица 11 – Образовательные технологии, применяемые в ходе преподавания дисциплины

Вид учебной работы	Применяемые образовательные технологии
Лекции	Проблемная лекция. Лекция-визуализация. Лекция-беседа.

Вид учебной работы	Применяемые образовательные технологии
	Лекция-дискуссия.
Практические занятия / Лабораторные работы	Решение практических задач. Тестирование.
Самостоятельная работа обучающихся	Проработка лекционного материала. Изучение рекомендуемой литературы. Подготовка к дискуссии. Выполнение практического задания / лабораторной работы. . Выполнение курсовой работы (курсового проекта) Подготовка докладов, рефератов Подготовка к лекциям. Подготовка к практическим занятиям. Изучение дополнительной литературы и самостоятельное формирование конспекта. Подготовка к экзамену/зачету
Консультации	Концентрация внимания на отдельных вопросах. Личностно-ориентированный подход. Диалог.
Промежуточная аттестация обучающихся	Зачет и экзамен (в устной или письменной форме).

## 7. РЕАЛИЗАЦИЯ ДИСЦИПЛИНЫ ПРИ ИСПОЛЬЗОВАНИИ ТЕХНОЛОГИЙ ЭЛЕКТРОННОГО ОБУЧЕНИЯ И (ИЛИ) ДИСТАНЦИОННЫХ ОБРАЗОВАТЕЛЬНЫХ ТЕХНОЛОГИЙ

В электронной информационно-образовательной среде БГТУ размещается электронный курс дисциплины, включающий в себя:

- сведения об авторе курса;
- краткое описание курса;
- рабочую программу дисциплины;
- полный перечень тем дисциплины;
- презентационные материалы для проведения занятий лекционного типа;
- лекции/краткий конспект лекций по каждой теме;
- методические указания по выполнению каждого практического задания;
- методические указания для выполнения курсового проекта;
- материалы и тестовые задания для текущего контроля успеваемости и промежуточной аттестации обучающихся.

Наименование электронного курса в электронной информационно-образовательной среде БГТУ — «Методы анализа данных и естественно-языковых текстов – автор Кузьменко А.А. по направлению подготовки 09.03.02 Информационные системы и технологии, профиль «Искусственный интеллект и обработка больших данных», форма обучения – очная.

Электронный курс предназначен для обеспечения обучающихся всеми необходимыми учебно-методическими материалами, а также проведения контрольно-оценочных мероприятий в процессе обучения. При необходимости осуществляется файловый обмен отчетами о выполнении обучающимися самостоятельной работы.

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

### **8.1. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся**

1. 1. Автоматическая обработка текстов на естественном языке и анализ данных: учеб. пособие / Большакова Е.И., Воронцов К.В., Ефремова Н.Э., Клышинский Э.С., Лукашевич Н.В., Сапин А.С. — М.: Изд-во НИУ ВШЭ, 2019. — 269 с

2. Кристофер Д. Маннинг, Прабхакар Рагхаван, Хайнрих Шютце. Введение в информационный поиск. : Пер. с англ. - М. : ООО "И.Д. Вильямс", 2021. - 428 е. : ил. [Электронный ресурс].

### **8.2. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины**

#### ***а) основная литература***

1. Боярский К.К. Введение в компьютерную лингвистику [Электронный ресурс] . учебное пособие / КК. Боярский. — Электрон. текстовые данные. — СПб. : Университет ИТМО, 2013. 73 с. 2227-8397.— Режим доступа: <http://www.iprbookshop.ru/71485.html>

2. Ингерсолл, Г.С. Обработка неструктурированных текстов. Поиск, организация и манипулирование [Электронный ресурс] / Г.С. Ингерсолл, Т.С. Мортон, Э.Л. Фэррис. — Электрон. дан. — Москва : ДМК Пресс, 2019. — 414 с. — Режим доступа: <https://e.lanbook.com/book/73069>

3. Риз, Р. Обработка естественного языка на Java [Электронный ресурс]: учеб. пособие — Электрон. дан. — Москва : ДМК Пресс, 2019. — 264 с. — Режим доступа: <https://e.lanbook.com/book/93272>

4. Павлов С.Н. Системы искусственного интеллекта. Часть 1 [Электронный ресурс] : учебное пособие / С.Н. Павлов. — Электрон. текстовые данные. — Томск: Томский государственный университет систем управления и радиоэлектроники, Эль Контент, 2021. 176 с. 978-5-4332-0013-5. — Режим доступа: <http://www.iprbookshop.ru/13974.html>

#### ***б) дополнительная литература***

1. Лисенкова И.М. Современный русский язык. Лексическая семантика. Синтагматика лексических единиц [Электронный ресурс] : учебное пособие / И.М. Лисенкова. — Электрон. текстовые данные. — М. : Русайнс, 2020 — 134

2. Бенгфорг Бенжамин. Прикладной анализ текстовых данных на Python. Машинное обучение и создание приложений обработки естественного языка [Электронный ресурс]: -СПб.: Питер, 2019- 368с.

3. Срини Джанарсанам. Практическое руководство по разработке чат-интерфейсов: пер. с англ. [Электронный ресурс]:- 2019-340с.

Батура, Т. В. Методы и системы автоматического реферирования текстов : монография / Т. В. Батура, А. М. Бакиева ; Ин-т систем информатики им. А. П. Ершова СО РАН. — Новосибирск : ИПЦ НГУ, 2019. — 110 с.2. ....

#### **б) справочная литература**

не предусмотрена

### **8.3. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», используемых при изучении дисциплины**

- 1). Сайт научной библиотеки БГТУ (<https://libri.tu-bryansk.ru>)
- 2). Электронно-библиотечная система «Лань» (<https://e.lanbook.com>).
- 3). Электронно-библиотечная система «IPRbooks» (<http://www.iprbookshop.ru>).
- 4). Электронно-библиотечная система ИД «Гребенников» (<https://grebennikon.ru>).
- 5). Единое окно доступа к информационным ресурсам (<http://window.edu.ru>).
- 6). Национальная электронная библиотека (<http://www.elibrary.ru>).
- 7). Федеральное хранилище «Единая коллекция цифровых образовательных ресурсов» (<http://school-collection.edu.ru>).
- 8). Федеральный Интернет-портал «Российское образование» (<http://www.edu.ru>).

### **8.4. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень программного обеспечения и (или) информационных справочных систем**

- 1). Операционная система класса Microsoft Windows.
- 2). Пакет офисных прикладных программ OpenOffice или Microsoft Office.
- 3). Справочная правовая система «КонсультантПлюс»

## **9. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ**

Для обеспечения обучения необходима следующая материально-техническая база:

- аудитория для проведения лекционных занятий и организации защиты курсовых работ/курсовых проектов, оборудованная персональными компьютерами, мультимедийным компьютерным проектором, средства звуковоспроизведения (по возможности), проекционным экраном, наличием доступа в информационно-коммуникационную сеть Интернет;
- компьютерный класс для проведения лабораторных работ с установленным комплектом программного обеспечения и доступом в информационно-коммуникационную сеть интернет, оборудованный мультимедийным ком-

пьютерным проектором, средства звуковоспроизведения (по возможности), проекционным экраном / лаборатория со специализированным оборудованием для проведения лабораторных работ;

- учебная аудитория, оснащенная комплектом мебели и доской, для проведения консультаций, зачета, зачета с оценкой, экзамена;
- компьютерные классы с постоянным доступом к информационно-телекоммуникационной сети «Интернет», а также читальные залы научной библиотеки БГТУ для самостоятельной работы обучающихся.

## **10. ОСОБЕННОСТИ ОРГАНИЗАЦИИ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ДЛЯ ИНВАЛИДОВ И ЛИЦ С ОГРАНИЧЕННЫМИ ВОЗМОЖНОСТЯМИ ЗДОРОВЬЯ**

Изучение дисциплины инвалидами и лицами с ограниченными возможностями здоровья организуется с учетом особенностей психофизического развития, индивидуальных возможностей и состояния здоровья таких обучающихся.

При проведении учебных занятий обеспечивается соблюдение следующих требований:

- учебные занятия проводятся для инвалидов и лиц с ограниченными возможностями здоровья в одной аудитории совместно с обучающимися, не имеющими ограниченных возможностей здоровья, если это не создает трудностей для обучающихся в ходе учебных занятий;
- присутствие ассистента из числа работников БГТУ или привлеченных лиц, оказывающего обучающимся необходимую техническую помощь с учетом их индивидуальных особенностей (занять рабочее место, передвигаться, прочесть и оформить задание, общаться с педагогическим работником и т. п.);
- обучающиеся с учетом их индивидуальных особенностей могут пользоваться необходимыми им техническими средствами;
- материально-технические условия должны обеспечивать возможность беспрепятственного доступа обучающихся в аудитории, туалетные и другие помещения, а также их пребывания в указанных помещениях (наличие пандусов, поручней, расширенных дверных проемов, лифтов, при отсутствии лифтов аудитория должна располагаться на первом этаже; наличие специальных кресел и других приспособлений).

Университетом созданы специальные условия для получения высшего образования обучающимися с ОВЗ:

1) для лиц с ограниченными возможностями здоровья по зрению:

- наличие альтернативной версии официального сайта организации в сети "Интернет" для слабовидящих;
- размещение в доступных для обучающихся, являющихся слепыми или слабовидящими, местах и в адаптированной форме (с учетом их особых потребностей) справочной информации о расписании учебных занятий (информация должна быть выполнена крупным рельефно-контрастным шрифтом (на бе-

лом или желтом фоне) и продублирована шрифтом Брайля);

- присутствие ассистента, оказывающего обучающемуся необходимую помощь;
- обеспечение выпуска альтернативных форматов печатных материалов (крупный шрифт или аудиофайлы);
- обеспечение доступа обучающегося, являющегося слепым и использующего собаку-проводника, к зданию организации;

2) для лиц с ограниченными возможностями здоровья по слуху:

- дублирование звуковой справочной информации о расписании учебных занятий визуальной (установка мониторов с возможностью трансляции субтитров (мониторы, их размеры и количество необходимо определять с учетом размеров помещения);
- обеспечение надлежащими звуковыми средствами воспроизведения информации;

3) для лиц с ограниченными возможностями здоровья, имеющих нарушения опорно-двигательного аппарата, материально-технические условия должны обеспечивать возможность беспрепятственного доступа обучающихся в учебные помещения, столовые, туалетные и другие помещения Университета, а также пребывания в указанных помещениях (наличие пандусов, поручней, расширенных дверных проемов, лифтов, локальное понижение стоек-барьеров; наличие специальных кресел и других приспособлений).

## **11. МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЕ**

### **11.1. Методические материалы для педагогических работников**

Основными формами организации обучения по дисциплине являются лекции, практические занятия и самостоятельная работа обучающихся.

**Организация теоретического обучения** предполагает использование инновационных технологий проведения занятий лекционного типа, к которым, в частности, относятся: проблемная лекция, лекция-визуализация, лекция-беседа, лекция-дискуссия, лекция-исследование.

1. *Проблемная лекция* предполагает преимущественно всесторонний анализ исторических и социокультурных, образовательных явлений, научный поиск истины. Проблемная лекция опирается на логику последовательно моделируемых проблемных ситуаций путем постановки проблемных вопросов или предъявления проблемных задач.

2. *Лекция-визуализация* реализует принцип наглядности и учит обучающихся преобразовывать устную и письменную информацию в визуальную форму, что формирует у них профессиональное мышление за счет систематизации и выделения наиболее значимых, существенных элементов содержания обучения.

3. *Лекция-беседа* является наиболее распространенной и сравнительно простой формой активного вовлечения обучающихся в учебный процесс. Такая

лекция предполагает непосредственный контакт (диалог) педагогического работника с аудиторией.

4. *Лекция-дискуссия*, в которой в отличие от лекции-беседы педагогический работник при изложении лекционного материала не только использует ответы обучающихся на свои вопросы, но и организует свободный обмен мнениями в интервалах между логическими разделами.

**Организация практических занятий по дисциплине** направлена на углубление научно-теоретических знаний обучающихся, формирование практических умений и овладение определенными методами самостоятельной работы.

Практические занятия представляют собой занятия по решению различных прикладных задач, образцы которых были даны на лекциях.

Задачи практических занятий:

- помочь обучающимся систематизировать, закрепить и углубить знания теоретического характера;
- научить обучающихся приемам решения задач из предметной области дисциплины;
- способствовать овладению навыками и умениями, входящих в структуру формируемых компетенций в результате освоения дисциплины;
- научить их работать с информацией, книгой, пользоваться справочной и научной и методической литературой;
- формировать умение учиться самостоятельно, т.е. овладевать методами, способами и приемами самообучения, саморазвития и самоконтроля.

Содержание практических работ составляют:

- устные экспресс-опросы;
- групповые дискуссии;
- выполнение практических заданий;
- письменное или компьютерное экспресс-тестирование и др.

Цели практических занятий наилучшим образом достигаются в том случае, если студент предварительно проработал тематику практического занятия. Поэтому преподаватель должен информировать студентов о теме следующего практического занятия, чтобы они могли целенаправленно самостоятельно заниматься в домашних условиях.

**Организация лабораторных занятий по дисциплине** направлена на следующие цели и задачи:

- углубление и закрепление знания теоретического курса путем практического изучения в лабораторных условиях изложенных в лекциях законов и положений;
- приобретение навыков в научном экспериментировании, анализе полученных результатов;
- формирование первичных навыков организации, планирования и проведения научных исследований.

Порядок подготовки лабораторного занятия:

- изучение требований программы дисциплины;



- формулировка цели и задач лабораторного занятия;
- разработка плана проведения лабораторного занятия;
- подбор содержания лабораторного занятия;
- разработка необходимых для лабораторного занятия инструкционных карт;
- моделирование лабораторного занятия;
- проверка специализированной лаборатории на соответствие санитарно-гигиеническим нормам, требованиям по безопасности и технической эстетике;
- проверка количества лабораторных мест, необходимых и достаточных для достижения поставленных целей обучения;
- проверка материально-технического обеспечения лабораторных занятий на соответствие требованиям программы дисциплины.

Формы проведения лабораторных занятий:

- фронтальная;
- по циклам;
- индивидуальная;
- смешанная (комбинированная).

При проведении лабораторных работ используют три подхода к их выполнению:

- на основе рецептурных действий обучающихся, когда они проявляют умение работать преимущественно в стандартных условиях, отраженных в руководстве по лабораторному практикуму;
- на основе частично поисковых действий, когда обучающиеся могут действовать достаточно самостоятельно, решать несложные творческие задачи при подсказке или непосредственном руководстве преподавателя;
- на основе активных творческих действий обучающихся, когда они проявляют способность действовать в условиях, близких к реальным, используя запас приобретенных знаний.

***Самостоятельная работа обучающихся*** предполагает аудиторную и внеаудиторную формы организации.

Основными видами самостоятельной работы обучающихся без участия педагогического работника являются: формирование и усвоение содержания конспекта лекций на базе рекомендованной лектором учебной литературы, включая информационные образовательные ресурсы (электронные учебники, электронные библиотеки и др.); подготовка к занятиям; составление аннотированного списка статей из соответствующих журналов по отраслям знаний и т.п.; текущий самоконтроль, выполнение курсового проекта/курсовой работы.

Выполнение курсового проекта/курсовой работы по дисциплине предусматривает информирование студентов о ее целях, структуре, выдачу методических указаний и задания, разъяснения по выбору варианта, ознакомление с порядком и сроками сдачи готовых материалов, проведение индивидуальных консультаций и разъяснение отдельных вопросов при необходимости.

Основными видами самостоятельной работы обучающихся с участием педагогического работника являются: текущие консультации, прием и разбор домашних заданий и др.

При подготовке к зачету и экзамену необходимо ориентироваться на конспекты лекций, рекомендуемую литературу, консультации преподавателя и др.

## 11.2. Методические материалы для обучающихся

Обучающимся, изучающим дисциплину, необходимо знать требования, предъявляемые к их различным видам учебных занятий, в том числе лекционным, практическим, индивидуальным и др. (таблица 12).

Таблица 12 – Методические рекомендации обучающимся по освоению дисциплины

Вид учебной работы	Организация деятельности обучающегося
Лекции	Изучение дисциплины следует начинать с прослушивания и конспектирования лекций, перечитывать конспект перед выполнением домашних заданий и практическими занятиями. Написание конспекта лекций: кратко, схематично, последовательно фиксировать основные положения, выводы, формулировки, обобщения; помечать важные мысли, выделять ключевые слова, термины. Проверка терминов, понятий с помощью энциклопедий, словарей, справочников с выписыванием толкований в тетрадь. Обозначить вопросы, термины, материал, который вызывает трудности, пометить и попытаться найти ответ в рекомендуемой литературе. Если самостоятельно не удастся разобраться в материале, необходимо сформулировать вопрос и задать педагогическому работнику на консультации, на практическом занятии. Над конспектами лекций надо работать систематически: первый просмотр рекомендуется сделать вечером того же дня, когда была прочитана лекция, затем просмотреть через 3-4 дня, и сделать это еще раз накануне практического занятия.
Практические занятия	Ознакомление с целью и задачами занятия. Конспектирование источников. Работа с конспектом лекций, подготовка ответов к контрольным вопросам, просмотр рекомендуемой литературы, работа с текстом. Прослушивание аудио- и видеозаписей по заданной теме. Выполнение (решение) практических заданий и задач по алгоритму, на основе частично поисковой и или исследовательской деятельности и др.
Лабораторные работы	Подготовка к эксперименту (ознакомление с целью и задачами, ходом лабораторной работы, работа с конспектом лекций, подготовка ответов к контрольным вопросам, просмотр рекомендуемой литературы, подготовка таблиц для фиксирования хода и результатов опытно-экспериментальной работы и др.). Проведение измерений (вводный и текущий инструктаж, проведение опытов и экспериментов). Обработка полученных результатов; формулировка выводов и написание отчета. Защита отчета по лабораторной работе.
Изучение дополнительной литературы и самостоятельное формирование конспекта	Ознакомление с основной и дополнительной литературой, включая справочные издания, зарубежные источники, конспект основных положений, терминов, сведений, требующих для запоминания и являющихся основополагающими в конкретной теме. Составление аннотаций к прочитанным источ-

Вид учебной работы	Организация деятельности обучающегося
	никам и др. Рефлексия собственных достижений
Выполнение курсового проекта	При выполнении курсового проекта/курсовой работы, обучающемуся следует придерживаться методических указаний. Предусмотрен следующий алгоритм действий: выбор варианта темы курсовой работы/курсового проекта, подбор и систематизация теоретического материала, являющегося основой для написания теоретического раздела/решения практических задач, проведение расчетов по исходным данным и анализ полученных значений, формулирование выводов по полученным результатам. Выполненная работа передается преподавателю на проверку. При необходимости осуществляется доработка отдельных частей работы с учетом требований и замечаний преподавателя.
Подготовка к зачету и экзамену	При подготовке к зачету/зачету с оценкой/экзамену необходимо ориентироваться на конспекты лекций, рекомендуемую литературу, шкалу оценивания и др.

## 12. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЕ

### 12.1. Виды и средства оценивания результатов освоения дисциплины

Виды и средства оценивания результатов освоения дисциплины представлены в таблице 13.

Таблица 13 – Виды и средства оценивания результатов освоения дисциплины

Код индикатора достижения компетенции	Оценочные средства текущего контроля успеваемости	Оценочные средства промежуточной аттестации обучающихся
ПК-3.1.	1. Устные экспресс-опросы (темы 1-12). 2. Экспресс-тестирование (комплекты тестов по темам 1-12). 3. Расчетно-графическая работа, курсовая работа, курсовой проект <i>(при наличии)</i> .	Вопросы к экзамену № 1-12.
ПК-3.2	1. Устные экспресс-опросы. (темы 1-12). 2. Экспресс-тестирование (комплекты тестов по темам 1-12). 3. Расчетно-графическая работа, курсовая работа, курсовой проект <i>(при наличии)</i> .	Вопросы к экзамену № 1-12.
ПК-3.3	4. Устные экспресс-опросы. (темы 1-12). 5. Экспресс-тестирование (комплекты тестов по темам 1-12). 6. Расчетно-графическая работа, курсовая работа, курсовой проект <i>(при наличии)</i> .	Вопросы к экзамену № 1-12.

### 12.2. Шкала оценивания при текущем контроле успеваемости

Оценивание отдельных видов работ в процессе изучения дисциплины рекомендуется осуществлять с использованием следующей шкалы:

– обучающийся ответил правильно на более, чем 90 % заданных вопросов или вопросов-тестов, выполнил и успешно защитил практические работы, показал отличное владение навыками применения полученных знаний и умений при решении профессиональных задач в рамках усвоенного учебного материала и т.д. – «отлично» (максимальный уровень освоения компетенций);

– обучающийся ответил правильно на 75-89% заданных вопросов или вопросов-тестов, выполнил и защитил практические работы с незначительными замечаниями, показал хорошее владение навыками применения полученных знаний и умений при решении профессиональных задач в рамках усвоенного учебного материала и т.д. – «хорошо» (средний уровень освоения компетенций);

– обучающийся ответил правильно на 60-74% заданных вопросов или вопросов-тестов, выполнил и защитил практические работы со значительными замечаниями, показал удовлетворительное владение навыками применения полученных знаний и умений при решении профессиональных задач в рамках усвоенного учебного материала и т.д. – «удовлетворительно» (минимальный уровень освоения компетенций);

– обучающийся ответил правильно на менее, чем 60% заданных вопросов или вопросов-тестов, не выполнил все или выполнил часть практических работ, не защитил или защитил их со значительными замечаниями, при выполнении задания обучающийся не продемонстрировал уровень самостоятельного владения умениями и навыками при решении профессиональных задач в рамках усвоенного учебного материала и т.д. – «неудовлетворительно» (минимальный уровень освоения компетенций не достигнут).

Критерии и шкала оценки доклада (реферата), его презентации по дисциплине представлены в таблице 14.

Таблица 14 – Критерии и шкала оценки доклада (реферата), его презентации по дисциплине

Оценка	Оцениваемые параметры
«отлично»	Теоретический вопрос раскрыт полностью без смысловых и логических ошибок. Задание решено верно. На защите ответ обучающегося полный и правильный. Обучающийся способен изложить решение задания, сделать собственные выводы, проанализировать основные показатели. В полном объеме представлен соответствующий графический материал.
«хорошо»	Теоретический вопрос раскрыт на достаточно высоком уровне без смысловых и логических ошибок. Задание решено верно. Имеются незначительные недочеты в определении единиц измерения, точности вычислений и т.п. На защите ответ обучающегося в целом полный и правильный. Обучающийся способен изложить решение задания, сделать собственные выводы, проанализировать основные показатели. В полном объеме представлен соответствующий графический материал.
«удовлетворительно»	Теоретический вопрос раскрыт на достаточном уровне, без существенных смысловых и логических ошибок. Задание решено верно, но имеются значительные недочеты в его решении, свя-

Оценка	Оцениваемые параметры
	занные с неполнотой ответа, с правильным исчислением одних данных и неверным – других и пр. На защите ответ неполный. Обучающийся способен четко изложить решение задания, но допускает неточности в формулировке собственных выводов и анализе основных показателей. В неполном объеме представлен графический материал.
«неудовлетворительно»	Теоретический вопрос не раскрыт или раскрыт не полностью при наличии разного рода неточностей и ошибок. Задание решено со значительными недочетами, с неполными ответа, с неправильным исчислением данных. На защите ответ обучающегося неполный. Обучающийся не способен четко изложить решение задания, допускает неточности в формулировке собственных выводов, не способен проанализировать основные показатели. Графический материал не представлен или представлен не в полном объеме.

В процесс преподавания дисциплины педагогическим работником формируется оценка, характеризующая текущую успеваемость обучающегося.

### 12.3. Шкала оценивания при промежуточной аттестации обучающихся

При проведении промежуточной аттестации обучающихся в форме зачета и экзамена используется шкала оценивания, представленная в таблице 15.

Таблица 15 – Шкала оценивания при промежуточной аттестации обучающихся

Уровень освоения (оценка)	Планируемые результаты освоения дисциплины
Высокий (зачтено / «отлично»)	Обучающийся глубоко и прочно усвоил теоретический и практический материал, уверенно это демонстрирует в ходе промежуточной аттестации. Исчерпывающе и логически стройно излагает учебный материал, умеет увязывать теорию с практикой, справляется с решением задач профессиональной направленности высокого уровня сложности, правильно обосновывает принятые решения. Свободно ориентируется в учебной и профессиональной литературе.
Повышенный (зачтено / «хорошо»)	Обучающийся знает теоретический и практический материал, грамотно и по существу излагает его в ходе промежуточной аттестации, не допуская существенных неточностей. Правильно применяет теоретические положения при решении практических задач профессиональной направленности разного уровня сложности, владеет необходимыми для этого навыками и приёмами. Достаточно хорошо ориентируется в учебной и профессиональной литературе.
Базовый (зачтено / «удовлетворительно»)	Обучающийся знает на базовом уровне теоретический и практический материал, допускает отдельные ошибки при его изложении в ходе промежуточной аттестации. Обучающийся испытывает определённые затруднения в применении теоретических положений при решении практических задач профессиональной направленности стандартного уровня сложности, владеет необходимыми для этого базовыми навыками и приё-

Уровень освоения (оценка)	Планируемые результаты освоения дисциплины
	мами. Демонстрирует достаточный уровень знания учебной литературы по дисциплине.
Низкий (не зачтено / «неудовлетворительно»)	Обучающийся не знает на пороговом уровне теоретический и практический материал, допускает грубые ошибки при его изложении на занятиях и в ходе промежуточной аттестации. Испытывает серьёзные затруднения в применении теоретических положений при решении практических задач профессиональной направленности стандартного уровня сложности, не владеет необходимыми для этого навыками и приёмами. Демонстрирует фрагментарные знания учебной литературы по дисциплине.

Уровень знаний, умений и навыков обучающегося при выполнении и защите курсовой работы (курсового проекта) оценивается по пятибалльной системе. Шкала оценивания представлена в таблице 16.

Таблица 16 – Шкала оценивания, применяемая при выполнении и защите курсовой работы (курсового проекта) для технических дисциплин

Уровень освоения (оценка)	Планируемые результаты освоения дисциплины
«отлично»	<p><b>а) Содержание работы:</b></p> <ul style="list-style-type: none"> <li>– работа полностью соответствует теме исследования;</li> <li>– грамотно обоснована актуальность работы;</li> <li>– обучающийся показывает глубокую общетеоретическую подготовку;</li> <li>– обучающийся корректно использует терминологический аппарат;</li> <li>– в работе используются актуальные источники, нормативные документы, законодательные акты;</li> <li>– обучающийся демонстрирует умение работать с различными видами источников информации, в том числе с данными, полученными экспериментальным путем и с электронными библиотечными системами вуза;</li> <li>– обучающийся проявляет умение обобщать, систематизировать и научно классифицировать материал;</li> <li>– исследование завершается научно-значимыми выводами и/или практическими рекомендациями.</li> </ul> <p><b>б) Владение навыками научного исследования:</b></p> <ul style="list-style-type: none"> <li>– обучающийся владеет методологическими подходами к изучению предмета исследования и конкретными методиками;</li> <li>– обучающийся умеет грамотно составить программу исследования (определить научную проблему, объект, предмет, цели, задачи, подобрать методы исследования), обосновать научную новизну и/или практическую значимость данного исследования;</li> <li>– обучающийся умеет делать аргументированные выводы, соответствующие поставленным целям и задачам;</li> <li>– обучающийся умеет предложить варианты использования результатов исследования в профессиональной деятельности.</li> </ul> <p><b>в) Оформление курсовой работы (проекта):</b></p> <ul style="list-style-type: none"> <li>– работа оформлена в соответствии с локальными актами.</li> </ul> <p><b>г) Защита курсовой работы (проекта):</b></p>

Уровень освоения (оценка)	Планируемые результаты освоения дисциплины
	<ul style="list-style-type: none"> <li>– обучающийся в устном выступлении на защите адекватно представляет результаты исследования;</li> <li>– обучающийся аргументированно отвечает на вопросы и ведет научную дискуссию;</li> <li>– обучающийся владеет научным стилем изложения;</li> <li>– обучающийся владеет понятийным аппаратом.</li> </ul>
«хорошо»	<p><b>а) Содержание работы:</b></p> <ul style="list-style-type: none"> <li>– полностью соответствует теме исследования;</li> <li>– актуальность работы обоснована недостаточно аргументированно;</li> <li>– обучающийся показывает достаточную общетеоретическую подготовку, допуская погрешности в использовании терминологического аппарата;</li> <li>– обзор теоретических и практических наработок по проблеме имеет описательный, а не аналитический характер;</li> <li>– источниковая база исследования недостаточно широкая;</li> <li>– обучающийся демонстрирует умение работать с различными видами источников, в том числе с данными, полученными экспериментальным путем;</li> <li>– обучающийся проявляет способности обобщать, систематизировать и научно классифицировать материал;</li> <li>– в работе отсутствуют научно-значимые выводы и/или практические результаты.</li> </ul> <p><b>б) Владение навыками научного исследования:</b></p> <ul style="list-style-type: none"> <li>– не обоснована научная новизна и практическая значимость данного исследования;</li> <li>– присутствуют отдельные недочеты в программе исследования (недостаточно аргументированно определена научная проблема, неверно сформулированы объект, предмет, цели, задачи, методы исследования подобраны не вполне корректно);</li> <li>– выводы исследования недостаточно аргументированны, не соответствуют поставленным целям и задачам.</li> </ul> <p><b>в) Оформление курсовой работы (проекта):</b></p> <ul style="list-style-type: none"> <li>– работа оформлена в соответствии с локальными актами.</li> </ul> <p><b>г) Защита курсовой работы (проекта):</b></p> <ul style="list-style-type: none"> <li>– обучающийся в устном выступлении на защите адекватно представляет результаты исследования;</li> <li>– обучающийся владеет научным стилем изложения;</li> <li>– обучающийся владеет понятийным аппаратом;</li> <li>– обучающийся во время защиты не смог ответить на ряд вопросов по предмету исследования.</li> </ul>
«удовлетворительно»	<p><b>а) Содержание работы:</b></p> <ul style="list-style-type: none"> <li>– частично соответствует теме исследования;</li> <li>– не обоснована актуальность работы;</li> <li>– обучающийся обнаружил удовлетворительные знания по предмету;</li> <li>– в работе отсутствует обзор теоретических и практических наработок по проблеме;</li> <li>– источниковая база исследования недостаточно широка, обучаю-</li> </ul>

Уровень освоения (оценка)	Планируемые результаты освоения дисциплины
	<p>щийся использует лишь данные научной литературы;  – обучающийся не сумел продемонстрировать умение работать с различными видами источников;  – в работе отсутствуют научно-значимые выводы или практические результаты.</p> <p><b>б) Оформление курсовой работы (проекта):</b>  – работа оформлена в соответствии с локальными актами.</p> <p><b>в) Защита курсовой работы (проекта):</b>  – в устном выступлении на защите обучающийся не может адекватно представить результаты исследования;  – обучающийся отступает от научного стиля изложения;  – обучающийся затрудняется в аргументации, отвечая на вопросы по теме работы.</p>
«неудовлетворительно»	– имеются принципиальные замечания по пяти и более параметрам курсовой работы (проекта); – обучающийся допустил грубые теоретические ошибки, не владеет навыками исследования.

#### 12.4. Оценивание окончательных результатов обучения по дисциплине

Итоговая оценка по дисциплине определяется с учетом результатов промежуточной аттестации обучающегося (зачета и экзамена) и оценок, полученных обучающимся в ходе текущего контроля успеваемости в семестре.

#### 12.5. Характеристика результатов обучения

Характеристики результатов обучения по дисциплине в зависимости от полученной обучающимся оценки приведены в таблице 18.

Таблица 18 – Характеристика результатов обучения по дисциплине

Оценка	Характеристика результатов обучения
Зачтено / «Отлично» (высокий уровень освоения всех индикаторов достижения компетенций в дисциплине)	Содержание дисциплины освоено полностью, все цели достигнуты, все предусмотренные программой обучения учебные задания выполнены
Зачтено / «Хорошо» (повышенный уровень освоения всех индикаторов достижения компетенций в дисциплине)	Содержание дисциплины освоено полностью, все предусмотренные программой обучения учебные задания выполнены с незначительными замечаниями
Зачтено / «Удовлетворительно» (базовый уровень освоения всех индикаторов достижения компетенций в дисциплине)	Содержание дисциплины освоено частично, большинство предусмотренных программой обучения учебных заданий выполнено, в них имеются ошибки
Не зачтено / «Неудовлетворительно» (низкий уровень освоения всех индикаторов достижения компетенций в дисциплине)	Содержание дисциплины не освоено, большинство предусмотренных программой обучения учебных заданий либо не выполнены, либо содержат грубые ошибки; дополнительная самостоятельная работа над материалом не привела к какому-либо значительному повышению качества выполнения учебных заданий



## **12.6. Контрольно-измерительные материалы для текущего контроля успеваемости и промежуточной аттестации обучающихся**

Контрольно-измерительные материалы для текущего контроля успеваемости и промежуточной аттестации обучающихся представлены в электронном курсе «Методы анализа данных и естественно-языковых текстов», размещенном в системе электронной поддержки учебных курсов на базе программного обеспечения Moodle со встроенной подсистемой тестирования (edu.tu-bryansk.ru), входящей в состав электронной информационно-образовательной среды БГТУ (<http://edu.tu-bryansk.ru>) и «Фонд оценочных средств по дисциплине «Методы анализа данных и естественно-языковых текстов».

## **13. ВОСПИТАТЕЛЬНАЯ РАБОТА**

В соответствии с Федеральным законом «Об образовании в Российской Федерации» воспитание - «деятельность, направленная на развитие личности, создание условий для самоопределения и социализации обучающихся на основе социокультурных, духовно-нравственных ценностей и принятых в российском обществе правил и норм поведения в интересах человека, семьи, общества и государства, формирование у обучающихся чувства патриотизма, гражданственности, уважения к памяти защитников Отечества и подвигам Героев Отечества, закону и правопорядку, человеку труда и старшему поколению, взаимного уважения, бережного отношения к культурному наследию и традициям многонационального народа Российской Федерации, природе и окружающей среде».

В учебном процессе воспитательная работа с обучающимися реализуется средствами учебных дисциплин.

Воспитательная деятельность в ходе преподавания дисциплины направлена на формирование у обучающегося системы убеждений, нравственных норм и общекультурных качеств, на оказание им помощи в жизненном самоопределении, нравственном, гражданском и профессиональном становлении, на создание условий для самореализации личности. Воспитательная работа также ориентирует обучающихся на будущую профессиональную деятельность, формируя не только личностные, но и профессионально значимые качества.

Воспитательные задачи во время учебных занятий выполняются в скрытой (контекстной) и открытой (целенаправленной) формах. Скрытая форма воспитательной работы представляет собой воздействие всего хода педагогического процесса на становление личностных качеств обучающихся. Например, соблюдение педагогическим работником трудовой дисциплины, демонстрация преданности науке, заинтересованность в успехе обучающихся, правильная речь, хорошие манеры и т.п. имеют положительное воспитательное значение и формируют у обучающихся добросовестность, исполнительность, трудолюбие, ответственность и другие положительные качества. Обучающиеся неосознанно перенимают данные черты у педагогического работника.

Воспитание в открытой форме – это целенаправленное воздействие содержанием учебной дисциплины на становление личности обучающегося.

Например, решение проблем и исследовательская работа формируют у обучающихся умение аргументировать, самостоятельно мыслить, стремление к научному поиску, развивают творчество, профессиональные умения.