

Документ подписан простой электронной подписью  
Информация о владельце:

ФИО: Исаев Игорь Магомедович

Должность: Проректор по учебной работе

Дата подписания: 28.08.2023 16:36:01

Уникальный идентификатор документа:

d7a26b9e8ca85e98ec3de2eb454b4659d061f249

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

**Федеральное государственное автономное образовательное учреждение  
высшего образования**

**«Национальный исследовательский технологический университет «МИСИС»**

## Рабочая программа дисциплины (модуля)

### Анализ естественного языка

Закреплена за подразделением Кафедра автоматизированного проектирования и дизайна

Направление подготовки 09.04.01 ИНФОРМАТИКА И ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА

Профиль Науки о данных

Квалификация **Магистр**

Форма обучения **очная**

Общая трудоемкость **7 ЗЕТ**

Часов по учебному плану 252

в том числе:

аудиторные занятия 51

самостоятельная работа 201

Формы контроля в семестрах:  
зачет с оценкой 3

#### Распределение часов дисциплины по семестрам

Семестр (<Курс>.<Семестр на курсе>)	3 (2.1)		Итого	
	19			
Неделя	УП	РП	УП	РП
Лекции	17	17	17	17
Практические	34	34	34	34
Итого ауд.	51	51	51	51
Контактная работа	51	51	51	51
Сам. работа	201	201	201	201
Итого	252	252	252	252

Программу составил(и):

*к.т.н., доцент, Калитин Д.В.*

Рабочая программа

**Анализ естественного языка**

Разработана в соответствии с ОС ВО:

Самостоятельно устанавливаемый образовательный стандарт высшего образования - магистратура Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский технологический университет «МИСиС» по направлению подготовки 09.04.01 ИНФОРМАТИКА И ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА (приказ от 05.03.2020 г. № 95 о.в.)

Составлена на основании учебного плана:

09.04.01 Информатика и вычислительная техника, 09.04.01-МИВТ-23-3.plx Науки о данных, утвержденного Ученым советом НИТУ МИСИС в составе соответствующей ОПОП ВО 22.06.2023, протокол № 5-23

Утверждена в составе ОПОП ВО:

09.04.01 Информатика и вычислительная техника, Науки о данных, утвержденной Ученым советом НИТУ МИСИС 22.06.2023, протокол № 5-23

Рабочая программа одобрена на заседании

**Кафедра автоматизированного проектирования и дизайна**

Протокол от 18.06.2020 г., №10

Руководитель подразделения Горбатов А.В.

**1. ЦЕЛИ ОСВОЕНИЯ**

1.1	Сформировать навыки обработки неструктурированного текста на естественном языке, а также навыки работы с современными математическими методами и компьютерными алгоритмами для решения этих задач.
1.2	В результате освоения модуля студент должен демонстрировать способность и готовность:
1.3	- разрабатывать сложные архитектуры нейронных сетей для решения задач обработки естественного языка;
1.4	- реализовывать глубокие нейронные сети с помощью библиотеки глубокого обучения Tensorflow;
1.5	- идентифицировать проблемы в обучении глубоких нейронных сетей для задач обработки естественного
1.6	языка и предлагать пути их решения.

**2. МЕСТО В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ**

Блок ОП:		Б1.В.ДВ.02
<b>2.1</b>	<b>Требования к предварительной подготовке обучающегося:</b>	
2.1.1	Глубокое обучение в науках о данных	
2.1.2	Инжиниринг данных	
2.1.3	Менеджмент для наук о данных	
2.1.4	Системы хранения и обработки данных	
2.1.5	Управление человеческими ресурсами в проектной деятельности	
2.1.6	Учебная практика по применению машинного обучения	
2.1.7	Алгоритмизация и программирование	
2.1.8	Высшая математика для машинного обучения	
2.1.9	Организация и технология научных исследований	
2.1.10	Основные алгоритмы машинного обучения	
2.1.11	Программирование на Python	
2.1.12	Современные методы решения инженерных задач	
<b>2.2</b>	<b>Дисциплины (модули) и практики, для которых освоение данной дисциплины (модуля) необходимо как предшествующее:</b>	
2.2.1	Подготовка к процедуре защиты и защита выпускной квалификационной работы	
2.2.2	Преддипломная практика	

**3. РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ, СООТНЕСЕННЫЕ С ФОРМИРУЕМЫМИ КОМПЕТЕНЦИЯМИ**

<b>ОПК-7: Способен адаптировать зарубежные комплексы обработки информации и автоматизированного проектирования к нуждам отечественных предприятий</b>	
<b>Знать:</b>	
ОПК-7-31 Функциональные требования к прикладному программному обеспечению для решения актуальных задач предприятий отрасли, национальные стандарты обработки информации и автоматизированного проектирования	
<b>ОПК-2: Способен проектировать и разрабатывать оригинальные алгоритмы и программные средства, в том числе с использованием современных интеллектуальных технологий, для решения профессиональных задач в условиях неопределенности и альтернативных решений в рамках междисциплинарных областей</b>	
<b>Знать:</b>	
ОПК-2-31 Современное программное и аппаратное обеспечение информационных и автоматизированных систем	
<b>ОПК-6: Способен осуществлять моделирование, анализ и эксперименты в целях проведения детального исследования, разрабатывать компоненты программно-аппаратных комплексов обработки информации и автоматизированного проектирования</b>	
<b>Знать:</b>	
ОПК-6-31 Аппаратные средства и платформы инфраструктуры информационных технологий, виды, назначение, архитектуру, методы разработки и администрирования программно-аппаратных комплексов объекта профессиональной деятельности	
<b>ОПК-1: Способен самостоятельно приобретать, развивать и применять математические, естественнонаучные, социально-экономические и профессиональные знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте</b>	
<b>Знать:</b>	
ОПК-1-31 Математические, естественнонаучные и социально-экономические методы для использования в профессиональной деятельности	

<b>УК-1: Способен осуществлять критический анализ новых и сложных инженерных объектов, процессов и систем в междисциплинарном контексте, проблемных ситуаций на основе системного подхода, выбрать и применить наиболее подходящие и актуальные методы из существующих аналитических, вычислительных и экспериментальных методов или новых и инновационных методов, вырабатывать стратегию действий</b>
<b>Знать:</b>
УК-1-31 Основные понятия и методы фундаментальных дисциплин
<b>ПК-1: Способен к проведению научно-исследовательских и опытно-конструкторских разработок при исследовании самостоятельных тем</b>
<b>Знать:</b>
ПК-1-31 Методы и средства разработки программного обеспечения, методы управления проектами разработки программного обеспечения, способы организации проектных данных, нормативно-технические документы (стандарты и регламенты) по разработке программных средств и проектов
<b>УК-2: Способен интегрировать знания и принимать решения в сложных ситуациях, формулировать суждения на основе неполной или ограниченной информации, управлять проектом на всех этапах его жизненного цикла</b>
<b>Знать:</b>
УК-2-31 Методы управления проектами; этапы жизненного цикла проекта
<b>ОПК-7: Способен адаптировать зарубежные комплексы обработки информации и автоматизированного проектирования к нуждам отечественных предприятий</b>
<b>Уметь:</b>
ОПК-7-У1 Приводить зарубежные комплексы обработки информации в соответствие с национальными стандартами, интегрировать с отраслевыми информационными системами
<b>ПК-1: Способен к проведению научно-исследовательских и опытно-конструкторских разработок при исследовании самостоятельных тем</b>
<b>Уметь:</b>
ПК-1-У1 Выбирать средства разработки, оценивать сложность проектов, планировать ресурсы, контролировать сроки выполнения и оценивать качество полученного результата
<b>ОПК-6: Способен осуществлять моделирование, анализ и эксперименты в целях проведения детального исследования, разрабатывать компоненты программно-аппаратных комплексов обработки информации и автоматизированного проектирования</b>
<b>Уметь:</b>
ОПК-6-У1 Анализировать техническое задание, разрабатывать и оптимизировать программный код для решения задач обработки информации и автоматизированного проектирования
<b>ОПК-2: Способен проектировать и разрабатывать оригинальные алгоритмы и программные средства, в том числе с использованием современных интеллектуальных технологий, для решения профессиональных задач в условиях неопределенности и альтернативных решений в рамках междисциплинарных областей</b>
<b>Уметь:</b>
ОПК-2-У1 Разрабатывать программное и аппаратное обеспечение информационных и автоматизированных систем для решения профессиональных задач
<b>УК-2: Способен интегрировать знания и принимать решения в сложных ситуациях, формулировать суждения на основе неполной или ограниченной информации, управлять проектом на всех этапах его жизненного цикла</b>
<b>Уметь:</b>
УК-2-У1 Разрабатывать и анализировать альтернативные варианты проектов для достижения намеченных результатов; разрабатывать проекты, определять целевые этапы и основные направления работ
<b>УК-1: Способен осуществлять критический анализ новых и сложных инженерных объектов, процессов и систем в междисциплинарном контексте, проблемных ситуаций на основе системного подхода, выбрать и применить наиболее подходящие и актуальные методы из существующих аналитических, вычислительных и экспериментальных методов или новых и инновационных методов, вырабатывать стратегию действий</b>
<b>Уметь:</b>
УК-1-У1 Применять знания, полученные в области фундаментальных наук, и использовать их в профессиональной деятельности
<b>ОПК-1: Способен самостоятельно приобретать, развивать и применять математические, естественнонаучные, социально-экономические и профессиональные знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте</b>
<b>Уметь:</b>
ОПК-1-У1 Решать нестандартные профессиональные задачи, в том числе в новой или незнакомой среде и в междисциплинарном контексте, с применением математических, естественнонаучных, социально-экономических и профессиональных знаний

<b>ПК-1: Способен к проведению научно-исследовательских и опытно-конструкторских разработок при исследовании самостоятельных тем</b>
<b>Владеть:</b>
ПК-1-В1 Методами разработки технического задания, составления планов, распределения задач, тестирования и оценки качества программных средств
<b>ОПК-7: Способен адаптировать зарубежные комплексы обработки информации и автоматизированного проектирования к нуждам отечественных предприятий</b>
<b>Владеть:</b>
ОПК-7-В1 Методами настройки интерфейса, разработки пользовательских шаблонов, подключения библиотек, добавления новых функций
<b>УК-1: Способен осуществлять критический анализ новых и сложных инженерных объектов, процессов и систем в междисциплинарном контексте, проблемных ситуаций на основе системного подхода, выбрать и применить наиболее подходящие и актуальные методы из существующих аналитических, вычислительных и экспериментальных методов или новых и инновационных методов, вырабатывать стратегию действий</b>
<b>Владеть:</b>
УК-1-В1 Методикой выбора методов решения задач профессиональной деятельности на основе теоретических знаний
<b>ОПК-1: Способен самостоятельно приобретать, развивать и применять математические, естественнонаучные, социально-экономические и профессиональные знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте</b>
<b>Владеть:</b>
ОПК-1-В1 Методами теоретического и экспериментального исследования объектов профессиональной деятельности, в том числе в новой или незнакомой среде и в междисциплинарном контексте
<b>ОПК-2: Способен проектировать и разрабатывать оригинальные алгоритмы и программные средства, в том числе с использованием современных интеллектуальных технологий, для решения профессиональных задач в условиях неопределенности и альтернативных решений в рамках междисциплинарных областей</b>
<b>Владеть:</b>
ОПК-2-В1 Методами модернизации программного и аппаратного обеспечения информационных и автоматизированных систем для решения профессиональных задач
<b>УК-2: Способен интегрировать знания и принимать решения в сложных ситуациях, формулировать суждения на основе неполной или ограниченной информации, управлять проектом на всех этапах его жизненного цикла</b>
<b>Владеть:</b>
УК-2-В1 Навыками разработки проектов в избранной профессиональной сфере; методами оценки эффективности проекта, а также потребности в ресурсах
<b>ОПК-6: Способен осуществлять моделирование, анализ и эксперименты в целях проведения детального исследования, разрабатывать компоненты программно-аппаратных комплексов обработки информации и автоматизированного проектирования</b>
<b>Владеть:</b>
ОПК-6-В1 Методами составления технической документации по использованию и настройке компонентов программно-аппаратного комплекса

#### 4. СТРУКТУРА И СОДЕРЖАНИЕ

Код занятия	Наименование разделов и тем /вид занятия/	Семестр / Курс	Часов	Формируемые индикаторы компетенций	Литература и эл. ресурсы	Примечание	КМ	Выполняемые работы
	<b>Раздел 1. Раздел 1. Задачи NLP</b>							
1.1	Задачи и ограничения, Сложности понимания, Классификация задач, Программное обеспечение. /Лек/	3	4	УК-1-31 ОПК-1-31	Л1.1 Л1.3Л2.1 Л2.2 Э1		КМ1	
	<b>Раздел 2. Раздел 2. Векторизация текста</b>							
2.1	5 NLP-операций для обработки текста /Лек/	3	1	УК-1-31 УК-2-31 ОПК-1-31	Л1.1 Л1.3Л2.1 Л2.2 Э1		КМ1	

2.2	Data Mining. Извлечение признаков (Feature Extraction) /Пр/	3	5	УК-1-31 УК-1-У1 УК-1-В1 ОПК-1-31 ОПК-1-У1 ОПК-1-В1 ОПК-6-31 ОПК-6-У1 ОПК-6-В1	Л1.1 Л1.3Л2.1 Л2.2Л3.1 Э1			
2.3	Векторизация: превращение слов в цифры /Пр/	3	7	УК-1-У1 УК-1-В1 ОПК-6-31 ОПК-6-У1 ОПК-6-В1 ОПК-7-31 ОПК-7-У1 ОПК-7-В1	Л1.1 Л1.3Л2.1 Л2.2Л3.1 Э1			
	<b>Раздел 3. Раздел 3. Нейронные сети в NLP</b>							
3.1	Нейронные сети с одним скрытым слоем. Векторные представления слов (word embeddings). Рекурсивные нейронные сети. /Лек/	3	3	УК-1-31 УК-2-31 ОПК-1-31 ОПК-2-31 ОПК-6-31 ПК-1-31 ОПК-7-31	Л1.1 Л1.3Л2.1 Л2.2 Э1		КМ1	
	<b>Раздел 4. Раздел 4. Типы обработки последовательностей</b>							
4.1	Понятие обработки естественного языка. Машинный перевод. Голосовые помощники. /Лек/	3	4	УК-1-31 УК-2-31 ОПК-1-31 ОПК-2-31 ОПК-6-31 ПК-1-31 ОПК-7-31	Л1.1 Л1.3Л2.1 Л2.2 Э1		КМ1	
4.2	Краткое изложение текста (Text Summarization) /Пр/	3	5	УК-1-У1 УК-1-В1 ОПК-1-У1 ОПК-1-В1 ОПК-2-У1 ОПК-2-В1 ОПК-6-У1 ОПК-6-В1 ПК-1-У1 ПК-1-В1 ОПК-7-У1 ОПК-7-В1	Л1.1 Л1.3Л2.1 Л2.2Л3.1 Э1			
4.3	Глубокое обучение в NLP. Векторное представление (text embeddings). GloVe. Вопросно-ответные (QA) системы. /Ср/	3	45	УК-1-31 УК-1-У1 ОПК-1-31 ОПК-1-У1 ОПК-2-31 ОПК-2-У1 ОПК-6-31 ОПК-6-У1 ПК-1-31 ПК-1-У1 ОПК-7-31 ОПК-7-У1	Л1.1 Л1.3Л2.1 Л2.2 Э1			
	<b>Раздел 5. Раздел 5. Механизм внимания</b>							
5.1	Обобщенный механизм внимания. Пример использования обобщенного механизма внимания для задачи машинного перевода /Лек/	3	1	УК-1-31 УК-2-31 ОПК-1-31 ОПК-2-31 ОПК-6-31 ПК-1-31 ОПК-7-31	Л1.1 Л1.3Л2.1 Л2.2 Э1		КМ1	

5.2	Базовая архитектура Seq2seq. Применение механизма внимания для Seq2seq. Self-Attention /Ср/	3	36	УК-1-31 УК-1-У1 ОПК-1-31 ОПК-1-У1 ОПК-2-31 ОПК-2-У1 ОПК-6-31 ПК-1-31 ПК-1-У1 ОПК-7-31	Л1.1 Л1.3Л2.1 Л2.2 Э1				
<b>Раздел 6. Раздел 6. Отказ от рекуррентных сетей</b>									
6.1	Рекуррентные нейронные сети /Лек/	3	1	УК-1-31 УК-2-31 ОПК-1-31 ПК-1-31	Л1.2Л2.1 Э1		КМ1		
6.2	Обратные методы решения задач /Пр/	3	5	УК-1-У1 УК-1-В1 ОПК-1-У1 ОПК-2-У1 ОПК-2-В1 ОПК-6-У1 ПК-1-У1 ОПК-7-У1 ОПК-7-В1	Л1.2Л2.1 Л2.2Л3.1 Э1				
<b>Раздел 7. Раздел 7. TRANSFORMER</b>									
7.1	Трансформеры как графовые нейронные сети. Трансформер по частям. /Лек/	3	3	УК-1-31 УК-2-31 ОПК-1-31 ОПК-6-31 ОПК-7-31	Л1.1 Л1.2Л2.1 Э1		КМ1		
7.2	Проблемы масштабирования и «подслой» прямого распространения /Пр/	3	5	УК-1-У1 ОПК-1-У1 ОПК-1-В1 ОПК-2-У1 ОПК-2-В1 ОПК-6-У1 ПК-1-У1 ПК-1-В1 ОПК-7-У1 ОПК-7-В1	Л2.1Л3.1 Э1				
7.3	Многоголовое внимание (Multi-Head Attention) /Ср/	3	25	УК-1-31 УК-1-У1 ОПК-1-31 ОПК-2-31 ОПК-2-У1 ПК-1-31 ПК-1-У1 ОПК-7-В1	Л1.2Л2.1 Э1				
<b>Раздел 8. Раздел 8. Обучение с подкреплением</b>									
8.1	Алгоритмы. Современные исследования. /Ср/	3	40	УК-1-31 УК-1-У1 УК-1-В1 УК-2-В1 ПК-1-31 ПК-1-У1 ПК-1-В1	Л1.2Л2.1Л3.1 Э1				
<b>Раздел 9. Раздел 9. Продвинутое нейронные сети (What's next?)</b>									
9.1	Нейросетевой логический элемент. Линейно разделимые данные. Решение задач, которые не являются линейно разделимыми. /Ср/	3	55	УК-1-31 УК-2-31 ОПК-1-31 ОПК-1-У1 ОПК-2-31 ОПК-2-У1 ОПК-2-В1 ОПК-6-31 ПК-1-31 ПК-1-У1 ОПК-7-31	Л1.1 Л1.2 Л1.3Л2.1 Л2.2 Э1				
<b>Раздел 10. Раздел 10. Практический кейс</b>									

10.1	Соревнование на Kaggle: Предсказание жанра фильма по его описанию и названию /Пр/	3	7	УК-1-У1 УК-1-В1 УК-2-У1 УК-2-В1 ОПК-1-У1 ОПК-1-В1 ОПК-2-У1 ОПК-2-В1 ОПК-6-У1 ОПК-6-В1 ПК-1-У1 ПК-1-В1 ОПК-7-У1 ОПК-7-В1	Л1.1 Л1.2 Л1.3Л2.1Л3.1 Э1			
------	---	---	---	---	---------------------------	--	--	--

### 5. ФОНД ОЦЕНОЧНЫХ МАТЕРИАЛОВ

#### 5.1. Контрольные мероприятия (контрольная работа, тест, коллоквиум, экзамен и т.п), вопросы для самостоятельной подготовки

Код КМ	Контрольное мероприятие	Проверяемые индикаторы компетенций	Вопросы для подготовки
КМ1	Устные и письменные опросы по темам лекций для проведения текущей и промежуточной аттестации.	ОПК-7-31;ОПК-6-31;ОПК-2-31;ОПК-1-31;УК-2-31;УК-1-31;ПК-1-31	<p>Перечень вопросов к устным и письменным опросам:</p> <ol style="list-style-type: none"> <li>1. Задачи обработки естественного языка.</li> <li>2. Возможности методов машинного обучения в обработке естественного языка.</li> <li>3. Приложения обработки естественного языка.</li> <li>4. Модель Bag-of-Words.</li> <li>5. Задача получения векторных представлений слов.</li> <li>6. Метод Word2Vec.</li> <li>7. Рекуррентные нейронные сети.</li> <li>8. Архитектура LSTM-сетей. Двухнаправленные LSTM-сети.</li> <li>9. Применение рекуррентных сетей для построения языковой модели.</li> <li>10. Метод построения контекстно-зависимых векторных представлений слов ELMo.</li> <li>11. Постановка задачи классификации текста. Приложения.</li> <li>12. Классификация текста с помощью LSTM-сетей.</li> <li>13. Классификация текста с помощью сверточных нейронных сетей.</li> <li>14. Метод классификации текста FastText.</li> <li>15. Архитектура нейронной сети Encoder-Decoder.</li> <li>16. Механизм внимания в архитектуре Encoder-Decoder.</li> <li>17. Модель информационного поиска. Новостная агрегация и рубрикация.</li> <li>18. Извлечение информации из текста. Основные типы ресурсов.</li> <li>19. Основные формальные модели: конечные автоматы, контекстно-свободные грамматики.</li> <li>20. Свойства естественного языка, создающие сложности для автоматической обработки: омонимия, отсутствие взаимоднозначного соответствия между формой и смыслом.</li> <li>21. Цепочка обработки: основные этапы обработки.</li> <li>22. Основные платформы и пакеты для разработки систем АОТ.</li> <li>23. Основные модели автоматического синтаксического анализа: непосредственные составляющие, зависимости.</li> <li>24. Контекстно-свободные грамматики. Унификационные грамматики.</li> <li>25. Синтаксический анализ: основные проблемы автоматического анализа (омонимия, типичные случаи синтаксической омонимии, синтаксические нули).</li> <li>26. Контекстно-свободные грамматики. базовые алгоритмы (нисходящий алгоритм, алгоритм спуска, алгоритм Кока-Янгера-Касами).</li> <li>27. Зависимостные грамматики. Алгоритмы анализа в терминах зависимостей. Универсальные зависимости (UD): основные стандарты морфологической и синтаксической разметки в терминах UD.</li> <li>28. Запуск системы синтаксического анализа в терминах UD (UD-pipe).</li> </ol>

<b>5.2. Перечень работ, выполняемых по дисциплине (Курсовая работа, Курсовой проект, РГР, Реферат, ЛР, ПР и т.п.)</b>			
Код работы	Название работы	Проверяемые индикаторы компетенций	Содержание работы
P1	Практическая работа № 1	ОПК-7-У1;ОПК-7-В1;ОПК-6-У1;ОПК-6-В1;ОПК-2-У1;ОПК-2-В1;ОПК-1-У1;ОПК-1-В1;УК-2-У1;УК-2-В1;УК-1-У1;УК-1-В1;ПК-1-У1;ПК-1-В1	<p>Реализовать несколько архитектур получения векторных представлений слов:</p> <ol style="list-style-type: none"> <li>1. Подготовить обучающую выборку.</li> <li>2. Реализовать составление словаря уникальных слов в выборке.</li> <li>3. Реализовать лемматизацию, токенизацию слов.</li> <li>4. Реализовать удаление стоп слов.</li> <li>5. Реализовать архитектуру CBOW.</li> <li>6. Реализовать архитектуру skip-gram.</li> <li>7. Реализовать процедуру получения семантически близких слов skipgram (поиск слова по окружению)</li> <li>8. Реализовать процедуру получения семантически близких слов cbow (поиск окружения по слову)</li> <li>9. Реализовать процедуру сравнения семантического расстояния между словами.</li> <li>10. Реализовать процедуру оценки важности слов в запросе.</li> </ol>
P2	Практическая работа № 2	ПК-1-В1;ПК-1-У1;УК-1-В1;УК-1-У1;УК-2-В1;УК-2-У1;ОПК-1-В1;ОПК-1-У1;ОПК-2-В1;ОПК-2-У1;ОПК-6-В1;ОПК-6-У1;ОПК-7-В1;ОПК-7-У1	<p>Реализовать архитектуру рекуррентной нейронной сети для классификации тональности текста:</p> <ol style="list-style-type: none"> <li>1. Подготовить обучающую выборку.</li> <li>2. Реализовать составление словаря уникальных слов в выборке.</li> <li>3. Реализовать лемматизацию, токенизацию слов.</li> <li>4. Реализовать удаление стоп слов.</li> <li>5. Реализовать способ получения векторных представлений слова на основе готового корпуса данных.</li> <li>6. Реализовать архитектуру рекуррентной нейронной сети.</li> <li>7. Реализовать способ представления предложения в виде вектора на основе векторов слов.</li> <li>8. Исследовать влияния количества подаваемых предложений на результаты работы метода.</li> <li>9. Оценить и сравнить время работы методов.</li> <li>10. Графически отобразить полученные результаты.</li> </ol>
P3	Практическая работа № 3	ПК-1-В1;ПК-1-У1;УК-1-В1;УК-1-У1;УК-2-В1;УК-2-У1;ОПК-1-В1;ОПК-1-У1;ОПК-2-В1;ОПК-2-У1;ОПК-6-В1;ОПК-6-У1;ОПК-7-В1;ОПК-7-У1	<p>Реализовать архитектуру Encoder-Decoder для машинного перевода:</p> <ol style="list-style-type: none"> <li>1. Подготовка обучающей выборки.</li> <li>2. Реализовать составление словаря уникальных слов в выборке.</li> <li>3. Реализовать лемматизацию, токенизацию слов, удаление стоп слов.</li> <li>4. Реализовать способ получения векторных представлений слова на основе готового корпуса данных.</li> <li>5. Выбор и реализация архитектуры Encoder.</li> <li>6. Выбор и реализация архитектуры Decoder.</li> <li>7. Выбор и реализация механизма внимания Decoder.</li> <li>8. Исследование влияния механизма внимания на итоговую точность модели.</li> <li>9. Оценка точности на обучающей, валидационной и тестовой выборке.</li> <li>10. Измерение скорости работы.</li> </ol>

### 5.3. Оценочные материалы, используемые для экзамена (описание билетов, тестов и т.п.)

Экзамен не предусмотрен.

### 5.4. Методика оценки освоения дисциплины (модуля, практики. НИР)

Промежуточный контроль успеваемости по дисциплине осуществляется при защите лабораторных работ и путем проведения тестов, входящих в состав курса.

Итоговый контроль осуществляется в виде дифференцированного зачета.

## 6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ

### 6.1. Рекомендуемая литература

#### 6.1.1. Основная литература

	Авторы, составители	Заглавие	Библиотека	Издательство, год
Л1.1	Головина Е.	Лингвистический анализ текста: учебное пособие	Электронная библиотека	Оренбург: Оренбургский государственный университет, 2012

	Авторы, составители	Заглавие	Библиотека	Издательство, год
Л1.2	Яхьяева Г. Э.	Основы теории нейронных сетей	Электронная библиотека	Москва: Национальный Открытый Университет «ИНТУИТ», 2016
Л1.3	Иссерс О. С.	Теоретическая и прикладная лингвистика: курс лекций: курс лекций	Электронная библиотека	Омск: Омский государственный университет им. Ф.М. Достоевского, 2017

#### 6.1.2. Дополнительная литература

	Авторы, составители	Заглавие	Библиотека	Издательство, год
Л2.1	Крутиков В. Н., Мешечкин В. В.	Анализ данных: учебное пособие	Электронная библиотека	Кемерово: Кемеровский государственный университет, 2014
Л2.2	Мыльников Л. А., Краузе Б., Кютц М., Баде К., Шмидт И. А.	Интеллектуальный анализ данных в управлении производственными системами (подходы и методы): монография	Электронная библиотека	Москва: Библио-Глобус, 2017

#### 6.1.3. Методические разработки

	Авторы, составители	Заглавие	Библиотека	Издательство, год
Л3.1	Алпайдин Э.	Машинное обучение: новый искусственный интеллект: пер. с англ.	Библиотека МИСиС	М.: Альпина Паблшер, 2017

#### 6.2. Перечень ресурсов информационно-телекоммуникационной сети «Интернет»

Э1	Электронный курс в LMS Canvas	lms.misis.ru
----	-------------------------------	--------------

#### 6.3 Перечень программного обеспечения

П.1	Microsoft Visual Studio 2015
П.2	Microsoft SQL server 2016
П.3	Microsoft Office
П.4	LMS Canvas
П.5	MS Teams
П.6	R Studio
П.7	Python
П.8	ОС Linux (Ubuntu) / Windows

#### 6.4. Перечень информационных справочных систем и профессиональных баз данных

И.1	Полнотекстовые российские научные журналы и статьи:
И.2	— Научная электронная библиотека eLIBRARY <a href="https://elibrary.ru/">https://elibrary.ru/</a>
И.3	— Полнотекстовые деловые публикации информагентств и прессы по 53 отраслям <a href="https://polpred.com/news">https://polpred.com/news</a>
И.4	— Российская Государственная библиотека <a href="https://www.rsl.ru">https://www.rsl.ru</a>
И.5	— Единое окно доступа к информационным ресурсам <a href="http://window.edu.ru">http://window.edu.ru</a>
И.6	Иностранные базы данных (доступ с IP адресов МИСиС):
И.7	— аналитическая база (индексы цитирования) Web of Science <a href="https://apps.webofknowledge.com">https://apps.webofknowledge.com</a>
И.8	— аналитическая база (индексы цитирования) Scopus <a href="https://www.scopus.com/">https://www.scopus.com/</a>
И.9	— наукометрическая система InCites <a href="https://apps.webofknowledge.com">https://apps.webofknowledge.com</a>
И.10	— научные журналы издательства Elsevier <a href="https://www.sciencedirect.com/">https://www.sciencedirect.com/</a>
И.11	— доступ к полным версиям книг издательства Springer на английском языке <a href="https://link.springer.com/">https://link.springer.com/</a>

### 7. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ

Ауд.	Назначение	Оснащение
Б-1004	Учебная аудитория:	доска аудиторная меловая, стационарные компьютеры 12 шт., пакет лицензионных программ MS Office, комплект учебной мебели

Любой корпус Мультимедийная	Учебная аудитория для проведения занятий лекционного типа и/или для проведения практических занятий:	комплект учебной мебели до 36 мест для обучающихся, мультимедийное оборудование, магнитно-маркерная доска, рабочее место преподавателя, ПК с доступом к ИТС «Интернет», ЭИОС университета через личный кабинет на платформе LMS Canvas, лицензионные программы MS Office, MS Teams, ESET Antivirus
Читальный зал электронных ресурсов		комплект учебной мебели на 55 мест для обучающихся, 50 ПК с доступом к ИТС «Интернет», ЭИОС университета через личный кабинет на платформе LMS Canvas, лицензионные программы MS Office, MS Teams, ESET Antivirus.

### 8. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ

Студенту рекомендуется внимательно слушать лектора, следить за тем, что написано на доске или представлено на слайдах презентации, анализировать получаемую им информацию. В случае, если материал лекции непонятен, следует задать вопрос в отведенное для вопросов время. Студенту также рекомендуется конспектировать материал лекции в тетради, что улучшает запоминание.

При выполнении практических работ студенту рекомендуется внимательно анализировать поставленную задачу, уделяя особое внимание критериям оценки точности решения задачи. Особое внимание следует уделять методологическим аспектам решения задач.