

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Волков В.В.

Должность: Ректор

Дата подписания: 24.08.2023 14:24:33

Уникальный программный ключ:

ed68fd4b85b778e0f0b1bfea5dbc56cf4148f1229917e799a70e51517ff6d591

**Автономная некоммерческая образовательная организация высшего образования
«Европейский университет в Санкт-Петербурге»**

УТВЕРЖДАЮ:
Ректор  В.В. Волков
« 15 » февраля 20 23 г.
Протокол Ученого Совета
№ 1 от 15 февраля 20 23 г.

Рабочая программа дисциплины
Обработка естественного языка/ Natural Language Processing

образовательная программа
направление подготовки
09.04.03 Прикладная информатика

направленность (профиль)
«Прикладной анализ данных»
программа подготовки – магистратура

язык обучения – русский
форма обучения - очная

квалификация (степень) выпускника
Магистр

Санкт-Петербург

Автор:

Левшун Д.С., к.т.н., доцент факультета социологии АНООВО «ЕУСПб»

Рецензент:

Тушканова О.Н., к.т.н., доцент факультета социологии АНООВО «ЕУСПб»

Рабочая программа дисциплины «**Обработка естественного языка/ Natural Language Processing**», входящей в образовательную программу уровня магистратуры «Прикладной анализ данных», утверждена на заседании Совета факультета социологии

Протокол заседания № 8 от 26 января 2023 года.

АННОТАЦИЯ РАБОЧЕЙ ПРОГРАММЫ ДИСЦИПЛИНЫ «Обработка естественного языка/ Natural Language Processing»

Дисциплина **«Обработка естественного языка/ Natural Language Processing»** является дисциплиной по выбору части, формируемой участниками образовательных отношений, основной профессиональной образовательной программы высшего образования «Прикладной анализ данных» по направлению подготовки 09.04.03 Прикладная информатика.

Дисциплина «Обработка естественного языка/ Natural Language Processing» знакомит магистрантов с основными методами и подходами к обработке естественного языка, в ходе изучения дисциплины проводится анализ принципов оценки качества методов обработки естественного языка, а также магистрантам предоставляется возможность овладеть практическими навыками обработки больших коллекций текстов.

Программой дисциплины предусмотрены следующие виды контроля: текущий контроль успеваемости, промежуточный контроль в форме зачета с оценкой.

Общая трудоемкость освоения дисциплины составляет 6 зачетных единиц, 216 часов.

Содержание

1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ	5
2. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОСВОЕНИЯ ДИСЦИПЛИНЫ	5
3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ	7
4. ОБЪЕМ ДИСЦИПЛИНЫ	8
5. СОДЕРЖАНИЕ И СТРУКТУРА ДИСЦИПЛИНЫ	8
5.1 Содержание дисциплины	8
5.2 Структура дисциплины	12
6. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ	12
6.1 Общие положения.....	12
6.2 Рекомендации по распределению учебного времени по видам самостоятельной работы и разделам дисциплины.....	13
6.3 Перечень основных вопросов по изучаемым темам для самостоятельной работы обучающихся по дисциплине.....	14
6.4 Перечень литературы для самостоятельной работы обучающегося:.....	15
6.5 Перечень учебно-методического обеспечения для самостоятельной работы.....	15
7. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕЙ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ	15
7.1 Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе текущей аттестации	15
7.2 Контрольные задания для текущей аттестации	20
7.3 Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе промежуточной аттестации	25
7.4 Типовые задания к промежуточной аттестации	27
7.5 Средства оценки индикаторов достижения компетенций	28
8. ОСНОВНАЯ И ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА	30
8.1. Основная литература.....	30
8.2 Дополнительная литература	31
9. ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ, ИСПОЛЬЗУЕМЫЕ ПРИ ОСУЩЕСТВЛЕНИИ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА	31
9.1 Программное обеспечение.....	31
9.2 Перечень информационно-справочных систем и профессиональных баз данных информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины:	31
9.3 Лицензионные электронные ресурсы библиотеки Университета.....	32
9.4 Электронная информационно-образовательная среда Университета.....	32
10. МАТЕРИАЛЬНО-ТЕХНИЧЕСКАЯ БАЗА, НЕОБХОДИМАЯ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА	33
ПРИЛОЖЕНИЕ 1	34

1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Цель освоения дисциплины «Обработка естественного языка/ Natural Language Processing» — знакомство с основными методами и приложениями автоматической обработки естественного языка (ОЕЯ), получение практических навыков работы с инструментами ОЕЯ.

Задачи:

1. Знакомство с основными методами и подходами к ОЕЯ.
2. Понимание принципов оценки качества методов ОЕЯ.
3. Овладение практическими навыками обработки больших коллекций текстов.

2. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОСВОЕНИЯ ДИСЦИПЛИНЫ

В результате изучения учебной дисциплины обучающийся должен овладеть следующими компетенциями: универсальными (УК) и профессиональными (ПК). Планируемые результаты формирования компетенций и индикаторы их достижения в результате освоения дисциплины представлены в Таблице 1.

Планируемые результаты освоения дисциплины, соотнесенные с индикаторами достижения компетенций обучающихся

Таблица 1

Код и наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (знать, уметь, владеть)
УК-1 Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий	ИД.УК-1.1. Анализирует проблемную ситуацию как систему, выявляя её составляющие и связи между ними ИД.УК-1.2. Определяет пробелы в информации, необходимой для решения проблемной ситуации, и проектирует процессы по их устранению ИД.УК-1.3. Критически оценивает надёжность источников информации, работает с противоречивой информацией из разных источников ИД.УК-1.4. Разрабатывает и содержательно аргументирует стратегию решения проблемной ситуации на основе системного и междисциплинарного подходов ИД.УК-1.5. Строит сценарии реализации стратегии, определяя возможные риски и предлагая пути их устранения	Знать: методы научного познания, в основе которых лежит рассмотрение объекта как системы: целостного комплекса взаимосвязанных элементов, методы и модели стратегического планирования З (УК-1)
		Уметь: с использованием методов системного подхода анализировать альтернативные варианты решения исследовательских задач, вырабатывать стратегию действий и оценивать социальную эффективность реализации стратегических планов У (УК-1)
		Владеть: целостной системой навыков методологического использования системного подхода при решении проблем, возникающих при выполнении исследовательских работ, навыками отстаивания своей точки зрения при выработке стратегических планов выполнения исследовательских работ В (УК-1)
УК-6 Способен определять и реализовывать приоритеты собственной деятельности и способы ее совершенствования на основе самооценки	ИД.УК-6.1. Оценивает свои ресурсы и их пределы (личностные, ситуативные, временные) для успешного выполнения порученного задания ИД.УК-6.2. Определяет образовательные потребности и способы совершенствования собственной (в т.ч. профессиональной) деятельности на основе самооценки ИД.УК-6.3. Выбирает и реализует с использованием инструментов	Знать: содержание процесса формирования целей профессионального и личностного развития З (УК-6)
		Уметь: применять полученные знания для формирования собственной жизненной стратегии с учётом индивидуально-личностных особенностей У (УК-6)

Код и наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (знать, уметь, владеть)
	<p>непрерывного образования возможности развития профессиональных компетенций и социальных навыков</p> <p>ИД.УК-6.4. Выстраивает гибкую профессиональную траекторию, с учётом накопленного опыта профессиональной деятельности, изменяющихся требований рынка труда и стратегии личного развития</p>	<p>Владеть: приёмами и технологиями формирования целей саморазвития на основе самооценки</p> <p>В (УК-6)</p>
<p>ПК-2 Способен организовать аналитическую работу в IT-проекте</p>	<p>ИД.ПК-2.1. Определяет состав аналитической группы проекта</p> <p>ИД.ПК-2.2. Представляет и организует обсуждение плана аналитических работ</p> <p>ИД.ПК-2.3. Распределяет роли и аналитические работы по участникам аналитической группы проекта, поддерживает обратную связь</p> <p>ИД.ПК-2.4. Ведет работы по достижению соглашений с владельцами ресурсов о выделении ресурсов для выполнения аналитических работ в проекте</p>	<p>Знать: принципы эффективного управления командой и управления ресурсами в профессиональной деятельности</p> <p>З (ПК-2)</p> <p>Уметь: эффективно планировать и организовывать работу аналитической группы проекта, в том числе в ходе проведения совещаний</p> <p>У (ПК-2)</p> <p>Владеть: навыками управления ресурсами, процессами и работой команды для решения поставленной задачи</p> <p>В (ПК-2)</p>
<p>ПК-5 Способен применить анализ данных к научным и общественным задачам</p>	<p>ИД.ПК-5.1. Анализирует новые направления научных исследований и ставит цели по решению актуальных общественных задач</p> <p>ИД.ПК-5.2. Обосновывает выбор методов анализа данных с учетом поставленных научных и общественных задач</p> <p>ИД.ПК-5.3. Формирует программы проведения прикладного анализа данных для решения различных задач научного характера</p> <p>ИД.ПК-5.4. Непосредственно принимает участие в реализации мероприятий в рамках выполнения прикладного анализа данных при решении научных и общественных задачах</p>	<p>Знать: методологический аппарат современного анализа данных в соответствии с целями и задачами прикладного исследования</p> <p>З (ПК-5)</p> <p>Уметь: проводить прикладной анализ данных в соответствии с целями научных исследований в области общественных и смежных наук</p> <p>У (ПК-5)</p> <p>Владеть: навыками проведения прикладных исследований, основанных на выбранной научной парадигме и самостоятельно сформированной методологии анализа данных с учетом новых тенденций в сфере общественных наук</p> <p>В (ПК-5)</p>
<p>ПК-6 Способен применить технологии машинного обучения к реальным общественным задачам</p>	<p>ИД.ПК-6.1. Анализирует новые направления развития технологий машинного обучения в различных областях</p> <p>ИД.ПК-6.2. Обосновывает перспективы применения технологий машинного обучения к решению конкретных общественных задач</p> <p>ИД.ПК-6.3. Формирует программы внедрения машинного обучения для решения определенных общественных задач</p> <p>ИД.ПК-6.4. Непосредственно принимает участие в реализации</p>	<p>Знать: методологический аппарат и направления развития технологий машинного обучения, в том числе для решения общественных задач</p> <p>З (ПК-6)</p> <p>Уметь: внедрять технологии машинного обучения в процесс решения различных общественных задач</p> <p>У (ПК-6)</p> <p>Владеть:</p>

Код и наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (знать, уметь, владеть)
	мероприятий в рамках решения общественных задач по средствам технологий машинного обучения	навыками применения актуальных технологий машинного обучения к реальным общественным задачам В (ПК-6)

В результате освоения дисциплины магистрант должен:

— **знать:** современные научные представления о мире и науке, методологию научного исследования и экспертно-аналитической работы; методы сбора и обработки данных; особенности применения различных теоретико-методологических концепций с использованием технологий прикладного анализа данных; общие правила ведения научных дискуссий;

— **уметь:** синтезировать новое профессиональное знание на базе применения знаний и аналитических навыков с использованием технологий прикладного анализа данных; использовать полученные знания и умения в профессиональной деятельности, деловой коммуникации и межличностном общении; использовать навыки научных исследований общественных процессов и отношений; разрабатывать программу научного исследования, правильно оформлять и представлять результаты исследований; анализировать и оценивать общественные процессы; выявлять необходимую информацию из текстов различной тематики и направленности, а также из иных источников; использовать имеющиеся знания для целей проведения научных дискуссий и участия в них;

— **владеть:** передовыми приёмами построения аналитического дискурса и аргументированного представления его результатов; навыками научных исследований общественных процессов и отношений, методами сбора и обработки данных, в том числе с использованием технологий прикладного анализа данных; углублёнными теоретическими знаниями и практическими навыками организации научных исследований; способностью к самостоятельному обучению новым методам исследования, к изменению научного и научно-производственного профиля своей профессиональной деятельности, к изменению социокультурных и социальных условий деятельности; навыками использования полученных знаний для формулировки собственной позиции по актуальным проблемам общественных наук; приёмами и методами ведения дискуссии по проблемам современной науки.

3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Дисциплина «**Обработка естественного языка/ Natural Language Processing**» является дисциплиной по выбору Блока 1 «Дисциплины (модули)» части, формируемой участниками образовательных отношений, образовательной программы «Прикладной анализ данных». Код дисциплины по Учебному плану Б1.В.ДВ.02.01. Курс читается в четвертом семестре, форма промежуточной аттестации – зачет с оценкой.

Для успешного освоения данной дисциплины требуются знания, полученные в рамках прохождения обучения по следующим дисциплинам:

- Б1.О.02 Основы языка программирования Python,
- Б1.О.05 Программирование в R,
- Б1.О.06 Введение в статистику,
- Б1.О.03 Практический минимум,
- Б1.О.08 Текстовые данные.

Знания, умения и навыки, полученные при освоении данной дисциплины, применяются магистрантами в процессе прохождения Б2.О.01(У) Технологической (проектно-технологической) практики и выполнения выпускной квалификационной работы.

4. ОБЪЕМ ДИСЦИПЛИНЫ

Общая трудоемкость освоения дисциплины составляет 6 (шесть) зачетных единиц, 216 часов.

Таблица 2

Объем дисциплины

Типы учебных занятий и самостоятельная работа		Объем дисциплины									
		Всего	Модуль								
	1		2	3	4	5	6	7	8	9	10
Контактная работа обучающихся с преподавателем в соответствии с УП:		28	-	-	-	28	-	-	-	-	-
Лекции (Л)		14	-	-	-	14	-	-	-	-	-
Лабораторные занятия (ЛЗ)		14	-	-	-	14	-	-	-	-	-
Самостоятельная работа (СР)		188	-	-	-	188	-	-	-	-	-
Промежуточная аттестация	форма	Зачет с оценкой	-	-	-	Зачет с оценкой	-	-	-	-	-
	час.	-	-	-	-	-	-	-	-	-	-
Общая трудоемкость дисциплины (час./з.е.)		216/6	-	-	-	216/6	-	-	-	-	-

5. СОДЕРЖАНИЕ И СТРУКТУРА ДИСЦИПЛИНЫ

Содержание дисциплины соотносится с планируемыми результатами обучения по дисциплине: через задачи, формируемые компетенции и их компоненты (знания, умения, навыки – далее ЗУВ) по средствам индикаторов достижения компетенций в соответствии с Таблицей 3.

5.1 Содержание дисциплины

Таблица 3

Содержание дисциплины

№ п/п	Наименование тем (разделов)	Содержание тем (разделов)	Коды компетенций	Индикаторы компетенций (в соот.с Таблицей 1)	Коды ЗУВ (в соот.с Таблицей 1)
1	Введение	Краткая история дисциплины, основные инструменты и приложения	УК-1 УК-6 ПК-2 ПК-5 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ПК-2.1. ИД.ПК-2.2. ИД.ПК-2.3. ИД.ПК-2.4. ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	З (УК-1) У (УК-1) В (УК-1) З (УК-6) У (УК-6) В (УК-6) З (ПК-2) У (ПК-2) В (ПК-2) З (ПК-5) У (ПК-5) В (ПК-5) З (ПК-6) У (ПК-6) В (ПК-6)
2	Морфологический анализ	Основные подходы, данные, инструменты.	УК-1 УК-6 ПК-2 ПК-5 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5.	З (УК-1) У (УК-1) В (УК-1) З (УК-6) У (УК-6)

№ п/п	Наименование тем (разделов)	Содержание тем (разделов)	Коды компетенций	Индикаторы компетенций (в соот.с Таблицей 1)	Коды ЗУВ (в соот.с Таблицей 1)
				ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ПК-2.1. ИД.ПК-2.2. ИД.ПК-2.3. ИД.ПК-2.4 ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	В (УК-6) З (ПК-2) У (ПК-2) В (ПК-2) З (ПК-5) У (ПК-5) В (ПК-5) З (ПК-6) У (ПК-6) В (ПК-6)
3	Языковые модели	Приложения, данные, оценки качества. Языковые модели на основе n-грамм, сглаживание. Нейронные языковые модели.	УК-1 УК-6 ПК-2 ПК-5 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ПК-2.1. ИД.ПК-2.2. ИД.ПК-2.3. ИД.ПК-2.4 ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	З (УК-1) У (УК-1) В (УК-1) З (УК-6) У (УК-6) В (УК-6) З (ПК-2) У (ПК-2) В (ПК-2) З (ПК-5) У (ПК-5) В (ПК-5) З (ПК-6) У (ПК-6) В (ПК-6)
4	Анализ тональности	Постановка задачи, приложения, данные, методы решения	УК-1 УК-6 ПК-2 ПК-5 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ПК-2.1. ИД.ПК-2.2. ИД.ПК-2.3. ИД.ПК-2.4 ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	З (УК-1) У (УК-1) В (УК-1) З (УК-6) У (УК-6) В (УК-6) З (ПК-2) У (ПК-2) В (ПК-2) З (ПК-5) У (ПК-5) В (ПК-5) З (ПК-6) У (ПК-6) В (ПК-6)

№ п/п	Наименование тем (разделов)	Содержание тем (разделов)	Коды компетенций	Индикаторы компетенций (в соот.с Таблицей 1)	Коды ЗУВ (в соот.с Таблицей 1)
5	Синтаксический анализ	Формализмы для представления синтаксической структуры, данные, подходы к решению, оценка качества.	УК-1 УК-6 ПК-2 ПК-5 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ПК-2.1. ИД.ПК-2.2. ИД.ПК-2.3. ИД.ПК-2.4. ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	З (УК-1) У (УК-1) В (УК-1) З (УК-6) У (УК-6) В (УК-6) З (ПК-2) У (ПК-2) В (ПК-2) З (ПК-5) У (ПК-5) В (ПК-5) З (ПК-6) У (ПК-6) В (ПК-6)
6	Извлечение информации	Именованные сущности, данные для обучения и тестирования. Методы на основе машинного обучения; методы на основе рекуррентных нейронных сетей.	УК-1 УК-6 ПК-2 ПК-5 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ПК-2.1. ИД.ПК-2.2. ИД.ПК-2.3. ИД.ПК-2.4. ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	З (УК-1) У (УК-1) В (УК-1) З (УК-6) У (УК-6) В (УК-6) З (ПК-2) У (ПК-2) В (ПК-2) З (ПК-5) У (ПК-5) В (ПК-5) З (ПК-6) У (ПК-6) В (ПК-6)
7	Автоматическое реферирование	Различные постановки задачи и методы, данные для обучения, автоматическое тестирование на основе метрики ROUGE, нейронные сети для автоматического реферирования.	УК-1 УК-6 ПК-2 ПК-5 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ПК-2.1. ИД.ПК-2.2. ИД.ПК-2.3. ИД.ПК-2.4. ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4.	З (УК-1) У (УК-1) В (УК-1) З (УК-6) У (УК-6) В (УК-6) З (ПК-2) У (ПК-2) В (ПК-2) З (ПК-5) У (ПК-5) В (ПК-5) З (ПК-6) У (ПК-6) В (ПК-6)

№ п/п	Наименование тем (разделов)	Содержание тем (разделов)	Коды компетенций	Индикаторы компетенций (в соот.с Таблицей 1)	Коды ЗУВ (в соот.с Таблицей 1)
				ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	
8	Вычислительная семантика	Семантические ресурсы (WordNet), дистрибутивные семантические модели, контекстуализированные векторные представления.	УК-1 УК-6 ПК-2 ПК-5 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ПК-2.1. ИД.ПК-2.2. ИД.ПК-2.3. ИД.ПК-2.4 ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	З (УК-1) У (УК-1) В (УК-1) З (УК-6) У (УК-6) В (УК-6) З (ПК-2) У (ПК-2) В (ПК-2) З (ПК-5) У (ПК-5) В (ПК-5) З (ПК-6) У (ПК-6) В (ПК-6)
9	Ответы на вопросы	Понимание прочитанного (reading comprehension) и ответы по базе знаний. Методы и данные.	УК-1 УК-6 ПК-2 ПК-5 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ПК-2.1. ИД.ПК-2.2. ИД.ПК-2.3. ИД.ПК-2.4 ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	З (УК-1) У (УК-1) В (УК-1) З (УК-6) У (УК-6) В (УК-6) З (ПК-2) У (ПК-2) В (ПК-2) З (ПК-5) У (ПК-5) В (ПК-5) З (ПК-6) У (ПК-6) В (ПК-6)
10	Машинный перевод (МП)	На основе правил, статистический МП, нейронный МП. Оценка качества (BLEU и другие метрики).	УК-1 УК-6 ПК-2 ПК-5 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ПК-2.1. ИД.ПК-2.2. ИД.ПК-2.3. ИД.ПК-2.4	З (УК-1) У (УК-1) В (УК-1) З (УК-6) У (УК-6) В (УК-6) З (ПК-2) У (ПК-2) В (ПК-2) З (ПК-5) У (ПК-5) В (ПК-5) З (ПК-6)

№ п/п	Наименование тем (разделов)	Содержание тем (разделов)	Коды компетенций	Индикаторы компетенций (в соот.с Таблицей 1)	Коды ЗУВ (в соот.с Таблицей 1)
				ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	У (ПК-6) В (ПК-6)

5.2 Структура дисциплины

Таблица 4

Структура дисциплины

№ п/п	Наименование тем (разделов)	Объем дисциплины, час.				Форма текущего контроля успеваемости *, промежуточной аттестации
		Всего	Контактная работа обучающихся с преподавателем по типам учебных занятий в соответствии с УП		СР	
			Л	ЛЗ		
Очная форма обучения						
Тема 1	Введение	20	1	2	17	О, Д
Тема 2	Морфологический анализ	22	1	2	19	КЗ, О
Тема 3	Языковые модели	22	2	1	19	КЗ
Тема 4	Анализ тональности	22	2	1	19	ПЗ
Тема 5	Синтаксический анализ	22	2	1	19	КЗ, ПЗ
Тема 6	Извлечение информации	22	1	2	19	ПЗ
Тема 7	Автоматическое реферирование	22	1	2	19	КЗ
Тема 8	Вычислительная семантика	21	1	1	19	ПЗ, Д
Тема 9	Ответы на вопросы	21	1	1	19	ПЗ
Тема 10	Машинный перевод (МП)	22	2	1	19	КЗ
Промежуточная аттестация		-	-	-	-	Зачет с оценкой
Всего:		216/6	14	14	188	-

*Примечание: формы текущего контроля успеваемости: диспут (Д), опрос (О), практическое задание (ПЗ), контрольное задание (КЗ).

6. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ

6.1 Общие положения

Знания и навыки, полученные в результате лекций и семинарских занятий, закрепляются и развиваются в результате повторения материала, усвоенного в аудитории, путем чтения текстов и исследовательской литературы (из списков основной и дополнительной литературы) и их анализа.

Самостоятельная работа является важнейшей частью процесса высшего образования. Ее следует осознанно организовать, выделив для этого необходимое время и соответствующим образом организовав рабочее пространство. Важнейшим элементом самостоятельной работы является проработка материалов прошедших занятий (анализ конспектов, чтение рекомендованной литературы) и подготовка к следующим

лекциям/семинарским занятиям. Литературу, рекомендованную в программе курса, следует, по возможности, читать в течение всего семестра, концентрируясь на обусловленных программой курса темах.

Существенную часть самостоятельной работы магистранта представляет самостоятельное изучение вспомогательных учебно-методических изданий, лекционных конспектов, интернет-ресурсов и пр. Подготовка к семинарским занятиям, контрольному тесту также является важной формой работы магистранта. Самостоятельная работа может вестись как индивидуально, так и при содействии преподавателя.

6.2 Рекомендации по распределению учебного времени по видам самостоятельной работы и разделам дисциплины

Тема 1. Введение:

1.1. Изучение вопросов, представленных в списке тем лекций. Повторение изученного на предыдущих лекциях материала при подготовке к последующим лекциям – 8 часов.

1.2. Подготовка к лабораторным занятиям по предложенным темам, самостоятельное изучение рекомендованной литературы, повторение материала лекций – 9 часов. Итого: 17 часов.

Тема 2. Морфологический анализ:

2.1. Изучение вопросов, представленных в списке тем лекций. Повторение изученного на предыдущих лекциях материала при подготовке к последующим лекциям – 9 часов.

2.2. Подготовка к лабораторным занятиям по предложенным темам, самостоятельное изучение рекомендованной литературы, повторение материала лекций – 10 часов. Итого: 19 часов.

Тема 3. Языковые модели:

3.1. Изучение вопросов, представленных в списке тем лекций. Повторение изученного на предыдущих лекциях материала при подготовке к последующим лекциям – 9 часов.

3.2. Подготовка к лабораторным занятиям по предложенным темам, самостоятельное изучение рекомендованной литературы, повторение материала лекций – 10 часов. Итого: 19 часов.

Тема 4. Анализ тональности:

4.1. Изучение вопросов, представленных в списке тем лекций. Повторение изученного на предыдущих лекциях материала при подготовке к последующим лекциям – 9 часов.

4.2. Подготовка к лабораторным занятиям по предложенным темам, самостоятельное изучение рекомендованной литературы, повторение материала лекций – 10 часов. Итого: 19 часов.

Тема 5. Синтаксический анализ:

5.1. Изучение вопросов, представленных в списке тем лекций. Повторение изученного на предыдущих лекциях материала при подготовке к последующим лекциям – 9 часов.

5.2. Подготовка к лабораторным занятиям по предложенным темам, самостоятельное изучение рекомендованной литературы, повторение материала лекций – 10 часов. Итого: 19 часов.

Тема 6. Извлечение информации:

6.1. Изучение вопросов, представленных в списке тем лекций. Повторение изученного на предыдущих лекциях материала при подготовке к последующим лекциям – 9 часов.

6.2. Подготовка к лабораторным занятиям по предложенным темам, самостоятельное изучение рекомендованной литературы, повторение материала лекций – 10 часов. Итого: 19 часов.

Тема 7. Автоматическое реферирование:

7.1. Изучение вопросов, представленных в списке тем лекций. Повторение изученного на предыдущих лекциях материала при подготовке к последующим лекциям – 9 часов.

7.2. Подготовка к лабораторным занятиям по предложенным темам, самостоятельное изучение рекомендованной литературы, повторение материала лекций – 10 часов. Итого: 19 часов.

Тема 8. Вычислительная семантика:

8.1. Изучение вопросов, представленных в списке тем лекций. Повторение изученного на предыдущих лекциях материала при подготовке к последующим лекциям – 9 часов.

8.2. Подготовка к лабораторным занятиям по предложенным темам, самостоятельное изучение рекомендованной литературы, повторение материала лекций – 10 часов. Итого: 19 часов.

Тема 9. Ответы на вопросы:

9.1. Изучение вопросов, представленных в списке тем лекций. Повторение изученного на предыдущих лекциях материала при подготовке к последующим лекциям – 9 часов.

9.2. Подготовка к лабораторным занятиям по предложенным темам, самостоятельное изучение рекомендованной литературы, повторение материала лекций – 10 часов. Итого: 19 часов.

Тема 10. Машинный перевод (МП):

10.1. Изучение вопросов, представленных в списке тем лекций. Повторение изученного на предыдущих лекциях материала при подготовке к последующим лекциям – 9 часов.

10.2. Подготовка к лабораторным занятиям по предложенным темам, самостоятельное изучение рекомендованной литературы, повторение материала лекций – 10 часов. Итого: 19 часов.

6.3 Перечень основных вопросов по изучаемым темам для самостоятельной работы обучающихся по дисциплине

Вопросы для самостоятельной подготовки по темам дисциплины:

1. Исправление опечаток.
2. Распознавание звучащей речи.
3. Морфологический анализ: основные подходы.
4. Нейронные языковые модели.
5. Анализ тональности: постановка задачи, приложения, данные, методы решения.
6. Синтаксический анализ: формализмы для представления синтаксической структуры, данные, подходы к решению, оценка качества.
7. Методы извлечения информации на основе машинного обучения.

8. Нейронные сети для автоматического реферирования.
9. Дистрибутивные семантические модели.
10. Машинный перевод: методы оценки качества.

6.4 Перечень литературы для самостоятельной работы обучающегося:

1. Ганегедара, Т. Обработка естественного языка с TensorFlow : монография / Т. Ганегедара ; пер. с англ. В. С. Яценкова. - Москва : ДМК Пресс, 2020. - 382 с. - ISBN 978-5-97060-756-5. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1094940> .
2. Дуглас А., Л. Люк, Д.А. Анализ сетей (графов) в среде R. Руководство пользователя / Д.А. Люк ; пер. с англ. А.В. Груздева. - Москва : ДМК Пресс, 2017. - 250 с. - ISBN 978-5-97060-428-1. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1028090>
3. Иванова, Н.Ю. Системное и прикладное программное обеспечение : учебное пособие / Н.Ю. Иванова, В.Г. Маняхина ; Московский педагогический государственный университет. – Москва : Прометей, 2011. – 202 с. : ил.,табл., схем. – Режим доступа: по подписке. – URL: <http://biblioclub.ru/index.php?page=book&id=105792>
4. Кравцова, Е. Д. Логика и методология научных исследований [Электронный ресурс] : учеб. пособие / Е. Д. Кравцова, А. Н. Городищева. – Красноярск : Сиб. федер. ун-т, 2014. – 168 с. - ISBN 978-5-7638-2946-4 - Режим доступа: <http://znanium.com/catalog.php?bookinfo=507377>

6.5 Перечень учебно-методического обеспечения для самостоятельной работы

Для обеспечения самостоятельной работы магистрантов по дисциплине «**Обработка естественного языка/ Natural Language Processing**» разработано учебно-методическое обеспечение в составе:

1. Контрольные задания для подготовки к процедурам текущего контроля (п. 7.2 Рабочей программы).
2. Типовые задания для подготовки к промежуточной аттестации (п. 7.4 Рабочей программы).
3. Рекомендуемые основная, дополнительная литература, Интернет-ресурсы и справочные системы (п. 8, 9 Рабочей программы).
4. Рабочая программа дисциплины размещена в электронной информационно-образовательной среде Университета на электронном учебно-методическом ресурсе АНООВО «ЕУСПб» — образовательном портале LMS Sakai — Sakai@EU.

7. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕЙ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ

7.1 Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе текущей аттестации

Информация о содержании и процедуре текущего контроля успеваемости, методике оценивания знаний, умений и навыков обучающегося в ходе текущего контроля доводятся научно-педагогическими работниками Университета до сведения обучающегося на первом занятии по данной дисциплине.

Текущий контроль предусматривает подготовку магистрантов к каждому лабораторному занятию, участие в опросах, диспутах, подготовку практических заданий, выполнение контрольных заданий, активное слушание на лекциях. Магистрант должен присутствовать на семинарских занятиях, отвечать на поставленные вопросы, показывая, что прочитал разбираемую литературу, представлять содержательные реплики по обсуждаемым вопросам.

Текущий контроль проводится в форме устных опросов и оценивания участия магистрантов в проходящих диспутах, оценивания выполненных практических заданий, контрольных работ, демонстрирующих степень знакомства с дополнительной литературой.

**Показатели, критерии и оценивание компетенций и индикаторов их
достижения в процессе текущей аттестации**

Наименование тем (разделов)	Коды компетенций	Индикаторы компетенций	Коды ЗУВ (в соот. с Таблицей 1)	Формы текущего контроля успеваемости	Результаты текущего контроля
Введение	УК-1 УК-6 ПК-2 ПК-5 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ПК-2.1. ИД.ПК-2.2. ИД.ПК-2.3. ИД.ПК-2.4. ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	З (УК-1) У (УК-1) В (УК-1) З (УК-6) У (УК-6) В (УК-6) З (ПК-2) У (ПК-2) В (ПК-2) З (ПК-5) У (ПК-5) В (ПК-5) З (ПК-6) У (ПК-6) В (ПК-6)	Опрос 1 Диспут 1	зачтено/ не зачтено зачтено/ не зачтено
Морфологический анализ	УК-1 УК-6 ПК-2 ПК-5 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ПК-2.1. ИД.ПК-2.2. ИД.ПК-2.3. ИД.ПК-2.4. ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	З (УК-1) У (УК-1) В (УК-1) З (УК-6) У (УК-6) В (УК-6) З (ПК-2) У (ПК-2) В (ПК-2) З (ПК-5) У (ПК-5) В (ПК-5) З (ПК-6) У (ПК-6) В (ПК-6)	Контрольное задание 1 Контрольное задание 2 Опрос 2 Контрольное задание 3	зачтено/ не зачтено зачтено/ не зачтено зачтено/ не зачтено зачтено/ не зачтено
Языковые модели	УК-1 УК-6 ПК-2 ПК-5 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ПК-2.1. ИД.ПК-2.2. ИД.ПК-2.3. ИД.ПК-2.4.	З (УК-1) У (УК-1) В (УК-1) З (УК-6) У (УК-6) В (УК-6) З (ПК-2) У (ПК-2) В (ПК-2) З (ПК-5) У (ПК-5) В (ПК-5) З (ПК-6)	Контрольное задание 4 Контрольное задание 5 Контрольное задание 6	зачтено/ не зачтено зачтено/ не зачтено зачтено/ не зачтено

Наименование тем (разделов)	Коды компетенций	Индикаторы компетенций	Коды ЗУВ (в соот. С Таблицей 1)	Формы текущего контроля успеваемости	Результаты текущего контроля
		ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	У (ПК-6) В (ПК-6)		
Анализ тональности	УК-1 УК-6 ПК-2 ПК-5 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ПК-2.1. ИД.ПК-2.2. ИД.ПК-2.3. ИД.ПК-2.4. ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	3 (УК-1) У (УК-1) В (УК-1) 3 (УК-6) У (УК-6) В (УК-6) 3 (ПК-2) У (ПК-2) В (ПК-2) 3 (ПК-5) У (ПК-5) В (ПК-5) 3 (ПК-6) У (ПК-6) В (ПК-6)	Практическое задание 1	зачтено/ не зачтено
Синтаксический анализ	УК-1 УК-6 ПК-2 ПК-5 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ПК-2.1. ИД.ПК-2.2. ИД.ПК-2.3. ИД.ПК-2.4. ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	3 (УК-1) У (УК-1) В (УК-1) 3 (УК-6) У (УК-6) В (УК-6) 3 (ПК-2) У (ПК-2) В (ПК-2) 3 (ПК-5) У (ПК-5) В (ПК-5) 3 (ПК-6) У (ПК-6) В (ПК-6)	Контрольное задание 7 Контрольное задание 8 Практическое задание 2	зачтено/ не зачтено зачтено/ не зачтено зачтено/ не зачтено
Извлечение информации	УК-1 УК-6 ПК-2 ПК-5 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4.	3 (УК-1) У (УК-1) В (УК-1) 3 (УК-6) У (УК-6) В (УК-6) 3 (ПК-2) У (ПК-2) В (ПК-2)	Практическое задание 3	зачтено/ не зачтено

Наименование тем (разделов)	Коды компетенций	Индикаторы компетенций	Коды ЗУВ (в соот. С Таблицей 1)	Формы текущего контроля успеваемости	Результаты текущего контроля
		ИД.ПК-2.1. ИД.ПК-2.2. ИД.ПК-2.3. ИД.ПК-2.4 ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	З (ПК-5) У (ПК-5) В (ПК-5) З (ПК-6) У (ПК-6) В (ПК-6)		
Автоматическое реферирование	УК-1 УК-6 ПК-2 ПК-5 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ПК-2.1. ИД.ПК-2.2. ИД.ПК-2.3. ИД.ПК-2.4 ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	З (УК-1) У (УК-1) В (УК-1) З (УК-6) У (УК-6) В (УК-6) З (ПК-2) У (ПК-2) В (ПК-2) З (ПК-5) У (ПК-5) В (ПК-5) З (ПК-6) У (ПК-6) В (ПК-6)	Контрольное задание 9	зачтено/ не зачтено
Вычислительная семантика	УК-1 УК-6 ПК-2 ПК-5 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ПК-2.1. ИД.ПК-2.2. ИД.ПК-2.3. ИД.ПК-2.4 ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	З (УК-1) У (УК-1) В (УК-1) З (УК-6) У (УК-6) В (УК-6) З (ПК-2) У (ПК-2) В (ПК-2) З (ПК-5) У (ПК-5) В (ПК-5) З (ПК-6) У (ПК-6) В (ПК-6)	Диспут 2 Практическое задание 4	зачтено/ не зачтено зачтено/ не зачтено
Ответы на вопросы	УК-1 УК-6 ПК-2 ПК-5 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5.	З (УК-1) У (УК-1) В (УК-1) З (УК-6) У (УК-6)	Практическое задание 5	зачтено/ не зачтено зачтено/ не зачтено

Наименование тем (разделов)	Коды компетенций	Индикаторы компетенций	Коды ЗУВ (в соот. С Таблицей 1)	Формы текущего контроля успеваемости	Результаты текущего контроля
		ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ПК-2.1. ИД.ПК-2.2. ИД.ПК-2.3. ИД.ПК-2.4 ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	В (УК-6) 3 (ПК-2) У (ПК-2) В (ПК-2) 3 (ПК-5) У (ПК-5) В (ПК-5) 3 (ПК-6) У (ПК-6) В (ПК-6)		
Машинный перевод (МП)	УК-1 УК-6 ПК-2 ПК-5 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ПК-2.1. ИД.ПК-2.2. ИД.ПК-2.3. ИД.ПК-2.4 ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	3 (УК-1) У (УК-1) В (УК-1) 3 (УК-6) У (УК-6) В (УК-6) 3 (ПК-2) У (ПК-2) В (ПК-2) 3 (ПК-5) У (ПК-5) В (ПК-5) 3 (ПК-6) У (ПК-6) В (ПК-6)	Контрольное задание 10	зачтено/ не зачтено

Таблица 6

Критерии оценивания

Формы текущего контроля успеваемости	Критерии оценивания
Опрос	<p>Ответ отсутствует или является односложным, или содержит существенные ошибки – не зачтено</p> <p>Магистрант в ответах демонстрирует знание всех теоретических положений, (развернуто) отвечает на все поставленные вопросы, предлагает обоснования при ответе на все или большинство поставленных вопросов; несущественные ошибки не снижают качество ответа — зачтено</p>
Диспут	<p>Пассивность, участие без представления аргументов и обоснования точки зрения, несформированность навыков профессиональной коммуникации в группе — не зачтено</p> <p>Представление аргументированной научной позиции, обоснование точки зрения в диспуте, демонстрация навыков профессиональной коммуникации в группе — зачтено</p>
Практическое задание	<p>магистрант выполняет задание частично или с существенными недочетами (некорректно сформулирован исследовательский вопрос, не определены основные агенты, некорректно выбраны методы</p>

Формы текущего контроля успеваемости	Критерии оценивания
	исследования, требования к содержанию, структуре, логике, аргументации, оформлению не выполнены) – не зачтено, полное и правильное выполнение задания в соответствии с требованиями к содержанию, структуре, логике, аргументации, оформлению с возможным небольшим количеством погрешностей (например, плохо выдержанная структура текста, недостаточная аргументация отдельных тезисов) – зачтено
Контрольное задание	магистрант выполняет задание частично или с существенными недочетами (некорректно сформулирован исследовательский вопрос, не определены основные агенты, некорректно выбраны методы исследования, требования к содержанию, структуре, логике, аргументации, оформлению не выполнены) – не зачтено, полное и правильное выполнение задания в соответствии с требованиями к содержанию, структуре, логике, аргументации, оформлению с возможным небольшим количеством погрешностей (например, плохо выдержанная структура текста, недостаточная аргументация отдельных тезисов) – зачтено

7.2 Контрольные задания для текущей аттестации

Примерный материал опросов, диспутов, практических заданий:

Тема 1. Введение.

Опрос 1.

1. Какие основные инструменты и приложения используются при обработке естественного языка?
2. История развития методов обработки естественного языка?

Диспут 1.

Как технологии обработки естественного языка меняют повседневную жизнь людей.

Тема 2. Морфологический анализ.

Контрольное задание 1.

Три слова в списке имеют одинаковый тип словоизменения, одно – отличается. Отметьте это слово.

- 1) чело
- 2) табло
- 3) сверло
- 4) кайло.

Контрольное задание 2.

Какие термины относятся к морфологической обработке? Выберите номера правильных ответов:

- 1) лемматизация;
- 2) лемминг;
- 3) стемминг;
- 4) прокрастинация;
- 5) парсинг.

Опрос 2.

1. Каковы основные подходы к морфологическому анализу?
2. Какова основная методика отбора данных для морфологического анализа?
3. Каковы основные инструменты морфологического анализа?

Контрольное задание 3.

Отметьте правильный набор граммем (в нотации НКРЯ/mystem) для выделенного слова в предложении:

Лично я ловлю покемонов не одобряю.

- 1) S,жен,неод=вин,ед
- 2) V,несов,пе=непрош,ед,изъяв,1-л
- 3) V,несов,пе=ед,пов,2-л
- 4) S,муж,неод=им,ед.

Тема 3. Языковые модели.

Контрольное задание 4.

Постройте биграммную языковую модель на основе корпуса:

<s> Вася любит мороженое </s>

<s> Лена любит малину </s>

<s> Вася любит Лену </s>

<s> Георгий ест мороженое </s>

<s> Лена рисует яблоко </s>

<s> Георгий любит Катю </s>

<s> Георгий любит смотреть, как Лена ест мороженое </s>

Упорядочите предложения по убыванию оценок вероятностей на основе построенной языковой модели.

1. <s> Лена любит мороженое </s>
2. <s> Лена рисует малину </s>
3. <s> Вася любит Катю </s>.

Контрольное задание 5.

Постройте биграммную языковую модель на основе корпуса:

<s> Вася любит мороженое </s>

<s> Лена любит малину </s>

<s> Вася любит Лену </s>

<s> Георгий ест мороженое </s>

<s> Лена рисует яблоко </s>

<s> Георгий любит Катю </s>

<s> Георгий любит смотреть, как Лена ест мороженое </s>

Вычислите перплексию модели на предложении:

<s> Георгий любит малину </s>

Обратите внимание, что для вычисления вероятности предложения (и, соответственно, перплексии) используются вероятности 4-х биграмм ($N=4$ в формуле вычисления перплексии).

Контрольное задание 6.

Пусть у нас есть корпус, содержащий 10000 предложений, размер словаря – 1500 уникальных слов (включая специальные «слова» – маркеры начала и конца предложений).

Некоторые частоты униграмм:

ем 100

дуриан 1

и 5000

не 3000

морщусь 50

и биграмм:

<s> ем 20
ем дуриан 0
дуриан и 0
и не 300
не морщусь 15
морщусь </s> 5.

Примените сглаживание Лапласа ($\alpha=1$, сглаживание «+1») для оценки вероятностей биграмм и оцените на их основе вероятность предложения <s> ем дуриан и не морщусь </s>
В качестве ответа введите натуральный логарифм оценки вероятности предложения.

Тема 4. Анализ тональности.

Практическое задание 1

Реализуйте анализатор тональности одним из способов. Данные – английские предложения из отзывов о фильмах из Stanford Sentiment Treebank (из датасета взяты только целые предложения; 5-уровневая разметка приведена к трехуровневой – негативный, нейтральный, позитивный). Вы можете использовать подход на основе словаря тонально окрашенных слов (например, SentiWords) или обучить классификатор на тренировочных данных. Вы можете использовать тренировочные данные и в первом случае – чтобы подобрать пороги для классификации предложений на основе весов словаря. Постройте матрицу ошибок (confusion matrix) на тестовом наборе. Оцените правильность (accuracy, доля правильно классифицированных предложений) классификатора на тестовом наборе. Проанализируйте неверно классифицированные предложения, сделайте предположения о причинах неверной классификации, предложите улучшения.

Тема 5. Синтаксический анализ.

Контрольное задание 7.

Выберите правильный разбор на составляющие предложения:

Советник губернатора Чукотского автономного округа Романа Абрамовича Роман Копин победил на повторных выборах главы администрации Чаунского района

1. [[[[Советник [[губернатора [Чукотского [автономного округа]]] [Романа Абрамовича]]] [Роман Копин]] [победил [на [повторных [выборах [главы [администрации [Чаунского района]]]]]]]]]]]

2. [[[[Советник [[губернатора [Чукотского [автономного округа]]] [Романа Абрамовича]]] [Роман Копин]] [[[[победил на] [повторных [выборах [главы [администрации [Чаунского района]]]]]]]]]]]

3. [[[[[[Советник губернатора] [Чукотского [автономного округа]]] [Романа Абрамовича]] [Роман Копин]] [победил [на [[повторных выборах] [главы [администрации [Чаунского района]]]]]]]]]]]

4. [[[[Советник [[губернатора [Чукотского [автономного округа]]] [Романа Абрамовича]]] [[Роман Копин] победил]] [на [[повторных выборах] [главы [администрации [Чаунского района]]]]]]]]]

[[[[Советник [[губернатора Чукотского] [[автономного округа] [Романа Абрамовича]]]]] [Роман Копин]] [победил [на [повторных [выборах [главы [администрации [Чаунского района]]]]]]]]]]]

Контрольное задание 8.

Выберите правильный разбор предложения в терминах зависимостей:

0 root 1 Активно 2 обсуждается 3 роль 4 нашей 5 страны 6 в 7 современном 8 быстро 9 меняющемся 10 мире, 11 проходящем 12 через 13 переломный 14 этап.

В приведенных ниже вариантах разбора каждая пара – зависимость (хозяин, слуга), знаки препинания не учитываются, используется традиционный подход при установлении зависимостей с участием предлогов (не как в universal dependencies).

1. (0, 3) (3, 2) (2, 1) (3, 5) (5, 4) (3, 6) (6, 10) (10, 7) (10, 9) (9, 8) (10, 11) (11, 12) (12, 14) (14, 13)
2. (0, 2) (2, 1) (2, 3) (3, 5) (5, 4) (3, 6) (6, 10) (10, 7) (10, 9) (9, 8) (10, 11) (11, 12) (12, 14) (14, 13)
3. (0, 2) (2, 1) (2, 3) (3, 5) (5, 4) (5, 6) (6, 10) (10, 7) (10, 9) (9, 8) (10, 11) (11, 12) (12, 14) (14, 13)
4. (0, 2) (2, 1) (2, 3) (3, 5) (5, 4) (3, 6) (6, 7) (7, 8) (8, 9) (9, 10) (10, 11) (11, 12) (12, 14) (14, 13)
5. (0, 2) (2, 3) (3, 5) (5, 4) (3, 6) (6, 10) (10, 7) (10, 9) (2, 8) (10, 11) (11, 1) (11, 12) (12, 14) (14, 13)

Практическое задание 2

Постройте деревья зависимостей для 2000 предложений «Войны и Мира» и «Братьев Карамазовых» с помощью библиотеки Stanza.

Оцените производительность библиотеки (предложения/с).

Проведите ручную оценку качества разбора на случайных 20 предложениях (по 10 из каждой книги).

Посчитайте среднюю глубину дерева разбора для каждого из романов. Посчитайте корреляцию между длиной предложения в словах и глубиной дерева разбора. Оцените на нескольких примерах, насколько глубина дерева соответствует субъективной сложности понимания предложения.

Тема 6. Извлечение информации.

Практическое задание 3

Примените и оцените модуль извлечения именованных сущностей Natasha (<https://github.com/natasha/natasha>). Для тестирования используйте текст и соответствующую разметку. Тестовые данные содержат только разметку для людей (PER) и организаций (ORG). Рассчитайте F1 для каждого типа сущностей и общее значение F1.

Тема 7. Автоматическое реферирование.

Контрольное задание 9.

Пусть у нас есть два «реферата-образца»:

1: карп лещ лещ щука сазан

2: лещ карп лещ сазан плотва

Упорядочите «рефераты» ниже по убыванию значения ROUGE-2.

- 1) лещ карп лещ сазан плотва
- 2) плотва сазан лещ карп сазан
- 3) щука сазан карп лещ сазан
- 4) карп карась лещ окунь лещ
- 5) щука сазан лещ карп сазан

Тема 8. Вычислительная семантика.

Диспут 2.

Возможные негативные последствия использования больших предобученных моделей для генерации текста.

Практическое задание 4

Реализуйте генератор юмора по мотивам работы Alessandro Valitutti et al. “Let Everything Turn Well in Your Wife”: Generation of Adult Humor Using Lexical Constraints, 2013. На входе генератора – новостной заголовок, в котором надо заменить одно слово. Предлагаемый алгоритм:

1. Проведите разбор предложения с помощью библиотеки Stanza. На основе результатов разбора выберете слово-кандидат на замену.
2. Найдите антоним для слова в WordNet (используйте интерфейс библиотеки NLTK).
3. Если антоним не нашлся, то найдите несколько слов, близких по звучанию или рифму с помощью datamuse api.
4. Получите вектора fasttext для начального слова и вариантов замены. Среди этих слов найдите самое далекое по косинусному расстоянию.

Оцените 20 модификаций по шкале от 0 (совсем не смешно) до 3 (очень смешно), приведите среднюю оценку.

Тема 9. Ответы на вопросы.

Практическое задание 5

Обучите модель для выделения ответа на вопрос из параграфа на тренировочной части русскоязычных данных TuDi QA. Оцените качество модели на данных для настройки (development set, EM/F1). Проанализируйте случаи, с которыми модель справилась хуже всего. Сделайте предположение, в чем сложность этих случаев.

Тема 10. Машинный перевод (МП).

Контрольное задание 10

Для фразы

Call me what instrument you will, though you can fret me, you cannot play upon me.

Есть два образцовых перевода (для простоты знаки препинания в образцах и вариантах перевода удалены):

- *назовите меня каким угодно инструментом вы хоть и можете меня терзать но играть на мне не можете*
- *объявите меня каким угодно инструментом вы можете расстроить меня но играть на мне нельзя*

Упорядочите варианты перевода ниже по убыванию BLEU-2 (метрика на основе униграмм и биграмм). Помните, что стандартная метрика BLEU не предполагает лемматизацию текстов.

- *позвони мне на каком инструменте вы будете хотя вы можете беспокоиться меня но вы не можете играть на мне*
- *назовите мне какой инструмент вы хотите хотя можете меня беспокоить но вы не можете играть на меня*
- *позвони мне какой инструмент ты будешь хотя ты можешь меня волновать но ты не можешь играть на меня*
- *назовите меня какой инструмент вы будете хотя вы можете раздражать меня все же вы не можете играть на меня*
- *считай меня чем тебе угодно ты можешь мучить меня но не играть мною*
- *назови меня каким угодно инструментом ты можешь меня расстроить но не играть на мне.*

7.3 Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе промежуточной аттестации

Форма промежуточной аттестации – зачет с оценкой, выставляемый на основе письменной работы (эссе).

Перед зачетом с оценкой проводится консультация, на которой преподаватель отвечает на вопросы магистрантов.

В результате промежуточного контроля знаний студенты получают аттестацию по дисциплине.

Таблица 7

Показатели, критерии и оценивание компетенций и индикаторов их достижения в процессе промежуточной аттестации

Форма промежуточной аттестации/вид промежуточной аттестации	Коды компетенций	Индикаторы компетенций (в соот. С Таблицей 1)	Коды ЗУВ (в соот. С Таблицей 1)	Критерии оценивания	Оценка
Зачет с оценкой/ Письменная работа (эссе)	УК-1 УК-6 ПК-2 ПК-5 ПК-6	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5. ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4. ИД.ПК-2.1. ИД.ПК-2.2. ИД.ПК-2.3. ИД.ПК-2.4 ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4. ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	З (УК-1) У (УК-1) В (УК-1) З (УК-6) У (УК-6) В (УК-6) З (ПК-2) У (ПК-2) В (ПК-2) З (ПК-5) У (ПК-5) В (ПК-5) З (ПК-6) У (ПК-6) В (ПК-6)	Эссе соответствует следующим требованиям: сформулирован исследовательский вопрос, корректно выбраны методы и собраны данные, тема раскрыта, соблюдены структура и научный стиль, сформулированы выводы, аргументация убедительна, правильно оформлен библиографический аппарат и т.д. Магистрант демонстрирует: глубокое усвоение программного материала; изложение данного материала исчерпывающе, последовательно, четко; умение делать обоснованные выводы; соблюдение норм устной и письменной литературной речи. Эссе представлено на защите на высоком профессиональном уровне	Зачтено, отлично
				Эссе соответствует следующим требованиям: сформулирован исследовательский вопрос, корректно выбраны методы и собраны данные, тема раскрыта, соблюдены структура и научный стиль, сформулированы выводы, аргументация убедительна, правильно оформлен библиографический аппарат и т.д. Магистрант демонстрирует: усвоение программного материала; изложение данного материала последовательно;	Зачтено, хорошо

Форма промежуточной аттестации/вид промежуточной аттестации	Коды компетенций	Индикаторы компетенций (в соот. С Таблицей 1)	Коды ЗУВ (в соот. С Таблицей 1)	Критерии оценивания	Оценка
				<p>умение делать обоснованные выводы; соблюдение норм устной и письменной литературной речи, но с присутствием некоторых неточностей в формулировках. Эссе представлено на защите на достаточно профессиональном уровне</p>	
				<p>Эссе соответствует следующим требованиям: сформулирован исследовательский вопрос, корректно выбраны методы и собраны данные, тема раскрыта, соблюдены структура и научный стиль, сформулированы выводы, аргументация убедительна, правильно оформлен библиографический аппарат и т.д. Магистрант демонстрирует: минимально достаточный уровень усвоения программного материала; изложение данного материала; умение делать обоснованные выводы; соблюдение норм устной и письменной литературной речи. Эссе представлено на защите на удовлетворительном уровне</p>	Зачтено, удовлетворительно
				<p>Представленное эссе не отвечает предъявляемым требованиям (либо не предоставление эссе); Магистрант демонстрирует: незнание значительной части программного материала: наличие существенных ошибок в определениях, формулировках, понимании теоретических положений; бессистемность при ответе на поставленный вопрос; отсутствие в ответе логически корректного анализа, аргументации, классификации; наличие нарушений норм устной и письменной литературной речи</p>	Не зачтено, неудовлетворительно

Результаты сдачи промежуточной аттестации по направлениям подготовки уровня магистратуры оцениваются по стобалльной системе оценки в соответствии с Положением о формах, периодичности и порядке организации и проведения текущего контроля

успеваемости и промежуточной аттестации обучающихся в АНООВО «ЕУСПб» следующим образом согласно таблице 7а.

Таблица 7а

Система оценки знаний обучающихся

Пятибалльная (стандартная) система	Стобалльная система оценки	Бинарная система оценки
5 (отлично)	100-81	зачтено
4 (хорошо)	80-61	
3 (удовлетворительно)	60-41	
2 (неудовлетворительно)	40 и менее	не зачтено

Результаты промежуточного контроля по дисциплине, выраженные в оценках «зачтено, отлично», «зачтено, хорошо», «зачтено, удовлетворительно», показывают уровень сформированности у обучающегося компетенций по дисциплине в соответствии с картами компетенций образовательной программы «Прикладной анализ данных» по направлению подготовки 09.04.03 Прикладная информатика (уровень магистратуры).

Результаты промежуточного контроля по дисциплине, выраженные в оценках «не зачтено, неудовлетворительно», показывают несформированность у обучающегося компетенций по дисциплине в соответствии с картами компетенций образовательной программы «Прикладной анализ данных» по направлению подготовки 09.04.03 Прикладная информатика (уровень магистратуры).

7.4 Типовые задания к промежуточной аттестации

Письменная работа (эссе) представляет текст на 3000 слов в формате статьи (самостоятельно и аргументированно выбрать журнал) или policy memo на русском или английском языке. Требуется использовать описательные статистики, простые графики и/или базовые регрессионные модели, чтобы показать необычный паттерн, неожиданное распределение результатов или другой феномен из гуманитарных или социальных наук, нуждающийся в дальнейшем изучении.

Примерные темы письменной работы (эссе):

1. Сравнение качества и производительности морфологических библиотек для русского языка.
2. Сравнение качества и производительности библиотек анализа тональности для русского языка.
3. Сравнение качества и производительности библиотек для выделения именованных сущностей для русского языка.
4. Сравнение качества онлайн сервисов машинного перевода для пары русский-английский (одно направление).
5. Сравнительный анализ произведений Льва Толстого и Федора Достоевского с помощью стилометрических методов.
6. Гендерное смещение (gender bias) в дистрибутивных моделях русского языка (проанализировать 2-3 статические модели отсюда: <https://rusvectors.org/ru/models/>).
7. Сравнительный анализ двух моделей вопросно-ответного поиска для русского языка с помощью инструмента CheckList (<https://github.com/marcotcr/checklist>).
8. Анализ качества кросс-языкового переноса моделей вопросно-ответного поиска на данных без дообучения. Исходный английский набор данных – SquAD, тестирование – на данных TyDi QA.
9. Исследование переносимости моделей распознавания юмора: исследовать 2-3 метода классификации на 2-3 англоязычных наборах данных.
10. Анализ существующих систем вопросно-ответного поиска по базам знаний (deeppavlov, Qanswer) с помощью тестового набора данных RuBQ.
11. Систематический анализ качества генерации текстов с помощью модели ruGPT-3 для различных сценариев (<https://github.com/sberbank-ai/ru-gpts>).

7.5 Средства оценки индикаторов достижения компетенций

Таблица 8

Средства оценки индикаторов достижения компетенций

Коды компетенций	Индикаторы компетенций (в соот.с Таблицей 1)	Средства оценки (в соот. С Таблицами 5, 7)
УК-1	ИД.УК-1.1. ИД.УК-1.2. ИД.УК-1.3. ИД.УК-1.4. ИД.УК-1.5.	Опрос, диспут, практическое задание, контрольное задание, письменная работа (эссе)
УК-6	ИД.УК-6.1. ИД.УК-6.2. ИД.УК-6.3. ИД.УК-6.4.	Опрос, диспут, практическое задание, контрольное задание, письменная работа (эссе)
ПК-2	ИД.ПК-2.1. ИД.ПК-2.2. ИД.ПК-2.3. ИД.ПК-2.4	Опрос, диспут, практическое задание, контрольное задание, письменная работа (эссе)
ПК-5	ИД.ПК-5.1. ИД.ПК-5.2. ИД.ПК-5.3. ИД.ПК-5.4.	Опрос, диспут, практическое задание, контрольное задание, письменная работа (эссе)
ПК-6	ИД.ПК-6.1. ИД.ПК-6.2. ИД.ПК-6.3. ИД.ПК-6.4.	Опрос, диспут, практическое задание, контрольное задание, письменная работа (эссе)

Таблица 9

Описание средств оценки индикаторов достижения компетенций

Средства оценки (в соот. С Таблицами 5, 7)	Рекомендованный план выполнения работы
Опрос	<p>Магистрант в ходе подготовки и участия в опросе показывает наличие практической базы знаний в рамках дисциплины, необходимой для выполнения следующих действий в области профессиональной деятельности:</p> <ol style="list-style-type: none"> 1. Анализировать проблемную ситуацию, определять пробелы в информации, оценивать надёжность источников информации, разрабатывать стратегию решения проблемной ситуации на основе системного и междисциплинарного подходов, строить сценарии реализации стратегии, определяя возможные риски и предлагая пути их устранения 2. Оценивать свои ресурсы для успешного выполнения задания. Определять образовательные потребности и реализовывать с использованием инструментов непрерывного образования возможности развития профессиональных компетенций и социальных навыков, что позволяет выстраивать гибкую профессиональную траекторию 3. Определять состав аналитической группы проекта, представлять и организовывать обсуждение плана аналитических работ, распределять роли и аналитические работы по участникам аналитической группы проекта, поддерживает обратную связь, определять необходимые ресурсы 4. Анализировать новые направления научных исследований, ставить цели, обосновывать выбор методов анализа данных, формировать программы и непосредственно принимать участие в реализации мероприятий в рамках выполнения прикладного анализа данных при решении научных и общественных задачах 5. Анализировать новые направления развития и обосновывать перспективы применения технологий машинного обучения к решению конкретных общественных задач, формировать программы и непосредственно принимает участие в реализации мероприятий в рамках решения общественных задач по средствам технологий машинного обучения

Средства оценки (в соот. С Таблицами 5, 7)	Рекомендованный план выполнения работы
Диспут	<p>Магистрант в ходе подготовки и участия в диспуте показывает наличие практической базы знаний в рамках дисциплины, необходимой для выполнения следующих действий в области профессиональной деятельности:</p> <ol style="list-style-type: none"> 1. Анализировать проблемную ситуацию, определять пробелы в информации, оценивать надёжность источников информации, разрабатывать стратегию решения проблемной ситуации на основе системного и междисциплинарного подходов, строить сценарии реализации стратегии, определяя возможные риски и предлагая пути их устранения 2. Оценивать свои ресурсы для успешного выполнения задания. Определять образовательные потребности и реализовывать с использованием инструментов непрерывного образования возможности развития профессиональных компетенций и социальных навыков, что позволяет выстраивать гибкую профессиональную траекторию 3. Определять состав аналитической группы проекта, представлять и организовывать обсуждение плана аналитических работ, распределять роли и аналитические работы по участникам аналитической группы проекта, поддерживает обратную связь, определять необходимые ресурсы 4. Анализировать новые направления научных исследований, ставить цели, обосновывать выбор методов анализа данных, формировать программы и непосредственно принимать участие в реализации мероприятий в рамках выполнения прикладного анализа данных при решении научных и общественных задачах 5. Анализировать новые направления развития и обосновывать перспективы применения технологий машинного обучения к решению конкретных общественных задач, формировать программы и непосредственно принимает участие в реализации мероприятий в рамках решения общественных задач по средствам технологий машинного обучения
Практическое задание	<p>Магистрант в ходе подготовки и выполнения практического задания показывает наличие практической базы знаний в рамках дисциплины, необходимой для выполнения следующих действий в области профессиональной деятельности:</p> <ol style="list-style-type: none"> 1. Анализировать проблемную ситуацию, определять пробелы в информации, оценивать надёжность источников информации, разрабатывать стратегию решения проблемной ситуации на основе системного и междисциплинарного подходов, строить сценарии реализации стратегии, определяя возможные риски и предлагая пути их устранения 2. Оценивать свои ресурсы для успешного выполнения задания. Определять образовательные потребности и реализовывать с использованием инструментов непрерывного образования возможности развития профессиональных компетенций и социальных навыков, что позволяет выстраивать гибкую профессиональную траекторию 3. Определять состав аналитической группы проекта, представлять и организовывать обсуждение плана аналитических работ, распределять роли и аналитические работы по участникам аналитической группы проекта, поддерживает обратную связь, определять необходимые ресурсы 4. Анализировать новые направления научных исследований, ставить цели, обосновывать выбор методов анализа данных, формировать программы и непосредственно принимать участие в реализации мероприятий в рамках выполнения прикладного анализа данных при решении научных и общественных задачах 5. Анализировать новые направления развития и обосновывать перспективы применения технологий машинного обучения к решению конкретных общественных задач, формировать программы и непосредственно принимает участие в реализации мероприятий в рамках решения общественных задач по средствам технологий машинного обучения
Контрольное задание	<p>Магистрант в ходе подготовки и выполнения контрольного задания показывает наличие практической базы знаний в рамках дисциплины, необходимой для выполнения следующих действий в области профессиональной деятельности:</p> <ol style="list-style-type: none"> 1. Анализировать проблемную ситуацию, определять пробелы в информации, оценивать надёжность источников информации, разрабатывать стратегию решения проблемной ситуации на основе системного и междисциплинарного подходов, строить

Средства оценки (в соот. С Таблицами 5, 7)	Рекомендованный план выполнения работы
	<p>сценарии реализации стратегии, определяя возможные риски и предлагая пути их устранения</p> <p>2. Оценивать свои ресурсы для успешного выполнения задания. Определять образовательные потребности и реализовывать с использованием инструментов непрерывного образования возможности развития профессиональных компетенций и социальных навыков, что позволяет выстраивать гибкую профессиональную траекторию</p> <p>3. Определять состав аналитической группы проекта, представлять и организовывать обсуждение плана аналитических работ, распределять роли и аналитические работы по участникам аналитической группы проекта, поддерживает обратную связь, определять необходимые ресурсы</p> <p>4. Анализировать новые направления научных исследований, ставить цели, обосновывать выбор методов анализа данных, формировать программы и непосредственно принимать участие в реализации мероприятий в рамках выполнения прикладного анализа данных при решении научных и общественных задачах</p> <p>Анализировать новые направления развития и обосновывать перспективы применения технологий машинного обучения к решению конкретных общественных задач, формировать программы и непосредственно принимает участие в реализации мероприятий в рамках решения общественных задач по средствам технологий машинного обучения</p>
Письменная работа (эссе)	<p>Магистрант в ходе подготовки и выполнения письменной работы (эссе) показывает наличие практической базы знаний в рамках дисциплины, необходимой для выполнения следующих действий в области профессиональной деятельности:</p> <p>1. Анализировать проблемную ситуацию, определять пробелы в информации, оценивать надёжность источников информации, разрабатывать стратегию решения проблемной ситуации на основе системного и междисциплинарного подходов, строить сценарии реализации стратегии, определяя возможные риски и предлагая пути их устранения</p> <p>2. Оценивать свои ресурсы для успешного выполнения задания. Определять образовательные потребности и реализовывать с использованием инструментов непрерывного образования возможности развития профессиональных компетенций и социальных навыков, что позволяет выстраивать гибкую профессиональную траекторию</p> <p>3. Определять состав аналитической группы проекта, представлять и организовывать обсуждение плана аналитических работ, распределять роли и аналитические работы по участникам аналитической группы проекта, поддерживает обратную связь, определять необходимые ресурсы</p> <p>4. Анализировать новые направления научных исследований, ставить цели, обосновывать выбор методов анализа данных, формировать программы и непосредственно принимать участие в реализации мероприятий в рамках выполнения прикладного анализа данных при решении научных и общественных задачах</p> <p>5. Анализировать новые направления развития и обосновывать перспективы применения технологий машинного обучения к решению конкретных общественных задач, формировать программы и непосредственно принимает участие в реализации мероприятий в рамках решения общественных задач по средствам технологий машинного обучения</p>

8. ОСНОВНАЯ И ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА

8.1. Основная литература

1. Ганегедара, Т. Обработка естественного языка с TensorFlow : монография / Т. Ганегедара ; пер. с англ. В. С. Яценкова. - Москва : ДМК Пресс, 2020. - 382 с. - ISBN 978-5-97060-756-5. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1094940> .
2. Тапскотт, Д. Викиномика: как массовое сотрудничество изменяет все / Д. Тапскотт, Э. Д. Уильямс. - Москва : Интеллектуальная Литература, 2020. - 456 с. - ISBN 978-5-6042878-7-3. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1220225> . – Режим доступа: по подписке.

8.2. Дополнительная литература

1. Богданов, Е. П. Интеллектуальный анализ данных : практикум для магистрантов направления 09.04.03 «Прикладная информатика» профиль подготовки «Информационные системы и технологии корпоративного управления» / Е. П. Богданов. - Волгоград : ФГБОУ ВО Волгоградский ГАУ, 2019. - 112 с. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1087885> . – Режим доступа: по подписке.

9. ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ, ИСПОЛЬЗУЕМЫЕ ПРИ ОСУЩЕСТВЛЕНИИ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА

9.1 Программное обеспечение

При осуществлении образовательного процесса магистрантами и профессорско-преподавательским составом используется следующее лицензионное программное обеспечение:

1. OS Microsoft Windows (OVS OS Platform)
2. MS Office (OVS Office Platform)
3. Adobe Acrobat Professional 11.0 MLP AOO License RU
4. Adobe CS5.5 Design Standart Win IE EDU CLP
5. ABBYY FineReader 11 Corporate Edition
6. ABBYY Lingvo x5
7. Adobe Photoshop Extended CS6 13.0 MLP AOO License RU
8. Adobe Acrobat Reader DC /Pro – бесплатно
9. Google Chrome – бесплатно
10. Opera – бесплатно
11. Mozilla – бесплатно
12. VLC – бесплатно

9.2 Перечень информационно-справочных систем и профессиональных баз данных информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины:

Информационно-справочные системы

1. Гарант.Ру. Информационно-правовой портал: <http://www.garant.ru>
2. Информационная система «Единое окно доступа к образовательным ресурсам»: <http://window.edu.ru/>
3. Открытое образование. Ассоциация «Национальная платформа открытого образования»: <http://npoed.ru>
4. Официальная Россия. Сервер органов государственной власти Российской Федерации: <http://www.gov.ru>
5. Официальный интернет-портал правовой информации. Государственная система правовой информации: <http://pravo.gov.ru>
6. Правовой сайт КонсультантПлюс: <http://www.consultant.ru/sys>
7. Российское образование. Федеральный портал: <http://www.edu.ru>

Профессиональные базы данных информационно-телекоммуникационной сети «Интернет»:

1. Google. Книги: <https://books.google.com>
2. Internet Archive: <https://archive.org>
3. Koob.ru. Электронная библиотека «Куб»: <http://www.koob.ru/philosophy/>
4. SOC.LIB.RU. Социология, психология, управление: <http://soc.lib.ru/>
5. Socioline.ru. Учебники, монографии по социологии: <http://socioline.ru>
6. Библиотека Гумер – гуманитарные науки: <http://www.gumer.info>

7. ЕНИП — Электронная библиотека «Научное наследие России»: <http://e-heritage.ru/>
8. Интелрос. Интеллектуальная Россия: <http://www.intelros.ru/>
9. Национальная электронная библиотека НЭБ: <http://www.rusneb.ru>
10. Неприкосновенный запас: <http://magazines.russ.ru/nz/>
11. Президентская библиотека: <http://www.prlib.ru>
12. Российская государственная библиотека: <http://www.rsl.ru/>
13. Российская национальная библиотека: <http://www.nlr.ru/poisk/>

9.3 Лицензионные электронные ресурсы библиотеки Университета

Профессиональные базы данных:

Полный перечень доступных обучающимся профессиональных баз данных представлен на официальном сайте Университета <https://eusp.org/library/electronic-resources>, включая следующие базы данных:

1. **East View** – 100 ведущих российских журналов по гуманитарным наукам (архив и текущая подписка): [https://dlib.eastview.com/browse](https://dlib.eastview.com/browse;);
2. **eLIBRARY.RU** — Российский информационно-аналитический портал в области науки, технологии, медицины и образования, содержащий рефераты и полные тексты научных статей и публикаций, наукометрическая база данных: [http://elibrary.ru](http://elibrary.ru;);
3. **Университетская информационная система РОССИЯ** — база электронных ресурсов для учебных программ и исследовательских проектов в области социально-гуманитарных наук: <http://www.uisrussia.msu.ru/>;
4. Электронные журналы по подписке (текущие номера научных зарубежных журналов).

Электронные библиотечные системы:

1. **Znanium.com** – Электронная библиотечная система (ЭБС) – <http://znanium.com/>;
2. Университетская библиотека онлайн – Электронная библиотечная система (ЭБС) – <http://biblioclub.ru/>

9.4 Электронная информационно-образовательная среда Университета

Образовательный процесс по дисциплине поддерживается средствами электронной информационно-образовательной среды Университета, которая включает в себя электронный учебно-методический ресурс АНООВО «ЕУСПб» — образовательный портал LMS Sakai — Sakai@EU, лицензионные электронные ресурсы библиотеки Университета, официальный сайт Университета (Европейский университет в Санкт-Петербурге [<https://eu.spb.ru/>]), локальную сеть и корпоративную электронную почту Университета, и обеспечивает:

- доступ к учебным планам, рабочим программам дисциплин (модулей), практик и к изданиям электронных библиотечных систем и электронным образовательным ресурсам, указанным в рабочих программах;
- фиксацию хода образовательного процесса, результатов промежуточной аттестации и результатов освоения основной образовательной программы;
- формирование электронного портфолио обучающегося, в том числе сохранение работ обучающегося, рецензий и оценок за эти работы со стороны любых участников образовательного процесса;
- взаимодействие между участниками образовательного процесса, в том числе синхронное и (или) асинхронное взаимодействие посредством сети «Интернет» (электронной почты и т.д.).

Каждый обучающийся в течение всего периода обучения обеспечен индивидуальным неограниченным доступом к электронным ресурсам библиотеки

Университета, содержащей издания учебной, учебно-методической и иной литературы по изучаемой дисциплине.

10. МАТЕРИАЛЬНО-ТЕХНИЧЕСКАЯ БАЗА, НЕОБХОДИМАЯ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА

В ходе реализации образовательного процесса используются специализированные многофункциональные аудитории для проведения занятий лекционного типа, занятий семинарского типа (практических занятий, лабораторных работ), групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, укомплектованные специализированной мебелью и техническими средствами обучения, служащими для представления учебной информации большой аудитории.

Проведение занятий лекционного типа обеспечивается демонстрационным оборудованием.

Помещения для самостоятельной работы оснащены компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду организации.

Для лиц с ограниченными возможностями здоровья и инвалидов предоставляется возможность присутствия в аудитории вместе с ними ассистента (помощника). Для слабовидящих предоставляется возможность увеличения текста на экране ПК. Для самостоятельной работы лиц с ограниченными возможностями здоровья в помещении для самостоятельной работы организовано одно место (ПК) с возможностями бесконтактного ввода информации и управления компьютером (специализированное лицензионное программное обеспечение – Camera Mouse, веб камера). Библиотека университета предоставляет удаленный доступ к электронным ресурсам библиотеки Университета с возможностями для слабовидящих увеличения текста на экране ПК. Лица с ограниченными возможностями здоровья могут при необходимости воспользоваться имеющимся в университете креслом-коляской. В учебном корпусе имеется адаптированный лифт. На первом этаже оборудован специализированный туалет. У входа в здание университета для инвалидов оборудована специальная кнопка, входная среда обеспечена информационной доской о режиме работы университета, выполненной рельефно-точечным тактильным шрифтом (азбука Брайля).