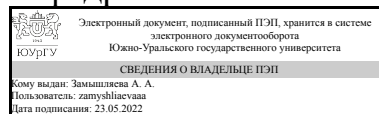


УТВЕРЖДАЮ:  
Заведующий выпускающей  
кафедрой



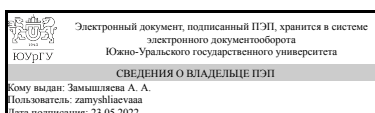
А. А. Замышляева

## РАБОЧАЯ ПРОГРАММА

**дисциплины** 1.Ф.М1.09 Анализ естественного языка при разработке интеллектуального цифрового ассистента: проектное обучение  
**для направления** 09.04.04 Программная инженерия  
**уровень** Магистратура  
**магистерская программа** Искусственный интеллект и машинное обучение в финтех индустрии  
**форма обучения** очная  
**кафедра-разработчик** Прикладная математика и программирование

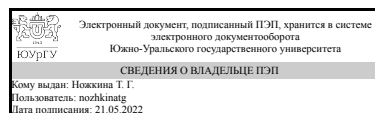
Рабочая программа составлена в соответствии с ФГОС ВО по направлению подготовки 09.04.04 Программная инженерия, утверждённым приказом Минобрнауки от 19.09.2017 № 932

Зав.кафедрой разработчика,  
д.физ.-мат.н., проф.



А. А. Замышляева

Разработчик программы,  
старший преподаватель



Т. Г. Ножкина

## 1. Цели и задачи дисциплины

Целью дисциплины является формирование базовых представлений, знаний и умений в анализе естественного языка. Основные задачи дисциплины: ознакомить студента с основными понятиями анализа и обработки текстов на естественном языке, дать понимание базовых подходов и методов при решении задач анализа естественного языка, получить практический опыт работы с различными алгоритмами машинного обучения и архитектурами искусственных нейронных сетей в рамках задач обработки естественного языка.

## Краткое содержание дисциплины

Изложены наиболее важные понятия, определения и методы машинного обучения и искусственных нейронных сетей в задачах анализа естественного языка. В курс входят следующие разделы: введение в анализ естественного языка, машинное обучение и глубокие нейронные сети для решения задач анализа и обработки естественного языка, построение диалоговых систем. На практике студенты применяют навыки построения моделей машинного обучения и искусственных нейронных сетей на языке Python для решения задач морфологического анализа, классификации и кластеризации текстовых документов, анализа тональности, определения семантической близости слов, машинного перевода, построения вопросно-ответных систем, автоматического реферирования текста, построения диалоговых систем.

## 2. Компетенции обучающегося, формируемые в результате освоения дисциплины

Планируемые результаты освоения ОП ВО (компетенции)	Планируемые результаты обучения по дисциплине
ПК-4 Способен разрабатывать и модернизировать программное обеспечение по анализу естественного языка методами искусственного интеллекта и машинного обучения для применения в финансовой индустрии	Знает: классические подходы к анализу текстовой информации на естественном языке Умеет: проводить анализ исходных данных конкретной задачи для выбора подходящего набора методов анализа и обработки текстов на естественном языке Имеет практический опыт: применения различных методов обработки и анализа текста и их применимость в реальных практических задачах

## 3. Место дисциплины в структуре ОП ВО

Перечень предшествующих дисциплин, видов работ учебного плана	Перечень последующих дисциплин, видов работ
Нет	Языковые модели для интеллектуального цифрового ассистента: проектное обучение, Производственная практика, преддипломная практика: проектное обучение (4 семестр), Производственная практика, научно-исследовательская работа: проектное обучение (3 семестр)

Требования к «входным» знаниям, умениям, навыкам студента, необходимым при освоении данной дисциплины и приобретенным в результате освоения предшествующих дисциплин:

Нет

#### 4. Объём и виды учебной работы

Общая трудоемкость дисциплины составляет 4 з.е., 144 ч., 32,25 ч. контактной работы

Вид учебной работы	Всего часов	Распределение по семестрам в часах	
		Номер семестра	
		2	
Общая трудоёмкость дисциплины	144	144	
<i>Аудиторные занятия:</i>	16	16	
Лекции (Л)	0	0	
Практические занятия, семинары и (или) другие виды аудиторных занятий (ПЗ)	16	16	
Лабораторные работы (ЛР)	0	0	
<i>Самостоятельная работа (СРС)</i>	111,75	111,75	
с применением дистанционных образовательных технологий	0		
Выполнение домашних заданий	32	32	
Подготовка к дифференцированному зачёту	19,75	19,75	
Изучение основной и дополнительной литературы по анализу и обработке естественного языка	60	60	
Консультации и промежуточная аттестация	16,25	16,25	
Вид контроля (зачет, диф.зачет, экзамен)	-	диф.зачет	

#### 5. Содержание дисциплины

№ раздела	Наименование разделов дисциплины	Объем аудиторных занятий по видам в часах			
		Всего	Л	ПЗ	ЛР
1	Введение в обработку естественного языка	2	0	2	0
2	Машинное обучение и глубокие нейронные сети для решения задач анализа и обработки естественного языка	10	0	10	0
3	Построение диалоговых систем	4	0	4	0

##### 5.1. Лекции

Не предусмотрены

##### 5.2. Практические занятия, семинары

№ занятия	№ раздела	Наименование или краткое содержание практического занятия, семинара	Кол-во



1	2	Текущий контроль	ДЗ-1. Реализация собственного POS- тэггера.	1	3	3 балла: задание выполнено полностью, 2 балла: задание выполнено полностью, но допущены незначительные ошибки, 1 балла: задание выполнено полностью, но допущены серьезные ошибки, 0 баллов: задание не выполнено	дифференцированный зачет
2	2	Текущий контроль	ДЗ-2. Классификация текстов методами машинного обучения	1	3	3 балла: задание выполнено полностью, 2 балла: задание выполнено полностью, но допущены незначительные ошибки, 1 балла: задание выполнено полностью, но допущены серьезные ошибки, 0 баллов: задание не выполнено	дифференцированный зачет
3	2	Текущий контроль	ДЗ-3. Классификация текстов на основе нейросетевых моделей языка	1	3	3 балла: задание выполнено полностью, 2 балла: задание выполнено полностью, но допущены незначительные ошибки, 1 балла: задание выполнено полностью, но допущены серьезные ошибки, 0 баллов: задание не выполнено	дифференцированный зачет
4	2	Текущий контроль	ДЗ-4. Кластеризация текстовой коллекции методами тематического моделирования	1	3	3 балла: задание выполнено полностью, 2 балла: задание выполнено полностью, но допущены незначительные ошибки, 1 балла: задание выполнено полностью, но допущены серьезные ошибки, 0 баллов: задание не выполнено	дифференцированный зачет
5	2	Текущий контроль	ДЗ-5. Классификация текстов с помощью различных архитектур глубоких нейронных сетей	1	3	3 балла: задание выполнено полностью, 2 балла: задание выполнено полностью, но допущены незначительные ошибки, 1 балла: задание выполнено полностью, но	дифференцированный зачет

						допущены серьезные ошибки, 0 баллов: задание не выполнено	
6	2	Текущий контроль	ДЗ-6. Реализация вопросно-ответной системы. Решение задачи автоматического реферирования	1	3	3 балла: задание выполнено полностью, 2 балла: задание выполнено полностью, но допущены незначительные ошибки, 1 балла: задание выполнено полностью, но допущены серьезные ошибки, 0 баллов: задание не выполнено	дифференцированный зачет
7	2	Текущий контроль	ДЗ-7. Разработка диалогового агента для чат-бота	2	3	3 балла: задание выполнено полностью, 2 балла: задание выполнено полностью, но допущены незначительные ошибки, 1 балла: задание выполнено полностью, но допущены серьезные ошибки, 0 баллов: задание не выполнено	дифференцированный зачет
8	2	Промежуточная аттестация	Итоговый тест	-	20	Компьютерный тест состоит из 20 вопросов. Каждый правильный ответ - 1 балл.	дифференцированный зачет

## 6.2. Процедура проведения, критерии оценивания

Вид промежуточной аттестации	Процедура проведения	Критерии оценивания
дифференцированный зачет	На зачете происходит оценивание учебной деятельности обучающихся по дисциплине на основе полученных оценок за контрольно-рейтинговые мероприятия текущего контроля. Студент может улучшить свой рейтинг, пройдя контрольное мероприятие промежуточной аттестации, которое не является обязательным. Контрольное мероприятие промежуточной аттестации проводится во время зачета в виде компьютерного тестирования. Тест состоит из 20 вопросов, время тестирования 60 минут.	В соответствии с пп. 2.5, 2.6 Положения

## 6.3. Паспорт фонда оценочных средств

Компетенции	Результаты обучения	№ КМ							
		1	2	3	4	5	6	7	8
ПК-4	Знает: классические подходы к анализу текстовой информации на естественном языке	+			+	+	+	+	+



		издательства Лань	с. <a href="https://e.lanbook.com/book/107901">https://e.lanbook.com/book/107901</a>
6	Дополнительная литература	Электронно-библиотечная система издательства Лань	Коэльо, Л. П. Построение систем машинного обучения на языке Python / Л. П. Коэльо, В. Ричарт ; перевод с английского А. А. Слинкин. — 2-е изд. — Москва : ДМК Пресс, 2016. — 302 с. <a href="https://e.lanbook.com/book/82818">https://e.lanbook.com/book/82818</a>

Перечень используемого программного обеспечения:

1. -Python(бессрочно)

Перечень используемых профессиональных баз данных и информационных справочных систем:

Нет

## 8. Материально-техническое обеспечение дисциплины

Вид занятий	№ ауд.	Основное оборудование, стенды, макеты, компьютерная техника, предустановленное программное обеспечение, используемое для различных видов занятий
Практические занятия и семинары	332 (36)	Проектор, экран, персональные компьютеры, программное обеспечение.
Зачет, диф.зачет	332 (36)	Персональные компьютеры.