

**Федеральное государственное автономное образовательное учреждение
высшего образования «Российский университет дружбы народов»**

Институт русского языка

(наименование основного учебного подразделения (ОУП)-разработчика ОП ВО)

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Машинное обучение. Инструменты распознавания речи

(наименование дисциплины/модуля)

Рекомендована МСЧН для направления подготовки/специальности:

45.04.01. Филология

(код и наименование направления подготовки/специальности)

Освоение дисциплины ведется в рамках реализации основной профессиональной образовательной программы высшего образования (ОП ВО):

Цифровые инновации в филологии (магистратура)

(наименование (профиль/специализация) ОП ВО)

2022 г.

1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Целью освоения дисциплины «Машинное обучение. Инструменты распознавания речи» является знакомство с базовыми понятиями машинного обучения, с основными алгоритмами машинного обучения, особенностями их применения в распознавании речи.

2. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Освоение дисциплины «Машинное обучение. Инструменты распознавания речи» направлено на формирование у обучающихся следующих компетенций (части компетенций):

Таблица 2.1. Перечень компетенций, формируемых у обучающихся при освоении дисциплины (результаты освоения дисциплины)

Шифр	Компетенция	Индикаторы достижения компетенции (в рамках данной дисциплины)
ОПК-4	Способен использовать цифровые технологии и методы в профессиональной деятельности в области филологии для изучения и моделирования объектов профессиональной деятельности, анализа данных, представления информации.	ОПК-4.1. Использует цифровые технологии и методы в профессиональной деятельности в области филологии для изучения и моделирования объектов профессиональной деятельности.
ПК-1	Способен применять информационные технологии в филологических областях знаний с использованием средств интеллектуального анализа данных и машинного обучения, компьютерной лингвистики и представления данных.	ПК-1.1. Знает области применения новых информационных технологий, включая использование интеллектуального анализа данных и машинного обучения компьютерной лингвистики и представления данных.
		ПК-1.2. Использует различные инструментальные средства, платформы для разработки приложений и прикладные программы в гуманитарных и филологических областях знаний.

В результате освоения дисциплины обучающийся должен:

Знать:

- основные принципы машинного обучения;
- возможности алгоритмов машинного обучения;
- классы задач, решаемых с помощью алгоритмов машинного обучения;
- принципы распознавания речи.

Уметь:

- применять на практике алгоритмы машинного обучения;

- обосновать применение того или иного алгоритма машинного обучения для решения конкретной задачи распознавания речи.

Владеть:

- базовым инструментарием машинного обучения;
- технологиями программной реализации алгоритмов машинного обучения;
- технологиями и приемами использования инструментов и алгоритмов машинного обучения для решения задач распознавания речи.

3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОП ВО

Дисциплина «Машинное обучение. Инструменты распознавания речи» относится к вариативной компоненте обязательной части Б1.О.02.13.

В рамках ОП ВО обучающиеся также осваивают другие дисциплины и/или практики, способствующие достижению запланированных результатов освоения дисциплины «Машинное обучение. Инструменты распознавания речи».

Таблица 3.1. Перечень компонентов ОП ВО, способствующих достижению запланированных результатов освоения дисциплины

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики*	Последующие дисциплины/модули, практики*
ОПК-4	Способен использовать цифровые технологии и методы профессиональной деятельности в области филологии для изучения и моделирования объектов профессиональной деятельности, анализа данных, представления информации.	Программирование на языке Python. Табличный редактор: продвинутый Excel.	Цифровые технологии компаративного исследования текста. Научно-исследовательская работа.
ПК-1	Способен применять информационные технологии в филологических областях знаний с использованием средств интеллектуального анализа данных и машинного обучения, компьютерной лингвистики и представления данных.	Программирование на языке Python.	Веб-семантика. Веб-разработка. Научно-исследовательская работа

* - заполняется в соответствии с матрицей компетенций и СУП ОП ВО

4. ОБЪЕМ ДИСЦИПЛИНЫ И ВИДЫ УЧЕБНОЙ РАБОТЫ

Общая трудоемкость дисциплины «Машинное обучение. Инструменты распознавания речи» составляет 4 зачетные единицы (144 ак.ч.а).

Таблица 4.1. Виды учебной работы по периодам освоения ОП ВО для ОЧНОЙ формы обучения

Вид учебной работы	ВСЕГО, ак.ч.	Семестр(-ы)			
		1	2	3	4
Контактная работа, ак.ч.	26			26	
в том числе:					
Лекции (ЛК)	13			13	
Лабораторные работы (ЛР)					
Практические/семинарские занятия (СЗ)	13			13	
Самостоятельная работа обучающихся, ак.ч.	92			92	
Контроль (экзамен), ак.ч.	26			26	
Общая трудоемкость дисциплины	ак.ч.	144		144	
	зач.ед.	4		4	

Таблица 4.2. Виды учебной работы по периодам освоения ОП ВО для ОЧНО-ЗАОЧНОЙ формы обучения*

Не предусмотрено ОП ВО.

Вид учебной работы	ВСЕГО, ак.ч.	Семестр(-ы)			
		1	2	3	4
Контактная работа, ак.ч.					
в том числе:					
Лекции (ЛК)					
Лабораторные работы (ЛР)					
Практические/семинарские занятия (СЗ)					
Самостоятельная работа обучающихся, ак.ч.					
Контроль (экзамен), ак.ч.					
Общая трудоемкость дисциплины	ак.ч.				
	зач.ед.				

* - заполняется в случае реализации программы в заочной форме

Таблица 4.3. Виды учебной работы по периодам освоения ОП ВО для ЗАОЧНОЙ формы обучения*

Вид учебной работы	ВСЕГО, ак.ч.	Семестр(-ы)			
		1	2	3	4
Контактная работа, ак.ч.	20				20
в том числе:					
Лекции (ЛК)	9				9
Лабораторные работы (ЛР)					
Практические/семинарские занятия (СЗ)	12				12
Самостоятельная работа обучающихся, ак.ч.	106				106
Контроль (экзамен), ак.ч.	18				18
Общая трудоемкость дисциплины	ак.ч.	144		144	
	зач.ед.	4		4	

* - заполняется в случае реализации программы в заочной форме

5. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

Таблица 5.1. Содержание дисциплины (модуля) по видам учебной работы

Наименование раздела дисциплины	Содержание раздела (темы)	Вид учебной работы*
Раздел 1. Введение в машинное обучение	Тема 1.1. Типы машинного обучения. Схема построения систем машинного обучения.	ЛК
	Тема 1.2. Методы качества модели, переобучение и недообучение, необходимый инструментарий.	ЛК
	Тема 1.3. Задачи по тестовому набору данных, валидационная выборка, метрики качества, библиотека scikit-learn	СЗ
Раздел 2. Обучение с учителем	Тема 2.1. Принципы обучения с учителем, подготовка признаков, задача классификации и метрика, задача регрессии и метрика	ЛК
	Тема 2.2. Задачи по обучающему набору данных, модели и алгоритмы, задачи классификации	СЗ
	Тема 2.3. Задачи по регрессии, дерево решений в регрессии, случайный лес в регрессии, линейная регрессия, выбор лучшей модели, отображение моделей	СЗ
Раздел 3. Обучение без учителя	Тема 3.1. Принципы обучения без учителя, кластеризация и метрика, поиск аномалий	ЛК
	Тема 3.2. Задачи по кластеризации, алгоритм k-средних, целевая функция, визуализация, интерпретация результата	
	Тема 3.3. Задачи по аномалии, диаграмма размаха, KNN для поиска аномалий	СЗ
Раздел 4. Основы глубокого обучения	Тема 4.1. Понятие нейронной сети, многослойная нейронная сеть. Применение библиотеки TensorFlow для обучения нейронной сети.	ЛК
	Тема 4.2. Задача генераторы текста, задача классификации документов	СЗ
Раздел 5. Распознавание речи	Тема 5.1. Классификация систем распознавания речи, введение в акустические и языковые модели, преобразование звука в вектор, извлечение признаков, кодирование и декодирование	ЛК
	Тема 5.2. Задача синтеза речи, задача распознавания речи Пакет SpeechRecognition Пакет pytsx3	СЗ

	Пакет vosk	
Раздел 6. Практическая часть	Тема 6.1. Реализация модели голосового управления	СЗ

* - заполняется только по **ОЧНОЙ** форме обучения: ЛК – лекции; ЛР – лабораторные работы; СЗ – семинарские занятия.

6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Таблица 6.1. Материально-техническое обеспечение дисциплины

Тип аудитории	Оснащение аудитории	Специализированное учебное/лабораторное оборудование, ПО и материалы для освоения дисциплины (при необходимости)
Лекционная	Аудитория для проведения занятий лекционного типа, оснащенная комплектом специализированной мебели; доской (экраном) и техническими средствами мультимедиа презентаций.	Проектор и/или большой экран/монитор, компьютер, ОС MS Windows 10, MS Office 2016 или MS Office 2019 или MS Office 365, портал Microsoft Azure, доступ к интернету
Лаборатория	Аудитория для проведения лабораторных работ, индивидуальных консультаций, текущего контроля и промежуточной аттестации, оснащенная комплектом специализированной мебели и оборудованием.	Проектор и/или большой экран/монитор, компьютер, ОС MS Windows 10, MS Office 2016 или MS Office 2019 или MS Office 365, портал Microsoft Azure, доступ к интернету
Семинарская	Аудитория для проведения занятий семинарского типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, оснащенная комплектом специализированной мебели и техническими средствами мультимедиа презентаций.	Проектор и/или большой экран/монитор, компьютер, ОС MS Windows 10, MS Office 2016 или MS Office 2019 или MS Office 365, портал Microsoft Azure, доступ к интернету
Компьютерный класс	Компьютерный класс для проведения занятий, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, оснащенная персональными компьютерами (в количестве 12 шт.), доской (экраном) и техническими средствами мультимедиа презентаций.	Язык программирования Python и программа Anaconda Доступ к сети Интернет.

Тип аудитории	Оснащение аудитории	Специализированное учебное/лабораторное оборудование, ПО и материалы для освоения дисциплины (при необходимости)
Для самостоятельной работы обучающихся	Аудитория для самостоятельной работы обучающихся (может использоваться для проведения семинарских занятий и консультаций), оснащенная комплектом специализированной мебели и компьютерами с доступом в ЭИОС.	Проектор и/или большой экран/монитор, компьютер, ОС MS Windows 10, MS Office 2016 или MS Office 2019 или MS Office 365, портал Microsoft Azure, доступ к интернету

* - аудитория для самостоятельной работы обучающихся указывается **ОБЯЗАТЕЛЬНО!**

7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Основная литература:

Таeho Jo. Machine Learning Foundations: Supervised, Unsupervised, and Advanced Learning. Springer, 2020.

Бринк Х., Ричардс Дж., Феверолф М. Машинное обучение.

Дополнительная литература:

Mohri M., Rostamizadeh A., Talwalkar A. Foundations of Machine Learning. MIT Press, 2012.

Рашка С. Python и машинное обучение. СПб.: Диалектика, 2020.

Саттон Р.С., Барто Э.Г. Обучение с подкреплением: Введение. 2-е изд. / пер. с англ. А. А. Слинкина. – М.: ДМК Пресс, 2020.

МООК «Машинное обучение». Режим доступа: <https://www.coursera.org/specializations/machine-learning-data-analysis>

Ресурсы информационно-телекоммуникационной сети «Интернет»:

1. ЭБС РУДН и сторонние ЭБС, к которым студенты университета имеют доступ на основании заключенных договоров:

- Электронно-библиотечная система РУДН – ЭБС РУДН <http://lib.rudn.ru/MegaPro/Web>

- ЭБС «Университетская библиотека онлайн» <http://www.biblioclub.ru>

- ЭБС Юрайт <http://www.biblio-online.ru>

- ЭБС «Консультант студента» www.studentlibrary.ru

- ЭБС «Лань» <http://e.lanbook.com/>

- ЭБС «Троицкий мост»

2. Базы данных и поисковые системы:

- электронный фонд правовой и нормативно-технической документации <http://docs.cntd.ru/>

- поисковая система Яндекс <https://www.yandex.ru/>

- поисковая система Google <https://www.google.ru/>
- реферативная база данных SCOPUS
<http://www.elsevierscience.ru/products/scopus/>
- <https://habr.com/>
- <http://www.opennet.org> ,
- <http://www.javascript.ru>
- www.phpclub.ru,
- www.w3c.org,
- <http://apache.org>
- Российская ассоциация цифровых гуманитарных наук (dhrussia.ru)

Учебно-методические материалы для самостоятельной работы обучающихся при освоении дисциплины/модуля:*

1. Курс лекций по дисциплине «Машинное обучение. Инструменты распознавания речи».

2. Презентационные материалы по дисциплине «Машинное обучение. Инструменты распознавания речи».

* - все учебно-методические материалы для самостоятельной работы обучающихся размещаются в соответствии с действующим порядком на странице дисциплины **в ТУИС!**

8. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ И БАЛЛЬНО-РЕЙТИНГОВАЯ СИСТЕМА ОЦЕНИВАНИЯ УРОВНЯ СФОРМИРОВАННОСТИ КОМПЕТЕНЦИЙ ПО ДИСЦИПЛИНЕ

Оценочные материалы и балльно-рейтинговая система* оценивания уровня сформированности компетенций (части компетенций) по итогам освоения дисциплины «Машинное обучение. Инструменты распознавания речи» представлены в Приложении к настоящей Рабочей программе дисциплины.

* - ОМ и БРС формируются на основании требований соответствующего локального нормативного акта РУДН.

РАЗРАБОТЧИКИ:

**Руководитель направления
развития Центра развития
цифровых технологий для
образовательных процессов
РУДН**

Должность, БУП



Подпись

Резаиан Н.

Фамилия И.О.

РУКОВОДИТЕЛЬ БУП:

**Директор
Института русского языка**

Наименование БУП



Подпись

Должикова А.В.

Фамилия И.О.

РУКОВОДИТЕЛЬ ОП ВО:

**Заведующий кафедрой
русского языка № 5
Института русского языка**

должность, БУП



Подпись

Куриленко В.Б.

Фамилия И.О.