

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Ястребов Олег Александрович

Должность: Ректор

Дата подписания: 29.05.2024 11:28:12

Уникальный программный ключ:

ca953a01204891083f939673078ef1a989dae18a

Федеральное государственное автономное образовательное учреждение высшего образования

«Российский университет дружбы народов имени Патриса Лумумбы»

Факультет искусственного интеллекта

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

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

ГЕНЕРАТИВНЫЕ МЕТОДЫ В NLP

(наименование дисциплины/модуля)

Рекомендована МССН для направления подготовки/специальности:

02.04.02 ФУНДАМЕНТАЛЬНАЯ ИНФОРМАТИКА И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

(код и наименование направления подготовки/специальности)

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

УПРАВЛЕНИЕ ДАННЫМИ И ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

(наименование (профиль/специализация) ОП ВО)

2024 г.

1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Дисциплина «Генеративные методы в NLP» входит в программу магистратуры «Управление данными и искусственный интеллект» по направлению 02.04.02 «Фундаментальная информатика и информационные технологии» и изучается в 3 семестре 2 курса. Дисциплину реализует NOT DEFINED. Дисциплина состоит из 4 разделов и 12 тем и направлена на изучение и приобретение знаний о принципах и методах использования генеративных моделей в обработке естественного языка.

Целью освоения дисциплины является овладение навыками разработки, применения и оценки генеративных моделей для решения различных задач в области обработки естественного языка.

2. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Освоение дисциплины «Генеративные методы в NLP» 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)
УК-7	Способен: искать нужные источники информации и данные, воспринимать, анализировать, запоминать и передавать информацию с использованием цифровых средств, а также с помощью алгоритмов при работе с полученными из различных источников данными с целью эффективного использования полученной информации для решения задач; проводить оценку информации, ее достоверность, строить логические умозаключения на основании поступающих информации и данных	УК-7.2 Умеет применять цифровые технологии для поиска, обработки, анализа, хранения и представления информации в области профессиональной деятельности; УК-7.3 Владеет навыками применения цифровых технологий и методов поиска, обработки, анализа, хранения и представления информации в области профессиональной деятельности;
ОПК-2	Способен применять компьютерные / суперкомпьютерные методы, современное программное обеспечение (в том числе отечественного производства) для решения задач профессиональной деятельности	ОПК-2.1 Знает основные положения и концепции в области программирования, архитектуру языков программирования, теории коммуникации, знает основную терминологию, знаком с перечнем ПО, включенного в Единый Реестр Российских программ; ОПК-2.2 Умеет анализировать типовые языки программирования, составлять программы;
ОПК-4	Способен оптимальным образом комбинировать существующие информационно-коммуникационные технологии для решения задач в области профессиональной деятельности с учетом требований информационной безопасности	ОПК-4.1 Знает принципы сбора и анализа информации, создания информационных систем на стадиях жизненного цикла; ОПК-4.2 Умеет осуществлять управление проектами информационных систем;

Шифр	Компетенция	Индикаторы достижения компетенции (в рамках данной дисциплины)
ОПК-5	Способен устанавливать и сопровождать программное обеспечение информационных систем, осуществлять эффективное управление разработкой программных средств и проектов	ОПК-5.2 Умеет реализовывать техническое сопровождение информационных систем и баз данных;
ПК-2	Способен к овладению методами математического моделирования объектов и процессов при разработке требований и проектировании программного обеспечения систем искусственного интеллекта	ПК-2.3 Имеет навыки применения методов математического моделирования объектов и процессов при разработке требований и проектировании программного обеспечения систем искусственного интеллекта;

3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОП ВО

Дисциплина «Генеративные методы в NLP» относится к обязательной части блока 1 «Дисциплины (модули)» образовательной программы высшего образования.

В рамках образовательной программы высшего образования обучающиеся также осваивают другие дисциплины и/или практики, способствующие достижению запланированных результатов освоения дисциплины «Генеративные методы в NLP».

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

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики*	Последующие дисциплины/модули, практики*
УК-7	Способен: искать нужные источники информации и данные, воспринимать, анализировать, запоминать и передавать информацию с использованием цифровых средств, а также с помощью алгоритмов при работе с полученными из различных источников данными с целью эффективного использования полученной информации для решения задач; проводить оценку информации, ее достоверность, строить логические умозаключения на основании поступающих информации и данных	Правовые и этические основы технологий искусственного интеллекта; <i>Современные устройства центров обработки больших данных**;</i> <i>Современные инструментальные средства разработки ПО для искусственного интеллекта**;</i> Алгоритмы и структуры данных; Морфологический и семантический анализ; Практикум по программированию;	Преддипломная практика;
ОПК-2	Способен применять компьютерные / суперкомпьютерные методы, современное программное обеспечение (в том числе	Методы глубокого обучения для обработки естественного языка; Методы машинного обучения; Практикум по программированию; Морфологический и	Технологическая (проектно-технологическая) практика (производственная);

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики*	Последующие дисциплины/модули, практики*
	отечественного производства) для решения задач профессиональной деятельности	семантический анализ; Введение в обработку естественного языка;	
ОПК-4	Способен оптимальным образом комбинировать существующие информационно-коммуникационные технологии для решения задач в области профессиональной деятельности с учетом требований информационной безопасности		
ОПК-5	Способен устанавливать и сопровождать программное обеспечение информационных систем, осуществлять эффективное управление разработкой программных средств и проектов	Методы машинного обучения; Глубинное обучение; Машинный перевод; Практикум по программированию; Обучение с подкреплением;	
ПК-2	Способен к овладению методами математического моделирования объектов и процессов при разработке требований и проектировании программного обеспечения систем искусственного интеллекта	Методы машинного обучения; Введение в обработку естественного языка; Глубинное обучение; Практикум по программированию; Обучение с подкреплением;	Технологическая (проектно-технологическая) практика (производственная); Преддипломная практика;

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

** - элективные дисциплины /практики

4. ОБЪЕМ ДИСЦИПЛИНЫ И ВИДЫ УЧЕБНОЙ РАБОТЫ

Общая трудоемкость дисциплины «Генеративные методы в NLP» составляет «4» зачетные единицы.

Таблица 4.1. Виды учебной работы по периодам освоения образовательной программы высшего образования для очной формы обучения.

Вид учебной работы	ВСЕГО, ак.ч.		Семестр(-ы)
			3
<i>Контактная работа, ак.ч.</i>	36		36
Лекции (ЛК)	18		18
Лабораторные работы (ЛР)	0		0
Практические/семинарские занятия (СЗ)	18		18
<i>Самостоятельная работа обучающихся, ак.ч.</i>	90		90
<i>Контроль (экзамен/зачет с оценкой), ак.ч.</i>	18		18
Общая трудоемкость дисциплины	ак.ч.	144	144
	зач.ед.	4	4

5. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

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

Номер раздела	Наименование раздела дисциплины	Содержание раздела (темы)		Вид учебной работы*
Раздел 1	Генеративные методы в NLP	1.1	Определение генеративных методов в обработке естественного языка	ЛК, СЗ
		1.2	Вероятностные модели в генеративных методах	ЛК, СЗ
		1.3	Применение генеративных методов в различных задачах NLP	ЛК, СЗ
Раздел 2	Генеративные модели для генерации текста	2.1	Модели языка в генеративных методах	ЛК, СЗ
		2.2	Методы генерации текста на основе генеративных моделей	ЛК, СЗ
		2.3	Оценка качества и сравнение различных генеративных моделей	ЛК, СЗ
Раздел 3	Генеративные методы для обработки речи	3.1	Применение генеративных моделей в обработке речи	ЛК, СЗ
		3.2	Генерация речи на основе генеративных моделей	ЛК, СЗ
		3.3	Оценка качества и сравнение различных генеративных моделей в обработке речи	ЛК, СЗ
Раздел 4	Применение генеративных методов в NLP	4.1	Генеративные методы для перевода текста	ЛК, СЗ
		4.2	Генеративные методы для аугментации данных	ЛК, СЗ
		4.3	Прочие приложения генеративных методов в NLP	ЛК, СЗ

* - заполняется только по **ОЧНОЙ** форме обучения: ЛК – лекции; ЛР – лабораторные работы; СЗ – практические/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.	
Компьютерный класс	Компьютерный класс для проведения занятий, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, оснащенная персональными компьютерами (в количестве 15 шт.), доской (экраном) и техническими средствами мультимедиа презентаций.	
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.	
Для самостоятельной работы	Аудитория для самостоятельной работы обучающихся (может использоваться для проведения семинарских занятий и консультаций), оснащенная комплектом специализированной мебели и компьютерами с доступом в ЭИОС.	

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

7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Основная литература:

1. Гольдберг Й. Нейросетевые методы в обработке естественного языка. – Litres, 2022.
2. Васильев Ю. Обработка естественного языка. Python и spaCy на практике. – "Издательский дом Питер", 2021.

Дополнительная литература:

1. Лейн Х., Хапке Х., Коул Х. Обработка естественного языка в действии. – "Издательский дом Питер", 2020.

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 «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/>

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

1. Курс лекций по дисциплине «Генеративные методы в NLP».

* - 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

«Генеративные методы в NLP» представлены в Приложении к настоящей Рабочей программе дисциплины.

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

РАЗРАБОТЧИК:

_____	_____	Воробьева Александра
<i>Должность, БУП</i>	<i>Подпись</i>	Андреевна
		<i>Фамилия И.О.</i>

РУКОВОДИТЕЛЬ БУП:

Заведующий кафедрой	_____	_____
<i>Должность БУП</i>	<i>Подпись</i>	<i>Фамилия И.О.</i>

РУКОВОДИТЕЛЬ ОП ВО:

_____	_____	Воробьева Александра
<i>Должность, БУП</i>	<i>Подпись</i>	Андреевна
		<i>Фамилия И.О.</i>