

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Ястребов Олег Александрович
Должность: Ректор
Дата подписания: 27.06.2022 11:51:29
Уникальный программный ключ:
ca953a0120d891083f939673078ef1a989dae18a

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

Факультет физико-математических и естественных наук

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Методы интеллектуального анализа текстов

Рекомендована МСЧН для направления подготовки:

09.04.03 Прикладная информатика

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

«Искусственный интеллект и анализ данных»

2022 г.

1. Цель освоения дисциплины

Целью освоения дисциплины является ознакомление студентов с методами интеллектуального анализа текстов, освоение принципов морфологического, синтаксического и семантического анализа, извлечения информации из текстов, классификации текстов.

Основными задачами освоения дисциплины являются получение навыков работы с библиотеками обработки текстов, применение инструментов обработки текстов к решению практических задач.

2. Требования к результатам освоения дисциплины

Освоение дисциплины «Методы интеллектуального анализа текстов» 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-1	Способность осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий	УК-1.1. Знает принципы сбора, отбора и обобщения информации.
		УК-1.2. Умеет соотносить разнородные явления и систематизировать их в рамках избранных видов профессиональной деятельности.
		УК-1.3. Имеет практический опыт работы с информационными источниками, опыт научного поиска, создания научных текстов.
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-7.1. 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
ОПК-1	Способность	ОПК-1.1. Обладает фундаментальными

Шифр	Компетенция	Индикаторы достижения компетенции (в рамках данной дисциплины)
	самостоятельно приобретать, развивать и применять математические, естественнонаучные, социально-экономические и профессиональные знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте	<p>знаниями в области математических и естественных наук, информатики и теории коммуникаций</p> <p>ОПК-1.2. Умеет осуществлять первичный сбор и анализ материала, интерпретировать различные математические и информационные объекты</p> <p>ОПК-1.3. Имеет практический опыт работы с решением математических и информационных задач и применяет его в профессиональной деятельности</p>
ОПК-2	Способность разрабатывать оригинальные алгоритмы и программные средства, в том числе с использованием современных интеллектуальных технологий, для решения профессиональных задач	<p>ОПК-2.1. Знает основные положения и концепции в области программирования, языков программирования, теории коммуникации, знает основную терминологию, знаком с перечнем ПО, включенного в Единый Реестр Российских программ</p> <p>ОПК-2.2. Умеет анализировать типовые языки программирования, составлять программы</p> <p>ОПК-2.3. Имеет практический опыт решения задач анализа и интеграции различных типов программного обеспечения, анализа типов коммуникации</p>
ОПК-3	Способность самостоятельно приобретать, развивать и применять математические, естественнонаучные, социально-экономические и профессиональные знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте	<p>ОПК-3.1. Знает методы информатики, методы системного и прикладного программирования, основные положения и концепции в области математических и информационных моделей</p> <p>ОПК-3.2. Умеет соотносить знания в области информатики и программирования, интерпретацию прочитанного, определять и создавать информационные ресурсы</p> <p>ОПК-3.3. Имеет практический опыт применения разработки программного обеспечения и тестирования программных продуктов</p>
ОПК-4	Способность применять на практике новые научные принципы и методы исследований	<p>ОПК-4.1. Знает новые научные принципы и методы исследований</p> <p>ОПК-4.2. Умеет применять на практике соотносить знания в области информатики и программирования, интерпретацию прочитанного, определять и создавать информационные ресурсы</p> <p>ОПК-4.3. Имеет практические навыки применения новых научных принципов и</p>

Шифр	Компетенция	Индикаторы достижения компетенции (в рамках данной дисциплины)
		методов исследования для решения профессиональных задач
ПК-1	Проведение работ по обработке и анализу научно-технической информации и результатов исследований	<p>ПК-1.1 Знает основы научно-исследовательской деятельности в области информационных и коммуникационных технологий, принципы построения научной работы, методы сбора и анализа полученного материала; владеет навыками подготовки научных обзоров, публикаций, рефератов и библиографий по тематике проводимых исследований на русском и иностранном языках; способен осуществлять подготовку к публикации материалов в научно-технических изданиях</p> <p>ПК-1.2 Умеет применять полученные знания в области математики и информатики, а также, решать стандартные задачи собственной научно-исследовательской деятельности; умеет решать научные задачи с пониманием существующих подходов к верификации моделей по тематике исследований в соответствии с выбранной методикой</p> <p>ПК-1.3 Знает основы ведения научной дискуссии и формы устного научного высказывания; умеет вести корректную дискуссию в области информационных технологий, задавать вопросы и отвечать на поставленные вопросы по теме научного исследования; владеет навыками выступлений и научной аргументации при анализе объекта научной и профессиональной деятельности; способен принимать участие в работе научных семинаров, научно-технических конференций</p>
ПК-2	Организационное и технологическое обеспечение проектирования и дизайна ИС	<p>ПК-2.1. Знает этапы жизненного цикла разработки программных систем, виды архитектур, методологии разработки программных систем, основные понятия и определения, относящиеся к концепции построения информационных систем, методы проектирования информационных систем</p> <p>ПК-2.2. Умеет применять в практической деятельности профессиональные стандарты в области информационных технологий, проводить анализ требований к информационной системе, разрабатывать варианты реализации информационной системы, производить оценку качества, надежности и эффективности информационных систем</p>

3. Место дисциплины в структуре ОП ВО

Дисциплина «Методы интеллектуального анализа текстов» относится к части, формируемой участниками образовательных отношений блока Б1 ОП ВО.

В рамках ОП ВО обучающиеся также осваивают другие дисциплины и/или практики, способствующие достижению запланированных результатов освоения дисциплины «Методы интеллектуального анализа текстов».

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-1	Способность осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий	Введение в компьютерные науки и искусственный интеллект	Научно-исследовательская работа Технологическая (проектно-технологическая) практика
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	Введение в компьютерные науки и искусственный интеллект Основы компьютерной лингвистики	Научно-исследовательская работа Технологическая (проектно-технологическая) практика

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики	Последующие дисциплины/модули, практики
ОПК-1	Способность самостоятельно приобретать, развивать и применять математические, естественнонаучные, социально-экономические и профессиональные знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте	Введение в компьютерные науки и искусственный интеллект Методы машинного обучения Основы компьютерной лингвистики	Научно-исследовательская работа Технологическая (проектно-технологическая) практика
ОПК-2	Способность разрабатывать оригинальные алгоритмы и программные средства, в том числе с использованием современных интеллектуальных технологий, для решения профессиональных задач	Языки программирования для задач искусственного интеллекта Интеллектуальный анализ больших данных Глубокое обучение и обучение с подкреплением Прикладные методы компьютерной лингвистики	Научно-исследовательская работа Технологическая (проектно-технологическая) практика
ОПК-3	Способность самостоятельно приобретать, развивать и применять математические, естественнонаучные, социально-экономические и профессиональные знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте	Интеллектуальные системы и их применение Распознавание образов и обработка изображений Глубокое обучение и обучение с подкреплением Коллективное поведение интеллектуальных систем	Научно-исследовательская работа Технологическая (проектно-технологическая) практика
ОПК-4	Способность	Интеллектуальные	Научно-

Шифр	Наименование компетенции	Предшествующие дисциплины/модули, практики	Последующие дисциплины/модули, практики
	применять на практике новые научные принципы и методы исследований	системы и их применение	исследовательская работа Технологическая (проектно-технологическая) практика
ПК-1	Проведение работ по обработке и анализу научно-технической информации и результатов исследований	Введение в компьютерные науки и искусственный интеллект	Научно-исследовательская работа Технологическая (проектно-технологическая) практика
ПК-2	Организационное и технологическое обеспечение проектирования и дизайна ИС	Введение в компьютерные науки и искусственный интеллект Интеллектуальные системы и их применение	Научно-исследовательская работа Технологическая (проектно-технологическая) практика

4. Объем дисциплины и виды учебной работы

Общая трудоемкость дисциплины «Методы интеллектуального анализа текстов» составляет 5 зачетных единиц.

Таблица 4.1. Виды учебной работы по периодам освоения ОП ВО

Вид учебной работы	ВСЕГО, ак.ч.	Семестр(-ы)
		1
Контактная работа, ак.ч.	54	54
в том числе:		
Лекции (ЛК)	18	18
Лабораторные работы (ЛР)	-	-
Практические/семинарские занятия (СЗ)	36	36
Самостоятельная работа обучающихся, ак.ч.	99	99
Контроль (экзамен/зачет с оценкой), ак.ч.	27	27
Общая трудоемкость дисциплины	ак.ч.	180
	зач.ед.	5

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 интеллектуальный анализ текстов	Тема 1.1. Основные определения. Постановки задач обработки текстов и подходы к их решению.	ЛК, СЗ
	Тема 1.2. Естественный язык. Типология естественных языков. Уровни анализа.	ЛК, СЗ

Наименование раздела дисциплины	Содержание раздела (темы)	Вид учебной работы
	Лингвистическое исследование.	
Раздел 2. Методы лингвистического анализа текстов	Тема 2.1. Методы морфологического анализа. Проблема морфологической многозначности. Методы снятия морфологической многозначности.	ЛК, СЗ
	Тема 2.2. Синтаксис естественных языков. Способы описания синтаксической структуры предложения. Методы синтаксического анализа.	ЛК, СЗ
	Тема 2.3. Способы формализации семантики. Методы семантического анализа. Реляционно-ситуационный анализ текста. Дистрибутивная семантика.	ЛК, СЗ
Раздел 3. Прикладные задачи анализа текстов	Тема 3.1. Извлечение информации из текстов. Классификация и кластеризация текстов.	ЛК, СЗ
	Тема 3.2. Методы машинного обучения для решения задач обработки естественного языка	ЛК, СЗ

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
Лекционная	Аудитория для проведения занятий лекционного типа, оснащенная комплектом специализированной мебели; доской (экраном) и техническими средствами мультимедиа презентаций.	Аудитория оснащена комплектом специализированной мебели. Рабочие места обучающихся, технические средства: интерактивная доска Samsung, рабочая станция Samsung; выход в интернет через ЛВС и Wi-Fi; Программное обеспечение: продукты Microsoft (ОС, пакет офисных приложений, в т.ч. MS Office/ Office 365, Teams, Skype)
Семинарская	Аудитория для проведения занятий семинарского типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, оснащенная комплектом специализированной мебели и техническими	Аудитория оснащена комплектом специализированной мебели. Рабочие места обучающихся, технические средства: интерактивная доска Samsung, рабочая станция Samsung; выход в интернет через ЛВС и Wi-Fi; Программное обеспечение: продукты Microsoft (ОС, пакет офисных приложений, в т.ч. MS Office/ Office 365, Teams, Skype)

Тип аудитории	Оснащение аудитории	Специализированное учебное/лабораторное оборудование, ПО и материалы для освоения дисциплины
	средствами мультимедиа презентаций.	
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.	Дисплейный класс оснащен комплектом специализированной мебели. Рабочие места обучающихся, технические средства: экран Prostar 153*20, переносной проектор, рабочее место обучающегося (моноблок Lenovo) - 12; выход в интернет через ЛВС и Wi-Fi;. Программное обеспечение: продукты Microsoft (ОС, пакет офисных приложений, в т.ч. MS Office/ Office 365, Teams, Skype) Операционная система Linux (дистрибутив Gentoo): - офисный пакет LibreOffice (лицензия MPL-2.0)

7. Учебно-методическое и информационное обеспечение дисциплины

Основная литература:

1. Осипов Г.С. Лекции по искусственному интеллекту. Изд. 2-е, испр. и доп. М.: Книжный дом «ЛИБРОКОМ», 2009. – 272 с.
2. Смирнов И.В. Введение в анализ естественных языков: учебно-методическое пособие. Москва: Изд-во РУДН, 2014. – 85 с.

Дополнительная литература:

1. Боярский К.К. Введение в компьютерную лингвистику. Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики. 2013
2. Jurafsky D., Martin J. H. Speech and language processing. – Pearson, 2014.

Ресурсы информационно-телекоммуникационной сети «Интернет»:

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

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

Смирнов И.В. Введение в анализ естественных языков: учебно-методическое пособие. Москва: Изд-во РУДН, 2014. – 85 с.

8. Оценочные материалы и бально-рейтинговая система оценивания уровня сформированности компетенций по дисциплине

Оценочные материалы и бально-рейтинговая система оценивания уровня сформированности компетенций по итогам освоения дисциплины «Методы интеллектуального анализа текстов» представлены в Приложении (Фонд оценочных средств) к настоящей Рабочей программе дисциплины.

Разработчик:

доцент кафедры
информационных технологий



И.В. Смирнов

Заведующий кафедрой

информационных технологий



Ю.Н. Орлов

Руководитель программы

доцент кафедры
информационных технологий



М.Б. Фомин