

Программа разработана экспертами
Федерального учебно-методического объединения
высшего образования по укрупненной группе
специальностей и направлений подготовки
45.00.00 Языкознание и литературоведение

Утверждена на заседании ФУМО
25 мая 2021 года

Примерная программа учебной дисциплины

**ИНСТРУМЕНТЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА
ДЛЯ АНАЛИЗА ЯЗЫКА И ТЕКСТА**

Уровень высшего образования:

БАКАЛАВРИАТ

Направление подготовки:

45.03.01 «ФИЛОЛОГИЯ»

45.03.02 «ЛИНГВИСТИКА»

Раздел 1. Характеристики учебных занятий

1.1. Цели и задачи учебных занятий

Целью данного курса является знакомство с технологиями автоматической обработки естественного языка с целью обучения применению таких технологий в различных сферах филологической и лингвистической деятельности. В рамках курса предусматривается ознакомление учащихся с основными областями и задачами применения технологий искусственного интеллекта, а также с современными программными средствами для решения базовых филологических и лингвистических прикладных задач. Результатом занятий должно стать приобретение студентами навыков работы с приложениями для обработки естественного языка, системами автоматического перевода, анализа тональности, а также получение представлений о возможном применении данных средств и ресурсов при проведении исследований.

1.2 Место дисциплины (модуля) в структуре образовательной программы, связь с другими дисциплинами (модулями) программы

Относится к вариативной части ОПОП ВО.

1.3 Требования подготовленности обучающегося к освоению содержания учебных занятий (пререквизиты)

Устанавливаются образовательной организацией.

1.4 Перечень результатов обучения

По окончании курса студент должен уметь работать с технологиями искусственного интеллекта в области обработки естественного языка и грамотно применять их в соответствии с типом задачи, знать различные типы электронных лингвистических ресурсов, иметь навыки работы с электронными словарями, лингвистическими базами данных, приложениями для автоматического анализа морфологии и синтаксиса, системами автоматизированного перевода.

Набор компетенций, соотнесенных с результатами обучения, определяется образовательной организацией.

1.5 Перечень рекомендуемых образовательных технологий

В преподавании дисциплины «Инструменты искусственного интеллекта для анализа языка и текста» используются разнообразные образовательные технологии как традиционного, так и инновационного характера, учитывающие смешанный, теоретико- и практикоориентированный характер дисциплины:

- лекции;
- практические занятия;
- дискуссии;
- выступления с докладами и сообщениями;
- аудиторные контрольные работы;
- внеаудиторные контрольные работы;
- тестирование.

Степень необходимости образовательной среды и ее выбор определяется образовательной организацией. Формы текущей аттестации определяются образовательной организацией.

1.6 Объем дисциплины (модуля) в зачетных единицах

2 з.е.

Раздел 2. Организация, структура и содержание учебных занятий

2.1 Организация учебных занятий

Предусмотрены учебные занятия с использованием дистанционных технологий.

2.2 Краткая аннотация содержания дисциплины (модуля)

Наименование темы (раздела, части)	Вид учебных занятий	Кол-во часов
1. Инструменты автоматической обработки языка. Приложения для анализа фонетики. Системы распознавания речи. Приложения для автоматического анализа морфологии и синтаксиса. Средства представления лексических знаний; онтологии.	Определяется образовательной организацией	6
2. Аспекты интеллектуального анализа текстов. Текст как объект интеллектуального анализа. Распознавание и извлечение информации из текста. Анализ тональности. Автоматическое аннотирование и реферирование текста.	Определяется образовательной организацией	4
3. ИИ в переводческой деятельности. История работы над программами для автоматического перевода. Автоматизированный перевод и машинный перевод. Системы автоматизированного перевода. Модели машинного перевода. Программные средства помощи переводчику.	Определяется образовательной организацией	4
4. Стилеметрия и возможности ИИ Определение авторства как филологическая задача. Признаки, используемые в стилеметрии. Методы и подходы в современной стилеметрии. Анализ стиля и искусственный интеллект.	Определяется образовательной организацией	2
5. ИИ в филологическом образовании Использование элементов ИИ в обучении языку. Компьютерные обучающие программы.	Определяется образовательной организацией	2
ИТОГО		18

Раздел 3. Обеспечение учебных занятий

3.1 Методические указания по освоению дисциплины

Преподавание дисциплины осуществляется в форме лекционных и практических занятий, которые сопровождаются демонстрацией наглядного материала – специализированных приложений, Интернет-ресурсов и презентаций о них. Во время занятий обучающиеся выполняют практические задания. Для закрепления пройденного материала предлагаются домашние задания по каждой из тем. Успешное овладение содержанием дисциплины «Инструменты искусственного интеллекта для анализа языка и текста» предполагает работу обучающихся в группах в аудитории, а также их самостоятельную работу. Дополнительные методические указания устанавливаются образовательной организацией.

3.2 Примерный перечень учебно-методического обеспечения самостоятельной работы обучающихся по дисциплине (модулю), в том числе примерный перечень учебной литературы и ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины (модуля)

Самостоятельная работа студентов должна включать усвоение теоретического материала, подготовку к практическим занятиям, выполнение творческих заданий, работу с электронным учебно-методическим комплексом, подготовку к текущему контролю знаний, к промежуточной аттестации (зачету).

Список рекомендованной литературы

- Большакова Е. И., Воронцов К. В., Ефремова Н. Э., Клышинский Э. С., Лукашевич Н. В., Сапин А. С. 2017. Автоматическая обработка текстов на естественном языке и анализ данных: учеб. пособие. М.: Изд-во НИУ ВШЭ, 2017.
- Андреев В.С. 2012. Моделирование индивидуального стиля (на основе лингвистических характеристик): монография. М.: Флинта – Наука.
- Бовтенко М.А. 2005. Компьютерная лингводидактика: учеб. пособие. М.: Флинта: Наука.
- Введение в электронные лингвистические ресурсы [Электронный ресурс] / сост. В. Е. Гольдин, О. Ю. Крючкова. Саратов: 2011.
- Герд А. С. 2005. Прикладная лингвистика. СПб: Изд-во Санкт-Петербургского ун-та.
- Добров Б. В., Иванов В. В., Лукашевич Н. В., Соловьев В. Д. 2009. Онтологии и тезаурусы: модели, инструменты, приложения. Изд-во ИНТУИТ.
- Зубов А.В., Зубова И.И. 2007. Основы искусственного интеллекта для лингвистов: учебное пособие для вузов. Москва: Логос.
- Леонтьева Н.Н. 2006. Автоматическое понимание текстов: системы, модели, ресурсы. М. Издательский центр «Академия».
- Мартыненко Г.Я. 1988 Основы стилеметрии. Л.: Изд-во Ленинградского ун-та.
- Марусенко М.А. 1990. Атрибуция анонимных и псевдонимных литературных произведений методами теории распознавания образов. Л.: Изд-во Ленинградского ун-та.
- Минин А. Я. 2016. Информационные технологии в образовании: Учебное пособие. Москва: Московский педагогический государственный университет.
- Осипов Г. С. 2011. Методы искусственного интеллекта. М.: Физматлит.
- Паттерсон Д., Гибсон А. 2018. Глубокое обучение с точки зрения практика. Москва: ДМК Пресс
- Прикладная и компьютерная лингвистика / Под ред. И. С. Николаева, О. В. Митрениной, Т. М. Ландо. Изд. 2-е. М.: ЛЕНАНД, 2017.
- Lucci S., Корес D. 2015. Artificial intelligence in the 21st century. Stylus Publishing, LLC.

Описание материально-технической базы, рекомендуемой для осуществления образовательного процесса по дисциплине (модулю)

Учебная аудитория с мультимедийным комплексом.

Описание материально-технической базы (в т.ч. программного обеспечения), рекомендуемой для адаптации электронных и печатных образовательных ресурсов для обучающихся из числа инвалидов и лиц с ОВЗ

Устанавливается образовательной организацией.

3.3 Методика проведения текущего контроля успеваемости и промежуточной аттестации и критерии оценивания

Для контроля усвоения данной дисциплины предусмотрен зачет. Мероприятия по текущему контролю знаний обучающихся проводятся в часы, отведенные для изучения дисциплины.

В течение семестра студентами выполняются практические и контрольные работы.

Порядок проведения зачета определяется ВУЗом.

3.4 Методические материалы для проведения текущего контроля успеваемости и промежуточной аттестации (контрольно-измерительные материалы, оценочные средства)

Примерные вопросы для самоконтроля:

1. История ИИ как научного направления.
2. Основные разделы теории и приложений ИИ
3. Взаимодействие искусственного интеллекта с другими научными направлениями
4. Обработка речевого сигнала в системах распознавания речи.
5. Синтаксический анализ текста.
6. Системы машинного перевода: история разработок, нерешенные проблемы, перспективы.
7. Требования, предъявляемые к системам представления и обработки знаний.
8. Семантические сети и графы. Фреймы.
9. Приобретение и формализация знаний. Трудности построения баз знаний.
10. Методы моделирования и обучения нейронных сетей.
11. Семантический анализ целого текста. Анализ тональности.

Примерный перечень вопросов к зачету (экзамену) по всему курсу:

1. Сферы использования ИИ в филологических и лингвистических исследованиях.
2. Основные приложения ИИ в филологии и языкознании. Автоматический анализ речевого сообщения. Системы распознавания речи.
3. Основные приложения ИИ в филологии и языкознании. Приложения для автоматического анализа морфологии и синтаксиса.
4. Основные приложения ИИ в филологии и языкознании. Средства представления лексических знаний; онтологии.
5. Текст как объект интеллектуального анализа. Распознавание и извлечение информации из текста. Вопросно-ответные системы.
6. Текст как объект интеллектуального анализа. Анализ тональности.
7. Текст как объект интеллектуального анализа. Автоматическое аннотирование и реферирование текста.
8. История работы над программами для автоматического перевода.
9. Автоматизированный перевод и машинный перевод. Принципы работы современных систем автоматического перевода.
10. Системы автоматизированного перевода. Модели машинного перевода. Программные средства помощи переводчику.
11. Определение авторства как филологическая задача. Признаки, используемые в стилиметрии.
12. Методы и подходы в современной стилиметрии. Анализ стиля и искусственный интеллект.

13. Использование элементов ИИ в обучении языку. Компьютерные обучающие программы.

3.5 Материально-техническое обеспечение

Минимально необходимый для реализации курса перечень материально-технического обеспечения включает лекционные аудитории (с компьютерным и видеопроекционным оборудованием для презентаций, средствами звуковоспроизведения и экраном, с выходом в Интернет).

3.6 Информационное обеспечение

Рекомендуемая обязательная литература

Зубов А.В., Зубова И.И. 2007. Основы искусственного интеллекта для лингвистов: учебное пособие для вузов. Москва: Логос.

Мартыненко Г.Я. 1988. Основы стилеметрии. Л.: Изд-во Ленинградского ун-та.

Рекомендуемая дополнительная литература

Большакова Е.И., Воронцов К.В., Ефремова Н.Э., Клышинский Э.С., Лукашевич Н.В., Сапин А.С. 2017. Автоматическая обработка текстов на естественном языке и анализ данных: учеб. пособие. М.: Изд-во НИУ ВШЭ.

Андреев В.С. 2012. Моделирование индивидуального стиля (на основе лингвистических характеристик): монография. М.: Флинта – Наука.

Бовтенко М.А. 2005. Компьютерная лингводидактика: учеб. пособие. М.: Флинта: Наука.

Введение в электронные лингвистические ресурсы [Электронный ресурс] / сост. В. Е. Гольдин, О. Ю. Крючкова. Саратов: 2011.

Герд А.С. 2005. Прикладная лингвистика. – СПб: Изд-во Санкт-Петербургского ун-та.

Добров Б.В., Иванов В.В., Лукашевич Н.В., Соловьев В.Д. 2009. Онтологии и тезаурусы: модели, инструменты, приложения. Изд-во ИНТУИТ.

Леонтьева Н. Н. 2006. Автоматическое понимание текстов: системы, модели, ресурсы. М.: Издательский центр «Академия»

Марусенко М.А. 1990. Атрибуция анонимных и псевдонимных литературных произведений методами теории распознавания образов. Л.: Изд-во Ленинградского ун-та.

Минин А. Я. 2016. Информационные технологии в образовании: Учебное пособие. Москва: Московский педагогический государственный университет.

Осипов Г. С. 2011. Методы искусственного интеллекта. М.: Физматлит.

Паттерсон Д., Гибсон А. 2018. Глубокое обучение с точки зрения практика. Москва: ДМК Пресс.

Прикладная и компьютерная лингвистика / Под ред. И. С. Николаева, О. В. Митрениной, Т. М. Ландо. Изд. 2-е. М.: ЛЕНАНД, 2017.

Lucci S., Kopec D. 2015. Artificial intelligence in the 21st century. Stylus Publishing, LLC.

Рекомендуемый перечень иных информационных источников

1. British National Corpus. <http://www.natcorp.ox.ac.uk/>
2. International Computer Archive of Modern and Medieval English <http://icame.uib.no/>
3. Атлас языков мира. <http://wals.info/>
4. База данных «Языки мира» www.dblang.ru
5. Коллекция словарей Института русского языка им. В.В. Виноградова РАН <http://slovari.ru>
6. Корпуса английского языка <https://www.english-corpora.org/>
7. Лингвистические данные Linguistic data consortium: <https://www ldc.upenn.edu/language-resources>
8. Лингвистический процессор «ЭТАП-3» <http://proling.iitp.ru/ru/etap>
9. Машинный фонд русского языка <http://cfri.ru>

10. Национальный корпус русского языка. <http://ruscorpora.ru/>
11. Новый частотный словарь русской лексики. <http://dict.ruslang.ru/freq.php>
12. Программное обеспечение в области автоматической обработки текста. www.aot.ru
13. Проект «Вавилонская башня» <http://starling.rinet.ru>
14. Русский ворднет <http://wordnet.ru/>
15. Системы автоматического аннотирования текстов www.copernic.com
16. Словари, созданные на основе Национального корпуса русского языка. <http://dict.ruslang.ru/>
17. Тезаурус английского языка WordNet <http://wordnet.princeton.edu>
18. Технологии Академии «Яндекс» <http://company.yandex.ru/technologies/search>
19. Упсальский корпус русского языка. <https://snd.gu.se/en/catalogue/study/ext0071>
20. Хельсинский аннотированный корпус русских текстов: <http://h248.it.helsinki.fi/hanco/>
21. Чешский национальный корпус. <https://ucnk.ff.cuni.cz/cs/>
22. Каталог лингвистических ресурсов CLARIN <https://www.clarin.eu/>
23. Каталог лингвистических ресурсов ELRA <http://www.elra.info/en/>.

Раздел 4. Разработчики программы

Рабочая группа ФУМО 45.00.00 по проблемам искусственного интеллекта в языкознании и литературоведении.

