

Документ подписан простой электронной подписью
Информация о владельце:

ФИО: Исаев Игорь Магомедович

Должность: Проректор по учебной и научной работе

Дата подписания: 28.09.2023 12:42:08

Уникальный идентификатор документа:

d7a26b9e8ca85e98ec3de2eb454b4659d061f249

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

**Федеральное государственное автономное образовательное учреждение
высшего образования**

«Национальный исследовательский технологический университет «МИСИС»

Рабочая программа дисциплины (модуля)

Управление машинным переводом

Закреплена за подразделением

Кафедра иностранных языков и коммуникативных технологий

Направление подготовки

45.04.02 ЛИНГВИСТИКА

Профиль

Цифровая лингвистика и локализация

Квалификация

Магистр

Форма обучения

очная

Общая трудоемкость

3 ЗЕТ

Часов по учебному плану

108

Формы контроля в семестрах:

в том числе:

зачет с оценкой 3

аудиторные занятия

26

самостоятельная работа

82

Распределение часов дисциплины по семестрам

Семестр (<Курс>.<Семестр на курсе>)	3 (2.1)		Итого	
	14			
Неделя	УП	РП	УП	РП
Вид занятий	УП	РП	УП	РП
Лекции	13	13	13	13
Практические	13	13	13	13
Итого ауд.	26	26	26	26
Контактная работа	26	26	26	26
Сам. работа	82	82	82	82
Итого	108	108	108	108

Программу составил(и):
асс., Гилин Михаил Игоревич

Рабочая программа

Управление машинным переводом

Разработана в соответствии с ОС ВО:

Самостоятельно устанавливаемый образовательный стандарт высшего образования - магистратура Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский технологический университет «МИСИС» по направлению подготовки 45.04.02 ЛИНГВИСТИКА (приказ от 02.04.2021 г. № 119 о.в.)

Составлена на основании учебного плана:

45.04.02 ЛИНГВИСТИКА, 45.04.02-МЛГ-23-3.plx Цифровая лингвистика и локализация, утвержденного Ученым советом НИТУ МИСИС в составе соответствующей ОПОП ВО 22.06.2023, протокол № 5-23

Утверждена в составе ОПОП ВО:

45.04.02 ЛИНГВИСТИКА, Цифровая лингвистика и локализация, утвержденной Ученым советом НИТУ МИСИС 22.06.2023, протокол № 5-23

Рабочая программа одобрена на заседании

Кафедра иностранных языков и коммуникативных технологий

Протокол от г., №

Руководитель подразделения Бондарева Лилия Владимировна, к.полит.н., доцент

1. ЦЕЛИ ОСВОЕНИЯ

1.1	Целью дисциплины принято освоение студентами инструментами тренировки лингвистических нейросетей, оценка эффективности и перетренировка для обеспечения лучших показателей по автоматическим метрикам качества.
-----	---

2. МЕСТО В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Блок ОП:		Б1.В
2.1	Требования к предварительной подготовке обучающегося:	
2.1.1	Контроль качества переводческих проектов	
2.1.2	Основы скорочтения	
2.1.3	Способы быстрого запоминания	
2.1.4	Нормативная база в области лингвистики	
2.1.5	Профильные интерфейсы лингвиста	
2.1.6	Современные подходы к управлению командами	
2.1.7	Форматирование и верстка	
2.2	Дисциплины (модули) и практики, для которых освоение данной дисциплины (модуля) необходимо как предшествующее:	
2.2.1	NLP-аналитика	
2.2.2	Подготовка к процедуре защиты и защита выпускной квалификационной работы	
2.2.3	Производственная практика (педагогическая)	
2.2.4	Редактирование медиаконтента	
2.2.5	Субтитрование и транскрибирование	

3. РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ, СООТНЕСЕННЫЕ С ФОРМИРУЕМЫМИ КОМПЕТЕНЦИЯМИ

ПК-2: Способен автоматизировать лингвистические и локализационные проекты, а также управлять производственным процессом перевода	
Знать:	
ПК-2-33	Подходы к оценке качества готовых лингвистических моделей
ПК-2-32	Принципы поиска данных для создания лингвистических корпусов
ПК-2-31	Принципы функционирования лингвистических нейросетей
Уметь:	
ПК-2-У3	Кастомизировать готовые модели
ПК-2-У4	Оценивать качество моделей с помощью автоматических метрик
ПК-2-У1	Искать данные для создания лингвистических корпусов
ПК-2-У2	Подготавливать корпуса

4. СТРУКТУРА И СОДЕРЖАНИЕ

Код занятия	Наименование разделов и тем /вид занятия/	Семестр / Курс	Часов	Формируемые индикаторы компетенций	Литература и эл. ресурсы	Примечание	КМ	Выполняемые работы
	Раздел 1. Вводный раздел							
1.1	Принципы функционирования лингвистических сетей /Лек/	3	5	ПК-2-31	Л2.1 Э1		КМ1	
1.2	Принципы функционирования лингвистических сетей /Ср/	3	6		Л2.1 Э1			
	Раздел 2. Поиск данных для тренировки лингвистических нейросетей							

2.1	Принципы поиска данных для дальнейшей тренировки моделей /Лек/	3	4	ПК-2-32	Л2.1 Э1		КМ2	
2.2	Поиск данных для дальнейшей тренировки моделей /Пр/	3	4	ПК-2-У1	Л2.1 Э1		КМ3	
2.3	Самостоятельный поиск данных /Ср/	3	22	ПК-2-У1	Л2.1 Э1			
Раздел 3. Оценка качества готовых моделей								
3.1	Подходы к оценке качества лингвистических моделей /Лек/	3	4	ПК-2-33	Л2.1 Э1		КМ4	
3.2	Оценка качества готовых моделей посредством автоматических метрик /Пр/	3	4	ПК-2-У4	Л2.1 Э1		КМ5	
3.3	Самостоятельная оценка качества лингвистических моделей /Ср/	3	22		Л2.1 Э1			
Раздел 4. Практический кейс								
4.1	Подготовка корпуса и кастомизация готовой модели вместе с преподавателем /Пр/	3	5	ПК-2-У2 ПК-2-У3	Л2.1 Э1			Р1
4.2	Самостоятельная командная работа по подготовке корпуса и кастомизации модели под ключ на выбранную командой тематику /Ср/	3	32	ПК-2-У2 ПК-2-У3	Л2.1 Э1			

5. ФОНД ОЦЕНОЧНЫХ МАТЕРИАЛОВ

5.1. Контрольные мероприятия (контрольная работа, тест, коллоквиум, экзамен и т.п), вопросы для самостоятельной подготовки

Код КМ	Контрольное мероприятие	Проверяемые индикаторы компетенций	Вопросы для подготовки
КМ1	Опрос	ПК-2-31	Как функционируют нейросети? На основе чего готовятся данные? Для чего нужна тренировка? Для чего нужна кастомизация?
КМ2	Опрос	ПК-2-32	Как осуществляется поиск данных для создания корпуса? Где собирается информация? Как она структурируется?
КМ3	Самостоятельная работа	ПК-2-У1	Самостоятельный поиск данных для создания лингвистического корпуса
КМ4	Опрос	ПК-2-33	Какие подходы к оценке качества готовых лингвистических моделей существуют? Как трактуются полученные оценки? Какие выводы можно сделать из той или иной оценки?
КМ5	Самостоятельная работа		Самостоятельная оценка качества моделей с помощью автоматических метрик

5.2. Перечень работ, выполняемых по дисциплине (Курсовая работа, Курсовой проект, РГР, Реферат, ЛР, ПР и т.п.)

Код работы	Название работы	Проверяемые индикаторы компетенций	Содержание работы
Р1	Практический кейс	ПК-2-У2;ПК-2-У3	Командная работа по подготовке корпуса и кастомизации модели в рамках выбранной командой тематики

5.3. Оценочные материалы, используемые для экзамена (описание билетов, тестов и т.п.)

Экзамен не предусмотрен.

5.4. Методика оценки освоения дисциплины (модуля, практики. НИР)

КМ1 - 10 баллов
 КМ2 - 10 баллов
 КМ3 - 20 баллов
 КМ4 - 10 баллов
 КМ5 - 20 баллов

Р1 - 30 баллов

В течение семестра студент получает баллы за выполненные задания. Итоговая оценка по дисциплине осуществляется посредством конвертации итогового балла (процента) в оценку по следующей схеме:

0-50% - Неудовлетворительно

51-69% - Удовлетворительно

70-84% - Хорошо

85-100% - Отлично

6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ**6.1. Рекомендуемая литература****6.1.2. Дополнительная литература**

	Авторы, составители	Заглавие	Библиотека	Издательство, год
Л2.1	Копотев М.	Введение в корпусную лингвистику: учебное пособие	Электронная библиотека	Прага: Animedia Company, 2014

6.2. Перечень ресурсов информационно-телекоммуникационной сети «Интернет»

Э1	МЛГ 23-3 Управление машинным переводом	lms.misis.ru
----	--	--------------

6.3 Перечень программного обеспечения

П.1	Microsoft SQL server 2016
П.2	Microsoft Office
П.3	MS Teams
П.4	Python
П.5	WinRAR
П.6	Zoom
П.7	Microsoft Edge

6.4. Перечень информационных справочных систем и профессиональных баз данных**7. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ**

Ауд.	Назначение	Оснащение
Б-825	Компьютерный класс	16 рабочих мест для обучающихся, рабочее место для преподавателя, моноблоки HP EliteOne 800 G5 23.8 Product No. 5NW34AV, 1 маркерная доска, комплект учебной мебели

8. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ

Студенты получают возможность освоить практически взаимодействие с нейросетями: их тренировку, оценку качества получившихся моделей и последующий файнтюнинг.