

МИНОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Институт прикладной математики и компьютерных наук

УТВЕРЖДАЮ

Директор института прикладной
математики и компьютерных наук

А.В. Замятин

« 16 » мая 2022 г.



Оценочные материалы текущего контроля и промежуточной аттестации по дисциплине
(Оценочные средства по дисциплине)

Обработка естественного языка - II

по направлению подготовки

01.04.02 Прикладная математика и информатика

Направленность (профиль) подготовки:

Big Data and Data Science

ОС составил(и):

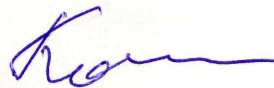
канд. техн. наук, доцент,
доцент кафедры теоретических основ информатики



М.С. Пожидаев

Рецензент:

д-р. техн. наук., профессор,
профессор кафедры теоретических основ информатики

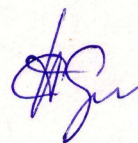


Ю.Л. Костюк

Оценочные средства одобрены на заседании учебно-методической комиссии
института прикладной математики и компьютерных наук (УМК ИПМКН)

Протокол от 12.05.2022 г. № 4

Председатель УМК ИПМКН,
д-р техн. наук, профессор



С.П. Сущенко

Оценочные средства (ОС) являются элементом системы оценивания сформированности компетенций у обучающихся в целом или на определенном этапе их формирования.

ОС разрабатываются в соответствии с рабочей программой (РП) дисциплины.

1. Компетенции и результаты обучения, формируемые в результате освоения дисциплины

Компетенция	Индикатор компетенции	Код и наименование результатов обучения (планируемые результаты обучения, характеризующие этапы формирования компетенций)	Критерии оценивания результатов обучения	
			Зачтено	Незачтено
УК-1. Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, выработать стратегию действий	ИУК-1.1 Выявляет проблемную ситуацию, на основе системного подхода осуществляет её многофакторный анализ и диагностику.	ОР-1.1.1. Умение применять основы прикладного системного анализа для выявления проблемных ситуаций и их причин.	Сформированные систематические знания	Отсутствие знаний
	ИУК-1.2 Осуществляет поиск, отбор и систематизацию информации для определения альтернативных вариантов стратегических решений в проблемной ситуации.	ОР-1.2.1. Навыки построения различных стратегий устранения обнаруженной проблемы и выбор наиболее перспективной из них.	Сформированные систематические знания	Отсутствие знаний
	ИУК-1.3 Предлагает и обосновывает стратегию действий с учетом ограничений, рисков и возможных последствий.	ОР-1.3.1. Способность выполнить необходимые действия на основе проведенных подготовительных мероприятий для устранения сложившейся проблемной ситуации.	Сформированные систематические умения	Отсутствие умений

УК-3. Способен организовывать и руководить работой команды, вырабатывая командную стратегию для достижения поставленной цели	ИУК-3.1 Формирует стратегию командной работы на основе совместного обсуждения целей и направлений деятельности для их реализации.	ОР-3.1.1. Навыки распределения ролей в команде на основе выявленных сильных и слабых сторон участников коллектива.	Сформированные систематические умения	Отсутствие умений
	ИУК-3.2 Организует работу команды с учетом объективных условий (технология, внешние факторы, ограничения) и индивидуальных возможностей членов команды	ОР-3.2.1. Умение планировать работы с построением графика сроков и оценок завершенности каждого этапа.		
	ИУК-3.3 Обеспечивает выполнение поставленных задач на основе мониторинга командной работы и своевременного реагирования на существенные отклонения.	ОР-3.3.1. Способность контроля прохождения этапов работ в соответствии с принятым графиком.		
ПК-6. Способен управлять получением, хранением, передачей, обработкой больших данных	ИПК-6.1 Осуществляет мониторинг и оценку производительности обработки больших данных.	ОР-6.1.1. Владение основами микросервисной архитектуры для построения горизонтально масштабируемых систем.	Сформированные систематические умения	Отсутствие умений
	ИПК-6.2 Использует методы и инструменты получения, хранения, передачи, обработки больших данных.	ОР-6.2.1. Навыки конфигурирования сетевых файловых хранилищ с поддержкой атомарных операций.		
	ИПК-6.3 Разрабатывает предложения по повышению производительности обработки больших данных.	ОР-6.3.1. Способность контроля целостности хранимой информации.		

2. Этапы формирования компетенций и виды оценочных средств

№	Этапы формирования компетенций (разделы дисциплины)	Код и наименование результатов обучения	Вид оценочного средства (тесты, задания, кейсы, вопросы и др.)
1.	Раздел 1. Введение, история развития дисциплины, решаемые задачи, подходы, методы и инструменты	ОР-1.1.1, ОР-1.2.1, ОР-1.3.1, ОР-3.1.1.	Выполнение практической работы №1
2.	Раздел 2. Предварительная обработка текстовых данных	ОР-1.1.1, ОР-3.2.1, ОР-6.1.1, ОР-6.3.1.	Выполнение практической работы №2
3.	Раздел 3. Вероятностные алгоритмы	ОР-1.2.1, ОР-1.3.1, ОР-6.1.1, ОР-6.3.1.	Выполнение практической работы №3
4.	Раздел 4. Формальные грамматики	ОР-1.1.1, ОР-1.3.1, ОР-3.2.1, ОР-3.3.1.	Выполнение практической работы №4
5.	Раздел 5. Векторное представление слов	ОР-1.1.1, ОР-1.2.1, ОР-3.2.1, ОР-3.3.1.	Выполнение практической работы №5
6.	Раздел 6. Модель Seq2seq	ОР-1.3.1, ОР-3.2.1, ОР-6.1.1, ОР-6.3.1.	Выполнение практической работы №6
7.	Раздел 7. Self-attention и Трансформер	ОР-1.1.1, ОР-3.3.1, ОР-6.1.1, ОР-6.3.1.	Выполнение практической работы №7
8.	Раздел 8. BERT и GPT-3	ОР-1.2.1, ОР-1.3.1, ОР-6.2.1, ОР-6.3.1.	Выполнение практической работы №8

3. Типовые контрольные задания или иные материалы, необходимые для оценки образовательных результатов обучения

3.1. Типовые задания для проведения текущего контроля успеваемости по дисциплине

Перечень практических работ:

Практическая работа № 1. Парсинг сайтов / использование api для получения текстовых данных.

Практическая работа № 2. Реализация стеммера Портера.

Практическая работа № 3. Использование библиотек для морфологического анализа, решение задачи частеречной разметки.

Практическая работа № 4. Векторное представление текста, word2vec, модели skip-gram и CBOW.

Практическая работа № 5. Тематическое моделирование с использованием библиотеки gensim.

Практическая работа № 6. Анализ тональности текстовых данных. Развертывание обученной модели в вебе.

Практическая работа № 7. Построение языковой модели, порождение текста.

Практическая работа № 8. Генерация подписи к изображению.

3.2. Типовые задания для проведения промежуточной аттестации по дисциплине

1. Сравните между собой работу аналитических формальных грамматик и вероятностных моделей.
2. Назовите сильные и слабые стороны моделей на основе искусственных нейронных сетей по сравнению с аналитическими алгоритмами.
3. Какие операции становятся возможными над словами, если выполнить переход от символического их представления к векторному?
4. Какие самые распространённые способы получения векторного представления слов можно выделить?
5. Какие проблемы рекуррентных моделей можно выделить для обоснования применения долгой краткосрочной памяти?
6. Какое ключевое преимущество можно назвать в механизме внутреннего внимания по сравнению с традиционным механизмом внимания?
7. На решение какой задачи ориентирован Трансформер в своей изначальной форме?
8. В чём идея fine tuning?
9. По каким причинам можно назвать модели BERT и GPT дополняющими друг друга?
10. В чём ключевая идея генеративных моделей?

4. Методические материалы, определяющие процедуры оценивания образовательных результатов обучения

4.1. Методические материалы для оценки текущего контроля успеваемости по дисциплине.

Текущий контроль успеваемости проводится во время сдачи практических работ. Каждая работа оценивается по следующим параметрам, согласно таблице раздела 1:

- полнота реализации практической работы,
- ответы на вопросы по практической работе,
- ответы на вопросы по теории из соответствующего раздела курса,
- умение исправлять ошибки и оперативно вносить изменения в практическую работу.

4.2. Методические материалы для проведения промежуточной аттестации по дисциплине.

«Зачтено» – студент выполнил все практические работы, ответил на все вопросы по практической работе;

«Не зачтено» – студент не сдал какие-либо практические работы, не ответил на вопросы по практической работе.

Во время зачета студент может повысить свою оценку, сдав заново соответствующую практическую работу.