



**МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
«Казанский государственный аграрный университет»
(ФГБОУ ВО Казанский ГАУ)**

Институт экономики
Кафедра цифровых технологий и прикладной информатики

УТВЕРЖДАЮ
Проректор по учебной
работе и цифровизации, доцент
_____ А.В. Дмитриев
«23» октября 2025 г.

**ФОНД ОЦЕНОЧНЫХ СРЕДСТВ
ДЛЯ ПРОВЕДЕНИЯ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ОБУЧАЮЩИХСЯ
ПО ДИСЦИПЛИНЕ
«Методы интеллектуального анализа естественного языка»
(Оценочные средства и методические материалы)**

приложение к рабочей программе дисциплины

Направление подготовки
09.04.03 Прикладная информатика

Направленность (профиль) подготовки
Анализ данных и искусственный интеллект

Форма обучения
очная

Паспорт
оценочных материалов для проведения текущего контроля и
промежуточной аттестации обучающихся по дисциплине
(модулю) Методы интеллектуального анализа естественного
языка

Перечень оценочных материалов и индикаторов достижения компетенций,
сформированность которых они контролируют

Наименование оценочного средства	Коды индикаторов достижения формируемых компетенции	Номер приложения
Собеседования по итогам выполнения лабораторных работ.	ИД-1 ОПК-2 ИД-2 ОПК-2 ИД-3 ОПК-2 ИД-1 ОПК-4 ИД-2 ОПК-4 ИД-3 ОПК-4	1
Зачет с оценкой	ИД-1 ОПК-2 ИД-2 ОПК-2 ИД-3 ОПК-2 ИД-1 ОПК-4 ИД-2 ОПК-4 ИД-3 ОПК-4	2

I. Текущий контроль

Приложение 1

Собеседования по итогам выполнения лабораторных работ.

1. Процедура выполнения лабораторных работ и собеседований по ним

Количество проводимых лабораторных работ в течение всего периода освоения дисциплины	8 работ
Формат проведения результатов	Бумажный отчет
Общее количество вопросов для проведения собеседования	4-10
Количество основных задаваемых при собеседовании вопросов	3-4
Формат проведения собеседования	Устно.

2. Шкала оценивания с учетом срока сдачи

Критерии оценивания	Балл
Студент правильно выполнил задание работы, продемонстрировал знания теоретического и практического материала	Отлично
Студент правильно выполнил задание работы, продемонстрировал неполные знания теоретического и практического материала	Хорошо
Студент выполнил задание работы, но допустил значительные неточности при выполнении, продемонстрировал неполные знания теоретического и практического материала	Удовлетворительно
Студент неправильно выполнил задание работы, не продемонстрировал знания теоретического и практического материала	Неудовлетворительно

3. Перечень лабораторных работ

Перечень лабораторных работ представлен в методическом указании для проведения лабораторных работ по данной дисциплине.

Список тем лабораторных работ:

1. Парсинг текстов. Базовая обработка текстов на Python. Регулярные выражения.
2. VOW представление текстов на естественном языке. Семантический спектр. Косинусное сходство.
3. Введение в нейронные сети. Обучение нейронной сети. Переобучение. Метрики качества
4. Полносвязные нейронные сети.
5. Классификация текстов с использованием полносвязных нейронных сетей
6. Классификация текстов с использованием эмбедингов, рекуррентных и сверточных нейронных сетей
7. Сегментация текста

8. Кластеризация текста.

Примерные вопросы при собеседовании.

1. Каким образом производилась предобработка текста.
2. Как формировалась обучающая и валидационная выборка.
3. Обоснуйте выбор архитектуры и параметров нейронной сети.
4. Аргументируйте выбор метрик качества предсказания модели и методов расчета ошибок предсказания.
5. Проанализируйте графики ошибки обучения модели на обучающей и валидационной выборке.
6. Какие способы повышения точности модели вы видите.
7. Проведите прогноз по построенным моделям.
8. Сравните качество вашей модели с предобученными эмбедами.
9. В чем основные недостатки вашей модели и какие существуют пути их решения.

Зачет с оценкой

1. Процедура проведения

К зачету допускаются учащиеся, успешно выполнившие все необходимые лабораторные работы и прошедшие собеседования по итогам их выполнения.

Общее количество вопросов к зачету	34 вопроса
Количество вопросов в билете	2 вопроса
Формат проведения	Устно и письменно

2. Шкала оценивания с учетом текущего контроля работы обучающегося в семестре

Критерии оценки уровня сформированности компетенций по дисциплине	Балл
Студент правильно ответил на основные и дополнительные вопросы, продемонстрировал знания теоретического и практического материала	Отлично
Студент правильно ответил на основные вопросы, продемонстрировал знания теоретического и практического материала	Хорошо
Студент ответил, большей частью правильно, на основные вопросы, продемонстрировал не полные знания теоретического и/или практического материала	Удовлетворительно
Студент неправильно выполнил задания и не продемонстрировал знания теоретического и практического материала	Неудовлетворительно

3. Вопросы и задачи (при необходимости) к зачету с оценкой

1. Автоматическая обработка естественного языка в кругу смежных дисциплин.
2. Особенности естественного языка и возможности его автоматической обработки.
3. Основные задачи автоматического анализа текстов и подходы к их решению.
4. Оценка систем автоматической обработки текстов.
5. Предобработка текста. Регулярные выражения.
6. Стеммеры, лемматизаторы, морфологические анализаторы.
7. Проблемы языковых моделей и способы их решения. Методы оценки языковых моделей.
8. Задачи разметки текста, применение разметки.
9. Классификация текстов: формулировка задачи и методы решения.
10. Наивный байесовский классификатор. Проблемы классификации текстов.
11. Информационный поиск и векторные модели текстов.
12. Задача парсинга, его применение. Синтаксис составляющих и синтаксис зависимостей.
13. Биологический и искусственный нейрон.
14. Основные функции активации нейронов. Преимущества нейронных сетей.
15. Классификации нейронных сетей, области применения и решаемые задачи.

16. Персептрон Розенблата.
17. Алгоритм обучения персептрона и правило Хебба.
18. Теорема о сходимости алгоритма обучения персептрона для линейноразделимых множеств. Проблема исключаящего «или».
19. Многослойный персептрон. Представление булевых функций.
20. Нейронные сети как универсальные аппроксиматоры.
21. Общая идея градиентных методов решения задач безусловной оптимизации. Метод наискорейшего спуска.
22. Алгоритм обратного распространения ошибки. Достоинства и недостатки алгоритма. Понятие паралича сети и причины его возникновения.
23. Эвристические приемы улучшения сходимости и качества градиентного обучения (нормализация, выбор функции активации, выбор начальных значений весов, порядок предъявления обучающих примеров, выбор величины шага, сокращение числа весов, выбивание из локальных минимумов, проблема переобучения и разделение выборки).
24. Методы упрощения структуры нейронной сети. Общие принципы обучения.
25. Задача кластеризации данных. Основные метрики для количественных и неколичественных переменных.
26. Подготовка и предобработка текстов перед подачей данных в нейросеть.
27. Предобработка данных в задачах кластеризации текстов.
28. Переобучение моделей машинного обучения. Причины, признаки. Способы решения данной проблемы.
29. Метрики качества. Методика выбора. Примеры.
30. Задача классификации текстов с использованием BOW. Реализация при помощи библиотеки Keras(либо TensorFlow, либо PyTorch).
31. Задача классификации текстов с использованием эмбеддингов. Реализация при помощи библиотеки Keras(либо TensorFlow, либо PyTorch).
32. Современные тенденции, направления решения задачи классификации текстов.
33. Косинусное расстояние между текстами. Семантический спектр. Пример реализации.
34. Морфологические анализаторы естественного языка на примере библиотек Python.