

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Соловьев Дмитрий Александрович
Должность: директор ФББОУ ВО «Саратовский университет»
Дата подписания: 14.09.2026 11:49:33
Уникальный программный ключ:
528682d78e671e56a607831e1ba2172f735a12

Приложение 1

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ



Федеральное государственное бюджетное образовательное
учреждение высшего образования
«Саратовский государственный университет
генетики, биотехнологии и инженерии имени Н.И. Вавилова»

УТВЕРЖДАЮ

И.о. заведующего кафедрой

 / Ключиков А.В./

« 16 » июня 2025 г.

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ

Дисциплина	ОБРАБОТКА ЕСТЕСТВЕННОГО ЯЗЫКА
Направление подготовки	38.04.05 Бизнес-информатика
Направленность (профиль)	Управление бизнес-анализом
Квалификация выпускника	Магистр
Нормативный срок обучения	2 года
Форма обучения	Заочная
Кафедра-разработчик	Цифровое управление процессами в АПК
Ведущий преподаватель	Шибайкин В.А., доцент

Разработчик: доцент, Шибайкин В.А.


(подпись)

Саратов 2025

Содержание

1. Перечень компетенций с указанием этапов их формирования в процессе освоения ОПОП.....	3
2. Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания.....	3
3. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы.....	5
4. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций.....	8

1. Перечень компетенций с указанием этапов их формирования в процессе освоения ОПОП

В результате изучения дисциплины «Обработка естественного языка» обучающиеся, в соответствии с ФГОС ВО по направлению подготовки 38.04.05 Бизнес-информатика, утвержденного приказом Министерства образования и науки РФ от 12.08.2020 г. № 990, формируют следующие компетенции, указанные в таблице 1.

Таблица 1

Формирование компетенций в процессе изучения дисциплины «Обработка естественного языка»

Компетенция		Индикаторы достижения компетенций	Этапы формирования компетенции в процессе освоения ОПОП (год)	Виды занятий для формирования компетенции	Оценочные средства для оценки уровня сформированности компетенции
Код	Наименование				
1	2	3	4	5	6
ПК-3	Способен выполнять технико-экономическое обоснование проектов по совершенствованию и регламентацию бизнес-процессов и ИТ-инфраструктуры предприятия	ПК-3.3 Решает аналитические задачи в сфере бизнеса с использованием современных технических средств и информационных технологий	2	лекции, лабораторные занятия	собеседование, доклад, типовое задание, тестирование

Профиль подготовки «Управление бизнес анализом»

Компетенция ПК-3 – также формируется в ходе освоения дисциплин:

Наука о данных для бизнеса, Управление цифровыми платформами и экосистемами современного бизнеса, Управление проектами (Agile и Scrum), Управление рисками, Управление изменениями, Ознакомительная практика, Научно-исследовательская работа, Подготовка к процедуре защиты и защита выпускной квалификационной работы.

2. Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания

Перечень оценочных материалов

Таблица 2

№ п/п	Наименование оценочного материала	Краткая характеристика оценочного материала	Представление оценочного средства в ОМ
1.	собеседование	средство контроля, организованное как специальная беседа педагогического работника с обучающимся на темы,	вопросы по темам дисциплины: – перечень вопросов

№ п/п	Наименование оценочного материала	Краткая характеристика оценочного материала	Представление оценочного средства в ОМ
		связанные с изучаемой дисциплиной и рассчитанной на выяснение объема знаний обучающегося по определенному разделу	для устного опроса – перечень вопросов для самостоятельной работы
2.	доклад	продукт самостоятельной работы студента, представляющий собой публичное выступление по представлению полученных результатов решения определенной учебно-практической, учебно-исследовательской или научной темы	темы докладов
3.	типовое задание	содержит средство проверки умений и навыков использования цифровых технологий, программного обеспечения для решения аналитических и управленческих задач.	банк типовых заданий

Программа оценивания контролируемой дисциплины

Таблица 3

№ п/п	Контролируемые разделы (темы дисциплины)	Код контролируемой компетенции (или ее части)	Наименование оценочного средства
1	2	3	4
1.	Работа с библиотеками обработки естественного языка, Реализация морфологического анализа, Построение синтаксического анализатора, Семантический анализ текста, Классификация текстов и извлечение сущностей	ПК-3	письменный опрос, типовое задание
2.	Разработка чат-бота, Применение методов анализа тональности	ПК-3	Доклад, письменный опрос, типовое задание

Описание показателей и критериев оценивания компетенций по дисциплине «Обработка естественного языка» на различных этапах их формирования, описание шкал оценивания

Таблица 4

Код компетенции, этапы освоения компетенции	Индикаторы достижения компетенции	Показатели и критерии оценивания результатов обучения			
		ниже порогового уровня (неудовлетворительно)	пороговый уровень (удовлетворительно)	продвинутый уровень (хорошо)	высокий уровень (отлично)
1	2	3	4	5	6

ПК-3 2 курс	ПК-3.3 Решает аналитические задачи в сфере бизнеса с использованием современных технических средств и информационных технологий	не выполнение простых задач (например, токенизация, извлечение ключевых слов) с минимальными ошибками, не завершение проекта или завершение проекта с минимальной функциональностью, не участие в обсуждениях по темам занятий	выполнение простых задач (например, токенизация, извлечение ключевых слов) с минимальным и ошибками, завершение проекта с ограниченной функциональностью и наличием ошибок в коде, регулярное присутствие, участие в обсуждениях, но без значительного вклада	знание ключевых концепций дисциплины и способность их объяснить, выполнение более сложных задач (например, анализ тональности, построение моделей) с незначительными ошибками, завершение проекта с хорошей функциональностью, но с некоторыми недочетами, активное участие в обсуждениях, готовность делиться своими мыслями и идеями. Отлично	знание понимания теоретических вопросов дисциплины и способность критически анализировать различные подходы, успешное выполнение сложных задач (например, создание чат-ботов, использование нейронных сетей) без ошибок, завершение высококачественного проекта с оригинальной идеей и без ошибок в коде, проактивное участие, ведение обсуждений, помощь однокурсникам и предложение новых идей.
----------------	--	--	---	---	---

3. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы

3.1. Доклады

Рекомендуемая тематика докладов по дисциплине приведена в таблице 5.

Таблица 5

Темы докладов, рекомендуемые к написанию при изучении дисциплины
«Обработка естественного языка»

№ п/п	Темы докладов
1	2
1.	Компьютерная лингвистика и обработка естественного языка: достижения и перспективы
2.	Изучение методов и технологий, используемых для автоматической классификации текстов на естественном языке, включая рубрицирование сообщений и распознавание спама.
3.	Исследование автоматических методов извлечения информации из текстов, включая примеры использования в различных областях.
4.	Рассмотрение задач и методов синтеза текста как обратного процесса к анализу, включая современные подходы
5.	Синтез и анализ текстов на естественном языке
6.	Обзор методик и инструментов для интеллектуального анализа текстов, а также их применение в различных сферах.
7.	Изучение процессов компрессии текста, включая сегментацию и выбор значимых фраз для формирования краткого изложения.

3.2. Типовое задание

Тематика типовых расчетов устанавливается в соответствии с разделами дисциплины. Типовой расчет подразделяется на два этапа. На первом этапе проводится теоретически обзор по теме занятия. На втором этапе решается задача по данной теме. Для каждой темы предусмотрен 2 варианта задания.

Пример типового задания

Тема 1. Компьютерная программа. Классификация текстов и извлечение сущностей:

1. Подготовить обучающую выборку.
2. Реализовать составление словаря уникальных слов в выборке.
3. Реализовать лемматизацию, токенизацию слов.
4. Реализовать удаление стоп слов.
5. Реализовать архитектуру CBOW
6. Реализовать архитектуру skip-gram.
7. Реализовать процедуру получения семантически близких слов skipgram (поиск слова по окружению)
8. Реализовать процедуру получения семантически близких слов cbow (поиск окружения по слову)
9. Реализовать процедуру сравнения семантического расстояния между словами.
10. Реализовать процедуру оценки важности слов в запросе

Практическая работа «Векторное представление слов»

- вектора для слов

- Используя `import gensim`, необходимо реализовать вычисление десяти самых близких по смыслу слов, находящихся в окрестности от результата операций сложения и вычитания в векторной модели. Каждому студенту преподавателем будет дана пара слов и необходимо найти такую линейную комбинацию исходных слов, чтобы в результате вычислений заданная пара попадала в первую десятку.

3.5 Промежуточная аттестация

В соответствии с учебным планом по направлению подготовки 38.04.05 Бизнес-информатика вид промежуточной аттестации – зачет

Целью проведения зачета по дисциплине является определение фактического уровня теоретических знаний и навыков обучающихся.

В вопросах на зачет расчетные задания отсутствуют.

Вопросы, выносимые на зачет

1. Что такое обработка естественного языка (ОЕЯ)?
2. Какие основные задачи решаются в ОЕЯ?
3. Какова роль предобработки данных в ОЕЯ?
4. Что такое токенизация и зачем она нужна?
5. Чем отличается стемминг от лемматизации?
6. Как работает метод *Bag of Words*?
7. Что такое *TF-IDF* и как он используется?
8. Каковы основные методы векторизации текста?
9. Что такое синтаксический анализ в ОЕЯ?
10. Что включает в себя обработка естественного языка?
11. Какие библиотеки вы можете использовать для токенизации текста?
12. Как реализовать метод *TF-IDF* на примере небольшого корпуса текстов?
13. Как создать простую нейронную сеть для обработки текста?
14. Как настроить параметры модели *LSTM* для обучения?
15. Каковы принципы обработки естественного языка?
16. Как использовать машинное обучение для классификации текстов?
17. Какие задачи можно решать с помощью обработки естественного языка?
18. Сформулируйте задачи обработки естественного языка
19. Сформулируйте принципы кодирования текстовых документов
20. Биологический нейрон и его математическая модель.
21. Как подготовить текстовые данные перед их анализом?
22. Какова цель обработки естественного языка ОЕЯ в AI?
23. Какие инструменты и библиотеки для ОЕЯ вы знаете?
24. Как установить и настроить библиотеку *sprasy*?
25. Какие языки программирования обычно используются в ОЕЯ?

26. Как проводить сбор и фильтрацию текстовых данных?
27. Какие метрики вы можете использовать для оценки модели ОЕЯ?
28. Какова структура типичного проекта ОЕЯ?
29. Сформулируйте методы машинного обучения для обработки текстовых документов
30. Модели нейронных сетей для обработки текстовых документов
31. Каковы основные типы синтаксических грамматик?
32. Каковы основные принципы работы RNN?
33. Каковы преимущества использования моделей трансформеров?
34. Что такое BERT и в каких задачах он используется?
35. Какие этические аспекты следует учитывать в ОЕЯ?
36. Как применять BERT для классификации текстов?
37. Как провести анализ тональности текстовых данных?
38. Как оценить производительность модели, использующей ОЕЯ?
39. Как использовать Python для предобработки текстов?
40. Как управлять данными и хранить их для дальнейшего анализа?
41. Каков процесс самообучения по методам в ОЕЯ?
42. Что такое LSTM и почему он популярен в ОЕЯ?
43. Какие шаги необходимо предпринять для разработки чат-бота?
44. Как эффективно представлять и делиться своими результатами с коллегами?
45. Какие наиболее распространенные модели ОЕЯ вы знаете?
46. Что включает в себя процесс разработки системы ОЕЯ?
47. Как выявить и минимизировать предвзятость в ваших моделях?
48. Как использовать графические представления для анализа результатов?
49. Как обеспечить безопасность и конфиденциальность данных при работе с ОЕЯ?

4. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций

4.1 Процедуры оценивания знаний, умений, навыков и (или) опыта деятельности

Контроль результатов обучения студентов, этапов и уровня формирования компетенций по дисциплине «Обработка естественного языка» осуществляется через проведение текущего, выходного контролей и контроля самостоятельной работы

Формы текущего итогового контроля и контрольные задания для текущего контроля разрабатываются кафедрой исходя из специфики дисциплины, и утверждаются на заседании кафедры.

4.2 Критерии оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы

Описание шкалы оценивания достижения компетенций по дисциплине приведено в таблице 6.

Таблица 6

Уровень освоения компетенции	Отметка по системе (зачет)	Описание
высокий	«зачтено»	Обучающийся обнаружил всестороннее, систематическое и глубокое знание понимание теоретических вопросов дисциплины, создания чат-ботов, использования нейронных сетей для обработки естественного языка, предоставил высококачественный проекта с оригинальной идеей и без ошибок в коде
высокий	«зачтено»	Обучающийся обнаружил полное знание понимание теоретических вопросов дисциплины и способность критически анализировать различные подходы, успешно выполняет сложные задачи (например, создание чат-ботов, использование нейронных сетей) без ошибок, завершение высококачественного проекта с оригинальной идеей и без ошибок в коде
базовый	«зачтено»	Обучающийся обнаружил полное знание ключевых концепций дисциплины и обладает способностью их объяснить и выполнить более сложные задачи (например, анализ тональности, построение моделей) с незначительными ошибками, завершение проекта с хорошей функциональностью, но с некоторыми недочетами.
пороговый	«зачтено»	Обучающийся обнаружил знания вопросов дисциплины, выполнение простых задач (например, токенизация, извлечение ключевых слов) с минимальными ошибками, завершение проекта с ограниченной функциональностью и наличием ошибок в коде,
-	«не зачтено»	Обучающийся не раскрывает основных понятий вопросов дисциплины не выполняет простые задания например, токенизация, извлечение ключевых слов с минимальными ошибками, не предоставил проект с минимальной функциональностью

4.2.1. Критерии оценки устного ответа при промежуточной аттестации

При ответе на вопрос обучающийся демонстрирует:

знания: ключевых понятий, задач и приложений обработки естественного языка, а также исторического контекста, методов токенизации, стемминга, лемматизации и других техник подготовки текстовых данных., основ синтаксического разбора и семантического анализа текстов., алгоритмов классификации, регрессии и кластеризации в контексте обработки текстов, этических аспектов обработки естественного языка, включая предвзятость и конфиденциальность данных.

умения: интерпретировать полученные результаты, умение выявлять и обсуждать этические вопросы, связанные с использованием технологий обработки естественного языка (ОЕЯ).

владение навыками: работать с библиотеками и инструментами для обработки текстов, разработки архитектуры приложений и систем, основанных на ОЕЯ, исследования и изучения новых технологий и методов в области ОЕЯ, представлять результаты работы и обсуждать их с коллегами и профессионалами в области.

Критерии оценки*

отлично	обучающийся демонстрирует: – знание материала , практики применения материала, исчерпывающе и последовательно, четко и логично излагает материал, хорошо ориентируется в материале, не затрудняется с ответом при видоизменении заданий; – умение <i>формулировать четкие и однозначные определения по вопросам, используя программные продукты и среды разработки;</i> – успешное и системное владение навыками чтения и оценки задач
хорошо	обучающийся демонстрирует: – знание материала, не допускает существенных неточностей; – в целом успешное, но содержащие отдельные пробелы, умение <i>формулировать четкие и однозначные определения по вопросам, используя программные продукты и среды разработки;</i> – в целом успешное, но содержащее отдельные пробелы или сопровождающееся отдельными ошибками владение навыками <i>чтения и оценки задач.</i>
удовлетворительно	обучающийся демонстрирует: – знания только основного материала, но не знает деталей, допускает неточности, допускает неточности в формулировках, нарушает логическую последовательность в изложении программного материала; – в целом успешное, но не системное умение формулировать четкие и однозначные определения по вопросам, используя программные продукты и среды разработки;; – в целом успешное, но не системное владение навыками чтения и оценки <i>задач</i>
неудовлетворительно	обучающийся:

	- не знает значительной части программного материала, плохо ориентируется в материале по современным технологиям обработки естественного языка; - не умеет <i>формулировать четкие и однозначные определения по вопросам, используя программные продукты и среды разработки;</i> - обучающийся не владеет навыками чтения и оценки <i>Задач</i>
--	---

4.2.2. Критерии оценки доклада

При написании доклада обучающийся демонстрирует:

знания: по вопросам доклада

умения: анализировать и содержательно интерпретировать полученные результаты исследований.

владение навыками: представлять результаты работы и обсуждать их с коллегами и профессионалами в области.

Критерии оценки доклада

отлично	обучающийся демонстрирует: - знание исследуемой темы (доклад структурирован; использованы различные точки зрения по рассматриваемому вопросу, сделаны и аргументированы основные выводы, прослушивается самостоятельность суждений, основные понятия вопроса изложены подробно) - логичность и структурированность изложения материала; - расширенную электронную презентацию к докладу на 5 слайдов
хорошо	обучающийся демонстрирует: - знание темы доклада (доклад структурирован; использованы различные точки зрения по рассматриваемому вопросу, сделаны и аргументированы основные выводы) - расширенную электронную презентацию к докладу менее 5 слайдов
удовлетворительно	обучающийся демонстрирует: - неполное знание материала (в материале представлена одна точка зрения, отсутствует самостоятельность суждений) - не представлена электронная презентация
неудовлетворительно	обучающийся: не выполнил доклад

4.2.3. Критерии оценки выполнения типового задания

При выполнении контрольных (самостоятельных) работ обучающийся демонстрирует:

знания: различных подходов к векторизации текстов, таких как Bag of Words, TF-IDF, Word2Vec и GloVe, принципов работы RNN, LSTM и современных архитектур, таких как трансформеры (BERT, GPT).

умения: умение реализовывать методы техники подготовки данных для анализа текстов, Bag of Words и TF-IDF на практике, использовать алгоритмы для решения задач классификации и регрессии в ОЕЯ, умение применять модели BERT и GPT для выполнения языковых задач

владение навыками: программирования на языке Python для решения задач в области ОЕЯ, собирать, обрабатывать и анализировать текстовые данные, для оценки эффективности ОЕЯ-моделей.

отлично	обучающийся демонстрирует: – знания современных методов и программных средств выявления тенденций изменения признаков – умения анализировать и правильно интерпретировать, применять инструментарий программы, использовать интегрированную среду разработки, применять инструменты для работы. – владеет навыками использования технологии обработки естественного языка.
хорошо	обучающийся демонстрирует: – современных методов и программных средств выявления тенденций изменения признаков, – умения анализировать и правильно интерпретировать, применять инструментарий программы, использовать интегрированную среду разработки, применять инструменты для работы, – владеет навыками использования технологии обработки естественного языка
удовлетворительно	обучающийся демонстрирует: – знания последовательности решения задания, не всех используемых современных методов и программных средств выявления тенденций изменения признаков. – умения применять инструментарий программы, проводить расчеты, не приводящие к правильному числовому ответу. – владеет навыками современных методов и программных средств выявления тенденций изменения признаков, не может самостоятельно составить выводов по результатам решения задачи.
неудовлетворительно	обучающийся: – не знает современных методов и программных средств выявления тенденций изменения признаков – не умеет применять инструментарий программы, проводить расчеты, – не владеет навыками использования технологии обработки естественного языка.

4.2.5. Критерии оценки собеседования

При устном собеседовании обучающийся демонстрирует:

знания: Знание методов токенизации, стемминга, лемматизации и других техник подготовки текстовых данных.

умения: оценивать производительность моделей и интерпретировать полученные результаты.

владение навыками: анализом решений с точки зрения достижения целевых показателей решений.

Критерии оценки собеседования

отлично	обучающийся демонстрирует:
----------------	----------------------------

	- высокий уровень знаний в области обработки естественного языка в части искусственного интеллекта, тема при собеседования раскрыта полностью; - умение проводить сравнительный анализ методов и инструментальных средств обработки естественного языка. - владеет навыками поиска современных средств обработки естественного языка.
хорошо	обучающийся демонстрирует: - обучающийся показывает хороший уровень в области обработки естественного языка в части искусственного интеллекта, тема при выполнении собеседования раскрыта полностью, но содержит неточности; - умение представлять информационные технологии и программные средства для обработки естественного языка; - владеет навыками использования специальной терминологии
Удовлетворительно	обучающийся демонстрирует: - обучающийся показывает средний уровень знаний по теме собеседования, тема раскрыта на 50 % - умение представлять информационных технологий и программных средств для обработки естественного языка - владеет навыками малой части использования специальных терминов.
Неудовлетворительно	- обучающийся показывает низкий уровень знаний по теме собеседования, тема при выполнении собеседования не раскрыта, содержит недостоверную информацию, отсутствует специальная терминология

Разработчик: доцент, Шибайкин В.А.


(подпись)