

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Антипова Наталья Викторовна
Должность: и.о. директора филиала
Дата подписания: 29.06.2026 19:39:11
Уникальный программный ключ:
fae5412acb1bf810dc69e6bc004ac45622b84b3a

Приложение 6
к основной профессиональной образовательной программе
по направлению подготовки 09.03.02
Информационные системы и технологии
направленность (профиль) программы
Системы и технологии искусственного интеллекта

**Министерство науки и высшего образования Российской Федерации
федеральное государственное бюджетное образовательное учреждение
высшего образования
«Российский экономический университет имени Г.В. Плеханова»**

Улан-Баторский филиал РЭУ им. Г.В. Плеханова



Одобрено
На заседании Совета Улан-Баторского
филиала РЭУ им. Г.В. Плеханова
Протокол № 3 от 23 октября 2025 г.
Председатель совета
Н.В. Антипова

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ

по дисциплине Технологии обработки естественного языка

Направление подготовки 09.03.02 Информационные системы и технологии

Направленность (профиль) программы Системы и технологии искусственного интеллекта

Уровень высшего образования Бакалавриат

Год начала подготовки 2026

Улан-Батор – 2025 г.

Рабочая программа одобрена на заседании междисциплинарной кафедры Улан-Баторского филиала РЭУ им. Г.В. Плеханова

протокол № 3 от 23 октября 2025 г.

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ

по учебной дисциплине *Технологии обработки естественного языка*

ПЕРЕЧЕНЬ КОМПЕТЕНЦИЙ С УКАЗАНИЕМ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ И ЭТАПОВ ИХ ФОРМИРОВАНИЯ ПО ДИСЦИПЛИНЕ

Формируемые компетенции (код и наименование компетенции)	Индикаторы достижения компетенций (код и наименование индикатора)	Результаты обучения (знания, умения)	Наименование контролируемых разделов и тем
ПК-1. Способен выявлять, формировать и согласовывать требования к результатам аналитических работ с применением технологий больших данных.	ПК-1.1. Выявляет требования заказчика к результатам анализа, определяет возможности применения анализа больших данных в предметной области и конкретных задачах заказчика.	ПК-1.1. 3-1. Знает регламенты организации по оформлению требований к результатам аналитических исследований с использованием технологий больших данных. ПК-1.1. 3-2. Знает современный опыт использования анализа больших данных. ПК-1.1. 3-3. Знает основы бизнес-интеллекта, типы систем бизнес-интеллекта. ПК-1.1. У-1. Умеет проводить сравнительный анализ методов и инструментальных средств анализа больших данных. ПК-1.1. У-2. Умеет проводить анализ больших данных в соответствии с утвержденными требованиями к результатам аналитического исследования.	Тема 1. Введение в обработку естественного языка Тема 2. Технологические решения обработки естественного языка
	ПК-1.2. Согласовывает с заказчиком и утверждает требования к результатам аналитического исследования.	ПК-1.2. 3-1. Знает предметную область анализа больших данных в соответствии с требованиями заказчика. ПК-1.2. У-1. Умеет подготавливать документы, регламентирующие требования к результатам аналитического исследования с использованием технологий больших данных в соответствии с существующими регламентами организации.	Тема 1. Введение в обработку естественного языка Тема 2. Технологические решения обработки естественного языка

ПК-4. Способен проводить аналитическое исследование с применением технологий больших данных в соответствии с требованиями заказчика.	ПК-4.1. Осуществляет выбор методов и инструментальных средств анализа больших данных для проведения аналитических работ.	ПК-4.1. 3-1. Знает статистические модели. ПК-4.1. 3-2. Знает алгоритмы машинного обучения: обучение с учителем, обучение без учителя, полууправляемое обучение, обучение с подкреплением. ПК-4.1. 3-3. Знает семантический анализ: обработка естественного языка, сентиментный анализ, анализ текста. ПК-4.1. 3-4. Знает нейронные сети: полносвязные, сверточные и рекуррентные нейронные сети, методы обучения нейронных сетей, нейросетевые методы понижения размерности. ПК-4.1. 3-5. Знает анализ изображений, анализ сетей, анализ пространственных данных, анализ временных рядов. ПК-4.1. У-1. Умеет проводить сравнительный анализ методов и инструментальных средств анализа больших данных. ПК-4.1. У-2. Умеет программировать на языках высокого уровня, ориентированных на работу с большими данными: для статистической обработки данных и работы с графикой, для работы с разрозненными фрагментами данных в больших массивах, для работы с базами структурированных и неструктурированных данных. ПК-4.1. У-3. Умеет решать задачи классификации, кластеризации, регрессии, прогнозирования, снижения размерности и ранжирования данных.	Тема 1. Введение в обработку естественного языка Тема 2. Технологические решения обработки естественного языка Тема 3. Большие текстовые данные на естественных языках
---	---	--	---

МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ, ХАРАКТЕРИЗУЮЩИЕ ЭТАПЫ ФОРМИРОВАНИЯ КОМПЕТЕНЦИЙ

Перечень учебных заданий на аудиторных занятиях

Темы групповых дискуссий

Индикаторы достижения: ПК-1. Способен выявлять, формировать и согласовывать требования к результатам аналитических работ с применением технологий больших данных.

Тема 2. Технологические решения обработки естественного языка

1. Развитие методов обработки текстовых данных – от классических до современных.
2. Перспективы создания универсального («сильного») ИИ.
3. Возможные расширения теста Тьюринга.
4. Возможности современных систем обработки естественного языка.
5. Способна ли система распознавания и генерации текста заменить менеджера по продажам?
6. Возможен ли переход к SmallData в контексте обработки естественного языка?

Критерии оценки (в баллах):

- 19-20 баллов выставляется студенту, если он верно и в полном объеме способен выявлять, формировать и согласовывать требования к результатам аналитических работ с применением технологий больших данных; ответ хорошо структурирован; прекрасно освоен понятийный аппарат; уровень освоения компетенций соответствует продвинутому уровню.
- 15-18 баллов выставляется студенту, если он с незначительными замечаниями способен выявлять, формировать и согласовывать требования к результатам аналитических работ с применением технологий больших данных; ответ хорошо структурирован; хорошо освоен понятийный аппарат; уровень освоения компетенций соответствует повышенному уровню.
- 10-14 баллов выставляется студенту, если он на базовом уровне, с ошибками, способен выявлять, формировать и согласовывать требования к результатам аналитических работ с применением технологий больших данных; ответ слабо структурирован; освоен понятийный аппарат; уровень освоения компетенций соответствует базовому уровню.
- 0-9 баллов выставляется студенту, если он не способен выявлять, формировать и согласовывать требования к результатам аналитических работ с применением технологий больших данных; ответ слабо структурирован; не освоен понятийный аппарат; компетенции не сформированы.

Задания для текущего контроля

Комплект тестов/тестовых заданий

Индикаторы достижения: ПК-1. Способен выявлять, формировать и согласовывать требования к результатам аналитических работ с применением технологий больших данных.

Тема 1. Введение в обработку естественного языка

1. К задачам обработки естественного языка относятся:
 - а) токенизация
 - б) выделение именованных сущностей
 - в) оптическое распознавание знаков

г) морфологическая разметка

2. Как кодируется предложение «Боб купил цветы, чтобы подарить Мэри» в словаре [«Боб», «пошел», «на», «рынок», «чтобы», «купить», «цветы», «купил», «подарить», «Мэри»]?

а) [1, 1, 1, 1, 1, 1, 1, 0, 0, 0]

б) [1, 0, 0, 0, 1, 0, 0, 1, 1, 1]

в) [1, 0, 0, 0, 0, 1, 0, 1, 1, 1]

г) [1, 0, 0, 0, 0, 1, 0, 2, 1, 1]

3. Какой из разделов науки о языке имеет дело с отношениями знаков к тому, что они обозначают?

а) семантика

б) синтаксис

в) прагматика

4. Какой из разделов науки о языке рассматривает способы сочетания знаков, ведущие в конечном счете к порождению текстов?

а) семантика

б) синтаксис

в) прагматика

5. Какой из разделов науки о языке занимается отношением знак – человек (коммуникант или реципиент)?

а) семантика

б) синтаксис

в) прагматика

6. Какие недостатки есть у традиционного подхода в обработке естественного языка?

а) необходимость конструирования признаков вручную

б) методы часто являются «черными ящиками», требующими интерпретации результатов работы

в) необходимость использования внешних ресурсов для нормальной работы методов

7. Что такое «глубокое обучение»:

а) подход в обучении, использующий «глубокие» деревья, то есть имеющие большое количество уровней

б) подход в обучении, основанный на нейросетях с большим количеством скрытых слоев

в) подход в обучении, использующий глубокое понимание фундаментальных принципов

г) один из вариантов традиционного подхода обработки естественного языка

8. Расширением какого класса нейросетей являются LSTM-сети (long short-term memory)?

а) CNN (convolutional neural networks)

б) RNN (recurrent neural networks)

в) SLTM (short long-term memory)

г) FCNN (fully connected neural networks)

9. К какому классу сетей относятся сети типа ResNet?

а) глубокие модели

б) очень глубокие модели

в) модели с краткосрочной памятью

г) модели с долгой краткосрочной памятью

10. Какая функция НЕ может быть использована в нейросети в качестве функции активации?

- а) ReLU
- б) SeLU
- в) Gamma
- г) softmax

Критерии оценки (в баллах): 0,5 балла за верный ответ на вопрос (верно выявляет, формирует и согласовывает требования к результатам аналитических работ с применением технологий больших данных)

Комплект заданий для контрольной работы

Индикаторы достижения: ПК-1. Способен выявлять, формировать и согласовывать требования к результатам аналитических работ с применением технологий больших данных; ПК-4. Способен проводить аналитическое исследование с применением технологий больших данных в соответствии с требованиями заказчика.

Тема 2. Технологические решения обработки естественного языка

1. Провести лемматизацию и частотный анализ текста.
2. Создать нейронную сеть, позволяющую распознать автора текста с точностью не менее 80%.
3. Для задачи распознавания построить не менее 5 архитектур нейронных сетей одного типа, сравнить точность.
4. Улучшить текущий скрипт чатбота, внедрив блок кода для присвоения словам вне словаря(out-of-vocabulary) метки «unknown» так, чтобы встретив в запросе незнакомое слово, исполнение кода не останавливалось, а продолжалось, игнорируя «unknown» слова.
5. Скорректировать модель чатбота для учета знаков препинания в диалоге.

Тема 3. Большие текстовые данные на естественных языках

6. Оценить общий объем данных, который могло бы вместить наиболее емкое теоретически возможное хранилище данных.
7. Предложить алгоритм классификации предложений по выбранному признаку.
8. Предложить алгоритм машинного перевода с английского на русский язык.
9. Оценить количество слов в орфографическом словаре русского языка.
10. Предложить стандарт формата цифрового следа пользователя интернета и метрику в пространстве таких следов.

Критерии оценки (в баллах): 10 баллов за полностью выполненное задание (верно выявляет, формирует и согласовывает требования к результатам аналитических работ с применением технологий больших данных).

Задания для творческого рейтинга

Тематика групповых проектов

Индикаторы достижения: ПК-1. Способен выявлять, формировать и согласовывать требования к результатам аналитических работ с применением технологий больших данных;

ПК-4. Способен проводить аналитическое исследование с применением технологий больших данных в соответствии с требованиями заказчика.

1. Оценить рейтинг статьи habr.ru по тексту и заголовку.
2. Генерировать заголовок статьи habr.ru по тексту.
3. Генерация случайной осмысленной психологической статьи.
4. Чат-бот, который имитирует вашу манеру переписки (или другого человека).
5. Чат-бот, который имитирует манеру общения какого-то персонажа из фильма.
6. Доработка текста так, чтобы стилистика письма стала лучше.

Критерии оценки (в баллах):

– 19-20 баллов выставляется студенту, если он верно и в полном объеме способен выявлять, формировать и согласовывать требования к результатам аналитических работ с применением технологий больших данных; проводить аналитическое исследование с применением технологий больших данных в соответствии с требованиями заказчика; делаются самостоятельные выводы, достигнуты все результаты, указанные в задании, качество оформления отчета соответствует требованиям и при защите обучающийся проявил отличное владение материалом работы и способность аргументировано отвечать на поставленные вопросы по теме работы. Уровень освоения компетенций соответствует продвинутому уровню.

– 15-18 баллов выставляется студенту, если он с незначительными замечаниями способен выявлять, формировать и согласовывать требования к результатам аналитических работ с применением технологий больших данных; проводить аналитическое исследование с применением технологий больших данных в соответствии с требованиями заказчика; в работе достигнуты все результаты, указанные в задании, качество оформления отчета соответствует требованиям и при защите обучающийся проявил хорошее владение материалом работы и способность аргументировано отвечать на поставленные вопросы по теме работы. Уровень освоения компетенций соответствует повышенному уровню.

– 10-14 балл выставляется студенту, если он на базовом уровне, с ошибками, способен выявлять, формировать и согласовывать требования к результатам аналитических работ с применением технологий больших данных; проводить аналитическое исследование с применением технологий больших данных в соответствии с требованиями заказчика; в работе достигнуты основные результаты, указанные в задании, качество оформления отчета в основном соответствует требованиям и при защите обучающийся проявил удовлетворительное владение материалом работы и способность отвечать на большинство поставленных вопросов по теме работы. Уровень освоения компетенций соответствует базовому уровню.

– 0-9 баллов выставляется студенту, если он не способен выявлять, формировать и согласовывать требования к результатам аналитических работ с применением технологий больших данных; проводить аналитическое исследование с применением технологий больших данных в соответствии с требованиями заказчика; в работе не достигнуты основные результаты, указанные в задании, качество оформления отчета не соответствует требованиям, при защите обучающийся проявил неудовлетворительное владение материалом работы и не смог ответить на большинство поставленных вопросов по теме работы.

МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ, ХАРАКТЕРИЗУЮЩИЕ ЭТАПЫ ФОРМИРОВАНИЯ КОМПЕТЕНЦИЙ ВО ВРЕМЯ ПРОВЕДЕНИЯ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

Структура экзаменационного задания

Наименование оценочного средства	Максимальное количество баллов
Вопрос 1	20
Вопрос 2	20

Задания, включаемые в экзаменационный билет

Типовой перечень вопросов к экзамену:

1. Что такое обработка естественного языка? Классификация задач обработки естественного языка.
2. Традиционный подход к обработке естественного языка. Недостатки традиционного подхода.
3. Применение глубокого обучения в обработке естественного языка. Примеры.
4. Что такое представление и значение слова? Классические подходы к представлению слов.
5. Подход word2vec к изучению представления слова.
6. Алгоритм skip-gram.
7. Алгоритм CBOW.
8. Сравнение продуктивности алгоритмов skip-gram и CBOW.
9. Современные алгоритмы, расширяющие skip-gram и CBOW.
10. Основы сверточных нейронных сетей (CNN). Возможности CNN.
11. Устройство CNN. Операция свертки. Операция субдискретизации.
12. Классификация предложений с помощью CNN.
13. Устройство рекуррентной нейронной сети (RNN).
14. Сети с обратным распространением.
15. Применение RNN.
16. Генерация текста с помощью RNN.
17. Рекуррентные нейронные сети с контекстными признаками (RNN-CF).
18. Устройство и принцип работы сетей с долгой краткосрочной памятью (LSTM).
19. Модификация LSTM – лучевой поиск.
20. Генерация текста на уровне слов вместо n-грамм.
21. История машинного перевода.
22. Общие принципы нейронного машинного перевода.
23. Подготовка данных для системы NMT.
24. Метрика BLEU для оценки систем машинного перевода.
25. Современные традиции в обработке естественного языка.

Показатели и критерии оценивания планируемых результатов освоения компетенций и результатов обучения, шкала оценивания

Шкала оценивания		Формируемые компетенции	Индикатор достижения компетенции	Критерии оценивания	Уровень освоения компетенций
85 – 100 баллов	«отлично»/ «зачтено»	ПК-1	ПК-1.1	Знает верно и в полном объеме: регламенты организации по оформлению требований к результатам аналитических исследований с использованием технологий больших данных, современный опыт использования анализа больших данных, основы бизнес-интеллекта, типы систем бизнес-интеллекта. Умеет верно и в полном объеме: проводить сравнительный анализ методов и инструментальных средств анализа больших данных, проводить анализ больших данных в соответствии с утвержденными требованиями к результатам аналитического исследования.	Продвинутый
			ПК-1.2	Знает верно и в полном объеме: предметную область анализа больших данных в соответствии с требованиями заказчика. Умеет верно и в полном объеме: подготавливать документы, регламентирующие требования к результатам аналитического исследования с использованием технологий больших данных в соответствии с существующими регламентами организации.	
		ПК-4	ПК-4.1	Знает верно и в полном объеме: статистические модели, алгоритмы машинного обучения: обучение с учителем, обучение без учителя, полууправляемое обучение, обучение с подкреплением, семантический анализ: обработка естественного языка, sentimentный анализ, анализ текста, нейронные сети: полносвязные, сверточные и рекуррентные нейронные сети, методы обучения нейронных сетей, нейросетевые методы понижения размерности, анализ изображений, анализ сетей, анализ пространственных данных, анализ временных рядов. Умеет верно и в полном объеме: проводить сравнительный анализ методов и инструментальных средств анализа больших данных, программировать на языках	

				высокого уровня, ориентированных на работу с большими данными: для статистической обработки данных и работы с графикой, для работы с разрозненными фрагментами данных в больших массивах, для работы с базами структурированных и неструктурированных данных, решать задачи классификации, кластеризации, регрессии, прогнозирования, снижения размерности и ранжирования данных.	
70 – 84 баллов	«хорошо»/ «зачтено»	ПК-1	ПК-1.1	Знает с незначительными замечаниями: регламенты организации по оформлению требований к результатам аналитических исследований с использованием технологий больших данных, современный опыт использования анализа больших данных, основы бизнес-интеллекта, типы систем бизнес-интеллекта. Умеет с незначительными замечаниями: проводить сравнительный анализ методов и инструментальных средств анализа больших данных, проводить анализ больших данных в соответствии с утвержденными требованиями к результатам аналитического исследования.	Повышенный
			ПК-1.2	Знает с незначительными замечаниями: предметную область анализа больших данных в соответствии с требованиями заказчика. Умеет с незначительными замечаниями: подготавливать документы, регламентирующие требования к результатам аналитического исследования с использованием технологий больших данных в соответствии с существующими регламентами организации.	
		ПК-4	ПК-4.1	Знает с незначительными замечаниями: статистические модели, алгоритмы машинного обучения: обучение с учителем, обучение без учителя, полууправляемое обучение, обучение с подкреплением, семантический анализ: обработка естественного языка, sentimentный анализ, анализ текста, нейронные сети: полносвязные, сверточные и рекуррентные нейронные сети, методы обучения нейронных сетей,	

				нейросетевые методы понижения размерности, анализ изображений, анализ сетей, анализ пространственных данных, анализ временных рядов. Умеет с незначительными замечаниями: проводить сравнительный анализ методов и инструментальных средств анализа больших данных, программировать на языках высокого уровня, ориентированных на работу с большими данными: для статистической обработки данных и работы с графикой, для работы с разрозненными фрагментами данных в больших массивах, для работы с базами структурированных и неструктурированных данных, решать задачи классификации, кластеризации, регрессии, прогнозирования, снижения размерности и ранжирования данных.	
50 – 69 баллов	«удовлетворительно»/ «зачтено»	ПК-1	ПК-1.1	Знает на базовом уровне, с ошибками: регламенты организации по оформлению требований к результатам аналитических исследований с использованием технологий больших данных, современный опыт использования анализа больших данных, основы бизнес-интеллекта, типы систем бизнес-интеллекта. Умеет на базовом уровне, с ошибками: проводить сравнительный анализ методов и инструментальных средств анализа больших данных, проводить анализ больших данных в соответствии с утвержденными требованиями к результатам аналитического исследования.	Базовый
			ПК-1.2	Знает на базовом уровне, с ошибками: предметную область анализа больших данных в соответствии с требованиями заказчика. Умеет на базовом уровне, с ошибками: подготавливать документы, регламентирующие требования к результатам аналитического исследования с использованием технологий больших данных в соответствии с существующими регламентами организации.	
		ПК-4	ПК-4.1	Знает с незначительными замечаниями: статистические модели, алгоритмы машинного	

				обучения: обучение с учителем, обучение без учителя, полуправляемое обучение, обучение с подкреплением, семантический анализ: обработка естественного языка, sentimentный анализ, анализ текста, нейронные сети: полносвязные, сверточные и рекуррентные нейронные сети, методы обучения нейронных сетей, нейросетевые методы понижения размерности, анализ изображений, анализ сетей, анализ пространственных данных, анализ временных рядов. Умеет с незначительными замечаниями: проводить сравнительный анализ методов и инструментальных средств анализа больших данных, программировать на языках высокого уровня, ориентированных на работу с большими данными: для статистической обработки данных и работы с графикой, для работы с разрозненными фрагментами данных в больших массивах, для работы с базами структурированных и неструктурированных данных, решать задачи классификации, кластеризации, регрессии, прогнозирования, снижения размерности и ранжирования данных.	
менее 50 баллов	«неудовлетворительно»/ «не зачтено»	ПК-1	ПК-1.1	Не знает на базовом уровне: регламенты организации по оформлению требований к результатам аналитических исследований с использованием технологий больших данных, современный опыт использования анализа больших данных, основы бизнес-интеллекта, типы систем бизнес-интеллекта. Не умеет на базовом уровне: проводить сравнительный анализ методов и инструментальных средств анализа больших данных, проводить анализ больших данных в соответствии с утвержденными требованиями к результатам аналитического исследования.	Компетенции не сформированы
			ПК-1.2	Не знает на базовом уровне: предметную область анализа больших данных в соответствии с требованиями заказчика. Не умеет на базовом уровне: подготавливать документы, регламентирующие требования к результатам аналитического	

				исследования с использованием технологий больших данных в соответствии с существующими регламентами организации.	
		ПК-4	ПК-4.1	Не знает с незначительными замечаниями: статистические модели, алгоритмы машинного обучения: обучение с учителем, обучение без учителя, полуправляемое обучение, обучение с подкреплением, семантический анализ: обработка естественного языка, sentimentный анализ, анализ текста, нейронные сети: полносвязные, сверточные и рекуррентные нейронные сети, методы обучения нейронных сетей, нейросетевые методы понижения размерности, анализ изображений, анализ сетей, анализ пространственных данных, анализ временных рядов. Не умеет с незначительными замечаниями: проводить сравнительный анализ методов и инструментальных средств анализа больших данных, программировать на языках высокого уровня, ориентированных на работу с большими данными: для статистической обработки данных и работы с графикой, для работы с разрозненными фрагментами данных в больших массивах, для работы с базами структурированных и неструктурированных данных, решать задачи классификации, кластеризации, регрессии, прогнозирования, снижения размерности и ранжирования данных.	