

---

**УТВЕРЖДЕНА**

Решением Ученого совета  
АНО ВО «Центральный университет»  
«25» май 2026 г.  
Протокол №2

**Рабочая программа дисциплины (модуля)  
«Обработка естественного языка»**

**Направление подготовки:** 02.03.01 Математика и компьютерные науки

**Направленность (профиль) подготовки:** Прикладной искусственный интеллект

**Квалификация (степень) выпускника:** бакалавр

**Форма обучения:** очная

**Срок освоения программы:** 4 года

**Год набора:** 2026

**Москва  
2026**

## Содержание

|                                                     |    |
|-----------------------------------------------------|----|
| 1. Краткая характеристика дисциплины (модуля) ..... | 3  |
| 2. Перечень планируемых результатов обучения .....  | 5  |
| 3. Тематический план .....                          | 10 |
| 4. Содержание дисциплины (модуля) .....             | 10 |
| 5. Учебно-методическое обеспечение .....            | 11 |
| 6. Материально-техническое обеспечение .....        | 11 |
| 7. Методические и оценочные материалы .....         | 13 |

## **1. Краткая характеристика дисциплины (модуля)**

Рабочая программа дисциплины (модуля) «Обработка естественного языка» составлена в соответствии с федеральным государственным образовательным стандартом высшего образования – бакалавриат по направлению 02.03.01 «Математика и компьютерные науки», профиль «Прикладной искусственный интеллект», утвержденный приказом Министерства науки и высшего образования Российской Федерации № 807 от 23.08.2017 года.

Изучение дисциплины (модуля) «Обработка естественного языка» позволяет эффективно решать задачи автоматизации взаимодействия между человеком и машиной в реальных проектах. Обработка естественного языка является ключевой технологией в разработке интеллектуальных систем: чат-ботов, голосовых помощников, поисковых систем и аналитических платформ.

### **Место дисциплины (модуля) в структуре образовательной программы**

Настоящая дисциплина (модуль) включена в учебный план по программе подготовки бакалавриата по направлению 02.03.01 «Математика и компьютерные науки», профиль «Прикладной искусственный интеллект» и входит в вариативную часть Блока 1, формируемую участниками образовательных отношений, как дисциплина по выбору.

Дисциплина (модуль) изучается на 3 курсе в 5 семестре, доступна для прохождения при условии успешного завершения дисциплины (модуля) «Глубокое обучение».

**Цель изучения дисциплины (модуля):** формирование систематических знаний, практических умений и навыков в области анализа, обработки и генерации естественного языка с использованием современных методов машинного обучения и нейросетевых архитектур для решения прикладных задач в области искусственного интеллекта.

### **Задачи изучения дисциплины (модуля):**

- формирование знаний принципов представления и предобработки текстовой информации в задачах обработки естественного языка;
- формирование знаний архитектур ключевых моделей обработки текста, включая RNN, LSTM, Transformer и современные языковые модели (BERT, GPT и др.);
- формирование знаний методов решения стандартных задач NLP: классификация текста, извлечение сущностей, машинный перевод, суммаризация, вопросно-ответные системы;
- формирование умений и навыков обучения, настройки и оценки качества моделей обработки естественного языка с использованием современных библиотек и фреймворков;
- формирование умений интегрировать языковые модели в сложные пайплайны и многомодальные системы, включая взаимодействие с визуальной информацией.

### **В результате освоения дисциплины (модуля) обучающийся должен:**

#### **знать:**

- основные задачи области NLP и типы текстовых данных, включая однок документарные и многок документарные задачи;
- принципы токенизации, векторных представлений текста и языкового моделирования;
- архитектуру Трансформера, механизм внимания и различия между энкодерными, декодерными и encoder-decoder моделями;
- подходы к использованию промптинга, обучения с учителем и адаптации моделей под домен;
- принципы построения систем с дополненной генерацией (RAG), метрики качества и основные риски NLP-систем.

#### **уметь:**

- формулировать NLP-задачу, выбирать подходящий бейзлайн и корректные метрики оценки;
- строить и оценивать системы текстовой классификации и семантического поиска на основе эмбедингов;
- извлекать структурированную информацию из текстов по заданной схеме с использованием современных моделей;
- дообучать энкодерные модели с учителем для решения прикладных задач;
- проектировать пайплайн с дополненной генерацией, анализировать ошибки и обосновывать выбор архитектуры.

***владеть:***

- стеком Python и библиотеками: pandas, scikit-learn, Hugging Face Transformers/Datasets, sentence-transformers, PyTorch;
- методами построения TF-IDF и эмбединговых бейзлайнов для поиска и классификации;
- практиками воспроизводимых экспериментов и автопроверяемых систем оценивания;
- техниками структурированного извлечения данных и проектирования прототипов RAG-систем;
- навыками анализа компромиссов между качеством, латентностью, стоимостью и безопасностью NLP-решений.

## 2. Перечень планируемых результатов обучения

Компетенции, формируемые в результате освоения дисциплины (модуля) при проведении учебных занятий в форме контактной работы обучающихся с педагогическими работниками Университета и в форме самостоятельной работы обучающихся:

| Код и содержание компетенции                                                             | Индикатор компетенции и перечень планируемых результатов обучения по дисциплине (модулю)                                                                                                                                                                                                                                                                                                                                                                                            |
|------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| ППК-Р4. Способен использовать базы данных при создании программных модулей и компонентов | ППК-Р4.1. Пишет программный код с использованием языков определения и манипулирования данными в базах данных (базовый)                                                                                                                                                                                                                                                                                                                                                              |
|                                                                                          | ППК-Р4.2. 3-1. Знает современные подходы к проектированию реляционных и нереляционных баз данных.                                                                                                                                                                                                                                                                                                                                                                                   |
|                                                                                          | ППК-Р4.2. Проектирует базы данных для программных модулей и компонентов (средний)<br>ППК-Р4.2. 3-1. Знает современные подходы к проектированию реляционных и нереляционных баз данных.                                                                                                                                                                                                                                                                                              |
| ППК-Р5. Способен разрабатывать требования и проектировать программное обеспечение        | ППК-Р4.3. Оптимизирует производительность работы с базами данных (средний)<br>ППК-Р4.3. 3-1. Знает внутреннее устройство СУБД выбранной архитектуры.                                                                                                                                                                                                                                                                                                                                |
|                                                                                          | ППК-Р5.1. Анализирует возможности реализации требований к компьютерному программному обеспечению (средний)<br>ППК-Р5.1. 3-1. Знает возможности существующей программно-технической архитектуры.<br>ППК-Р5.1. 3-2. Знает возможности современных и перспективных средств разработки программных продуктов, технических средств.<br>ППК-Р5.1. У-3. Умеет вырабатывать варианты реализации требований к компьютерному программному обеспечению.                                        |
|                                                                                          | ППК-Р5.2. Разрабатывает технические спецификации на программные компоненты и их взаимодействие (средний)<br>ППК-Р5.2. 3-1. Знает методы и средства проектирования компьютерного программного обеспечения.<br>ППК-Р5.2. 3-2. Знает методы и средства проектирования программных интерфейсов.<br>ППК-Р5.2. У-1. Умеет выбирать средства реализации требований к компьютерному программному обеспечению.<br>ППК-Р5.2. У-3. Умеет проводить оценку и обоснование рекомендуемых решений. |
|                                                                                          | ППК-Р5.3. Проектирует компьютерное программное обеспечение (продвинутый)<br>ППК-Р5.3. 3-1. Знает принципы построения и виды архитектуры компьютерного программного обеспечения.<br>ППК-Р5.3. 3-2. Знает типовые решения, библиотеки программных модулей, шаблоны, классы объектов, используемые при разработке компьютерного программного обеспечения.<br>ППК-Р5.3. У-2. Умеет проектировать структуры данных.                                                                      |
|                                                                                          | ППК-Р5.4. Выполняет логическое проектирование системы (продвинутый)<br>ППК-Р5.4. 3-3. Знает базовые технологии взаимодействия и                                                                                                                                                                                                                                                                                                                                                     |

|                                                                                                   |                                                                                                                                                                                                                                                                                                             |
|---------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                                                                                   | <p>интеграции систем и компонентов.</p> <p>ППК-Р5.4. З-4. Знает методы функциональной декомпозиции ИТ-систем.</p> <p>ППК-Р5.4. У-1. Умеет описывать интеграции со смежными системами на логическом уровне.</p>                                                                                              |
| ППК-Р6. Способен участвовать в промышленной разработке программного обеспечения                   | <p>ППК-Р6.1. Работает в соответствии с промышленными методологиями разработки (средний)</p> <p>ППК-Р6.1. З-1. Знает принципы Agile и их применение в промышленных проектах.</p> <p>ППК-Р6.1. У-2. Умеет работать в команде с использованием инструментов управления проектами.</p>                          |
|                                                                                                   | <p>ППК-Р6.2. Использует инструменты промышленной разработки (средний)</p> <p>ППК-Р6.2. З-1. Знает принципы Continuous Integration and Continuous Delivery (CI/CD).</p> <p>ППК-Р6.2. У-1. Умеет настраивать потоки работ CI/CD.</p> <p>ППК-Р6.2. У-2. Умеет настраивать мониторинг в продуктивной среде.</p> |
|                                                                                                   | <p>ППК-Р6.3. Разрабатывает масштабируемый и поддерживаемый код (продвинутый)</p> <p>ППК-Р6.3. З-1. Знает принципы чистого кода, SOLID, DRY, KISS и др.</p> <p>ППК-Р6.3. У-1. Умеет разрабатывать модульный и тестируемый программный код.</p>                                                               |
|                                                                                                   | <p>ППК-Р6.4. Участвует в развертывании и поддержке программного обеспечения (продвинутый)</p> <p>ППК-Р6.4. З-1. Знает стратегии развертывания промышленного программного обеспечения.</p> <p>ППК-Р6.4. У-2. Умеет обнаруживать и устранять инциденты с работой продуктивного программного обеспечения.</p>  |
| ППК-Р7. Способен применять искусственный интеллект (ИИ) для генерации и отладки программного кода | <p>ППК-Р7.1. Применяет ИИ-инструменты для генерации программного кода (базовый)</p> <p>ППК-Р7.1. З-1. Знает принципы работы современных генеративных ИИ-моделей для генерации кода.</p> <p>ППК-Р7.1. У-1. Умеет формулировать корректные текстовые запросы (промты) для генерации кода.</p>                 |
|                                                                                                   | <p>ППК-Р7.2. Использует ИИ для анализа и отладки кода (средний)</p> <p>ППК-Р7.2. З-1. Знает методы ИИ-анализа кода.</p> <p>ППК-Р7.2. У-2. Умеет интерпретировать рекомендации ИИ по исправлению кода.</p>                                                                                                   |
|                                                                                                   | <p>ППК-Р7.3. Оптимизирует код с помощью ИИ (продвинутый)</p> <p>ППК-Р7.3. З-1. Знает методы ИИ-оптимизации.</p> <p>ППК-Р7.3. У-2. Умеет проверять корректность оптимизаций, предложенных ИИ.</p>                                                                                                            |
|                                                                                                   | <p>ППК-Р7.4. Оценивает этические и профессиональные аспекты применения ИИ в разработке (продвинутый)</p> <p>ППК-Р7.4. З-1. Знает этические нормы использования ИИ (конфиденциальность, плагиат кода и т.п.).</p> <p>ППК-Р7.4. З-2. Знает лицензионные ограничения сгенерированного кода.</p>                |

|                                                                                                                                                         |                                                                                                                                                                                                                                                                                                                                |
|---------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| ППК-ДА4. Способен выявлять бизнес-проблемы или бизнес-возможности                                                                                       | ППК-ДА4.1. Собирает информацию о бизнес-проблемах или бизнес-возможностях (средний)<br>ППК-ДА4.1. 3-3. Знает инструменты, техники анализа бизнес-ситуации и предметной области.<br>ППК-ДА4.1. У-1. Умеет собирать, классифицировать, систематизировать и обеспечивать хранение и актуализацию информации для бизнес-анализа.   |
|                                                                                                                                                         | ППК-ДА4.2. Выявление истинных бизнес-проблем или бизнес-возможностей (продвинутый)<br>ППК-ДА4.2. 3-2. Знает инструменты, техники анализа бизнес-ситуации и предметной области, включая анализ данных.<br>ППК-ДА4.2. У-1. Умеет выявлять и классифицировать бизнес-проблемы или бизнес-возможности.                             |
| ППК-ДА5. Способен применять методы статистического анализа и теорию эксперимента для планирования, проведения и интерпретации результатов экспериментов | ППК-ДА5.1. Разрабатывает дизайн эксперимента, включая формирование гипотез, определение метрик и размера выборки (базовый)<br>ППК-ДА5.1. 3-1. Знает основные принципы планирования экспериментов (рандомизация, контрольные группы, мощность теста).<br>ППК-ДА5.1. У-2. Умеет выбирать метрики для оценки эффекта воздействия. |
|                                                                                                                                                         | ППК-ДА5.2. Проводит статистический анализ данных эксперимента (проверка гипотез, расчёт доверительных интервалов) (средний)<br>ППК-ДА5.2. 3-1. Знает методы проверки гипотез (t-тест, $\chi^2$ , ANOVA).<br>ППК-ДА5.2. У-2. Умеет интерпретировать p-value и уровень значимости.                                               |
|                                                                                                                                                         | ППК-ДА5.3. Интерпретирует результаты экспериментов и формулирует выводы для принятия бизнес-решений (продвинутый)<br>ППК-ДА5.3. У-2. Умеет формулировать рекомендации на основе статистических выводов.                                                                                                                        |
| МО-2. Способен проектировать и внедрять системы автоматизации жизненного цикла машинного обучения (MLOps)                                               | МО-2.1. Проектирует и настраивает CI/CD-пайплайны для машинного обучения (базовый)<br>МО-2.1. 3-1. Знает принципы CI/CD применительно к моделям машинного обучения (тестирование, сборка, развертывание).<br>МО-2.1. У-1. Умеет настраивать CI/CD-пайплайны для МО с использованием инструментов.                              |
|                                                                                                                                                         | МО-2.2. Организует версионирование моделей, данных и экспериментов (средний)<br>МО-2.2. 3-1. Понимает системы версионирования моделей и данных<br>МО-2.2. У-2. Умеет использовать инструменты MLOps для воспроизводимости и масштабируемости экспериментов                                                                     |
|                                                                                                                                                         | МО-2.3. Внедряет системы мониторинга моделей в продуктивную эксплуатацию (средний)<br>МО-2.3. 3-1. Знает метрики мониторинга моделей МО<br>МО-2.3. У-1. Умеет развертывать системы мониторинга моделей МО                                                                                                                      |

|                                                                                                                          |                                                                                                                                                                                                                                                                                  |
|--------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                                                                                                          | <p>МО-2.4. Автоматизирует процессы переобучения и развертывания моделей (продвинутый)</p> <p>МО-2.4. 3-1. Знает стратегии автоматизации переобучения</p> <p>МО-2.4. У-1. Умеет автоматизировать переобучение моделей и А/В-тестирование развертываний</p>                        |
| РГ-1. Способен руководить процессами разработки компьютерного программного обеспечения                                   | <p>РГ-1.1. Руководит разработкой программного кода (средний)</p> <p>РГ-1.1. У-1. Способен распределять задачи на разработку программного кода между исполнителями</p>                                                                                                            |
|                                                                                                                          | <p>РГ-1.2. Руководит проверкой работоспособности компьютерного программного обеспечения (средний)</p> <p>РГ-1.2. У-2. Умеет оценивать результаты проверки работоспособности компьютерного программного обеспечения</p>                                                           |
|                                                                                                                          | <p>РГ-1.3. Руководит интеграцией программных модулей и компонентов компьютерного программного обеспечения (продвинутый)</p> <p>РГ-1.3. 3-2. Знает методы и программные интерфейсы взаимодействия компьютерного программного обеспечения с внешними программными компонентами</p> |
| РГ-2. Способен организовать процессы разработки компьютерного программного обеспечения                                   | <p>РГ-2.1. Управляет проектированием компьютерного программного обеспечения (средний)</p> <p>РГ-2.1. 3-1. Знает принципы построения архитектуры компьютерного программного обеспечения и виды архитектуры программного обеспечения</p>                                           |
|                                                                                                                          | <p>РГ-2.2. Управляет процессом разработки компьютерного программного обеспечения (продвинутый)</p> <p>РГ-2.2. У-1. Умеет составлять планы процесса разработки программного продукта</p>                                                                                          |
| РГ-4. Способен организовать промышленную разработку программного обеспечения                                             | <p>РГ-4.1. Внедряет промышленные методологии разработки (средний)</p> <p>РГ-4.1. 3-1. Знает принципы Agile и их применение в промышленных проектах</p>                                                                                                                           |
|                                                                                                                          | <p>РГ-4.3. Организует разработку масштабируемого и поддерживаемого кода (продвинутый)</p> <p>У-1. Умеет управлять разработкой модульного и тестируемого программного кода</p>                                                                                                    |
| РГ-5. Способен организовать процессы применения искусственного интеллекта (ИИ) для генерации и отладки программного кода | <p>РГ-5.1. Организует применение ИИ-инструментов для генерации программного кода (средний)</p> <p>РГ-5.1. 3-1. Знает принципы работы современных генеративных ИИ-моделей для генерации кода</p>                                                                                  |
|                                                                                                                          | <p>РГ-5.2. Организует применение ИИ для анализа и отладки кода (продвинутый)</p> <p>РГ-5.2. 3-1. Знает методы ИИ-анализа кода</p>                                                                                                                                                |
|                                                                                                                          | <p>РГ-5.3. Организует оптимизацию кода с помощью ИИ (продвинутый)</p> <p>РГ-5.3. 3-1. Знает методы ИИ-оптимизации</p>                                                                                                                                                            |
|                                                                                                                          | <p>РГ-5.4. Оценивает этические и профессиональные аспекты применения ИИ в разработке (продвинутый)</p> <p>РГ-5.4. 3-1. Знает этические нормы использования ИИ (конфиденциальность, плагиат кода и т.п.)</p>                                                                      |
|                                                                                                                          |                                                                                                                                                                                                                                                                                  |

|                                                                                                                                                                                                                                                                                       |                                                                                                                                                                                                                                                            |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>УК-4. (УНК-01.) Способен осуществлять деловую коммуникацию в устной и письменной формах на государственном языке Российской Федерации и иностранном(ых) языке(ах)<br/>(соответствует УНК-01 – Способен применять коммуникационные компетенции в профессиональной деятельности)</p> | <p>УНК-01.01. Осуществляет коммуникации в профессиональной деятельности (базовый)<br/>УНК-01.01.01. У-1 Владеет навыками презентации и публичной дискуссии<br/>УНК-01.01.01. У-3 Способен формулировать и понимать технологические и бизнес-требования</p> |
|                                                                                                                                                                                                                                                                                       | <p>УНК-01.02. Демонстрирует владение профессиональной культурой (базовый)<br/>УНК-01.02.01. У-3. Умеет грамотно оформлять документацию и вести коммуникацию в соответствии с профессиональными стандартами</p>                                             |
| <p>УК-3. (УНК-02.) Способен осуществлять социальное взаимодействие и реализовывать свою роль в команде<br/>(соответствует УНК-02 – Способен применять методы коллективной работы в профессиональной деятельности)</p>                                                                 | <p>УНК-02.01. Принимает участие в групповом взаимодействии в ходе профессиональной деятельности (базовый)<br/>УНК-02.01.01. У-5. Умеет работать в команде, проявлять инициативу и поддерживать коллег</p>                                                  |
|                                                                                                                                                                                                                                                                                       | <p>УНК-02.02. Проявляет лидерство и осуществляет наставничество (базовый)<br/>УНК-02.02.01. У-3. Умеет делиться знаниями и опытом, оказывать поддержку и осуществлять наставничество коллегам</p>                                                          |

### 3. Тематический план

| №п/<br>п                                  | Наименование<br>раздела дисциплины<br>(модуля) | Трудоемкость, академические часы |                         |          |                               | ТКУ<br>(текущий<br>контроль<br>успеваемости) |
|-------------------------------------------|------------------------------------------------|----------------------------------|-------------------------|----------|-------------------------------|----------------------------------------------|
|                                           |                                                | Очная форма                      |                         |          |                               |                                              |
|                                           |                                                | Аудиторная работа                |                         | Контроль | Самостояте<br>льная<br>работа |                                              |
|                                           |                                                | Лекции                           | Практические<br>занятия |          |                               |                                              |
| 1                                         | Основы NLP                                     | 10                               | 10                      |          | 40                            | Домашние<br>задания,<br>Тесты                |
| 2                                         | Трансформеры и LLM                             | 8                                | 8                       | 2        | 34                            | Домашние<br>задания,<br>Тесты                |
| 3                                         | Прикладной NLP                                 | 12                               | 12                      | 2        | 48                            | Домашние<br>задания,<br>Тесты                |
|                                           | Экзамен                                        |                                  |                         | 4        |                               |                                              |
| Итого:                                    |                                                | 30                               | 30                      | 8        | 122                           |                                              |
| Объем дисциплины (модуля)<br>(в ак. ч.)   |                                                | 190                              |                         |          |                               |                                              |
| Объем дисциплины (модуля)<br>(в зач. ед.) |                                                | 5                                |                         |          |                               |                                              |

### 4. Содержание дисциплины (модуля)

| №п/п | Наименование раздела дисциплины (модуля) | Содержание дисциплины (модуля) по темам                                                                                                                                                  |
|------|------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1    | Основы NLP                               | Введение в область NLP. Основы обработки текста. Текстовые представления и семантическая близость. Извлечение структурированной информации из текста. Нейросетевые подходы в области NLP |
| 2    | Трансформеры и LLM                       | Архитектура трансформера. Трансформеры в дискриминативных и генеративных задачах. Адаптация и пост-обучение больших языковых моделей                                                     |
| 3    | Прикладной NLP                           | Поиск и ранжирование. Основы методов дополненной генерации (RAG). Оценка качества в области NLP. Оптимизация больших языковых моделях NLP-системы.                                       |

## 5. Учебно-методическое обеспечение

Университет располагает полным набором лицензионного и свободно распространяемого программного обеспечения, включая продукты отечественного производства.

Каждый студент в течение всего периода обучения получает индивидуальный неограниченный доступ к электронно-библиотечной системе и электронной информационно-образовательной среде университета. Эти системы предоставляют возможность доступа к ресурсам из любой точки, где есть подключение к сети Интернет, как на территории университета, так и за его пределами.

Студентам обеспечен удаленный доступ к современным профессиональным базам данных и информационным справочным системам.

### *Основная литература:*

1. Бессмертный, И. А. Интеллектуальные системы : учебник и практикум для среднего профессионального образования / И. А. Бессмертный, А. Б. Нугуманова, А. В. Платонов. — 2-е изд. — Москва : Издательство Юрайт, 2026. — 250 с. — (Профессиональное образование). — ISBN 978-5-534-20730-9. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/586775>.

2. Кревецкий, А. В. Основы технологий искусственного интеллекта : учебное пособие / А. В. Кревецкий, Н. И. Роженцова, Ю. А. Ипатов ; под общ. ред. А. В. Кревцового. - Йошкар-Ола : Поволжский государственный технологический университет, 2023. - 272 с. - ISBN 978-5-8158-2358-7. - Текст : электронный. - URL: <https://znanium.ru/catalog/product/2133953>.

### *Дополнительная литература:*

1. Платонов, А. В. Машинное обучение : учебник для вузов / А. В. Платонов. — 2-е изд. — Москва : Издательство Юрайт, 2026. — 78 с. — (Высшее образование). — ISBN 978-5-534-22200-5. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/600893>.

2. Гасанов, Э. Э. Интеллектуальные системы. Теория хранения и поиска информации : учебник для вузов / Э. Э. Гасанов, В. Б. Кудрявцев. — 2-е изд., испр. и доп. — Москва : Издательство Юрайт, 2026. — 271 с. — (Высшее образование). — ISBN 978-5-534-08684-3. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/584574>.

## 6. Материально-техническое обеспечение

Университет располагает материально-технической базой, соответствующей действующим противопожарным правилам и нормам и обеспечивающей проведение всех видов дисциплинарной и междисциплинарной подготовки, практической и научно-исследовательской работ обучающихся, предусмотренных учебным планом.

Помещения, которые представляют собой учебные аудитории для проведения занятий лекционного типа, занятий семинарского (практического) типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, а также помещения для самостоятельной работы и помещения для хранения и профилактического обслуживания учебного оборудования. Помещения укомплектованы специализированной мебелью и техническими средствами обучения, служащими для представления учебной информации большой аудитории.

Изучение дисциплины (модуля) обеспечивается в учебных аудиториях, оснащенных:

- столами и стульями;
- компьютерной техникой;

- механическими калькуляторами;
- специализированным оборудованием, включая демонстрационное оборудование.

Помещения для самостоятельной работы обучающихся, в том числе приспособленные для использования инвалидами и лицами с ограниченными возможностями здоровья, оснащены компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду Университета.

Обучающимся предоставляется доступ (в том числе удаленный) к ресурсам информационно-телекоммуникационной сети «Интернет», электронным ресурсам (в том числе электронным библиотечным системам, современным профессиональным базам данных и информационным справочным системам):

| №  | Наименование портала<br>(издания, курса, документа)                            | Ссылка                                                                          |
|----|--------------------------------------------------------------------------------|---------------------------------------------------------------------------------|
| 1. | Научная электронная библиотека elibrary.ru библиотека                          | <a href="https://elibrary.ru/defaultx.asp">https://elibrary.ru/defaultx.asp</a> |
| 2. | База данных для IT-специалистов                                                | <a href="https://habr.com">https://habr.com</a>                                 |
| 3. | База данных ScienceDirect                                                      | <a href="https://www.sciencedirect.com">https://www.sciencedirect.com</a>       |
| 4. | Официальный сайт Министерства науки и высшего образования Российской Федерации | <a href="https://minobrnauki.gov.ru/">https://minobrnauki.gov.ru/</a>           |
| 5. | Федеральный портал «Российское образование»                                    | <a href="https://www.edu.ru/">https://www.edu.ru/</a>                           |
| 6. | Информационная система "Единое окно доступа к образовательным ресурсам"        | <a href="http://window.edu.ru/">http://window.edu.ru/</a>                       |
| 7. | Единая коллекция цифровых образовательных ресурсов                             | <a href="http://school-collection.edu.ru/">http://school-collection.edu.ru/</a> |
| 8. | Федеральный центр информационно - образовательных ресурсов                     | <a href="http://fcior.edu.ru/">http://fcior.edu.ru/</a>                         |

Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), в том числе комплект лицензионного программного обеспечения, современные профессиональные базы данных и информационные справочные системы:

| Наименование ПО                                                  | Производство  | Лицензионное / свободно распространяемое |
|------------------------------------------------------------------|---------------|------------------------------------------|
| <b>Операционные системы:</b>                                     |               |                                          |
| Microsoft Imagine (Windows Client, Server)                       | зарубежное    | лицензионное                             |
| <b>Браузеры:</b>                                                 |               |                                          |
| Яндекс.Браузер                                                   | отечественное | свободно распространяемое                |
| Google Chrome                                                    | зарубежное    | свободно распространяемое                |
| <b>Офисные приложения:</b>                                       |               |                                          |
| Microsoft Imagine (Visio, OneNote)                               | зарубежное    | лицензионное                             |
| TeXstudio                                                        | зарубежное    | свободно распространяемое                |
| Adobe Acrobat Reader                                             | зарубежное    | свободно распространяемое                |
| <b>Программное обеспечение для планирования и учета времени:</b> |               |                                          |
| Toggle app                                                       | зарубежное    | свободно распространяемое                |
| <b>Системы управления проектами:</b>                             |               |                                          |
| Microsoft Imagine (Project)                                      | зарубежное    | лицензионное                             |
| <b>Системы управления базами данных:</b>                         |               |                                          |
| Microsoft Imagine (SQL Server)                                   | зарубежное    | лицензионное                             |
| <b>Системы резервного копирования (backup):</b>                  |               |                                          |
| Acronis Backup Advanced for HyperV                               | зарубежное    | лицензионное                             |
| <b>Справочно-правовые системы:</b>                               |               |                                          |

|                                                                     |               |                           |
|---------------------------------------------------------------------|---------------|---------------------------|
| КонсультантПлюс: справочно-правовая система                         | отечественное | лицензионное              |
| <b>Средства антивирусной защиты:</b>                                |               |                           |
| Kaspersky Endpoint Security для бизнеса Стандартный Russian Edition | отечественное | лицензионное              |
| <b>Среды разработки:</b>                                            |               |                           |
| Visual Studio Code                                                  | зарубежное    | свободно распространяемое |
| Bash (Unix shell)                                                   | зарубежное    | свободно распространяемое |
| Anaconda                                                            | зарубежное    | свободно распространяемое |
| Robotic Operating System                                            | зарубежное    | свободно распространяемое |
| CopelliaSim                                                         | зарубежное    | свободно распространяемое |
| Google Colaboratory                                                 | зарубежное    | свободно распространяемое |
| <b>Пакеты программных средств и библиотек:</b>                      |               |                           |
| AutoPsy                                                             | зарубежное    | свободно распространяемое |
| Interactive Disassembler (IDA)                                      | зарубежное    | свободно распространяемое |
| <b>Системы управления библиографической информацией:</b>            |               |                           |
| Zotero                                                              | зарубежное    | свободно распространяемое |
| <b>Сервисы и службы:</b>                                            |               |                           |
| Bind                                                                | зарубежное    | свободно распространяемое |
| Docker                                                              | зарубежное    | свободно распространяемое |

## 7. Методические и оценочные материалы

### Методические указания для обучающихся по освоению дисциплины (модуля)

В процессе изучения дисциплины (модуля) «Обработка естественного языка» в рамках текущего контроля успеваемости используются такие виды учебной работы, как лекции, семинары и домашние задания, тесты, а также различные виды самостоятельной работы обучающихся по заданию преподавателя, направленные на развитие навыков профессиональной лексики, закрепление практических профессиональных компетенций, поощрение инициатив.

*Лекция* – систематическое, последовательное, монологическое изложение преподавателем учебного материала, как правило, теоретического характера.

В процессе лекций рекомендуется вести конспект лекций: кратко и схематично фиксировать основные идеи, выводы и обобщения лекции; выделять важные мысли, ключевые слова и термины. Необходимо отметить вопросы или материалы, которые вызывают затруднения, и попытаться найти ответы в рекомендованной литературе. Если разобраться в материале не удастся, следует сформулировать вопрос и задать его преподавателю на консультации или во время семинарского (практического) занятия.

*Семинар* – это форма учебной деятельности, проводимая в учебном заведении под руководством преподавателя, где студенты активно участвуют в обсуждениях, практических заданиях и других формах взаимодействия.

Для успешной подготовки к семинару рекомендуется заранее ознакомиться с темой занятия и основными материалами, чтобы иметь возможность активно участвовать в обсуждении. Также полезно подготовить вопросы и идеи для обсуждения, что поможет глубже понять материал и продемонстрировать заинтересованность.

*Домашнее задание* – набор задач по темам недели.

При работе над домашними заданиями важно внимательно ознакомиться с требованиями и сроками выполнения. Рекомендуется разбивать задания на этапы, чтобы избежать перегрузки и лучше усвоить материал. Использовать различные источники информации, включая учебники и онлайн-ресурсы, для более глубокого понимания темы.

*Тест* – особая форма проверки знаний. Проводится после освоения одной или нескольких тем и свидетельствует о качестве понимания основных понятий изучаемого материала. Тестовые задания составлены к ключевым понятиям, основным разделам, важным терминологическим категориям изучаемой дисциплины (модуля).

Для подготовки к тесту необходимо знать терминологический аппарат дисциплины (модуля), понимать смысл научных категорий и уметь их использовать в профессиональной лексике. Владение понятийным аппаратом, включённым в тестовые задания, позволяет преподавателю быстро проверить уровень понимания студентами важных методологических категорий.

*Соревнование* – организованное мероприятие, в рамках которого участники соперничают друг с другом для достижения определенной цели, демонстрируя свои навыки, знания или способности в заданной области.

*Бонусные баллы* – это оценки, которые студенты могут получить за выполнение дополнительных заданий.

Формат бонусных баллов позволяет студентам улучшить общую оценку по дисциплине (модулю) и стимулирует углубленное изучение материала.

*Самостоятельная работа* – работа студентов, направленная на углубленное изучение отдельных тем и вопросов учебной дисциплины (модуля).

В процессе самостоятельной работы студенты взаимодействуют с рекомендованными материалами при минимальном участии преподавателя. Задачи студента включают работу с конспектами лекций (обработка текста), повторное изучение учебных материалов планов и тезисов ответов, изучение дополнительных тем, выполнение учебно-исследовательских заданий и другое.

### **Система оценивания результатов обучения по дисциплине (модулю)**

**Критерии получения уровня и оценивания сформированности компетенций по дисциплине (модулю) «Обработка естественного языка»**

Оценивание уровня учебных достижений, обучающихся по дисциплине (модулю), осуществляется в виде текущего контроля успеваемости и промежуточной аттестации.

**Промежуточная аттестация** по дисциплине (модулю) осуществляется в форме *экзамена*, при этом проводится оценка компетенций, сформированных по дисциплине.

Для оценивания текущего контроля успеваемости и промежуточной аттестации используется десятибалльная шкала оценивания, которая соотносится с традиционной пятибалльной шкалой следующим образом, если иное не указано в силлабусе дисциплины:

| Десятибалльная оценка | Пятибалльная оценка | Общая характеристика результата обучения по дисциплине (модулю)                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|-----------------------|---------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 10                    | Отлично             | Студент полностью владеет знаниями, изложенными в рабочей программе, и глубоко осмысляет дисциплину. Он самостоятельно и логически последовательно отвечает на все вопросы, акцентируя внимание на наиболее важном. Умеет анализировать, сравнивать, классифицировать, обобщать, конкретизировать и систематизировать изученный материал, выделяя ключевые моменты и устанавливая причинно-следственные связи. Четко формулирует ответы, уверенно интерпретирует результаты анализов и других исследований, а также решает сложные |
| 9                     | Отлично             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 8                     | Отлично             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |

| Десятибалльная оценка | Пятибалльная оценка | Общая характеристика результата обучения по дисциплине (модулю)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|-----------------------|---------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                       |                     | задачи. Студент хорошо знаком с методами исследования, необходимыми для практической деятельности, и умеет связывать теоретические аспекты дисциплины (модуля) с практическими задачами.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 7                     | Хорошо              | Студент обладает знаниями предмета почти в полном объеме рабочей программы и самостоятельно, логически последовательно и всесторонне отвечает на все вопросы, акцентируя внимание на наиболее значимых моментах. Он умеет анализировать, сравнивать, классифицировать, обобщать, конкретизировать и систематизировать изученный материал, выделяя его ключевые аспекты и устанавливая причинно-следственные связи. Формулирует свои ответы, уверенно интерпретирует результаты анализов и других исследований, а также решает сложные ситуационные задачи. Студент хорошо знаком с методами исследования, необходимыми для практической деятельности, и умеет связывать теоретические аспекты предмета с практическими задачами. |
| 6                     | Хорошо              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 5                     | Удовлетворительно   | Студент обладает базовыми знаниями по дисциплине (модулю), но испытывает трудности при самостоятельных ответах и использует неточные формулировки. В ходе ответов он допускает ошибки, касающиеся сути вопросов. Студент способен решать только самые простые задачи и владеет лишь минимальным набором методов исследования.                                                                                                                                                                                                                                                                                                                                                                                                    |
| 4                     | Удовлетворительно   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 3                     | Не сдан             | Студент не овладел обязательным минимумом знаний по предмету и не может ответить на вопросы, даже если преподаватель задает дополнительные наводящие вопросы.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 2                     | Не сдан             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 1                     | Не сдан             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |

Дисциплина (модуль) «Обработка естественного языка» оценивается следующим образом:

| Активность       | Вес | Количество | Описание                                                                                                               |
|------------------|-----|------------|------------------------------------------------------------------------------------------------------------------------|
| Домашние задания | 40% | 5          | Набор задач по темам недели                                                                                            |
| Тест             | 30% | 15         | Форма проверки знаний студентов через ответы на вопросы с выбором из нескольких вариантов ответов                      |
| Экзамен          | 30% | 1          | Письменная или устная работа над заданием, направленным на проверку полученных знаний и навыков по дисциплине (модулю) |

В рамках изучения дисциплины (модуля) возможно получение бонусных баллов.

**Формула расчёта итоговой оценки по дисциплине (модулю) «Обработка естественного языка»:**  $0,4 \times \text{среднее за домашние задания} + 0,3 \times \text{среднее за тесты} + 0,3 \times \text{Экзамен}$ .

**Текущий контроль успеваемости обучающихся по дисциплине (модулю)**

## Примерные домашние задания

### Домашнее задание 1.

#### Задание 1.

1. Объясните, что такое цифровой текст и опишите основные этапы его предобработки: токенизация, нормализация, удаление стоп-слов.
2. Реализуйте на Python пайплайн предобработки текста для набора коротких отзывов (например, из IMDb или Yelp) с использованием библиотек nltk или spaCy.
3. Проведите анализ: сравните количество токенов до и после предобработки, визуализируйте самые частотные слова.

#### Задание 2.

1. Опишите, как текст может быть представлен в виде векторов: методы Bag-of-Words, TF-IDF, эмбединги (Word2Vec, GloVe).
2. Реализуйте вычисление сходства между текстами с помощью TF-IDF и косинусной меры. Протестируйте на парах предложений (похожие/разные).
3. Проведите эксперимент: сравните результаты семантического поиска с использованием TF-IDF и предобученных эмбедингов (например, sentence-transformers). Прокомментируйте различия.

#### Задание 3.

1. Опишите задачу извлечения структурированной информации из текста (например, извлечение имён, дат, организаций).
2. Реализуйте извлечение именованных сущностей (NER) с использованием модели spaCy или transformers на наборе новостных текстов.
3. Оцените точность извлечения вручную на 10 примерах, проанализируйте типичные ошибки.

### Домашнее задание 2.

#### Задание 1.

1. Объясните архитектуру Трансформера: механизм внимания, multi-head attention, позиционные кодировки.
2. Сравните энкодерные (BERT), декодерные (GPT) и encoder-decoder (T5) архитектуры: в каких задачах каждая из них применяется?
3. Визуализируйте веса внимания в предобученной модели (например, BERT) с помощью библиотеки BertViz на примере простого предложения.

#### Задание 2.

1. Изучите подходы к адаптации больших языковых моделей: fine-tuning, prompt tuning, LoRA.
2. Выберите модель (например, distilbert-base-uncased или bloomz) и дообучите её на задаче классификации текстов (например, SST-2) с использованием Hugging Face Transformers.
3. Сравните качество до и после дообучения, опишите используемый пайплайн и гиперпараметры.

#### Задание 3.

1. Реализуйте генерацию текста с помощью предобученной LLM (например, gpt2, Llama-3-8b через Hugging Face).
2. Проведите эксперимент с разными стратегиями генерации: greedy, beam search, sampling (с разными temperature, top\_k, top\_p).

3. Проанализируйте, как параметры влияют на разнообразие и качество генерации.

### Домашнее задание 3.

#### Задание 1.

1. Опишите принцип работы систем семантического поиска: индексация, эмбединги, поиск по векторам.
2. Реализуйте простой поисковый движок на основе эмбедингов (sentence-transformers) для набора документов (например, абстракты научных статей).
3. Оцените качество поиска: проверьте, насколько релевантны топ-3 результата для нескольких запросов.

#### Задание 2.

1. Объясните концепцию дополненной генерации (RAG): как внешние данные используются для улучшения генерации.
2. Реализуйте простую RAG-систему: используйте векторный поиск для извлечения контекста и LLM (например, gpt-3.5-turbo или Llama-3) для генерации ответа.
3. Протестируйте на вопросах, требующих знаний вне модели (например, "Какие меры были приняты на G7 в 2023 году?"). Сравните ответы с и без контекста.

#### Задание 3.

1. Опишите ключевые метрики оценки NLP-моделей: точность, F1, BLEU, ROUGE, perplexity, а также современные подходы (LLM-as-a-Judge).
2. Выберите задачу (например, суммаризация) и оцените два варианта выхода модели по нескольким метрикам.
3. Проведите анализ: сравните автоматические метрики с субъективной оценкой (по шкале 1–5). Сделайте вывод о надёжности метрик.

### Примерные задания для теста

#### Тест 1.

##### Вопрос 1.

Что из перечисленного является основной целью токенизации?

- A. Уменьшение размера текста
- B. Преобразование текста в последовательность числовых векторов
- C. Разбиение текста на слова, подслова или символы
- D. Удаление стоп-слов и пунктуации

Ответ: C.

##### Вопрос 2.

Какой метод представления текста учитывает частоту слова в документе и его редкость в корпусе?

- A. Bag-of-Words
- B. One-Hot Encoding
- C. TF-IDF
- D. Word2Vec

Ответ: C.

##### Вопрос 3.

Что такое эмбединг слова?

- A. Словарная статья из википедии
- B. Числовой вектор, представляющий семантику слова
- C. Индекс слова в словаре
- D. Частота слова в тексте

Ответ: В.

**Вопрос 4.**

Какой из следующих подходов позволяет определить, что слова "автомобиль" и "машина" близки по смыслу?

- A. TF-IDF
- B. Эмбединги (например, Word2Vec или Sentence-BERT)
- C. One-Hot Encoding
- D. Подсчёт символов

Ответ: В.

**Вопрос 5.**

Какая задача заключается в извлечении имен, дат и организаций из текста?

- A. Суммаризация
- B. Классификация текста
- C. Извлечение именованных сущностей (NER)
- D. Машинный перевод

Ответ: С.

**Тест 2.**

**Вопрос 1.**

Что является ключевой идеей механизма внимания в архитектуре Трансформера?

- A. Учёт всех слов в последовательности при вычислении представления каждого слова
- B. Использование рекуррентных связей между токенами
- C. Классификация слов по частям речи
- D. Преобразование текста в изображение

Ответ: А.

**Вопрос 2.**

Какая архитектура используется в моделях типа BERT?

- A. Декодерная (decoder-only)
- B. Энкодерная (encoder-only)
- C. Encoder-decoder
- D. CNN + RNN

Ответ: В.

**Вопрос 3.**

Какая модель является примером генеративной архитектуры с декодерной структурой?

- A. T5
- B. BERT
- C. RoBERTa
- D. GPT-3

Ответ: D.

**Вопрос 4.**

Что такое fine-tuning языковой модели?

- A. Генерация текста с использованием промптов
- B. Обучение модели с нуля на новом корпусе
- C. Дообучение предобученной модели на задаче с размеченными данными
- D. Извлечение эмбедингов без обучения

Ответ: С.

**Вопрос 5.**

Какой метод позволяет адаптировать LLM с меньшим количеством параметров для обновления?

- A. LoRA (Low-Rank Adaptation)
- B. TF-IDF

- C. PCA
- D. Bag-of-Words

Ответ: A.

### **Кейс от ООО «КНС ГРУПП»**

Студентам предлагается разработать локальный ИИ-сервис транскрибации и диаризации речи для EdTech-аналитики (Voice Detection).

#### **1. Описание задания:**

Создание высокопроизводительного локального бэкенд-сервиса на FastAPI для автоматической обработки аудиозаписей образовательных лекций и уроков. Пайплайн системы включает пакетную почанковую (chunk-by-chunk) обработку аудио, распознавание речи (ASR) с помощью библиотеки faster-whisper и выделение спикеров (Voice Diarization) с сохранением сквозного контекста голосов на протяжении всей записи.

*Проблема:* EdTech-платформы и методисты обрабатывают терабайты аудиозаписей занятий. Готовые облачные решения накладывают ограничения на конфиденциальность данных и неэффективны по стоимости. При этом стандартные open-source решения часто требуют загрузки всего аудио целиком в оперативную память, что приводит к падению сервисов на длинных лекциях. Почанковая обработка решает проблему утилизации ресурсов, но создаёт сложный инженерный вызов: как сопоставить и не потерять ID уникальных спикеров (учителя и разных студентов) между изолированными чанками аудио.

*Почему важно для бизнеса:* Создание легко масштабируемого, автономного (on-premise) сервиса аналитики аудио позволяет EdTech-компаниям бесшовно размечать структуру уроков, оценивать вовлечённость студентов и контролировать качество преподавания без передачи данных на сторонние API.

#### **2. Этапы работы:**

- Разработать архитектуру FastAPI сервиса, принимающего аудиофайл и выполняющего его пакетную нарезку на чанки.
- Интегрировать модуль транскрибации на базе faster-whisper для почанкового перевода речи в текст.
- Реализовать модуль Voice Detection (Diarization) для распознавания спикеров внутри чанков.
- Разработать механизм сквозного трекинга спикеров (сопоставление эмбедингов голосов) между чанками, чтобы уникальные ID студентов сохранялись едиными для всей лекции.
- Создать демо-интерфейс для визуализации таймлайна урока с разметкой ролей.

#### **3. Используемые ресурсы:**

- реальные аудиозаписи лекций и практических занятий с качественной звуковой дорожкой;
- базовый скрипт-пример интеграции модуля транскрибации faster-whisper.

#### **4. Итоговый результат:**

Полностью локальный работающий бэкенд-сервис на FastAPI, который принимает на вход аудиозапись занятия, производит пакетную обработку по чанкам и возвращает структурированный таймлайн с разметкой спикеров (кто и когда говорил) и соответствующим текстом транскрипции. К сервису прилагается базовый веб-интерфейс (демо) для визуального отображения цветной ленты таймлайна.

#### **5. По итогам выполнения проекта студенты должны представить:**

Демо работающей системы — живой показ прототипа на тестовых аудиозаписях с демонстрацией кода, разбором пайплайна, бэкенд-архитектуры и результатов сквозного трекинга спикеров.

*Удовлетворительно:* Сервис на FastAPI стабильно принимает аудио, делит его на чанки и транскрибирует через faster-whisper без падения по памяти. Реализовано бинарное

разделение речи на Преподавателя (Teacher) и Студентов (Students). Система точно отделяет голос лектора от общей массы реплик учащихся (все студенты могут иметь единый маркер ролей). Демо-интерфейс визуализирует этот таймлайн.

*Хорошо:* Сервис на FastAPI стабильно принимает аудио, делит его на чанки и транскрибирует через faster-whisper. Реализовано разделение речи на Преподавателя и индивидуальных Студентов (более 2 уникальных спикеров). Внедрена кластеризация голосовых профилей для различения учащихся, однако сквозное сопоставление идентичности спикера на протяжении всей записи реализовано частично (ID спикера может меняться между удаленными сегментами). Демо-интерфейс визуализирует таймлайн с несколькими участниками.

*Отлично (Продвинутый Voice Detection):* Реализован продвинутый пайплайн: система способна самостоятельно в динамическом режиме определять точное количество уникальных спикеров на записи (N студентов + 1 учитель). Внедрён алгоритм сквозного сопоставления векторов/профилей голосов (эмбеддингов) между изолированными чанками, благодаря чему ID конкретного студента не «бьётся» и остаётся единым на протяжении всего хронометража лекции. Сервис выдаёт чистый и точный итоговый лог-таймлайн.

### Задания для промежуточной аттестации по дисциплине (модулю)

| № п/п | Задание                                                                         | Ответ                                                      | Компетенция | Индикатор      |
|-------|---------------------------------------------------------------------------------|------------------------------------------------------------|-------------|----------------|
| 1.    | Назовите одно из основных понятий в классических методах обработки изображений. | фильтр / порог / морфология                                | ППК-Р5      | ППК-Р5.1. 3-1  |
| 2.    | Укажите одно из базовых операций с изображениями в CV до эпохи нейронных сетей. | свертка / бинаризация / эрозия                             | ППК-Р5      | ППК-Р5.1. 3-1  |
| 3.    | Назовите одну из архитектур CNN для задачи классификации.                       | ResNet / VGG / AlexNet                                     | ППК-Р5      | ППК-Р5.3. 3-2  |
| 4.    | Укажите одно из понятий в эффективном пайплайне обучения сети.                  | оптимизатор / регуляризация / аугментация                  | МО-2        | МО-2.1. 3-1    |
| 5.    | Назовите одну из трансформерных архитектур в CV.                                | ViT / Deit / Swin                                          | ППК-Р5      | ППК-Р5.3. 3-2  |
| 6.    | Укажите одно из понятий в задаче keypoint detection.                            | ключевые точки / позы / landmarks                          | ППК-Р5      | ППК-Р5.1. 3-1  |
| 7.    | Назовите одну из методов сегментации изображений.                               | семантическая сегментация / экземплярная сегментация / SAM | ППК-Р5      | ППК-Р5.3. 3-2  |
| 8.    | Укажите одно из подходов в anchor-based детекции объектов.                      | Faster R-CNN / SSD / YOLO                                  | ППК-Р5      | ППК-Р5.1. 3-1  |
| 9.    | Назовите одну из методов в anchor-free детекции.                                | CenterNet / CornerNet / FCOS                               | ППК-Р5      | ППК-Р5.1. 3-1  |
| 10.   | Укажите одно из понятий в Face Recognition.                                     | верификация / идентификация / Re-ID                        | ППК-ДА4     | ППК-ДА4.1. У-1 |

|     |                                                                                     |                                                  |         |                |
|-----|-------------------------------------------------------------------------------------|--------------------------------------------------|---------|----------------|
| 11. | Назовите одну из задач в Multi-Object Tracking.                                     | трекинг / MOT / Kalman filter                    | МО-2    | МО-2.3. 3-1    |
| 12. | Укажите одно из этапов в пайплайне OCR.                                             | детекция текста / распознавание / постобработка  | ППК-P5  | ППК-P5.1. У-1  |
| 13. | Назовите одно из методов Self-Supervised Learning.                                  | contrastive learning / masked modeling / BYOL    | ППК-P7  | ППК-P7.1. 3-1  |
| 14. | Укажите одно из понятий в Vision Language Models.                                   | CLIP / VLMs / multimodal                         | ППК-P7  | ППК-P7.1. 3-1  |
| 15. | Назовите одно из современных подходов в генерации изображений.                      | GAN / diffusion models / Stable Diffusion        | ППК-P7  | ППК-P7.1. 3-1  |
| 16. | Укажите одно из основных понятий в классических методах обработки изображений.      | гистограмма / фильтр Гаусса / edge detection     | ППК-P5  | ППК-P5.1. 3-1  |
| 17. | Назовите одну из предпосылок для применения трансформеров в CV.                     | self-attention / патчи / позиционное кодирование | ППК-P5  | ППК-P5.3. 3-2  |
| 18. | Укажите одно из фундаментальных моделей в CV.                                       | CLIP / SAM / Grounding DINO                      | ППК-P5  | ППК-P5.3. 3-2  |
| 19. | Назовите одно из понятий в задаче pose estimation.                                  | ключевые точки / скелет / 3D pose                | ППК-P5  | ППК-P5.1. 3-1  |
| 20. | Укажите одно из методов для эффективного извлечения признаков в Retrieval.          | triplet loss / metric learning / embeddings      | ППК-P4  | ППК-P4.1. У-1  |
| 21. | Какой язык запросов используется для управления данными в реляционных базах данных? | SQL                                              | ППК-P4  | ППК-P4.1. 3-1  |
| 22. | Какая система управления версиями моделей рекомендуется в MLOps?                    | DVC / MLflow                                     | МО-2    | МО-2.2. 3-1    |
| 23. | Какой принцип Agile используется для планирования спринтов в разработке CV?         | Scrum / Kanban                                   | ППК-P6  | ППК-P6.1. 3-1  |
| 24. | Что означает аббревиатура CI/CD в контексте машинного обучения?                     | Continuous Integration / Continuous Delivery     | МО-2    | МО-2.1. 3-1    |
| 25. | Какой инструмент используется для контейнеризации приложений компьютерного зрения?  | Docker                                           | ППК-P6  | ППК-P6.2. У-2  |
| 26. | Назовите принцип чистого кода, требующий избегания дублирования.                    | DRY (Don't Repeat Yourself)                      | ППК-P6  | ППК-P6.3. 3-1  |
| 27. | Какой метод используется для проверки гипотез в эксперименте с моделями?            | A/B тестирование / t-тест                        | ППК-ДА5 | ППК-ДА5.2. У-1 |
| 28. | Какая метрика используется для оценки качества детекции объектов?                   | mAP / IoU                                        | ППК-ДА5 | ППК-ДА5.1. У-2 |

|     |                                                                                        |                                                        |               |                   |
|-----|----------------------------------------------------------------------------------------|--------------------------------------------------------|---------------|-------------------|
| 29. | Какой инструмент ИИ помогает генерировать код для обработки изображений?               | GitHub Copilot / ChatGPT                               | ППК-Р7        | ППК-Р7.1. У-1     |
| 30. | Что необходимо учитывать при использовании ИИ для написания кода с точки зрения этики? | конфиденциальность / авторство / плагиат               | ППК-Р7        | ППК-Р7.4. 3-1     |
| 31. | Какой документ описывает требования к программному обеспечению для CV?                 | Техническое задание / Spec                             | ППК-Р5        | ППК-Р5.1. У-2     |
| 32. | Какой метод используется для декомпозиции ИТ-системы на подсистемы?                    | Функциональная декомпозиция                            | ППК-Р5        | ППК-Р5.4. У-1     |
| 33. | Как называется процесс улучшения кода без изменения его функциональности?              | Рефакторинг                                            | ППК-Р6        | ППК-Р6.3. У-3     |
| 34. | Какая платформа <u>常用于</u> для развертывания моделей в облаке?                         | AWS / Azure / GCP                                      | ППК-Р6        | ППК-Р6.4. У-1     |
| 35. | Какой метод используется для мониторинга дрейфа данных в продакшене?                   | Data Drift Detection                                   | МО-2          | МО-2.3. 3-1       |
| 36. | Как называется процесс автоматического переобучения модели?                            | AutoML / Retraining Pipeline                           | МО-2          | МО-2.4. У-1       |
| 37. | Какой навык необходим для формулирования бизнес-требований для CV проекта?             | Коммуникация / Сбор требований                         | УК-4 (УНК-01) | УНК-01.01.01. У-3 |
| 38. | Какое действие проявляет лидерство в команде разработчиков?                            | Наставничество / Распределение задач                   | УК-3 (УНК-02) | УНК-02.02.01. У-3 |
| 39. | Какой инструмент используется для управления задачами в команде?                       | Jira / Trello                                          | РГ-4          | РГ-4.1. У-3       |
| 40. | Что такое технический долг в разработке ПО?                                            | Накопленные проблемы кода / Необходимость рефакторинга | РГ-4          | РГ-4.3. У-3       |
| 41. | Какой тип базы данных лучше подходит для хранения неструктурированных изображений?     | NoSQL / Object Storage                                 | ППК-Р4        | ППК-Р4.1. 3-1     |
| 42. | Какой метод используется для оптимизации запросов к базе данных с метаданными?         | Индексирование                                         | ППК-Р4        | ППК-Р4.3. У-1     |
| 43. | Какой принцип SOLID гарантирует, что класс должен иметь одну ответственность?          | Single Responsibility Principle (SRP)                  | ППК-Р6        | ППК-Р6.3. 3-1     |
| 44. | Какой метод используется для выбора метрик эффективности модели?                       | Анализ бизнес-целей / ROC-AUC                          | ППК-ДА5       | ППК-ДА5.1. У-2    |
| 45. | Как называется процесс сбора информации о бизнес-проблеме заказчика?                   | Интервью / Анализ требований                           | ППК-ДА4       | ППК-ДА4.1. У-1    |

|     |                                                                            |                                  |               |                   |
|-----|----------------------------------------------------------------------------|----------------------------------|---------------|-------------------|
| 46. | Какой инструмент используется для поиска уязвимостей в коде с помощью ИИ?  | SAST / AI Security Tools         | ППК-Р7        | ППК-Р7.2. У-1     |
| 47. | Какой метод используется для оптимизации кода модели с помощью ИИ?         | AI Code Optimization / Profiling | ППК-Р7        | ППК-Р7.3. 3-1     |
| 48. | Как называется план процесса разработки программного продукта?             | Roadmap / План проекта           | РГ-2          | РГ-2.2. У-1       |
| 49. | Какой навык необходим для работы в команде над CV проектом?                | Коллаборация / Поддержка коллег  | УК-3 (УНК-02) | УНК-02.01.01. У-5 |
| 50. | Какой метод используется для адаптации к изменениям в требованиях проекта? | Гибкая методология / Agile       | УК-4 (УНК-01) | УНК-01.01.01. У-4 |