

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ
ФЕДЕРАЦИИ
федеральное государственное автономное образовательное учреждение высшего
образования
"САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
АЭРОКОСМИЧЕСКОГО ПРИБОРОСТРОЕНИЯ"

Кафедра № 63

УТВЕРЖДАЮ

Ответственный за образовательную
программу

к.ф.н., доц.

(должность, уч. степень, звание)

М.А. Чиханова

(инициалы, фамилия)



(подпись)

«27» июня 2024 г

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

«Информационный поиск и извлечение информации»
(Наименование дисциплины)


Код направления подготовки/ специальности	45.03.02
Наименование направления подготовки/ специальности	Лингвистика
Наименование направленности	Теоретическая и прикладная лингвистика
Форма обучения	заочная
Год приема	2024

Санкт-Петербург– 2024

Лист согласования рабочей программы дисциплины

Программу составил (а)

доц. каф. 63, к.т.н., доц.
(должность, уч. степень, звание)

 27.05.2024
(подпись, дата)

Троицкий Д.И.
(инициалы, фамилия)

Программа одобрена на заседании кафедры № 63

«27» мая 2024 г, протокол № 10

Заведующий кафедрой № 63


к.ф.н., доц.
(уч. степень, звание)

 27.05.2024
(подпись, дата)

М.А. Чиханова
(инициалы, фамилия)

Заместитель декана факультета №6 по методической работе

проф., д.и.н., доц.
(должность, уч. степень, звание)

 27.05.2024
(подпись, дата)

Л.Ю. Гусман
(инициалы, фамилия)

Аннотация

Дисциплина «Информационный поиск и извлечение информации» входит в образовательную программу высшего образования – программу бакалавриата по направлению подготовки/ специальности 45.03.02 «Лингвистика» направленности «Теоретическая и прикладная лингвистика». Дисциплина реализуется кафедрой «№63».

Дисциплина нацелена на формирование у выпускника следующих компетенций:

УК-1 «Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач»

УК-2 «Способен определять круг задач в рамках поставленной цели и выбирать оптимальные способы их решения, исходя из действующих правовых норм, имеющихся ресурсов и ограничений»

ПК-1 «Способность работать с основными информационно-поисковыми и экспертными системами, системами представления знаний, синтаксического и морфологического анализа, автоматического синтеза и распознавания речи, обработки лексикографической информации и автоматизированного перевода, автоматизированными системами идентификации и верификации личности»

ПК-5 «Способность решать стандартные задачи профессиональной деятельности на основе информационной и библиографической культуры с применением информационно-лингвистических технологий и с учетом основных требований информационной безопасности»

ПК-8 «Способность решать основные типы задач в области лингвистического обеспечения информационных и других прикладных систем»

Содержание дисциплины охватывает круг вопросов, связанных с основными концепциями и методами информационного поиска, включая алгоритмы ранжирования, модели представления информации и технологии обработки естественного языка. Особое внимание уделяется методам поиска по изображениям и оценке качества поиска. Курс включает практические задания и проекты для разработки и реализации эффективных поисковых стратегий и адаптации систем поиска к различным областям применения.

Преподавание дисциплины предусматривает следующие формы организации учебного процесса: лекции, лабораторные работы, самостоятельная работа обучающегося.

Программой дисциплины предусмотрены следующие виды контроля: текущий контроль успеваемости, промежуточная аттестация в форме зачета.

Общая трудоемкость освоения дисциплины составляет 2 зачетных единицы, 72 часа.

Язык обучения по дисциплине «русский»

1. Перечень планируемых результатов обучения по дисциплине

1.1. Цели преподавания дисциплины

Цель преподавания дисциплины «Информационный поиск и извлечение информации» — предоставить студентам глубокие знания о методах и технологиях информационного поиска и извлечения данных. Курс охватывает изучение основных алгоритмов и моделей информационного поиска, технологий обработки естественного языка, поиска по изображениям, а также методов оценки качества информационного поиска. Студенты также научатся разрабатывать и реализовывать поисковые стратегии, адаптировать системы поиска к различным областям применения и использовать современные технологии для улучшения релевантности и эффективности поисковых запросов.

1.2. Дисциплина входит в состав части, формируемой участниками образовательных отношений, образовательной программы высшего образования (далее – ОП ВО).

1.3. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения ОП ВО.

В результате изучения дисциплины обучающийся должен обладать следующими компетенциями или их частями. Компетенции и индикаторы их достижения приведены в таблице 1.

Таблица 1 – Перечень компетенций и индикаторов их достижения

Категория (группа) компетенции	Код и наименование компетенции	Код и наименование индикатора достижения компетенции
Универсальные компетенции	УК-1 Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	УК-1.3.1 знать методики поиска, сбора и обработки информации, в том числе с использованием информационных технологий, включая интеллектуальные УК-1.У.3 уметь оценивать информацию на достоверность; сохранять и передавать данные с использованием цифровых средств УК-1.В.1 владеть навыками критического анализа и синтеза информации, в том числе с помощью цифровых инструментов
Универсальные компетенции	УК-2 Способен определять круг задач в рамках поставленной цели и выбирать оптимальные способы их решения, исходя из действующих правовых норм, имеющихся ресурсов и ограничений	УК-2.В.1 владеть навыками выбора оптимального способа решения задач с учетом действующих правовых норм
Профессиональные компетенции	ПК-1 Способность работать с основными информационно-поисковыми и экспертными системами, системами представления знаний, синтаксического и морфологического	ПК-1.3.1 знать основные информационно-поисковые и экспертные системы ПК-1.У.1 уметь работать с формальными системами обработки естественного языка ПК-1.В.1 владеть навыками работы с системами представления знаний, синтаксического и морфологического анализа, автоматического синтеза и распознавания речи, автоматизированными системами идентификации и верификации

	анализа, автоматического синтеза и распознавания речи, обработки лексикографической информации и автоматизированного перевода, автоматизированными системами идентификации и верификации личности	личности
Профессиональные компетенции	ПК-5 Способность решать стандартные задачи профессиональной деятельности на основе информационной и библиографической культуры с применением информационно-лингвистических технологий и с учетом основных требований информационной безопасности	ПК-5.У.1 уметь практически реализовывать основные алгоритмы информационного поиска
Профессиональные компетенции	ПК-8 Способность решать основные типы задач в области лингвистического обеспечения информационных и других прикладных систем	ПК-8.3.1 знать терминологию, алгоритмы и схемы, приемы и законы создания и чтения чертежей, и документации по аппаратным и программным компонентам информационных систем ПК-8.У.1 уметь определять назначение технического описания информационных систем, нормативной документации и методических указаний к выполнению работы ПК-8.В.1 владеть навыками демонстрации результатов анализа и расчётов, выполненных согласно изученному техническому описанию, ПО, нормативной документации; владеть навыками проведения научных исследований, связанных с автоматизированной обработкой текста

2. Место дисциплины в структуре ОП

Дисциплина может базироваться на знаниях, ранее приобретенных обучающимися при изучении следующих дисциплин:

- основы проектной деятельности;
- информатика.

Знания, полученные при изучении материала данной дисциплины, имеют как самостоятельное значение, так и используются при изучении других дисциплин:

- информационные технологии в лингвистике;
- особенности перевода в аэрокосмическом приборостроении;
- выпускная квалификационная работа.

3. Объем и трудоемкость дисциплины

Данные об общем объеме дисциплины, трудоемкости отдельных видов учебной работы по дисциплине (и распределение этой трудоемкости по семестрам) представлены в таблице 2.

Таблица 2 – Объем и трудоемкость дисциплины

Вид учебной работы	Всего	Трудоемкость по семестрам
		№5
1	2	3
Общая трудоемкость дисциплины, ЗЕ/ (час)	2/ 72	2/ 72
Из них часов практической подготовки	8	8
Аудиторные занятия, всего час.	24	24
в том числе:		
лекции (Л), (час)	10	10
практические/семинарские занятия (ПЗ), (час)		
лабораторные работы (ЛР), (час)	14	14
курсовой проект (работа) (КП, КР), (час)		
экзамен, (час)		
Самостоятельная работа, всего (час)	48	48
Вид промежуточной аттестации: зачет, дифф. зачет, экзамен (Зачет, Дифф. зач, Экз.**)	Зачет	Зачет

4. Содержание дисциплины

4.1. Распределение трудоемкости дисциплины по разделам и видам занятий.

Разделы, темы дисциплины и их трудоемкость приведены в таблице 3.

Таблица 3 – Разделы, темы дисциплины, их трудоемкость

Разделы, темы дисциплины	Лекции (час)	ПЗ (СЗ) (час)	ЛР (час)	КП (час)	СРС (час)
Семестр 5					
Раздел 1. Введение в информационный поиск Тема 1.1. Основные концепции информационного поиска 1.1.1. Определение информационного поиска 1.1.2. История развития и основные этапы	2		3		9

Раздел 2. Алгоритмы информационного поиска Тема 2.1. Основные методы ранжирования 2.1.1. TF-IDF и его применение 2.1.2. BM25 и другие алгоритмы ранжирования Тема 2.2. Алгоритмы машинного обучения в информационном поиске 2.2.1. Обучение на основе обратной связи 2.2.2. Методы ранжирования с использованием ML	2		3		10
Раздел 3. Модели информационного поиска Тема 3.1. Поисковые системы и их компоненты 3.1.1. Компоненты поисковой системы 3.1.2. Архитектура и принципы работы поисковых систем Тема 3.2. Модели представления информации 3.2.1. Модель мешка слов и её ограничения 3.2.2. Векторные модели и методы их построения	2		3		9
Раздел 4. Технологии обработки естественного языка в информационном поиске Тема 4.1. Основы анализа текста 4.1.1. Токенизация и лемматизация 4.1.2. Извлечение ключевых слов и фраз Тема 4.2. Использование NLP для улучшения поисковых запросов 4.2.1. Разрешение омонимии и синонимии 4.2.2. Поиск с использованием семантических анализаторов	2		3		10
Раздел 5. Поиск по изображениям Тема 5.1. Основные методы и алгоритмы поиска по изображениям 5.1.1. Content-based поиск 5.1.2. Поиск похожих изображений Тема 5.2. Применение машинного обучения в поиске по изображениям 5.2.1. Обучение моделей на больших наборах данных изображений 5.2.2. Алгоритмы распознавания и классификации изображений	2		2		10
Итого в семестре:	10		14		48
Итого	10	0	14	0	48

Практическая подготовка заключается в непосредственном выполнении обучающимися определенных трудовых функций, связанных с будущей профессиональной деятельностью.

4.2. Содержание разделов и тем лекционных занятий.

Содержание разделов и тем лекционных занятий приведено в таблице 4.

Таблица 4 – Содержание разделов и тем лекционного цикла

Номер раздела	Название и содержание разделов и тем лекционных занятий
1.	Тема 1.1. Основные концепции информационного поиска 1.1.1. Определение информационного поиска 1.1.2. История развития и основные этапы
2.	Тема 2.1. Основные методы ранжирования 2.1.1. TF-IDF и его применение 2.1.2. BM25 и другие алгоритмы ранжирования Тема 2.2. Алгоритмы машинного обучения в информационном поиске 2.2.1. Обучение на основе обратной связи 2.2.2. Методы ранжирования с использованием ML

3.	Тема 3.1. Поисковые системы и их компоненты 3.1.1. Компоненты поисковой системы 3.1.2. Архитектура и принципы работы поисковых систем Тема 3.2. Модели представления информации 3.2.1. Модель мешка слов и её ограничения 3.2.2. Векторные модели и методы их построения
4.	Тема 4.1. Основы анализа текста 4.1.1. Токенизация и лемматизация 4.1.2. Извлечение ключевых слов и фраз Тема 4.2. Использование NLP для улучшения поисковых запросов 4.2.1. Разрешение омонимии и синонимии 4.2.2. Поиск с использованием семантических анализаторов
5.	Тема 5.1. Основные методы и алгоритмы поиска по изображениям 5.1.1. Content-based поиск 5.1.2. Поиск похожих изображений Тема 5.2. Применение машинного обучения в поиске по изображениям 5.2.1. Обучение моделей на больших наборах данных изображений 5.2.2. Алгоритмы распознавания и классификации изображений

4.3. Практические (семинарские) занятия

Темы практических занятий и их трудоемкость приведены в таблице 5.

Таблица 5 – Практические занятия и их трудоемкость

№ п/п	Темы практических занятий	Формы практических занятий	Трудоемкость, (час)	Из них практической подготовки, (час)	№ раздела дисциплины
Учебным планом не предусмотрено					
Всего					

4.4. Лабораторные занятия

Темы лабораторных занятий и их трудоемкость приведены в таблице 6.

Таблица 6 – Лабораторные занятия и их трудоемкость

№ п/п	Наименование лабораторных работ	Трудоемкость, (час)	Из них практической подготовки, (час)	№ раздела дисциплины
Семестр 5				
1.	Исследование основных концепций информационного поиска: анализ и сравнение подходов.	3	3	1
2.	Реализация и оценка алгоритмов ранжирования TF-IDF и BM25.	3	3	2
3.	Построение и тестирование векторной модели представления текста.	3	3	3
4.	Применение методов NLP для улучшения качества поиска: токенизация, лемматизация и извлечение ключевых слов.	3	3	4
5.	Реализация и тестирование алгоритмов поиска по изображениям, включая content-based поиск и поиск похожих изображений.	2	2	5
Всего		14	14	

4.5. Курсовое проектирование/ выполнение курсовой работы

Учебным планом не предусмотрено

4.6. Самостоятельная работа обучающихся

Виды самостоятельной работы и ее трудоемкость приведены в таблице 7.

Таблица 7 – Виды самостоятельной работы и ее трудоемкость

Вид самостоятельной работы	Всего, час	Семестр 5, час
1	2	3
Изучение теоретического материала дисциплины (ТО)	10	10
Курсовое проектирование (КП, КР)		
Расчетно-графические задания (РГЗ)		
Выполнение реферата (Р)		
Подготовка к текущему контролю успеваемости (ТКУ)	10	10
Домашнее задание (ДЗ)	4	4
Контрольные работы заочников (КРЗ)	4	4
Подготовка к промежуточной аттестации (ПА)	10	10
Всего:	48	48

5. Перечень учебно-методического обеспечения

для самостоятельной работы обучающихся по дисциплине (модулю)

Учебно-методические материалы для самостоятельной работы обучающихся указаны в п.п. 7-11.

6. Перечень печатных и электронных учебных изданий

Перечень печатных и электронных учебных изданий приведен в таблице 8.

Таблица 8– Перечень печатных и электронных учебных изданий

Шифр/ URL адрес	Библиографическая ссылка	Количество экземпляров в библиотеке (кроме электронных экземпляров)
https://elar.urfu.ru/bitstream/10995/58206/1/978-5-321-01618-3_2009.pdf?ysclid=lz19uq9h2m966178905	Васина Е.Ю. В 19 Профессиональный поиск научно-технической информации. Индекс научного цитирования: учеб. пособие / Е.Ю. Васина ; Урал. гос. техн. ун-т – УПИ, Ин-т образоват. информ. технологий, Секция информатизации библиотечных дел. – Екатеринбург:	

7. Перечень электронных образовательных ресурсов информационно-телекоммуникационной сети «Интернет»

Перечень электронных образовательных ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины приведен в таблице 9.

Таблица 9 – Перечень электронных образовательных ресурсов информационно-телекоммуникационной сети «Интернет»

URL адрес	Наименование
www.google.com .	Поисковая система Google

8. Перечень информационных технологий

8.1. Перечень программного обеспечения, используемого при осуществлении образовательного процесса по дисциплине.

Перечень используемого программного обеспечения представлен в таблице 10.

Таблица 10– Перечень программного обеспечения

№ п/п	Наименование
	Не предусмотрено

8.2. Перечень информационно-справочных систем,используемых при осуществлении образовательного процесса по дисциплине

Перечень используемых информационно-справочных систем представлен в таблице 11.

Таблица 11– Перечень информационно-справочных систем

№ п/п	Наименование
	Не предусмотрено

9. Материально-техническая база

Состав материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине, представлен в таблице 12.

Таблица 12 – Состав материально-технической базы

№ п/п	Наименование составной части материально-технической базы	Номер аудитории (при необходимости)
1	Компьютерный класс	34-09
2	Мультимедийная лекционная аудитория	34-10

10. Оценочные средства для проведения промежуточной аттестации

10.1. Состав оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине приведен в таблице 13.

Таблица 13 – Состав оценочных средств для проведения промежуточной аттестации

Вид промежуточной аттестации	Перечень оценочных средств
Зачет	Список вопросов; Тесты; Задачи.

10.2. В качестве критериев оценки уровня сформированности (освоения) компетенций обучающимися применяется 5-балльная шкала оценки сформированности компетенций, которая приведена в таблице 14. В течение семестра может использоваться 100-балльная шкала модульно-рейтинговой системы Университета, правила использования которой, установлены соответствующим локальным нормативным актом ГУАП.

Таблица 14 –Критерии оценки уровня сформированности компетенций

Оценка компетенции	Характеристика сформированных компетенций
5-балльная шкала	
«отлично» «зачтено»	<ul style="list-style-type: none"> – обучающийся глубоко и всесторонне усвоил программный материал; – уверенно, логично, последовательно и грамотно его излагает; – опираясь на знания основной и дополнительной литературы, тесно привязывает усвоенные научные положения с практической деятельностью направления; – умело обосновывает и аргументирует выдвигаемые им идеи; – делает выводы и обобщения; – свободно владеет системой специализированных понятий.
«хорошо» «зачтено»	<ul style="list-style-type: none"> – обучающийся твердо усвоил программный материал, грамотно и по существу излагает его, опираясь на знания основной литературы; – не допускает существенных неточностей; – увязывает усвоенные знания с практической деятельностью направления; – аргументирует научные положения; – делает выводы и обобщения; – владеет системой специализированных понятий.
«удовлетворительно» «зачтено»	<ul style="list-style-type: none"> – обучающийся усвоил только основной программный материал, по существу излагает его, опираясь на знания только основной литературы; – допускает несущественные ошибки и неточности; – испытывает затруднения в практическом применении знаний направления; – слабо аргументирует научные положения; – затрудняется в формулировании выводов и обобщений; – частично владеет системой специализированных понятий.
«неудовлетворительно» «не зачтено»	<ul style="list-style-type: none"> – обучающийся не усвоил значительной части программного материала; – допускает существенные ошибки и неточности при рассмотрении проблем в конкретном направлении; – испытывает трудности в практическом применении знаний; – не может аргументировать научные положения; – не формулирует выводов и обобщений.

10.3. Типовые контрольные задания или иные материалы.

Вопросы (задачи) для экзамена представлены в таблице 15.

Таблица 15 – Вопросы (задачи) для экзамена

№ п/п	Перечень вопросов (задач) для экзамена	Код индикатора
	Учебным планом не предусмотрено	

Вопросы (задачи) для зачета / дифф. зачета представлены в таблице 16.

Таблица 16 – Вопросы (задачи) для зачета / дифф. зачета

№ п/п	Перечень вопросов (задач) для зачета / дифф. зачета	Код индикатора
1	Опишите основные принципы информационного поиска.	УК-1.3.1
2	Как изменялись подходы к информационному поиску с момента его возникновения?	УК-1.3.1
3	Объясните, что такое информационный поиск и каковы его ключевые задачи.	УК-1.3.1
4	Какие основные этапы развития информационного поиска можно выделить?	УК-1.3.1
5	В чем заключается отличие между TF-IDF и BM25?	УК-1.3.1
6	Как метод TF-IDF помогает в ранжировании документов?	УК-1.3.1
7	Опишите, как BM25 улучшает результаты поиска по сравнению с TF-IDF.	УК-1.3.1
8	Какие алгоритмы машинного обучения применяются для ранжирования в информационном поиске?	УК-1.3.1
9	Как обучение на основе обратной связи влияет на качество поиска?	УК-1.3.1
10	Какие преимущества и недостатки имеет использование машинного обучения в информационном поиске?	УК-1.3.1
11	Объясните, что такое модель мешка слов и в чем ее ограничения.	УК-1.3.1
12	Как векторные модели представляют текст и какие методы их построения существуют?	УК-1.3.1
13	Опишите основные компоненты поисковой системы и их функции.	УК-1.3.1
14	Как архитектура поисковой системы влияет на её производительность?	УК-1.3.1
15	Какие существуют методы анализа текста в обработке естественного языка?	УК-1.3.1
16	В чем заключаются основные задачи токенизации и лемматизации?	УК-1.3.1
17	Как извлечение ключевых слов и фраз улучшает поисковые запросы?	УК-1.3.1
18	Какие методы NLP могут использоваться для разрешения омонимии и синонимии?	УК-1.3.1
19	Как семантические анализаторы помогают в улучшении поиска?	УК-1.3.1
20	Что такое content-based поиск и как он работает?	УК-1.3.1
21	Опишите основные подходы к поиску похожих изображений.	УК-1.У.3
22	Какие алгоритмы машинного обучения применяются для распознавания и классификации изображений?	УК-1.У.3
23	Как обучение моделей на больших наборах данных изображений помогает в поиске по изображениям?	УК-1.У.3
24	Какие метрики используются для оценки качества информационного поиска?	УК-1.У.3
25	Как точность и полнота влияют на оценку качества поиска?	УК-1.У.3
26	Объясните, что такое F-мера и как она вычисляется.	УК-1.У.3
27	Что такое ROC-кривая и как она применяется в оценке поиска?	УК-1.У.3
28	Какие методы можно использовать для оптимизации запросов в информационном поиске?	УК-1.У.3
29	Как адаптивные алгоритмы помогают улучшить релевантность результатов поиска?	УК-1.У.3
30	В чем заключаются основные цели и задачи лабораторной работы по теме TF-IDF?	УК-1.У.3

31	Как в лабораторной работе можно реализовать алгоритм BM25?	УК-1.У.3
32	Какие подходы можно использовать для построения векторной модели представления текста?	УК-1.У.3
33	Как токенизация и лемматизация могут быть использованы для улучшения результатов поиска в лабораторной работе?	УК-1.У.3
34	Какие методы content-based поиска наиболее эффективны для практического применения?	УК-1.У.3
35	Как в лабораторной работе можно оценить качество поиска по изображениям?	УК-1.У.3
36	Какие данные необходимы для обучения моделей распознавания изображений в лабораторных условиях?	УК-1.У.3
37	Как можно использовать метрики оценки для улучшения качества поисковых систем?	УК-1.У.3
38	Какие алгоритмы машинного обучения наиболее подходящие для информационного поиска и почему?	УК-1.У.3
39	Как можно применить методы NLP для повышения точности поиска в лабораторной работе?	УК-1.У.3
40	В чем заключается специфика оценки качества поиска по изображениям?	УК-1.У.3
41	Как в лабораторной работе можно исследовать влияние различных алгоритмов ранжирования?	УК-1.В.1
42	Какие методы могут быть применены для анализа и улучшения работы поисковых систем?	УК-1.В.1
43	Как в лабораторных условиях можно тестировать и улучшать методы разрешения омонимии и синонимии?	УК-1.В.1
44	Какие подходы можно использовать для оценки эффективности методов поиска по изображениям?	УК-1.В.1
45	Как в лабораторной работе можно исследовать влияние различных методов токенизации на результаты поиска?	УК-1.В.1
46	Какие особенности имеет разработка и реализация поисковых стратегий для различных областей применения?	УК-1.В.1
47	Как можно использовать результаты лабораторных работ для улучшения реальных поисковых систем?	УК-1.В.1
48	Какие данные и методы анализа необходимы для оценки работы поисковых систем?	УК-1.В.1
49	Как лабораторные работы помогают понять принципы работы алгоритмов ранжирования?	УК-1.В.1
50	Как разработать эффективную поисковую стратегию для задачи на основе реальных данных?	УК-1.В.1
51	Какие критерии важны для оценки качества поисковых запросов в лабораторных условиях?	УК-1.В.1
52	Как улучшение NLP технологий может повлиять на результаты информационного поиска?	УК-1.В.1
53	Какие современные технологии могут быть использованы для улучшения эффективности поисковых систем?	УК-1.В.1
54	Каковы основные этапы и принципы развития технологий информационного поиска?	УК-1.В.1
55	Что такое "информационный запрос" и как он формируется?	УК-1.В.1

56	Как концепция "информационного поиска" изменялась с течением времени?	УК-1.В.1
57	Объясните, как TF-IDF используется для определения релевантности документов.	УК-1.В.1
58	Какие основные принципы лежат в основе алгоритма BM25?	УК-1.В.1
59	Как можно улучшить результаты поиска, используя метод TF-IDF?	УК-1.В.1
60	Опишите, как алгоритмы машинного обучения могут улучшить качество поиска.	УК-1.В.1
61	Какие типы обратной связи используются в обучении моделей для информационного поиска?	УК-2.В.1
62	Как машинное обучение может быть интегрировано в традиционные методы информационного поиска?	УК-2.В.1
63	Какие ограничения существуют у методов ранжирования, основанных на машинном обучении?	УК-2.В.1
64	В чем заключаются ограничения модели мешка слов и как они могут быть преодолены?	УК-2.В.1
65	Как векторные модели помогают в анализе текстов и их поиске?	УК-2.В.1
66	Опишите, как компоненты поисковой системы взаимодействуют для выполнения запроса.	УК-2.В.1
67	Какие основные принципы архитектуры поисковой системы обеспечивают её эффективность?	УК-2.В.1
68	Какие ключевые задачи решаются в процессе анализа текста в NLP?	УК-2.В.1
69	Как токенизация и лемматизация влияют на обработку и анализ текстовых данных?	УК-2.В.1
70	В чем заключается важность извлечения ключевых слов и фраз для поисковых систем?	УК-2.В.1
71	Как разрешение омонимии и синонимии влияет на точность поиска?	УК-2.В.1
72	Какие методы семантического анализа могут улучшить результаты поиска?	УК-2.В.1
73	Опишите, как работает content-based поиск и его применение на практике.	УК-2.В.1
74	Как можно использовать алгоритмы поиска похожих изображений для улучшения систем поиска?	УК-2.В.1
75	Какие современные методы машинного обучения применяются в распознавании и классификации изображений?	УК-2.В.1
76	Как обучение на больших наборах изображений помогает в улучшении поиска по изображениям?	УК-2.В.1
77	Какие метрики наиболее важны для оценки качества информационного поиска и почему?	УК-2.В.1
78	Как можно измерить и интерпретировать точность и полноту поиска?	УК-2.В.1
79	Объясните, как вычисляется F-мера и её значение для оценки качества поиска.	УК-2.В.1
80	Как ROC-кривая помогает в оценке эффективности поисковых систем?	УК-2.В.1
81	Какие методы оптимизации запросов могут быть применены для улучшения поиска?	ПК-1.3.1
82	Как адаптивные алгоритмы могут улучшить релевантность результатов поиска?	ПК-1.3.1

83	Какие ключевые аспекты следует учитывать при реализации алгоритма TF-IDF в лабораторных условиях?	ПК-1.3.1
84	Как можно эффективно протестировать алгоритм BM25 на наборе данных?	ПК-1.3.1
85	Какие подходы следует использовать для построения векторной модели текста?	ПК-1.3.1
86	Как методы NLP, такие как токенизация и лемматизация, применяются в лабораторных работах?	ПК-1.3.1
87	Как можно улучшить методы content-based поиска в лабораторных условиях?	ПК-1.3.1
88	Какие методы можно использовать для оценки качества поиска по изображениям в лабораторных экспериментах?	ПК-1.3.1
89	Какие данные и подходы необходимы для обучения моделей распознавания изображений?	ПК-1.3.1
90	Как можно использовать метрики для анализа и улучшения качества поиска?	ПК-1.3.1
91	Какие алгоритмы машинного обучения наиболее эффективны для конкретных задач информационного поиска?	ПК-1.3.1
92	Как методы NLP могут быть применены для улучшения точности поисковых систем?	ПК-1.3.1
93	Как можно исследовать влияние различных алгоритмов ранжирования на качество поиска в лабораторной работе?	ПК-1.3.1
94	Какие методы могут быть использованы для анализа и улучшения работы поисковых систем?	ПК-1.3.1
95	Как в лабораторных условиях можно тестировать и совершенствовать методы разрешения омонимии и синонимии?	ПК-1.3.1
96	Какие подходы можно использовать для оценки и улучшения поиска по изображениям?	ПК-1.3.1
97	Как в лабораторных условиях можно исследовать влияние методов токенизации на результаты поиска?	ПК-1.3.1
98	Какие особенности разработки поисковых стратегий для различных приложений следует учитывать?	ПК-1.3.1
99	Как лабораторные работы могут помочь в понимании работы алгоритмов ранжирования?	ПК-1.3.1
100	Какие критерии оценки качества поисковых запросов важны в лабораторных экспериментах?	ПК-1.3.1
101	Как улучшение технологий NLP может повлиять на эффективность поисковых систем?	ПК-1.У.1
102	Какие современные технологии можно применить для улучшения систем поиска?	ПК-1.У.1
103	Какие основные принципы ранжирования документов в информационных системах?	ПК-1.У.1
104	Каковы основные этапы процесса информационного поиска от получения запроса до выдачи результатов?	ПК-1.У.1
105	Объясните, как алгоритмы ранжирования учитывают релевантность документа.	ПК-1.У.1
106	Как можно улучшить точность результатов поиска, используя методы анализа текста?	ПК-1.У.1

107	В чем различие между TF-IDF и BM25 в контексте поиска информации?	ПК-1.У.1
108	Как алгоритм BM25 учитывает частоту и длину документа при ранжировании?	ПК-1.У.1
109	Какие техники машинного обучения наиболее эффективно применяются для улучшения поиска?	ПК-1.У.1
110	Как методы машинного обучения помогают в адаптации поисковых систем к пользовательским запросам?	ПК-1.У.1
111	Объясните, как алгоритмы обучения на основе обратной связи могут улучшить результаты поиска.	ПК-1.У.1
112	Какие методы машинного обучения позволяют решать задачи ранжирования в информационном поиске?	ПК-1.У.1
113	Каковы ограничения модели мешка слов при обработке сложных текстов?	ПК-1.У.1
114	Какие преимущества и недостатки векторных моделей по сравнению с моделью мешка слов?	ПК-1.У.1
115	Как можно использовать модель мешка слов в современных поисковых системах?	ПК-1.У.1
116	Как структура архитектуры поисковой системы влияет на её способность обрабатывать запросы?	ПК-1.У.1
117	Какие принципы лежат в основе эффективного построения архитектуры поисковой системы?	ПК-1.У.1
118	Как методы анализа текста помогают в улучшении качества поиска в системах?	ПК-1.У.1
119	Какие современные методы NLP можно применить для обработки естественного языка в поисковых системах?	ПК-1.У.1
120	Как токенизация и лемматизация влияют на обработку текстов в различных языках?	ПК-1.У.1
121	Объясните, как извлечение ключевых слов и фраз помогает в улучшении поисковых запросов.	ПК-1.В.1
122	Как разрешение омонимии и синонимии помогает улучшить точность результатов поиска?	ПК-1.В.1
123	Какие подходы к семантическому анализу могут улучшить поиск информации?	ПК-1.В.1
124	В чем состоит суть content-based поиска и как его можно улучшить?	ПК-1.В.1
125	Как алгоритмы поиска похожих изображений помогают в создании систем визуального поиска?	ПК-1.В.1
126	Какие современные технологии машинного обучения используются для классификации и распознавания изображений?	ПК-1.В.1
127	Как можно применять обучение на больших наборах изображений для улучшения поиска?	ПК-1.В.1
128	Какие метрики наиболее полезны для оценки качества поисковых систем и как их интерпретировать?	ПК-1.В.1
129	Как можно использовать точность и полноту для оценки эффективности поисковых систем?	ПК-1.В.1
130	Объясните, как F-мера может помочь в оценке качества поиска и какие её ограничения?	ПК-1.В.1
131	Как ROC-кривая может быть использована для сравнения различных поисковых систем?	ПК-1.В.1

132	Какие методы можно использовать для улучшения запросов в поисковых системах?	ПК-1.В.1
133	Как адаптивные алгоритмы могут повысить релевантность поиска в реальных системах?	ПК-1.В.1
134	Каковы ключевые аспекты реализации алгоритма TF-IDF в лабораторных условиях?	ПК-1.В.1
135	Какие шаги необходимо предпринять для эффективного тестирования алгоритма BM25?	ПК-1.В.1
136	Как построить векторную модель представления текста для информационного поиска?	ПК-1.В.1
137	Как токенизация и лемматизация могут быть использованы в лабораторной работе для улучшения поиска?	ПК-1.В.1
138	Какие методы могут быть применены для улучшения content-based поиска в лабораторных условиях?	ПК-1.В.1
139	Как можно оценить и улучшить качество поиска по изображениям в лабораторной работе?	ПК-1.В.1
140	Какие данные необходимы для обучения моделей распознавания изображений и как их подготовить?	ПК-1.В.1
141	Как использовать метрики для анализа качества поисковых систем и какие данные для этого нужны?	ПК-5.У.1
142	Какие алгоритмы машинного обучения наиболее эффективны для информационного поиска и почему?	ПК-5.У.1
143	Как методы NLP могут быть применены для улучшения точности результатов поиска?	ПК-5.У.1
144	Какие практические шаги можно предпринять для тестирования алгоритмов ранжирования в лабораторных условиях?	ПК-5.У.1
145	Как улучшить работу поисковых систем с помощью методов анализа и обработки текста?	ПК-5.У.1
146	Какие подходы следует использовать для анализа методов разрешения омонимии и синонимии?	ПК-5.У.1
147	Как можно оценить и улучшить поиск по изображениям на основе лабораторных экспериментов?	ПК-5.У.1
148	Какие особенности следует учитывать при разработке поисковых стратегий для различных областей?	ПК-5.У.1
149	Как лабораторные работы помогают в понимании работы и улучшении алгоритмов ранжирования?	ПК-5.У.1
150	Как современные технологии NLP могут улучшить работу поисковых систем?	ПК-5.У.1
151	Какие современные методы и технологии можно применить для повышения эффективности поиска?	ПК-5.У.1
152	Какое значение имеет историческое развитие информационного поиска для современных технологий?	ПК-5.У.1
153	В чем заключается основное различие между информационным поиском и извлечением данных?	ПК-5.У.1
154	Как можно улучшить понимание запроса пользователя в системах информационного поиска?	ПК-5.У.1
155	Опишите, как алгоритм TF-IDF влияет на поиск и ранжирование документов.	ПК-5.У.1

156	Как методы BM25 и TF-IDF могут быть комбинированы для улучшения поиска?	ПК-5.У.1
157	В чем заключается разница между традиционными и современными подходами к информационному поиску?	ПК-5.У.1
158	Как обучение с обратной связью улучшает результаты информационного поиска?	ПК-5.У.1
159	Какие методы ранжирования на основе машинного обучения наиболее эффективны в различных сценариях поиска?	ПК-5.У.1
160	Как векторные модели помогают в устранении неоднозначностей в текстах?	ПК-5.У.1
161	Какие проблемы могут возникнуть при применении модели мешка слов в информационном поиске?	ПК-8.3.1
162	Как компоненты поисковой системы взаимодействуют для формирования релевантных результатов?	ПК-8.3.1
163	Какие принципы архитектуры поисковой системы необходимо учитывать при её разработке?	ПК-8.3.1
164	Как токенизация и лемматизация влияют на эффективность обработки текстов в NLP?	ПК-8.3.1
165	Как можно использовать извлечение ключевых слов для улучшения поисковых систем?	ПК-8.3.1
166	Какие методы NLP можно применить для обработки омонимии и синонимии в запросах?	ПК-8.3.1
167	Как семантический анализ может улучшить поиск и понимание текстов?	ПК-8.3.1
168	В чем заключаются особенности реализации content-based поиска и его применения?	ПК-8.3.1
169	Как алгоритмы поиска похожих изображений могут быть адаптированы для различных типов данных?	ПК-8.3.1
170	Какие методы машинного обучения наиболее эффективны для поиска и классификации изображений?	ПК-8.3.1
171	Как можно использовать большие наборы данных для обучения моделей поиска по изображениям?	ПК-8.3.1
172	Как метрики, такие как точность и полнота, помогают в оценке качества поиска?	ПК-8.3.1
173	Какие методы оценки качества поиска могут быть использованы для улучшения поиска?	ПК-8.3.1
174	Как можно интерпретировать результаты F-меры в контексте информационного поиска?	ПК-8.3.1
175	Как ROC-кривая может быть использована для анализа и оценки эффективности поисковых систем?	ПК-8.3.1
176	Какие методы оптимизации запросов могут быть применены для повышения качества поиска?	ПК-8.3.1
177	Как адаптивные алгоритмы могут быть использованы для улучшения релевантности поиска?	ПК-8.3.1
178	Как реализовать и протестировать алгоритм TF-IDF в лабораторной работе?	ПК-8.3.1
179	Какие подходы следует применять для оценки и улучшения алгоритма BM25?	ПК-8.3.1

180	Как построить и применить векторные модели для поиска в текстовых данных?	ПК-8.З.1
181	Как токенизация и лемматизация могут быть использованы для повышения точности поиска?	ПК-8.У.1
182	Какие методы можно использовать для улучшения content-based поиска в лабораторных условиях?	ПК-8.У.1
183	Как оценить качество поиска по изображениям в лабораторных экспериментах?	ПК-8.У.1
184	Какие данные необходимы для обучения моделей распознавания изображений и как их использовать?	ПК-8.У.1
185	Как метрики оценки могут быть применены для анализа и улучшения систем поиска?	ПК-8.У.1
186	Как можно использовать современные методы машинного обучения для улучшения поиска?	ПК-8.У.1
187	Какие методы NLP могут быть использованы для повышения точности и эффективности поиска?	ПК-8.У.1
188	Как тестирование различных алгоритмов ранжирования может помочь в улучшении поиска?	ПК-8.У.1
189	Какие методы можно использовать для анализа и улучшения работы поисковых систем?	ПК-8.У.1
190	Как методы разрешения омонимии и синонимии могут быть протестированы и улучшены?	ПК-8.У.1
191	Какие подходы применяются для оценки и улучшения систем поиска по изображениям?	ПК-8.У.1
192	Как токенизация и лемматизация влияют на результаты поиска и их улучшение?	ПК-8.У.1
193	Какие особенности необходимо учитывать при разработке поисковых стратегий для различных областей?	ПК-8.У.1
194	Как лабораторные работы помогают понять принципы и методы информационного поиска?	ПК-8.У.1
195	Какие критерии оценки качества поисковых запросов наиболее важны в лабораторных условиях?	ПК-8.У.1
196	Как улучшение технологий обработки естественного языка может повлиять на качество поиска?	ПК-8.У.1
197	Какое влияние на качество поиска оказывает история развития информационного поиска?	ПК-8.У.1
198	Какие ключевые концепции определяют эффективность информационного поиска?	ПК-8.У.1
199	В чем заключаются основные различия между методами TF-IDF и BM25 в контексте ранжирования документов?	ПК-8.У.1
200	Как методы машинного обучения могут улучшить ранжирование результатов поиска?	ПК-8.У.1
201	Опишите, как алгоритмы машинного обучения адаптируются к обратной связи от пользователей.	ПК-8.В.1
202	Какие методы ранжирования на основе машинного обучения наиболее популярны в информационном поиске?	ПК-8.В.1
203	Как модели представления текста могут быть улучшены для повышения точности поиска?	ПК-8.В.1

204	Какие основные компоненты и принципы работы поисковых систем влияют на их производительность?	ПК-8.В.1
205	Как NLP технологии могут использоваться для оптимизации поисковых запросов?	ПК-8.В.1
206	Опишите, как токенизация и лемматизация помогают в обработке текстовых данных для поиска.	ПК-8.В.1
207	Как методы разрешения омонимии и синонимии улучшают результаты информационного поиска?	ПК-8.В.1
208	Какие подходы существуют для анализа и извлечения ключевых слов из текста?	ПК-8.В.1
209	Как content-based поиск применяется для улучшения результатов поиска по изображениям?	ПК-8.В.1
210	Какие алгоритмы можно использовать для поиска похожих изображений в больших базах данных?	ПК-8.В.1
211	Опишите, как обучение на больших наборах данных изображений может улучшить алгоритмы распознавания.	ПК-8.В.1
212	Какие метрики оценки поиска являются наиболее информативными для анализа качества результатов?	ПК-8.В.1
213	Как можно измерить и интерпретировать точность и полноту результатов поиска?	ПК-8.В.1
214	В чем заключается вычисление и применение F-меры для оценки эффективности поисковых систем?	ПК-8.В.1
215	Как ROC-кривая может использоваться для оценки качества поиска и что она показывает?	ПК-8.В.1
216	Какие методы оптимизации запросов помогут повысить релевантность результатов поиска?	ПК-8.В.1
217	Как адаптивные алгоритмы могут быть использованы для улучшения качества поиска в различных приложениях?	ПК-8.В.1
218	Какие ключевые аспекты необходимо учитывать при реализации алгоритма TF-IDF в реальных условиях?	ПК-8.В.1
219	Как можно протестировать и оценить эффективность алгоритма BM25 на различных наборах данных?	ПК-8.В.1
220	Какие методы можно использовать для построения и оценки векторных моделей представления текста?	ПК-8.В.1

Перечень тем для курсового проектирования/выполнения курсовой работы представлены в таблице 17.

Таблица 17 – Перечень тем для курсового проектирования/выполнения курсовой работы

№ п/п	Примерный перечень тем для курсового проектирования/выполнения курсовой работы
	Учебным планом не предусмотрено

Вопросы для проведения промежуточной аттестации в виде тестирования представлены в таблице 18.

Таблица 18 – Примерный перечень вопросов для тестов

№ п/п	Примерный перечень вопросов для тестов	Код индикатора

1	Опишите основные принципы информационного поиска.	УК-1.3.1
2	Как изменялись подходы к информационному поиску с момента его возникновения?	УК-1.3.1
3	Объясните, что такое информационный поиск и каковы его ключевые задачи.	УК-1.3.1
4	Какие основные этапы развития информационного поиска можно выделить?	УК-1.3.1
5	В чем заключается отличие между TF-IDF и BM25?	УК-1.3.1
6	Как метод TF-IDF помогает в ранжировании документов?	УК-1.3.1
7	Опишите, как BM25 улучшает результаты поиска по сравнению с TF-IDF.	УК-1.3.1
8	Какие алгоритмы машинного обучения применяются для ранжирования в информационном поиске?	УК-1.3.1
9	Как обучение на основе обратной связи влияет на качество поиска?	УК-1.3.1
10	Какие преимущества и недостатки имеет использование машинного обучения в информационном поиске?	УК-1.3.1
11	Объясните, что такое модель мешка слов и в чем ее ограничения.	УК-1.3.1
12	Как векторные модели представляют текст и какие методы их построения существуют?	УК-1.3.1
13	Опишите основные компоненты поисковой системы и их функции.	УК-1.3.1
14	Как архитектура поисковой системы влияет на её производительность?	УК-1.3.1
15	Какие существуют методы анализа текста в обработке естественного языка?	УК-1.3.1
16	В чем заключаются основные задачи токенизации и лемматизации?	УК-1.3.1
17	Как извлечение ключевых слов и фраз улучшает поисковые запросы?	УК-1.3.1
18	Какие методы NLP могут использоваться для разрешения омонимии и синонимии?	УК-1.3.1
19	Как семантические анализаторы помогают в улучшении поиска?	УК-1.3.1
20	Что такое content-based поиск и как он работает?	УК-1.3.1
21	Опишите основные подходы к поиску похожих изображений.	УК-1.У.3
22	Какие алгоритмы машинного обучения применяются для распознавания и классификации изображений?	УК-1.У.3
23	Как обучение моделей на больших наборах данных изображений помогает в поиске по изображениям?	УК-1.У.3
24	Какие метрики используются для оценки качества информационного поиска?	УК-1.У.3
25	Как точность и полнота влияют на оценку качества поиска?	УК-1.У.3
26	Объясните, что такое F-мера и как она вычисляется.	УК-1.У.3
27	Что такое ROC-кривая и как она применяется в оценке поиска?	УК-1.У.3
28	Какие методы можно использовать для оптимизации запросов в информационном поиске?	УК-1.У.3
29	Как адаптивные алгоритмы помогают улучшить релевантность результатов поиска?	УК-1.У.3
30	В чем заключаются основные цели и задачи лабораторной работы по теме TF-IDF?	УК-1.У.3
31	Как в лабораторной работе можно реализовать алгоритм BM25?	УК-1.У.3

32	Какие подходы можно использовать для построения векторной модели представления текста?	УК-1.У.3
33	Как токенизация и лемматизация могут быть использованы для улучшения результатов поиска в лабораторной работе?	УК-1.У.3
34	Какие методы content-based поиска наиболее эффективны для практического применения?	УК-1.У.3
35	Как в лабораторной работе можно оценить качество поиска по изображениям?	УК-1.У.3
36	Какие данные необходимы для обучения моделей распознавания изображений в лабораторных условиях?	УК-1.У.3
37	Как можно использовать метрики оценки для улучшения качества поисковых систем?	УК-1.У.3
38	Какие алгоритмы машинного обучения наиболее подходящие для информационного поиска и почему?	УК-1.У.3
39	Как можно применить методы NLP для повышения точности поиска в лабораторной работе?	УК-1.У.3
40	В чем заключается специфика оценки качества поиска по изображениям?	УК-1.У.3
41	Как в лабораторной работе можно исследовать влияние различных алгоритмов ранжирования?	УК-1.В.1
42	Какие методы могут быть применены для анализа и улучшения работы поисковых систем?	УК-1.В.1
43	Как в лабораторных условиях можно тестировать и улучшать методы разрешения омонимии и синонимии?	УК-1.В.1
44	Какие подходы можно использовать для оценки эффективности методов поиска по изображениям?	УК-1.В.1
45	Как в лабораторной работе можно исследовать влияние различных методов токенизации на результаты поиска?	УК-1.В.1
46	Какие особенности имеет разработка и реализация поисковых стратегий для различных областей применения?	УК-1.В.1
47	Как можно использовать результаты лабораторных работ для улучшения реальных поисковых систем?	УК-1.В.1
48	Какие данные и методы анализа необходимы для оценки работы поисковых систем?	УК-1.В.1
49	Как лабораторные работы помогают понять принципы работы алгоритмов ранжирования?	УК-1.В.1
50	Как разработать эффективную поисковую стратегию для задачи на основе реальных данных?	УК-1.В.1
51	Какие критерии важны для оценки качества поисковых запросов в лабораторных условиях?	УК-1.В.1
52	Как улучшение NLP технологий может повлиять на результаты информационного поиска?	УК-1.В.1
53	Какие современные технологии могут быть использованы для улучшения эффективности поисковых систем?	УК-1.В.1
54	Каковы основные этапы и принципы развития технологий информационного поиска?	УК-1.В.1
55	Что такое "информационный запрос" и как он формируется?	УК-1.В.1
56	Как концепция "информационного поиска" изменялась с течением времени?	УК-1.В.1

57	Объясните, как TF-IDF используется для определения релевантности документов.	УК-1.В.1
58	Какие основные принципы лежат в основе алгоритма BM25?	УК-1.В.1
59	Как можно улучшить результаты поиска, используя метод TF-IDF?	УК-1.В.1
60	Опишите, как алгоритмы машинного обучения могут улучшить качество поиска.	УК-1.В.1
61	Какие типы обратной связи используются в обучении моделей для информационного поиска?	УК-2.В.1
62	Как машинное обучение может быть интегрировано в традиционные методы информационного поиска?	УК-2.В.1
63	Какие ограничения существуют у методов ранжирования, основанных на машинном обучении?	УК-2.В.1
64	В чем заключаются ограничения модели мешка слов и как они могут быть преодолены?	УК-2.В.1
65	Как векторные модели помогают в анализе текстов и их поиске?	УК-2.В.1
66	Опишите, как компоненты поисковой системы взаимодействуют для выполнения запроса.	УК-2.В.1
67	Какие основные принципы архитектуры поисковой системы обеспечивают её эффективность?	УК-2.В.1
68	Какие ключевые задачи решаются в процессе анализа текста в NLP?	УК-2.В.1
69	Как токенизация и лемматизация влияют на обработку и анализ текстовых данных?	УК-2.В.1
70	В чем заключается важность извлечения ключевых слов и фраз для поисковых систем?	УК-2.В.1
71	Как разрешение омонимии и синонимии влияет на точность поиска?	УК-2.В.1
72	Какие методы семантического анализа могут улучшить результаты поиска?	УК-2.В.1
73	Опишите, как работает content-based поиск и его применение на практике.	УК-2.В.1
74	Как можно использовать алгоритмы поиска похожих изображений для улучшения систем поиска?	УК-2.В.1
75	Какие современные методы машинного обучения применяются в распознавании и классификации изображений?	УК-2.В.1
76	Как обучение на больших наборах изображений помогает в улучшении поиска по изображениям?	УК-2.В.1
77	Какие метрики наиболее важны для оценки качества информационного поиска и почему?	УК-2.В.1
78	Как можно измерить и интерпретировать точность и полноту поиска?	УК-2.В.1
79	Объясните, как вычисляется F-мера и её значение для оценки качества поиска.	УК-2.В.1
80	Как ROC-кривая помогает в оценке эффективности поисковых систем?	УК-2.В.1
81	Какие методы оптимизации запросов могут быть применены для улучшения поиска?	ПК-1.3.1
82	Как адаптивные алгоритмы могут улучшить релевантность результатов поиска?	ПК-1.3.1
83	Какие ключевые аспекты следует учитывать при реализации алгоритма TF-IDF в лабораторных условиях?	ПК-1.3.1

84	Как можно эффективно протестировать алгоритм BM25 на наборе данных?	ПК-1.3.1
85	Какие подходы следует использовать для построения векторной модели текста?	ПК-1.3.1
86	Как методы NLP, такие как токенизация и лемматизация, применяются в лабораторных работах?	ПК-1.3.1
87	Как можно улучшить методы content-based поиска в лабораторных условиях?	ПК-1.3.1
88	Какие методы можно использовать для оценки качества поиска по изображениям в лабораторных экспериментах?	ПК-1.3.1
89	Какие данные и подходы необходимы для обучения моделей распознавания изображений?	ПК-1.3.1
90	Как можно использовать метрики для анализа и улучшения качества поиска?	ПК-1.3.1
91	Какие алгоритмы машинного обучения наиболее эффективны для конкретных задач информационного поиска?	ПК-1.3.1
92	Как методы NLP могут быть применены для улучшения точности поисковых систем?	ПК-1.3.1
93	Как можно исследовать влияние различных алгоритмов ранжирования на качество поиска в лабораторной работе?	ПК-1.3.1
94	Какие методы могут быть использованы для анализа и улучшения работы поисковых систем?	ПК-1.3.1
95	Как в лабораторных условиях можно тестировать и совершенствовать методы разрешения омонимии и синонимии?	ПК-1.3.1
96	Какие подходы можно использовать для оценки и улучшения поиска по изображениям?	ПК-1.3.1
97	Как в лабораторных условиях можно исследовать влияние методов токенизации на результаты поиска?	ПК-1.3.1
98	Какие особенности разработки поисковых стратегий для различных приложений следует учитывать?	ПК-1.3.1
99	Как лабораторные работы могут помочь в понимании работы алгоритмов ранжирования?	ПК-1.3.1
100	Какие критерии оценки качества поисковых запросов важны в лабораторных экспериментах?	ПК-1.3.1
101	Как улучшение технологий NLP может повлиять на эффективность поисковых систем?	ПК-1.У.1
102	Какие современные технологии можно применить для улучшения систем поиска?	ПК-1.У.1
103	Какие основные принципы ранжирования документов в информационных системах?	ПК-1.У.1
104	Каковы основные этапы процесса информационного поиска от получения запроса до выдачи результатов?	ПК-1.У.1
105	Объясните, как алгоритмы ранжирования учитывают релевантность документа.	ПК-1.У.1
106	Как можно улучшить точность результатов поиска, используя методы анализа текста?	ПК-1.У.1
107	В чем различие между TF-IDF и BM25 в контексте поиска информации?	ПК-1.У.1

108	Как алгоритм BM25 учитывает частоту и длину документа при ранжировании?	ПК-1.У.1
109	Какие техники машинного обучения наиболее эффективно применяются для улучшения поиска?	ПК-1.У.1
110	Как методы машинного обучения помогают в адаптации поисковых систем к пользовательским запросам?	ПК-1.У.1
111	Объясните, как алгоритмы обучения на основе обратной связи могут улучшить результаты поиска.	ПК-1.У.1
112	Какие методы машинного обучения позволяют решать задачи ранжирования в информационном поиске?	ПК-1.У.1
113	Каковы ограничения модели мешка слов при обработке сложных текстов?	ПК-1.У.1
114	Какие преимущества и недостатки векторных моделей по сравнению с моделью мешка слов?	ПК-1.У.1
115	Как можно использовать модель мешка слов в современных поисковых системах?	ПК-1.У.1
116	Как структура архитектуры поисковой системы влияет на её способность обрабатывать запросы?	ПК-1.У.1
117	Какие принципы лежат в основе эффективного построения архитектуры поисковой системы?	ПК-1.У.1
118	Как методы анализа текста помогают в улучшении качества поиска в системах?	ПК-1.У.1
119	Какие современные методы NLP можно применить для обработки естественного языка в поисковых системах?	ПК-1.У.1
120	Как токенизация и лемматизация влияют на обработку текстов в различных языках?	ПК-1.У.1
121	Объясните, как извлечение ключевых слов и фраз помогает в улучшении поисковых запросов.	ПК-1.В.1
122	Как разрешение омонимии и синонимии помогает улучшить точность результатов поиска?	ПК-1.В.1
123	Какие подходы к семантическому анализу могут улучшить поиск информации?	ПК-1.В.1
124	В чем состоит суть content-based поиска и как его можно улучшить?	ПК-1.В.1
125	Как алгоритмы поиска похожих изображений помогают в создании систем визуального поиска?	ПК-1.В.1
126	Какие современные технологии машинного обучения используются для классификации и распознавания изображений?	ПК-1.В.1
127	Как можно применять обучение на больших наборах изображений для улучшения поиска?	ПК-1.В.1
128	Какие метрики наиболее полезны для оценки качества поисковых систем и как их интерпретировать?	ПК-1.В.1
129	Как можно использовать точность и полноту для оценки эффективности поисковых систем?	ПК-1.В.1
130	Объясните, как F-мера может помочь в оценке качества поиска и какие её ограничения?	ПК-1.В.1
131	Как ROC-кривая может быть использована для сравнения различных поисковых систем?	ПК-1.В.1
132	Какие методы можно использовать для улучшения запросов в поисковых системах?	ПК-1.В.1

133	Как адаптивные алгоритмы могут повысить релевантность поиска в реальных системах?	ПК-1.В.1
134	Каковы ключевые аспекты реализации алгоритма TF-IDF в лабораторных условиях?	ПК-1.В.1
135	Какие шаги необходимо предпринять для эффективного тестирования алгоритма BM25?	ПК-1.В.1
136	Как построить векторную модель представления текста для информационного поиска?	ПК-1.В.1
137	Как токенизация и лемматизация могут быть использованы в лабораторной работе для улучшения поиска?	ПК-1.В.1
138	Какие методы могут быть применены для улучшения content-based поиска в лабораторных условиях?	ПК-1.В.1
139	Как можно оценить и улучшить качество поиска по изображениям в лабораторной работе?	ПК-1.В.1
140	Какие данные необходимы для обучения моделей распознавания изображений и как их подготовить?	ПК-1.В.1
141	Как использовать метрики для анализа качества поисковых систем и какие данные для этого нужны?	ПК-5.У.1
142	Какие алгоритмы машинного обучения наиболее эффективны для информационного поиска и почему?	ПК-5.У.1
143	Как методы NLP могут быть применены для улучшения точности результатов поиска?	ПК-5.У.1
144	Какие практические шаги можно предпринять для тестирования алгоритмов ранжирования в лабораторных условиях?	ПК-5.У.1
145	Как улучшить работу поисковых систем с помощью методов анализа и обработки текста?	ПК-5.У.1
146	Какие подходы следует использовать для анализа методов разрешения омонимии и синонимии?	ПК-5.У.1
147	Как можно оценить и улучшить поиск по изображениям на основе лабораторных экспериментов?	ПК-5.У.1
148	Какие особенности следует учитывать при разработке поисковых стратегий для различных областей?	ПК-5.У.1
149	Как лабораторные работы помогают в понимании работы и улучшении алгоритмов ранжирования?	ПК-5.У.1
150	Как современные технологии NLP могут улучшить работу поисковых систем?	ПК-5.У.1
151	Какие современные методы и технологии можно применить для повышения эффективности поиска?	ПК-5.У.1
152	Какое значение имеет историческое развитие информационного поиска для современных технологий?	ПК-5.У.1
153	В чем заключается основное различие между информационным поиском и извлечением данных?	ПК-5.У.1
154	Как можно улучшить понимание запроса пользователя в системах информационного поиска?	ПК-5.У.1
155	Опишите, как алгоритм TF-IDF влияет на поиск и ранжирование документов.	ПК-5.У.1
156	Как методы BM25 и TF-IDF могут быть комбинированы для улучшения поиска?	ПК-5.У.1

157	В чем заключается разница между традиционными и современными подходами к информационному поиску?	ПК-5.У.1
158	Как обучение с обратной связью улучшает результаты информационного поиска?	ПК-5.У.1
159	Какие методы ранжирования на основе машинного обучения наиболее эффективны в различных сценариях поиска?	ПК-5.У.1
160	Как векторные модели помогают в устранении неоднозначностей в текстах?	ПК-5.У.1
161	Какие проблемы могут возникнуть при применении модели мешка слов в информационном поиске?	ПК-8.3.1
162	Как компоненты поисковой системы взаимодействуют для формирования релевантных результатов?	ПК-8.3.1
163	Какие принципы архитектуры поисковой системы необходимо учитывать при её разработке?	ПК-8.3.1
164	Как токенизация и лемматизация влияют на эффективность обработки текстов в NLP?	ПК-8.3.1
165	Как можно использовать извлечение ключевых слов для улучшения поисковых систем?	ПК-8.3.1
166	Какие методы NLP можно применить для обработки омонимии и синонимии в запросах?	ПК-8.3.1
167	Как семантический анализ может улучшить поиск и понимание текстов?	ПК-8.3.1
168	В чем заключаются особенности реализации content-based поиска и его применения?	ПК-8.3.1
169	Как алгоритмы поиска похожих изображений могут быть адаптированы для различных типов данных?	ПК-8.3.1
170	Какие методы машинного обучения наиболее эффективны для поиска и классификации изображений?	ПК-8.3.1
171	Как можно использовать большие наборы данных для обучения моделей поиска по изображениям?	ПК-8.3.1
172	Как метрики, такие как точность и полнота, помогают в оценке качества поиска?	ПК-8.3.1
173	Какие методы оценки качества поиска могут быть использованы для улучшения поиска?	ПК-8.3.1
174	Как можно интерпретировать результаты F-меры в контексте информационного поиска?	ПК-8.3.1
175	Как ROC-кривая может быть использована для анализа и оценки эффективности поисковых систем?	ПК-8.3.1
176	Какие методы оптимизации запросов могут быть применены для повышения качества поиска?	ПК-8.3.1
177	Как адаптивные алгоритмы могут быть использованы для улучшения релевантности поиска?	ПК-8.3.1
178	Как реализовать и протестировать алгоритм TF-IDF в лабораторной работе?	ПК-8.3.1
179	Какие подходы следует применять для оценки и улучшения алгоритма BM25?	ПК-8.3.1
180	Как построить и применить векторные модели для поиска в текстовых данных?	ПК-8.3.1

181	Как токенизация и лемматизация могут быть использованы для повышения точности поиска?	ПК-8.У.1
182	Какие методы можно использовать для улучшения content-based поиска в лабораторных условиях?	ПК-8.У.1
183	Как оценить качество поиска по изображениям в лабораторных экспериментах?	ПК-8.У.1
184	Какие данные необходимы для обучения моделей распознавания изображений и как их использовать?	ПК-8.У.1
185	Как метрики оценки могут быть применены для анализа и улучшения систем поиска?	ПК-8.У.1
186	Как можно использовать современные методы машинного обучения для улучшения поиска?	ПК-8.У.1
187	Какие методы NLP могут быть использованы для повышения точности и эффективности поиска?	ПК-8.У.1
188	Как тестирование различных алгоритмов ранжирования может помочь в улучшении поиска?	ПК-8.У.1
189	Какие методы можно использовать для анализа и улучшения работы поисковых систем?	ПК-8.У.1
190	Как методы разрешения омонимии и синонимии могут быть протестированы и улучшены?	ПК-8.У.1
191	Какие подходы применяются для оценки и улучшения систем поиска по изображениям?	ПК-8.У.1
192	Как токенизация и лемматизация влияют на результаты поиска и их улучшение?	ПК-8.У.1
193	Какие особенности необходимо учитывать при разработке поисковых стратегий для различных областей?	ПК-8.У.1
194	Как лабораторные работы помогают понять принципы и методы информационного поиска?	ПК-8.У.1
195	Какие критерии оценки качества поисковых запросов наиболее важны в лабораторных условиях?	ПК-8.У.1
196	Как улучшение технологий обработки естественного языка может повлиять на качество поиска?	ПК-8.У.1
197	Какое влияние на качество поиска оказывает история развития информационного поиска?	ПК-8.У.1
198	Какие ключевые концепции определяют эффективность информационного поиска?	ПК-8.У.1
199	В чем заключаются основные различия между методами TF-IDF и BM25 в контексте ранжирования документов?	ПК-8.У.1
200	Как методы машинного обучения могут улучшить ранжирование результатов поиска?	ПК-8.У.1
201	Опишите, как алгоритмы машинного обучения адаптируются к обратной связи от пользователей.	ПК-8.В.1
202	Какие методы ранжирования на основе машинного обучения наиболее популярны в информационном поиске?	ПК-8.В.1
203	Как модели представления текста могут быть улучшены для повышения точности поиска?	ПК-8.В.1
204	Какие основные компоненты и принципы работы поисковых систем влияют на их производительность?	ПК-8.В.1

205	Как NLP технологии могут использоваться для оптимизации поисковых запросов?	ПК-8.В.1
206	Опишите, как токенизация и лемматизация помогают в обработке текстовых данных для поиска.	ПК-8.В.1
207	Как методы разрешения омонимии и синонимии улучшают результаты информационного поиска?	ПК-8.В.1
208	Какие подходы существуют для анализа и извлечения ключевых слов из текста?	ПК-8.В.1
209	Как content-based поиск применяется для улучшения результатов поиска по изображениям?	ПК-8.В.1
210	Какие алгоритмы можно использовать для поиска похожих изображений в больших базах данных?	ПК-8.В.1
211	Опишите, как обучение на больших наборах данных изображений может улучшить алгоритмы распознавания.	ПК-8.В.1
212	Какие метрики оценки поиска являются наиболее информативными для анализа качества результатов?	ПК-8.В.1
213	Как можно измерить и интерпретировать точность и полноту результатов поиска?	ПК-8.В.1
214	В чем заключается вычисление и применение F-меры для оценки эффективности поисковых систем?	ПК-8.В.1
215	Как ROC-кривая может использоваться для оценки качества поиска и что она показывает?	ПК-8.В.1
216	Какие методы оптимизации запросов помогут повысить релевантность результатов поиска?	ПК-8.В.1
217	Как адаптивные алгоритмы могут быть использованы для улучшения качества поиска в различных приложениях?	ПК-8.В.1
218	Какие ключевые аспекты необходимо учитывать при реализации алгоритма TF-IDF в реальных условиях?	ПК-8.В.1
219	Как можно протестировать и оценить эффективность алгоритма BM25 на различных наборах данных?	ПК-8.В.1
220	Какие методы можно использовать для построения и оценки векторных моделей представления текста?	ПК-8.В.1

Перечень тем контрольных работ по дисциплине обучающихся заочной формы обучения, представлены в таблице 19.

Таблица 19 – Перечень контрольных работ

№ п/п	Перечень контрольных работ
1.	Алгоритмы машинного обучения в информационном поиске.
2.	Поисковые системы и их компоненты.
3.	Применение машинного обучения в поиске по изображениям.

10.4. Методические материалы, определяющие процедуры оценивания индикаторов, характеризующих этапы формирования компетенций, содержатся в локальных нормативных актах ГУАП, регламентирующих порядок и процедуру проведения текущего контроля успеваемости и промежуточной аттестации обучающихся ГУАП.

11. Методические указания для обучающихся по освоению дисциплины

11.1. Методические указания для обучающихся по освоению лекционного материала

Основное назначение лекционного материала – логически стройное, системное, глубокое и ясное изложение учебного материала. Назначение современной лекции в рамках дисциплины не в том, чтобы получить всю информацию по теме, а в освоении фундаментальных проблем дисциплины, методов научного познания, новейших достижений научной мысли. В учебном процессе лекция выполняет методологическую, организационную и информационную функции. Лекция раскрывает понятийный аппарат конкретной области знания, её проблемы, дает цельное представление о дисциплине, показывает взаимосвязь с другими дисциплинами.

Планируемые результаты при освоении обучающимися лекционного материала:

- получение современных, целостных, взаимосвязанных знаний, уровень которых определяется целевой установкой к каждой конкретной теме;
- получение опыта творческой работы совместно с преподавателем;
- развитие профессионально-деловых качеств, любви к предмету и самостоятельного творческого мышления.
- появление необходимого интереса, необходимого для самостоятельной работы;
- получение знаний о современном уровне развития науки и техники и о прогнозе их развития на ближайшие годы;
- научиться методически обрабатывать материал (выделять главные мысли и положения, приходить к конкретным выводам, повторять их в различных формулировках);
- получение точного понимания всех необходимых терминов и понятий.

Лекционный материал может сопровождаться демонстрацией слайдов и использованием раздаточного материала при проведении коротких дискуссий об особенностях применения отдельных тематик по дисциплине.

Структура предоставления лекционного материала:

Основные концепции информационного поиска. Определение информационного поиска. История развития и основные этапы.

Основные методы ранжирования. TF-IDF и его применение. BM25 и другие алгоритмы ранжирования. Алгоритмы машинного обучения в информационном поиске. Обучение на основе обратной связи Методы ранжирования с использованием ML

Поисковые системы и их компоненты. Компоненты поисковой системы. Архитектура и принципы работы поисковых систем. Модели представления информации. Модель мешка слов и её ограничения. Векторные модели и методы их построения

Основы анализа текста. Токенизация и лемматизация. Извлечение ключевых слов и фраз. Использование NLP для улучшения поисковых запросов. Разрешение омонимии и синонимии. Поиск с использованием семантических анализаторов

Основные методы и алгоритмы поиска по изображениям. Content-based поиск. Поиск похожих изображений. Применение машинного обучения в поиске по изображениям. Обучение моделей на больших наборах данных изображений. Алгоритмы распознавания и классификации изображений

11.2. Методические указания для обучающихся по выполнению лабораторных работ

В ходе выполнения лабораторных работ обучающийся должен углубить и закрепить знания, практические навыки, овладеть современной методикой и техникой эксперимента в соответствии с квалификационной характеристикой обучающегося. Выполнение лабораторных работ состоит из экспериментально-практической, расчетно-аналитической частей и контрольных мероприятий.

Выполнение лабораторных работ обучающимся является неотъемлемой частью изучения дисциплины, определяемой учебным планом, и относится к средствам, обеспечивающим решение следующих основных задач обучающегося:

- приобретение навыков исследования процессов, явлений и объектов, изучаемых в рамках данной дисциплины;
- закрепление, развитие и детализация теоретических знаний, полученных на лекциях;
- получение новой информации по изучаемой дисциплине;
- приобретение навыков самостоятельной работы с лабораторным оборудованием и приборами.

Задание и требования к проведению лабораторных работ

Примеры заданий на лабораторные работы.

ЛР №1. Исследование основных концепций информационного поиска:

Задание: Проведите анализ различных подходов к информационному поиску, таких как ключевые слова, запросы и метаданные. Сравните их эффективность на наборе текстовых данных. Опишите, как каждый подход влияет на результаты поиска и предложите рекомендации для улучшения поисковой системы.

ЛР №2. Реализация и оценка алгоритмов ранжирования TF-IDF и BM25:

Задание: Реализуйте алгоритмы TF-IDF и BM25 для ранжирования документов по запросу. Используйте выборку текстов для оценки эффективности алгоритмов. Сравните результаты ранжирования, определите их сильные и слабые стороны, и предложите возможные улучшения.

ЛР 3. Построение и тестирование векторной модели представления текста:

Задание: Постройте векторную модель представления текстов (например, с использованием word2vec или GloVe). Проведите тестирование модели на наборе текстов, оцените её способность захватывать семантические связи между словами, и проанализируйте результаты.

ЛР №4. Применение методов NLP для улучшения поисковых запросов:

Задание: Реализуйте и протестируйте методы NLP, такие как токенизация, лемматизация и извлечение ключевых слов, для улучшения качества поисковых запросов. Оцените влияние этих методов на результаты поиска и предложите улучшения для повышения их эффективности.

ЛР №5. Реализация и тестирование алгоритмов поиска по изображениям:

Задание: Реализуйте алгоритмы поиска по изображениям, включая content-based поиск и поиск похожих изображений. Используйте набор изображений для оценки точности и эффективности алгоритмов. Сравните результаты и предложите способы их улучшения.

Требования к проведению ЛР:

- Используйте предоставленные наборы данных и инструменты.
- Реализуйте и протестируйте алгоритмы или методы в соответствии с заданиями.
- Соблюдайте точность и корректность выполнения задач, предоставляя подробные объяснения своих действий и решений.
- Работайте в установленный срок и следуйте предоставленным инструкциям и рекомендациям.

Структура и форма отчета о лабораторной работе

Отчет о лабораторной работе должен содержать следующие разделы:

Введение: Обзор задачи и цели работы.

Методы: Описание используемых алгоритмов, моделей и подходов.

Результаты: Презентация полученных данных и результатов экспериментов.

Анализ: Оценка результатов, обсуждение выявленных проблем и возможных улучшений.

Заключение: Итоги работы и рекомендации.

Требования к оформлению отчета о лабораторной работе

Отчет должен быть написан в ясном и понятном стиле, без грамматических и орфографических ошибок.

Используйте стандартный формат документа (например, А4, шрифт 12, межстрочный интервал 1,5).

Включите диаграммы, графики и таблицы, если это необходимо для иллюстрации результатов.

Следуйте структуре отчета, предоставленной выше, и обеспечьте четкость и логичность изложения материала.

11.3. Методические указания для обучающихся по прохождению самостоятельной работы

В ходе выполнения самостоятельной работы, обучающийся выполняет работу по заданию и при методическом руководстве преподавателя, но без его непосредственного участия.

Для обучающихся по заочной форме обучения, самостоятельная работа может включать в себя контрольную работу.

В процессе выполнения самостоятельной работы, у обучающегося формируется целесообразное планирование рабочего времени, которое позволяет им развивать умения и навыки в усвоении и систематизации приобретаемых знаний, обеспечивает высокий уровень успеваемости в период обучения, помогает получить навыки повышения профессионального уровня.

Методическими материалами, направляющими самостоятельную работу обучающихся, являются:

- учебно-методический материал по дисциплине;
- методические указания по выполнению контрольных работ (для обучающихся по заочной форме обучения).

11.4. Методические указания для обучающихся по прохождению текущего контроля успеваемости.

Текущий контроль успеваемости предусматривает контроль качества знаний обучающихся, осуществляемого в течение семестра с целью оценивания хода освоения дисциплины.

Текущий контроль успеваемости проводится в середине семестра в форме тестирования. Каждый билет включает в себя 5 вопросов. За каждый вопрос можно получить от 0 до 10 баллов (всего от 0 до 50 баллов). Полученная сумма баллов сохраняется до конца семестра и суммируется с суммой баллов, полученных при прохождении промежуточной аттестации.

11.5. Методические указания для обучающихся по прохождению промежуточной аттестации.

Промежуточная аттестация обучающихся предусматривает оценивание промежуточных и окончательных результатов обучения по дисциплине. Она включает в себя:

- зачет – это форма оценки знаний, полученных обучающимся в ходе изучения учебной дисциплины в целом или промежуточная (по окончании семестра) оценка знаний обучающимся по отдельным разделам дисциплины с аттестационной оценкой «зачтено» или «не зачтено».

Промежуточная аттестация проводится в виде тестирования. В билет включается текст на перевод объемом 80...100 слов, оцениваемый по шкале от 0 до 50 баллов. Полученная оценка суммируется с баллами, набранными в течение семестра.

Лист внесения изменений в рабочую программу дисциплины

Дата внесения изменений и дополнений. Подпись внесшего изменения	Содержание изменений и дополнений	Дата и № протокола заседания кафедры	Подпись зав. кафедрой