

*Федеральное государственное автономное образовательное учреждение высшего образования
«Российский университет дружбы народов»*

Факультет физико-математических и естественных наук

Рекомендовано МССН
02.00.00 «Компьютерные и
информационные науки»

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Наименование дисциплины

Методы интеллектуального анализа текстов

Рекомендуется для направления подготовки

02.04.02 Фундаментальная информатика и информационные технологии

Направленность программы (профиль)

«Управление инфокоммуникациями и интеллектуальные системы»

Квалификация (степень) выпускника магистр

1. Цели и задачи дисциплины: Целью дисциплины является ознакомление студентов с методами интеллектуального анализа текстов, освоение принципов морфологического, синтаксического и семантического анализа, извлечения информации из текстов, классификации текстов. Основной задачей дисциплины является получение навыков работы с библиотеками обработки текстов.

2. Место дисциплины в структуре ОП ВО:

Дисциплина Методы интеллектуального анализа текстов относится к обязательной части блока 1 учебного плана.

В таблице № 1 приведены предшествующие и последующие дисциплины, направленные на формирование компетенций дисциплины в соответствии с матрицей компетенций ОП ВО.

Таблица № 1

Предшествующие и последующие дисциплины, направленные на формирование компетенций

| № п/п | Шифр и наименование компетенции | Предшествующие дисциплины | Последующие дисциплины (группы дисциплин) |
|---|---------------------------------|--|---|
| Общекультурные компетенции | | | |
| | УК: 1, 7 | Дизайн интерактивных систем, Алгоритмические основы мультимедийных технологий | Интеллектуальные динамические системы |
| Общепрофессиональные компетенции | | | |
| | ОПК: 2, 3, 4, 5 | Дизайн интерактивных систем, Алгоритмические основы мультимедийных технологий | Интеллектуальные динамические системы |
| Профессиональные компетенции (вид профессиональной деятельности) | | | |
| | ПК: 1, 2 | Дизайн интерактивных систем, Алгоритмические основы мультимедийных технологий | Интеллектуальные динамические системы |
| Профессионально-специализированные компетенции специализации | | | |
| | | | |

УК-1: Способен осуществлять поиск, критический анализ проблемных ситуаций на основе системного подхода, выработать стратегию действий;

УК-2 Способен управлять проектом на всех этапах его жизненного цикла

УК-7: Способен: искать нужные источники информации и данные, воспринимать, анализировать, запоминать и передавать информацию с использованием цифровых средств, а также с помощью алгоритмов при работе с полученными из различных источников данными с целью эффективного использования полученной информации для решения задач; проводить оценку информации, ее достоверность, строить логические умозаключения на основании поступающих информации и данных;

ОПК-2: Способен применять компьютерные / суперкомпьютерные методы, современное программное обеспечение (в том числе отечественного производства) для решения задач профессиональной деятельности.

ОПК-3: Способен проводить анализ математических моделей, создавать инновационные методы решения прикладных задач профессиональной деятельности в области информатики и математического моделирования

ОПК-4: Способен оптимальным образом комбинировать существующие информационно-

коммуникационные технологии для решения задач в области профессиональной деятельности с учетом требований информационной безопасности;

ОПК-5 Способен устанавливать и сопровождать программное обеспечение информационных систем, осуществлять эффективное управление разработкой программных средств и проектов.

ПК-1: Проведение работ по обработке и анализу научно-технической информации и результатов исследований.

ПК-2: Организационное и технологическое обеспечение проектирования и дизайна ИС.

3. Требования к результатам освоения дисциплины:

Процесс изучения дисциплины направлен на формирование следующих компетенций:

УК-1; УК-7; ОПК-2; ОПК-3; ОПК-4; ОПК-5; ПК-1; ПК-2

УК-1: Способен осуществлять поиск, критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий:

- УК-1.1 Знает принципы сбора, отбора и обобщения информации
- УК-1.2 Умеет соотносить разнородные явления и систематизировать их в рамках избранных видов профессиональной деятельности
- УК-1.3 Имеет практический опыт работы с информационными источниками, опыт научного поиска, создания научных текстов.

УК-7: Способен: искать нужные источники информации и данные, воспринимать, анализировать, запоминать и передавать информацию с использованием цифровых средств, а также с помощью алгоритмов при работе с полученными из различных источников данными с целью эффективного использования полученной информации для решения задач; проводить оценку информации, ее достоверность, строить логические умозаключения на основании поступающих информации и данных;

- УК-7.1 Знает принципы применения цифровых технологий для сбора, отбора и обобщения информации
- УК-7.2 Умеет применять цифровые технологии для поиска, обработки, анализа, хранения и представления информации в области фундаментальной информатики и информационных технологий
- УК-7.3 Владеет навыками применения цифровых технологий и методов поиска, обработки, анализа, хранения и представления информации в области фундаментальной информатики и информационных технологий

ОПК-2: Способен применять компьютерные / суперкомпьютерные методы, современное программное обеспечение (в том числе отечественного производства) для решения задач профессиональной деятельности.

- ОПК-2.1 Знает основные положения и концепции в области программирования, архитектуру языков программирования, теории коммуникации, знает основную терминологию, знаком с перечнем ПО, включенного в Единый Реестр Российских программ
- ОПК-2.2 Умеет анализировать типовые языки программирования, составлять программы
- ОПК-2.3 Имеет практический опыт решения задач анализа, интеграции различных типов программного обеспечения, анализа типов коммуникации

ОПК-3: Способен проводить анализ математических моделей, создавать инновационные методы решения прикладных задач профессиональной деятельности в области информатики и математического моделирования

- ОПК-3.1 Знает методы теории алгоритмов, методы системного и прикладного программирования, основные положения и концепции в области математических, информационных и имитационных моделей
- ОПК-3.2 Умеет соотносить знания в области программирования, интерпретацию прочитанного, определять и создавать информационные ресурсы глобальных

сетей, образовательного контента, средств тестирования систем

- ОПК-3.3 Имеет практический опыт применения и разработки программного обеспечения, тестирования программных продуктов

ОПК-4: Способен оптимальным образом комбинировать существующие информационно-коммуникационные технологии для решения задач в области профессиональной деятельности с учетом требований информационной безопасности

- ОПК-4.1 Знает принципы сбора и анализа информации, создания информационных систем на стадиях жизненного цикла
- ОПК-4.2 Умеет осуществлять управление проектами информационных систем
- ОПК-4.3 Имеет практический опыт анализа и интерпретации информационных систем

ОПК-5 Способен устанавливать и сопровождать программное обеспечение информационных систем, осуществлять эффективное управление разработкой программных средств и проектов.

- ОПК-5.1 Знает методику установки и администрирования информационных систем и баз данных. Знаком с перечнем ПО, входящим в Единый реестр российских программ
- ОПК-5.2 Умеет реализовывать техническое сопровождение информационных систем и баз данных
- ОПК-5.3 Имеет практические навыки установки и инсталляции программных комплексов

ПК-1: Проведение работ по обработке и анализу научно-технической информации и результатов исследований.

- ПК-1.3 Умеет применять полученные знания в области фундаментальных научных основ математики и информатики, а также решать стандартные задачи собственной научно-исследовательской деятельности; умеет решать научные задачи с пониманием существующих подходов к верификации моделей по тематике исследований в соответствии с выбранной методикой

ПК-2: Организационное и технологическое обеспечение проектирования и дизайна ИС.

- ПК-2.1 Знает этапы жизненного цикла разработки программных систем, понятие архитектуры и виды архитектур, методологии разработки программных систем, современные CASE средства; умеет самостоятельно выбирать подходящее CASE-средство для решения задач на каждом этапе жизненного цикла разработки программных систем; владеет навыками использования CASE-инструментов для разработки программных систем
- ПК-2.2 Знает основные понятия и определения, относящиеся к концепции построения интерактивных систем; основные методы сбора информации о пользователях и модели поведения пользователя; методы проектирования интерактивных систем; умеет строить модели поведения пользователей; производить постановку задачи по разработке интерактивной системы, разрабатывать спецификации интерактивных систем; производить оценку качества дизайна интерактивных систем; владеет методами оценки эффективности интерактивных систем
- ПК-2.3 Знает основы проектирования и элементы архитектурных решений информационных систем; умеет применять в практической деятельности профессиональные стандарты в области информационных технологий; имеет практический опыт составления технического задания на разработку информационной системы
- ПК-2.4 Умеет проводить анализ требований к информационной системе; разрабатывать варианты реализации информационной системы; проводить оценку качества, надежности и эффективности информационной системы
- ПК-2.5 Знает основы программирования; современные методики тестирования

разрабатываемых информационных систем; современные инструменты и методы верификации программного кода.

- ПК-2.7 Методики описания и моделирования бизнес-процессов; средства моделирования бизнес-процессов; основы реинжиниринга бизнес-процессов организации.
- ПК-2.8 Знает устройство и функционирование современных информационных систем; современные стандарты взаимодействия информационных систем; программные средства и платформы инфраструктуры информационных технологий организации; современные подходы и стандарты автоматизации организации (например CRM, ERP, ITIL)

В результате изучения дисциплины студент должен:

Знать: основные методы, алгоритмы и программные инструменты интеллектуального анализа текстов.

Уметь: применять на практике полученные знания для решения задач по обработке больших массивов текстов.

Владеть: навыками использования программных инструментов анализа текстов для решения конкретных задач.

4. Объем дисциплины и виды учебной работы

Общая трудоемкость дисциплины составляет 4 зачетных единицы.

| № | Вид учебной работы | Всего часов | Семестры |
|-------|---|-------------|---------------------|
| | | | Семестр 3, модуль 5 |
| 1. | Аудиторные занятия (всего) | 54 | 54 |
| 1.1 | Лекции | 18 | 18 |
| 1.2.1 | <i>Практические занятия (ПЗ)</i> | 36 | 36 |
| 1.2.2 | <i>Семинары (С)</i> | | |
| 1.2.3 | <i>Лабораторные работы (ЛР)</i> | | |
| 2. | Самостоятельная работа студентов (ак. часов) | 90 | 90 |
| 3. | Общая трудоемкость (ак. часов) | 144 | 144 |
| 4. | Общая трудоемкость (зачетных единиц) | 4 | 4 |

5. Содержание дисциплины

5.1. Содержание разделов дисциплины

| Название разделов дисциплины | Краткое содержание тем дисциплины |
|---|--|
| 1. Введение в интеллектуальный анализ текстов | 1. Основные определения. Постановки задач обработки текстов и подходы к их решению. |
| | 2. Естественный язык. Типология естественных языков. Уровни анализа. Лингвистическое исследование. |
| 2. Методы лингвистического анализа текстов | 1. Методы морфологического анализа. Проблема морфологической многозначности. Методы снятия морфологической многозначности. |

| | |
|--|--|
| | 2. Синтаксис естественных языков. Способы описания синтаксической структуры предложения. Методы синтаксического анализа. |
| | 3. Способы формализации семантики. Методы семантического анализа. Реляционно-ситуационный анализ текста. Дистрибутивная семантика. |
| 3. Прикладные задачи интеллектуального анализа текстов | 1. Извлечение информации из текстов. Классификация и кластеризация текстов. |
| | 2. Методы машинного обучения для решения задач обработки естественного языка. |

5.2. Разделы дисциплин и виды занятий

| № п/п | Наименование раздела дисциплины | Лекц | Практ зан. | Лаб. зан. | Семи н | СРС | Всего час. |
|-------|---|------|------------|-----------|--------|-----|------------|
| 1. | Введение в интеллектуальный анализ текстов | 4 | 8 | | | 20 | 32 |
| 2. | Методы лингвистического анализа текстов | 8 | 16 | | | 50 | 74 |
| 3. | Прикладные задачи интеллектуального анализа текстов | 6 | 12 | | | 20 | 38 |
| | Итого: | 18 | 36 | | | 90 | 144 |

6. Лабораторный практикум (при наличии)

Не предусмотрено

7. Практические занятия (семинары) (при наличии)

| № п/п | № раздела дисциплины | Наименование лабораторных работ | Трудоемкость (час.) |
|-------|----------------------|--|---------------------|
| 1. | 1 | Исследование классификации естественных языков. | 8 |
| 2. | 2 | Разработка алгоритма морфологического анализа текстов, основанного на словарном подходе. | 16 |
| 3. | 3 | Разработка метода извлечения информации из текстов. | 12 |

8. Материально-техническое обеспечение дисциплины:

Мультимедийная учебная аудитория для проведения учебных занятий (в том числе для практического и лекционного типов занятий, индивидуальных и групповых консультаций, текущего контроля и промежуточной аттестации).

Компьютерные (дисплейные) классы с доступом к сети Интернет и электронно-образовательной среде Университета для выполнения обучающимися практических заданий по дисциплине и для проведения обучающимися самостоятельной работы, компьютерного тестирования обучающихся (при необходимости).

9. Информационное обеспечение дисциплины

а) программное обеспечение: ОС Windows, MS Office (программа корпоративного лицензирования (Microsoft Subscription) Enrollment for Education Solutions), браузер Firefox (лицензия MPL-2.0), браузер Chrome (лицензия Google Chrome Terms of Service), Dev-C++ (лицензия GNU GPL), свободно распространяемые анализаторы естественных языков

б) базы данных, информационно-справочные и поисковые системы:

1. Сайт библиотеки РУДН <http://lib.rudn.ru/>
2. ТУИС <http://esystem.pfur.ru/>

10. Учебно-методическое обеспечение дисциплины:

а) основная литература

1. Николаев И.С., Митренина О.В., Ландо Т.М. (Ред.). Прикладная и компьютерная лингвистика URSS. 2016. 320 с.
2. Баранов А.Н. Введение в прикладную лингвистику. - М.: Либроком, 2013. - 368 с.

б) дополнительная литература

Jurafsky D., Martin J. H. Speech and language processing. – Pearson, 2014.

11. Методические указания для обучающихся по освоению дисциплины (модуля)

Учебным планом на изучение дисциплины отводится один семестр. В течение семестра выполняются практические работы и контрольные мероприятия. В конце семестра производится итоговый контроль знаний.

11.1 Структура лабораторных (практических) занятий

1. Задания по практическим работам выполняются или в компьютерных классах, или удаленно в MS Teams в соответствии с календарным планом и методическими указаниями по выполнению практических работ по дисциплине.
2. По результатам выполнения каждой практической работы студентом готовится компьютерная программа с корректным набором данных.

11.2. Самостоятельная работа студента

1. Часть практических работ предусматривает задания для индивидуальной самостоятельной работы студента (домашние задания), обязательные для выполнения.
2. Выполнение заданий для самостоятельной работы позволяет студенту приобрести дополнительные навыки и закрепить знания по изучаемой теме.

12. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине

ФОС по дисциплине представлен в приложении к данной программе.

Разработчик:

доцент кафедры
информационных технологий,
к.ф.-м.н., доцент

Заведующий кафедрой

информационных технологий, д.ф.-м.н.

Руководитель программы

Заведующий кафедрой прикладной
информатики и теории вероятностей,
д.т.н., профессор

 И.В. Смирнов
 Ю.Н. Орлов
 К.Е. Самуйлов

Федеральное государственное автономное образовательное учреждение высшего образования «Российский университет дружбы народов»

Факультет физико-математических и естественных наук

Кафедра информационных технологий

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

ПО УЧЕБНОЙ ДИСЦИПЛИНЕ

Методы интеллектуального анализа текстов

(наименование дисциплины)

Рекомендуется для направления подготовки

02.04.02 — Фундаментальная информатика и информационные технологии

(код и наименование направления подготовки)

«Управление инфокоммуникациями и интеллектуальные системы»

(наименование профиля подготовки)

магистр

Квалификация (степень) выпускника

Паспорт фонда оценочных средств по дисциплине
Методы интеллектуального анализа текстов

Направление: 02.04.02 — «Фундаментальная информатика и информационные технологии»

| Код контролируемой компетенции или ее части | Контролируемый раздел дисциплины | Контролируемая тема дисциплины | Наименование оценочного средства | | | Баллы темы | Баллы раздела |
|---|---|---|----------------------------------|--------------------------|----------------|------------|---------------|
| | | | Текущий контроль | Промежуточная аттестация | | | |
| | | | Работа на практических занятиях | Выполнение ДЗ | Экзамен /Зачет | | |
| УК: 1, 7 | Раздел 1. Введение в интеллектуальный анализ текстов | Тема 1: Основные определения. Постановки задач обработки текстов и подходы к их решению. Тема 2. Естественный язык. Типология естественных языков. Уровни анализа. Лингвистическое исследование. | 5 | 13 | 2 | 20 | 20 |
| ОПК: 2, 3, 4, 5 ПК: 1, 2 | Раздел 2. Методы лингвистического анализа текстов | Тема 1: Методы морфологического анализа. Проблема морфологической многозначности. Методы снятия морфологической многозначности. Тема 2. Синтаксис естественных языков. Способы описания синтаксической структуры предложения. Методы синтаксического анализа. Тема 3. Способы формализации семантики. Методы семантического анализа. Реляционно-ситуационный анализ текста. Дистрибутивная семантика. | 15 | 39 | 6 | 60 | 60 |
| ОПК: 3, 4, ПК: 1, 2 | Раздел 3. Прикладные задачи интеллектуального анализа текстов | Тема 1: Извлечение информации из текстов. Классификация и кластеризация текстов. Тема 2. Методы машинного обучения для решения задач обработки естественного языка. | 5 | 13 | 2 | 20 | 20 |
| | | ИТОГО: | 25 | 65 | 10 | 100 | 100 |

Процесс изучения дисциплины направлен на формирование следующих компетенций:

УК-1; УК-7; ОПК-2; ОПК-3; ОПК-4; ОПК-5; ПК-1; ПК-2

УК-1: Способен осуществлять поиск, критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий:

- УК-1.1 Знает принципы сбора, отбора и обобщения информации
- УК-1.2 Умеет соотносить разнородные явления и систематизировать их в рамках избранных видов профессиональной деятельности
- УК-1.3 Имеет практический опыт работы с информационными источниками, опыт научного поиска, создания научных текстов.

УК-7: Способен: искать нужные источники информации и данные, воспринимать, анализировать, запоминать и передавать информацию с использованием цифровых средств, а также с помощью алгоритмов при работе с полученными из различных источников данными с целью эффективного использования полученной информации для решения задач; проводить оценку информации, ее достоверность, строить логические умозаключения на основании поступающих информации и данных;

- УК-7.1 Знает принципы применения цифровых технологий для сбора, отбора и обобщения информации
- УК-7.2 Умеет применять цифровые технологии для поиска, обработки, анализа, хранения и представления информации в области фундаментальной информатики и информационных технологий
- УК-7.3 Владеет навыками применения цифровых технологий и методов поиска, обработки, анализа, хранения и представления информации в области фундаментальной информатики и информационных технологий

ОПК-2: Способен применять компьютерные / суперкомпьютерные методы, современное программное обеспечение (в том числе отечественного производства) для решения задач профессиональной деятельности.

- ОПК-2.1 Знает основные положения и концепции в области программирования, архитектуру языков программирования, теории коммуникации, знает основную терминологию, знаком с перечнем ПО, включенного в Единый Реестр Российских программ
- ОПК-2.2 Умеет анализировать типовые языки программирования, составлять программы
- ОПК-2.3 Имеет практический опыт решения задач анализа, интеграции различных типов программного обеспечения, анализа типов коммуникации

ОПК-3: Способен проводить анализ математических моделей, создавать инновационные методы решения прикладных задач профессиональной деятельности в области информатики и математического моделирования

- ОПК-3.1 Знает методы теории алгоритмов, методы системного и прикладного программирования, основные положения и концепции в области математических, информационных и имитационных моделей
- ОПК-3.2 Умеет соотносить знания в области программирования, интерпретацию прочитанного, определять и создавать информационные ресурсы глобальных сетей, образовательного контента, средств тестирования систем
- ОПК-3.3 Имеет практический опыт применения и разработки программного обеспечения, тестирования программных продуктов

ОПК-4: Способен оптимальным образом комбинировать существующие информационно-коммуникационные технологии для решения задач в области профессиональной деятельности с учетом требований информационной безопасности

- ОПК-4.1 Знает принципы сбора и анализа информации, создания информационных систем на стадиях жизненного цикла
- ОПК-4.2 Умеет осуществлять управление проектами информационных систем
- ОПК-4.3 Имеет практический опыт анализа и интерпретации информационных

систем

ОПК-5 Способен устанавливать и сопровождать программное обеспечение информационных систем, осуществлять эффективное управление разработкой программных средств и проектов.

- ОПК-5.1 Знает методику установки и администрирования информационных систем и баз данных. Знаком с перечнем ПО, входящим в Единый реестр российских программ
- ОПК-5.2 Умеет реализовывать техническое сопровождение информационных систем и баз данных
- ОПК-5.3 Имеет практические навыки установки и инсталляции программных комплексов

ПК-1: Проведение работ по обработке и анализу научно-технической информации и результатов исследований.

- ПК-1.3 Умеет применять полученные знания в области фундаментальных научных основ математики и информатики, а также решать стандартные задачи собственной научно-исследовательской деятельности; умеет решать научные задачи с пониманием существующих подходов к верификации моделей по тематике исследований в соответствии с выбранной методикой

ПК-2: Организационное и технологическое обеспечение проектирования и дизайна ИС.

- ПК-2.1 Знает этапы жизненного цикла разработки программных систем, понятие архитектуры и виды архитектур, методологии разработки программных систем, современные CASE средства; умеет самостоятельно выбирать подходящее CASE-средство для решения задач на каждом этапе жизненного цикла разработки программных систем; владеет навыками использования CASE-инструментов для разработки программных систем
- ПК-2.2 Знает основные понятия и определения, относящиеся к концепции построения интерактивных систем; основные методы сбора информации о пользователях и модели поведения пользователя; методы проектирования интерактивных систем; умеет строить модели поведения пользователей; производить постановку задачи по разработке интерактивной системы, разрабатывать спецификации интерактивных систем; производить оценку качества дизайна интерактивных систем; владеет методами оценки эффективности интерактивных систем
- ПК-2.3 Знает основы проектирования и элементы архитектурных решений информационных систем; умеет применять в практической деятельности профессиональные стандарты в области информационных технологий; имеет практический опыт составления технического задания на разработку информационной системы
- ПК-2.4 Умеет проводить анализ требований к информационной системе; разрабатывать варианты реализации информационной системы; проводить оценку качества, надежности и эффективности информационной системы
- ПК-2.5 Знает основы программирования; современные методики тестирования разрабатываемых информационных систем; современные инструменты и методы верификации программного кода.
- ПК-2.7 Методики описания и моделирования бизнес-процессов; средства моделирования бизнес-процессов; основы реинжиниринга бизнес-процессов организации.
- ПК-2.8 Знает устройство и функционирование современных информационных систем; современные стандарты взаимодействия информационных систем; программные средства и платформы инфраструктуры информационных технологий организации; современные подходы и стандарты автоматизации организации (например CRM, ERP, ITIL)

Примерный перечень оценочных средств

по дисциплине Методы интеллектуального анализа текстов

| п/п | Наименование оценочного средства | Краткая характеристика оценочного средства | Представление оценочного средства в фонде |
|-------------------------------|----------------------------------|---|---|
| <i>Аудиторная работа</i> | | | |
| 1 | Практические занятия | Выполнение практического задания по разделу дисциплины. | Примерные темы практических заданий |
| 3 | Экзамен\Зачет | Форма проверки качества выполнения студентами лабораторных работ, усвоения учебного материала и выполнения в процессе обучения всех учебных поручений в соответствии с утвержденной программой. | Примеры вопросов |
| <i>Самостоятельная работа</i> | | | |
| 1 | Домашнее задание | Продукт самостоятельной работы студента, представляющий собой выполнение практической задачи. | Примерные темы домашних заданий |

Примерный перечень вопросов промежуточного и итогового контроля знаний

по дисциплине Методы интеллектуального анализа текстов

1. Анализ естественного языка. Основные определения и постановки задач.
2. Естественный язык. Типы классификаций естественных языков. Строение языка и виды анализа. Лингвистический эксперимент.
3. Морфологический анализ. Снятие морфологической многозначности.
4. Синтаксический анализ. Transition-based методы построения дерева зависимостей.
5. Семантический анализ. Дистрибутивная семантика. Предикатно-аргументная семантика и установление семантически х ролей.
6. Извлечение информации из текстов. Определения и основные подходы.
7. Проблема недостатка размеченных данных и её решение. Машинное обучение в NLP. Кроме основного вопроса на контроле задаются дополнительные вопросы по всем темам.

Примерные темы практических занятий

по дисциплине Методы интеллектуального анализа текстов

1. Разработка токенизатора.
2. Использование библиотек морфологического анализа текстов.
3. Применение библиотек синтаксического анализа текстов.
4. Применение языковых моделей для решения различных задач обработки текстов.

Примерные темы домашних заданий по дисциплине Методы интеллектуального анализа текстов

1. Выявление депрессии у автора текста
 2. Чат-бот, автоматизирующий часто задаваемые вопросы (FAQ) с заданного сайта.
 3. Автоматическая генерация заголовков новостей.
 5. Выявление фейковых новостей.
 6. Извлечение именованных сущностей из текстов.
 8. Кластеризация коллекций текстов.
 9. Атрибуция текстов.
 10. Определение несвязности текста.
- Задачи необходимо выполнять на основе готовых анализаторов текстов.

Критерии оценки

по дисциплине Методы интеллектуального анализа текстов

Итоговая оценка выставляется по сумме набранных баллов за контрольные работы, лабораторные и домашние работы, а в случае выполнения дополнительных заданий, за реферат и доклад по материалам реферата. Для учащихся, которые набрали от 31 до 50 баллов, проводится экзамен (зачет), в рамках которого за ответы на вопросы билетов учащиеся могут набрать недостающее до зачета (51) число баллов.

95-100 баллов:

- полные ответы на вопросы контрольных работ;
- полное выполнение лабораторных работ;
- высокий уровень подготовки реферата (в случае выполнения этого задания);
- высокий уровень подготовки доклада (в случае выполнения этого задания);
- активное участие в мероприятиях, предусмотренных программой дисциплины;
- систематизированное, глубокое и полное освоение навыков и компетенций по всем разделам программы дисциплины;
- использование научной терминологии, стилистически грамотное, логически правильное изложение ответа на вопросы, умение делать обоснованные выводы;
- умение эффективно использовать методiku программы дисциплины в постановке и решении научных и профессиональных задач;
- выраженная способность самостоятельно и творчески решать поставленные задачи;
- полная самостоятельность и творческий подход при изложении материала по программе дисциплины;
- полное и глубокое усвоение основной и дополнительной литературы, рекомендованной программой дисциплины и преподавателем.

86- 94 балла:

- хорошие ответы на вопросы контрольных работ;
- полное выполнение лабораторных работ;
- хороший уровень подготовки реферата (в случае выполнения этого задания);
- хороший уровень подготовки доклада (в случае выполнения этого задания);
- участие в мероприятиях, предусмотренных программой дисциплины;
- систематизированное, глубокое и полное освоение навыков и компетенций по всем разделам программы дисциплины;

- использование научной терминологии, стилистически грамотное, логически правильное изложение ответа на вопросы, умение делать обоснованные выводы;
- умение эффективно использовать методику программы дисциплины в постановке и решении научных и профессиональных задач;
- способность самостоятельно решать поставленные задачи в нестандартных производственных ситуациях;
- усвоение основной и дополнительной литературы, рекомендованной программой дисциплины и преподавателем.

69-85 баллов:

- правильные ответы на часть вопросов контрольных работ;
- частичное выполнение лабораторных работ;
- подготовка реферата (в случае выполнения этого задания);
- подготовка доклада (в случае выполнения этого задания);
- участие в мероприятиях, предусмотренных программой дисциплины;
- систематизированное и полное освоение навыков и компетенций по всем разделам программы дисциплины;
- умение использовать методику программы дисциплины в постановке и решении научных и профессиональных задач;
- усвоение основной литературы, рекомендованной программой дисциплины.

51-68 баллов:

- неполные ответы на часть вопросов контрольных работ;
- частичное выполнение лабораторных работ;
- участие в мероприятиях, предусмотренных программой дисциплины;
- систематизированное и полное освоение навыков и компетенций по всем разделам программы дисциплины;
- удовлетворительное умение использовать методику программы дисциплины в постановке и решении научных и профессиональных задач;
- удовлетворительное усвоение основной литературы;

31 - 50 баллов – НЕ ЗАЧТЕНО:

- частично правильные ответы на часть вопросов контрольных работ;
- частичное выполнение лабораторных работ;
- недостаточно полный объем навыков и компетенции в рамках программы дисциплины;
- неумение использовать в практической деятельности научной терминологии, изложение ответа на вопросы с существенными стилистическими и логическими ошибками;
- слабое умение использовать методику программы дисциплины в постановке и решении научных и профессиональных задач;
- удовлетворительное усвоение основной литературы.

0-30 баллов, НЕ ЗАЧТЕНО:

- отсутствие умений, навыков, знаний и компетенции в рамках программы дисциплины;
- невыполнение лабораторных заданий; отказ от ответа по программе дисциплины;
- игнорирование занятий по дисциплине по неуважительной причине.