

МИНОБРНАУКИ РОССИИ

Ярославский государственный университет им. П.Г. Демидова

Кафедра вычислительных и программных систем

УТВЕРЖДАЮ

Декан факультета ИВТ

 Д.Ю. Чалый

«23» мая 2023 г.

Рабочая программа дисциплины

«Автоматический анализ текстов»

Направление подготовки

09.04.03 Прикладная информатика

Профиль

«Искусственный интеллект в корпоративных информационных системах»

Квалификация выпускника

Магистр

Форма обучения

очная

Программа рассмотрена
на заседании кафедры
от «21» апреля 2023 г.,
протокол № 8

Программа одобрена НМК
факультета ИВТ
протокол № 6 от
«28» апреля 2023 г.

Ярославль

1. Цели освоения дисциплины

Целями дисциплины «Автоматический анализ текстов» являются формирование у студентов системного представления о методах автоматической обработки естественно-языкового текста, ознакомлении студентов с теоретическими концепциями, практическими методами работы систем автоматической обработки естественно-языковых текстов. Дисциплина «Автоматический анализ текстов» является основой для дальнейшего изучения курсов, посвященных методам обработки естественного языка.

Для освоения дисциплины необходимо умение осваивать и использовать программные инструменты, знание основ статистики.

2. Место дисциплины в структуре ОП магистратуры

Дисциплина «Автоматический анализ текстов» относится к базовой части ОП магистратуры

Для освоения данной дисциплины студентам необходимы знания, полученные при изучении дисциплин «Основы программирования», «Высокоуровневое программирование», «Языки программирования».

Знания и навыки, полученные при изучении дисциплины «Автоматический анализ текстов», используются учащимися при изучении последующих дисциплин, таких как «Компьютерное моделирование», «Интеллектуальные системы», «Программная инженерия».

3. Планируемые результаты обучения по дисциплине, соотнесенные с планируемыми результатами освоения ОП магистратуры.

Процесс изучения дисциплины направлен на формирование следующих элементов компетенций в соответствии с ФГОС ВО, ОП ВО и приобретения следующих знаний, умений, навыков и (или) опыта деятельности:

Код компетенции	Формулировка компетенции	Перечень планируемых результатов обучения
Общепрофессиональные компетенции		
ПК-4. Способен адаптировать и применять методы и алгоритмы машинного обучения для решения прикладных задач в различных предметных областях	ПК-4.1. Ставит задачи по адаптации или совершенствованию методов и алгоритмов для решения комплекса задач предметной области	Знать: <ul style="list-style-type: none">- устройство, назначение и возможности применения технологий;- основные понятия ООАП; Уметь: <ul style="list-style-type: none">- разрабатывать объектно-ориентированные модели Владеть: <ul style="list-style-type: none">-- навыками работы с научно-технической литературой и технической документацией по программному обеспечению.

4. Объем структура и содержание дисциплины «Автоматический анализ текстов»

Общая трудоемкость дисциплины составляет 4 зачетные единицы, 144 часов.

№ п/п	Темы (разделы) дисциплины, их содержание	Семестр	Виды учебных занятий, включая самостоятельную работу студентов, и их трудоемкость (в академических часах)						Формы текущего контроля успеваемости Форма промежуточной аттестации (по семестрам)
			Контактная работа						
			лекции	практические	лабораторные	консультации	аттестационные испытания	самостоятельная работа	
1	Введение. Обработка естественного языка (Natural Language Processing, NLP). Области применения NLP. Методы NLP. Лингвистические ресурсы.	1	2	4				17	
2	Общая схема предварительной обработки текстов. Преобразование формата. Удаление шума. Выделение единиц текста. Вычисление признаков каждого токена. Маркировка.	1	2	4				18	
3	Морфологический анализ. Четкая морфология на основе словаря. Нечеткая морфология на основе системы правил. Обзор модулей морфологического анализа.	1	2	4				18	
4	Программные инструменты для автоматического анализа естественного языка. Stanza, spaCy. Обзор API. Примеры использования.	1	2	4				17	
5	Эксперименты и оценка качества решения задач. Сбор и разметка корпусов текстов. Обработка текста и расчет числовых параметров. Статистические характеристики для оценки результатов: точность, полнота, F-мера.	1	2	4				17	
6	Классификация текстов. Виды	1	2	4				20,7	

задач классификации. Модели текста. Методы и инструменты классификации.								
Всего за 1 семестр		12	24				107,7	зачет
Всего								

Содержание разделов дисциплины:

Введение. Обработка естественного языка (Natural Language Processing, NLP). Области применения NLP. Методы NLP. Лингвистические ресурсы.
Общая схема предварительной обработки текстов. Преобразование формата. Удаление шума. Выделение единиц текста. Вычисление признаков каждого токена. Маркировка.
Морфологический анализ. Четкая морфология на основе словаря. Нечеткая морфология на основе системы правил. Обзор модулей морфологического анализа.
Программные инструменты для автоматического анализа естественного языка. Stanza, spaCy. Обзор API. Примеры использования.
Эксперименты и оценка качества решения задач. Сбор и разметка корпусов текстов. Обработка текста и расчет числовых параметров. Статистические характеристики для оценки результатов: точность, полнота, F-мера.
Классификация текстов. Виды задач классификации. Модели текста. Методы и инструменты классификации.

5. Образовательные технологии, используемые при осуществлении образовательного процесса

В процессе обучения используются следующие образовательные технологии:

Вводная лекция – дает первое целостное представление о дисциплине и ориентирует студента в системе изучения данной дисциплины. Студенты знакомятся с назначением и задачами курса, его ролью и местом в системе учебных дисциплин и в системе подготовки в целом. Дается краткий обзор курса, история развития науки и практики, достижения в этой сфере, имена известных ученых, излагаются перспективные направления исследований. На этой лекции высказываются методические и организационные особенности работы в рамках данной дисциплины, а также дается анализ рекомендуемой учебно-методической литературы.

Академическая лекция (или лекция общего курса) – последовательное изложение материала, осуществляемое преимущественно в виде монолога преподавателя. Требования к академической лекции: современный научный уровень и насыщенная информативность, убедительная аргументация, доступная и понятная речь, четкая структура и логика, наличие ярких примеров, научных доказательств, обоснований, фактов.

Лабораторная работа – организация учебной работы с реальными материальными и информационными объектами, экспериментальная работа с аналоговыми моделями реальных объектов.

6. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень лицензионного программного обеспечения и информационных справочных систем (при необходимости)

В процессе осуществления образовательного процесса используются: для разработки документов, презентаций, для работы с электронными таблицами

OfficeStd 2013 RUS OLP NL Acdmc 021-10232

LibreOffice (свободное)

- для поиска учебной литературы библиотеки ЯрГУ – Автоматизированная библиотечная информационная система "БУКИ-NEXT" (АБИС "Буки-Next");
- для проведения лабораторных занятий используется CASE-среда Enterprise Architect (разработчик Sparx Systems).

7. Перечень основной и дополнительной учебной литературы, ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины

Основная литература

1. Проблемы компьютерной лингвистики и типологии : Сборник научных трудов / «ВОРОНЕЖСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ФАКУЛЬТЕТ РОМАНО-ГЕРМАНСКОЙ ФИЛОЛОГИИ НАУЧНО-МЕТОДИЧЕСКИЙ ЦЕНТР КОМПЬЮТЕРНОЙ ЛИНГВИСТИКИ. – Воронеж : Издательский дом ВГУ, 2017. – 246 с. URL: https://www.elibrary.ru/download/elibrary_35022773_12374093.pdf
2. Мамаев, И. Д. Русско-английский словарь основных терминов компьютерной лингвистики / И. Д. Мамаев // Лексикографическая копилка : Сборник научных статей / Под научной редакцией В.В. Гончаровой. – Санкт-Петербург : Санкт-Петербургский государственный экономический университет, 2020. – С. 83-92. URL: https://www.elibrary.ru/download/elibrary_43856875_79359109.pdf

б) дополнительная литература:

Пентус, А. Е. Математическая теория формальных языков / Пентус А. Е. , Пентус М. Р. - Москва : Национальный Открытый Университет "ИНТУИТ", 2016. (Основы информатики и математики) - ISBN 5-9556-0062-0. - Текст : электронный // ЭБС "Консультант студента" : [сайт]. - URL : <https://www.studentlibrary.ru/book/ISBN5955600620.html>

в) ресурсы сети «Интернет»

1. 1. Пакет анализа естественного языка Stanza: официальный сайт. URL: <https://stanfordnlp.github.io/stanza/>
2. Пакет анализа естественного языка SpaCy: официальный сайт. URL: <https://spacy.io/>

8. Материально-техническая база, необходимая для осуществления образовательного процесса по дисциплине

Материально-техническая база, необходимая для осуществления образовательного процесса по дисциплине включает в свой состав специальные помещения:

- учебные аудитории для проведения занятий лекционного типа и практических занятий (семинаров);
- учебные аудитории для проведения лабораторных занятий;
- учебные аудитории для проведения групповых и индивидуальных консультаций,
- учебные аудитории для проведения текущего контроля и промежуточной аттестации;
- помещения для самостоятельной работы;
- помещения для хранения и профилактического обслуживания технических средств обучения.

Специальные помещения укомплектованы средствами обучения, служащими для представления учебной информации большой аудитории.

Для проведения занятий лекционного типа предлагаются наборы демонстрационного

оборудования и учебно-наглядных пособий, хранящиеся на электронных носителях и

обеспечивающие тематические иллюстрации, соответствующие рабочим программам дисциплин.

Помещения для лабораторных занятий и самостоятельной работы обучающихся оснащены компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду организации.

Число посадочных мест в лекционной аудитории больше либо равно списочному составу потока, а в аудитории для практических занятий (семинаров) – списочному составу группы обучающихся.

**Приложение №1 к рабочей программе дисциплины
«Автоматический анализ текстов»**

**Фонд оценочных средств
для проведения текущей и промежуточной аттестации студентов
по дисциплине**

**1. Типовые контрольные задания или иные материалы,
необходимые для оценки знаний, умений, навыков и (или) опыта
деятельности,**

характеризующих этапы формирования компетенций

**1.1 Контрольные задания и иные материалы,
используемые в процессе текущей аттестации**

Задания для самостоятельной работы

Задания для выполнения рефератов

№	Тема	Описание темы	Источники
1	Статья I. Национальный корпус русского языка	Структура, состав, функциональные возможности корпуса. Примеры использования.	https://ruscorpora.ru/new/index.html
2	Тезаурусы WordNet и RussNet. Алгоритмы построения тезаурусов типа WordNet	Описание WordNet. и RussNet. Основные элементы: существительные, прилагательные, глаголы (с примерами). Описание RussNet, Основные элементы и связи. Алгоритмы автоматизации построения тезауруса: предварительная обработка корпуса текстов, выделение терминов, определение отношений.	http://nsu.ru/xmlui/bitstream/handle/nsu/9086/louk_book.pdf http://www.nsu.ru/xmlui/bitstream/handle/nsu/412/Text_BimenovaZB.pdf http://project.phil.spbu.ru/RussNet/index_ru.shtml https://wordnet.princeton.edu/
3	Алгоритмы вида «обучение с учителем» для выделения ключевых слов из текстов KEA и Maui	Шаги алгоритма. Расчет числовых характеристик слов-кандидатов. Методы выбора ключевых слов. Загрузка и использование KEA.	Официальный сайт KEA http://community.nzdl.org/kea/index.html Русскоязычная статья о KEA https://openbooks.itmo.ru/ru/file/6522/6522.pdf Статья об алгоритме Maui https://www.aclweb.org/anthology/D09-1137.pdf
4	Алгоритмы вида «обучение без учителем» для выделения ключевых слов из текстов TextRank и Topical PageRank	Шаги алгоритма. Расчет числовых характеристик слов-кандидатов. Методы выбора ключевых слов. Загрузка и использование TextRank.	TextRank https://github.com/summanlp/textrank Статья об алгоритме https://www.aclweb.org/anthology/W04-3252.pdf Статья об алгоритме Topical PageRank https://www.aclweb.org/anthology/D10-1036.pdf
5	Векторное представление слов	Алгоритмы word2vec. Описание и использование алгоритмов	Статья автора алгоритма word2vec https://papers.nips.cc/paper/5021-distributed-representations-of-words-and-phrases-and-their-compositionality.pdf Пример системы, реализующей алгоритм

			https://www.kaggle.com/pierremegret/gensim-word2vec-tutorial
6	Алгоритм синтаксического парсера для английского языка	Подробное описание алгоритма. Примеры применения алгоритма	https://explosion.ai/blog/parsing-english-in-python
7	Анализ тональности текстов	Методы анализа тональности. Словарные ресурсы для анализа тональности. Оценка результатов. Открытые программные системы для определения тональности. Примеры их использования	Статьи http://www.dialog-21.ru/media/1451/50.pdf vestnik.psu.ru/docs/2012/1/1/2012111.doc с http://www.isa.ru/aidt/images/documents/2014-01/25_33.pdf Учебник https://www.hse.ru/data/2017/07/22/1173852769/NLP_and_DA.pdf
8	Извлечение структурированных данных из текста с помощью Томита-парсер	Методы извлечения информации. Извлечение именованных сущностей. Извлечение атрибутов понятий. Извлечение фактов. Применение Томита-парсера	Томита-парсер https://yandex.ru/dev/tomita/?turbo=true Учебник https://www.hse.ru/data/2017/07/22/1173852769/NLP_and_DA.pdf
9	Классификация полнотекстовых документов. Алгоритмы классификации с учителем	Представление данных в задачах классификации текстов. Меры сходства и различий между образами документов. Отбор терминов для классификации. Признаки документов. Оценка результатов. Пример одного из алгоритмов классификации с учителем. Описание алгоритма, пример применения.	Учебник http://window.edu.ru/resource/465/78465/files/miem_lingvistika.pdf Статья на русском языке https://cyberleninka.ru/article/n/metody-avtomaticheskoy-klassifikatsii-tekstov
10	Классификация полнотекстовых документов. Алгоритмы классификации без учителя	Понятие кластеризации. Виды алгоритмов кластеризации: k-средних, иерархические, эвристические. Оценка результатов. Пример одного из алгоритмов классификации без учителя. Описание алгоритма, пример применения.	Учебник http://window.edu.ru/resource/465/78465/files/miem_lingvistika.pdf Статья на русском языке http://www.dialog-21.ru/digest/2001/articles/kirichenko/
11	Методы автоматической рубрикации.	Методы автоматической рубрикации. Оценка качества рубрикации. Пример системы рубрикации текста. Алгоритм работы системы, оценка результатов работы. Корпус текстов, который может использоваться для оценки работы системы.	https://cyberleninka.ru/article/n/avtomaticheskaya-rubrikatsiya-tekstov-metody-i-problemy
12	Машинный перевод	Лингвистические стратегии машинного перевода и поколения систем машинного перевода. Автоматический перевод, основанный на правилах. Оценки качества машинного перевода. Статистический машинный перевод. Программные инструменты, предоставляющие API для машинного перевода	http://bwbooks.net/index.php?id1=4&category=lingvistika&author=leontevann&book=2006
13	Автоматическая проверка орфографии с помощью	Методы проверки орфографии. API Яндекс Спеллер. Демонстрация	Яндекс // Спеллер // URL: https://tech.yandex.ru/speller/

	Яндекс Спеллер	работы	Пикалёв Я. С., Вовнянко А. С., Денищенко И. Я. АНАЛИЗ АВТОМАТИЧЕСКИХ СИСТЕМ ПРОВЕРКИ ПРАВОПИСАНИЯ РУССКОГО ЯЗЫКА //Проблемы искусственного интеллекта. – 2018. – №. 2 (9). https://cyberleninka.ru/article/n/analiz-avtomaticheskikh-sistem-proverki-pravopisaniya-russkogo-yazyka
14	Автоматическая проверка орфографии с помощью LanguageTool	Методы проверки орфографии. API LanguageTool. Демонстрация работы	LanguageTool // LanguageTool ПО для проверки грамматики и орфографии // URL: https://languagetool.org/ru/ Пикалёв Я. С., Вовнянко А. С., Денищенко И. Я. АНАЛИЗ АВТОМАТИЧЕСКИХ СИСТЕМ ПРОВЕРКИ ПРАВОПИСАНИЯ РУССКОГО ЯЗЫКА //Проблемы искусственного интеллекта. – 2018. – №. 2 (9). https://cyberleninka.ru/article/n/analiz-avtomaticheskikh-sistem-proverki-pravopisaniya-russkogo-yazyka

Требования к оформлению и защите рефератов.

Реферат оформляется в электронном виде как файл формата pdf, защита осуществляется в виде доклада, сопровождаемого презентацией. Правила оформления реферата <https://kursach37.com/oformlenie-referata-pogosti/>.

Критерии оценки

«Отлично» – оцениваются рефераты, содержание которых основано на глубоком и всестороннем знании темы, изученной литературы, изложено логично, аргументировано и в полном объеме. Основные понятия, выводы и обобщения сформулированы убедительно и доказательно.

«Хорошо» – оцениваются рефераты, основанные на твердом знании исследуемой темы. Возможны недостатки в систематизации или в обобщении материала, неточности в выводах. Студент твердо знает основные категории, умело применяет их для изложения материала.

«Удовлетворительно» – оцениваются рефераты, которые базируются на знании основ предмета, но имеются значительные пробелы в изложении материала, затруднения в его изложении и систематизации, выводы слабо аргументированы, в содержании допущены теоретические ошибки.

«Неудовлетворительно» – оцениваются рефераты, в которых обнаружено неверное изложение основных вопросов темы, обобщений и выводов нет. Текст реферата целиком или в значительной части дословно переписан из первоисточника без ссылок на него.

Пример задания для выполнения лабораторных работ

Установите одну из библиотек по обработке текста поддерживающую работу с русским языком. Создайте файл, содержащий текст рассказа или статьи. Например, выберите рассказ А. П. Чехова из электронного ресурса (<http://chehov-lit.ru/chehov/text/rasskazy.htm>). Решите следующие задачи:

- а) определите количество слов текста;
- б) определите количество предложений в тексте;
- в) определите количество знаков пунктуации в тексте;

- г) определите долю существительных и долю глаголов относительно всех слов текста; д) введите слово и выведите список предложений, в которых оно встречается;
- е) определите количество различных слов в тексте;
- ж) выведите предложения, которые начинаются и заканчиваются на одно и то же слово;
- з) выведите пары предложений, которые начинаются одним и тем же словом или словосочетанием.

Критерии оценки

«Отлично» – решены все задачи.

«Хорошо» – решено 6-7 задач.

«Удовлетворительно» – решено 4-5 задач

«Неудовлетворительно» – решено менее 4 задач.

Вопросы к зачету

1. Перечислите области применения автоматической обработки естественного языка.
2. Сформулируйте задачи информационного поиска.
3. Сформулируйте постановку задачи реферирования и аннотирования текста.
4. Опишите задачи по обработке текста в аналитических системах.
5. Сформулируйте постановку задачи автоматического редактирования текста.
6. Опишите задачи по обработке текста в вопросно-ответных и диалоговых системах.
7. Опишите задачи по обработке текста в обучении естественному языку.
8. Сформулируйте постановку задачи автоматической генерации текста
9. Опишите задачи по обработке текста в распознавании и синтезе звучащей речи.
10. Назовите виды лингвистических ресурсов.
11. Опишите общую схему предварительной обработки текстов.
12. Опишите особенности преобразование формата сырого текста.
13. Опишите особенности удаления шума из текста.
14. Перечислите единицы текста и особенности их выделения.
15. Перечислите группы признаков и отдельные признаки токенов.
16. Дайте определение морфологического анализа.
17. Опишите метод четкой морфологии на основе словаря.
18. Опишите метод нечеткой морфологии на основе системы правил.
19. Назовите модули морфологического анализа.
20. Дайте определение стемминга и лемматизации.
21. Опишите алгоритм работы стеммера.
22. Опишите алгоритм работы лемматизатора.
23. Назовите синтаксические маркеры слов и токенов предложений.
24. Назовите программные инструменты для автоматического анализа естественного языка.
25. Опишите пайплайн обработки текста и приведите пример его организации в одной из программных библиотек.
26. Приведите пример токенизации текста.
27. Приведите пример определения морфологических характеристик текста.
28. Приведите пример определения леммы слов предложения.
29. Приведите пример определения роли токенов в предложении.
30. Назовите алгоритмы семантического анализа, используемые в программном инструменте.
31. Опишите и приведите пример определения именованных сущностей.
32. Опишите и приведите пример определения тональности слова.
33. Опишите процесс проведения экспериментов в области автоматического анализа текстов.
34. Сформулируйте особенности сбора и разметки корпуса текстов.
35. Опишите деление результатов эксперимента на группы для расчета метрик качества.
36. Дайте определение метрикам качества: точность, полнота, F-мера.

Критерии оценки

«Отлично» – ответ на вопросы показывает всестороннее знание темы, изученной литературы, изложен логично, аргументировано и в полном объеме. Основные понятия, выводы и обобщения сформулированы убедительно и доказательно.

Продемонстрированы полные и глубокие навыки практического применения программного инструмента для обработки текста.

«Хорошо» – ответ на вопросы основан на твердом знании темы. Возможны недостатки в систематизации или в обобщении материала, неточности в выводах. Продемонстрированы хорошие навыки практического применения программного инструмента для обработки текста.

«Удовлетворительно» – ответ на вопросы базируется на знании основ предмета, но имеются значительные пробелы в изложении материала, затруднения в его изложении и систематизации, выводы слабо аргументированы, в содержании допущены теоретические ошибки. Продемонстрированы элементарные навыки практического применения программного инструмента для обработки текста для решения простых задач.

«Неудовлетворительно» – оценивается ответ на вопросы, в котором обнаружено неверное изложение темы, систематизации знаний, обобщений и выводов нет. Навыки практического применения программного инструмента для обработки текста слабые и отрывочные или отсутствуют.

2. Перечень компетенций, этапы их формирования, описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкалы оценивания

2.1 Шкала оценивания сформированности компетенций и ее описание

Оценивание уровня сформированности компетенций в процессе освоения дисциплины осуществляется по следующей трехуровневой шкале:

Пороговый уровень - предполагает отражение тех ожидаемых результатов, которые определяют минимальный набор знаний и (или) умений и (или) навыков, полученных студентом в результате освоения дисциплины. Пороговый уровень является обязательным уровнем для студента к моменту завершения им освоения данной дисциплины.

Продвинутый уровень - предполагает способность студента использовать знания, умения, навыки и (или) опыт деятельности, полученные при освоении дисциплины, для решения профессиональных задач. Продвинутый уровень превосходит пороговый уровень по нескольким существенным признакам.

Высокий уровень - предполагает способность студента использовать потенциал интегрированных знаний, умений, навыков и (или) опыта деятельности, полученных при освоении дисциплины, для творческого решения профессиональных задач и самостоятельного поиска новых подходов в их решении путем комбинирования и использования известных способов решения применительно к конкретным условиям. Высокий уровень превосходит пороговый уровень по всем существенным признакам.

3. Методические рекомендации преподавателю по процедуре оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций

Целью процедуры оценивания является определение степени овладения студентом ожидаемыми результатами обучения (знаниями, умениями, навыками и (или) опытом деятельности).

Процедура оценивания степени овладения студентом ожидаемыми результатами обучения осуществляется с помощью методических материалов, представленных в разделе «Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций»

3.1 Критерии оценивания степени овладения знаниями, умениями, навыками и (или) опытом деятельности, определяющие уровни сформированности компетенций

Пороговый уровень (общие характеристики):

- владение основным объемом знаний по программе дисциплины;
- знание основной терминологии данной области знаний, стилистически грамотное, логически правильное изложение ответа на вопросы без существенных ошибок;
- владение инструментарием дисциплины, умение его использовать в решении стандартных (типовых) задач;
- способность самостоятельно применять типовые решения в рамках рабочей программы дисциплины;
- усвоение основной литературы, рекомендованной рабочей программой дисциплины;
- знание базовых теорий, концепций и направлений по изучаемой дисциплине;
- самостоятельная работа на практических и лабораторных занятиях, периодическое участие в групповых обсуждениях, достаточный уровень культуры исполнения заданий.

Продвинутый уровень (общие характеристики):

- достаточно полные и систематизированные знания в объеме программы дисциплины;
- использование основной терминологии данной области знаний, стилистически грамотное, логически правильное изложение ответа на вопросы, умение делать выводы;
- владение инструментарием дисциплины, умение его использовать в решении учебных и профессиональных задач;
- способность самостоятельно решать сложные задачи (проблемы) в рамках рабочей программы дисциплины;
- усвоение основной и дополнительной литературы, рекомендованной рабочей программой дисциплины;
- умение ориентироваться в базовых теориях, концепциях и направлениях по изучаемой дисциплине и давать им сравнительную оценку;
- самостоятельная работа на практических и лабораторных занятиях, участие в групповых обсуждениях, высокий уровень культуры исполнения заданий.

Высокий уровень (общие характеристики):

- систематизированные, глубокие и полные знания по всем разделам дисциплины;
- точное использование терминологии данной области знаний, стилистически грамотное, логически правильное изложение ответа на вопросы, умение делать обоснованные выводы;
- безупречное владение инструментарием дисциплины, умение его использовать в постановке и решении научных и профессиональных задач;

- способность самостоятельно и творчески решать сложные задачи (проблемы) в рамках рабочей программы дисциплины;
- полное и глубокое усвоение основной и дополнительной литературы, рекомендованной рабочей программой дисциплины;
- умение ориентироваться в основных теориях, концепциях и направлениях по изучаемой дисциплине и давать им критическую оценку;
- активная самостоятельная работа на практических и лабораторных занятиях, творческое участие в групповых обсуждениях, высокий уровень культуры исполнения заданий.

3.2 Описание процедуры выставления оценки

В зависимости от уровня сформированности каждой компетенции по окончании освоения дисциплины студенту выставляется оценка «зачтено» или «не зачтено».

Показатели и критерии, используемые при выставлении оценки подробно описаны в разделе «Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций».

Пороговый уровень формирования компетенций соответствует оценке «зачтено» за выполнение лабораторных работ в течение семестра и сдачу зачета в конце семестра.

Оценка «зачет» выставляется студенту, у которого каждая компетенция (полностью или частично формируемая данной дисциплиной) сформирована не ниже, чем на пороговом уровне.

Оценка «незачтено» выставляется студенту, у которого хотя бы одна компетенция (полностью или частично формируемая данной дисциплиной) сформирована ниже, чем на пороговом уровне.

Приложение №2 к рабочей программе дисциплины «Автоматический анализ текстов»

Методические указания для студентов по освоению дисциплины

Основной формой изложения учебного материала по дисциплине «Автоматический анализ текстов» являются лекции, проводимые в виде электронных презентаций, что позволяет сделать материал лекций более наглядными, улучшает информативность и понимание изучаемого курса.

По большинству тем предусмотрены лабораторные работы, на которых происходит закрепление лекционного материала путем применения его к конкретным задачам и получение навыков работы непосредственно с CASE- системой, путем выполнения заданий по построению моделей определенного типа. Выполнение предлагаемых в процессе изучения курса лабораторных работ позволяет не только понять и закрепить теоретический материал, но и приобрести навык анализа предметной области и построения объектно-ориентированных моделей с применением современных технологий.

Для успешного освоения дисциплины очень важно решение достаточно большого количества задач, как в аудитории, так и самостоятельно в качестве домашних заданий. Примеры решения задач разбираются и обсуждаются на лекциях и лабораторных занятиях. Основная цель решения задач – помочь усвоить фундаментальные понятия и основы концепции объектно-ориентированного анализа и проектирования. Для решения всех задач необходимо знать и понимать лекционный материал. Поэтому в процессе изучения дисциплины рекомендуется регулярное повторение пройденного лекционного материала. Материал, законспектированный на лекциях, необходимо дома еще раз прорабатывать и при необходимости дополнять информацией, полученной при выполнении лабораторных работ или из учебной литературы.

Большое внимание должно быть уделено выполнению домашней работы. В качестве заданий для самостоятельной работы дома студентам предлагаются задачи, аналогичные разобранным на лекциях и лабораторных занятиях или немного более сложные, которые являются результатом объединения нескольких базовых задач.

Для проверки и контроля усвоения теоретического материала, приобретенных практических навыков работы с CASE - системой, в течение обучения проводятся мероприятия текущей аттестации в виде двух контрольных работ. Для более успешной подготовки к контрольным работам перед их проведением преподавателем проводятся консультации.

В конце семестра студенты сдают зачет. Зачет по итогам семестра выставляется по результатам написания лабораторных и контрольных работ, а так же краткого собеседования по вопросам теоретического материала.

Освоить вопросы, излагаемые в процессе изучения дисциплины «Автоматический анализ текстов» самостоятельно студенту достаточно сложно. Это связано с отсутствием опыта в проведении анализа предметной области и построения объектно-ориентированных моделей. Поэтому посещение всех аудиторных занятий является совершенно необходимым. Без упорных и регулярных занятий в течение семестра сдать зачет по итогам изучения дисциплины студенту будет сложно.

Учебно-методическое обеспечение самостоятельной работы студентов по дисциплине

Для самостоятельной работы особенно рекомендуется использовать учебную литературу, указанную в разделе № 7 данной рабочей программы.

Также для подбора учебной литературы рекомендуется использовать широкий спектр интернет-ресурсов:

1. Электронно-библиотечная система «Университетская библиотека online»

(www.biblioclub.ru) - электронная библиотека, обеспечивающая доступ к наиболее востребованным материалам-первоисточникам, учебной, научной и художественной литературе ведущих издательств (*регистрация в электронной библиотеке – только в сети университета. После регистрации работа с системой возможна с любой точки доступа в Internet.).

2. Для самостоятельного подбора литературы в библиотеке ЯрГУ рекомендуется использовать:

1. Личный кабинет (http://lib.uniyar.ac.ru/opac/bk_login.php) дает возможность получения online доступа к списку выданной в автоматизированном режиме литературы, просмотра и копирования электронных версий изданий сотрудников университета (учеб. и метод. пособия, тексты лекций и т.д.) Для работы в «Личном кабинете» необходимо зайти на сайт Научной библиотеки ЯрГУ с любой точки, имеющей доступ в Internet, в пункт меню «Электронный каталог»; пройти процедуру авторизации, выбрав вкладку «Авторизация», и заполнить представленные поля информации.

2. Электронная библиотека учебных материалов ЯрГУ

(http://www.lib.uniyar.ac.ru/opac/bk_cat_find.php) содержит более 2500 полных текстов учебных и учебно-методических материалов по основным изучаемым дисциплинам, изданных в университете. Доступ в сети университета, либо по логину/паролю.

3. Электронная картотека «[Книгообеспеченность](#)»

(http://www.lib.uniyar.ac.ru/opac/bk_bookreq_find.php) раскрывает учебный фонд научной библиотеки ЯрГУ, предоставляет оперативную информацию о состоянии книгообеспеченности дисциплин основной и дополнительной литературой, а также цикла дисциплин и специальностей. Электронная картотека «[Книгообеспеченность](#)» доступна в сети университета и через Личный кабинет.