

Министерство науки и высшего образования Российской Федерации
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ (НИ ТГУ)

Институт прикладной математики и компьютерных наук

УТВЕРЖДЕНО:

Директор

А. В. Замятин

Рабочая программа дисциплины

Постреляционные модели данных

по направлению подготовки

01.04.02 Прикладная математика и информатика

Направленность (профиль) подготовки:

Интеллектуальный анализ больших данных

Форма обучения

Очная

Квалификация

Магистр

Год приема

2024

СОГЛАСОВАНО:

Руководитель ОП

А.В. Замятин

Председатель УМК

С.П. Сущенко

1. Цель и планируемые результаты освоения дисциплины

Целью освоения дисциплины является формирование следующих компетенций:

– ПК-4 – способность определять проблемную ситуацию, ставить задачи анализа данных в области профессиональной деятельности, выбирать методы и использовать специализированные знания для её решения.

Результатами освоения дисциплины являются следующие индикаторы достижения компетенций:

ИПК-4.1 Определяет проблемную ситуацию и ставит задачу анализа данных в области профессиональной деятельности.

ИПК-4.2 Собирает и обрабатывает данные в области профессиональной деятельности, знает и использует специфику таких данных.

ИПК-4.3 Выбирает методы анализа данных в области профессиональной деятельности.

ИПК-4.4 Разрабатывает концепцию и техническое задание на решение задачи в области профессиональной деятельности на основе анализа данных.

2. Задачи освоения дисциплины

– Ознакомить с современными структурами и инструментами хранения данных, а также научить правильно выбирать базу данных в зависимости от поставленных задач и эффективно работать с ней с использованием всех ее преимуществ.

3. Место дисциплины в структуре образовательной программы

Дисциплина относится к части образовательной программы, формируемой участниками образовательных отношений, предлагается обучающимся на выбор. Дисциплина входит в модуль «Специализация».

4. Семестр(ы) освоения и форма(ы) промежуточной аттестации по дисциплине

Второй семестр, экзамен

5. Входные требования для освоения дисциплины

Для успешного освоения дисциплины требуются результаты обучения по следующим дисциплинам: «Математические методы и модели для компьютерных наук», «Алгоритмы и структуры данных», знание Python .

6. Язык реализации

Русский

7. Объем дисциплины

Общая трудоемкость дисциплины составляет 3 з.е., 108 часов, из которых:

-лекции: 16 ч.

-лабораторные: 16 ч.

в том числе практическая подготовка: 0 ч.

Объем самостоятельной работы студента определен учебным планом.

8. Содержание дисциплины, структурированное по темам

Раздел 1. История развития и становления концепции баз данных.

Реляционные базы данных.

- 1.1. Историческое развитие систем хранения данных.
- 1.2. Основные понятия баз данных, структур данных и систем управления базами данных.
- 1.3. Реляционная модель. Основные понятия и термины реляционной модели.
Реляционные СУБД.
- 1.4. Понятие нормальной формы.
- 1.5. SQL - стандартный язык запросов к реляционным СУБД. Оптимизация SQL запросов.
- 1.6. Понятие целостности данных. Обработка транзакций.
- 1.7. Недостатки реляционной модели данных.

Раздел 2. Объектно-ориентированные СУБД. Постреляционные базы данных (NoSQL).

- 2.1. Объектно-ориентированные СУБД.
- 2.2. Постреляционные базы данных (NoSQL). Теорема CAP.
- 2.3. Основные виды баз данных NoSQL.
- 2.4. Преимущества и недостатки баз данных NoSQL. Область применения.
- 2.5. Основные принципы, лежащие в основе темпоральных моделях данных. Понятие времени в темпоральных моделях данных. Модели, используемые в темпоральных баз данных (TRM, HDM)
- 2.6. Темпоральность в реляционной СУБД.
- 2.7. Многомерное моделирование данных.

Раздел 3. Документно-ориентированные базы данных. Графовые базы данных.

- 3.1. Документно-ориентированная модель данных. Методы работы с данными. Структуры хранения данных.
- 3.2. База данных MongoDB.
- 3.3. Elasticsearch.
- 3.4. Графовая модель данных. Методы работы с данными. Структуры хранения данных.
- 3.5. База данных Neo4J.

Раздел 4. Базы данных «ключ-значение». Колоночные базы данных.

- 4.1. Модель данных «ключ-значение». Методы работы с данными. Структуры хранения данных.
- 4.2. База данных Redis.
- 4.3. Колоночные БД. Модель данных. Методы работы с данными. Структуры хранения данных.
- 4.4. База данных Cassandra.

9. Текущий контроль по дисциплине

Текущий контроль по дисциплине проводится путем проведения лабораторных занятий и защиты отчета по лабораторной работе, а также путем защиты индивидуальных проектов.

Типовое задание для лабораторной работы.

1. Скачать и установить требуемое программное обеспечение базу данных.
2. Создать базу данных.
3. Отработать стандартные CRUD операции базе данных.
4. Отработать операции выборки данных и агрегации.
5. Осуществить импорт и экспорт данных.
6. Подключиться к базе данных из среды Python с использованием драйвера соответствующей базы данных.

Индивидуальное задание представляет собой ознакомление с одной noSQL базой данных, относящихся к одному из основных типов и предполагает следующие действия:

1. Скачать и установить требуемое программное обеспечение базу данных.

2. Создать базу данных.
3. Отработать стандартные CRUD операции базе данных.
4. Отработать операции выборки данных и агрегации.
5. Подготовку отчета и презентации
6. Публичная защита работы в конце семестра.

Базу данных студент выбирает самостоятельно и согласовывает с преподавателем.

10. Порядок проведения и критерии оценивания промежуточной аттестации

Экзамен во втором семестре проводится в письменной форме по билетам. Билет для экзамена состоит из двух теоретических и одного практического вопроса.

Типовые задания для проведения промежуточной аттестации по дисциплине

Билет №1

Вопросы:

1. Сравните реляционные и постреляционные базы данных. Опишите сферы применения постреляционных баз данных.
2. Назовите основные типы и структуры данных в DB Redis.
3. Создайте коллекцию в MongoDB. Импортируйте данные из внешнего набора данных.(располагаемого по ссылке..)

Билет №2

Вопросы:

1. Основные положения документ-ориентированных баз данных Достоинства и недостатки Mongo DB..
2. Назовите основные типы постреляционных баз данных.
3. Создайте базу данных в Neo4j. Импортируйте данные из внешнего набора данных.(располагаемого по ссылке.) Свяжите узлы.

11. Учебно-методическое обеспечение

- а) Электронный учебный курс по дисциплине в электронном университете «Moodle»
- б) Оценочные материалы текущего контроля и промежуточной аттестации по дисциплине.

12. Перечень учебной литературы и ресурсов сети Интернет

- а) основная литература:
 - Садаладж, Фаулер. NoSQL. Методология разработки нереляционных баз данных. – Диалектика, 2020.
 - Джим Р. Уилсон, Эрик Редмонд. Семь баз данных за семь недель. Введение в современные базы данных и идеологию NoSQL. – ДМК Пресс, 2013.
 - Карпентер, Хьюитт. Cassandra. Полное руководство. – ДМК Пресс, 2017.
 - Кайл Бэнкер. MongoDB в действии. – ДМК Пресс, 2012.
 - Шукла, Кумар. Elasticsearch, Kibana, Logstash и поисковые системы нового поколения. – Питер, 2019.

13. Перечень информационных технологий

- а) лицензионное и свободно распространяемое программное обеспечение:
 - Oracle DataModeler
 - MongoDB Compass
 - Neo4j Desktop

б) информационные справочные системы:

- <https://www.mongodb.com/>
- <https://www.redis.com/>
- <http://cassandra.apache.org/>
- <https://neo4j.com/>
- <https://www.elastic.co/>
- <http://www.oracle.com/>

14. Материально-техническое обеспечение

Аудитории для проведения занятий лекционного типа.

Аудитории для проведения занятий семинарского типа, индивидуальных и групповых консультаций, текущего контроля и промежуточной аттестации.

Помещения для самостоятельной работы, оснащенные компьютерной техникой и доступом к сети Интернет, в электронную информационно-образовательную среду и к информационным справочным системам.

15. Информация о разработчиках

Мокина Елена Евгеньевна, старший преподаватель кафедры теоретических основ информатики.