**Санкт-Петербургский государственный университет**

**Р А Б О Ч А Я П Р О Г Р А М М А**

**УЧЕБНОЙ ДИСЦИПЛИНЫ**

Обработка и оптимизация запросов в базах данных

Processing and Optimization of Database Queries

**Язык(и) обучения**

русский

Трудоемкость в зачетных единицах: 4

Регистрационный номер рабочей программы: 027274

Санкт-Петербург

2020

**Раздел 1. Характеристики учебных занятий**

**1.1. Цели и задачи учебных занятий**

Сформировать у слушателей навык о написания сложных запросов, процедур и функций в реляционных СУБД, применение методов индексирования для организации быстрого доступа к данным.

**1.2. Требования подготовленности обучающегося к освоению содержания учебных занятий (пререквизиты)**

Знание основ информатики, программирования, курса «Базы данных и СУБД» в пределах бакалаврской подготовки.

**1.3. Перечень результатов обучения (learning outcomes)**

В процессе изучения дисциплины «Обработка и оптимизация запросов в базах данных» обучаемые приобретают следующие

***знания***

* основные методы получения, хранения и переработки информации;
* основные принципы построения индексов;
* применение различных видов индексирования.

***умения***

* использовать средства современных СУБД для организации хранения различных видов информации;
* строить и анализировать план запроса;
* анализировать необходимость применения различных методов индексирования для ускорения доступа к данным;
* создавать и реализовывать новые схемы индексирования и применять их к данным;
* делать анализ и грамотную оценку эффективности применения индексов.

***навыки***

* применения средств современных СУБД для эффективной работы с данными;
* осуществления целенаправленного поиска информации о технологических достижениях в сети Интернет и из других источников.

Знать содержание дисциплины «Обработка и оптимизация запросов в базах данных», в частности, знать различные способы представления данных средствами современных СУБД, применять современные технологии для организации доступа к данным, иметь возможность организовывать новые эффективные структуры для ускорения доступа к данным.

Дисциплина развивает следующие компетенции:

* ПКА-1 – способен находить, формулировать и решать актуальные проблемы фундаментальной и прикладной информатики и информационных технологий;
* ПКА-2 – способен проектировать, разрабатывать и внедрять программные продукты и программные комплексы различного назначения;
* ПКП-1 – способен демонстрировать фундаментальные знания математических и естественных наук, программирования и информационных технологий;
* ПКП-8 – способен использовать основные концептуальные положения функционального, логического, объектно-ориентированного и визуального направлений программирования, методы, способы и средства разработки программ в рамках этих направлений.

**1.4. Перечень и объём активных и интерактивных форм учебных занятий**

В качестве основных интерактивных форм (общее количество 20 часов) предполагается

* проведение лекционных занятий (10 часов), на которых обучающиеся активно участвуют в обсуждении;
* проведение практических занятий (10 часов), на которых обучающиеся будут применять различные методы написания и оптимизации запросов к данным в различных реляционных СУБД.

**Раздел 2. Организация, структура и содержание учебных занятий**

**2.1. Организация учебных занятий**

**2.1.1 Основной курс**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Трудоёмкость, объёмы учебной работы и наполняемость групп обучающихся | | | | | | | | | | | | | | | | | | |
| Код модуля в составе дисциплины,  практики и т.п. | Контактная работа обучающихся с преподавателем | | | | | | | | | | | | Самостоятельная работа | | | | Объём активных и интерактивных  форм учебных занятий | Трудоёмкость |
| лекции | семинары | консультации | практические  занятия | лабораторные работы | контрольные работы | коллоквиумы | текущий контроль | промежуточная  аттестация | итоговая аттестация | под руководством преподавателя | в присутствии  преподавателя | сам. раб. с использованием  методических материалов | текущий контроль (сам.раб.) | промежуточная аттестация (сам.раб.) | итоговая аттестация  (сам.раб.) |
| ОСНОВНАЯ ТРАЕКТОРИЯ | | | | | | | | | | | | | | | | | | |
| Форма обучения: очная | | | | | | | | | | | | | | | | | | |
| Семестр 3 | 16 | 16 | 2 |  |  |  |  |  | 2 |  |  |  | 72 |  | 36 |  | 20 | 4 |
|  | 2-25 | 2-25 | 2-25 |  |  |  |  |  | 1-25 |  |  |  | 1-1 |  | 1-1 |  |  |  |
| ИТОГО | 16 | 16 | 2 |  |  |  |  |  | 2 |  |  |  | 72 |  | 36 |  |  | 4 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Виды, формы и сроки текущего контроля успеваемости и промежуточной аттестации | | | | | | |
| Код модуля в составе дисциплины, практики и т.п. | Формы текущего контроля успеваемости | | Виды промежуточной аттестации | | Виды итоговой аттестации  (только для программ итоговой аттестации и дополнительных образовательных программ) | |
| Формы | Сроки | Виды | Сроки | Виды | Сроки |
| ОСНОВНАЯ ТРАЕКТОРИЯ | | | | | | |
| Форма обучения: очная | | | | | | |
| Семестр 3 |  |  | экзамен, устно, традиционная форма | по графику промежуточной аттестации |  |  |

**2.2. Структура и содержание учебных занятий**

Применение различных типов данных для хранения информации на основе СУБД

Организация хранения при помощи средств современных СУБД иерархически организованных данных, географических данных, текстовых данных большого объема.

1. Применение одномерных способов индексирования

Построение индексов на основе B-деревьев, битовых шкал, хеширования. Сравнение эффективности рассматриваемых технологий.

1. Применение пространственных способов индексирования

Построение пространственных индексов на основе R-дерева и Spatial-grid.

1. Способы многомерного индексирования данных

Построение структур для многоключевого поиска информации. Метрические индексы.

1. Применение полнотекстового индексирования

Построение и настройка полнотекстовых индексов, ведение словаря для поиска текстовой информации.

1. Изучение статистик объектов баз данных.
2. Построение плана запроса, определение «узких» мест. VIII. Императивный и декларативный подход к написанию запросов. Использование подсказок для оптимизатора запросов.
3. Применение процедур и функций в реляционных СУБД.
4. Модели оценки стоимости запросов в реляционной СУБД.

**Раздел 3. Обеспечение учебных занятий**

**3.1. Методическое обеспечение**

**3.1.1 Методические указания по освоению дисциплины**

Успешное освоение дисциплины возможно благодаря посещению лекций, участию в обсуждении рассматриваемых вопросов, самостоятельной работе, включающей в себя чтение специальной литературы по разделам темы, выполнению докладов на заданную тему, применению изученных навыков на практических занятиях. В течение семестра обучающимся предлагаются для выполнения практические задания по работе с базой данных. Предлагаются различные наборы данных (4 набора), для которых необходимо разработать процедуры/функции заданного функционала, построить индексные структуры в зависимости от характера данных, описать зависимость производительности от наличия индексных структур.

**3.1.2 Методическое обеспечение самостоятельной работы**

Самостоятельная работа обучающихся в рамках данной дисциплины является важным компонентом обучения, предусмотренным компетентностно-ориентированным учебным планом и рабочей программой учебной дисциплины.

Настоящей программой предусмотрены формы самостоятельной работы с использованием методических материалов по тематике курса и источников, указанных в обязательной, дополнительной литературе и интернет-источниках, указанных с данной программе.

**3.1.3 Методика проведения текущего контроля успеваемости и промежуточной аттестации и критерии оценивания**

Экзамен проводится в форме представления обучающимся реализаций практических заданий по четырем наборам данных и отчетов о проделанной работе.

За каждое практическое задание обучающийся может получить максимум 25 баллов. Оценка может быть снижена за не полностью описанные возможности индексов в разных СУБД, отсутствие сравнения производительности. Баллы за все домашние задания суммируются, получившийся балл используется для выставления оценки по следующей шкале:

|  |  |  |
| --- | --- | --- |
| Итоговый балл | Оценка СПбГУ при  проведении экзамена | Оценка  ECTS |
| 90-100 | отлично | A |
| 86-89 | хорошо | B |
| 80-85 | хорошо | C |
| 76-79 | удовлетворительно | D |
| 70-75 | удовлетворительно | E |
| менее 70 | неудовлетворительно | F |

**3.1.4 Методические материалы для проведения текущего контроля успеваемости и промежуточной аттестации (контрольно-измерительные материалы, оценочные средства)**

Задания на темы курса:

1. Применение одномерных способов индексирования.

Построение индексов на основе B-деревьев, битовых шкал, хеширования. Сравнение эффективности рассматриваемых технологий.

1. Применение пространственных способов индексирования

Построение пространственных индексов на основе R-дерева и Spatial-grid.

1. Способы многомерного индексирования данных

Построение структур для многоключевого поиска информации. Метрические индексы.

1. Применение полнотекстового индексирования

Построение и настройка полнотекстовых индексов, ведение словаря для поиска текстовой информации.

1. Изучение статистик объектов баз данных.
2. Построение плана запроса, определение «узких» мест. VIII. Императивный и декларативный подход к написанию запросов. Использование подсказок для оптимизатора запросов.
3. Применение процедур и функций в реляционных СУБД.
4. Модели оценки стоимости запросов в реляционной СУБД.

**3.1.5 Методические материалы для оценки обучающимися содержания и качества учебного процесса**

Оценка обучающимися содержания и качества учебного процесса по дисциплине осуществляется в установленном в СПбГУ порядке

**3.2. Кадровое обеспечение**

**3.2.1 Образование и (или) квалификация штатных преподавателей и иных лиц, допущенных к проведению учебных занятий**

К преподаванию дисциплины могут быть допущены преподаватели, имеющие диплом о высшем образовании по соответствующему направлению.

**3.2.2 Обеспечение учебно-вспомогательным и (или) иным персоналом**

Специальных требований нет.

**3.3. Материально-техническое обеспечение**

**3.3.1 Характеристики аудиторий (помещений, мест) для проведения занятий**

Для проведения занятий требуется аудитория, оснащенная проектором и компьютерами.

**3.3.2 Характеристики аудиторного оборудования, в том числе неспециализированного компьютерного оборудования и программного обеспечения общего пользования**

MS Office.

**3.3.3 Характеристики специализированного оборудования**

Необходимо наличие компьютерных классов с высокопроизводительными многопроцессорными компьютерами/параллельными кластерами.

**3.3.4 Характеристики специализированного программного обеспечения**

На компьютерах должны быть установлены СУБД Microsoft SQL Server, Express Edition, Postgres.

**3.3.5 Перечень и объёмы требуемых расходных материалов**

Не требуются.

**3.4. Информационное обеспечение**

**3.4.1 Список обязательной литературы**

1. Дейт К. Введение в системы баз данных. М.: Вильямс, 8-е изд., 2005.
2. Новиков, Борис Асенович. Настройка приложений баз данных: учебное пособие для вузов / Б. А. Новиков, Г. Р. Домбровская. - СПб.: БХВ-Петербург, 2012. - 240 с.

**3.4.2 Список дополнительной литературы**

1. Кузнецов, Сергей Дмитриевич. Основы баз данных: учебное пособие / С. Д. Кузнецов. - 2-е изд., испр. - М.: Интернет-Университет Информационных Технологий; М.: БИНОМ. Лаборатория знаний, 2005-2007. - 484 с.

**3.4.3 Перечень иных информационных источников**

1. http://www.vldb.org/ — сайт конференции Very large data bases
2. http://syrcodis.ispras.ru/ — сайт Spring Colloquium for Young Researchers on Database and Information Systems
3. http://ieee-icde2014.eecs.northwestern.edu/index.html — сайт International Conference on Data Engineering.
4. http://www.edbt.org/ — Extending Database Technology

**Раздел 4. Разработчики программы**

Михайлова Елена Георгиевна, доцент Кафедры ИАС СПбГУ, e.mikhaylova@spbu.ru.