

Документ подписан простой электронной подписью Информация о владельце: ФИО: Таскаев Сергей Валерьевич Должность: Ректор Дата подписания: 21.05.2025 09:19:01 Уникальный программный ключ (специальности) 09.04.04 04c19ed8bfb98f3b6cb77a486b9a8788b8322323	МИНОВЕРНАУКИ РОССИИ Федеральное государственное бюджетное образовательное учреждение высшего образования «Челябинский государственный университет» (ФГБОУ ВО «ЧелГУ»)	Рабочая программа дисциплины "Методы и системы обработки больших данных" по направлению подготовки (специальности) 09.04.04 "Программная инженерия" направленности (профилю) Искусственный интеллект и инженерия данных ФГБОУ ВО «ЧелГУ»	стр. 1
--	--	--	--------

Рабочая программа дисциплины (модуля)*

Методы и системы обработки больших данных

Направление подготовки (специальность)

09.04.04 Программная инженерия

Направленность (профиль)

Искусственный интеллект и инженерия данных

Присваиваемая квалификация (степень)

магистр

Форма обучения

очная

Год(ы) набора

*Рабочая программа дисциплины (модуля) адаптирована для инклюзивного обучения инвалидов и лиц с ограниченными возможностями здоровья

Челябинск 2023 г.

09.04.04 Программная инженерия, Искусственный интеллект и инженерия данных, магистр, *Методы и системы обработки больших данных*, 2024, очная

Проректор по учебной работе утверждено 21.02.2024 А.А. Саламатов

Ученым советом института информационных технологий

Протокол заседания № 6 от 14.02.2024

Председатель Ученого совета
института информационных
технологий

согласовано

Ю. В. Петриченко

Заседанием кафедры информационных технологий и экономической информатики

Протокол заседания № 6 от 14.02.2024

И. о. заведующего кафедрой

согласовано

С.А. Скрипов

Автор (составитель)

А.В. Вохминцев

Структура рабочей программы соответствует приказу ректора ФГБОУ ВО «ЧелГУ» от «13» апреля 2021 г. № 247-1



Содержание

1. Цели освоения дисциплины
2. Место дисциплины в структуре ОПОП
3. Компетенции обучающегося, формируемые в результате освоения дисциплины (модуля)
4. Объем дисциплины (модуля)
5. Структура и содержание дисциплины (модуля)
6. Фонд оценочных средств
 - 6.1. Перечень видов оценочных средств
 - 6.2. Типовые контрольные задания и иные материалы для текущей аттестации
 - 6.3. Типовые контрольные вопросы и задания для промежуточной аттестации
 - 6.4. Критерии оценивания
7. Учебно-методическое и информационное обеспечение дисциплины (модуля)
 - 7.1. Рекомендуемая литература
 - 7.2. Перечень ресурсов информационно-телекоммуникационной сети "Интернет"
 - 7.3. Перечень информационных технологий
8. Материально-техническое обеспечение дисциплины (модуля)
9. Методические указания для обучающихся по освоению дисциплины (модуля)
10. Специальные условия освоения дисциплины обучающимися с инвалидностью и ограниченными возможностями здоровья



1. ЦЕЛИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Цель курса - дать представление о моделях, методах и алгоритмах, используемых в современных системах хранения и обработки больших данных. Эти знания позволят понять как работают СУБД и другие программные системы. Это позволит в будущем принимать правильные проектные решения в управлении центрами обработки данных в больших компаниях.

Результаты обучения по дисциплине направлены на достижение и формирование индикаторов:

ПК-2.1 Знает: фундаментальные знания в области разработки систем управления большими данными; основные положения и концепции в области хранения и обработки больших данных

ПК-2.2 Умеет: осуществлять первичный сбор и анализ материала в области разработки систем управления большими данными; анализировать типовые решения в области хранения и обработки больших данных, реализовывать техническое сопровождение информационных систем и баз данных, используемых для решения задач в области хранения и обработки больших данных, программировать системы хранения и обработки больших данных

ПК-2.3 Имеет практический опыт: анализа и оптимизации найденных решений в области разработки систем управления большими данными; интеграции различных типов программного обеспечения в области хранения и обработки больших данных

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП

Цикл (раздел) ОПОП: Б1.В.02

2.1 Требования к предварительной подготовке обучающегося:

Дисциплина базируется на знаниях, полученных при изучении таких дисциплин

Разработка систем искусственного интеллекта на языке Python

2.2 Дисциплины и практики, для которых освоение данной дисциплины (модуля) необходимо как предшествующее:

Знания, полученные при изучении данной дисциплины, могут быть использованы при

Анализ и прогнозирование временных рядов методами искусственного интеллекта

3. КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)

ПК-2: Способен разрабатывать системы хранения и обработки больших данных

Знать:

фундаментальные знания в области разработки систем управления большими данными

Уметь:

осуществлять первичный сбор и анализ материала в области разработки систем управления большими данными

Владеть:

навыком анализа и оптимизации найденных решений в области разработки систем управления большими данными

В результате освоения дисциплины обучающийся должен

3.1 Знать:

3.2 Уметь:

3.3 Владеть:



4. ОБЪЕМ ДИСЦИПЛИНЫ (МОДУЛЯ)

Общая трудоемкость	3 ЗЕТ
Часов по учебному плану : 108 в том числе : аудиторные занятия : 48 самостоятельная работа : 53,75 : контактная работа: 54,25 ИКР: 6,25	Виды контроля в семестрах: зачеты 2

5. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Код занятия	Наименование разделов и тем /вид занятия/	Семестр / Курс	Часов	Литература
	Раздел 1. Обработка запросов в системах баз данных			
1.1	Общая схема обработки запроса в реляционной СУБД. Реляционная алгебра и реляционные операции. Представление реляционного выражения в виде логического плана. Общая схема компиляции запроса. Разбор запроса. Синтаксический анализ запроса. Грамматика SQL. Построение дерева разбора. Разрешение представлений. Верификация запроса. Конверсия запроса в логический план. Логическая оптимизация. Алгебраические законы. Оптимизация операций выборки, проекции и удаления дубликатов. Оптимизация композиции операций выборки и прямого произведения. Система хранения данных. Поля, записи, блоки. Использование буферного пула. Индексы. В-деревья. Оптимизация физического плана. Оценка стоимости реляционных операций. Статистические характеристики данных. Гистограммы. Выбор порядка соединений. Исполнитель запросов. Итераторы. Алгоритмы соединений. 2 Обработка потоков данных. Система управления потоками данных. Активные запросы. Окна данных. Скетчинг данных. Фильтры Блума. 2 Подсчет количества различных элементов в потоке с помощью алгоритма Флажолле-Мартена. Модель параллельной обработки больших данных MapReduce. Структура вычислительного кластера. Распределенная файловая система. Схема работы MapReduce: функции высшего порядка Map и Reduce; процедура Shuffle. Программный стек Hadoop: набор инфраструктурных программных библиотек и утилит Hadoop Common, распределенная файловая система HDFS, система для планирования заданий и управления кластером YARN, 2 платформа распределенных вычислений Hadoop MapReduce. Экосистема Hadoop: Pig - реализация реляционной алгебры на базе Hadoop, Hive - реализация SQL на базе Hadoop. Программная платформа Spark: Spark SQL - выполнение SQL-запросов над данными; Spark Streaming - обработка потоковых данных; Spark Mllib - набор библиотек машинного обучения; Spark GraphX - распределенная обработка графов. /Лек/	2	32	Л2.1 Л2.2 Л2.3 Э1 Э2 Э3 Э4 Э5



Рабочая программа дисциплины "Методы и системы обработки больших данных" по направлению подготовки (специальности) 09.04.04 "Программная инженерия" направленности (профилю) Искусственный интеллект и инженерия данных ФГБОУ ВО «ЧелГУ»		стр. 5		
1.2	1) Вычислить результат выражения реляционной алгебры - 5 примеров; 2) Написать выражение реляционной алгебры, вычисляющее ответ на запрос к 4 базе данных "Поставки" - 8 примеров; 3) Написать выражение реляционной алгебры, соответствующее логическому плану - 5 примеров. 1) Построить дерево разбора для SQL-запроса - 4 примера; 2) Конвертировать дерево разбора в логический план - 5 примеров; 3) 4) Выполнить оптимизацию логического плана - 4 примера. 1) Вычислить количество блоков, необходимых для хранения отношения - 4 примера; 2) Вычислить количество блоков, необходимое для хранения 4 индексного файла - 4 примера; 3) Вычислить оценку для размера результата реляционного выражения - 4 примера. 1) Вычислить оценку размера естественного соединения отношений с использованием гистограмм - 4 примера; 2) Вычислить оценку эффективности алгоритма соединения - 5 примеров; 3) С использованием 4 метода динамического программирования определить оптимальный порядок соединений - 4 примера. /Пр/	2	16	Л2.1 Л2.2 Л2.3 Э1 Э2 Э3 Э4 Э5
1.3	Повторить пройденные материалы, подготовиться к решению задач. /Ср/	2	53,75	Л2.1 Л2.2 Л2.3 Э1 Э2 Э3 Э4 Э5
Раздел 2. Иная контактная работа				
2.1	Индивидуальные консультации, текущий контроль /ИКР/	2	6,25	Л2.1 Л2.2 Л2.3 Э1 Э2 Э3 Э4 Э5

6. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

6.1. Перечень видов оценочных средств

Тест.

6.2. Типовые контрольные задания и иные материалы для текущей аттестации

Тест содержит три задания. Каждый правильный ответ оценивается 1 баллом. Продолжительность тестирования 30 мин. 3 балла: даны верные ответы на все вопросы теста.
2 балла: даны верные ответы на 2 вопроса теста.
1 балл: дан верный ответ на 1 вопрос теста. 0 баллов: верные ответы отсутствуют

6.3. Типовые контрольные вопросы и задания для промежуточной аттестации

Тест содержит 12 равнозначных вопросов и рассчитан на 45 мин. Дается две попытки. Вторая попытка не зависит от первой. Количество баллов равно количеству правильных ответов на вопросы теста.

6.4. Критерии оценивания

Оценка за дисциплину рассчитывается на основе полученных оценок за контрольно-рейтинговые мероприятия текущего контроля и промежуточной аттестации (зачет). Допускается выставление зачета автоматом (на основе текущего рейтинга). На зачете студенты выполняют итоговый тест, который содержит 12 равнозначных вопросов. Тест рассчитан на 45 минут. Количество баллов за тест равно количеству верных ответов. Для повышения оценки допускается вторая попытка для написания теста, при этом учитывается лучший результат из двух попыток. При оценивании результатов учебной деятельности обучающегося по дисциплине используется балльно-рейтинговая система оценивания результатов учебной деятельности обучающихся (утверждена приказом ректора от 24.05.2019 г. No 179). Зачтено: Величина рейтинга обучающегося по дисциплине 60...100 %. Незачтено: Величина рейтинга обучающегося по дисциплине 0...59 %.

7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

7.1. Рекомендуемая литература

7.1.2. Дополнительная литература

	Авторы, составители	Заглавие	Издательство, год	Ресурс
Л2.1	Маккинни У.	Python и анализ данных (https://e.lanbook.com/book/131721)	Москва : ДМК Пресс, 2020	ЭБС



Рабочая программа дисциплины "Методы и системы обработки больших данных" по направлению подготовки (специальности) 09.04.04 "Программная инженерия" направленности (профилю) Искусственный интеллект и инженерия данных ФГБОУ ВО «ЧелГУ»

стр. 6

	Авторы, составители	Заглавие	Издательство, год	Ресурс
Л2.2	Богданов Е.П.	Интеллектуальный анализ данных: учебное пособие (https://znanium.com/catalog/document?id=357344)	Волгоград : ФГБОУ ВПО Волгоградский государственный аграрный университет, 2019	ЭБС
Л2.3	Радченко И. А., Николаев И. Н.	Технологии и инфраструктура Big Data: учебное пособие (https://e.lanbook.com/book/136430)	Санкт-Петербург : НИУ ИТМО, 2018	ЭБС

7.2. Перечень ресурсов информационно-телекоммуникационной сети "Интернет"

Э1	Лекториум - просветительский проект: массовые открытые онлайн-курсы, открытый видеоархив лекций вузов России https://www.lektorium.tv
Э2	Научная электронная библиотека eLIBRARY.RU» - раздел "Журналы открытого доступа" (https://elibrary.ru/projects/subscription/rus_titles_free.asp)
Э3	Лань [Электронный ресурс] : электронно-библиотечная система (ЭБС) / издательство Лань http://e.lanbook.com
Э4	Юрайт [Электронный ресурс] : электронно-библиотечная система (ЭБС) / издательство Юрайт. https://urait.ru/
Э5	Znanium.com [Электронный ресурс] : электронно-библиотечная система (ЭБС) / Науч. электрон. б-ка http://znanium.com/

7.3 Перечень информационных технологий

7.3.1 Программное обеспечение

LMS Moodle

MS Office365

7.3.2 Профессиональные базы данных и информационно-справочные системы

1. Национальная электронная библиотека (НЭБ) (<https://rusneb.ru/>) Национальная электронная библиотека (НЭБ) : объединенный электронный каталог фондов российских библиотек : сайт. – URL: <http://нэб.рф>. – Режим доступа: из читальных залов библиотеки ЧелГУ. – Текст : электронный.

2. Научная электронная библиотека eLIBRARY.RU (<https://elibrary.ru/defaultx.asp?>) eLIBRARY.RU : научная электронная библиотека : сайт. – Москва, 2000 – . – URL: <https://elibrary.ru>. – Режим доступа: для зарегистрир. пользователей. – Текст : электронный.

3. EBSCO Information Services-EBSCOhost Research Databases

8. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Для реализации дисциплины используются учебные аудитории для проведения занятий лекционного типа, занятий семинарского типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, а также помещения для самостоятельной работы.

Учебные аудитории укомплектованы специализированной мебелью и техническими средствами обучения: доска, парты, мультимедийное и аудиооборудование.

Для проведения занятий лекционного типа предлагаются наборы демонстрационного оборудования и учебно-наглядных пособий: цифровые образовательные ресурсы, а также используется переносное и / или стационарное мультимедийное оборудование (экран, ноутбук, проектор, колонки).

Для семинарских занятий используются аудитории оснащенные обычной доской, партами, переносным мультимедийным и аудиооборудованием (в случае необходимости).

Помещения для самостоятельной работы обучающихся оснащены компьютерной техникой с подключением к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду университета.

9. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ (МОДУЛЯ)

Успешное изучение курса требует от обучающихся посещения лекций, активной работы на семинарах, выполнения всех учебных заданий преподавателя, ознакомления с базовыми учебниками, основной и дополнительной



литературой. Запись лекции – одна из форм активной работы студентов, требующая навыков и умения кратко, схематично, последовательно и логично фиксировать основные положения, выводы, обобщения, формулировки. Последующая работа над текстом лекции воскрешает в памяти ее содержание, позволяет развивать экономическое мышление. В конце лекции преподаватель оставляет время (5 минут) для того, чтобы студенты имели возможность задать уточняющие вопросы по изучаемому материалу.

Основным методом обучения является самостоятельная работа студентов с учебно-методическими материалами, научной литературой. При изучении дисциплины необходимо изучить вопросы, которые преподаватель вынес на самостоятельное изучение, быть готовым к обсуждению этих вопросов. Дискуссия – коллективная форма устного представления информации. Обычно дискуссии готовит один или несколько человек, представляющих основные вопросы темы и точки зрения. Остальные участники дискуссии высказывают свои мнения и суждения. Дискуссию организует ведущий (чаще преподаватель) в обязанность которого входит предоставление слова разным участникам, сдерживание эмоциональных реакций участников и подведение итогов обсуждения.

К промежуточной аттестации необходимо готовиться целенаправленно, регулярно, систематически и с первых дней обучения по данной дисциплине. После этого у обучающегося должно сформироваться четкое представление об объеме и характере знаний и умений, которыми надо будет овладеть по дисциплине. Систематическое выполнение учебной работы на лекциях и семинарских занятиях позволит успешно освоить дисциплину.

В случае применения при обучении дисциплины электронного обучения, дистанционных образовательных технологий общение обучающихся и преподавателя осуществляется в режиме реального времени (онлайн-лекции (вебинары), чаты, видео-конференции и др.) или отложенного времени (система дистанционного обучения Moodle, MS Office365, форумы, электронная почта и др.).

Большую часть времени обучающиеся самостоятельно работают с учебно-методическими материалами. Студенты имеют возможность консультироваться с преподавателем по всем вопросам, возникающим в ходе самостоятельной работы посредством электронной почты, социальных сетей, системы дистанционного обучения Moodle и т.п.

Доступ обучающегося к учебным ресурсам в режиме отложенного времени, самостоятельной работы осуществляется через сеть Интернет в удобном для него месте, времени и темпе.

При обучении лиц с ограниченными возможностями здоровья электронное обучение, дистанционные образовательные технологии предусматривают возможность приема-передачи информации в доступных для них формах.

Реализация дисциплины с применением электронного обучения, дистанционных образовательных технологий (далее – ЭО, ДОТ) осуществляется на основании «Положения о реализации основных и дополнительных образовательных программ с применением электронного обучения и дистанционных образовательных технологий в федеральном государственном бюджетном образовательном учреждении высшего образования «Челябинский государственный университет», «Положения о порядке зачета обучающимися по основным профессиональным образовательным программам высшего образования в ФГБОУ ВО «ЧелГУ» результатов освоения в организациях, осуществляющих образовательную деятельность, учебных предметов, курсов, дисциплин (модулей), практик, дополнительных образовательных программ» посредством электронной информационно-образовательной среды ФГБОУ ВО «ЧелГУ». В исключительных случаях (форс-мажор и т.п.) при реализации образовательной деятельности с применением ЭО, ДОТ могут применять компоненты, не входящие в перечень электронной информационно-образовательной среды.

10. СПЕЦИАЛЬНЫЕ УСЛОВИЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ ОБУЧАЮЩИМИСЯ С ИНВАЛИДНОСТЬЮ И ОГРАНИЧЕННЫМИ ВОЗМОЖНОСТЯМИ ЗДОРОВЬЯ

Освоение дисциплины инвалидами и лицами с ограниченными возможностями здоровья осуществляется с использованием специальных технических средств и голо информационных технологий, предоставляемых Ресурсным учебно-методическим центром по обучению инвалидов и лиц с ограниченными возможностями здоровья ЧелГУ по запросу обучающегося.

1. Мобильные специальные технические средства для лиц с нарушениями зрения: портативный компьютер с вводом/выводом шрифтом Брайля с синтезатором речи «ElBraille-W14J G2»; ноутбуки с программой экранного доступа NVDA; электронные увеличители для удаленного просмотра; видеувеличители портативные; тифлоплеер; цифровые диктофоны.

2. Мобильные специальные технические средства для лиц с нарушениями слуха: система свободного звукового поля со встроенной совместимостью с FM-устройствами; радиоклассы «Сонет-PCM» с передатчиком, заушным индуктором и индукционной петлей; система информационная для слабослышащих переносная «Исток» А2 со встроенным плеером – звуковым информатором; документ-камера; программируемые слуховые аппараты индивидуального пользования.

3. Ассистивные информационные технологии: программное обеспечение экранного доступа с синтезом речи NVDA; программы экранного увеличения; программы речевого синтеза для компьютеров и ноутбуков; программы речевого синтеза для мобильных устройств; экранная клавиатура; экранная лупа.

При необходимости для обучающихся с нарушениями зрения на рабочих местах для проведения практических или



лабораторных занятий устанавливается специальное программное обеспечение (программа речевой навигации NVDA, речевые синтезаторы, экранные лупы).

В учебные аудитории обеспечивается беспрепятственный доступ для обучающихся инвалидов и обучающихся с ограниченными возможностями здоровья. В каждой аудитории, где обучаются инвалиды и лица с ограниченными возможностями здоровья, предусматривается соответствующее количество мест для обучающихся с учетом нарушений их здоровья.

Для освоения дисциплины инвалидам и лицам с ограниченными возможностями здоровья предоставляется доступ к печатным источникам, имеющимся в научной библиотеке ЧелГУ, с помощью специальных технических средств; доступ к электронным источникам, представленным в форме электронного документа в фонде научной библиотеки ЧелГУ или электронно-библиотечных системах, с помощью специальных технических и программных средств (рабочее место для незрячего пользователя с программным обеспечением экранного доступа с синтезом речи NVDA, рабочее место с компьютерным роллером и клавиатурой Clevy с большими кнопками и с разделяющей клавиши накладкой).

Учебно-методические материалы для обучающихся из числа инвалидов и лиц с ограниченными возможностями здоровья предоставляются в формах, адаптированных к ограничениям их здоровья и восприятия информации:

Для лиц с нарушениями зрения:

- в печатной форме увеличенным шрифтом,
- в форме электронного документа,
- в форме аудиофайла,
- в печатной форме шрифтом Брайля.

Для лиц с нарушениями слуха:

- в печатной форме,
- в форме электронного документа.

Для лиц с нарушениями опорно-двигательного аппарата:

- в печатной форме,
- в форме электронного документа,
- в форме аудиофайла.

Данный перечень может быть конкретизирован в зависимости от контингента обучающихся.

Для инвалидов и лиц с ограниченными возможностями здоровья освоение дисциплины может быть частично или полностью осуществлено с использованием дистанционных образовательных технологий (Moodle, Adobe Connect Pro и пр.).

В освоении дисциплины инвалидами и лицами с ограниченными возможностями здоровья используется индивидуальная работа. Под индивидуальной работой подразумевается две формы взаимодействия с преподавателем: индивидуальная учебная работа (консультации), т.е. дополнительное разъяснение учебного материала и углубленное изучение материала с теми обучающимися, которые в этом заинтересованы, и индивидуальная воспитательная работа. Индивидуальные консультации направлены на индивидуализацию обучения и установлению воспитательного контакта между преподавателем и обучающимся инвалидом или обучающимся с ограниченными возможностями здоровья.

При проведении процедуры оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья по дисциплине обеспечивается выполнение следующих дополнительных требований в зависимости от индивидуальных особенностей, обучающихся:

- а) инструкция по порядку проведения процедуры оценивания предоставляется в доступной форме (устно, в письменной форме, в письменной форме шрифтом Брайля, устно с использованием услуг сурдопереводчика);
- б) доступная форма предоставления заданий оценочных средств (в печатной форме, в печатной форме увеличенным шрифтом, в печатной форме шрифтом Брайля, в форме электронного документа, задания зачитываются ассистентом, задания предоставляются с использованием сурдоперевода);
- в) доступная форма предоставления ответов на задания (письменно на бумаге, набор ответов на компьютере, письменно шрифтом Брайля, с использованием услуг ассистента, устно).

При проведении процедуры оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья предусматривается использование технических средств, необходимых им в связи с их индивидуальными особенностями. Эти средства могут быть предоставлены ЧелГУ или могут использоваться собственные технические средства. При необходимости инвалидам и лицам с ограниченными возможностями здоровья предоставляется дополнительное время для подготовки ответа на задания, процедура оценивания результатов обучения по дисциплине может проводиться в несколько этапов.

Проведение процедуры оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья допускается с использованием дистанционных образовательных технологий.