

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«Национальный исследовательский ядерный университет «МИФИ»
Обнинский институт атомной энергетики –
филиал федерального государственного автономного образовательного учреждения высшего образования
«Национальный исследовательский ядерный университет «МИФИ»
(ИАТЭ НИЯУ МИФИ)

Утверждено на заседании
УМС ИАТЭ НИЯУ МИФИ
Протокол №2-8/2021 От 30.08.2021

РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ

Визуализация данных: Metabase

Шифр, название дисциплины

01.04.02 «Прикладная математика и информатика»

Шифр, название специальности/направления подготовки

Математическое моделирование и прикладной анализ данных

Название программы магистратуры

магистр

(Квалификация (степень) выпускника)

Форма обучения: очная

г. Обнинск 2024 г.

Программа составлена в соответствии с требованиями образовательного стандарта высшего образования национального исследовательского ядерного университета «МИФИ» по направлению подготовки 01.04.02 – Прикладная математика и информатика. (квалификация (степень) магистр).

Программу составил:

_____ С.В. Ермаков, доцент, к.ф.-м.н, доцент

Рецензент:

_____ Г.Е. Деев, доцент, к.ф.-м.н, доцент

Программа рассмотрена на заседании ОИКС

(протокол № 5/7 от «30» июля от 2024 г.)

Руководитель направления подготовки 01.04.02
«Прикладная математика и информатика»

_____ Ермаков С.В.

« ____ » _____ 2024 г.

1. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы

В результате освоения ООП магистратуры обучающийся должен овладеть следующими результатами обучения по дисциплине:

Коды компетенций	Результаты освоения ООП <i>Содержание компетенций*</i>	Перечень планируемых результатов обучения по дисциплине**
УК-1	Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий	З-УК-1 Знать: методы системного и критического анализа; методики разработки стратегии действий для выявления и решения проблемной ситуации У-УК-1 Уметь: применять методы системного подхода и критического анализа проблемных ситуаций; разрабатывать стратегию действий, принимать конкретные решения для ее реализации В-УК-1 Владеть: методологией системного и критического анализа проблемных ситуаций; методиками постановки цели, определения способов ее достижения, разработки стратегий действий.
ОПК-1	Способен решать актуальные задачи фундаментальной и прикладной математики	З-ОПК-1 Знать актуальные задачи фундаментальной и прикладной математики, методы математического моделирования. У-ОПК-1 Уметь использовать методы математического моделирования для решения задач фундаментальной и прикладной математики. В-ОПК-1 Владеть методами математического моделирования и основами их использования.

2. Место дисциплины в структуре ООП магистратуры

Дисциплина реализуется в рамках общенаучного модуля.

Для освоения дисциплины необходимы компетенции, сформированные в рамках изучения следующих дисциплин: SQL для анализа данных

Дисциплина изучается на 1 курсе в 1 семестре.

3. Объем дисциплины в зачетных единицах с указанием количества академических часов, выделенных на контактную работу обучающихся с преподавателем (по видам занятий) и на самостоятельную работу обучающихся

Общая трудоемкость (объем) дисциплины составляет 6 зачетных единиц (з.е.), 216 академических часа.

3.1. Объем дисциплины по видам учебных занятий (в часах)

	Семестр		
	№ 1	№ 2	Всего
	Количество часов на вид работы:		
Контактная работа обучающихся с преподавателем			
Аудиторные занятия (всего)	64		64
В том числе:			
лекции	16		16
практические занятия	16		16
лабораторные занятия			
Промежуточная аттестация			
В том числе:			
зачет			
экзамен	36		36
Самостоятельная работа обучающихся (всего)	148		148
В том числе:			
проработка учебного (теоретического) материала	37		37
выполнение индивидуальных заданий	37		37
подготовка ко всем видам контрольных испытаний текущего контроля успеваемости (в течение семестра)	37		37
подготовка ко всем видам контрольных испытаний промежуточной аттестации (по окончании семестра)	37		37
Всего (часы):	216		216
Всего (зачетные единицы):	6		6

4. Содержание дисциплины, структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий

4.1. Разделы дисциплины и трудоемкость по видам учебных занятий (в академических часах)

№ п/п	Наименование раздела /темы дисциплины	Общая трудоём- кость всего (в часах)	Виды учебных занятий, включая самостоятельную работу обучающихся и трудоёмкость (в часах)				Формы текущего контроля успевае- мости
			Аудиторные учебные занятия			СРО	
			Лек	Сем/Пр	Лаб		
1.			16	16	-	148	

1.1.	Знакомимся с Metabase	5	1	1		16	
1.2.	Как зайти в Metabase	5	1	1	-	16	
1.3.	Визуализируем наши SQL-запросы	10	3	3	-	16	
1.4.	Разбор домашнего задания	6	1	1	-	16	
1.5.	Работа с Questions	8	2	2		16	
1.6.	Filter Fields	8	2	2	-	16	
1.7.	Построение дашбордов	8	2	2	-	16	
1.8.	Дополнительные фишки	8	2	2	-	16	
1.9.	Итоговая работа	8	2	2		20	

4.2. Содержание дисциплины, структурированное по разделам (темам)

Лекционный курс

№	Наименование раздела /темы дисциплины	Содержание
1.1.	Знакомимся с Metabase	что такое Metabase?
1.2.	Как зайти в Metabase	работаем с нашим экземпляром Metabase
1.3.	Визуализируем наши SQL-запросы	Как в Metabase писать SQL-запросы Какие основные типы визуализаций есть в Metabase Какие есть фишки для форматирования визуализаций Научимся выставлять цели, настраивать прогресс выполнения
1.4.	Разбор домашнего задания	Разбор практики
1.5.	Работа с Questions	Что такое Questions

		Как джойнить таблицы Как фильтровать данные Как делать агрегации Как написать SQL-запрос, даже если не знаешь SQL
1.6.	Filter Fields	поля для фильтрации или Filter Fields.
1.7.	Построение дашбордов	Построение дашбордов
1.8.	Дополнительные фишки	Шеринг отдельных элементов Скачивание в файл Коллекции
1.9	Итоговая работа	Визуализация аналитического исследования

Практические/семинарские занятия

№	Наименование раздела /темы дисциплины	Содержание
1.1.	Знакомимся с Metabase	что такое Metabase?
1.2.	Как зайти в Metabase	работаем с нашим экземпляром Metabase
1.3.	Визуализируем наши SQL-запросы	Как в Metabase писать SQL-запросы Какие основные типы визуализаций есть в Metabase Какие есть фишки для форматирования визуализаций Научимся выставлять цели, настраивать прогресс выполнения
1.4.	Разбор домашнего задания	Разбор практики
1.5.	Работа с Questions	Что такое Questions Как джойнить таблицы Как фильтровать данные Как делать агрегации Как написать SQL-запрос, даже если не знаешь SQL
1.6.	Filter Fields	поля для фильтрации или Filter Fields.
1.7.	Построение дашбордов	Построение дашбордов
1.8.	Дополнительные фишки	Шеринг отдельных элементов

		Скачивание в файл Коллекции
1.9	Итоговая работа	Визуализация аналитического исследования

Лабораторные занятия

Не предусмотрены.

5. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине

В качестве учебно-методических материалов используется рекомендованная литература.

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

6.1. Паспорт фонда оценочных средств по дисциплине

№ п/п	Контролируемые разделы (темы) дисциплины (результаты по разделам)	Код контролируемой компетенции (или её части) / и ее формулировка	Наименование оценочного средства
1.3	Пространство обобщенных функций D'	УК-1:	Контрольная работа № 1
2.1.	Дифференцирование обобщенных функций.	ОПК-1	Контрольная работа № 1
2.3.	Свертка обобщенных функций	ОПК-1:	Контрольная работа № 2
2.4.	Преобразование Фурье обобщенных функций медленного роста.	ОПК-1.	Контрольная работа № 2
3.1.	Фундаментальные решения линейных дифференциальных операторов.	УК-1	Контрольная работа № 2

6.2. Типовые контрольные задания или иные материалы

6.2.1. Экзамен

В экзаменационном билете два теоретических вопроса и один практический

Теоретические вопросы билета:

1. Что такое реляционная база данных и какие её основные характеристики?
2. В чем отличие между первичным и внешним ключом?
3. Как нормализуются данные в базе данных, и зачем это нужно?
4. Как написать SQL-запрос для выборки всех данных из таблицы?
5. Как ограничить количество строк в результате запроса с помощью оператора LIMIT?
6. Что такое оператор WHERE и как он используется для фильтрации данных?

7. Как использовать операторы поиска Google для нахождения SQL-запросов и решений задач по базам данных?
8. Какие ресурсы и форумы для SQL и работы с базами данных вы используете для поиска решений?
9. В чем разница между INNER JOIN, LEFT JOIN и RIGHT JOIN? Приведите примеры использования каждого типа объединения.
10. Как объединить более двух таблиц в SQL? Приведите пример запроса с несколькими объединениями.
11. Как использовать оператор GROUP BY для группировки данных? Приведите пример запроса с агрегатной функцией COUNT.
12. Чем отличается HAVING от WHERE в SQL, и в каком случае следует использовать HAVING?
13. Что такое подзапрос, и как он используется в SQL? Приведите пример.
14. Чем CTE (Common Table Expressions) отличается от обычных подзапросов?
15. В чем отличие между подзапросами с корреляцией и без?
16. Что такое оконные функции в SQL? Приведите пример использования функции ROW_NUMBER().
17. Как использовать оконные функции для вычисления скользящих средних?
18. Какие основные принципы код-стайла следует соблюдать при написании SQL-запросов?
19. Почему важно использовать алиасы для таблиц и столбцов в SQL-запросах?
20. Чем отличаются команды DDL (Data Definition Language) от DML (Data Manipulation Language)? Приведите примеры каждой.
21. Как создать таблицу в SQL с использованием команды CREATE TABLE? Приведите пример.
22. Как подключить базу данных к Metabase и начать работать с запросами?
23. Как построить дашборд в Metabase, добавив несколько визуализаций?
24. Что такое "Questions" в Metabase и как с ними работать?
25. Как работать с данными в формате JSON в SQL? Приведите пример извлечения данных из JSON-колонки.
26. Как использовать функции SQL для обработки неструктурированных данных?
27. Что такое инкрементальное обновление данных, и как оно помогает при работе с большими объемами данных?
28. Как настроить инкрементальную загрузку данных с использованием временных меток?
29. Как использовать оператор ROLLUP в SQL для создания иерархической агрегации данных?
30. В чем отличие между CUBE и ROLLUP в SQL? Приведите примеры использования.

Критерий оценки – правильность и полнота ответа на вопросы. Оценка выставляется по шкале от 0 до 40 баллов: теоретические вопросы – 30 баллов, 10 баллов – дополнительные вопросы. Экзамен считается сданным при оценке не ниже 25 баллов.

6.2.2. Контрольная работа № 1

Постройте визуализацию опираясь на условия ниже:

Выделим такие метрики:

1. DAU (также хочется видеть динамику по сравнению с предыдущими периодами)

2. MAU
3. Распределение активации аккаунтов (подтвердили/не подтвердили аккаунт)
4. Распределение регистраций по компаниям
5. Воронка "Сделал попытку решить задачу - Решил задачу успешно - Пополнил кошелек"
6. Распределение - на что люди тратят CodeCoins (покупают задачи, подсказки, решения, тесты)
7. Распределение первых и повторных покупок - сколько человек купило первый раз, а сколько совершают уже не первую покупку
8. Rolling retention по заходам на платформу с разбивкой по когортам

Очень важно, чтобы везде был удобный фильтр по дате - чтобы мы могли ежедневно ставить удобный для нас период и смотреть статистику за разное время.

6.2.2. Контрольная работа № 2

Постройте визуализацию опираясь на условия ниже:

1. Нам важно понимать, какие пользователи приходят к нам просто так, а какие - участники каких-то компаний. Поэтому хочется иметь график, который максимально наглядно покажет соотношение пользователей у нас на платформе.
2. Одна из важнейших метрик - активация аккаунта. Если у нас зарегистрировалось 1000 человек, но аккаунт подтвердили всего 10, то нам толку мало. Поэтому хотелось бы видеть нечто, где мы будем явно видеть долю активированных и неактивированных аккаунтов.
3. Также у нас есть гипотеза показывать разные рекламные предложения нашим пользователям в зависимости от времени суток, чтобы лучше их конвертировать в покупку. А для этого хочется видеть, как от времени суток вообще зависит активность. Поэтому сделайте, пожалуйста, визуализацию, на которой будет видно количество попыток решения задач в разбивкой по часам в сутках. При этом очень важно также, чтобы мы видели и среднее значение на том же графике. Так будет проще сравнивать и оценивать.
4. И последнее - мы бы хотели видеть «воронку» жизни пользователя на нашей платформе. То есть как регистрация конвертируется в активную регистрацию, затем активная регистрация в попытку решения задачи, попытка - в успешное решение, а успешное решение - в покупку задачи (возможно, какой-то другой).

б) критерии оценивания компетенций (результатов) – правильная работа кода программы, понимание алгоритма метода оптимизации, умение вывести необходимые для алгоритма формулы.

в) описание шкалы оценивания:

Каждая задача оценивается по шкале от 0 до 10 баллов.

Контрольная работа считается выполненной успешно при суммарной оценке не ниже 18 баллов.

6.3. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций

Форма аттестации	Наименование оценочного средства	Баллы
Экзамен (100 баллов)	Контрольная работа № 1	30
	Контрольная работа № 2	30
	Ответы на экзаменационный билет	40

7. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины

а) основная учебная литература:

1. Сергей Малиновский — *SQL для аналитиков* — 2020 — 280 стр.
2. Екатерина Баранова — *Метабаза: визуализация и создание дашбордов* — 2020 — 220 стр.
3. Виталий Яковлев — *Основы BI-систем: от отчетности до дашбордов* — 2022 — 260 стр.

б) дополнительная учебная литература:

- 1) **Юрий Лаврентьев** — *Практическое руководство по меткам и фильтрам в BI* — 2021 — 300 стр.

8. Перечень ресурсов* информационно-телекоммуникационной сети «Интернет» (далее - сеть «Интернет»), необходимых для освоения дисциплины

-

9. Методические указания для обучающихся по освоению дисциплины

Вид учебного занятия	Организация деятельности студента
Лекция	Написание конспекта лекций: кратко, схематично, последовательно фиксировать основные положения, выводы, формулировки, обобщения; помечать важные мысли, выделять ключевые слова, термины. Проверка терминов, понятий с помощью энциклопедий, словарей, справочников с выписыванием толкований в тетрадь. Обозначить вопросы, термины, материал, который вызывает трудности, пометить и попытаться найти ответ в рекомендуемой литературе. Если самостоятельно не удастся разобраться в материале, необходимо сформулировать вопрос и задать преподавателю на консультации, на практическом занятии.
Практические занятия	Проработка рабочей программы, уделяя особое внимание целям и задачам, структуре и содержанию дисциплины. Работа с конспектом лекций, просмотр рекомендуемой литературы. Изучение выбранной предметной области на примерах решения задач семинарских занятий, индивидуальных домашних заданий.
Курсовая работа	Не предусмотрена
Контрольная работа	Ознакомиться с основной и дополнительной литературой, включая справочные издания, зарубежные источники, основополагающие термины. Попрактиковаться в решении аналогичных домашних задач по всем темам

	контрольных работ.
Лабораторная работа	Не предусмотрена.
Подготовка к экзамену	При подготовке к экзамену необходимо ориентироваться на конспекты лекций и рекомендуемую литературу.

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

Издательская система LaTeX для подготовки докладов, презентаций и учебного материала.

11. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине

Видеопроектор, компьютер, издательская система LaTeX для подготовки докладов, презентаций и учебного материала.

12. Иные сведения и (или) материалы

12.1. Перечень образовательных технологий, используемых при осуществлении образовательного процесса по дисциплине

Часов в интерактивной форме – 8.

В ходе практических занятий происходит публичное обсуждение каждой решаемой задачи. При этом студенты высказывают свои мнения по выбору наиболее простого способа поиска оптимального решения.

После решения домашних работ на консультациях проводится разбор допущенных студентами ошибок.

12.2. Формы организации самостоятельной работы обучающихся (темы, выносимые для самостоятельного изучения; вопросы для самоконтроля; типовые задания для самопроверки)

Некоторые темы изучаются студентами самостоятельно. Для изучения используется приведённая в списке основная и дополнительная литература. Контроль освоения материала осуществляется при проверке контрольных работ, домашнего задания и на экзамене.

№	Тема и часть, изучаемая (осваиваемая) самостоятельно
1.1	Инструменты для визуализации данных
1.2	ETL-процессы и интеграция с Metabase
1.3	SQL для аналитики в BI-системах
1.4	API Metabase и автоматизация
1.5	Аналитика в реальном времени
1.6	Продвинутые фильтры и параметры
1.7	Методы очистки и подготовки данных для BI
1.8	Обработка больших данных в BI
1.9	Метрики и KPI для дашбордов
1.10	Использование метаданных в Metabase

Вопросы и задания для самоконтроля по всем темам:

1. Как зайти в Metabase?
2. Как визуализировать SQL-запросы в Metabase?
3. Как выполнять разбор домашнего задания в Metabase?
4. Что такое "Questions" в Metabase и как с ними работать?
5. Что такое "Filter Fields" и как их использовать в Metabase?
6. Как построить дашборды в Metabase и какие элементы можно добавлять?
7. Какие дополнительные фишки Metabase могут улучшить работу с данными?

12.3. Краткий терминологический словарь

Filter Fields	это поля, которые позволяют добавлять фильтры в запросы или дашборды в Metabase. Фильтры позволяют пользователям динамически изменять отображаемые данные, выбирая определённые значения или диапазоны для анализа.
Хранимые процедуры	это заранее определённые SQL-запросы или наборы команд, которые сохраняются в базе данных и могут быть выполнены по запросу. Хранимые процедуры используются для автоматизации повторяющихся операций, улучшения производительности запросов и обеспечения централизованного управления бизнес-логикой на уровне базы данных.