

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ М.В. ЛОМОНОСОВА

ФАКУЛЬТЕТ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И КИБЕРНЕТИКИ

Направление подготовки

«ПРИКЛАДНАЯ МАТЕМАТИКА И ИНФОРМАТИКА»

Магистерская программа

«МНОГОМАСШТАБНОЕ МОДЕЛИРОВАНИЕ И МЕТОДЫ АНАЛИЗА
ДАННЫХ В ЕСТЕСТВЕННОНАУЧНЫХ ИССЛЕДОВАНИЯХ»

Цель программы - подготовка специалистов в области математического моделирования, анализа данных и программного обеспечения для проведения комплексных исследований в области естественных наук. Выпускники программы владеют теоретическими основами многомасштабного моделирования, навыками построения физико-математических моделей, методами анализа и управления данными, включая большие данные, программными методами и средствами построения многоуровневых информационных систем, с базами знаний, по направлениям исследований в области естествознания.

Мы учим:

- разрабатывать, реализовывать и тестировать **многомасштабные модели** сложных физических явлений и процессов;
- анализировать и обосновывать модели, методы, алгоритмы для решения задач в междисциплинарных областях, применять методы **обработки и анализа данных**;
- проводить анализ предметной области, проектировать архитектуру и **реализовывать приложения и базы данных** для прикладных задач и научных сервисов;
- выбирать и применять **методы машинного обучения** к задачам кластеризации, классификации, распознавания образов, анализировать и обосновывать модели и методы анализа данных в междисциплинарных исследованиях;
- применять **параллельную обработку данных** для научных исследований и проведение расчетов на высокопроизводительных вычислительных комплексах при решении широкого класса исследовательских и практических задач;

Дисциплины курса

Модуль «Моделирование в задачах естествознания»

Многомасштабное моделирование в задачах естествознания

Курс посвящен созданию многомасштабных моделей и многоуровневых информационных систем для исследования научных проблем, включающих в себя явления несопоставимых пространственных и/или временных масштабов, изучение которых невозможно без учета всех факторов, играющих ключевые роли в таких задачах.

Курс дает необходимые знания, формирует интегрированное представление и навыки по применению формализованной технологии многомасштабного математического моделирования, классических методов физико-математического моделирования, передовых технологий анализа данных и информационных технологий для решения актуальных прикладных задач в области естествознания.

Нейросетевое моделирование

Курс дает представление и формирует навыки необходимые при создании математических моделей, алгоритмов и программного обеспечения для разработки и адаптации глубоких нейронных сетей, включая импульсные. Рассматриваются аналоговые реализации нейросетевых алгоритмов. Приводятся примеры, включая схематические и программные решения построения глубоких нейронных сетей на мемристорных кроссбарах в задачах распознавания изображений, объектов, другое.

Методы создания нейроморфных систем

Курс дает представление о теоретических подходах и современных методах, применяемых при создании нейроморфных систем работающих на новых физических принципах. Рассматриваются вопросы разработки прототипов нейроморфных вычислительных устройств, используемых не только для обучения нейросетей и распознавания объектов, но и для проведения полноценных симуляций.

Интервальный анализ и его приложения

Курс посвящен основам интервального анализа, дает необходимые знания о современных подходах к решению задач с интервальными неопределенностями и неоднозначностями в данных, возникающими в постановке задачи или на промежуточных стадиях процесса решения, формирует представления о самых актуальных направлениях использования интервальных вычислений.

Модуль «Системы и средства представления и хранения информации»

Платформы для проектирования баз знаний

Курс дает необходимые элементы анализа и проектирования моделей данных для реляционных и нереляционных СУБД, вместе с практическими навыками работы с языками SQL, Java, Python, необходимыми для дальнейшей работы с Apache Spark, Apache Hadoop, Hive SQL и хранилищами данных. В курсе рассматриваются SQL для аналитики и основы оптимизации запросов для работы с большими объемами данных.

Разработка распределенных приложений в научных исследованиях

Курс посвящен проектированию и разработке современных распределенных веб-приложений и распределенных NoSQL баз данных для научных сервисов. Практическая часть курса включает практику с наиболее распространенными продуктами каждого из типов NoSQL баз данных: MongoDB, Apache Cassandra, Elasticsearch, Redis, Hazelcast а также принципы и практику разработки распределенных высоконагруженных REST Web сервисов для научных приложений.

Модуль «Методы и технологии анализа данных в задачах естествознания»

Современные подходы к обработке и анализу данных

Математические основы анализа данных, необходимые для дальнейшей работы с моделями машинного обучения (элементы линейной алгебры, математической статистики и численных методов) с практическими заданиями по анализу данных на Python (NumPy, Pandas)

Теоретические основы машинного обучения

В курсе рассматриваются основные задачи обучения по прецедентам: классификация, кластеризация, регрессия, понижение размерности. Изучаются методы их решения, как классические, так и новые, созданные за последние годы. Нейронные сети. Нейронные сети глубокого обучения. Рекомендательные системы. Упор делается на глубокое понимание математических основ, взаимосвязей, достоинств и ограничений рассматриваемых методов

Параллельная обработка данных в научных исследованиях

Курс посвящен принципам и инструментарию параллельных вычислений и работе с данными в научных исследованиях с использованием OpenMP, MPI, GPU (Nvidia CUDA)

Машинное обучение на больших массивах данных

В курсе рассматривается набор задач, решаемых разнообразными пакетами машинного обучения, где требуется проводить обучение или анализ больших объемов данных. Рассматриваются задачи анализа текстов, распознавание образов в изображениях, алгоритмы на графах и другие

Спецсеминар «Многомасштабное моделирование и методы анализа данных в задачах естествознания»

В первом семестре на спецсеминаре сотрудники и преподаватели кафедры рассказывают про свои научные проекты, в которых в том числе могут участвовать студенты в рамках своих выпускных работ. Также мы планируем приглашать наших выпускников и коллег из партнерских организаций с целью трудоустройства студентов и сотрудничества.

Дисциплины блока по выбору студента

Многомасштабное моделирование в задачах материаловедения

Рассматривается применение многомасштабного подхода к задачам вычислительного материаловедения на примере множества реализованных проектов ФИЦ ИУ РАН.

Хранилища данных и OLAP

В данном курсе рассматриваются принципы организаций многомерных и реляционных хранилищ данных, как традиционного способа для анализа и визуализации. Практическая часть включает проектирования хранилища, загрузку и трансформацию данных, визуализацию данных для интерактивного анализа с использованием Tableau.

Математические методы обработки изображений

Программирование графических процессоров

Курс посвящен низкоуровневой работе с CUDA на C++, с помощью которой можно реализовывать высокопроизводительные расчеты в различных задачах моделирования в том случае, когда в готовом пакете нет такой возможности. Курс дает глубокое понимание работы

с современным кластером GPU для его оптимального использования при расчетах.

Хранение и обработка больших данных

В курсе рассматриваются основные паттерны хранения и обработки данных большого объема на базе платформы Apache Hadoop: HDFS, YARN, Map Reduce (Java), Hive SQL. Особое внимание уделяется практике на Apache Spark (Java, Python), пакетной и потоковой обработке и анализу данных. Практические задания выполняются на учебном кластере Cloudera CDH с реальными данными.

Математические модели полимерных композиционных материалов

В курсе анализируются различные математические модели, применяемые для изучения структурных особенностей и свойств полимерных композиционных материалов, а также рассматриваются методы и подходы к изучению процессов деградации механических свойств, прочности и разрушения элементов композитных конструкций. Представлен новый междисциплинарный подход, базирующийся на интеграции многомасштабного моделирования, методов анализа данных и информационных технологий, который дает новое представление о механизмах химических превращений в полимерных композитах при внешних воздействиях, об особенностях их создания и сроках эксплуатации.

Большинство обязательных курсов и курсов по выбору имеют уклон в математическое моделирование с той или иной интенсивностью программирования на Python или C++, либо уклон в промышленное программирование на различных СУБД, платформе Java, языкам Scala, Kotlin и др. Таким образом, студентам предоставляется выбрать наиболее оптимальную комбинацию навыков и знаний для дальнейшей успешной карьеры.