

Федеральное государственное бюджетное образовательное учреждение
высшего образования
Московский государственный университет имени М. В. Ломоносова
Механико-математический факультет

УТВЕРЖДАЮ
И. о. декана механико-математического
факультета МГУ

_____ /В. Н. Чубариков /

«__» _____ 2018 г.
М.П.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Наименование дисциплины:	Аналитика больших данных
Уровень высшего образования:	Специалитет
Направление подготовки / специальность:	01.05.01 «Фундаментальные математика и механика»
Направленность (профиль)/специализация ОПОП:	Специализация «Фундаментальная математика»
Форма обучения:	Очная
Язык преподавания:	Русский
Автор (авторы) программы:	Главацкий Сергей Тимофеевич, доцент, к.ф.-м.н.

Рабочая программа рассмотрена и одобрена
На заседании кафедры теоретической информатики

Москва 2018

Рабочая программа дисциплины разработана в соответствии с самостоятельно установленным МГУ образовательным стандартом (ОС МГУ) для реализуемых основных профессиональных образовательных программ высшего образования по направлению подготовки / специальности «Фундаментальные математика и механика» (программы специалитета) от 22 июля 2011 года № 729 (в редакции приказов МГУ от 22 ноября 2011 года № 1066, от 21 декабря 2011 года № 1228, от 30 декабря 2011 года № 1289, от 30 июня 2016 года № 746).

Содержимое

1.	Место дисциплины в структуре ОПОП ВО	3
2.	Объем дисциплины	3
3.	Формат обучения	3
4.	Преподаватели	3
5.	Входные требования для освоения дисциплины	3
6.	Результаты обучения по дисциплине	3
7.	Содержание дисциплины	6
8.	Ресурсное обеспечение	7
8.1.	Список основной литературы	7
8.2.	Список дополнительной литературы (при наличии)	7
8.3.	Список программного обеспечения	7
8.4.	Список баз данных и информационных справочных систем	7
8.5.	Список ресурсов сети «Интернет»	7
8.6.	Материально-техническое обеспечение	7
9.	Фонд оценочных средств	7
9.1.	Текущий контроль успеваемости	7
9.2.	Промежуточная аттестация	8

1. Место дисциплины в структуре ОПОП ВО

Относится к вариативной части.

2. Объем дисциплины

Объем дисциплины (модуля) составляет 3 з.е., в том числе 36 академических часа, отведенных на контактную работу обучающихся с преподавателем, 72 академических часа на самостоятельную работу обучающихся.

3. Формат обучения

Очный с применением электронного обучения.

4. Преподаватели

Дисциплину ведут преподаватели кафедры теоретической информатики.

5. Входные требования для освоения дисциплины

Отсутствуют

6. Результаты обучения по дисциплине

Компетенции выпускников	Индикаторы достижения компетенций, реализуемые в настоящей дисциплине	Планируемые результаты обучения по дисциплине, соотнесенные с индикаторами достижения компетенций
Компетенция УК-1 Способность формулировать научно обоснованные гипотезы, создавать теоретические модели явлений и процессов, применять методологию научного познания в профессиональной деятельности.	Индикатор УК-1.1 Способен формулировать научно обоснованные гипотезы.	Уметь формулировать научно обоснованные гипотезы.
	Индикатор УК-1.2 Умеет создавать теоретические модели явлений и процессов, применять методологию научного познания в профессиональной деятельности.	Знать основные принципы математического моделирования. Уметь создавать теоретические модели явлений и процессов, применять методологию научного познания в профессиональной деятельности.
Компетенция ОПК-1 Готовность использовать фундаментальные знания в области математического	Индикатор ОПК-1.1. Способен использовать фундаментальные знания в области математического анализа, комплексного и функционального анализа, алгебры, линейной алгебры,	Знать основные понятия в области математического анализа, комплексного и функционального анализа, алгебры, линейной алгебры,

<p>анализа, комплексного и функционального анализа, алгебры, линейной алгебры, аналитической геометрии, дифференциальной геометрии и топологии, дифференциальных уравнений и уравнений в частных производных, дискретной математики, теории вероятностей, математической статистики и случайных процессов, численных методов, теоретической механики, механики сплошной среды, теории управления и оптимизации в будущей профессиональной деятельности</p>	<p>аналитической геометрии, дифференциальной геометрии и топологии, дифференциальных уравнений и уравнений в частных производных, дискретной математики, теории вероятностей, математической статистики и случайных процессов, численных методов, теоретической механики, механики сплошной среды, теории управления и оптимизации в будущей профессиональной деятельности</p>	<p>аналитической геометрии, дифференциальной геометрии и топологии, дифференциальных уравнений и уравнений в частных производных, дискретной математики, теории вероятностей, математической статистики и случайных процессов, численных методов, теоретической механики, механики сплошной среды, теории управления и оптимизации</p> <p>Уметь применять знания в области математического анализа, комплексного и функционального анализа, алгебры, линейной алгебры, аналитической геометрии, дифференциальной геометрии и топологии, дифференциальных уравнений и уравнений в частных производных, дискретной математики, теории вероятностей, математической статистики и случайных процессов, численных методов, теоретической механики, механики сплошной среды, теории управления и оптимизации</p> <p>Владеть основными методами математического анализа, комплексного и функционального анализа, алгебры, линейной алгебры, аналитической геометрии, дифференциальной геометрии и топологии, дифференциальных уравнений и уравнений в частных производных, дискретной математики, теории вероятностей, математической статистики и случайных процессов, численных методов, теоретической механики, механики сплошной среды, теории управления и оптимизации.</p>
<p>Компетенция ОПК-3. Способность находить, анализировать, реализовывать</p>	<p>Индикатор ОПК-3.1. Способен находить, анализировать, реализовывать программно и использовать на</p>	<p>Знать основные методы оценки применимости математических моделей и алгоритмов к решению задач.</p>

программно и использовать на практике математические алгоритмы, в том числе с применением современных вычислительных систем	практике математические алгоритмы, в том числе с применением современных вычислительных систем	Уметь реализовывать программно основные математические алгоритмы. Владеть основными методами анализа математических алгоритмов.
Компетенция ПК-1. Способность к самостоятельному анализу поставленной задачи, выбору корректного метода ее решения, построению алгоритма и его реализации, обработке и анализу полученной информации	Индикатор ПК-1.1. Способен анализировать поставленные задачи, выбирать и реализовывать методы решения задач математики и механики.	Знать основные методы решения задач, рассматриваемых в дисциплине. Уметь выбирать метод решения конкретной задачи с учетом ограничений на область применимости методов.
Компетенция ПК-5. Умение ориентироваться в современных методах и алгоритмах компьютерной математики	Индикатор ПК-5.1. Ориентируется в современных методах и алгоритмах компьютерной математики	Знать современные методы и алгоритмы компьютерной математики, их достоинства и недостатки, области применимости.
Компетенция СПК-1. Владение специальными разделами фундаментальной математики, методами анализа и решения задач специализации	Индикатор СПК-1.1. Владеет методами анализа и решения задач специального раздела фундаментальной математики	Знать: методы анализа и решения задач специального раздела фундаментальной математики Уметь: обоснованно выбирать методы решения задач специального раздела фундаментальной математики Владеть: техниками применения методов анализа и решения задач специального раздела фундаментальной математики

7. Содержание дисциплины

Учебный план дисциплины						
"Аналитика больших данных"						
№ п/п	Наименование разделов и тем дисциплины	Номинальные трудозатраты обучающегося				
		Всего ак. ч.	Контактная работа (работа во взаимодействии с преподавателем) Виды контактной работы, ак. ч.			Самостоят. работа, ак. ч.
			Ауд., ак. ч.	Лекции, ак. ч.	Семинары, практ., ак. ч.	
1	Устройство виртуальной машины Java.	4	2	2		2
2	Функциональное программирование (Scala).	4	2	2		2
3	Источники данных и работа с ними.	10	4	4		6
4	Инструменты для написания кода.	10	4	4		6
5	Apache Spark и обработка данных (Scala + Python).	8	2	2		6
6	Экосистема Hadoop.	8	2	2		6
7	Hadoop как хранилище данных.	8	2	2		6
8	Apache Airflow для оркестрации конвейеров.	8	2	2		6
9	Облачные хранилища.	4	2	2		2
10	Распределённые системы.	4	2	2		2
11	Архитектура обработки больших данных.	8	2	2		6
12	Алгоритмы параллельной обработки данных.	8	2	2		6
13	Программные платформы для систем больших данных.	8	2	2		6
14	Оборудование для построения ЦОД	10	4	4		6
15	Промежуточная аттестация - устный экзамен.	6	2	2		4
	Итого, ак. ч.	108	36	36	0	72
	Итого, з. е.	3				

8. Ресурсное обеспечение

8.1. Список основной литературы

1. Анналин, Ын., Теоретический минимум по Big Data. СПб.: Издательство: Питер, 2019, 208 с. ISBN: 978-5-4461-1040-7.
2. Базы данных в высокопроизводительных информационных системах : учебное пособие / авт.-сост. Е.И. Николаев ; Министерство образования и науки РФ, Федеральное государственное автономное образовательное учреждение высшего образования «Северо-Кавказский федеральный университет». - Ставрополь : СКФУ, 2016 - 163 с. : ил. - Библиогр.: с.161
3. Онлайн курс Big Data Fundamentals via PySpark, автор Upendra Kumar Devisetty Science Analyst at CyVerse <https://www.datacamp.com/courses/big-data-fundamentalsvia-pyspark>

8.2. Список дополнительной литературы (при наличии)

1. Minelli M., Chambers M., Dhiraj A. Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses. John Wiley & Sons, 2012.
2. Сенько, А. Работа с BigData в облаках. Обработка и хранение данных с примерами из Microsoft Azure. СПб: Издательство: Питер, 2018, 448 с. ISBN: 978-5-4461- 0578-6.
3. Чак Лэм, Hadoop в действии - ДМК Пресс, 2012

8.3. Список программного обеспечения

Не требуется

8.4. Список баз данных и информационных справочных систем

1. Du D.-Z., Ko K.-I. Theory of computational complexity. – John Wiley & Sons, 2000.
2. Sipser M. Introduction to the theory of computation. – Thomson Course Technology, 2nd ed., 2006.
3. Лавров, Игорь Андреевич. Математическая логика: учебное пособие для студентов высших учебных заведений, обучающихся по техническим и естественнонаучным специальностям / И. А. Лавров; под ред. Л. Л. Максимовой. Москва: Академия, 2006, 39, [1] с.: ил; 22 см. (Университетский учебник. Прикладная математика и информатика. ISBN 5-7695-2735-8.
4. Мендельсон, Э. Введение в математическую логику / Э. Мендельсон; под ред. С.И. Адян; пер. с англ. Ф.А. Кабакова. - Москва: Наука, 1971 - 320 с.: ил.; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=458257>.

8.5. Список ресурсов сети «Интернет»

1. <http://window.edu.ru/> Электронная библиотека учебно-методической литературы для общего и профессионального образования;
2. <http://elibrary.ru/> Научная электронная библиотека;
3. <http://www.scholar.ru/> Научные статьи, диссертации и авторефераты из электронных научных библиотек

8.6. Материально-техническое обеспечение

1. Презентационное оборудование (мультимедиапроектор, экран, компьютер для управления) для проведения лекционных занятий
2. Компьютерный класс (с выходом в Internet) для организации самостоятельной работы и организации практических занятий обучающихся

9. Фонд оценочных средств

Фонд оценочных средств (ФОС), оценочные и методические материалы) для оценивания результатов обучения по дисциплине (модулю).

9.1. Текущий контроль успеваемости

Типовые контрольные задания или иные материалы для проведения текущего контроля успеваемости, критерии и шкалы оценивания.

Примерные темы для опросов:

1. Устройство виртуальной машины Java.
2. Функциональное программирование (Scala).
3. Источники данных и работа с ними.
4. Инструменты для написания кода.
5. Apache Spark и обработка данных (Scala + Python).
6. Экосистема Hadoop.
7. Hadoop как хранилище данных.
8. Apache Airflow для оркестрации конвейеров.
9. Облачные хранилища.
10. Распределённые системы.
11. Архитектура обработки больших данных.
12. Алгоритмы параллельной обработки данных.
13. Программные платформы для систем больших данных.
14. Оборудование для построения ЦОД. Примеры контрольных работ

9.2. Промежуточная аттестация

Типовые контрольные задания или иные материалы для проведения промежуточной аттестации по дисциплине (модулю), критерии и шкалы оценивания.

Примерные вопросы для проведения экзамена:

1. Понятие Большие данные. Роль цифровой информации в 21 веке.
2. Виды массивов данных.
3. Базовые принципы обработки больших данных.
4. Технологии обработки больших данных: NoSQL, MapReduce, Hadoop, R.
5. Технологии Business Intelligence и реляционные системы управления базами данных.
6. Прогнозирование и предвидение: общее и особенное.
7. Виды прогнозов.
8. Опишите методики анализа больших данных.
9. Процесс аналитики анализа больших данных.
10. Дайте характеристику Big Data на мировом рынке.
11. Охарактеризуйте Big Data в России.
12. Определите понятие Data Mining.
13. Вопросы безопасности больших данных.
14. В чем состоит когнитивный анализ данных.
15. Какие модели данных вы знаете?
16. Основные описательные статистики.
17. Определите различия между параметрическими, непараметрическими и номинальными методами.
18. Опишите основную идею корреляционного анализа.
19. Регрессионный анализ.
20. Основная идея дисперсионного анализа.
21. Сущность кластерного анализа.
22. Дискриминантный анализ: модель и общая процедура выполнения.
23. Цели факторного анализа.
24. Программные средства анализа данных: Statistica, SPSS, Excel; их преимущества и недостатки.
25. Языки для Big Data.
26. Фреймворки для Big Data.
27. Базы данных для Big Data.
28. Аналитические платформы для Big Data.