

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Казанский национальный исследовательский технологический университет»
(ФГБОУ ВО «КНИТУ»)

УТВЕРЖДАЮ
Проректор по УР
А.В. Бурмистров
« 03 » 09 2019 г.

РАБОЧАЯ ПРОГРАММА

По дисциплине «Большие данные»
Направление подготовки 09.03.02 «Информационные системы и технологии»
Профиль подготовки Информационные системы и технологии
Квалификация выпускника бакалавр
Форма обучения очная
Институт, факультет Институт технологий легкой промышленности, моды и
дизайна, Факультет дизайна и программной инженерии
Кафедра-разработчик рабочей программы Информатики и прикладной
математики
Курс, семестр 3, 5

	Часы	Зачетные единицы
Лекции	18	0,5
Практические занятия		
Семинарские занятия		
Лабораторные занятия	36	1
Самостоятельная работа	90	2,5
Всего	144	4
Форма аттестации	Зачет с оценкой	

Казань, 2019 г.

Рабочая программа составлена с учетом требований Федерального государственного образовательного стандарта высшего образования № 926 от 19.09.2017 по направлению 09.03.02 «Информационные системы и технологии» для профиля «Информационные системы и технологии», на основании учебного плана набора обучающихся 2019 года.

Разработчики программы:

ст. преп. каф. ИПМ

ст. преп. каф. ИПМ




Шайдуллина Н. К.

Шайдуллин М. Р.

Рабочая программа рассмотрена и одобрена на заседании кафедры ИПМ 02.09.2019 г., протокол № 7.

Зав. кафедрой ИПМ, профессор



Н.К. Нуриев

УТВЕРЖДЕНО

Начальник УМЦ, доцент



Л.А. Китаева

1. Цели освоения дисциплины

Целями освоения дисциплины «Большие данные» являются

- а) формирование у студентов навыков разработки и использования систем обработки и анализа больших массивов данных;
- б) изучение специализированных программных систем, отвечающих за обработку больших данных.

2. Место дисциплины (модуля) в структуре образовательной программы

Дисциплина «Большие данные» относится к формируемой участниками образовательных отношений части ООП и формирует у бакалавров по направлению подготовки 09.03.02 «Информационные системы и технологии» набор знаний, умений, навыков и компетенций.

Для успешного освоения дисциплины бакалавр по направлению подготовки 09.03.02 «Информационные системы и технологии» должен освоить материал предшествующих дисциплин:

- а) информатика;
- б) информационные технологии;
- в) технологии программирования;
- г) дискретная математика.

Дисциплина «Большие данные» является предшествующей и необходима для успешного усвоения последующих дисциплин:

- а) введение в распределенные системы;
- б) управление информационными процессами;
- в) методы искусственного интеллекта

Знания, полученные при изучении дисциплины «Большие данные» могут быть использованы при прохождении практик и выполнении выпускной квалификационной работы.

3. Компетенции обучающегося и индикаторы достижения компетенции, формируемые в результате освоения дисциплины

ПК-3 Способность выполнять работы по обеспечению функционирования баз данных и обеспечению их информационной безопасности.

ПК-3.1 Знает теорию баз данных, основы программирования, возможности информационных систем, Инструменты и методы проектирования структур баз данных.

ПК-3.2 Умеет применять методы разграничения полномочий пользователей и управления доступом к ресурсам в защищенных операционных системах; разрабатывать структуру баз данных.

ПК-3.3 Владеет моделями защиты информационных систем; навыками разработки структуры баз данных информационных систем в соответствии с архитектурной спецификацией.

В результате освоения дисциплины обучающийся должен:

1) знать: теорию баз данных, основы программирования, возможности информационных систем, инструменты и методы проектирования структур баз данных.

2) уметь: применять методы разграничения полномочий пользователей и управления доступом к ресурсам в защищенных операционных системах; разрабатывать структуру баз данных.

3) владеть: моделями защиты информационных систем; навыками разработки структуры баз данных информационных систем в соответствии с архитектурной спецификацией.

4. Структура и содержание дисциплины «Большие данные»

Общая трудоемкость дисциплины составляет 4 зачетные единиц, 144 часа.

№ п/п	Раздел дисциплины	Семестр	Виды учебной работы (в часах)				Оценочные средства для проведения промежуточной аттестации по разделам
			Лекции	Семинар (Практические занятия, лабораторные практикумы)	Лабораторные работы	СРС	
1	Жизненный цикл анализа больших данных.	4	2			10	Тест
2	Корреляция и регрессия	4	2		4	10	Лабораторная работа, тест

3	Классификация и кластеризация данных	4	4	8	20	Лабораторная работа, тест
4	Готовые решения анализа данных	4	4	10	20	Лабораторная работа, тест
5	Подготовка данных		2	4	10	Лабораторная работа, тест
6	Парадигма Map Reduce.		4	10	20	Лабораторная работа, тест
	Итого		18	36	90	
Форма аттестации			Зачет с оценкой			

5. Содержание лекционных занятий по темам с указанием формируемых компетенций

№ п/п	Раздел дисциплины	Часы	Тема лекционного занятия	Краткое содержание	Индикаторы достижения компетенции
1	Жизненный цикл анализа больших данных.	2	Введение в большие данные. Жизненный цикл анализа больших данных.	Введение в большие данные. Жизненный цикл анализа больших данных. Песочница.	ПК-3.1 ПК-3.2 ПК-3.3
2	Корреляция и регрессия	2	Корреляция и регрессионный анализ.	Коэффициент корреляции. Графическое представление. Постановка задачи регрессионного анализа. Линейная регрессия. Метод наименьших квадратов. Их роль в аналитике больших данных.	ПК-3.1 ПК-3.2 ПК-3.3
3	Классификация и кластеризация данных	4	Задачи классификации и кластеризации.	Задачи классификации и кластеризации. Ассоциативные правила.	ПК-3.1 ПК-3.2 ПК-3.3
4	Готовые решения анализа данных	4	Языки Python и R.	Библиотеки Python. Синтаксис языка R, основные типы данных.	ПК-3.1 ПК-3.2 ПК-3.3
5	Подготовка данных	2	Подготовка данных.	Визуализация данных. Понимание данных.	ПК-3.1 ПК-3.2 ПК-3.3
6	Парадигма Map Reduce.	4	Парадигма Map Reduce.	Реализация Hadoop.	ПК-3.1 ПК-3.2 ПК-3.3

6. Содержание практических занятий

Учебным планом программы 09.03.02 проведение практических (семинарских) занятий по дисциплине «Большие данные» не предусмотрено.

7. Содержание лабораторных занятий

Цель проведения лабораторных занятий – формирование у студентов навыков разработки и использования систем обработки и анализа больших массивов данных; изучение специализированных программных систем, отвечающих за обработку больших данных.

Занятия проводятся в компьютерном классе.

№ п/п	Раздел дисциплины	Часы	Наименование лабораторной работы	Индикаторы достижения компетенции
1	Корреляция и регрессия	4	Задачи в области больших данных, решаемые методом регрессионного анализа.	ПК-3.1 ПК-3.2 ПК-3.3
2	Классификация и кластеризация данных	8	Постановка задачи классификации. Постановка задачи кластеризации. Задача построения ассоциативных правил.	ПК-3.1 ПК-3.2 ПК-3.3
3	Готовые решения анализа данных	10	Python и R: библиотеки, готовые решения.	ПК-3.1 ПК-3.2 ПК-3.3
4	Подготовка данных	4	Методы предварительной подготовки данных.	ПК-3.1 ПК-3.2 ПК-3.3
5	Парадигма Map Reduce.	10	Изучение и конфигурирование программного комплекса Apache Hadoop. Размещения набора данных по заданной тематике.	ПК-3.1 ПК-3.2 ПК-3.3

8. Самостоятельная работа бакалавра/магистранта/аспиранта

№ п/п	Темы, выносимые на самостоятельную работу	Часы	Форма СРС	Индикаторы достижения компетенции
1	Жизненный цикл анализа больших данных.	10	Изучение лекционного материала и рекомендуемой литературы.	ПК-3.1 ПК-3.2 ПК-3.3
2	Корреляция и регрессия	10	Изучение лекционного материала и рекомендуемой литературы, оформление отчета по лабораторной работе.	ПК-3.1 ПК-3.2 ПК-3.3
3	Классификация и кластеризация данных	20	Изучение лекционного материала и рекомендуемой литературы, оформление отчета по лабораторной работе.	ПК-3.1 ПК-3.2 ПК-3.3
4	Готовые решения анализа данных	20	Изучение лекционного материала и рекомендуемой литературы, оформление отчета по лабораторной работе.	ПК-3.1 ПК-3.2 ПК-3.3
5	Подготовка данных	10	Изучение лекционного материала и рекомендуемой литературы, оформление отчета по	ПК-3.1 ПК-3.2 ПК-3.3

			лабораторной работе.	
6	Парадигма Map Reduce.	20	Изучение лекционного материала и рекомендуемой литературы, оформление отчета по лабораторной работе.	ПК-3.1 ПК-3.2 ПК-3.3

9. Использование рейтинговой системы оценки знаний.

При оценке результатов деятельности студентов в рамках дисциплины «Большие данные» используется рейтинговая система. Рейтинговая оценка формируется на основании текущего и промежуточного контроля. Максимальное и минимальное количество баллов по различным видам учебной работы описано в положении о рейтинговой системе.

При изучении указанной дисциплины предусматривается выполнение 5 лабораторных работ, тестирования. В результате максимальный текущий рейтинг составит 60 баллов. За тест студент может получить максимальное кол-во баллов – 40.

Оценочные средства	Кол-во	Min, баллов	Max, баллов
Лабораторная работа	5	36	60
Тест	1	24	40
Итого:		60	100

10. Оценочные средства для текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины

Оценочные средства для проведения текущего контроля успеваемости, промежуточной аттестации обучающихся и итоговой (государственной итоговой) аттестации разрабатываются согласно положению о Фондах оценочных средств, рассматриваются как составная часть рабочей программы и оформляются отдельным документом.

11. Информационно-методическое обеспечение дисциплины

11.1 Основная литература

При изучении дисциплины «Большие данные» в качестве основных источников информации рекомендуется использовать следующую литературу.

Основные источники информации	Кол-во экз.
1. Дадян, Э. Г. Методы, модели, средства хранения и обработки данных: учебник / Э.Г. Дадян, Ю.А. Зеленков. — Москва: Вузовский учебник: ИНФРА-М, 2017. — 168 с. - ISBN 978-5-9558-0490-3.	ЭБС «Знаниум» https://znanium.com/catalog/product/543943 Доступ из любой точки интернета после регистрации с IP адресов КНИТУ
2. Форман, Д. Много цифр. Анализ больших данных при помощи Excel / Форман Д.; Пер. с англ. Соколовой А. - Москва :Альпина Пабли., 2016. - 461 с. ISBN 978-5-9614-5032-3.	ЭБС «Знаниум» https://znanium.com/catalog/product/551044 Доступ из любой точки интернета после регистрации с IP адресов КНИТУ
3. Соколов, Г. А. Введение в регрессионный анализ и планирование регрессионных экспериментов в экономике : учеб. пособие. — М. : ИНФРА-М, 2019. — 202 с. — (Высшее образование: Магистратура). - ISBN 978-5-16-003646-5.	ЭБС «Знаниум» https://znanium.com/catalog/product/1001125 Доступ из любой точки интернета после регистрации с IP адресов КНИТУ

11.2 Дополнительная литература

В качестве дополнительных источников информации рекомендуется использовать следующую литературу:

Дополнительные источники информации	Кол-во экз.
1. Полищук, Ю. В. Базы данных и их безопасность / Ю.В. Полищук, А.С. Боровский. - Москва : ИНФРА-М, 2020. - 210 с.	ЭБС «Знаниум» https://znanium.com/catalog/product/1011088 Доступ из любой точки интернета после регистрации с IP адресов КНИТУ
2. Основы проектирования баз данных: Учебное пособие / Голицына О.Л., Партыка Т.Л., Попов И.И., - 2-е изд. - Москва :Форум, НИЦ ИНФРА-М, 2016. - 416 с.	ЭБС «Знаниум» https://znanium.com/catalog/product/552969 Доступ из любой точки интернета после регистрации с IP адресов КНИТУ
3 Карпова Т.С. Базы данных. Модели, разработка, реализация [Электронный ресурс]/ Карпова Т.С.— Электрон. текстовые данные.— Москва: Интернет-Университет Информационных Технологий (ИНТУИТ), 2016.— 403 с.	ЭБС «IPRbooks» http://www.iprbookshop.ru/73728.html Доступ из любой точки интернета после регистрации с IP адресов КНИТУ

11.3 Электронные источники информации

При изучении дисциплины «Большие данные» использование электронных источников информации:

1. Электронный каталог УНИЦ КНИТУ – Режим доступа: <http://ruslan.kstu.ru/>
2. ЭБС «Знаниум» - Режим доступа: <https://znanium.com>
3. ЭБС «IPRbooks» - Режим доступа: <http://www.iprbookshop.ru>
4. «Открытое образование», Каталог курсов, MOOK: «Введение в инженерию больших данных» <https://openedu.ru/course/misis/ABD/>

Согласовано:

Зав. сектором ОКУФ



11.4. Современные профессиональные базы данных и информационные справочные системы.

1. Центральный журнал по математике «Zentralblatt MATH». – Доступ свободный: <https://zbmath.org/>.

2. Общероссийский портал Math-Net.Ru. – Доступ свободный: <http://www.mathnet.ru/>.

3. Сайт о программировании metanit.com. – Доступ свободный: <https://metanit.com/>.

12. Материально-техническое обеспечение дисциплины (модуля).

Учебные аудитории для проведения учебных занятий оснащены оборудованием:

1. парты,
2. стулья,
3. доска;

техническими средствами обучения:

1. проектор,
2. персональные компьютеры, с возможностью подключения к сети

«Интернет» и обеспечены доступом в электронную информационную среду КНИТУ.

Помещения для самостоятельной работы оснащены компьютерной техникой:

1. персональные компьютеры, с возможностью подключения к сети «Интернет» и обеспечены доступом в электронную информационную среду КНИТУ. Допускается замена оборудования его виртуальными аналогами.

Лицензированное свободно распространяемое программное обеспечение, используемое в учебном процессе при освоении дисциплины «Большие данные»:

1. Visual Studio 2015
2. Microsoft Office
3. Apache Hadoop

13. Образовательные технологии

Количество занятий, проводимых в интерактивных формах – 18 часов.

Основные интерактивные формы проведения учебных занятий:

- работа в малых группах;
- дискуссия;
- изучение и закрепление нового материала на интерактивной лекции;
- эвристическая беседа;
- системы дистанционного обучения.