

Министерство науки и высшего образования РФ
Федеральное государственное автономное образовательное учреждение
высшего образования
«СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»
РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Б1.В.05 Анализ больших данных

наименование дисциплины (модуля) в соответствии с учебным планом

Направление подготовки / специальность

01.04.02 Прикладная математика и информатика

Направленность (профиль)

01.04.02.08 Анализ данных и математическое моделирование

Форма обучения

очная

Год набора

2022

Красноярск 2022

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Программу составили _____

должность, инициалы, фамилия

1 Цели и задачи изучения дисциплины

1.1 Цель преподавания дисциплины

Ознакомление студентов с существующими методами и систематизация хранения, обработки и анализа больших данных

1.2 Задачи изучения дисциплины

Изучить принципы работы с большими данными, получить представление о компонентах экосистемы Hadoop и облачных платформах для реализации решения по BigData;

овладеть стандартными инструментами работы с большими данными

1.3 Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения образовательной программы

Код и наименование индикатора достижения компетенции	Запланированные результаты обучения по дисциплине
ПК-2: Способен управлять этапами жизненного цикла методологической и технологической инфраструктуры анализа данных в организации.	
ПК-2.1: Знает современные методы и инструментальные средства анализа данных; методы интерпретации и визуализации данных; источники информации, в том числе информации, необходимой для обеспечения деятельности в предметной области; теоретические основы информатики и исследования операций; методы извлечения информации и знаний из гетерогенных, мульти структурированных, неструктурированных источников, в том числе при потоковой обработке.	

<p>ПК-2.2: Умеет применять современные методы и инструментальные средства анализа данных; методы интерпретации и визуализации данных; умеет пользоваться источниками информации, в том числе источниками информации, необходимыми для обеспечения деятельности в</p>	
<p>предметной области.</p>	
<p>ПК-2.3: Способен анализировать и использовать современные методы и инструментальные средства анализа данных для решения практических и научных задач; способен применять методы интерпретации и визуализации данных для решения научных и практических задач.</p>	
<p>ПК-3: Способен управлять разработкой продуктов, услуг и решений на основе данных.</p>	
<p>ПК-3.1: Знает: состояние и перспективы развития информационных технологий, технологий данных в России и в мире; современные и перспективные методы сбора, хранения и передачи данных; источники данных, интенсивность генерации данных источниками; технические средства и среды сбора, хранения и обработки данных; современные и перспективные средства визуализации и интерпретации данных; исследование операций; машинное обучение; математическое моделирование; методы сравнительного анализа.</p>	

ПК-3.2: Способен проводить аналитические и поисковые исследования по тематике информационных технологий,	
технологий данных.	
ПК-4: Способен разрабатывать и внедрять новые методы и технологии исследования данных.	
ПК-4.1: Знает: состояние и перспективы развития информационных технологий, технологий данных в России и в мире; современные и перспективные методы сбора, хранения и передачи данных; источники данных, интенсивность генерации данных источниками; технические средства и среды сбора, хранения и обработки данных; современные и перспективные средства визуализации и интерпретации данных; исследование операций; машинное обучение; математическое моделирование; методы сравнительного анализа.	
ПК-4.2: Способен проводить аналитические и поисковые исследования по тематике информационных технологий, технологий данных.	

<p>ПК-4.3: Знает: состояние и перспективы развития информационных технологий, технологий данных в России и в мире; существующие и перспективные методы и программный инструментарий технологий данных; существующие и перспективные методы и программный инструментарий технологий данных; предметная область использования технологий больших данных; современный опыт использования технологий</p>	
<p>данных; математическое моделирование; методы сравнительного анализа методов, моделей, алгоритмов, технологий и инструментальных средств; существующие и перспективные математические методы и инструментальные средства анализа данных; существующие и перспективные методы обработки данных; методы машинного обучения; современные и перспективные методы сбора данных; источники данных, интенсивность генерации данных источниками; технические средства и среды сбора, хранения и обработки данных; современные и перспективные средства визуализации и интерпретации данных; применение технологий данных в практических задачах предметной области; показатели эффективности технологий данных.</p>	

ПК-4.4: Способен проводить аналитические исследования по тематике информационных технологий, технологий данных; выполнять научные исследования.	
УК-2: Способен управлять проектом на всех этапах его жизненного цикла.	
УК-2.1: Формулирует на основе поставленной проблемы проектную задачу и способ ее решения через реализацию проектного управления.	
УК-2.2: Разрабатывает концепцию проекта в рамках обо-значенной проблемы: формулирует цель, задачи, обосновывает актуальность, значимость, ожидаемые результаты и возможные сферы их применения.	
УК-2.3: Планирует необходимые ресурсы, в том числе с учетом их заменимости.	
УК-2.4: Разрабатывает план реализации проекта.	
УК-2.5: Осуществляет мониторинг хода реализации проекта, корректирует отклонения, вносит дополнительные изменения в план реализации проекта, уточняет зоны ответственности участников проекта.	
УК-3: Способен организовывать и руководить работой команды, вырабатывая командную стратегию для достижения поставленной цели.	
УК-3.1: Вырабатывает стратегию сотрудничества и на ее основе организует отбор членов команды для достижения поставленной цели.	
УК-3.2: Планирует и корректирует работу команды с учетом интересов, особенностей поведения и мнений ее членов.	

УК-3.3: Разрешает конфликты и противоречия при деловом общении на основе учета интересов всех сторон.	
УК-3.4: Организует дискуссии по заданной теме и обсуждение результатов работы команды с привлечением оппонентов разработанным идеям.	
УК-3.5: Планирует командную работу, распределяет поручения и делегирует полномочия членам команды.	

1.4 Особенности реализации дисциплины

Язык реализации дисциплины: Английский.

Дисциплина (модуль) реализуется без применения ЭО и ДОТ.

2. Объем дисциплины (модуля)

Вид учебной работы	Всего, зачетных единиц (акад. час)	е
		1
Контактная работа с преподавателем:	1 (36)	
занятия лекционного типа	0,5 (18)	
практические занятия	0,5 (18)	
Самостоятельная работа обучающихся:	2 (72)	
курсовое проектирование (КП)	Нет	
курсовая работа (КР)	Нет	

3 Содержание дисциплины (модуля)

3.1 Разделы дисциплины и виды занятий (тематический план занятий)

		Контактная работа, ак. час.							
№ п/п	Модули, темы (разделы) дисциплины	Занятия лекционного типа		Занятия семинарского типа				Самостоятельная работа, ак. час.	
				Семинары и/или Практические занятия		Лабораторные работы и/или Практикумы			
		Всего	В том числе в ЭИОС	Всего	В том числе в ЭИОС	Всего	В том числе в ЭИОС	Всего	В том числе в ЭИОС
1. Batch processing									
	1. Distributed file systems	2							
	2. MapReduce Computation Model	2							
	3. Beyond MapReduce, Spark	2							
	4. Distributed file systems			2					
	5. MapReduce Computation Model			2					
	6. Beyond MapReduce, Spark			2					
	7. Batch processing							24	
2. Stream Data Processing									
	1. Schemes for receiving and processing sweat data	2							
	2. Stream processing models and their characteristics	2							
	3. Streaming frameworks (open source Apache products)	2							
	4. Schemes for receiving and processing sweat data			2					
	5. Stream processing models and their characteristics			2					

6. Streaming frameworks (open source Apache products)			2					
7. Stream Data Processing							24	
3. Big data storage								
1. Types of data warehouses (BigTable-like, Dynamo-like, SQL over BigData)	2							
2. Architecture of data processing systems	4							
3. Types of data warehouses (BigTable-like, Dynamo-like, SQL over BigData)			2					
4. Architecture of data processing systems			4					
5. Big data storage							24	
Bcero	18		18				72	

4 Учебно-методическое обеспечение дисциплины

4.1 Печатные и электронные издания:

1. Efron B., Hastie T. Computer Age Statistical Inference: Algorithms, Evidence, and Data Science: monographs(Cambridge: Cambridge University Press).
2. Rajasekaran S., Reif J. Handbook of Parallel Computing. Models, Algorithms and Applications(London: Chapman & Hall/CRC).

4.2 Лицензионное и свободно распространяемое программное обеспечение, в том числе отечественного производства (программное обеспечение, на которое университет имеет лицензию, а также свободно распространяемое программное обеспечение):

1. Языки программирования PHP и Python

4.3 Интернет-ресурсы, включая профессиональные базы данных и информационные справочные системы:

1. • электронные информационно-справочные ресурсы научной библиотеки СФУ (<http://bik.sfu-kras.ru>)

5 Фонд оценочных средств

Оценочные средства находятся в приложении к рабочим программам дисциплин.

6 Материально-техническая база, необходимая для осуществления образовательного процесса по дисциплине (модулю)

Учебные лаборатории и классы, оснащенные современными компьютерами, объединенными в локальные вычислительные сети с выходом в Интернет, а также периферийным и проекционным оборудованием.