

Министерство науки и высшего образования Российской Федерации  
ФГБОУ ВО «БАЙКАЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

УТВЕРЖДАЮ  
Проректор по учебной работе  
д.э.н., доц. Бубнов В.А



22.06.2020г.

**Рабочая программа дисциплины**  
Б1.В.13. Анализ больших данных

Направление подготовки: 38.03.05 Бизнес-информатика  
Направленность (профиль): Цифровая экономика  
Квалификация выпускника: бакалавр  
Форма обучения: очная

Курс	4
Семестр	41
Лекции (час)	28
Практические (сем, лаб.) занятия (час)	42
Самостоятельная работа, включая подготовку к экзаменам и зачетам (час)	110
Курсовая работа (час)	
Всего часов	180
Зачет (семестр)	41
Экзамен (семестр)	

Иркутск 2020

Программа составлена в соответствии с ФГОС ВО по направлению 38.03.05  
Бизнес-информатика.

Автор В.В. Братищенко

Рабочая программа обсуждена и утверждена на заседании кафедры  
математических методов и цифровых технологий

Заведующий кафедрой А.В. Родионов

Дата актуализации рабочей программы: 30.06.2021

Дата актуализации рабочей программы: 30.06.2022

Дата актуализации рабочей программы: 30.06.2023

## 1. Цели изучения дисциплины

Целью изучения дисциплины является получение компетенций в области применения технологии больших данных.

Задачами изучения дисциплины являются

- знакомство с особенностями размещения и обработки больших данных,
- овладение методами применения структур хранения распределенных данных,
- изучение технологий обработки распределенных данных,
- изучение типовых моделей исследования больших данных.

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

### Компетенции обучающегося, формируемые в результате освоения дисциплины

Код компетенции по ФГОС ВО	Компетенция
ПК-3	выбор рациональных информационных систем и информационно-коммуникативных технологий решения для управления бизнесом
ПК-13	умение проектировать и внедрять компоненты ИТ-инфраструктуры предприятия, обеспечивающие достижение стратегических целей и поддержку бизнес-процессов
ПК-18	способность использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации по теме исследования

### Структура компетенции

Компетенция	Формируемые ЗУНы
ПК-3 выбор рациональных информационных систем и информационно-коммуникативных технологий решения для управления бизнесом	З. Знает теоретические основы выбора рациональных информационных систем и информационно-коммуникативных технологий решения для управления бизнесом У. Умеет выбирать рациональные информационные системы и информационно-коммуникативные технологии решения для управления бизнесом Н. Владеет навыками выбора рациональных информационных систем и информационно-коммуникативных технологий решения для управления бизнесом
ПК-13 умение проектировать и внедрять компоненты ИТ-инфраструктуры предприятия, обеспечивающие достижение стратегических целей и поддержку бизнес-процессов	З. Знает теоретические основы проектирования и внедрения компонентов ИТ-инфраструктуры предприятия, обеспечивающие достижение стратегических целей и поддержку бизнес-процессов У. Умеет проектировать и внедрять компоненты ИТ-инфраструктуры предприятия, обеспечивающие достижение стратегических целей и поддержку бизнес-процессов Н. Владеет навыками проектирования и внедрения компонентов ИТ-инфраструктуры предприятия, обеспечивающие достижение стратегических целей и поддержку бизнес-процессов

ПК-18 способность использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации по теме исследования	<p>З. Знает, как использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации по теме исследования</p> <p>У. Умеет использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации по теме исследования</p> <p>Н. Владеет навыками использования соответствующего математического аппарата и инструментальных средств для обработки, анализа и систематизации информации по теме исследования</p>
---	--

### 3. Место дисциплины (модуля) в структуре образовательной программы

Принадлежность дисциплины - БЛОК 1 ДИСЦИПЛИНЫ (МОДУЛИ): Вариативная часть.

Предшествующие дисциплины (освоение которых необходимо для успешного освоения данной): "Программирование", "Теория вероятностей и математическая статистика", "Интернет-технологии", "Распределенные системы"

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

Общая трудоемкость дисциплины составляет 5 зач. ед., 180 часов.

Вид учебной работы	Количество часов
Контактная(аудиторная) работа	
Лекции	28
Практические (сем, лаб.) занятия	42
Самостоятельная работа, включая подготовку к экзаменам и зачетам	110
Всего часов	180

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

#### 5.1. Содержание разделов дисциплины

№ п/п	Раздел и тема дисциплины	Семестр	Лекции	Семинар Лаборат. Практич.	Самостоят. раб.	В интерактивной форме	Формы текущего контроля успеваемости
1	Общие сведения о больших данных	41	2	4	12		Тест 1
2	Структуры распределенных данных	41	4	8	16		Тест 2. Задание 1. Объединение и анализ данных в Power Pivot
3	Технологии хранения и использования	41	4	4	12		Тест 3. Задание 5. Технология

№ п/п	Раздел и тема дисциплины	Семестр	Лекции	Семинар Лаборат. Практич.	Самостоят. раб.	В интерактивной форме	Формы текущего контроля успеваемости
	больших данных						MapReduce
4	Многомерный анализ данных (OLAP-технология)	41	4	6	12		Задание 2. Объединение и анализ данных в Power BI. Задание 3. Выбор данных в Интернете и анализ данных в Power BI. Задание 4. Облачные технологии многомерного анализа на примере Yandex Datalens
5	Распределенная обработка данных	41	6	8	26		Тест 4
6	Задачи и модели исследования зависимостей данных	41					Тест 5. Задание 6. Исследование и трансформация данных. Задание 8. Подготовка данных и решение задачи кластеризации. Задание 11. Анализ временных рядов. Задание 10. Исследование ассоциаций. Задание 12. Решение задачи регрессии. Задание 7. Решение задачи классификации. Задание 9. Сбор, предобработка данных. Решение задачи классификации текстов
6.1	Введение в искусственный интеллект и основные методы машинного обучения для работы с табличными данными	41	2	0	4		
6.2	Обучение с подкреплением	41	3	6	12		

№ п/п	Раздел и тема дисциплины	Семестр	Лекции	Семинар Лаборат. Практич.	Самостоят. раб.	В интерактивной форме	Формы текущего контроля успеваемости
6.3	Системы глубокого обучения	41	3	6	16		
	ИТОГО		28	42	110		

## 5.2. Лекционные занятия, их содержание

№ п/п	Наименование разделов и тем	Содержание
1	Общие сведения о больших данных	Области применения больших данных. Структурирование больших данных. Технологии анализа больших данных. Этапы анализа: выбор целей и задач анализа, сбор данных, подготовка данных, предварительное исследование данных, выбор задач и моделей исследования зависимостей данных, развертывание и внедрение
2	Структуры распределенных данных	Структурированные данные и неструктурированные данные. Данные на естественном языке. Машинные данные. Графовые, или сетевые, данные. Аудио, видео и графика. Поточковые данные
3	Многомерный анализ данных (OLAP-технология)	Программные средства интеграции данных из разных источников. Многомерный анализ данных
4	Технологии хранения и использования больших данных	Распределенные файловые системы. Базы данных. Инфраструктура распределенного программирования. Инфраструктура интеграции данных. Инфраструктуры машинного обучения. Безопасность
5	Распределенная обработка данных	MapReduce – модель распределенной обработки данных, предложенная компанией Google для обработки больших объемов данных на компьютерных кластерах Hadoop Distributed File System (HDFS) – распределённая файловая система, позволяющая хранить информацию практически неограниченного объёма. Hadoop YARN – фреймворк для управления ресурсами кластера и менеджмента задач, в том числе включает фреймворк MapReduce
6	Задачи и модели исследования зависимостей данных	Основные задачи систем искусственного интеллекта. Классификация, кластеризация, регрессия. Типы машинного обучения: с учителем, без учителя, с частичным привлечением учителя, обучение с подкреплением. Классификация на примере алгоритма k-ближайших соседей (kNN) Метрики оценки классификации: полнота, точность, F1, ROC, AUC. Валидационная и тестовая выборка. Кросс-валидация. Работа с категориальными признаками. Линейные модели для классификации. Перцептрон, логистическая регрессия, полносвязные нейронные сети, стохастический градиентный спуск и обратное распространение градиента. Алгоритмы, основанные на применении решающих деревьев. Критерии разделения узла: информационный выигрыш, критерий Джини. Ансамбли решающих деревьев: случайный лес, градиентный бустинг.

№ п/п	Наименование разделов и тем	Содержание
		Кластеризация. k-means, DBSCAN, агломеративная кластеризация. Метрики оценки кластеризации. Классификация и кластеризация текстов.

### 5.3. Семинарские, практические, лабораторные занятия, их содержание

№ раздела и темы	Содержание и формы проведения
1	Разработка архитектуры хранения и обработки распределенных данных. Выбор технологии хранения распределенных данных, предварительная обработка данных, определение целей и задач обработки данных, выбор задач и моделей исследования зависимостей, планирование архитектуры распределенной обработки данных
2	Создание структур распределенных данных. Разработка требований к системе распределенных данных. Предварительное исследование данных. Определение источников данных и технологии сбора данных. Выбор форм представления и форматов данных
3	Создание технологии сбора и хранения данных. Работа в распределенных файловых средах. Использование реляционных СУБД в распределенных средах. Базы данных «ключ-значение», документные базы данных, графовые базы данных
4	Многомерный анализ данных (OLAP-технология). Сбор данных из разных табличных источников и многомерный анализ данных. Применение Power Pivot, Power BI Desktop. Облачные технологии интеграции и многомерного анализа данных на примере Yandex DataLens
5	Распределенная обработка данных. Применение технологии MapReduce для решения задач обработки распределенных данных. Инструментальные среды проектирования распределенной обработкой
6	Применение задач и моделей исследования зависимостей. Исходные данные исследования зависимостей. Предварительная обработка исходных данных. Задачи и модели классификации. Использование и сравнение алгоритмов классификации: kNN, решающие деревья и их ансамбли, логистическая регрессия. Задачи и модели кластеризации. Задачи и модели ассоциаций. Задачи и модели анализа текстов

## 6. Фонд оценочных средств для проведения промежуточной аттестации по дисциплине (полный текст приведен в приложении к рабочей программе)

### 6.1. Текущий контроль

№ п/п	Этапы формирования компетенций (Тема из рабочей программы дисциплины)	Перечень формируемых компетенций по ФГОС ВО	(ЗУНы: З.1...З.п, У.1...У.п, Н.1...Н.п)	Контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы (Наименование оценочного средства)	Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания (по 100-балльной шкале)
1	1. Общие сведения о больших данных	ПК-3	З.Знает теоретические основы выбора рациональных информационных систем и информационно-коммуникативных технологий решения для управления бизнесом	Тест 1	Доля правильных ответов (6)
2	2. Структуры распределенных данных	ПК-13	З.Знает теоретические основы проектирования и внедрения компонентов ИТ-инфраструктуры предприятия, обеспечивающие достижение стратегических целей и поддержку бизнес-процессов У.Умеет проектировать и внедрять компоненты ИТ-инфраструктуры предприятия, обеспечивающие достижение стратегических целей и поддержку бизнес-процессов Н.Владеет навыками проектирования и внедрения компонентов ИТ-инфраструктуры предприятия, обеспечивающие достижение стратегических целей и поддержку бизнес-процессов	Задание 1. Объединение и анализ данных в Power Pivot	Корректность обработки (5)
3		ПК-3	З.Знает теоретические основы выбора рациональных информационных систем и информационно-коммуникативных технологий решения для управления	Тест 2	Доля правильных ответов (6)



№ п/п	Этапы формирования компетенций (Тема из рабочей программы дисциплины)	Перечень формируемых компетенций по ФГОС ВО	(ЗУНы: (З.1...З.п, У.1...У.п, Н.1...Н.п))	Контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы (Наименование оценочного средства)	Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания (по 100-балльной шкале)
			бизнесом		
4	3. Технологии хранения и использования больших данных	ПК-3	<p>З.Знает теоретические основы выбора рациональных информационных систем и информационно-коммуникативных технологий решения для управления бизнесом</p> <p>У.Умеет выбирать рациональные информационные системы и информационно-коммуникативные технологии решения для управления бизнесом</p> <p>Н.Владеет навыками выбора рациональных информационных систем и информационно-коммуникативных технологий решения для управления бизнесом</p>	Задание 5. Технология MapReduce	Эффективность предложенных решений (7)
5		ПК-13	<p>З.Знает теоретические основы проектирования и внедрения компонентов ИТ-инфраструктуры предприятия, обеспечивающие достижение стратегических целей и поддержку бизнес-процессов</p>	Тест 3	Доля правильных ответов (6)
6	4. Многомерный анализ данных (OLAP-технология)	ПК-18	<p>З.Знает, как использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации по теме исследования</p>	Задание 2. Объединение и анализ данных в Power BI	Корректность и полнота анализа (5)

№ п/п	Этапы формирования компетенций (Тема из рабочей программы дисциплины)	Перечень формируемых компетенций по ФГОС ВО	(ЗУНы: (З.1...З.п, У.1...У.п, Н.1...Н.п))	Контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы (Наименование оценочного средства)	Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания (по 100-балльной шкале)
			У. Умеет использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации по теме исследования Н. Владеет навыками использования соответствующего математического аппарата и инструментальных средств для обработки, анализа и систематизации информации по теме исследования		
7		ПК-18	З. Знает, как использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации по теме исследования У. Умеет использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации по теме исследования Н. Владеет навыками использования соответствующего математического аппарата и инструментальных средств для обработки, анализа и систематизации информации по теме	Задание 3. Выбор данных в Интернете и анализ данных в Power BI	Корректность анализа и применения инструментов визуализации (6)

№ п/п	Этапы формирования компетенций (Тема из рабочей программы дисциплины)	Перечень формируемых компетенций по ФГОС ВО	(ЗУНы: (З.1...З.п, У.1...У.п, Н.1...Н.п)	Контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы (Наименование оценочного средства)	Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания (по 100-балльной шкале)
			исследования		
8		ПК-13	З.Знает теоретические основы проектирования и внедрения компонентов ИТ-инфраструктуры предприятия, обеспечивающие достижение стратегических целей и поддержку бизнес-процессов У.Умеет проектировать и внедрять компоненты ИТ-инфраструктуры предприятия, обеспечивающие достижение стратегических целей и поддержку бизнес-процессов Н.Владеет навыками проектирования и внедрения компонентов ИТ-инфраструктуры предприятия, обеспечивающие достижение стратегических целей и поддержку бизнес-процессов	Задание 4. Облачные технологии многомерного анализа на примере Yandex Datalens	Корректность и полнота анализа (6)
9	5. Распределенная обработка данных	ПК-3	З.Знает теоретические основы выбора рациональных информационных систем и информационно-коммуникативных технологий решения для управления бизнесом	Тест 4	Доля правильных ответов (6)
10	6. Задачи и модели исследования зависимостей данных	ПК-18	З.Знает, как использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и	Задание 6. Исследование и трансформация данных	Корректность и обоснованность трансформации (6)

№ п/п	Этапы формирования компетенций (Тема из рабочей программы дисциплины)	Перечень формируемых компетенций по ФГОС ВО	(ЗУНы: (З.1...З.п, У.1...У.п, Н.1...Н.п)	Контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы (Наименование оценочного средства)	Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания (по 100-балльной шкале)
			<p>систематизации информации по теме исследования У. Умеет использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации по теме исследования Н. Владеет навыками использования соответствующего математического аппарата и инструментальных средств для обработки, анализа и систематизации информации по теме исследования</p>		
11		ПК-13	<p>З. Знает теоретические основы проектирования и внедрения компонентов ИТ-инфраструктуры предприятия, обеспечивающие достижение стратегических целей и поддержку бизнес-процессов У. Умеет проектировать и внедрять компоненты ИТ-инфраструктуры предприятия, обеспечивающие достижение стратегических целей и поддержку бизнес-процессов Н. Владеет навыками проектирования и внедрения компонентов ИТ-инфраструктуры предприятия,</p>	Задание 7. Решение задачи классификации 1) Сбор, предобработка данных для классификации	Полнота решения (3)

№ п/п	Этапы формирования компетенций (Тема из рабочей программы дисциплины)	Перечень формируемых компетенций по ФГОС ВО	(ЗУНы: (З.1...З.п, У.1...У.п, Н.1...Н.п)	Контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы (Наименование оценочного средства)	Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания (по 100-балльной шкале)
			обеспечивающие достижение стратегических целей и поддержку бизнес-процессов		
12		ПК-18	З.Знает, как использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации по теме исследования У.Умеет использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации по теме исследования Н.Владеет навыками использования соответствующего математического аппарата и инструментальных средств для обработки, анализа и систематизации информации по теме исследования	Задание 7. Решение задачи классификации 2) Решение задачи классификации	Корректность применения моделей (4)
13		ПК-18	З.Знает, как использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации по теме исследования У.Умеет использовать соответствующий математический аппарат и	Задание 8. Подготовка данных и решение задачи кластеризации	Правильность оценки результатов кластеризации (7)

№ п/п	Этапы формирования компетенций (Тема из рабочей программы дисциплины)	Перечень формируемых компетенций по ФГОС ВО	(ЗУНы: (З.1...З.п, У.1...У.п, Н.1...Н.п)	Контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы (Наименование оценочного средства)	Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания (по 100-балльной шкале)
			инструментальные средства для обработки, анализа и систематизации информации по теме исследования Н. Владеет навыками использования соответствующего математического аппарата и инструментальных средств для обработки, анализа и систематизации информации по теме исследования		
14		ПК-13	З. Знает теоретические основы проектирования и внедрения компонентов ИТ-инфраструктуры предприятия, обеспечивающие достижение стратегических целей и поддержку бизнес-процессов У. Умеет проектировать и внедрять компоненты ИТ-инфраструктуры предприятия, обеспечивающие достижение стратегических целей и поддержку бизнес-процессов Н. Владеет навыками проектирования и внедрения компонентов ИТ-инфраструктуры предприятия, обеспечивающие достижение стратегических целей и поддержку бизнес-процессов	Задание 9. Сбор, предобработка данных. Решение задачи классификации текстов 1) Сбор, предобработка текстов	Полнота решения (3)
15		ПК-18	З. Знает, как использовать	Задание 9. Сбор, предобработка	Корректность применения

№ п/п	Этапы формирования компетенций (Тема из рабочей программы дисциплины)	Перечень формируемых компетенций по ФГОС ВО	(ЗУНы: З.1...З.п, У.1...У.п, Н.1...Н.п)	Контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы (Наименование оценочного средства)	Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания (по 100-балльной шкале)
			соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации по теме исследования У. Умеет использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации по теме исследования Н. Владеет навыками использования соответствующего математического аппарата и инструментальных средств для обработки, анализа и систематизации информации по теме исследования	данных. Решение задачи классификации текстов 2) Решение задачи классификации текстов	моделей (4)
16		ПК-13	З. Знает теоретические основы проектирования и внедрения компонентов ИТ-инфраструктуры предприятия, обеспечивающие достижение стратегических целей и поддержку бизнес-процессов У. Умеет проектировать и внедрять компоненты ИТ-инфраструктуры предприятия, обеспечивающие достижение стратегических целей и поддержку бизнес-процессов	Задание 10. Исследование ассоциаций	Корректность интерпретации результатов моделирования (5)

№ п/п	Этапы формирования компетенций (Тема из рабочей программы дисциплины)	Перечень формируемых компетенций по ФГОС ВО	(ЗУНы: (З.1...З.п, У.1...У.п, Н.1...Н.п)	Контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы (Наименование оценочного средства)	Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания (по 100-балльной шкале)
			Н. Владеет навыками проектирования и внедрения компонентов ИТ-инфраструктуры предприятия, обеспечивающие достижение стратегических целей и поддержку бизнес-процессов		
17		ПК-18	З. Знает, как использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации по теме исследования У. Умеет использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации по теме исследования Н. Владеет навыками использования соответствующего математического аппарата и инструментальных средств для обработки, анализа и систематизации информации по теме исследования	Задание 11. Анализ временных рядов	Умение применять модели временных рядов (5)
18		ПК-18	З. Знает, как использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации	Задание 12. Решение задачи регрессии	Обоснованность выбора модели (4)



№ п/п	Этапы формирования компетенций (Тема из рабочей программы дисциплины)	Перечень формируемых компетенций по ФГОС ВО	(ЗУНы: (З.1...З.п, У.1...У.п, Н.1...Н.п)	Контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы (Наименование оценочного средства)	Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания (по 100-балльной шкале)
			информации по теме исследования У. Умеет использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации по теме исследования Н. Владеет навыками использования соответствующего математического аппарата и инструментальных средств для обработки, анализа и систематизации информации по теме исследования		
19		ПК-18	З. Знает, как использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации по теме исследования	Тест 5	Доля правильных ответов (6)
				<b>Итого</b>	<b>100</b>

## 6.2. Промежуточный контроль (зачет, экзамен)

Рабочим учебным планом предусмотрен Зачет в семестре 41.

### ВОПРОСЫ ДЛЯ ПРОВЕРКИ ЗНАНИЙ:

1-й вопрос билета (30 баллов), вид вопроса: Тест/проверка знаний. Критерий: Доля правильных ответов.

**Компетенция: ПК-13 умение проектировать и внедрять компоненты ИТ-инфраструктуры предприятия, обеспечивающие достижение стратегических целей и поддержку бизнес-процессов**

**Знание:** Знает теоретические основы проектирования и внедрения компонентов ИТ-инфраструктуры предприятия, обеспечивающие достижение стратегических целей и поддержку бизнес-процессов

1. Базы данных NoSQL. Особенности, классификация
2. Понятие Больших данных. Особенности сбора, хранения, обработки и анализа Больших данных
3. Структурированные данные и неструктурированные данные. Данные на естественном языке. Машинные данные. Графовые, или сетевые, данные. Аудио, видео и графика. Поточковые данные
4. Требования к распределенным информационным системам

**Компетенция: ПК-18 способность использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации по теме исследования**

**Знание:** Знает, как использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации по теме исследования

5. Распределенная обработка и модели для решения задач анализа текстов
6. Распределенная обработка и модели для решения задач классификации
7. Распределенная обработка и модели для решения задач кластеризации

**Компетенция: ПК-3 выбор рациональных информационных систем и информационно-коммуникативных технологий решения для управления бизнесом**

**Знание:** Знает теоретические основы выбора рациональных информационных систем и информационно-коммуникативных технологий решения для управления бизнесом

8. Документо-ориентированные базы данных. Возможности СУБД MongoDB
9. Работа с документо-ориентированными БД.
10. Система Apache Hadoop
11. Средства построения распределенных информационных систем
12. Технология Map-Reduce

#### ТИПОВЫЕ ЗАДАНИЯ ДЛЯ ПРОВЕРКИ УМЕНИЙ:

2-й вопрос билета (30 баллов), вид вопроса: Задание на умение. Критерий: Корректность и эффективность решения.

**Компетенция: ПК-13 умение проектировать и внедрять компоненты ИТ-инфраструктуры предприятия, обеспечивающие достижение стратегических целей и поддержку бизнес-процессов**

**Умение:** Умеет проектировать и внедрять компоненты ИТ-инфраструктуры предприятия, обеспечивающие достижение стратегических целей и поддержку бизнес-процессов

**Задача № 1.** Определить набор коллекций документов MongoDB для предложенной предметной области

**Компетенция: ПК-18 способность использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации по теме исследования**

**Умение:** Умеет использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации по теме исследования

**Задача № 2.** Построить модель и обучить ее по предложенному набору данных

**Компетенция: ПК-3 выбор рациональных информационных систем и информационно-коммуникативных технологий решения для управления бизнесом**

**Умение:** Умеет выбирать рациональные информационные системы и информационно-коммуникативные технологии решения для управления бизнесом

Задача № 3. Определить технологию Map-Reduce для решения предложенной задачи обработки данных

#### ТИПОВЫЕ ЗАДАНИЯ ДЛЯ ПРОВЕРКИ НАВЫКОВ:

3-й вопрос билета (40 баллов), вид вопроса: Задание на навыки. Критерий: Корректность и эффективность решения.

**Компетенция: ПК-13 умение проектировать и внедрять компоненты ИТ-инфраструктуры предприятия, обеспечивающие достижение стратегических целей и поддержку бизнес-процессов**

Навык: Владеет навыками проектирования и внедрения компонентов ИТ-инфраструктуры предприятия, обеспечивающие достижение стратегических целей и поддержку бизнес-процессов

Задание № 1. Для предложенной предметной области предложить архитектуру ИТ-инфраструктуры для решения задач анализа данных

**Компетенция: ПК-18 способность использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации по теме исследования**

Навык: Владеет навыками использования соответствующего математического аппарата и инструментальных средств для обработки, анализа и систематизации информации по теме исследования

Задание № 2. Для предложенной предметной области выполнить анализ тенденций, закономерностей и интерпретацию результатов моделирования на основании подобранной модели

**Компетенция: ПК-3 выбор рациональных информационных систем и информационно-коммуникативных технологий решения для управления бизнесом**

Навык: Владеет навыками выбора рациональных информационных систем и информационно-коммуникативных технологий решения для управления бизнесом

Задание № 3. Для предложенной предметной области разработать схему сбора и хранения данных, продумать доступ к данным для различных приложений

#### ОБРАЗЕЦ БИЛЕТА

Министерство науки и высшего образования  
Российской Федерации  
Федеральное государственное бюджетное  
образовательное учреждение  
высшего образования  
**«БАЙКАЛЬСКИЙ ГОСУДАРСТВЕННЫЙ  
УНИВЕРСИТЕТ»  
(ФГБОУ ВО «БГУ»)**

Направление - 38.03.05 Бизнес-  
информатика  
Профиль - Цифровая экономика  
Кафедра математических методов и  
цифровых технологий  
Дисциплина - Анализ больших данных

#### БИЛЕТ № 1

1. Тест (30 баллов).
2. Построить модель и обучить ее по предложенному набору данных (30 баллов).
3. Для предложенной предметной области выполнить анализ тенденций, закономерностей и интерпретацию результатов моделирования на основании подобранной модели (40 баллов).

Составитель \_\_\_\_\_ В.В. Братищенко

## **7. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины (модуля)**

### **а) основная литература:**

1. Воронова Л.И., Воронов В.И. Big Data. Методы и средства анализа. учебное пособие. Электронный ресурс/ В.И. Воронов.- Москва: Московский технический университет связи и информатики, 2016.-33 с.
2. Чубукова И. А. Data Mining. учеб. пособие/ И. А. Чубукова.- М.: БИНОМ. Лаборатория знаний, 2006.-382 с.
3. Билл Фрэнкс Революция в аналитике [Электронный ресурс] : как в эпоху Big Data улучшить ваш бизнес с помощью операционной аналитики / Фрэнкс Билл. — Электрон. текстовые данные. — М. : Альпина Паблишер, 2017. — 320 с. — 978-5-9614-5302-7. — Режим доступа: <http://www.iprbookshop.ru/58563.html>
4. Воронов, В. И. Data Mining - технологии обработки больших данных : учебное пособие / В. И. Воронов, Л. И. Воронова, В. А. Усачев. — Москва : Московский технический университет связи и информатики, 2018. — 47 с. — ISBN 2227-8397. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/81324.html>
5. Воронова Л.И. Big Data. Методы и средства анализа [Электронный ресурс] : учебное пособие / Л.И. Воронова, В.И. Воронов. — Электрон. текстовые данные. — М. : Московский технический университет связи и информатики, 2016. — 33 с. — 2227-8397. — Режим доступа: <http://www.iprbookshop.ru/61463.html>
6. Чубукова И.А. Data Mining [Электронный ресурс] / И.А. Чубукова. — Электрон. текстовые данные. — М. : Интернет-Университет Информационных Технологий (ИНТУИТ), 2016. — 470 с. — 978-5-94774-819-2. — Режим доступа: <http://www.iprbookshop.ru/56315.html>

### **б) дополнительная литература:**

1. Технологии анализа данных: Data Mining, Visual Mining, Text Mining, OLAP. учеб. пособие [для вузов]. рек. УМО вузов по унив. политехн. образованию. 2-е изд., перераб. и доп./ А. А. Барсегян [и др.].- СПб.: БХВ-Петербург, 2008.-375 с.
2. Билл, Фрэнкс Революция в аналитике : как в эпоху Big Data улучшить ваш бизнес с помощью операционной аналитики / Фрэнкс Билл ; перевод И. Евстигнеева ; под редакцией В. Мылов. — Москва : Альпина Паблишер, 2017. — 320 с. — ISBN 978-5-9614-5302-7. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/58563.html>
3. Кухаренко Б.Г. Интеллектуальные системы и технологии [Электронный ресурс] : учебное пособие / Б.Г. Кухаренко. — Электрон. текстовые данные. — М. : Московская государственная академия водного транспорта, 2015. — 116 с. — 2227-8397. — Режим доступа: <http://www.iprbookshop.ru/47933.html>

## **8. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины (модуля), включая профессиональные базы данных и информационно-справочные системы**

Для освоения дисциплины обучающемуся необходимы следующие ресурсы информационно-телекоммуникационной сети «Интернет»:

– Сайт Байкальского государственного университета, адрес доступа: <http://bgu.ru/>, доступ круглосуточный неограниченный из любой точки Интернет

## **9. Методические указания для обучающихся по освоению дисциплины (модуля)**

Изучать дисциплину рекомендуется в соответствии с той последовательностью, которая обозначена в ее содержании. Для успешного освоения курса обучающиеся должны иметь первоначальные знания в области информационных технологий, баз данных, интернет-технологий.

На лекциях преподаватель озвучивает тему, знакомит с перечнем литературы по теме, обосновывает место и роль этой темы в данной дисциплине, раскрывает ее практическое значение. В ходе лекций студенту необходимо вести конспект, фиксируя основные понятия и проблемные вопросы.

Практические (семинарские) занятия по своему содержанию связаны с тематикой лекционных занятий. Начинать подготовку к занятию целесообразно с конспекта лекций. Задание на практическое (семинарское) занятие сообщается обучающимся до его проведения. На семинаре преподаватель организует обсуждение этой темы, выступая в качестве организатора, консультанта и эксперта учебно-познавательной деятельности обучающегося.

Изучение дисциплины (модуля) включает самостоятельную работу обучающегося.

Основными видами самостоятельной работы студентов с участием преподавателей являются:

- текущие консультации;
- коллоквиум как форма контроля освоения теоретического содержания дисциплин: (в часы консультаций, предусмотренные учебным планом);
- прием и разбор домашних заданий (в часы практических занятий);
- прием и защита лабораторных работ (во время проведения занятий);
- выполнение курсовых работ в рамках дисциплин (руководство, консультирование и защита курсовых работ в часы, предусмотренные учебным планом) и др.

Основными видами самостоятельной работы студентов без участия преподавателей являются:

- формирование и усвоение содержания конспекта лекций на базе рекомендованной лектором учебной литературы, включая информационные образовательные ресурсы (электронные учебники, электронные библиотеки и др.);
- самостоятельное изучение отдельных тем или вопросов по учебникам или учебным пособиям;
- написание рефератов, докладов;
- подготовка к семинарам и лабораторным работам;
- выполнение домашних заданий в виде решения отдельных задач, проведения типовых расчетов, расчетно-компьютерных и индивидуальных работ по отдельным разделам содержания дисциплин и др.

## **10. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения**

В учебном процессе используется следующее программное обеспечение:

- Python,
- Scilab,
- Модули: matplotlib, Mlxtend, PrefixSpan, NumPy, pandas, scikit-learn, SciPy, TensorFlow, языка Python,
- Модули: pymongo, PyMySQL, PyQt, SQL Alchemy языка Python,
- Надстройка интеллектуального анализа данных для MS Office,
- MongoDB,
- MS SQL Server и программы администрирования,
- ActivePython x64,

- Sql datamining,

#### **11. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю):**

В учебном процессе используется следующее оборудование:

- Помещения для самостоятельной работы, оснащенные компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду вуза,
- Учебные аудитории для проведения: занятий лекционного типа, занятий семинарского типа, практических занятий, выполнения курсовых работ, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, укомплектованные специализированной мебелью и техническими средствами обучения,
- Мультимедийный класс,
- Компьютерный класс,
- Наборы демонстрационного оборудования и учебно-наглядных пособий