Министерство науки и высшего образования Российской Федерации

ФГБОУ ВО «БАЙКАЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

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

_____

25.06.2021г.

Рабочая программа дисциплины
Б1.У.16. 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)	28
Самостоятельная работа, включая подготовку к экзаменам и зачетам (час)	124
Курсовая работа (час)	
Всего часов	180
Зачет (семестр)	41
Экзамен (семестр)	

Иркутск 2021

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: 30.06.2022
Дата актуализации рабочей программы: 30.06.2023
Дата актуализации рабочей программы: 28.06.2024

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-9	Способен проводить анализ данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта

Структура компетенции

Компетенция	Формируемые ЗУНы
ПК-9 Способен проводить анализ данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта	З. Знать способы и методы анализа данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта У. Уметь проводить анализ данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта Н. Владеть навыками анализа данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта

3. Место дисциплины (модуля) в структуре образовательной программы

Принадлежность дисциплины - БЛОК 1 ДИСЦИПЛИНЫ (МОДУЛИ): Часть, формируемая участниками образовательных отношений.

Предшествующие дисциплины (освоение которых необходимо для успешного освоения данной): "Программирование", "Теория вероятностей и математическая статистика", "Операционные системы", "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	28

Самостоятельная работа, включая подготовку к экзаменам и зачетам	124
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	Семе-стр	Лек-ции	Семинар Лаборат. Практич.	Само-стоят. раб.	В интера-ктивной форме	Формы текущего контроля успеваемости
1	Общие сведения о больших данных	41	2	4	12		Тест 1
2	Структуры распределенных данных	41	4	4	16		Задание 1. Объединение и анализ данных в Power Pivot. Тест 2
3	Технологии хранения и использования больших данных	41	4	4	18		Задание 6. Технология MapReduce. Тест 3
4	Многомерный анализ данных (OLAP-технология)	41	4	4	18		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. Задание 5. Yandex Datalens ABC-XYZ анализ
5	Распределенная обработка данных	41	6	4	26		Тест 4
6	Задачи и модели исследования зависимостей данных	41					Задание 7. Исследование и трансформация данных. Задание 8. Решение задачи классификации. Задание 9. Решение задачи регрессии. Задание 10. Подготовка данных и решение задачи

№ п/п	Раздел и тема дисциплины	Семе-стр	Лек-ции	Семинар Лаборат. Практич.	Само-стоят. раб.	В интера-ктивной форме	Формы текущего контроля успеваемости
							кластеризации. Задание 11. Исследование ассоциаций. Задание 12. Анализ временных рядов. Задание 13. Сбор, предобработка данных. Решение задачи классификации текстов. Тест 5
6.1	Введение в искусственный интеллект и основные методы машинного обучения для работы с табличными данными	41	2	0	6		
6.2	Обучение с подкреплением	41	3	4	12		
6.3	Системы глубокого обучения	41	3	4	16		
	ИТОГО		28	28	124		

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. Потоковые 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 объёмов данных на компьютерных кластерах Hadoop Distributed File System (HDFS) – распределённая

№ п/п	Наименование разделов и тем	Содержание
		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. 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 Piwot, 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

№ раздела и темы	Содержание и формы проведения
	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.n, У.1…У.n, Н.1…Н.n)	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-9	З.Знать способы и методы анализа данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта	Тест 1	Доля правильных ответов (6)
2	2. Структуры распределенных данных	ПК-9	З.Знать способы и методы анализа данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта У.Уметь проводить анализ данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта Н.Владеть навыками анализа данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта	Задание 1. Объединение и анализ данных в Power Pivot	Корректность обработки (7)
3		ПК-9	З.Знать способы и методы анализа данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта	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.n, У.1…У.n, Н.1…Н.n)	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)
4	3. Технологии хранения и использования больших данных	ПК-9	З.Знать способы и методы анализа данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта У.Уметь проводить анализ данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта Н.Владеть навыками анализа данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта	Задание 6. Технология MapReduce	Эффективность предложенных решений (7)
5		ПК-9	З.Знать способы и методы анализа данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта	Тест 3	Доля правильных ответов (6)
6	4. Многомерный анализ данных (OLAP-технология)	ПК-9	З.Знать способы и методы анализа данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта У.Уметь проводить анализ данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта Н.Владеть навыками анализа данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта	Задание 2. Объединение и анализ данных в Power BI	Корректность и полнота анализа (7)
7		ПК-9	З.Знать способы и методы анализа данных по созданию,	Задание 3. Выбор данных в Интернете и анализ данных в	Корректность анализа и применения

№ п/п	Этапы формирования компетенций (Тема из рабочей программы дисциплины)	Перечень формируемых компетенций по ФГОС ВО	(ЗУНы: (З.1…З.n, У.1…У.n, Н.1…Н.n)	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)
			развитию, выводу на рынок, продажам и продвижению ИТ-продукта У.Уметь проводить анализ данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта Н.Владеть навыками анализа данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта	Power BI	инструментов визуализации (7)
8		ПК-9	З.Знать способы и методы анализа данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта У.Уметь проводить анализ данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта Н.Владеть навыками анализа данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта	Задание 4. Облачные технологии многомерного анализа на примере Yandex Datalens	Корректность и полнота анализа (7)
9		ПК-9	З.Знать способы и методы анализа данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта У.Уметь проводить анализ данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-	Задание 5. Yandex Datalens ABC-XYZ анализ	Полнота анализа данных (7)

№ п/п	Этапы формирования компетенций (Тема из рабочей программы дисциплины)	Перечень формируемых компетенций по ФГОС ВО	(ЗУНы: (З.1…З.n, У.1…У.n, Н.1…Н.n)	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)
			продукта Н.Владеть навыками анализа данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта		
10	5. Распределенная обработка данных	ПК-9	З.Знать способы и методы анализа данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта	Тест 4	Доля правильных ответов (6)
11	6. Задачи и модели исследования зависимостей данных	ПК-9	З.Знать способы и методы анализа данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта У.Уметь проводить анализ данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта Н.Владеть навыками анализа данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта	Задание 7. Исследование и трансформация данных	Корректность и обоснованность трансформации (7)
12		ПК-9	З.Знать способы и методы анализа данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта У.Уметь проводить анализ данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта Н.Владеть навыками анализа данных по	Задание 8. Решение задачи классификации	Корректность применения моделей (7)

№ п/п	Этапы формирования компетенций (Тема из рабочей программы дисциплины)	Перечень формируемых компетенций по ФГОС ВО	(ЗУНы: (З.1…З.n, У.1…У.n, Н.1…Н.n)	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)
			созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта		
13		ПК-9	З.Знать способы и методы анализа данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта У.Уметь проводить анализ данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта Н.Владеть навыками анализа данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта	Задание 9. Решение задачи регрессии	Обоснованность выбора модели (7)
14		ПК-9	З.Знать способы и методы анализа данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта У.Уметь проводить анализ данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта Н.Владеть навыками анализа данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта	Задание 10. Подготовка данных и решение задачи кластеризации	Правильность оценки результатов кластеризации (7)
15		ПК-9	У.Уметь проводить анализ данных по созданию, развитию, выводу на рынок, продажам и	Задание 11. Исследование ассоциаций	Корректность интерпретации результатов моделирования (0)

№ п/п	Этапы формирования компетенций (Тема из рабочей программы дисциплины)	Перечень формируемых компетенций по ФГОС ВО	(ЗУНы: (З.1…З.n, У.1…У.n, Н.1…Н.n)	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)
			продвижению ИТ-продукта Н.Владеть навыками анализа данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта		
16		ПК-9	З.Знать способы и методы анализа данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта У.Уметь проводить анализ данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта Н.Владеть навыками анализа данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта	Задание 12. Анализ временных рядов	Умение применять модели временных рядов (0)
17		ПК-9	З.Знать способы и методы анализа данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта У.Уметь проводить анализ данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта Н.Владеть навыками анализа данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта	Задание 13. Сбор, предобработка данных. Решение задачи классификации текстов	Корректность применения моделей (0)
18		ПК-9	З.Знать способы и	Тест 5	Доля правильных

№ п/п	Этапы формирования компетенций (Тема из рабочей программы дисциплины)	Перечень формируемых компетенций по ФГОС ВО	(ЗУНы: (З.1…З.n, У.1…У.n, Н.1…Н.n)	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)
			методы анализа данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта		ответов (6)
				Итого	**100**

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-9 Способен проводить анализ данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта
Знание: Знать способы и методы анализа данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта
1. Базы данных NoSQL. Особенности, классификация
2. Документо-ориентированные базы данных. Возможности СУБД MongoDB
3. Понятие Больших данных. Особенности сбора, хранения, обработки и анализа Больших данных
4. Работа с документо-ориентированными БД.
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
8. Система Apache Hadoop
9. Средства построения распределенных информационных систем
10. Структурированные данные и неструктурированные данные. Данные на естественном языке. Машинные данные. Графовые, или сетевые, данные. Аудио, видео и графика. Потоковые данные
11. Технология Map-Reduce
12. Требования к распределенным информационным системам

ТИПОВЫЕ ЗАДАНИЯ ДЛЯ ПРОВЕРКИ УМЕНИЙ:

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

Компетенция: ПК-9 Способен проводить анализ данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта

Умение: Уметь проводить анализ данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта
Задача № 1. Определить набор коллекций документов MongoDB для предложенной предметной области
Задача № 2. Определить технологию Map-Reduce для решения предложенной задачи обработки данных
Задача № 3. Построить модель и обучить ее по предложенному набору данных

ТИПОВЫЕ ЗАДАНИЯ ДЛЯ ПРОВЕРКИ НАВЫКОВ:

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

Компетенция: ПК-9 Способен проводить анализ данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта
Навык: Владеть навыками анализа данных по созданию, развитию, выводу на рынок, продажам и продвижению ИТ-продукта
Задание № 1. Для предложенной предметной области выполнить анализ тенденций, закономерностей и интерпретацию результатов моделирования на основании подобранной модели
Задание № 2. Для предложенной предметной области предложить архитектуру ИТ-инфраструктуры для решения задач анализа данных
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

Заведующий кафедрой _____ А.В. Родионов

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