

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

УТВЕРЖДАЮ

Проректор по учебной работе
д.филос.н., доц. Атанов А.А.



29.05.2025г.

Рабочая программа дисциплины
Б1.У.17. Архитектура Big Data систем

Направление подготовки: 38.03.05 Бизнес-информатика
Направленность (профиль): Автоматизация и цифровая трансформация
бизнеса
Квалификация выпускника: бакалавр
Форма обучения: очно-заочная

Курс	4
Семестр	41
Лекции (час)	14
Практические (сем, лаб.) занятия (час)	0
Самостоятельная работа, включая подготовку к экзаменам и зачетам (час)	94
Курсовая работа (час)	
Всего часов	108
Зачет (семестр)	41
Экзамен (семестр)	

Иркутск 2025

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

1. Цели изучения дисциплины

Целью изучения дисциплины является получение компетенций в области применения технологии Big Data.

Задачами изучения дисциплины являются

- знакомство с архитектурой 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-4	Способен проектировать, разрабатывать и внедрять компоненты ИТ-инфраструктуры предприятия, обеспечивающие совершенствование и поддержку бизнес-процессов, в том числе с применением инновационных цифровых технологий

Структура компетенции

Компетенция	Формируемые ЗУНы
ПК-4 Способен проектировать, разрабатывать и внедрять компоненты ИТ-инфраструктуры предприятия, обеспечивающие совершенствование и поддержку бизнес-процессов, в том числе с применением инновационных цифровых технологий	З. Знать современные тенденции развития и применения инновационных цифровых компонент ИТ-инфраструктуры предприятия У. Уметь проектировать, разрабатывать и внедрять инновационные цифровые компоненты ИТ-инфраструктуры предприятия Н. Владеть навыками проектирования, разработки и внедрения инновационных цифровых компонент ИТ-инфраструктуры предприятия

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): "Базы данных", "Операционные системы", "Теория вероятностей и математическая статистика"

Дисциплины, использующие знания, умения, навыки, полученные при изучении данной: "Использование систем искусственного интеллекта для анализа бизнеса"

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 3 зач. ед., 108 часов.

Вид учебной работы	Количество часов
Контактная(аудиторная) работа	
Лекции	14
Практические (сем, лаб.) занятия	0
Самостоятельная работа, включая подготовку к экзаменам и зачетам	94
Всего часов	108

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	Экосистема больших данных	41	1	0	8		Тест 1. Технология больших данных. Тест 2. Общие сведения
2	Оперативная аналитическая обработка данных (OLAP)	41	1	0	8		Тест 3. Многомерный анализ данных. Тест 4. Построение дашбордов
3	Средства OLAP-анализа в MS Office	41	1	0	10		Задание 1. Объединение и анализ данных в Power Pivot. Тест 5. OLAP в Power Pivot MS Excel
4	Интеграция и обработка данных с использованием Power BI	41	1	0	12		Задание 2. Объединение и анализ данных в Power BI. Задание 3. Выбор данных в Интернете и анализ данных в Power BI. Тест 6. OLAP-анализ с использованием Power BI
5	Облачные ресурсы анализа данных	41	1	0	10		Задание 4. Облачные технологии многомерного анализа на примере Yandex Datalens. Тест 7. OLAP-анализ с использованием

№ п/п	Раздел и тема дисциплины	Семе- стр	Лек- ции	Семинар Лаборат. Практич.	Само- стоят. раб.	В интера- ктивной форме	Формы текущего контроля успеваемости
							Yandex Datalens
6	Классификация объектов по результатам аналитической обработки	41	1	0	10		Задание 5. Yandex Datalens ABC-XYZ анализ. Тест 8. ABC-XYZ анализ
7	Распределенное хранение и обработка данных	41	1	0	8		Задание 6. Распределенная обработка в Питоне
8	Задачи и модели исследования зависимостей данных	41		0			Задание 7. Исследование и трансформация данных. Задание 8. Решение задачи классификации. Задание 9. Решение задачи регрессии. Тест 9. Трансформация данных. Тест 10. Задачи и модели Data Mining
8.1	Введение в искусственный интеллект и основные методы машинного обучения для работы с табличными данными	41	1	0	4		
8.2	Трансформация данных	41	2	0	6		
8.3	Задачи регрессии и классификации данных	41	2	0	10		
8.4	Задача кластеризации данных	41	2	0	8		
	ИТОГО		14		94		

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
10	Преобразование и очистка данных	Предварительная обработка данных является важным этапом аналитики. Собранные в одну таблицу данные из разных источников могут отличаться по форматам и классификаторам,

№ п/п	Наименование разделов и тем	Содержание
		содержать ошибки и пропуски, измеряться в разных единицах и противоречить друг другу. Они могут быть излишне детальными или наоборот содержать агрегированные значения. Предварительную обработку рассматривают как отдельный этап анализа – ETL – Extract, Transform, Load — «извлечение, преобразование, загрузка».
11	Задачи классификации, регрессии и кластеризации	Решение задачи классификации – отнесения объекта (например, заемщика) по его атрибутам (данные заемщика) к некоторому классу (надежный или ненадежный). Модель классификации настраивается по данным наблюдений атрибутов и классов объектов (кредитная история). Регрессионные модели для определения средних значений выходных переменных по значениям входных переменных. Задача кластеризации – объединение объектов в кластеры по степени близости характеристик.
2	OLAP-технология анализа	Структуры данных OLAP, показатели и измерения, операции OLAP-анализа
3	Дизайн дашбордов	Общие сведения. Назначение дашборда. Показатели и измерения. Визуальные компоненты дашборда. Объединение визуальных компонент в дашборд.
4	Применение Power Pivot для анализа данных	Создание модели данных. Вычисление полей и показателей. Построение сводных таблиц и диаграмм
5	Анализ данных с применением PowerBI	Определение источников данных, загрузка и связывание таблиц, вычисление полей и показателей, использование визуальных компонентов
6	Облачные технологии анализа	Применение Yandex DataLens для определения источников, наборов данных, вычислений, визуальных элементов для построения аналитических панелей - дашбордов
7	ABC-XYZ анализ	Выполнение ABC-XYZ анализа в PowerBI и Datalens
8	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
9	Исследование зависимостей. Обзор	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. Обучающая и тестовая выборка. Работа с категориальными признаками. Перцептрон, логистическая регрессия, полносвязные нейронные сети. Алгоритмы, основанные на применении решающих деревьев. Ансамбли решающих деревьев: случайный лес, градиентный бустинг. Линейные и

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

5.3. Семинарские, практические, лабораторные занятия, их содержание

№ раздела и темы	Содержание и формы проведения
3	Разработка модели данных, вычислений и аналитической панели в Power Pivot. Загрузка и соединение таблиц, определение вычисляемых полей и показателей, задание иерархий, построение аналитической панели, включающей сводную таблицу, диаграммы, срезы
4	Разработка модели данных, вычислений и аналитической панели в PowerBI. Загрузка и соединение таблиц, определение вычисляемых полей и показателей, задание иерархий, построение аналитической панели, включающей сводную таблицу, диаграммы, срезы
5	Разработка модели данных, вычислений и аналитической панели в Datalens. Загрузка и соединение таблиц, определение вычисляемых полей и показателей, задание иерархий, построение аналитической панели, включающей сводную таблицу, диаграммы, срезы
6	ABC-XYZ анализ. Создание визуальных элементов и показателей для ABC-XYZ анализа в среде Datalens
7	Распределенная обработка данных. Применение технологии MapReduce для решения задач обработки распределенных данных. Инструментальные среды проектирования распределенной обработкой
8	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. Экосистема больших данных	ПК-4	З.Знать современные тенденции развития и применения инновационных	Тест 1. Технология больших данных	Доля правильных ответов (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)
			цифровых компонент ИТ-инфраструктуры предприятия		
2		ПК-4	З.Знать современные тенденции развития и применения инновационных цифровых компонент ИТ-инфраструктуры предприятия	Тест 2. Общие сведения	Доля правильных ответов (3)
3	2. Оперативная аналитическая обработка данных (OLAP)	ПК-4	З.Знать современные тенденции развития и применения инновационных цифровых компонент ИТ-инфраструктуры предприятия	Тест 3. Многомерный анализ данных	Доля правильных ответов (3)
4		ПК-4	З.Знать современные тенденции развития и применения инновационных цифровых компонент ИТ-инфраструктуры предприятия	Тест 4. Построение дашбордов	Доля правильных ответов (3)
5	3. Средства OLAP-анализа в MS Office	ПК-4	З.Знать современные тенденции развития и применения инновационных цифровых компонент ИТ-инфраструктуры предприятия У.Уметь проектировать, разрабатывать и внедрять инновационные цифровые компоненты ИТ- инфраструктуры предприятия Н.Владеть навыками проектирования, разработки и внедрения инновационных цифровых компонент ИТ-инфраструктуры предприятия	Задание 1. Объединение и анализ данных в Power Pivot	Корректность запросов. Соответствие анализа постановке задачи (7)
6		ПК-4	З.Знать современные тенденции развития и применения инновационных	Тест 5. OLAP в Power Pivot MS Excel	Доля правильных ответов (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)
			цифровых компонент ИТ-инфраструктуры предприятия		
7	4. Интеграция и обработка данных с использованием Power BI	ПК-4	З.Знать современные тенденции развития и применения инновационных цифровых компонент ИТ-инфраструктуры предприятия У.Уметь проектировать, разрабатывать и внедрять инновационные цифровые компоненты ИТ- инфраструктуры предприятия Н.Владеть навыками проектирования, разработки и внедрения инновационных цифровых компонент ИТ-инфраструктуры предприятия	Задание 2. Объединение и анализ данных в Power BI	Полнота представленных данных (8)
8		ПК-4	З.Знать современные тенденции развития и применения инновационных цифровых компонент ИТ-инфраструктуры предприятия У.Уметь проектировать, разрабатывать и внедрять инновационные цифровые компоненты ИТ- инфраструктуры предприятия Н.Владеть навыками проектирования, разработки и внедрения инновационных цифровых компонент ИТ-инфраструктуры предприятия	Задание 3. Выбор данных в Интернете и анализ данных в Power BI	Корректность создания БД (8)
9		ПК-4	З.Знать современные	Тест 6. OLAP-анализ с	Доля правильных

№ 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)
			тенденции развития и применения инновационных цифровых компонент ИТ-инфраструктуры предприятия	использованием Power BI	ответов (3)
10	5. Облачные ресурсы анализа данных	ПК-4	З.Знать современные тенденции развития и применения инновационных цифровых компонент ИТ-инфраструктуры предприятия У.Уметь проектировать, разрабатывать и внедрять инновационные цифровые компоненты ИТ- инфраструктуры предприятия Н.Владеть навыками проектирования, разработки и внедрения инновационных цифровых компонент ИТ-инфраструктуры предприятия	Задание 4. Облачные технологии многомерного анализа на примере Yandex Datalens	Полнота описания процесса показателями и признаками (8)
11		ПК-4	З.Знать современные тенденции развития и применения инновационных цифровых компонент ИТ-инфраструктуры предприятия	Тест 7. OLAP-анализ с использованием Yandex Datalens	Доля правильных ответов (3)
12	6. Классификация объектов по результатам аналитической обработки	ПК-4	З.Знать современные тенденции развития и применения инновационных цифровых компонент ИТ-инфраструктуры предприятия У.Уметь проектировать, разрабатывать и внедрять инновационные цифровые компоненты ИТ- инфраструктуры	Задание 5. Yandex Datalens ABC-XYZ анализ	Корректность (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)
			предприятия Н. Владеть навыками проектирования, разработки и внедрения инновационных цифровых компонент ИТ-инфраструктуры предприятия		
13		ПК-4	З. Знать современные тенденции развития и применения инновационных цифровых компонент ИТ-инфраструктуры предприятия	Тест 8. ABC-XYZ анализ	Доля правильных ответов (3)
14	7. Распределенное хранение и обработка данных	ПК-4	З. Знать современные тенденции развития и применения инновационных цифровых компонент ИТ-инфраструктуры предприятия У. Уметь проектировать, разрабатывать и внедрять инновационные цифровые компоненты ИТ- инфраструктуры предприятия Н. Владеть навыками проектирования, разработки и внедрения инновационных цифровых компонент ИТ-инфраструктуры предприятия	Задание 6. Распределенная обработка в Питоне	Корректность программы (8)
15	8. Задачи и модели исследования зависимостей данных	ПК-4	З. Знать современные тенденции развития и применения инновационных цифровых компонент ИТ-инфраструктуры предприятия У. Уметь проектировать, разрабатывать и внедрять инновационные	Задание 7. Исследование и трансформация данных	Полнота преобразований (8)

№ 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)
			цифровые компоненты ИТ- инфраструктуры предприятия Н. Владеть навыками проектирования, разработки и внедрения инновационных цифровых компонент ИТ-инфраструктуры предприятия		
16		ПК-4	З. Знать современные тенденции развития и применения инновационных цифровых компонент ИТ-инфраструктуры предприятия У. Уметь проектировать, разрабатывать и внедрять инновационные цифровые компоненты ИТ- инфраструктуры предприятия Н. Владеть навыками проектирования, разработки и внедрения инновационных цифровых компонент ИТ-инфраструктуры предприятия	Задание 8. Решение задачи классификации	Корректность применения моделей (8)
17		ПК-4	З. Знать современные тенденции развития и применения инновационных цифровых компонент ИТ-инфраструктуры предприятия У. Уметь проектировать, разрабатывать и внедрять инновационные цифровые компоненты ИТ- инфраструктуры предприятия	Задание 9. Решение задачи регрессии	Корректность моделей и интерпретации результатов (8)

№ 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. Владеть навыками проектирования, разработки и внедрения инновационных цифровых компонент ИТ-инфраструктуры предприятия		
18		ПК-4	З. Знать современные тенденции развития и применения инновационных цифровых компонент ИТ-инфраструктуры предприятия	Тест 9. Трансформация данных	Доля правильных ответов (3)
19		ПК-4	З. Знать современные тенденции развития и применения инновационных цифровых компонент ИТ-инфраструктуры предприятия	Тест 10. Задачи и модели Data Mining	Доля правильных ответов (3)
				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-4 Способен проектировать, разрабатывать и внедрять компоненты ИТ-инфраструктуры предприятия, обеспечивающие совершенствование и поддержку бизнес-процессов, в том числе с применением инновационных цифровых технологий

Знание: Знать современные тенденции развития и применения инновационных цифровых компонент ИТ-инфраструктуры предприятия

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 кластеризации

7. Система Apache Hadoop
8. Средства построения распределенных информационных систем
9. Структурированные данные и неструктурированные данные. Данные на естественном языке. Машинные данные. Графовые, или сетевые, данные. Аудио, видео и графика. Поточковые данные
10. Технология Map-Reduce
11. 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-4 Способен проектировать, разрабатывать и внедрять компоненты ИТ-инфраструктуры предприятия, обеспечивающие совершенствование и поддержку бизнес-процессов, в том числе с применением инновационных цифровых технологий

Умение: Уметь проектировать, разрабатывать и внедрять инновационные цифровые компоненты ИТ-инфраструктуры предприятия

Задача № 1. Для предложенного датасета выполнить исследование и трансформацию данных

Задача № 2. Для предложенного датасета построить дашборд в среде Power Pivot BI

Задача № 3. Для предложенного датасета построить дашборд в среде Power Pivot MS Excel

Задача № 4. Для предложенного датасета построить дашборд в среде Yandex Datalens

Задача № 5. Для предложенного датасета решить задачу классификации

Задача № 6. Для предложенного датасета решить задачу кластеризации

Задача № 7. Для предложенного датасета решить задачу регрессии

Задача № 8. Используя функции map и reduce выполнить имитацию распределенной обработки данных о продажах товаров для вычисления указанных показателей для измерения

ТИПОВЫЕ ЗАДАНИЯ ДЛЯ ПРОВЕРКИ НАВЫКОВ:

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

Компетенция: ПК-4 Способен проектировать, разрабатывать и внедрять компоненты ИТ-инфраструктуры предприятия, обеспечивающие совершенствование и поддержку бизнес-процессов, в том числе с применением инновационных цифровых технологий

Навык: Владеть навыками проектирования, разработки и внедрения инновационных цифровых компонент ИТ-инфраструктуры предприятия

Задание № 1. Для предложенной предметной области разработать решение задачи классификации, регрессии или кластеризации

Задание № 2. Для предложенной предметной области разработать схему анализа показателей

ОБРАЗЕЦ БИЛЕТА

Министерство науки и высшего образования
Российской Федерации
Федеральное государственное бюджетное

Направление - 38.03.05 Бизнес-информатика
Профиль - Автоматизация и цифровая

БИЛЕТ № 1

1. Тест (30 баллов).
2. Для предложенного датасета решить задачу классификации (30 баллов).
3. Для предложенной предметной области разработать схему анализа показателей (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. Руссо, М. Подробное руководство по DAX: бизнес-аналитика с Microsoft Power BI, SQL Server Analysis Services и Excel / М. Руссо, А. Феррари ; перевод А. Ю. Гинько. — Москва : ДМК Пресс, 2021. — 776 с. — ISBN 978-5-97060-859-3. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: <https://www.iprbookshop.ru/124728.html> (дата обращения: 14.10.2022).
7. Феррари, А. Анализ данных при помощи Microsoft Power BI и Power Pivot для Excel / А. Феррари, М. Руссо ; перевод А. Ю. Гинько. — Москва : ДМК Пресс, 2020. — 288 с. — ISBN 978-5-97060-858-6. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: <https://www.iprbookshop.ru/124710.html> (дата обращения: 14.10.2022).

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. Братищенко В.В. Реляционные и документационные базы данных.- Иркутск: Изд-во БГУ, 2020.- 132 с.
3. [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>](http://www.iprbookshop.ru/58563.html)
4. [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>](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
- Yandex DataLens — это сервис для бизнес-аналитики, адрес доступа: <https://yandex.cloud/ru/docs/datalens/>. доступ неограниченный

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,
- Orange 3,

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