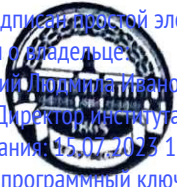


Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Хоружий Людмила Ивановна
Должность: Директор института экономики и управления АПК
Дата подписания: 25.07.2022 19:17:36
Уникальный программный ключ:
1e90b132d9b04dce67585160b0150aa12c61e6a9



МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
**«РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ –
МСХА имени К.А. ТИМИРЯЗЕВА»**
(ФГБОУ ВО РГАУ - МСХА имени К.А. Тимирязева)

Институт экономики и управления АПК
Кафедра статистики и кибернетики

УТВЕРЖДАЮ:
Директор института экономики и управ-
ления АПК

Л.И. Хоружий
“ 25 ” июля 2022 г.



РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
Б1.В.ДВ.03.01 Хранилища и системы интеллектуального анализа данных
на иностранном языке

для подготовки бакалавров

ФГОС ВО

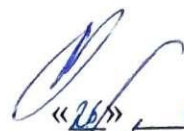
Направление: 09.03.02 «Информационные системы и технологии»
Направленность:
Большие данные и машинное обучение (Machine Learning & Big Data)

Курс 3
Семестр 5


Форма обучения очная
Год начала подготовки 2022

Москва, 2022

Разработчик (и): Харитонов А.Е., к.э.н., доцент
(ФИО, ученая степень, ученое звание)


«26» 08 2022 г.

Рецензент: Коломеева Е.С., к.э.н.
(ФИО, ученая степень, ученое звание)


(подпись)
«26» 08 2022 г.

Программа составлена в соответствии с требованиями ФГОС ВО, профессионального стандарта и учебного плана по направлению подготовки 09.03.02 «Информационные системы и технологии».

Программа обсуждена на заседании кафедры статистики и кибернетики протокол № 11 от «26» августа 2022 г.

И.о.зав. кафедрой Уколова А.В., к.э.н., доцент
(ФИО, ученая степень, ученое звание)


(подпись)
«26» 08 2022 г.

Согласовано:

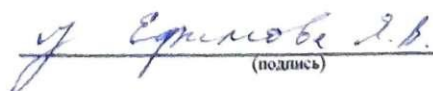
Председатель учебно-методической
комиссии института экономики и управления АПК
Корольков А.Ф., к.э.н., доцент
(ФИО, ученая степень, ученое звание)


(подпись)
«26» 08 2022 г.

И.о.заведующего выпускающей кафедрой
статистики и кибернетики
Уколова А.В., к.э.н., доцент
(ФИО, ученая степень, ученое звание)


(подпись)
«26» 08 2022 г.

Заведующий отделом комплектования ЦНБ


(подпись)

СОДЕРЖАНИЕ

АННОТАЦИЯ.....	4
1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ.....	6
2. МЕСТО ДИСЦИПЛИНЫ В УЧЕБНОМ ПРОЦЕССЕ	6
3. ПЕРЕЧЕНЬ ПЛАНИРУЕМЫХ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ, СООТНЕСЕННЫХ С ПЛАНИРУЕМЫМИ РЕЗУЛЬТАТАМИ ОСВОЕНИЯ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ.....	6
4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ	7
4.1 РАСПРЕДЕЛЕНИЕ ТРУДОЁМКОСТИ ДИСЦИПЛИНЫ ПО ВИДАМ РАБОТ	7
ПО СЕМЕСТРАМ	7
4.2 СОДЕРЖАНИЕ ДИСЦИПЛИНЫ.....	11
4.3 ПРАКТИЧЕСКИЕ ЗАДАНИЯ.....	14
5. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ	18
6. ТЕКУЩИЙ КОНТРОЛЬ УСПЕВАЕМОСТИ И ПРОМЕЖУТОЧНАЯ АТТЕСТАЦИЯ ПО ИТОГАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ	18
6.1. ТИПОВЫЕ КОНТРОЛЬНЫЕ ЗАДАНИЯ ИЛИ ИНЫЕ МАТЕРИАЛЫ, НЕОБХОДИМЫЕ ДЛЯ ОЦЕНКИ ЗНАНИЙ, УМЕНИЙ И НАВЫКОВ И (ИЛИ) ОПЫТА ДЕЯТЕЛЬНОСТИ	18
6.2. ОПИСАНИЕ ПОКАЗАТЕЛЕЙ И КРИТЕРИЕВ КОНТРОЛЯ УСПЕВАЕМОСТИ, ОПИСАНИЕ ШКАЛ ОЦЕНИВАНИЯ	30
7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ.....	31
7.1 ОСНОВНАЯ ЛИТЕРАТУРА	31
7.2 ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА.....	32
7.3 МЕТОДИЧЕСКИЕ УКАЗАНИЯ, РЕКОМЕНДАЦИИ И ДРУГИЕ МАТЕРИАЛЫ К ЗАНЯТИЯМ.....	32
8. ПЕРЕЧЕНЬ РЕСУРСОВ ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ «ИНТЕРНЕТ», НЕОБХОДИМЫХ ДЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ.....	32
9. ПЕРЕЧЕНЬ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ И ИНФОРМАЦИОННЫХ СПРАВОЧНЫХ СИСТЕМ.....	33
10. ОПИСАНИЕ МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЙ БАЗЫ, НЕОБХОДИМОЙ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ.....	33
11. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ОБУЧАЮЩИМСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ.	35
12. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПРЕПОДАВАТЕЛЯМ ПО ОРГАНИЗАЦИИ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ.....	35

Аннотация

рабочей программы учебной дисциплины

Б1.В.ДВ.03.01 «Хранилища и системы интеллектуального анализа данных на иностранном языке»

для подготовки бакалавров по направлению 09.03.02 «Информационные системы и технологии» по направленности Большие данные и машинное обучение (Machine Learning & Big Data) и

Цель освоения дисциплины: Целью дисциплины «Хранилища и системы интеллектуального анализа данных на иностранном языке» является освоение студентами теоретических и практических знаний и приобретение умений и навыков в области проведения аналитического исследования с применением технологий больших данных и выполнения работ по обеспечению функционирования баз данных и обеспечению их информационной безопасности, а также осуществлять деловую коммуникацию в устной и письменной формах на государственном языке Российской Федерации и иностранном языке.

Место дисциплины в учебном плане: включена в часть, формируемую участниками образовательных отношений (дисциплина по выбору) учебного плана по направлению подготовки 09.03.02 «Информационные системы и технологии».

Требования к результатам освоения дисциплины: в результате освоения дисциплины формируются следующие компетенции (индикаторы): УК-4 (УК-4.2, УК-4.3), ПКос-3 (ПКос-3.1, ПКос-3.3), ПКос-9 (ПКос-9.1, ПКос-9.2, ПКос-9.3).

Краткое содержание дисциплины:

Отличия баз данных от Хранилищ данных. Причины появления Хранилищ данных. Хранилище - надстройка над существующими базами данных. Отличительная особенность Хранилищ. Основные требования к данным, вводимым в Хранилище. Задачи построения Хранилища. Интеграция с различными источниками данных. Предварительная обработка данных для их дальнейшего использования в алгоритмах Data Mining. Заполнение пропусков, сэмплинг, редактирование выбросов. Квантование. Конечные классы. Разбиение на множества. Редактирование выбросов. Сглаживание. Группировка. Дата и время. Дополнение данных. Замена. Кросс-таблица. Объединение. Параметры полей. Разгруппировка. Свёртка столбцов. Скользящее окно. Слияние. Соединение. Сортировка. Фильтр строк. Дубликаты и противоречия во входных данных. Автокорреляция исходных переменных. Корреляционный анализ. Коэффициент корреляции Пирсона. Линейная регрессия. Разбиение на множества. Отбор факторов и защита от переобучения. F-тест. Коэффициент детерминации. Кластеризация. Центры кластеров. Автоопределение числа кластеров. Нормализация данных. Алгоритм CLOPE. Коэффициент отталкивания. алгоритм кластеризации EM. Медианная модификация. Модель ARIMAX. Среднеквадратическая ошибка на обучающем множестве. Минимальная достоверность правила. Минимальный лифт правила. Максимальное число следствий. Использование нейросети для задачи регрессии. Разбиение на множества. Метод вали-

дации. Настройка параметров Нейросети. Параметры обучения. Критерии останова. Стратегия оптимизации. Использование нейросети для задачи классификации.

Differences between databases and data warehouses. Reasons for the emergence of Data Warehouses. Storage - add-on over existing databases. Distinctive feature of Vaults. Basic requirements for data entered into the Warehouse. Tasks of building the Repository. Integration with various data sources. Data preprocessing for further use in Data Mining algorithms. Gap filling, sampling, outlier editing. Quantization. Final classes. Splitting into sets. Editing outliers. Smoothing. Grouping. Date and time. Data completion. Replacement. Cross table. An association. Field options. Ungrouping. Collapse columns. Sliding window. Merging. Compound. Sorting. Row filter. Duplicates and contradictions in the input data. Autocorrelation of original variables. Correlation analysis. Pearson's correlation coefficient. Linear regression. Splitting into sets. Selection of factors and protection from overfitting. F-test. Determination coefficient. Clustering. cluster centers. Auto-detection of the number of clusters. Data normalization. CLOPE algorithm. repulsion coefficient. EM clustering algorithm. Median modification. ARIMAX model. Root-mean-square error on the training set. The minimum certainty of the rule. Minimum lift rules. The maximum number of consequences. Using a neural network for a regression problem. Splitting into sets. validation method. Setting the parameters of the Neural Network. Learning options. Stop criteria. Optimization strategy. Using a neural network for a classification task.

Общая трудоемкость дисциплины составляет: 3 зачетные единицы (108 часов).

Промежуточный контроль: зачет с оценкой.

1. Цель освоения дисциплины

Целью дисциплины «Хранилища и системы интеллектуального анализа данных на иностранном языке» является освоение студентами теоретических и практических знаний и приобретение умений и навыков в области проведения аналитического исследования с применением технологий больших данных и выполнения работ по обеспечению функционирования баз данных и обеспечению их информационной безопасности, а также осуществлять деловую коммуникацию в устной и письменной формах на государственном языке Российской Федерации и иностранном языке.

2. Место дисциплины в учебном процессе

Дисциплина «Хранилища и системы интеллектуального анализа данных на иностранном языке» включена в часть, формируемую участниками образовательных отношений (дисциплина по выбору) учебного плана. Дисциплина «Хранилища и системы интеллектуального анализа данных на иностранном языке» реализуется в соответствии с требованиями ФГОС ВО, профессионального стандарта, ОПОП ВО и Учебного плана по направлению 09.03.02 «Информационные системы и технологии».

Предшествующими курсами, на которых непосредственно базируется дисциплина «Хранилища и системы интеллектуального анализа данных на иностранном языке» являются «Введение в компьютерные науки на иностранном языке», «Математический анализ», «Математическая статистика», «Теория вероятностей», «Алгоритмизация и программирование», «Основы науки о данных (Data Science)», «Анализ экономических данных с использованием современных информационных технологий на иностранном языке», «Анализ экономических данных с использованием современных информационных технологий».

Дисциплина «Хранилища и системы интеллектуального анализа данных на иностранном языке» является основополагающей для изучения следующих дисциплин: «Администрирование информационных систем», «Многомерные статистические методы», «Методы искусственного интеллекта», «Большие данные».

Рабочая программа дисциплины «Хранилища и системы интеллектуального анализа данных на иностранном языке» для инвалидов и лиц с ограниченными возможностями здоровья разрабатывается индивидуально с учетом особенностей психофизического развития, индивидуальных возможностей и состояния здоровья таких обучающихся.

3. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы

Образовательные результаты освоения дисциплины обучающимся, представлены в таблице 1.

4. Структура и содержание дисциплины

4.1 Распределение трудоёмкости дисциплины по видам работ по семестрам

Общая трудоёмкость дисциплины составляет 3 зач.ед. (108 часов), их распределение по видам работ семестрам представлено в таблице 2.

Таблица 1

Требования к результатам освоения учебной дисциплины

№ п/п	Код компетенции	Содержание компетенции (или её части)	Индикаторы компетенций	В результате изучения учебной дисциплины обучающиеся должны:		
				знать	уметь	владеть
1.	УК-4	Способен осуществлять деловую коммуникацию в устной и письменной формах на государственном языке Российской Федерации и иностранном(ых) языке(ах)	УК-4.2 Уметь: применять на практике деловую коммуникацию в устной и письменной формах, методы и навыки делового общения на русском и иностранном языках		применять на практике деловую коммуникацию в сфере работы с хранилищами данных и их анализов в устной и письменной формах на русском и иностранном языках	
2.			УК-4.3 Владеть: навыками чтения и перевода текстов на иностранном языке в профессиональном общении; навыками деловых коммуникаций в устной и письменной форме на русском и иностранном языках; методикой составления суждения в межличностном деловом общении на русском и иностранном языках			навыками чтения и перевода текстов на иностранном языке в области хранилищ и интеллектуального анализа данных.
3.	ПКос-3	Способность выполнять работы по обеспечению функ-	ПКос-3.1 Знать: понятие «целостность	понятие «целостность баз данных»; причины		

		ционирования баз данных и обеспечению их информационной безопасности	баз данных»; причины нарушения целостности баз данных; способы достижения безопасности баз данных	нарушения целостности баз и хранилищ данных.		
4.			ПКос-3.3 Иметь навыки: сопровождения базы данных в оперативном и обеспечивающем режимах			сопровождения базы и хранилища данных в оперативном и обеспечивающем режимах
5.	ПКос-9	Способен проводить аналитическое исследование с применением технологий больших данных	ПКос-9.1 Знать: предметную область анализа, типы больших данных, источники и методы извлечения информации, теоретические и прикладные основы анализа, технологии хранения и обработки, современные методы и инструментальные средства анализа больших данных	типы данных, источники и методы извлечения информации, теоретические и прикладные основы интеллектуального анализа, технологии хранения и обработки, современные методы и инструментальные средства анализа больших данных		
6.			ПКос-9.2 Уметь: оценивать соответствие наборов данных задачам анализа больших данных; использовать инструментальные сред-		оценивать наборы данных методами интеллектуального анализа данных; использовать инструментальные средства для извлечения, преобразования, хранения и	

			ства для извлечения, преобразования, хранения и обработки данных из разнородных источников; разрабатывать и оценивать модели больших данных; автоматизировать процесс анализа больших данных; визуализировать результаты анализа больших данных		обработки данных из разнородных источников; визуализировать результаты анализа больших данных	
7.			ПКос-9.3 Иметь навыки: выбора источников данных, оценки соответствия набора данных предметной области и задачам аналитических работ; получения и фильтрации, извлечения, проверки, очистки, агрегации и разработки представления больших объемов данных из гетерогенных источников			выбора источников данных, оценки соответствия набора данных предметной области и задачам аналитических работ; получения и фильтрации, извлечения, проверки, очистки, агрегации и разработки представления больших объемов данных из гетерогенных источников

Распределение трудоёмкости дисциплины по видам работ по семестрам

Вид учебной работы	Трудоёмкость, 5 семестр час. всего/*
Общая трудоёмкость дисциплины по учебному плану	108
1. Контактная работа:	50,35
Аудиторная работа	50,35
<i>в том числе:</i>	
<i>лекции (Л)</i>	16
<i>практические занятия (ПЗ)</i>	34/4
<i>контактная работа на промежуточном контроле (КРА)</i>	0,35
2. Самостоятельная работа (СРС)	57,65
<i>самостоятельное изучение разделов, самоподготовка (проработка и повторение лекционного материала и материала учебников и учебных пособий, подготовка к практическим занятиям и т.д.)</i>	
<i>Подготовка к зачету (контроль)</i>	9
Вид промежуточного контроля:	зачет с оценкой

* в том числе практическая подготовка.

4.2 Содержание дисциплины

Тематический план учебной дисциплины

Наименование разделов и тем дисциплин (укрупнено)	Всего	Аудиторная работа			Внеаудиторная работа СР
		Л	ПЗ всего /*	ПКР	
Тема 1 Технология хранения данных при принятии решений Theme 1 Technology of data storage in decision-making	18	2	6/2		10
Тема 2 Предобработка и консолидация данных Theme 2 Data preprocessing and consolidation	11	1	4/2		6
Тема 3 Трансформация данных Theme 3 Data transformation	7	1	2		4
Тема 4 Предварительный анализ данных Theme 4 Preparatory data analysis	10	2	4		4
Тема 5 Регрессия Theme 5 Linear Regression	11,65	2	4		5,65
Тема 6 Кластеризация Theme 6 Clustering	16	2	6		8
Тема 7 ARIMAX Theme 7 ARIMAX	10	2	2		6
Тема 8 Ассоциативные правила	12	2	2		8

Наименование разделов и тем дисциплин (укрупнёно)	Всего	Аудиторная работа			Внеаудиторная работа СР
		Л	ПЗ всего /*	ПКР	
Theme 8 Association Rules					
Тема 9 Нейросеть Theme 9 Neural network	12	2	4		6
Контактная работа на промежуточном контроле (КРА)	0,35			0,35	
Всего за 5 семестр	108	16	34/4	0,35	57,65
Итого по дисциплине	108	16	34/4	0,35	57,65

* в том числе практическая подготовка

Тема 1 Технология хранения данных при принятии решений Theme 1 Technology of data storage in decision-making

Отличия баз данных от Хранилищ данных. Причины появления Хранилищ данных. Хранилище - надстройка над существующими базами данных. Отличительная особенность Хранилищ. Основные требования к данным, вводимым в Хранилище. Задачи построения Хранилища. Интеграция с различными источниками данных. SOAP-запрос. REST-запрос. Выполнение программы. Извлечение XML.

Differences of databases from Data Warehouses. The reasons for the emergence of Data-storage-data. Storage - add-on over existing databases. A distinctive feature of the Vaults. The main requirements for data entered in the Warehouse. Warehouse building Tasks. Integration with various data sources. SOAP request. REST request. Program execution. XML extraction.

Тема 2 Предобработка и консолидация данных Theme 2 Data preprocessing and consolidation

Предварительная обработка данных для их дальнейшего использования в алгоритмах Data Mining. Заполнение пропусков, сэмплинг. Квантование. Конечные классы. Разбиение на множества. Редактирование выбросов. Сглаживание.

Data preprocessing for their further use in Data Mining algorithms. Gap filling, sampling. Quantization. Final classes. Splitting into sets. Editing outliers. Smoothing.

Тема 3 Трансформация данных Theme 3 Data transformation

Набор компонентов для первоначальной подготовки и простой обработки исходных наборов данных. Группировка. Дата и время. Дополнение данных. Замена. Кросс-таблица. Объединение. Параметры полей. Разгруппировка. Свёртка столбцов. Скользящее окно. Слияние. Соединение. Сортировка. Фильтр строк.

A set of components for initial preparation and simple processing of initial data sets. Grouping. Date and time. Data completion. Replacement. Cross table. An association. Field options. Ungrouping. Collapse columns. Sliding window. Merging. Compound. Sorting. Row filter.

Тема 4 Предварительный анализ данных Theme 4 Preparatory data analysis

Дубликаты и противоречия во входных данных. Автокорреляция исходных переменных. Корреляционный анализ. Коэффициент корреляции Пирсона. Коэффициент Тау-в Кендалла. Экстремум взаимокорреляционной функции. Коэффициент корреляции Спирмена.

Duplicates and contradictions in the input data. Autocorrelation of source variables. Correlation analysis. Pearson's correlation coefficient. Kendall's Tau-b coefficient. Extremum of the cross-correlation function. Spearman's correlation coefficient.

Тема 5 Линейная регрессия

Theme 5 Linear Regression

Линейная регрессия. Настройки нормализации. Разбиение на множества. Настройка линейной регрессии. Детальные настройки. Отбор факторов и защита от переобучения. F-тест. Коэффициент детерминации. Скорректированный коэффициент детерминации. Информационный критерий Акаике. Информационный критерий Байеса. Информационный критерий Ханнана-Квина. Порог значимости при добавлении фактора. Порог значимости при исключении фактора.

Linear regression. Normalization settings. Splitting into sets. Setting up a linear regression. Detailed settings. Selection of factors and protection against overfitting. F-test. Determination coefficient. Adjusted coefficient of determination. Information criterion Akaike. Bayes information criterion. Hannan-Queen information criterion. Threshold of significance when adding a factor. Threshold of significance when excluding a factor.

Тема 6 Кластеризация

Theme 6 Clustering

Кластеризация. Центры кластеров. Автоопределение числа кластеров. Нормализация данных. Алгоритм CLOPE. Коэффициент отталкивания. алгоритм кластеризации EM. Медианная модификация.

Clustering, cluster centers. Auto-detection of the number of clusters. Data normalization. CLOPE algorithm. repulsion coefficient. EM clustering algorithm. median modification.

Тема 7 ARIMAX

Theme 7 ARIMAX

Модель ARIMAX для анализа временных рядов как интегрированная авторегрессия, скользящее среднее с возможностью учета дополнительных внешних факторов. Порядок сезонной AR части. Порядок сезонного интегрирования. Порядок сезонной MA части. Период сезонной составляющей. Среднеквадратическая ошибка на обучающем множестве. Средняя абсолютная ошибка на обучающем множестве. Средняя относительная ошибка на обучающем множестве. Коэффициент детерминации. Скорректированный коэффициент детерминации.

ARIMAX model for time series analysis as an integrated autoregression, moving average with the possibility of taking into account additional external factors. The order of the seasonal AR part. The order of seasonal integration. The order of the seasonal MA part. The period of the seasonal component. Root-mean-square error on the training set. Average absolute error on the training set. Average relative error on the training set. Determination coefficient. Adjusted coefficient of determination.

Тема 8 Ассоциативные правила

Theme 8 Association Rules

Ассоциативные правила в транзакционных данных. Минимальная поддержка. Минимальная достоверность правила. Минимальный лифт правила. Максимальное число следствий. Понятия и методы выявления закономерностей в интеллектуальном анализе данных. Использование выявленных закономерностей для предсказания неизвестных значений (прогностическое моделирование). Анализ исключений, предназначенный для выявления и толкования аномалий в найденных закономерностях.

Association rules in transactional data. Minimum support. The minimum certainty of the rule. Minimum lift rules. The maximum number of consequences. Concepts and methods of identifying patterns in data mining. Applying the discovered patterns for prediction of unknown values (predictive modeling). Analysis of exceptions designed to detect and interpret anomalies in the discovered patterns.

Тема 9 Нейросеть

Theme 9 Neural network

Использование нейросети для задачи регрессии. Разбиение на множества. Метод валидации. Настройка параметров Нейросети. Параметры обучения. Критерии останова. Стратегия оптимизации. Использование нейросети для задачи классификации.

Using a neural network for a regression problem. Splitting into sets. validation method. Setting the parameters of the Neural Network. Learning options. Stop criteria. Optimization strategy. Using a neural network for a classification task.

4.3 Лекции /практические занятия

Таблица 4

Содержание лекции /практические занятия и контрольные мероприятия

№ п/п	№ раздела	№ и название практических занятий	Формируемые компетенции (индикаторы)	Вид контрольного мероприятия	Кол-во Часов/ из них практическая подготовка
1	Тема 1 Технология хранения данных при принятии решений Theme 1 Technology of data storage in decision-making	Лекция 1. Технология хранения данных при принятии решений Lecture 1 Technology of data storage in decision-making	ПКос-3.1 ПКос-9.1 УК-4.2		2
		Практическая работа № 1. Принципы работы с хранилищами данных Practical lesson № 1. Principles of working with data stores	ПКос-3.3 ПКос-9.2 ПКос-9.3 УК-4.2 УК-4.3	Чтение, перевод, защита работы	2
		Практическая работа № 2. Интеграция с различными базами данных Practical lesson № 2. Integration with various databases	ПКос-3.3 ПКос-9.2 ПКос-9.3 УК-4.2	Защита работы	2/2

№ п/п	№ раздела	№ и название практических занятий	Формируемые компетенции (индикаторы)	Вид контрольного мероприятия	Кол-во Часов/ из них практическая подготовка
		Практическая работа № 3. Импорт данных в различных форматах Practical lesson № 3. Import data in various formats	ПКос-3.3 ПКос-9.2 ПКос-9.3 УК-4.2	Защита работы	2
2	Тема 2 Предобработка и консолидация данных Theme 2 Data preprocessing and consolidation	Лекция 2 Предобработка и консолидация данных Lecture 2 Data preprocessing and consolidation	ПКос-3.1 ПКос-9.1 УК-4.2		1
		Практическая работа № 4. Подготовка данных для анализа Practical lesson № 4. Data preparation for analysis	ПКос-9.2 ПКос-9.3 УК-4.2 ПКос-3.3	Защита работы	4/2
3	Тема 3 Трансформация данных Theme 3 Data transformation	Лекция 3 Трансформация данных Lecture 3 Data transformation	ПКос-3.1 ПКос-9.1 УК-4.2		1
		Практическая работа № 5. Сортировка, объединение и разделение наборов данных Practical lesson № 5. Sorting, merging and splitting datasets	ПКос-9.2 ПКос-9.3 УК-4.2	Защита работы	
4	Тема 4 Предварительный анализ данных Theme 4 Preparatory data analysis	Лекция 4 Методы анализа данных Lecture 4 Methods of data analysis	ПКос-9.1 УК-4.2		2
		Практическая работа № 6. Выявление дубликатов и противоречий во входных данных Practical lesson № 6. Identification of duplicates and contradictions in the input data	ПКос-9.2 ПКос-9.3 УК-4.2	Защита работы	2
		Практическая работа № 7. Предварительный анализ данных Practical lesson № 7. Preparatory data analysis	ПКос-9.2 ПКос-9.3 УК-4.2 УК-4.3	Чтение, перевод, защита работы	2
5	Тема 5 Регрессия Theme 5 Linear Regression	Лекция 5 Регрессионный анализ Lecture 5 Regression analysis	ПКос-9.1 УК-4.2		2
		Практическая работа № 8. Парная линейная регрессия Practical lesson № 8. Paired Linear Regression	ПКос-9.2 ПКос-9.3 УК-4.2 УК-4.3	Чтение, перевод, защита работы	2
		Практическая работа № 9. Множественная линейная регрессия Practical lesson № 9. Multiple	ПКос-9.2 ПКос-9.3 УК-4.2	Защита работы	2

№ п/п	№ раздела	№ и название практических занятий	Формируемые компетенции (индикаторы)	Вид контрольного мероприятия	Кол-во Часов/ из них практическая подготовка
		Linear Regression			
6	Тема 6 Кластеризация Theme 6 Clustering	Лекция 6 Кластерный анализ Lecture 6 Cluster analysis	ПКос-9.1 УК-4.2		2
		Практическая работа № 10. Алгоритм CLOPE. Practical lesson № 10. CLOPE algorithm.	ПКос-9.2 ПКос-9.3 УК-4.2	Защита работы	2
		Практическая работа № 11. Алгоритм кластеризации EM Practical lesson № 11. EM Clustering Algorithm	ПКос-9.2 ПКос-9.3 УК-4.2	Защита работы	2
7	Тема 7 ARIMAX Theme 7 ARIMAX	Лекция 7 Метод ARIMAX Lecture 7 ARIMAX method	ПКос-9.1 УК-4.2		2
		Практическая работа № 12. Анализ временных рядов методом ARIMAX Practical lesson № 12. Time series analysis using the ARIMAX method	ПКос-9.2 ПКос-9.3 УК-4.2	Защита работы	2
8	Тема 8 Ассоциативные правила Theme 8 Association Rules	Лекция 8 Ассоциативные правила Lecture 8 Association rules	ПКос-9.1 УК-4.2		2
		Практическая работа № 13 Поиск ассоциативных правил Practical lesson № 13. Search for association rules	ПКос-9.2 ПКос-9.3 УК-4.2 УК-4.3	Чтение, перевод, защита работы	2
9	Тема 9 Нейросеть Theme 9 Neural network	Лекция 8 Нейронные сети Lecture 8 Neural networks	ПКос-9.1 УК-4.2		2
		Практическая работа № 14 Применение нейронной сети для регрессии Practical lesson № 14. Applying a Neural Network for Regression	ПКос-9.2 ПКос-9.3 УК-4.2	Защита работы	2
		Итоговый тест по дисциплине	ПКос-3.1 ПКос-3.3 ПКос-9.1 ПКос-9.2 ПКос-9.3 УК-4.2	Тестирование	2

Таблица 5

Перечень вопросов для самостоятельного изучения дисциплины

№ п/п	№ раздела и темы	Перечень рассматриваемых вопросов для самостоятельного изучения
1.	Тема 1 Технология хранения данных при принятии решений Theme 1 Technology of data storage in decision-making	Отличительная особенность Хранилищ. Основные требования к данным, вводимым в Хранилище. Задачи построения Хранилища. A distinctive feature of the Vaults. The main requirements for data entered in the Warehouse. Warehouse building Tasks. (УК-4.2, УК-4.3, ПКос-3.1, ПКос-3.3)
2.	Тема 2 Предобработка и консолидация данных Theme 2 Data preprocessing and consolidation	Квантование. Конечные классы. Сглаживание. Quantization. Final classes. Smoothing. (УК-4.2, УК-4.3, ПКос-3.1, ПКос-3.3)
3	Тема 3 Трансформация данных Theme 3 Data transformation	Дополнение данных. Замена. Кросс-таблица. Объединение. Параметры полей. Разгруппировка. Свёртка столбцов. Скользящее окно. Слияние. Соединение. Data completion. Replacement. Cross table. An association. Field options. Ungrouping. Collapse columns. Sliding window. Merging. Compound. (УК-4.2, УК-4.3, ПКос-3.1, ПКос-3.3, ПКос-9.1, ПКос-9.2)
4	Тема 4 Предварительный анализ данных Theme 4 Preparatory data analysis	Автокорреляция исходных переменных. Коэффициент Тау-б Кендалла. Экстремум взаимокорреляционной функции. Коэффициент корреляции Спирмена. Autocorrelation of source variables. Kendall's Tau-b coefficient. Extremum of the cross-correlation function. Spearman's correlation coefficient. (УК-4.2, УК-4.3, ПКос-9.1, ПКос-9.2, ПКос-9.3)
5	Тема 5 Регрессия Theme 5 Linear Regression	Отбор факторов и защита от переобучения. Информационный критерий Акаике. Информационный критерий Байеса. Информационный критерий Ханнана-Квина. Порог значимости при добавлении фактора. Порог значимости при исключении фактора. Selection of factors and protection against overfitting. Information criterion Akaike. Bayes information criterion. Hannan-Queen information criterion. Threshold of significance when adding a factor. Threshold of significance when excluding a factor. (УК-4.2, УК-4.3, ПКос-9.1, ПКос-9.2, ПКос-9.3)
6	Тема 6 Кластеризация Theme 6 Clustering	Коэффициент отталкивания. Медианная модификация. Repulsion coefficient. Median modification. (УК-4.2, УК-4.3, ПКос-9.1, ПКос-9.2, ПКос-9.3)
9	Тема 7 ARIMAX Theme 7 ARIMAX	Порядок сезонной AR части. Порядок сезонного интегрирования. Порядок сезонной MA части. Период сезонной составляющей. The order of the seasonal AR part. The order of seasonal integration. The order of the seasonal MA part. The period of the seasonal component. (УК-4.2, УК-4.3, ПКос-3.1, ПКос-3.3, ПКос-9.1, ПКос-9.2, ПКос-9.3)
10	Тема 8 Ассоциативные правила Theme 8 Association Rules	Анализ исключений, предназначенный для выявления и толкования аномалий в найденных закономерностях. Analysis of exceptions designed to detect and interpret anomalies in the discovered patterns. (УК-4.2, УК-4.3, ПКос-9.1, ПКос-9.2, ПКос-9.3)

№ п/п	№ раздела и темы	Перечень рассматриваемых вопросов для самостоятельного изучения
11	Тема 9 Нейросеть Theme 9 Neural network	Метод валидации. Настройка параметров Нейросети. Параметры обучения. Критерии останова. Стратегия оптимизации. Validation method. Setting the parameters of the Neural Network. Learning options. Stop criteria. Optimization strategy. (УК-4.2, УК-4.3 ПКос-9.1, ПКос-9.2, ПКос-9.3)

5. Образовательные технологии

Таблица 6

Применение активных и интерактивных образовательных технологий

№ п/п	Тема и форма занятия	Наименование используемых активных и интерактивных образовательных технологий
	Лекция 6 Кластерный анализ Lecture 6 Cluster analysis	Л Лекция-визуализация
1.	Практическая работа № 5. Сортировка, объединение и разделение наборов данных Practical lesson № 5. Sorting, merging and splitting datasets	ПЗ Деловая игра

6. Текущий контроль успеваемости и промежуточная аттестация по итогам освоения дисциплины

6.1. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений и навыков и (или) опыта деятельности

Пример теста

Example of the Test

1. What are the objectives of decision support systems:

- a) data input;
- b) data storage;
- c) data analysis.

2. What task, solved with the help of decision support systems, is the main:

- a) data input;
- b) data storage;
- c) data analysis.

3. With the help of which class of analysis tasks is carried out the grouping and generalization of the necessary data analytics:

- a) information retrieval;
- b) operational analytical;
- c) intellectual.

4. With what class of analysis tasks is the data search based on predefined queries:

- a) information retrieval;
- b) operational analytical;
- c) intellectual.

5. With the help of which class of analysis tasks is the search for functional regularities in the accumulated data, construction of models and rules:

- a) information retrieval;
- b) operational analytical;
- c) intellectual.

6. Which subsystems are included in the generalized architecture of the decision support system:

- a) information retrieval system;
- b) data entry subsystem;
- c) storage subsystem;
- d) analysis subsystem;
- f) report subsystem.

7. In which subsystem is the operational processing of data based on the use of DBMS:

- a) information retrieval system;
- b) data entry subsystem;
- c) storage subsystem;
- d) analysis subsystem;
- f) report subsystem.

8. In which subsystem of the analysis subsystem, methods and algorithms for data mining are implemented:

- a) subsystem of information retrieval;
- b) operational analysis subsystem;
- c) intellectual analysis subsystem.

9. In which systems, the presence of human factors when inputting data increases the likelihood of erroneous data and can create local problems in systems:

- a) decision support system;
- b) OLTP system;
- c) database management system;
- d) economic information system;
- f) information retrieval system.

10. What systems are required to ensure the formation of arbitrary queries to the database:

- a) decision support system;
- b) OLTP system;
- c) database management system;
- d) economic information system;
- f) information retrieval system.

11. For which systems is the priority high performance and data availability:

- a) decision support system;
- b) OLTP system;
- c) database management system;
- d) economic information system;
- f) information retrieval system.

12. A data warehouse is

- a) subject-oriented, integrated, unchanging, chronologically supported data set organized to support decision-making;
- b) a model of a certain subject area, consisting of related data on objects, their properties and characteristics;
- c) a system that has the properties of input, storage and analysis of data, relating to a certain subject area, with a view to finding solutions.

13. Select the data store properties:

- a) subject orientation;
- b) integration;
- c) chronological support;
- d) unchangeability;
- e) redundancy;
- f) priority of system characteristics.

14. The main purpose of OLAP-systems:

- a) detection of hidden knowledge and patterns;
- b) support of analytical activities;
- c) support for requests of analyst users;
- d) input and storage of data.

15. Intelligent data analysis is

- a) research and detection in raw data of hidden knowledge that were not previously known, non-trivial, practically useful;
- b) The technology of operative analytical data processing, using methods and tools for collecting, storing and analyzing multidimensional data in support of decision-making processes;
- c) a subject-oriented, integrated, variable, chronological support data set organized to support decision-making;
- d) a model of a certain subject area, consisting of related data on objects, their properties and characteristics.

Вопросы к зачету с оценкой

1. What DBMS levels are characteristic now for large organizations?
2. Describe the applicability of relational DBMS in PPR systems.
3. Describe the notion of "data warehouse".
4. What features have non-normalized databases for modeling multidimensionality?
5. Describe the "Star" scheme.
6. Describe the scheme "Snowflake".
7. What are multi-dimensional databases?
8. Give comparative characteristics of multidimensional and relational DBMS.

9. What operations are provided for working with measurements in multidimensional database models?
10. Describe the notion of a data showcase.
11. Describe the technology of OLAP.
12. What is the classification of intelligent systems?
13. What are the methods of the IAP and what are their names?
14. What are the characteristics of the IAP methods?
15. Describe the areas of application of the IAP methods.
16. What are the stages of research using the methods of data mining?
17. What is the stage of preliminary processing of data using data mining?
18. Describe the method for finding the rule.
19. Describe the method of clustering.
20. What is the difference in the purposes and algorithms of statistical and intellectual approaches.
21. Describe neural network analysis methods.
22. Tell us about the classification of objects using the Kohonen network.
23. How are logical rules built in algorithms "What if something"?
24. How is the phase dimension determined to determine the number of variables representing a one-dimensional time series?
25. How is the degree of randomness of the time series determined?
26. Tell us about the horizon of forecasting and what does it depend on?
27. Classification of intellectual systems.
28. Methods of data mining. and their purpose.
29. Features of data mining methods.
30. Areas of application of data mining methods.
31. Stages of research by the methods of data mining.

Примерные тексты для перевода

Текст 1

What is a data warehouse?

A data warehouse, or enterprise data warehouse (EDW), is a system that aggregates data from different sources into a single, central, consistent data store to support data analysis, data mining, artificial intelligence (AI), and machine learning. A data warehouse system enables an organization to run powerful analytics on huge volumes (petabytes and petabytes) of historical data in ways that a standard database cannot.

Data warehousing systems have been a part of business intelligence (BI) solutions for over three decades, but they have evolved recently with the emergence of new data types and data hosting methods. Traditionally, a data warehouse was hosted on-premises—often on a mainframe computer—and its functionality was focused on extracting data from other sources, cleansing and preparing the data, and loading and maintaining the data in a relational database. More recently, a data warehouse might be hosted on a dedicated appliance or in the cloud, and most data warehouses have added analytics capabilities and data visualization and presentation tools.

Текст 2

What is data preparation?

Data preparation is, as its name suggests, a preparatory stage of big data analytics that involves cleaning and transforming the existing raw data into a format suitable for further processing. In a nutshell, data preparation refers to the process of ensuring that the data is formatted and prepared for analysis.

In order to turn raw data into valuable insights and minimize or avoid errors caused by low data quality and integrity, it is important to place data in the proper format.

Data preparation may require to remove duplicate rows or columns, merge multiple datasets together, perform string manipulation on columns containing text, apply custom logic or functions to transform the dataset from one representation into another.

Data preparation is a fundamental step in data analysis. This process is often overlooked, but it is a critical step that can dramatically improve the analysis and results. Before making any business decisions, be sure that the data is clean and accurate. There are many advantages of cleaned and correct data, including –

Access to high-quality data: Data preparation assists organizations in obtaining clean, easy-to-use, structured data and maintaining its consistency.

Improved Analysis: Day to day operations in an organization are dependent on data analysis, so clean and prepared data helps in better analysis. It reduces errors and eliminates anomalies or illogical values.

Make interpretation easier: Prepared data is simple to understand, and anyone can work with it effectively. It enables organizations to interpret data systematically.

Increased flexibility: Access to high-quality data increases flexibility. Flexibility improves the competitiveness of an organization.

Better decision making: Data preparation leads to better decision making. It is crucial to have access to accurate and reliable information to make more effective, fact-based decisions.

Increased productivity: Employees tend to be more productive when they get access to high-quality, error-free data.

Elimination of errors: During data preparation, sporadic errors can be detected in advance, and data errors can be avoided.

Текст 3

Regression models are so popular because they can incorporate many predictors in a single model (multivariate: allows to test the impact of one predictor while the impact of (all) other predictors is controlled for) extremely flexible and can be fitted to different types of predictors and dependent variables provide output that can be easily interpreted conceptually relative simple and not overly complex from a mathematical perspective

The most widely use regression models are
linear regression (dependent variable is numeric, no outliers)
logistic regression (dependent variable is binary)
ordinal regression (dependent variable represents an ordered factor, e.g. Likert items)

multinomial regression (dependent variable is categorical)
The major difference between these types of models is that they take different types of dependent variables: linear regressions take numeric, logistic regressions

take nominal variables, ordinal regressions take ordinal variables, and Poisson regressions take dependent variables that reflect counts of (rare) events. Robust regression, in contrast, is a simple multiple linear regression that is able to handle outliers due to a weighing procedure.

If regression models contain a random effect structure which is used to model nestedness or dependence among data points, the regression models are called mixed-effect models. regressions that do not have a random effect component to model nestedness or dependence are referred to as fixed-effect regressions (we will have a closer look at the difference between fixed and random effects below).

There are two basic types of regression models:

fixed-effects regression models

mixed-effects regression models

Fixed-effects regression models are models that assume a non-hierarchical data structure, i.e. data where data points are not nested or grouped in higher order categories (e.g. students within classes). The first part of this tutorial focuses on fixed-effects regression models while the second part focuses on mixed-effects regression models.

Текст 4

Association rule mining is one of the major concepts of Data mining and Machine learning, it is simply used to identify the occurrence pattern in a large dataset. We establish a set of rules to find out how the positioning of different items is affecting each other. These patterns can be of any type like Telephone contacting pattern, suspicious activity pattern, pattern in symptoms of any disease, and customer shopping pattern. Here we will focus on customer shopping patter, a more suitable term is Market Basket Analysis.

Market Basket Analysis is one of the most popular techniques in finding the best product placement in the store and deciding offers which increases the overall sales. The idea is to bring the set of products together that have some kind of interdependency in terms of their use. Doing so can surely boost the sale because placing them together will remind or encourage customers about their need for the associated product. To solve this, we develop all the possible sets of product's association rules and find out the most effective ones. Now the question is how do we develop and find the effectiveness of these association rules? The answer is the Apriori Algorithm.

Пример работ

Практическая работа № 1. Принципы работы с хранилищами данных
На основе исходных данных спроектировать хранилище данных. Установить связь между измерениями, определить факты, отражающие сущность события, соответствующие измерениям. Изобразить схему «снежинка» и загрузить данные используя аналитическую платформу Loginom. Представить данные в виде многомерных кубов, называемых также OLAP-кубами или гиперкубами. По итогам выполнения работы предоставить отчет с выводами.

Вопросы к защите:

1. Что Вы понимаете под хранилищем данных?
2. Что Вы понимаете под базой данных?
3. Основные отличия баз данных от хранилищ.
4. Модели организации баз данных

5. Модели организации хранилищ данных

6. Что такое OLAP-кубам?

Practical lesson № 1. Principles of working with data warehouses

Based on the initial data, design a data warehouse. Establish a connection between measurements, determine the facts that reflect the essence of the event, corresponding to the measurements. Draw a snowflake diagram and upload data using the Loginom analytical platform. Represent data as multidimensional cubes, also called OLAP cubes or hypercubes. Based on the results of the work, provide a report with conclusions.

Questions:

1. What do you mean by data warehouse?

2. What do you mean by a database?

3. The main differences between databases and storage.

4. Database organization models

5. Data Warehouse Management Models

6. What is an OLAP cube?

Практическая работа № 2. Интеграция с различными базами данных

Подключить Loginom к имеющимся в исходных данных базам данным. Перед тем как использовать в сценариях узлы взаимодействия с SOAP-сервисами/REST-сервисами и импорта/экспорта из бизнес приложений, хранилищ данных, баз данных или иных источников, в пакете необходимо создать подключение к источнику данных. Подключение содержит в себе все необходимые параметры для соединения с источником данных, такие как пароль/логин, расположение источника и др. Созданное подключение в дальнейшем может использоваться в пакете многократно. По итогам выполнения работы предоставить отчет с выводами.

Вопросы к защите:

1. С какими базами данных можно настроить подключение Loginom

2. Какие возможности есть в меню редактирования подключения?

3. Каким образом осуществляется навигация по подключениям?

Practical lesson № 2. Integration with various databases

Connect Loginom to the databases available in the source data. Before you can use SOAP/REST service interaction nodes and import/export nodes from business applications, data stores, databases, or other sources in your scripts, you must create a data source connection in your package. The connection contains all the necessary parameters for connecting to the data source, such as password / login, source location, etc. The created connection can later be used repeatedly in the package. At the end of the work, provide a report with conclusions.

Questions:

1. Which databases can be used to set up a Loginom connection?

2. What options are there in the connection editing menu?

3. How are connections navigated?

Практическая работа № 3. Импорт данных в различных форматах

Требуется импортировать исходные данные в Loginom в виде текстовых файлов, электронных таблиц, баз данных и других. Настроить параметры загрузки.

Определить нужный формат импортируемых данных. По итогам выполнения работы предоставить отчет с выводами.

Вопросы к защите:

1. Какие типы файлов можно импортировать в Loginom?
2. Какие возможности есть в меню редактирования импорта?
3. В чем особенность импорта электронных таблиц?

Practical lesson № 3. Import data in various formats

It is required to import the source data into Loginom in the form of text files, spreadsheets, databases, and others. Set download options. Determine the desired format of the imported data. At the end of the work, provide a report with conclusions.

Questions:

1. What types of files can be imported into Loginom?
2. What options are there in the import editing menu?
3. What is special about importing spreadsheets?

Практическая работа № 4. Подготовка данных для анализа

По результатам загрузки данных прошлого занятия настроить автоматическое заполнение пропущенных значений в наборах данных. Для каждого столбца исходного набора данных пользователь может выбрать наиболее подходящий метод заполнения пропусков. Пропусками считаются Null-значения. Настроить программу для автоматической корректировки выбросов и экстремальных значений в наборах данных. Для каждого поля исходного набора данных критерии определения выбросов и экстремальных значений задаются пользователем с помощью указания допустимого стандартного отклонения или интерквартильного размаха. Осуществить разбиение данных на обучающую и тестовую выборки. По итогам выполнения работы предоставить отчет с выводами.

Вопросы к защите:

1. Методы заполнения пропущенных данных.
2. Что понимается под выбросами?
3. Методы определения выбросов.
4. Методы разбиения множества на выборки

Practical lesson № 4. Data preparation for analysis

Based on the results of loading data from the last lesson, set up automatic filling in of missing values in data sets. For each column of the original dataset, the user can choose the most appropriate method for filling in the gaps. Null values are considered gaps. Set up a program to automatically correct for outliers and extremes in datasets. For each field of the original data set, the criteria for determining outliers and extreme values are set by the user by specifying an acceptable standard deviation or interquartile range. Split the data into training and test sets. At the end of the work, provide a report with conclusions.

Questions:

1. Methods for filling in missing data.
2. What is meant by emissions?
3. Methods for determining outliers.

4. Methods for partitioning a set into samples

Практическая работа № 5. Сортировка, объединение и разделение наборов данных

По данным прошлого занятия провести группировку данных, разгруппировку, замену данных, слияние и свертку столбцов. Построить кросс-таблицу. Настроить параметры полей. Осуществить сортировку и фильтрацию строк. По итогам выполнения работы предоставить отчет с выводами.

Вопросы к защите:

1. Что Вы понимаете под группировкой данных?
2. Для чего используется замена данных?
3. Способы слияния и свертки столбцов?
4. Что такое кросс-таблица?

Practical lesson № 5. Sorting, merging and splitting datasets

According to the last lesson, group data, ungroup, replace data, merge and collapse columns. Build a crosstab. Customize margin settings. Sorting and filtering rows. At the end of the work, provide a report with conclusions.

Questions:

1. What do you understand by data grouping?
2. What is data replacement used for?
3. Ways to merge and collapse columns?
4. What is a crosstab?

Практическая работа № 6. Выявление дубликатов и противоречий во входных данных

По данным прошлого занятия провести анализ исходных данных на наличие дублирующих, противоречивых записей. По итогам выполнения работы предоставить отчет с выводами.

Вопросы к защите:

1. Что Вы понимаете под противоречивыми записями?
2. Для чего выявлять дубликаты в исходных данных?
3. Способы появления противоречивых и дублирующих данных?
4. Методы выявления противоречивых и дублирующих данных.

Practical lesson № 6. Identification of duplicates and contradictions in the input data

According to the last lesson, analyze the source data for the presence of duplicate, conflicting records. At the end of the work, provide a report with conclusions.

Questions:

1. What do you understand by conflicting records?
2. Why identify duplicates in source data?
3. How do conflicting and duplicate data appear?
4. Methods for identifying conflicting and duplicate data.

Практическая работа № 7. Предварительный анализ данных

По данным прошлого занятия провести анализ данных основе корреляционного анализа. Построить матрицу парных коэффициентов корреляции. Определить взаимосвязи между признаками. Проверить исходные данные на нали-

чие автокорреляции. По итогам выполнения работы предоставить отчет с выводами.

Вопросы к защите:

1. Что такое корреляция?
2. Что показывает коэффициент корреляции?
3. Методы отбора факторов для построения регрессионных моделей?
4. Что такое автокорреляция?
5. Способы выявления автокорреляции.
6. Способы устранения автокорреляции.

Practical lesson № 7. Preparatory data analysis

According to the last lesson, analyze the data on the basis of correlation analysis. Build a matrix of paired correlation coefficients. Determine relationships between features. Check the original data for autocorrelation. Based on the results of the work, provide a report with conclusions.

Questions:

1. What is a correlation?
2. What does the correlation coefficient show?
3. Methods for selecting factors for building regression models?
4. What is autocorrelation?
5. Methods for detecting autocorrelation.
6. Ways to eliminate autocorrelation.

Практическая работа № 8. Парная линейная регрессия

На основе результатов прошлого занятия построить уравнение парной линейной регрессии. Сделать прогноз по уравнению. Определить качество построенной модели по тестовым данным. Сравнить метрики качества моделей без обучения и с обучением. Оформить отчет с выводами.

Вопросы к защите:

1. Назовите цель и задачи построения парной линейной модели регрессии.
2. Дайте интерпретацию коэффициента регрессии и условного начала в уравнении парной линейной регрессии.
3. Как проводить прогноз в системе Loginom по парной линейной модели?
4. Особенности построения парной линейной регрессии в системе Loginom.

Practical lesson № 8. Paired Linear Regression

Based on the results of the last lesson, build a paired linear regression equation. Make a prediction using the equation. Determine the quality of the built model according to the test data. Compare the quality metrics of untrained and trained models. Prepare a report with conclusions.

Questions:

1. Name the purpose and objectives of building a paired linear regression model.
2. Give an interpretation of the regression coefficient and the conditional beginning in the paired linear regression equation.
3. How to make a forecast in the Loginom system using a paired linear model?
4. Features of building a paired linear regression in the Loginom system.

Практическая работа № 9. Множественная линейная регрессия

На основе результатов занятия №7 построить множественную регрессионную модель регрессии. Построить прогноз, сделать выводы. Определить качество построенной модели по тестовым данным. Сравнить метрики качества моделей без обучения и с обучением. В результате выполнения работы необходимо оформить отчет с выводами.

Вопросы к защите:

1. Как проводить прогноз в системе Loginom по множественной линейной модели.
2. Стандартизированные коэффициенты множественной линейной регрессии. Интерпретация, особенности расчета.
3. Способы отбора факторов в уравнение множественной линейной регрессии

Practical lesson № 9. Multiple Linear Regression

Based on the results of lesson No. 7, build a multiple regression model. Build a forecast, draw conclusions. Determine the quality of the constructed model from the test data. Compare the quality metrics of models without training and with training. As a result of the work, it is necessary to draw up a report with conclusions.

Questions:

1. How to make a forecast in the Loginom system using a multiple linear model.
2. Standardized multiple linear regression coefficients. Interpretation, features of the calculation.
3. Methods for selecting factors in the equation of multiple linear regression

Практическая работа № 10. Алгоритм CLOPE.

Загрузить исходные данные о посещениях веб-ресурсов в Loginom. Применить к ним алгоритм CLOPE. Задача состоит в получении такого разбиения всего множества транзакций, чтобы похожие транзакции оказались в одном кластере, а отличающиеся друг от друга - в разных кластерах. В результате выполнения работы необходимо оформить отчет с выводами.

Вопросы к защите:

1. Что такое кластерный анализ?
2. В чем особенность алгоритма CLOPE?
3. Интерпретация результатов алгоритма CLOPE.

Practical lesson № 10. CLOPE algorithm.

Upload initial data on visits to web resources in Loginom. Apply the CLOPE algorithm to them. The task is to obtain such a partition of the entire set of transactions so that similar transactions end up in one cluster, and different ones - in different clusters. As a result of the work, it is necessary to draw up a report with conclusions.

Questions:

1. What is cluster analysis?
2. What is the peculiarity of the CLOPE algorithm?

3. Interpretation of the results of the CLOPE algorithm.

Практическая работа № 11. Алгоритм кластеризации EM

Загрузить исходные данные в Loginom. Применить к ним алгоритм кластеризации EM. Определить оптимальное число кластеров. Определить вероятность отнесения объекта к каждому кластеру. В результате выполнения работы необходимо оформить отчет с выводами.

Вопросы к защите:

1. Что такое кластерный анализ?
2. В чем особенность алгоритма кластеризации EM?
3. Интерпретация результатов алгоритма кластеризации EM.

Practical lesson № 11. EM Clustering Algorithm

Upload initial data to Loginom. Apply the EM clustering algorithm to them. Determine the optimal number of clusters. Determine the probability of assigning an object to each cluster. As a result of the work, it is necessary to draw up a report with conclusions.

Questions:

1. What is cluster analysis?
2. What is the feature of the EM clustering algorithm?
3. Interpretation of the results of the EM clustering algorithm.

Практическая работа № 12. Анализ временных рядов методом ARIMAX

На сайте <https://rosstat.gov.ru/> выбрать данные с 1990 года и применить к ним метод ARIMAX для прогнозирования. Оценить качество построенной модели. Оценить качества прогноза. Осуществить прогноз по ретроспективным данным и сравнить с результатами работы модели. В результате выполнения работы нужно оформить отчет с выводами.

Вопросы к защите:

1. Назовите методы, применяемые для выравнивания рядов динамики
2. В чем особенность метода ARIMAX?
3. Определение параметров метода ARIMAX.

Practical lesson № 12. Time series analysis using the ARIMAX method

On the site <https://rosstat.gov.ru/>, select data from 1990 and apply the ARIMAX method to them for forecasting. Evaluate the quality of the constructed model. Assess the quality of the forecast. Make a forecast based on historical data and compare with the results of the model. As a result of the work, you need to draw up a report with conclusions.

Questions:

1. Name the methods used to align time series
2. What is the peculiarity of the ARIMAX method?
3. Determining the parameters of the ARIMAX method.

Практическая работа № 13 Поиск ассоциативных правил

По набору данных о покупках клиентов определить наиболее часто встречающиеся товары и наибольшую вероятность совместных покупок товаров. Какие товары не встречаются в одном чеке? Оформить отчет с выводами.

Вопросы к защите:

1. Цель поиска ассоциативных правил
2. Примеры применения метода поиска ассоциативных правил
3. Особенности применения метода поиска ассоциативных правил
4. Качество выявленных взаимосвязей
5. Расчет и интерпретация ошибок

Practical lesson № 13. Search for association rules

Based on a set of customer purchases, determine the most common products and the greatest likelihood of joint purchases of goods. What products are not found in one check? Make a report with conclusions.

Questions:

1. The purpose of the search for associative rules
2. Examples of application of the association rule search method
3. Features of application of the search method of associative rules
4. The quality of the identified relationships
5. Calculation and interpretation of errors

Практическая работа № 14 Применение нейронной сети для регрессии

Построить модель регрессии на основе искусственных нейронных сетей. Сравнить результаты классификации при разном уровне скрытых слоев. Сравнить качество классификации с предыдущими методами. Сравнить результаты построенных моделей. Оформить отчет с выводами.

1. Суть нейронных сетей
2. Обучение модели
3. Суть искусственного нейрона
4. Определение количества слоев
5. Определение числа нейронов

Practical lesson № 14. Applying a Neural Network for Regression

Build a regression model based on artificial neural networks. Compare classification results at different levels of hidden layers. Compare classification quality with previous methods. Compare the results of the constructed models. Prepare a report with conclusions.

1. The essence of neural networks
2. Model training
3. The essence of an artificial neuron
4. Determining the number of layers
5. Determining the number of neurons

6.2. Описание показателей и критериев контроля успеваемости, описание шкал оценивания

Текущий контроль знаний, умений и навыков проводится в форме тестирования и теоретическими вопросами. Оценка работ проводится по стобалльной шкале. Индивидуальные задачи, выполняемые каждым студентом на практике оцениваются по итогам защиты по аналогичной шкале. Ликвидация сту-

дентами текущих задолженностей производится также в форме выполнения индивидуальной задачи по соответствующей теме и дальнейшей ее защиты преподавателю кафедры с оценкой по стобалльной шкале.

Для получения зачета с оценкой необходимо набрать более 60%. Вид промежуточного контроля по данному направлению – зачет с оценкой.

Градация оценок:

0 – 60% - «неудовлетворительно»;

60 – 75 – «удовлетворительно»;

75 – 85 – «хорошо»;

85 – 100 – «отлично»

Формы контроля: тестовый контроль, индивидуальное собеседование, защита выполнения практического задания по индивидуальному варианту. В итоговую сумму баллов входят результаты всех контролируемых видов вашей деятельности – посещение занятий, выполнение заданий, прохождение тестов, активность на лабораторных занятиях и т.п.

В итоговый рейтинг входит: 30% - результат выполнения контрольных мероприятий (тест, самостоятельные работы и др.), 60% - баллы за сданные индивидуальные работы и 10% - посещение занятий.

При изучении каждого раздела дисциплины проводится промежуточный контроль знаний с целью проверки и коррекции хода освоения теоретического материала и практических умений и навыков.

7. Учебно-методическое и информационное обеспечение дисциплины

7.1 Основная литература

1. Митина, О. А. Технологии организации, обработки и хранения статистических данных : учебное пособие / О. А. Митина, И. А. Юрченков. — Москва : РТУ МИРЭА, 2019. — 163 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/171511> (дата обращения: 27.11.2022). — Режим доступа: для авториз. пользователей.

2. Прокопенко, Н. Ю. Аналитические информационные системы поддержки принятия решений : учебное пособие / Н. Ю. Прокопенко. — Нижний Новгород : ННГАСУ, 2020. — 142 с. — ISBN 978-5-528-00395-5. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/164866> (дата обращения: 27.11.2022). — Режим доступа: для авториз. пользователей.

3. Скляр, А. Я. Технология хранения и интерактивных обработки данных : учебное пособие / А. Я. Скляр. — Москва : РТУ МИРЭА, 2020. — 69 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/163914> (дата обращения: 27.11.2022). — Режим доступа: для авториз. пользователей.

4. Орешков, В. И. Хранилища данных и OLAP-технологии : учебное пособие / В. И. Орешков. — Рязань : РГРТУ, 2017. — 64 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL:

<https://e.lanbook.com/book/167981> (дата обращения: 27.11.2022). — Режим доступа: для авториз. пользователей.

5. Data Science / Francesco Palumbo, Angela Montanari, Maurizio Vichi. Springer International Publishing AG, 2017 – Текст : электронный // Springer: электронно-библиотечная система. URL: <https://link.springer.com/book/10.1007/978-3-319-55723-6#editorsandaffiliations> (дата обращения: 27.11.2022).

7.2 Дополнительная литература

1. Бессмертный, И. А. Интеллектуальные системы : учебник и практикум для вузов / И. А. Бессмертный, А. Б. Нугуманова, А. В. Платонов. — Москва : Издательство Юрайт, 2022. — 243 с. — (Высшее образование). — ISBN 978-5-534-01042-8. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/490020> (дата обращения: 27.11.2022).

2. Станкевич, Л. А. Интеллектуальные системы и технологии : учебник и практикум для среднего профессионального образования / Л. А. Станкевич. — Москва : Издательство Юрайт, 2022. — 397 с. — (Профессиональное образование). — ISBN 978-5-534-11659-5. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/495988> (дата обращения: 27.11.2022).

3. New Advances in Statistics and Data Science / Ding-Geng, Chen Zhezhen, Jin Gang, Li Yi, Li Aiyi, Liu Yichuan, Zhao. Springer International Publishing AG, 2017 – Текст : электронный // Springer: электронно-библиотечная система. URL: <https://link.springer.com/book/10.1007/978-3-319-69416-0#editorsandaffiliations> (дата обращения: 27.11.2022).

4. Точилкина, Т. Е. Хранилища данных и средства бизнес-аналитики : учебное пособие / Т. Е. Точилкина, А. А. Громова. — Москва : Финансовый университет, 2017. — 161 с. — ISBN 978-5-7942-1387-4. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/208367> (дата обращения: 27.11.2022). — Режим доступа: для авториз. пользователей.

7.3 Методические указания, рекомендации и другие материалы к занятиям

1. Харитонова, А.Е. Хранилища и системы интеллектуального анализа данных: методические указания / А.Е. Харитонова. – М.: РГАУ-МСХА им. К.А. Тимирязева, 2016. – 25 с.

8. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины

1. Loginom. URL: <https://loginom.ru/>(открытый доступ)
2. Справочник Loginom. URL: <https://help.loginom.ru/userguide/>(открытый доступ)
3. Официальный сайт Росстата. URL: <https://rosstat.gov.ru/> (открытый доступ)

4. Официальный сайт Центрального Банка России. URL: <http://www.cbr.ru> (открытый доступ)
5. Bureau of Economic Analysis. URL: <http://www.bea.gov> (открытый доступ)
6. Московская международная валютная биржа. <http://www.micex.ru> (открытый доступ)
7. Официальный сайт Всемирного банка . URL: <http://www.worldbank.org> (открытый доступ)
8. Официальный сайт Министерства финансов РФ. URL: <http://www.minfin.gov.ru> (открытый доступ)
9. Официальный сайт Национального бюро экономических исследований США. URL: <http://www.nber.org> (открытый доступ)

9. Перечень программного обеспечения и информационных справочных систем

Таблица 9

Перечень программного обеспечения

№ п/п	Наименование раздела учебной дисциплины	Наименование программы	Тип программы	Автор	Год разработки
1	Тема 1 Технология хранения данных при принятии решений Theme 1 Technology of data storage in decision-making Тема 2 Предобработка и консолидация данных Theme 2 Data preprocessing and consolidation Тема 3 Трансформация данных Theme 3 Data transformation Тема 4 Предварительный анализ данных Theme 4 Preparatory data analysis Тема 5 Регрессия Theme 5 Linear Regression Тема 6 Кластеризация Theme 6 Clustering Тема 7 ARIMAX Theme 7 ARIMAX Тема 8 Ассоциативные правила Theme 8 Association Rules Тема 9 Нейросеть Theme 9 Neural network	Loginom	Аналитическая платформа	Loginom Company	Текущая версия

10. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине

Таблица 10

Сведения об обеспеченности специализированными аудиториями, кабинетами, лабораториями

Наименование специальных помещений и помещений для самостоятельной работы (№ учебного корпуса, № аудитории)	Оснащенность специальных помещений и помещений для самостоятельной работы
1	2
<p><i>учебная аудитория для проведения занятий лекционного типа, учебная аудитория для проведения занятий семинарского типа, учебная аудитория для групповых и индивидуальных консультаций, учебная аудитория для текущего контроля и промежуточной аттестации</i></p>	<ol style="list-style-type: none"> 1. Экран с электроприводом 1 шт. (Инв. №558771/2) 2. Проектор 1 шт. (без инв. №) – приобретался не за счет средств вуза 3. Вандалоустойчивый шкаф 1 шт. (Инв.№558850/7) 4. Системный блок с монитором 1 шт. (Инв. №558777/9) 5. Стенд «Сергеев Сергей Степанович 1910-1999» 1 шт. (Инв.№591013/25) 6. Огнетушитель порошковый 1 шт. (Инв. №559527) 7. Подвесное крепление к огнетушителю 1 шт. (Инв. № 559528) 8. Жалюзи 2шт. (Инв. №1107-221225, Инв. №1107-221225) 9. Лавка 20 шт. 10. Стол аудиторный 20 шт. 11. Стол для преподавателя 1 шт. 12. Стул 2 шт. 13. Доска маркерная 1 шт. 14. Трибуна напольная 1 шт. (без инв. №)
<p><i>учебная аудитория для проведения занятий семинарского типа, учебная аудитория для групповых и индивидуальных консультаций, учебная аудитория для текущего контроля и промежуточной аттестации, помещение для самостоятельной работы</i></p>	<ol style="list-style-type: none"> 1. Системный блок Intel Core Intel Core i3-2100/4096Mb/500Gb/DVD-RW 10 шт. (Инв.№601997, Инв.№601998, Инв.№601999, Инв.№602000, Инв.№602001, Инв.№602002, Инв.№602003, Инв.№602004, Инв.№602005, Инв.№602006) 2. Монитор 10 шт. (без инв. №) - приобретались не за счет средств вуза 3. Шкаф 2 шт. (Инв.№594166, Инв.№594167) 4. Тумба 1 шт. (Инв.№594168) 5. Подвесное крепление к огнетушителю 1 шт. (Инв. № 559528) 6. Огнетушитель порошковый 1 шт. (Инв. №559527) 7. Жалюзи 1 шт. (Инв.№551557) 8. Доска магнитно-маркерная 1 шт. 9. Стол 5 шт. 10. Стол компьютерный 12 шт. 11. Стул офисный 21 шт. 12. Сейф 1 шт. (без Инв.№).
<p><i>учебная аудитория для проведения занятий лекционного типа, учебная аудитория для проведения занятий семинарского типа, учебная аудитория для групповых и индивидуальных консультаций, учебная аудитория для текущего контроля и промежуточной аттестации, помещение для самостоятельной работы</i></p>	<ol style="list-style-type: none"> 1. Трибуна напольная 1 шт. (Инв.№ 599205) 2. Шкаф для документов 3 шт. (Инв.№593633, Инв.№593634, Инв.№559548/18) 3. Вешалка напольная 2 шт. (Инв.№1107-333144, Инв.№1107-333144) 4. Жалюзи 1 шт. (Инв.№591110) 5. Доска магнитно-маркерная 1 шт. 6. Стол 15 шт. 7. Скамейка 14 шт. 8. Стол эрго 1 шт. 9. Стул 2 шт.
<p><i>Центральная научная библиотека имени Н.И. Железнова</i></p>	<p>Читальные залы библиотеки</p>
<p><i>Студенческое общежитие</i></p>	<p>Комната для самоподготовки</p>

11. Методические рекомендации обучающимся по освоению дисциплины

Все виды учебных работ должны быть выполнены точно в сроки, предусмотренные программой обучения. Если студент не выполнил какое-либо из учебных заданий по неуважительной причине (пропустил тестовый контроль, не выполнили домашнего задания, выполнил работу не по своему варианту и т.п.), то за данный вид учебной работы баллы рейтинга не начисляются, а подготовленные позже положенного срока работы оцениваются с понижающим коэффициентом. Если же невыполнение учебных работ произошло по уважительной причине, то следует представить преподавателю подтверждающий документ, и защитить пропущенные занятия в часы, отведенные для еженедельных консультаций.

Виды и формы отработки пропущенных занятий

Студент, пропустивший занятия обязан выполнить самостоятельно индивидуальную работу, выполняемую на занятиях по своему варианту.

12. Методические рекомендации преподавателям по организации обучения по дисциплине

Курс должен давать не абстрактно-формальные, а прикладные знания. Данная цель может быть реализована только при условии соблюдения в учебных планах преемственности учебных дисциплин. Базовые знания для изучения дисциплины дают такие предметы, как экономическая теория, информатика.

Преподаватель должен указывать, в какой последовательности следует изучать материал дисциплины, обращать внимание на особенности изучения отдельных тем и разделов, помогать отбирать наиболее важные и необходимые сведения из учебных пособий, а также давать объяснения вопросам программы курса, которые обычно вызывают затруднения. При этом преподавателю необходимо учитывать следующие моменты:

1. Не следует перегружать студентов творческими заданиями.
2. Чередовать творческую работу на занятиях с заданиями во внеаудиторное время.
3. Давать студентам четкий инструктаж по выполнению самостоятельных заданий: цель задания; условия выполнения; объем; сроки; требования к оформлению.
4. Осуществлять текущий учет и контроль за самостоятельной работой.

5. Давать оценку обобщать уровень усвоения навыков самостоятельной, творческой работы.

Программу разработал (и):

Харитонов А.Е., к.э.н., доцент



(подпись)

РЕЦЕНЗИЯ

на рабочую программу дисциплины Б1.В.ДВ.03.01 «Хранилища и системы интеллектуального анализа данных на иностранном языке»

ОПОП ВО по направлению 09.03.02 «Информационные системы и технологии», направленность «Большие данные и машинное обучение (Machine Learning & Big Data)»

(квалификация выпускника – бакалавр)

Коломеева Елена Сергеевна, доцент кафедры финансов ФГБОУ ВО г. Москвы «РГАУ-МСХА имени К.А. Тимирязева», кандидатом экономических наук (далее по тексту рецензент), проведена рецензия рабочей программы дисциплины «Хранилища и системы интеллектуального анализа данных на иностранном языке» ОПОП ВО по направлению 09.03.02 «Информационные системы и технологии», направленность «Большие данные и машинное обучение (Machine Learning & Big Data)» (бакалавриат) разработанной в ФГБОУ ВО «Российский государственный аграрный университет – МСХА имени К.А. Тимирязева», на кафедре статистики и кибернетики (разработчик – Харитоновна Анна Евгеньевна, кандидат экономических наук, доцент кафедры статистики и кибернетики).

Рассмотрев представленные на рецензирование материалы, рецензент пришел к следующим выводам:

1. Предъявленная рабочая программа дисциплины «Хранилища и системы интеллектуального анализа данных на иностранном языке» (далее по тексту Программа) соответствует требованиям ФГОС ВО по направлению 09.03.02 «Информационные системы и технологии». Программа содержит все основные разделы, соответствует требованиям к нормативно-методическим документам.

2. Представленная в Программе актуальность учебной дисциплины в рамках реализации ОПОП ВО не подлежит сомнению – дисциплина относится к дисциплинам по выбору части, формируемой участниками образовательных отношений учебного цикла – Б1.В.ДВ.

3. Представленные в Программе цели дисциплины соответствуют требованиям ФГОС ВО направления 09.03.02 «Информационные системы и технологии».

4. В соответствии с Программой за дисциплиной «Хранилища и системы интеллектуального анализа данных на иностранном языке» закреплено 3 компетенции (7 индикаторов). Дисциплина «Хранилища и системы интеллектуального анализа данных на иностранном языке» и представленная Программа способна реализовать их в объявленных требованиях.

5. Общая трудоёмкость дисциплины «Хранилища и системы интеллектуального анализа данных на иностранном языке» составляет 3 зачётных единицы (108 часов/из них практическая подготовка 4 ч.).

6. Информация о взаимосвязи изучаемых дисциплин и вопросам исключения дублирования в содержании дисциплин соответствует действительности. Дисциплина «Хранилища и системы интеллектуального анализа данных на иностранном языке» взаимосвязана с другими дисциплинами ОПОП ВО и Учебного плана по направлению 09.03.02 «Информационные системы и технологии» и возможность дублирования в содержании отсутствует.

7. Представленная Программа предполагает использование современных образовательных технологий, используемые при реализации различных видов учебной работы. Формы образовательных технологий соответствуют специфике дисциплины.

8. Программа дисциплины «Хранилища и системы интеллектуального анализа данных на иностранном языке» предполагает 2 часа занятий в интерактивной форме.

9. Виды, содержание и трудоёмкость самостоятельной работы студентов, представленные в Программе, соответствуют требованиям к подготовке выпускников, содержащимся во ФГОС ВО направления 09.03.02 «Информационные системы и технологии».

10. Представленные и описанные в Программе формы текущей оценки знаний (опрос, как в форме обсуждения отдельных вопросов, так и участие в деловых играх), соответствуют специфике дисциплины и требованиям к выпускникам.

Форма промежуточного контроля знаний студентов, предусмотренная Программой, осуществляется в форме зачета с оценкой, что соответствует статусу дисциплины, как дисциплины по выбору части, формируемой участниками образовательных отношений учебного цикла – Б1.В.ДВ ФГОС ВО направления 09.03.02 «Информационные системы и технологии».

Формы оценки знаний, представленные в Программе, соответствуют специфике дисциплины и требованиям к выпускникам.

11. Учебно-методическое обеспечение дисциплины представлено: основной литературой – 5 источников (базовый учебник), дополнительной литературой – 4 наименования, Интернет-ресурсы – 9 источников и соответствует требованиям ФГОС ВО направления 09.03.02 «Информационные системы и технологии».

12. Материально-техническое обеспечение дисциплины соответствует специфике дисциплины «Хранилища и системы интеллектуального анализа данных на иностранном языке» и обеспечивает использование современных образовательных, в том числе интерактивных методов обучения.

13. Методические рекомендации студентам и методические рекомендации преподавателям по организации обучения по дисциплине дают представление о специфике обучения по дисциплине «Хранилища и системы интеллектуального анализа данных на иностранном языке».

ОБЩИЕ ВЫВОДЫ

На основании проведенного рецензирования можно сделать заключение, что характер, структура и содержание рабочей программы дисциплины «Хранилища и системы интеллектуального анализа данных на иностранном языке» ОПОП ВО по направлению 09.03.02 «Информационные системы и технологии», направленность **«Большие данные и машинное обучение (Machine Learning & Big Data)»** (квалификация выпускника – бакалавр), разработанная Харитоновой А. Е., к.э.н., доцентом кафедры статистики и кибернетики, соответствует требованиям ФГОС ВО, современным требованиям экономики, рынка труда и позволит при её реализации успешно обеспечить формирование заявленных компетенций.

Рецензент: Коломеева Е.С., доцент кафедры финансов ФГБОУ ВО «Российский государственный аграрный университет – МСХА имени К.А. Тимирязева», кандидат экономических наук _____ « 26 » _____ 2022 г.

(подпись)