

**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ**

**Федеральное государственное автономное
образовательное учреждение высшего образования
«Национальный исследовательский Нижегородский государственный университет
им. Н.И. Лобачевского»**

Дзержинский филиал ННГУ

УТВЕРЖДЕНО

решением Ученого совета ННГУ

(протокол от «30» ноября 2022 г. № 13)

Рабочая программа дисциплины

ТЕХНОЛОГИИ ОБРАБОТКИ БОЛЬШИХ ДАННЫХ

Уровень высшего образования

БАКАЛАВРИАТ

Направление подготовки

09.03.03 ПРИКЛАДНАЯ ИНФОРМАТИКА

Направленность (профиль) образовательной программы

**ИТ-СЕРВИСЫ И ТЕХНОЛОГИИ ОБРАБОТКИ ДАННЫХ В ЭКОНОМИКЕ
И ФИНАНСАХ**

Год набора: 2023

Квалификация

БАКАЛАВР

Форма обучения

ОЧНАЯ, ОЧНО-ЗАОЧНАЯ

Дзержинск
2022 г.

1. Место дисциплины в структуре ОПОП

Дисциплина Б1.В.10 «Технологии обработки больших данных» относится к части ООП направления подготовки 09.03.03 Прикладная информатика, формируемой участниками образовательных отношений.

Целью изучения дисциплины «Технологии обработки больших данных» является приобретение обучающимися теоретических и практических основ обработки и глубокого анализа данных, необходимых в современных условиях при описании, анализе и прогнозировании реальных экономических процессов.

2. Планируемые результаты обучения по дисциплине, соотнесенные с планируемыми результатами освоения образовательной программы (компетенциями и индикаторами достижения компетенций)

Формируемые компетенции (код, содержание компетенции)	Планируемые результаты обучения по дисциплине (модулю), в соответствии с индикатором достижения компетенции		Наименование оценочного средства
	Индикатор достижения компетенции (код, содержание индикатора)	Результаты обучения по дисциплине	
УК-1 Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	УК-1.1. Демонстрирует знание принципов сбора, отбора и обобщения информации, базирующихся на системном подходе	Знать: понятие большие данные, методы сбора и обобщения информации. Уметь: обрабатывать и управлять большими объемами постоянно обновляющейся информации. Владеть: навыками сбора и обработки больших данных.	Собеседование, тестирование, отчеты по лабораторным работам
	УК-1.2. Демонстрирует умение соотносить разнородные явления и систематизировать их в рамках избранных видов профессиональной деятельности.	Знать: виды обработки больших данных: синхронная и асинхронная, пакетная и поточная. Уметь: решать стандартные задачи профессиональной деятельности на основе теории вероятностей и математической статистики. Владеть: навыками применения методов и средств обеспечения интеллектуального анализа данных.	Собеседование, тестирование, отчеты по лабораторным работам
	УК-1.3. Демонстрирует наличие практического опыта работы с информационными источниками, опыта научного поиска и представ-	Знать: технологии работы с информационными источниками. Уметь: работать с информационными источниками. Владеть: практическим опытом работы с информационными источниками, навыками представления результатов обработки	Собеседование, отчеты по лабораторным работам

Формируемые компетенции (код, содержание компетенции)	Планируемые результаты обучения по дисциплине (модулю), в соответствии с индикатором достижения компетенции		Наименование оценочного средства
	Индикатор достижения компетенции (код, содержание индикатора)	Результаты обучения по дисциплине	
	ления научных результатов.	данных	
ПК-9 Способен моделировать прикладные (бизнес) процессы и объекты предметной области	ПК-9.1. Демонстрирует знание методических основ моделирования процессов и объектов предметной области	Знать: методики и алгоритмы анализа и обработки больших данных с использованием современных языков программирования высокого уровня Уметь: применять методы математического анализа и моделирования к анализу больших данных. Владеть: приемами и алгоритмами обработки данных, и построения моделей восстановления данных.	Собеседование, тестирование, отчеты по лабораторным работам
	ПК-9.2 Демонстрирует умение применения знаний к моделированию прикладных процессов и объектов предметной области при разработке программного обеспечения ИС	Знать: математические основы построения моделей анализа и обработки данных, способы извлечения дополнительной информации из массив данных. Уметь: обрабатывать данные, выявлять скрытые зависимости и закономерности в больших данных, извлеченных непосредственно или полученных в результате эксперимента. Владеть: навыками разработки моделей машинного обучения.	Собеседование, тестирование, отчеты по лабораторным работам
	ПК-9.3 Имеет практический опыт моделирования процессов и объектов на примере конкретной предметной области.	Знать: современные инструменты, пакеты программ, и технологии обработки и анализа данных. Уметь: настраивать программное обеспечение, обеспечивающее доступ к удаленным данным, в том числе с применением облачных технологий. Владеть: навыками и средствами экспериментов с использованием современных средств работы с большими данными.	Собеседование, отчеты по лабораторным работам

3. Структура и содержание дисциплины

3.1. Трудоемкость дисциплины

	очная форма обучения	очно-заочная форма обучения
Общая трудоемкость	3 ЗЕТ	3 ЗЕТ
Часов по учебному плану	108	108
в том числе		
аудиторные занятия (контактная работа):	49	19
- занятия лекционного типа	16	8
- занятия семинарского типа		
- лабораторные занятия	32	10
- КСР	1	1
самостоятельная работа	59	89
Промежуточная аттестация – зачёт	–	–

3.2. Содержание дисциплины

Наименование и краткое содержание разделов и тем дисциплины	Всего (часы)			в том числе												Самостоятельная работа обучающегося, часы		
				Контактная работа (работа во взаимодействии с преподавателем), часы														
				из них														
	Занятия лекционного типа			Занятия семинарского типа			Занятия лабораторного типа			Всего								
Очная	Очно-заочная	Заочная	Очная	Очно-заочная	Заочная	Очная	Очно-заочная	Заочная	Очная	Очно-заочная	Заочная	Очная	Очно-заочная	Заочная	Очная	Очно-заочная	Заочная	
Тема 1. Введение в предмет «Технологии обработки больших данных»	11	10		2	1							2	1		9	9		
Тема 2. Хранилища данных (ХД) и OLAP-системы	16	19		2	1				4	2		6	3		10	16		
Тема 3. Задачи и практическое применение Data Mining	22	20		4	2				8	2		12	4		10	16		
Тема 4. Методы кластеризации	22	20		4	2				8	2		12	4		10	16		

Наименование и краткое содержание разделов и тем дисциплины	Всего (часы)			В том числе															Самостоятельная работа обучающегося, часы		
				Контактная работа (работа во взаимодействии с преподавателем), часы																	
				из них																	
	Занятия лекционного типа			Занятия семинарского типа			Занятия лабораторного типа			Всего											
Очная	Очно-заочная	Заочная	Очная	Очно-заочная	Заочная	Очная	Очно-заочная	Заочная	Очная	Очно-заочная	Заочная	Очная	Очно-заочная	Заочная	Очная	Очно-заочная	Заочная				
данных																					
Тема 5. Основные задачи в проблеме распознавания образов	20	19		2	1				8	2		10	3		10	16					
Тема 6. Работа со сложными системами отклонений	16	19		2	1				4	2		6	3		10	16					
КСР	1	1										1	1								
Промежуточная аттестация																					
Итого	108	108		16	8				32	10		49	19		59	89					

Содержание дисциплины по темам

Тема 1. Введение в предмет «Технологии обработки больших данных»

Основные понятие Data Mining и возникновение, перспективы, проблемы Data mining. Системы поддержки принятия решений (СППР). Задачи систем поддержки принятия решений. База данных – СППР. Неэффективность использования OLTP-систем для анализа данных.

Тема 2. Хранилища данных (ХД) и OLAP-системы

Понятие данных. Значение понятий объект и атрибут, выборка, зависимая и независимая переменная. Типы шкал. Концепция хранилища данных. Организация ХД. Очистка данных. Хранилища данных и анализ. Многомерная модель данных. Определение OLAP-систем. Концептуальное многомерное представление данных. Архитектура OLAP-систем: MOLAP, ROLAP, HOLAP.

Тема 3. Задачи и практическое применение Data Mining.

Интеллектуальный анализ данных. Модели и методы DM. Процесс обнаружения знаний. Инструменты Data Mining. Поиск ассоциативных правил. Постановка задачи. Сиквенциальный анализ. Разновидности поиска ассоциативных правил. Методы представления результатов. Алгоритмы поиска ассоциативных правил.

Тема 4. Методы кластеризации данных.

Постановка задачи кластеризации. Меры близости, основанные на расстояниях, используемых в алгоритмах кластеризации. Базовые алгоритмы кластеризации. Адаптивные методы кластеризации. Классификация и регрессия. Постановка задачи. Представление резуль-

татов. Методы построения правил классификации. Методы построения математических функций. Сущность прогнозирования. Роль и значение прогнозирования. Понятия временных рядов, его компоненты и параметры. Задача визуализация данных.

Тема 5. Основные задачи в проблеме распознавания образов.

Проблема распознавания образов. Классификация и характеристика основных задач распознавания образов. Формирование информативного признакового пространства. Основные понятия. Критерия информативности. Логико-эвристические методы формирования признакового пространства. Статистические методы формирования признакового пространства. Алгоритмы распознавания, основанные на вычислении оценок (АВО). Необходимые определения. Этапы задания АВО. Теоремы о вычислении оценок. Задачи, решаемые алгоритмами вычисления оценок.

Тема 6. Работа со сложными системами отклонений.

Мера важности объектов в сложных системах. Важность объекта в системе. Мера важности. Примеры введения меры важности объектов в сложных системах. Мера важности признаков в задаче распознавания образов. Вычислительные алгоритмы получения информационных весов и оценки их сложности. Некоторые применения мера важности признаков.

Лабораторные работы организуются, в том числе в форме практической подготовки, которая предусматривает участие обучающихся в выполнении отдельных элементов работ, связанных с будущей профессиональной деятельностью.

Практическая подготовка предусматривает: выполнение проекта по профилю профессиональной деятельности и направленности образовательной программы.

На проведение практических занятий (лабораторных работ) в форме практической подготовки отводится 6 часов.

Практическая подготовка направлена на формирование и развитие:

- практических навыков в соответствии с профилем ОП: формирование и анализ требований к информатизации и автоматизации прикладных процессов, формализация предметной области проекта;
- компетенций – ПК-9 способен моделировать прикладные (бизнес) процессы и объекты предметной области

Текущий контроль успеваемости реализуется в рамках занятий лабораторного типа.

4. Учебно-методическое обеспечение самостоятельной работы обучающихся

Основные виды самостоятельной работы студентов:

- работа с основной и дополнительной литературой;
- изучение категориального аппарата дисциплины;
- самостоятельное изучение тем дисциплины;
- подготовка к зачету.

Работа с основной и дополнительной литературой

Изучение рекомендованной литературы следует начинать с учебников и учебных пособий, затем переходить к научным монографиям и материалам периодических изданий. Работа с литературой предусматривает конспектирование наиболее актуальных и познавательных материалов. Это не только мобилизует внимание, но и способствует более глубокому осмыслению материала, его лучшему запоминанию, а также позволяет студентам проводить система-

тизацию и сравнительный анализ изучаемой информации. Таким образом, конспектирование – одна из основных форм самостоятельного труда, которая требует от студента активно работать с учебной литературой и не ограничиваться конспектом лекций.

Студент должен уметь самостоятельно подбирать необходимую литературу для учебной и научной работы, уметь обращаться с предметными каталогами и библиографическим справочником библиотеки.

Изучение категориального аппарата дисциплины

Изучение и осмысление экономических категорий требует проработки лекционного материала, выполнения практических заданий, изучение словарей, энциклопедий, справочников.

Индивидуальная самостоятельная работа студента направлена на овладение и грамотное применение экономической терминологии в области компьютерного моделирования.

Самостоятельное изучение тем дисциплины

Особое место отводится самостоятельной проработке студентами отдельных разделов и тем изучаемой дисциплины. Такой подход вырабатывает у студентов инициативу, стремление к увеличению объема знаний, умений и навыков, всестороннего овладения способами и приемами профессиональной деятельности.

Изучение вопросов определенной темы направлено на более глубокое усвоение основных категорий экономической теории, понимание экономических процессов, происходящих в обществе, совершенствование навыка анализа теоретического и эмпирического материала.

Подготовка к зачету

Промежуточная аттестация студентов по дисциплине проходит в виде зачета и предусматривает оценку. Условием успешного прохождения промежуточной аттестации является систематическая работа студента в течение семестра. В этом случае подготовка к экзамену является систематизацией всех полученных знаний по данной дисциплине.

Рекомендуется внимательно изучить перечень вопросов к зачету, а также использовать в процессе обучения программу, учебно-методический комплекс, другие методические материалы.

Желательно спланировать троекратный просмотр материала перед зачетом. Во-первых, внимательное чтение с осмыслением, подчеркиванием и составлением краткого плана ответа. Во-вторых, повторная проработка наиболее сложных вопросов. В-третьих, быстрый просмотр материала или планов ответов для его систематизации в памяти.

Контрольные вопросы и задания для проведения текущего контроля и промежуточной аттестации по итогам освоения дисциплины приведены в п. 5.2.

Для обеспечения самостоятельной работы обучающихся используется электронный курс Технологии обработки больших данных, созданный в системе электронного обучения ННГУ – <https://e-learning.unn.ru/>.

5. Фонд оценочных средств для промежуточной аттестации по дисциплине (модулю), включающий

5.1. Описание шкал оценивания результатов обучения по дисциплине

Уровень сформированности компетенций (индикатора достижения компетенций)	Шкала оценивания сформированности компетенций						
	плохо	неудовлетворительно	удовлетворительно	хорошо	очень хорошо	отлично	превосходно
	не зачтено		зачтено				
<u>Знания</u>	Отсутствие знаний теоретического материала. Невозможность оценить полноту знаний вследствие отказа обучающегося от ответа	Уровень знаний ниже минимальных требований. Имели место грубые ошибки.	Минимально допустимый уровень знаний. Допущено много негрубых ошибок.	Уровень знаний в объеме, соответствующем программе подготовки. Допущено несколько негрубых ошибок	Уровень знаний в объеме, соответствующем программе подготовки. Допущено несколько незначительных ошибок	Уровень знаний в объеме, соответствующем программе подготовки, без ошибок.	Уровень знаний в объеме, превышающем программу подготовки.
<u>Умения</u>	Отсутствие минимальных умений. Невозможность оценить наличие умений вследствие отказа обучающегося от ответа	При решении стандартных задач не продемонстрированы основные умения. Имели место грубые ошибки.	Продemonстрированы основные умения. Решены типовые задачи с негрубыми ошибками. Выполнены все задания, но не в полном объеме.	Продemonстрированы все основные умения. Решены все основные задачи с негрубыми ошибками. Выполнены все задания, в полном объеме, но некоторые с недочетами.	Продemonстрированы все основные умения. Решены все основные задачи. Выполнены все задания, в полном объеме, но некоторые с недочетами.	Продemonстрированы все основные умения, решены все основные задачи с отдельными несущественными недочетами, выполнены все задания в полном объеме.	Продemonстрированы все основные умения. Решены все основные задачи. Выполнены все задания, в полном объеме без недочетов
<u>Навыки</u>	Отсутствие владения материалом. Невозможность оценить наличие навыков вследствие отказа обучающегося от ответа	При решении стандартных задач не продемонстрированы базовые навыки. Имели место грубые ошибки.	Имеется минимальный набор навыков для решения стандартных задач с некоторыми недочетами	Продemonстрированы базовые навыки при решении стандартных задач с некоторыми недочетами	Продemonстрированы базовые навыки при решении стандартных задач без ошибок и недочетов.	Продemonстрированы навыки при решении нестандартных задач без ошибок и недочетов.	Продemonстрирован творческий подход к решению нестандартных задач

Шкала оценки при промежуточной аттестации

Оценка		Уровень подготовки
зачтено	Превосходно	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «превосходно»
	Отлично	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «отлично», при этом хотя бы одна компетенция сформирована на уровне «отлично»
	Очень хорошо	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «очень хорошо», при этом хотя бы одна компетенция сформирована на уровне «очень хорошо»
	Хорошо	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «хорошо», при этом хотя бы одна компетенция сформирована на уровне «хорошо»
	Удовлетворительно	Все компетенции (части компетенций), на формирование которых направлена дисциплина, сформированы на уровне не ниже «удовлетворительно», при этом хотя бы одна компетенция сформирована на уровне «удовлетворительно»
не зачтено	Неудовлетворительно	Хотя бы одна компетенция сформирована на уровне «неудовлетворительно», ни одна из компетенций не сформирована на уровне «плохо»
	Плохо	Хотя бы одна компетенция сформирована на уровне «плохо»

5.2. Типовые контрольные задания или иные материалы, необходимые для оценки результатов обучения.

5.2.1. Контрольные вопросы

Вопрос	Код компетенции
1. Задача анализа данных. Технологии обработки данных. Отличия от других видов анализа данных.	УК-1
2. Постановка и порядок решение задачи интеллектуального анализа данных на примере информационной модели торговой компании.	УК-1
3. Закономерности в данных: неочевидные, объективные, полезные. Области применения интеллектуального анализа данных.	УК-1
4. Подготовка данных для анализа. Обработка пропущенных и недостающих данных. Анализ исключений.	УК-1
5. Основные концепции баз данных. Модель данных. СУБД. Использование баз данных в интеллектуальном анализе данных.	ПК-9
6. Постановка и порядок решения задачи интеллектуального анализа данных на примере хранилища слабоструктурированных текстовых документов.	ПК-9
7. Введение в многомерные базы данных. Особенности базы данных для оперативной аналитической обработки данных. Порядок и особенности построения хранилищ данных.	ПК-9

Вопрос	Код компетенции
8. Виды и особенности шкал измерений данных.	УК-1
9. Введение в OLAP. Типовая модель данных для OLAP. Особенности приложений для оперативной аналитической обработки данных.	ПК-9
10. Постановка и порядок решения задачи интеллектуального анализа данных на примере документной базы данных.	УК-1
11. Особенности использования данных при интеллектуальном анализе данных. Данные, информация и знания.	УК-1
12. Основные задачи интеллектуального анализа данных.	УК-1
13. Стадии интеллектуального анализа данных. Начальные этапы. Вычисление и анализ показателей описательной статистики.	УК-1
14. Характеристики временных рядов, используемые в ходе интеллектуального анализа данных.	ПК-9
15. Построение и использование моделей для интеллектуального анализа данных.	ПК-9
16. Решение задач классификации в ходе интеллектуального анализа данных: цель и основные методы.	ПК-9
17. Преимущественные особенности методов анализа данных: статистических, машинного обучения, оперативного анализа, интеллектуального анализа.	ПК-9
18. Решение задач кластеризация в ходе интеллектуального анализа данных: цель и основные методы.	ПК-9
19. Сущность и методы решения задачи прогнозирования при интеллектуальном анализе данных.	ПК-9
20. Комплексный подход к внедрению Data Mining, OLAP и хранилищ данных в СППР.	ПК-9

5.2.2. Типовые тестовые задания для оценки сформированности компетенции УК-1

1. Для машинного обучения подходят данные

A) Любых форматов в цифровом виде

B) Числовые типа int

C) Бинарные

D) Предварительно подготовленные, очищенные от ошибок, пропусков и выбросов, а также нормализованные и представленные в виде числовых векторов

2. Укажите фактор, способствовавший появлению тренда больших данных

A) маркетинговые кампании крупных корпораций

B) снижение издержек на хранение данных

C) появление новых технологий обработки потоковых данных

D) выпуск баз данных с обработкой данных в памяти

3. Выберите верный ответ

A) большие данные – это обработка или хранение более 1 Тб информации

B) проблема больших данных – это такая проблема, когда при существующих технологиях хранения и обработки сущностная обработка данных затруднена или невозможна

C) большие данные – это огромная PR-акция крупных вендоров и не более того

D) большие данные – это явление, когда цифровые данные наиболее полно представляют изучаемый объект

4. Отметьте те из вариантов, в которых данные структурированы:

A) данные о продажах компании, представленные в виде помесечных отчётов в формате MS Word

B) таблица с ежедневными показаниями температуры помещения за год в файле формата csv

C) текст педагогической поэмы А.С. Макаренко, представленный в формате PDF

D) библиотека фильмов, представленных в формате mpeg4 на одном жестком диске

5. Какие из следующих средств разумно использовать для анализа данных, представленных единственным csv-файлом размера более 100Гб:

A) Hadoop

B) Data Warehouse

C) "Песочница"

D) Python

ПК-9

6. Недостатками подхода, при котором имеется хранилище данных (ХД) и витрины данных (ВД) для информации по разделам данной области, являются:

A) увеличение объема данных, хранимых в ВД

B) дополнительные затраты на разработку систем поддержки принятия решений с ХД и ВД

C) увеличение нагрузки на основное ХД

D) избыточность (данные хранятся как в ХД, так и в ВД)

7. Данные в хранилище данных делятся на следующие категории:

A) детальные данные

B) агрегированные данные

C) метаданные

D) временные данные

8. Агрегированные данные в хранилище данных подразделяются на следующие типы:

A) аддитивные

B) транзитивные

C) полуаддитивные

D) неаддитивные

9. Метаданные описывают:

A) объекты предметной области, информация о которых хранится в хранилище данных

B) категории пользователей, использующих данные

C) местоположение серверов, рабочих станций и оперативные источники данных

D) системных администраторов

Е) размещенные на серверах и рабочих станциях программных средств и распределение данных

10. Данные, поступающие из оперативных источников данных в хранилища данных, образуют следующие информационные потоки:

- А) входной
- Вб) виртуальных данных
- С) метаданных
- Д) обобщения

5.2.3. Типовые задания для оценки сформированности компетенции УК-1

Задание 1

1. Проверить нулевую гипотезу о том, что заданное значение $a_0=10$ является математическим ожиданием нормально распределённой случайной величины при 5%-м уровне значимости для двусторонней критической области, если в результате обработки выборки объёма $n=10$ получено выборочное среднее $\bar{x}=12$, а выборочное среднее квадратическое отклонение равно $s_1=1$.

2. При уровне значимости $\alpha=0,1$ проверить гипотезу о равенстве дисперсий двух нормально распределённых случайных величин X и Y на основе выборочных данных при альтернативной гипотезе $H_1: \sigma_x^2 \neq \sigma_y^2$.

x_i	20	22	23	24	26
n_i	3	4	2	2	4

y_i	18	19	20	22	23
n_i	6	3	4	2	5

ПК-9

Задание 2

В таблице приведены данные по двум экономическим показателям

1. Рассчитайте параметры уравнений линейной, степенной, показательной и гиперболической парных регрессий.

2. Оцените тесноту связи с помощью показателей корреляции и детерминации.

3. Оцените качество уравнений с помощью средней ошибки аппроксимации.

4. Оцените статистическую надёжность результатов регрессионного моделирования с помощью F-критерия Фишера. По значениям характеристик, рассчитанных в пп. 3, 4 и данном пункте, выберите лучшее уравнение регрессии и дайте обоснование выбору.

5. Рассчитайте ожидаемое значение результата, если прогнозное значение фактора увеличится на 10 % от его среднего уровня.

6. Оцените полученные результаты.

Номер района	Потребительские расходы в расчёте на душу населения, тыс. руб., у	Среднемесячный доход на душу населения, тыс. руб., х
1	10,07	18,47
2	12,00	18,67
3	10,33	18,17
4	13,83	22,40

Номер района	Потребительские расходы в расчёте на душу населения, тыс. руб., у	Среднемесячный доход на душу населения, тыс. руб., х
5	15,07	26,53
6	16,73	25,90
7	11,83	21,07
8	13,87	22,93
9	16,70	27,77
10	13,43	19,23
11	6,93	19,47
12	15,40	31,63
13	12,27	29,60
14	13,30	27,70
15	11,40	18,73
16	11,80	22,17
17	18,60	23,50

Задание 3

Администрация банка изучает динамику депозитов физических лиц за ряд лет (млрд. долл. В сопоставимых ценах). Исходные данные представлены в таблице.

Время, лет	1	2	3	4	5	6	7
Депозиты физических лиц, х	2	6	7	3	10	12	13

Известно, что $\Sigma x^2 = 511$

1. Постройте уравнение линейного тренда и дайте интерпретацию его параметров.
2. Определите коэффициент детерминации для линейного тренда.
3. Администрация банка предполагает, что среднегодовой абсолютный прирост депозитов физических лиц составляет не менее 2,5 млрд. долл. Подтверждается ли это полученными Вами результатами?

Задание 4

При уровне значимости $\alpha = 0,05$ методом дисперсионного анализа проверьте нулевую гипотезу о влиянии фактора на качество объекта на основании пяти измерений для трёх уровней фактора Ф:

Номер измерения	Φ_1	Φ_2	Φ_3
1	12	10	20
2	16	8	26
3	15	7	28
4	17	5	24
5	14	9	27

5.2.3 Вопросы для собеседования

Вопрос	Код компетенции
1. Технологии BigData: дать определение для BigData, назначение BigData, история появления и основные принципы BigData. Достоин-	УК-1

Вопрос	Код компетенции
ства и недостатки BigData.	
2. Технологии управления знаниями, визуализации знаний и интеллектуальные карты. Дать определение понятиям, назначение технологии, привести примеры программного обеспечения для визуализации знаний и построения интеллектуальных карт.	УК-1
3. Данные, информация, знания, модели. Наука о данных.	ПК-9
4. Критерии больших данных. Источники больших данных.	УК-1
5. Технологии обработки больших данных: NoSQL	ПК-9
6. Возможные этапы работы с большими данными.	УК-1
7. Примеры и истории успеха работы с большими данными: торговля, финансы, кадры.	УК-1
8. Обзор подходов к работе с данными: от языка простых запросов до методов анализа больших данных.	ПК-9
9. Интеллектуальный анализ данных: краткий обзор подходов.	УК-1
10. Деревья принятия решений.	ПК-9
11. Технологии обработки больших данных: MapReduce	ПК-9
12. Визуализация больших данных.	ПК-9
13. Специфика хранения и обработки больших данных.	ПК-9
14. Технологии обработки больших данных: Hadoop, R.	ПК-9
15. Методика извлечения знаний Knowledge Discovery in Databases (KDD). Этапы	ПК-9
16. Data Mining. Постановка основных задач.	ПК-9

6. Учебно-методическое и информационное обеспечение дисциплины

Основная литература

1. Макшанов, А. В. Технологии интеллектуального анализа данных : учебное пособие / А. В. Макшанов, А. Е. Журавлев. — 2-е изд., стер. — Санкт-Петербург : Лань, 2019. — 212 с. — ISBN 978-5-8114-4493-9. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/120063> (дата обращения: 31.10.2022). — Режим доступа: для авториз. пользователей.

2. Вагин, В. Н. Знания и убеждения в интеллектуальном анализе данных : монография / В. Н. Вагин. — Москва : ФИЗМАТЛИТ, 2019. — 536 с. — ISBN 978-5-9221-1841-5. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/143885> (дата обращения: 31.10.2022). — Режим доступа: для авториз. пользователей.

3. Храмов, А. Г. Методы и алгоритмы интеллектуального анализа данных : учебное пособие / А. Г. Храмов. — Самара : Самарский университет, 2019. — 176 с. — ISBN 978-5-7883-1414-3. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/148603> (дата обращения: 31.10.2022). — Режим доступа: для авториз. пользователей.

Дополнительная литература

1. Нестеров, С. А. Основы интеллектуального анализа данных. Лабораторный практикум : учебное пособие / С. А. Нестеров. — Санкт-Петербург : Лань, 2020. — 40 с. — ISBN 978-5-8114-4509-7. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/130181> (дата обращения: 31.10.2022). — Режим доступа: для авториз. пользователей.

4. Станкевич, Л. А. Интеллектуальные системы и технологии : учебник и практикум для среднего профессионального образования / Л. А. Станкевич. — Москва : Издательство Юрайт, 2020. — 397 с. — (Профессиональное образование). — ISBN 978-5-534-11659-5. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/457149> (дата обращения: 31.10.2022).

5. Железнов, М. М. Методы и технологии обработки больших данных : учебно-методическое пособие / М. М. Железнов. — Москва : МИСИ – МГСУ, 2020. — 46 с. — ISBN 978-5-7264-2193-3. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/145102> (дата обращения: 31.10.2022). — Режим доступа: для авториз. пользователей.

Программное обеспечение и Интернет-ресурсы

1. Электронная библиотека учебников [Электронный ресурс]. - Режим доступа: <http://studentam.net> — Загл. с экрана. [Дата обращения: 31.10.2022]

2. Российская государственная библиотека [Электронный ресурс]. - Режим доступа: <http://www.rsl.ru> — Загл. с экрана. [Дата обращения: 31.10.2022]

3. Научная электронная библиотека [Электронный ресурс]. - Режим доступа: <http://elibrary.ru/> — Загл. с экрана. [Дата обращения: 31.01.2022]

4. Программное обеспечение: менеджер пакетов Anaconda, Spyder (open-source IDE для Python).

7. Материально-техническое обеспечение дисциплины

Помещения представляют собой учебные аудитории для проведения учебных занятий, предусмотренных программой, оснащенные оборудованием и техническими средствами обучения: компьютерного класса, имеющего компьютеры, объединенные сетью с выходом в Интернет, презентационная техника (телевизор, компьютер/ноутбук).

Помещения для самостоятельной работы обучающихся оснащены компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечены доступом в электронную информационно-образовательную среду.

Специальные условия организации обучения по дисциплине для инвалидов и лиц с ограниченными возможностями здоровья

Организация обучения по дисциплине инвалидов и лиц с ограниченными возможностями здоровья осуществляется с учетом особенностей психофизического развития, индивидуальных возможностей и состояния здоровья при наличии таких обучающихся путем создания специальных условий для получения образования.

Профессорско-преподавательский состав знакомится с психолого-физиологическими особенностями обучающихся инвалидов и лиц с ограниченными возможностями здоровья, индивидуальными программами реабилитации инвалидов (при наличии).

В соответствии с Методическими рекомендациями по организации образовательного процесса для обучения инвалидов и лиц с ограниченными возможностями здоровья в образовательных организациях высшего образования, в том числе оснащенности образовательного процесса, утв. Минобрнауки РФ 08.04.2014 АК-44/05вн при изучении дисциплины предполагается использование социально-активных и рефлексивных методов обучения, технологий социокультурной реабилитации с целью оказания помощи в установлении полноценных межличностных отношений с другими студентами, создании комфортного психологического климата в студенческой группе.

При освоении дисциплины используются различные сочетания видов учебной работы с методами и формами активизации познавательной деятельности обучающихся для достижения запланированных результатов обучения и формирования компетенций. Форма проведения промежуточной аттестации для обучающихся-инвалидов и лиц с ограниченными возможностями здоровья устанавливается с учетом индивидуальных психофизиологических

особенностей. По личной просьбе обучающегося с ограниченными возможностями здоровья, изложенной в форме письменного заявления, по дисциплине предусматриваются:

- замена устного ответа на письменный ответ при сдаче зачета или экзамена;
- увеличение продолжительности времени на подготовку к ответу на зачете или экзамене;
- при подведении результатов промежуточной аттестации студентов выставляется максимальное количество баллов за посещаемость аудиторных занятий.

Программа составлена в соответствии с требованиями ФГОС ВО/ОС ННГУ по направлению 09.03.03 Прикладная информатика (приказ №349-ОД от 21.06.2021).

Автор(ы): к.т.н. Кечкина Н.И.

Рецензент:

Программа одобрена Методической комиссией Дзержинского филиала ННГУ от 10.11.2022 года, протокол № 12.