

**Министерство науки и высшего образования РФ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Национальный исследовательский университет «МЭИ»**

Направление подготовки/специальность: 27.04.04 Управление в технических системах

Наименование образовательной программы: Управление и информатика в технических системах

Уровень образования: высшее образование - магистратура

Форма обучения: Очная

**Оценочные материалы
по дисциплине
Интеллектуальные информационные системы**

**Москва
2022**

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ РАЗРАБОТАЛ:

Преподаватель

(должность)

	Подписано электронной подписью ФГБОУ ВО «НИУ «МЭИ»	
	Сведения о владельце ЦЭП МЭИ	
	Владелец	Мохов А.С.
	Идентификатор	R55ae9104-MokhovAS-2434a28b

(подпись)

А.С. Мохов

(расшифровка
подписи)

СОГЛАСОВАНО:

Руководитель
образовательной
программы

(должность, ученая степень, ученое
звание)

	Подписано электронной подписью ФГБОУ ВО «НИУ «МЭИ»	
	Сведения о владельце ЦЭП МЭИ	
	Владелец	Бобряков А.В.
	Идентификатор	R2c90f415-BobriakovAV-70dec1fa

(подпись)

А.В.

Бобряков

(расшифровка
подписи)

Заведующий
выпускающей кафедры

(должность, ученая степень, ученое
звание)

	Подписано электронной подписью ФГБОУ ВО «НИУ «МЭИ»	
	Сведения о владельце ЦЭП МЭИ	
	Владелец	Бобряков А.В.
	Идентификатор	R2c90f415-BobriakovAV-70dec1fa

(подпись)

А.В.

Бобряков

(расшифровка
подписи)

ОБЩАЯ ЧАСТЬ

Оценочные материалы по дисциплине предназначены для оценки: достижения обучающимися запланированных результатов обучения по дисциплине, этапа формирования запланированных компетенций и уровня освоения дисциплины.

Оценочные материалы по дисциплине включают оценочные средства для проведения мероприятий текущего контроля успеваемости и промежуточной аттестации.

Формируемые у обучающегося компетенции:

1. ПК-2 Способен разрабатывать и применять информационные системы обработки и анализа данных для автоматизации процессов управления в сложных технических и организационно-технических системах

ИД-1 Демонстрирует умение организовывать экспериментальные исследования и сбор экспертной информации, проводить анализ и предварительную обработку данных с применением автоматизированных информационных систем, выбирать обоснованные способы обеспечения защиты данных

ИД-2 Может разрабатывать информационные и информационно-аналитические системы автоматизации процессов управления в сложных технических и организационно-технических системах

и включает:

для текущего контроля успеваемости:

Форма реализации: Компьютерное задание

1. Защита лабораторной работы №1 (Лабораторная работа)
2. Защита лабораторной работы №2 (Лабораторная работа)
3. Защита лабораторной работы №3 (Лабораторная работа)
4. Контрольная работа №1 (Тестирование)
5. Контрольная работа №2 (Тестирование)

БРС дисциплины

3 семестр

Раздел дисциплины	Веса контрольных мероприятий, %					
	Индекс КМ:	КМ-1	КМ-2	КМ-3	КМ-4	КМ-5
	Срок КМ:	3	6	12	13	15
Подготовка данных в процессе Data Mining в условиях больших данных						
Введение в проблему больших данных		+	+			
Обработка данных в условиях Big Data		+	+			
Модель MapReduce обработки больших объемов данных						
Технологии работы с большими данными				+	+	
Файловая система HDFS				+	+	

Технология MapReduce			+	+	
Современные технологии обработки больших объемов данных					
Инструменты хранения и обработки больших коллекций				+	+
Технологии обработки Big Data				+	+
Вес КМ:	10	20	30	10	30

\$Общая часть/Для промежуточной аттестации\$

СОДЕРЖАНИЕ ОЦЕНОЧНЫХ СРЕДСТВ ТЕКУЩЕГО КОНТРОЛЯ

I. Оценочные средства для оценки запланированных результатов обучения по дисциплине, соотнесенных с индикаторами достижения компетенций

Индекс компетенции	Индикатор	Запланированные результаты обучения по дисциплине	Контрольная точка
ПК-2	ИД-1 _{ПК-2} Демонстрирует умение организовывать экспериментальные исследования и сбор экспертной информации, проводить анализ и предварительную обработку данных с применением автоматизированных информационных систем, выбирать обоснованные способы обеспечения защиты данных	Знать: методы снижения размерности выборок большого объема для проведения их анализа Уметь: применять современные технологии для решения задачи обработки и анализа больших данных с целью увеличения качества	Контрольная работа №1 (Тестирование) Защита лабораторной работы №1 (Лабораторная работа) Защита лабораторной работы №3 (Лабораторная работа)
ПК-2	ИД-2 _{ПК-2} Может разрабатывать информационные и информационно-аналитические системы автоматизации процессов управления в сложных технических и организационно-технических системах	Знать: современные подходы и инструменты работы с большим объемом данных Уметь: применять модель MapReduce для проведения вычислений над большими объемами данных	Защита лабораторной работы №2 (Лабораторная работа) Контрольная работа №2 (Тестирование)

II. Содержание оценочных средств. Шкала и критерии оценивания

КМ-1. Контрольная работа №1

Формы реализации: Компьютерное задание

Тип контрольного мероприятия: Тестирование

Вес контрольного мероприятия в БРС: 10

Процедура проведения контрольного мероприятия: Выполнение тестовых заданий в СДО «Прометей».

Краткое содержание задания:

Дать правильные ответы на вопросы тестирования, связанные с методами снижения размерности и сэмплинга в задачах обработки больших данных. Знать способы выявления информативных признаков и алгоритмы методов сэмплинга.

Контрольные вопросы/задания:

Знать: методы снижения размерности выборок большого объема для проведения их анализа	1.Что такое «Информативный признак»? 2.Какие этапы процедуры снижения размерности данных вы знаете?
--	--

Описание шкалы оценивания:

Оценка: 5

Нижний порог выполнения задания в процентах: 85

Описание характеристики выполнения знания: Оценка "отлично" выставляется если задание выполнено в полном объеме или выполнено преимущественно верно.

Оценка: 4

Нижний порог выполнения задания в процентах: 65

Описание характеристики выполнения знания: Оценка "хорошо" выставляется если задание выполнено с небольшими ошибками.

Оценка: 3

Нижний порог выполнения задания в процентах: 50

Описание характеристики выполнения знания: Оценка "удовлетворительно" выставляется если задание преимущественно выполнено, часть ответов дана с ошибками.

КМ-2. Защита лабораторной работы №1

Формы реализации: Компьютерное задание

Тип контрольного мероприятия: Лабораторная работа

Вес контрольного мероприятия в БРС: 20

Процедура проведения контрольного мероприятия: Выполнение общего задания в программной среде Jupiter Notebook. Демонстрация выполнения работы программы с комментариями по реализации. Внесение изменений в программу в соответствии с индивидуальным дополнительным заданием. Демонстрация работы программы с внесенными изменениями.

Краткое содержание задания:

Защита лабораторной работы №1: «Влияние методов сэмплинга и снижения размерности на качество решения задач машинного обучения». Загрузка требуемых модулей и

библиотек. Загрузка выборки. Проведение сэмпинга и снижения размерности данных. Оценка качества классификации на обработанной выборке.

Контрольные вопросы/задания:

Знать: методы снижения размерности выборок большого объема для проведения их анализа	1. Назовите основные группы методов сэмпинга. 2. В чем недостатки работы с данными большой размерности?
--	--

Описание шкалы оценивания:

Оценка: 5

Нижний порог выполнения задания в процентах: 85

Описание характеристики выполнения знания: Оценка "отлично" выставляется если задание выполнено в полном объеме или выполнено преимущественно верно, даны правильные ответы на дополнительные вопросы.

Оценка: 4

Нижний порог выполнения задания в процентах: 65

Описание характеристики выполнения знания: Оценка "хорошо" выставляется если задание выполнено с небольшими ошибками, ответы на дополнительные вопросы преимущественно правильные.

Оценка: 3

Нижний порог выполнения задания в процентах: 50

Описание характеристики выполнения знания: Оценка "удовлетворительно" выставляется если задание преимущественно выполнено, ответы на дополнительные вопросы неточные, неполные.

КМ-3. Защита лабораторной работы №2

Формы реализации: Компьютерное задание

Тип контрольного мероприятия: Лабораторная работа

Вес контрольного мероприятия в БРС: 30

Процедура проведения контрольного мероприятия: Выполнение общего задания в программной среде Jupiter Notebook. Демонстрация выполнения работы программы с комментариями по реализации. Внесение изменений в программу в соответствии с индивидуальным дополнительным заданием. Демонстрация работы программы с внесенными изменениями.

Краткое содержание задания:

Защита лабораторной работы №2: «Применение модели MapReduce для решения задач анализа текстовых данных». Загрузка требуемых модулей и библиотек. Загрузка выборки. Реализация алгоритмов обработки данных и классификации с использованием технологии MapReduce. Анализ полученных результатов.

Контрольные вопросы/задания:

Уметь: применять модель MapReduce для проведения вычислений над большими объемами данных	1. Как реализовать разделение данных между узлами на шаге Map? 2. Как реализовать объединение данных на шаге Reduce?
--	---

Описание шкалы оценивания:

Оценка: 5

Нижний порог выполнения задания в процентах: 85

Описание характеристики выполнения знания: Оценка "отлично" выставляется если задание выполнено в полном объеме или выполнено преимущественно верно, даны правильные ответы на дополнительные вопросы.

Оценка: 4

Нижний порог выполнения задания в процентах: 65

Описание характеристики выполнения знания: Оценка "хорошо" выставляется если задание выполнено с небольшими ошибками, ответы на дополнительные вопросы преимущественно правильные.

Оценка: 3

Нижний порог выполнения задания в процентах: 50

Описание характеристики выполнения знания: Оценка "удовлетворительно" выставляется если задание преимущественно выполнено, ответы на дополнительные вопросы неточные, неполные.

КМ-4. Контрольная работа №2

Формы реализации: Компьютерное задание

Тип контрольного мероприятия: Тестирование

Вес контрольного мероприятия в БРС: 10

Процедура проведения контрольного мероприятия: Выполнение тестовых заданий в СДО «Прометей».

Краткое содержание задания:

Дать правильные ответы на вопросы тестирования, связанные с знанием умением работы с инструментами обработки больших данных – экосистемы Hadoop, технологии MapReduce, файловой системы HDFS.

Контрольные вопросы/задания:

Знать: современные подходы и инструменты работы с большим объемом данных	1. В чем особенность файловой системы HDFS? 2. Какие NoSql СУБД вы знаете?
--	---

Описание шкалы оценивания:

Оценка: 5

Нижний порог выполнения задания в процентах: 85

Описание характеристики выполнения знания: Оценка "отлично" выставляется если задание выполнено в полном объеме или выполнено преимущественно верно.

Оценка: 4

Нижний порог выполнения задания в процентах: 65

Описание характеристики выполнения знания: Оценка "хорошо" выставляется если задание выполнено с небольшими ошибками.

Оценка: 3

Нижний порог выполнения задания в процентах: 50

Описание характеристики выполнения знания: Оценка "удовлетворительно" выставляется если задание преимущественно выполнено, часть ответов дана с ошибками.

КМ-5. Защита лабораторной работы №3

Формы реализации: Компьютерное задание

Тип контрольного мероприятия: Лабораторная работа

Вес контрольного мероприятия в БРС: 30

Процедура проведения контрольного мероприятия: Выполнение общего задания в системе Hadoop. Демонстрация выполнения работы программы с комментариями по реализации. Внесение изменений в программу в соответствии с индивидуальным дополнительным заданием. Демонстрация работы программы с внесенными изменениями.

Краткое содержание задания:

Защита лабораторной работы №3: «Применение современных библиотек для анализа больших объемов данных». Загрузка требуемых модулей и библиотек. Загрузка выборки. Настройка системы Hadoop. Реализация в Hadoop процесса обработки данных.

Контрольные вопросы/задания:

Уметь: применять современные технологии для решения задачи обработки и анализа больших данных с целью увеличения качества	1.Как запустить программный модуль на исполнение в системе Hadoop? 2.Как записать данные в файловую систему HDFS Hadoop?
---	---

Описание шкалы оценивания:

Оценка: 5

Нижний порог выполнения задания в процентах: 85

Описание характеристики выполнения знания: Оценка "отлично" выставляется если задание выполнено в полном объеме или выполнено преимущественно верно, даны правильные ответы на дополнительные вопросы.

Оценка: 4

Нижний порог выполнения задания в процентах: 65

Описание характеристики выполнения знания: Оценка "хорошо" выставляется если задание выполнено с небольшими ошибками, ответы на дополнительные вопросы преимущественно правильные.

Оценка: 3

Нижний порог выполнения задания в процентах: 50

Описание характеристики выполнения знания: Оценка "удовлетворительно" выставляется если задание преимущественно выполнено, ответы на дополнительные вопросы неточные, неполные.

СОДЕРЖАНИЕ ОЦЕНОЧНЫХ СРЕДСТВ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

3 семестр

Форма промежуточной аттестации: Зачет с оценкой

Пример билета

1. «Большие данные» - определение, особенности обработки.
2. Методы снижения размерности данных.

Процедура проведения

Зачет с оценкой проводится в устной форме по билетам. На подготовку ответа студенту отводится 45 минут.

1. Перечень компетенций/индикаторов и контрольных вопросов проверки результатов освоения дисциплины

1. Компетенция/Индикатор: ИД-1_{ПК-2} Демонстрирует умение организовывать экспериментальные исследования и сбор экспертной информации, проводить анализ и предварительную обработку данных с применением автоматизированных информационных систем, выбирать обоснованные способы обеспечения защиты данных

Вопросы, задания

1. Методы уменьшения объема выборок.
2. Методы снижения размерности признакового пространства выборок.
3. Активное обучение (Active learning) – определение, цели использования.

Материалы для проверки остаточных знаний

1. Что НЕ относится к признакам Big Data согласно концепции «V V V»?

Ответы:

Velocity (скорость)

Variety (многообразие)

Volume (объем)

Visible (обозримый)

Верный ответ: Velocity (скорость) Variety (многообразие) Volume (объем) *Visible (обозримый)

2. Что не относится к методам undersampling?

Ответы:

Tomek Links

RandomUnderSampler

SMOTE

Condensed Nearest Neighbors

Верный ответ: Tomek Links RandomUnderSampler * SMOTE Condensed Nearest Neighbors

3. Для чего используется метод Tomek Links?

Ответы:

Снижение размерности

Уменьшение количества наблюдений

Подсчет меры информативности

Верный ответ: Снижение размерности * Уменьшение количества наблюдений
Подсчет меры информативности

4.Активное обучение (active learning) подразумевает следующее:

Ответы:

Система запрашивает разметку конкретных объектов

Эксперт размечает все данные

Объекты для разметки выбираются случайным образом

Объекты для разметки выбираются на основе правил

Верный ответ: *Система запрашивает разметку конкретных объектов Эксперт размечает все данные Объекты для разметки выбираются случайным образом

*Объекты для разметки выбираются на основе правил

2. Компетенция/Индикатор: ИД-2_{ПК-2} Может разрабатывать информационные и информационно-аналитические системы автоматизации процессов управления в сложных технических и организационно-технических системах

Вопросы, задания

1.Компоненты экосистемы Hadoop – назначение, особенности.

2.Технология MapReduce. Определение, назначение, особенности.

3.Файловая система HDFS. Определение, назначение, особенности.

Материалы для проверки остаточных знаний

1.Какого шага нет в MapReduce?

Ответы:

Map

Reduce

Shuffle

Rewind

Верный ответ: Map Reduce Shuffle *Rewind

2.Что происходит на шаге Map в MapReduce?

Ответы:

Разбиение данных по узлам

Объединение данных с разных узлов

Отправка данных на сервер

Загрузка данных с сервера

Верный ответ: * Разбиение данных по узлам Объединение данных с разных узлов

Отправка данных на сервер Загрузка данных с сервера

3.Что происходит на шаге Reduce в MapReduce?

Ответы:

Разбиение данных по узлам

Объединение данных с разных узлов

Отправка данных на сервер

Загрузка данных с сервера

Верный ответ: Разбиение данных по узлам * Объединение данных с разных узлов

Отправка данных на сервер Загрузка данных с сервера

4.Что не относится к методам выявления информативных признаков?

Ответы:

Метод главных компонент

Tomek Links

Mutual Information

Коэффициент Жаккара

Верный ответ: Метод главных компонент *Tomek Links Mutual Information

Коэффициент Жаккара

5.Что не входит в архитектуру HDFS?

Ответы:

Data Node
Secondary Node
Name Node
Training Node

Верный ответ: Data Node Secondary Node Name Node * Training Node

6. Что из перечисленного является файловой системой?

Ответы:

HDFS
NoSQL
MapReduce
Hadoop

Верный ответ: *HDFS NoSQL MapReduce Hadoop

7. Сколько Name Node должно существовать в HDFS?

Ответы:

0
1
2

Сколько угодно

Верный ответ: 0 * 1 2 Сколько угодно

8. Сколько Data Node может существовать в HDFS?

Ответы:

0
1

По числу узлов

Верный ответ: 0 1 * По числу узлов

9. Что из перечисленного является NoSQL СУБД?

Ответы:

MSSQL
Cassandra
My SQL

Верный ответ: MSSQL *Cassandra My SQL

10. Что из перечисленного не является NoSQL СУБД?

Ответы:

MSSQL
Cassandra
HBase

Верный ответ: * MSSQL Cassandra HBase

11. Что не входит в принцип «BASE» NoSQL СУБД?

Ответы:

базовая доступность (basic availability)
гибкое состояние (soft state)
согласованность в конечном счёте (eventual consistency)
атомарность (atomicity)

Верный ответ: базовая доступность (basic availability) гибкое состояние (soft state)
согласованность в конечном счёте (eventual consistency) * атомарность (atomicity)

II. Описание шкалы оценивания

Оценка: 5

Нижний порог выполнения задания в процентах: 85

Описание характеристики выполнения знания: Работа выполнена в рамках "продвинутого" уровня. Ответы даны верно, четко сформулированные особенности практических решений.

Оценка: 4

Нижний порог выполнения задания в процентах: 65

Описание характеристики выполнения знания: Работа выполнена в рамках "базового" уровня. Большинство ответов даны верно. В части материала есть незначительные недостатки.

Оценка: 3

Нижний порог выполнения задания в процентах: 50

Описание характеристики выполнения знания: Работа выполнена в рамках "порогового" уровня. Основная часть задания выполнена верно. на вопросы углубленного уровня.

III. Правила выставления итоговой оценки по курсу

Оценка определяется в соответствии с Положением о балльно-рейтинговой системе для студентов НИУ «МЭИ» на основании семестровой и аттестационной составляющих.