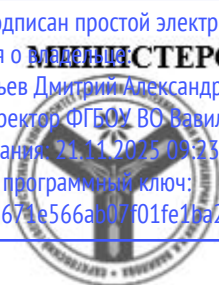


Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Соловьев Дмитрий Александрович
Должность: ректор ФГБОУ ВО Вавиловский университет
Дата подписания: 21.11.2025 09:23:48
Уникальный программный ключ:
528682d78e674e566a007f01fe1ba2172f735a12



МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение высшего образования

«Саратовский государственный университет генетики, биотехнологии и инженерии имени Н.И. Вавилова»

**ОЦЕНОЧНЫЕ МАТЕРИАЛЫ
для проверки сформированности компетенций**

Дисциплина	Big Data
Направление подготовки	09.04.03 Прикладная информатика
Направленность (профиль)	Проектирование информационных систем
Квалификация выпускника	Бакалавр
Нормативный срок обучения	4 года
Форма обучения	Очная

Разработчик: *доцент Шибайкин В.А.*


(подпись)

Саратов 2024

ОГЛАВЛЕНИЕ

1. Перечень компетенций с указанием этапов их формирования в процессе освоения ОПОП	3
2. Сценарии выполнения заданий	3
3. Система оценивания выполнения заданий	4
4. Описание дополнительных материалов и оборудования, необходимых для выполнения заданий	5
5. Задания для проверки уровня сформированности компетенций с указанием типа заданий (с ключами к оцениванию заданий).....	6

1. Перечень компетенций с указанием этапов их формирования в процессе освоения ОПОП

В результате изучения дисциплины «Big Data» обучающиеся, в соответствии с ФГОС ВО по направлению подготовки 09.04.03 Прикладная информатика, утвержденного приказом Министерства образования и науки РФ от 19.09.2017 № 916, формируют следующие компетенции, указанные в таблице:

Код компетенции	Наименование компетенции	Этапы формирования компетенции в процессе освоения ОПОП (семестр)
ОПК-3	Способен анализировать профессиональную информацию, выделять в ней главное, структурировать, оформлять и представлять в виде аналитических обзоров с обоснованными выводами и рекомендациями	1
ОПК-8	Способен осуществлять эффективное управление разработкой программных средств и проектов.	1

2. Сценарии выполнения заданий

№ п/п	Тип задания	Последовательность действий при выполнении задания
1. Задания закрытого типа		
1.1	Задание закрытого типа на установление соответствия	1. Внимательно прочитать текст задания и понять, что в качестве ответа ожидаются пары элементов. 2. Внимательно прочитать оба списка: список 1 – вопросы, утверждения, факты, понятия и т.д.; список 2 – утверждения, свойства объектов и т.д. 3. Сопоставить элементы списка 1 с элементами списка 2, сформировать пары элементов. 4. Записать попарно буквы и цифры (в зависимости от задания) вариантов ответа (например, А1 или Б4).
1.2	Задание закрытого типа на установление последовательности	1. Внимательно прочитать текст задания и понять, что в качестве ответа ожидается последовательность элементов. 2. Внимательно прочитать предложенные варианты ответа. 3. Построить верную последовательность из предложенных элементов. 4. Записать буквы/цифры (в зависимости от задания) вариантов ответа в нужной последовательности без пробелов и знаков препинания (например, БВА или 135).
2. Задания открытого типа		
2.1	Задание открытого типа с кратким ответом	1. Внимательно прочитать текст задания и понять суть вопроса. 2. Продумать краткий ответ.

№ п/п	Тип задания	Последовательность действий при выполнении задания
		3. Записать ответ в виде слова, словосочетания или числа. 4. В случае расчетной задачи, записать ответ в виде числа.
2.2	Задание открытого типа с развернутым ответом	1. Внимательно прочитать текст задания и понять суть вопроса. 2. Продумать логику и полноту ответа. 3. Записать ответ, используя четкие компактные формулировки. 4. В случае расчетной задачи, записать решение и ответ.
3. Задания комбинированного типа		
3.1	Задание комбинированного типа с выбором одного верного ответа из предложенных и обоснованием выбора	1. Внимательно прочитать текст задания и понять, что в качестве ответа ожидается только один из предложенных вариантов. 2. Внимательно прочитать предложенные варианты ответа. 3. Выбрать один ответ, наиболее верный. 4. Записать только номер (или букву) выбранного варианта ответа. 5. Записать аргументы, обосновывающие выбор ответа.
3.2	Задание комбинированного типа с выбором нескольких верных ответов из предложенных и обоснованием выбора	1. Внимательно прочитать текст задания и понять, что в качестве ответа ожидается несколько из предложенных вариантов. 2. Внимательно прочитать предложенные варианты ответа. 3. Выбрать несколько ответов, наиболее верных. 4. Записать только номера (или буквы) выбранных вариантов ответа. 5. Записать аргументы, обосновывающие выбор ответов.

3. Система оценивания выполнения заданий

№ п/п	Указания по оцениванию	Характеристика правильности ответа
1. Задания закрытого типа		
1.1	Задание закрытого типа на установление соответствия считается верным, если правильно установлены все соответствия (позиции из одного столбца верно сопоставлены с позициями другого).	«верно» / «неверно»
1.2	Задание закрытого типа на установление последовательности считается верным, если правильно указана вся последовательность цифр.	«верно» / «неверно»
2. Задания открытого типа		
2.1	Задание открытого типа с кратким ответом оценивается по следующим критериям: 1) Правильность ответа (отсутствие фактических и грамматических ошибок). 2). Сопоставимость с эталонным ответом в случае расчетной задачи.	«верно» / «неверно»

№ п/п	Указания по оцениванию	Характеристика правильности ответа
2.2	Задание открытого типа с развернутым ответом оценивается по следующим критериям. 1) Правильность ответа (отсутствие фактических ошибок). 2) Полнота ответа (раскрытие объема используемых понятий). 3) Обоснованность ответа (наличие аргументов). 4) Логика изложения ответа (грамотная последовательность излагаемого материала). 5. Сопоставимость с эталонным ответом.	«верно» / «неверно»
3. Задания комбинированного типа		
3.1	Задание комбинированного типа с выбором одного верного ответа из предложенных с обоснованием выбора ответа считается верным, если правильно указана цифра (буква) и приведены корректные аргументы, используемые при выборе ответа	«верно» / «неверно»
3.2	Задание комбинированного типа с выбором нескольких вариантов ответа из предложенных с обоснованием выбора ответов считается верным, если правильно указаны цифры (буквы) и приведены корректные аргументы, используемые при выборе ответа.	«верно» / «неверно»

4. Описание дополнительных материалов и оборудования, необходимых для выполнения заданий

Для выполнения заданий дополнительные материалы и оборудование не требуются.

**5. Задания для проверки уровня сформированности компетенций с указанием типа заданий
(с ключами к оцениванию заданий)**

Номер задания	Формулировка задания	Тип задания	Ключ к оцениванию задания
1 семестр			
ОПК-3 Способен анализировать профессиональную информацию, выделять в ней главное, структурировать, оформлять и представлять в виде аналитических обзоров с обоснованными выводами и рекомендациями;			
1.	<i>Прочитайте текст, выберите один правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа:</i> Выберите метод являющийся ключевым для анализа больших данных с целью выделения главной информации, обоснуйте выбор: 1) Использование только Excel для всех видов анализа. 2) Применение методов машинного обучения и облачных вычислений для обработки и анализа данных. 3) Ручной ввод данных без использования автоматизированных систем. 4) Использование исключительно бумажных носителей для хранения данных. 5) Проведение анализа данных без предварительной обработки и очистки.	Задание комбинированного типа с выбором одного верного ответа из предложенных и обоснованием выбора	2 Обоснование: Облачные платформы (вычисления) легко масштабируемы (позволяют гибко увеличивать или уменьшать ресурсы в зависимости от задач), а машинное обучение помогает находить скрытые закономерности в данных. Это делает анализ больших данных более эффективным (помогает принимать более правильные решения на основе данных)

Номер задания	Формулировка задания	Тип задания	Ключ к оцениванию задания
2.	<i>Прочитайте текст, выберите один правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа:</i> Выберите наиболее эффективный подход для структурирования и представления аналитических обзоров на основе больших данных: 1) Представление данных в виде неструктурированного текста без графиков и таблиц. 2) Использование визуализации данных и интерактивных дашбордов для наглядного представления информации. 3) Ограничение анализа только текстовыми отчетами без использования цифровых инструментов. 4) Представление данных без выводов и рекомендаций. 5) Использование исключительно устных презентаций без письменных отчетов.	Задание комбинированного типа с выбором одного верного ответа из предложенных и обоснованием выбора	2 Обоснование: Визуализация помогает сделать сложные данные понятными и удобными для анализа. Графики и диаграммы позволяют быстро увидеть закономерности (тренды) и необычные отклонения.
3.	<i>Прочитайте текст и запишите развернутый обоснованный ответ:</i> Опишите основные характеристики Big Data.	Задание открытого типа с развернутым ответом	Основные характеристики Big Data включают объем (volume), скорость (velocity) и многообразие (variety).
4.	<i>Прочитайте текст и запишите развернутый обоснованный ответ:</i> Объясните, почему традиционные методы обработки данных не подходят для работы с Big Data.	Задание открытого типа с развернутым ответом	Обычные методы обработки данных не подходят для работы с большими данными (Big Data), потому что они не справляются с огромными объемами, высокой скоростью поступления и разными форматами данных.
5.	<i>Прочитайте текст и запишите развернутый обоснованный ответ:</i> Сформулируйте основные методы и наименования инструментов, используемых для анализа Big Data.	Задание открытого типа с развернутым ответом	Традиционные базы данных плохо работают с неструктурированной информацией и не обеспечивают нужную скорость анализа. Для работы с Big Data используют машинное обучение, облачные технологии

Номер задания	Формулировка задания	Тип задания	Ключ к оцениванию задания
			и специальные облачные инструменты (Hadoop, Spark и NoSQL базы данных).
6.	<i>Прочитайте текст и запишите развернутый обоснованный ответ:</i> Сформулируйте методы и технологии, используемые для процесса анализа больших данных (Big Data). Приведите примеры инструментов, которые применяются в этих процессах.	Задание открытого типа с развернутым ответом	MapReduce лежит в основе многих современных систем обработки больших данных, таких как Hadoop, Spark и др. (Hadoop, Microsoft Azure).
7.	<i>Прочитайте текст и запишите развернутый обоснованный ответ:</i> Опишите основные вызовы, с которыми сталкиваются специалисты при классификации больших данных. Какие методы и инструменты могут быть использованы для преодоления этих вызовов?	Задание открытого типа с развернутым ответом	При работе с Big Data часто возникают трудности: - огромное количество признаков; - шум и пропуски; - сложно выбрать алгоритм.
8.	<i>Прочитайте текст и запишите развернутый обоснованный ответ:</i> Объясните, как технологии Интернета вещей (IoT) интегрируются с системами Big Data. Приведите примеры применения такой интеграции в реальных бизнес-кейсах.	Задание открытого типа с развернутым ответом	Технологии Интернета вещей (IoT) объединяются с системами Big Data: данные с разных датчиков и устройств собираются и отправляются в одну систему для обработки. Это используется, например, в умных городах для управления трафиком и энергопотреблением, а также на производствах для предсказания поломок оборудования и отслеживания процессов.
9.	<i>Прочитайте текст и установите соответствие:</i> Соотнесите программное обеспечение для анализа данных с их основными характеристиками или функциями: А) Apache Hadoop Б) 1С: Аналитика В) Yandex DataLens Г) Apache Spark	Задание закрытого типа на установление соответствия	А – 2 Б – 1 В – 3 Г – 4

Номер задания	Формулировка задания	Тип задания	Ключ к оцениванию задания
	1) Инструмент визуализации данных, который позволяет создавать интерактивные и наглядные дашборды для анализа данных. 2) Платформа для распределенной обработки больших данных, которая обеспечивает хранение и обработку данных на кластерах компьютеров. 3) Инструмент бизнес-аналитики, который предоставляет возможности для визуализации данных и интеграции с различными источниками данных. 4) Фреймворк для обработки данных в реальном времени, который поддерживает распределенные вычисления и обработку больших данных.		
10.	<i>Прочитайте текст и установите последовательность. Запишите соответствующую последовательность цифр слева направо:</i> Последовательность этапов анализа данных с использованием нейронных сетей в контексте работы с большими данными (Big Data) можно представить следующим образом: 1) Обработка и очистка данных: удаление пропусков, аномалий, нормализация данных. 2) Сбор данных: определение источников данных и их извлечение. 3) Обучение нейронной сети: настройка гиперпараметров, выбор архитектуры сети, запуск процесса обучения. 4) Анализ результатов: интерпретация полученных данных, построение выводов и рекомендаций. 5) Предобработка данных: преобразование данных в формат, пригодный для подачи в нейронную сеть. 6) Валидация модели: проверка качества работы модели на тестовых данных.	Задание закрытого типа на установление последовательности	2, 1, 5, 3, 6, 4,

Номер задания	Формулировка задания	Тип задания	Ключ к оцениванию задания
ОПК-8 Способен осуществлять эффективное управление разработкой программных средств и проектов.			
1.	<i>Прочитайте текст и установите последовательность. Запишите соответствующую последовательность цифр слева направо:</i> Компания занимается розничной торговлей и хочет внедрить решение на основе нейронных сетей для прогнозирования спроса на товары с использованием больших данных (исторические данные продаж, поведенческие метрики пользователей, сезонность, рекламные кампании и др.). Установите правильную последовательность выполнения ключевых этапов проекта по внедрению нейросетевого решения: 1) Выбор архитектуры нейронной сети и гиперпараметров модели. 2) Согласование KPI проекта с заинтересованными сторонами (точность прогноза, сроки, бюджет). 3) Подготовка и очистка данных из различных источников (продажи, логи, CRM, маркетинговые каналы). 4) Тестирование модели на тестовой выборке и оценка точности прогноза. 5) Разработка технического задания на создание интеграционных модулей между системами компании и моделью прогнозирования. 6) Обучение модели на исторических данных. 7) Внедрение модели в производственную среду и мониторинг ее работы. 8) Анализ бизнес-задачи и формулировка целевых метрик прогнозирования. 9) Создание пайплайна обработки данных и автоматизации обучения модели.	Задание закрытого типа на установление последовательности	8, 2, 5, 3, 1, 6, 4, 9, 10, 7

Номер задания	Формулировка задания	Тип задания	Ключ к оцениванию задания
	10) Проведение пилотного запуска и сбор обратной связи от пользователей.		
2.	<i>Прочитайте текст и установите соответствие:</i> Компания занимается электронной коммерцией и хочет внедрить решение на основе Data Mining для прогнозирования поведения клиентов с целью повышения конверсии. Соотнесите этапы проекта (обозначены цифрами) и их содержательное описание (обозначено буквами), чтобы показать понимание жизненного цикла внедрения аналитических решений. А) Выбор инструментов и сред анализа, удовлетворяющих бизнес-требованиям. Б) Анализ полученных результатов, интерпретация закономерностей и принятие решений. В) Определение ключевых метрик успеха, согласование целей с заинтересованными сторонами Г) Извлечение данных из CRM, логов, баз данных, их объединение и нормализация. Д) Обучение модели на тренировочной выборке, подбор алгоритмов и параметров. Е) Тестирование модели на тестовой выборке, анализ точности и адекватности. Ж) Создание автоматизированного пайплайна и интеграции модели в существующую систему. 3) Устранение пропусков, выбросов, стандартизация форматов данных. 1) Формулировка целей и задач проекта 2) Сбор и подготовка данных 3) Выбор программного обеспечения 4) Обработка и очистка данных 5) Построение модели 6) Оценка качества модели	Задание закрытого типа на установление соответствия	А - 3 Б – 8 В - 1 Г – 2 Д – 5 Е – 6 Ж – 7 3 – 4

Номер задания	Формулировка задания	Тип задания	Ключ к оцениванию задания
	7) Интеграция модели в бизнес-процессы 8) Анализ результатов и принятие решений		
3.	<i>Прочитайте текст и установите соответствие:</i> Компания занимается обработкой больших объемов пользовательских данных с целью прогнозирования поведения клиентов. Соотнесите ключевые этапы работы с большими данными (обозначены цифрами) и их содержательное описание (обозначены буквами), чтобы показать понимание жизненного цикла проекта по анализу и извлечению Big Data. А) Процесс очистки, фильтрации, нормализации и подготовки данных к анализу. Б) Сохранение обработанных данных в хранилище или систему аналитики. В) Сбор информации из разных источников (CRM, логи, веб-ресурсы и др.). Г) Применение статистических и машинных методов для выявления закономерностей и тенденций. Д) Процесс обнаружения скрытых закономерностей, ассоциаций и моделей в больших массивах данных. 1. Извлечение данных (Extract) 2. Преобразование данных (Transform) 3. Загрузка данных (Load) 4. Анализ данных 5. Data Mining	Задание закрытого типа на установление соответствия	А – 2 Б – 3 В – 1 Г – 4 Д – 5
4.	<i>Прочитайте текст и запишите развернутый обоснованный ответ:</i> Перечислите основные функции нейронных сетей в обработке и анализе Big Data. Сформулируйте словосочетаниями минимум 4 функции.	Задание открытого типа с развернутым ответом	1. Классификация и кластеризация данных. 2. Прогнозирование и регрессия. 3. Обнаружение аномалий. 4. Обработка естественного языка (NLP). 5. Обработка изображений и видео. 6. Рекомендательные системы.

Номер задания	Формулировка задания	Тип задания	Ключ к оцениванию задания
5.	<i>Прочитайте текст и запишите развернутый обоснованный ответ:</i> Сформулируйте одним предложением отличие между Big Data и классическим анализом данных.	Задание открытого типа с развернутым ответом	Big Data обрабатывает неструктурированные данные – их нужно подготовить и обработать с помощью специальных алгоритмов. А при классическом анализе данных работают с уже готовыми и упорядоченными наборами информации.
6.	<i>Прочитайте текст и запишите развернутый обоснованный ответ:</i> Сформулируйте 3 основные цели на которые может быть направлен процесс извлечения данных согласно методологии работы с большими данными.	Задание открытого типа с развернутым ответом	1. Архивация (перевод данных из физических форматов в цифровые). 2. Изменение формата данных (адаптация под новые системы). 3. Анализ данных (подготовка к исследованию и выявлению закономерностей).
7.	<i>Прочитайте текст и запишите развернутый обоснованный ответ:</i> Сформулируйте и обоснуйте подход, требующийся для реализации функциональности программной системы внедрения алгоритмов машинного обучения.	Задание открытого типа с развернутым ответом	Система, которая способна самостоятельно изменять свои алгоритмы (адаптивная). Это важно при работе с данными, которые не структурированы и постоянно меняются.
8.	<i>Прочитайте текст и запишите развернутый обоснованный ответ:</i> Какую роль играет процесс ETL при организации эффективного потока данных в проекте разработки информационной системы, ориентированной на работу с большими данными?	Задание открытого типа с развернутым ответом	ETL (Extract – извлечение, Transform – преобразование, Load – загрузка) – это процесс извлечения, преобразования и загрузки данных из разных источников. Он нужен, чтобы подготовить данные для анализа и создания информационных систем.
9.	<i>Прочитайте текст и запишите развернутый обоснованный ответ:</i> Сформулируйте минимум 3 задачи, которые могут быть решены с помощью инструментов Data Mining.	Задание открытого типа с развернутым ответом	1. Классификация. 2. Прогнозирование. 3. Кластеризация. 4. Выявление ассоциаций. 5. Анализ отклонений. 6. Анализ последовательностей.
10.	<i>Прочитайте текст и запишите развернутый обоснованный ответ:</i> Каким образом применение технологии Data Mining может способствовать достижению бизнес-целей	Задание открытого типа с развернутым ответом	Data Mining помогает находить скрытые закономерности и тенденции в больших данных. Это позволяет принимать взвешенные управленческие решения и улучшать аналитические системы

Номер задания	Формулировка задания	Тип задания	Ключ к оцениванию задания
	проекта при разработке программных средств анализа больших данных?		
11.	<i>Прочитайте текст, выберите все правильные варианты ответов и запишите аргументы, обосновывающие выбор ответа:</i> Выберите основные задачи, которые решаются на этапе преобразования больших данных: 1) Удаление дубликатов и аномалий 2) Объединение данных из разных источников 3) Приведение данных к единому формату 4) Загрузка данных в облачное хранилище 5) Выявление скрытых закономерностей 6) Нормализация и фильтрация данных	Задания комбинированного типа с выбором нескольких верных ответов из предложенных и обоснованием выбора	1, 3, 6 Обоснование: №1 Удаление дубликатов и аномалий (выбросов) поможет правильно улучшить качество обработки информации. № 3. Данные могут приходить из разных источников и в разных форматах: текст, изображения, числа и так далее. Чтобы их можно было анализировать вместе, нужно привести их к единому виду: например, все даты должны быть в одном формате, а текст – переведён в числа. № 6. Нормализация приводит данные к единой шкале, а фильтрация убирает ненужную информацию