**МИНИСТЕРСТВО науки и высшего образования Российской Федерации**

федеральное Государственное бюджетное образовательное учреждение высшего образования

**«ТЮМЕНСКИЙ ИНДУСТРИАЛЬНЫЙ университет»**

|  |
| --- |
| **Утверждаю:**  \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  «\_\_\_» \_\_\_\_\_\_\_\_\_\_\_\_ |

**ФОНД ОЦЕНОЧНЫХ СРЕДСТВ**

|  |  |
| --- | --- |
| дисциплины: | **Управление проектами в Big Data** |
| направление подготовки: | **01.04.02 Прикладная математика и информатика** |
| направленность (профиль): | **Машинное обучение и анализ данных** |
| форма обучения: | **очная, очно-заочная** |

Фонд оценочных средств рассмотрен

на заседании кафедры бизнес – информатики и математики

Протокол № \_\_\_ от «\_\_\_\_\_» \_\_\_\_\_\_\_\_\_\_\_\_\_

**Процесс изучения дисциплины направлен на формирование следующих компетенций:**

|  |  |
| --- | --- |
| Код и наименование компетенции | Код и наименование индикатора достижения компетенции (ИДК) |
| УК-2. Способен управлять проектом на всех этапах его жизненного цикла | УК-2.1. Способен планировать этапы жизненного цикла управления проектами |
|
|
| УК-2.2. Способен разрабатывать план реализации проекта с учетом действующих стандартов |
|
|
| УК-2.3. Способен управлять проектами и оценивать их эффективность. |
|
|
| УК-3. Способен организовывать и руководить работой команды, вырабатывая командную стратегию для достижения поставленной цели | УК-3.1. Способен организовывать и управлять командной работой |
|
|
| УК-3.2. Обладает коммуникативными навыками коллективной работы над проектом |
|
|
| ПКС-1. Способен управлять этапами жизненного цикла методологической и технологической инфраструктуры анализа больших данных | ПКС-1.1. Способен проводить сравнительный анализ и выбор источников данных, средств хранения и обработки данных, оценку условий их приобретения и использования |
|
|
| ПКС-1.2. Способен оценивать экономические показатели (затраты на внедрение/модификацию, эффективность и пр.) технологий больших данных |
|
|

**Для оценки УК-2** Способен управлять проектом на всех этапах его жизненного цикла

1. Выберите понятие: программа проектов это -

1. совокупность проектов, находящихся в компетенции одного центра ответственности
2. группа взаимосвязанных проектов и различных мероприятий, объединенных общей целью и условиями их выполнения
3. -комплекс взаимосвязанных мероприятий, предназначенных для достижения поставленных целей с установленными требованиями к качеству результата в течение заданного времени и при установленном бюджете

**Правильны ответ 2**

**2.** Назвать тип структурной декомпозиции работ

1. Продуктовая СДР
2. Функциональная СДР
3. Организационная СДР

**Правильны ответ 2**

3.Управление риском проекта это

1. системное применение политики, процедур и методов управления к задачам определения ситуации, идентификации, анализа, оценки, обработки, мониторинга риска и обмена информацией, для обеспечения снижения потерь и увеличения рентабельности.
2. системное применение политики, процедур и методов управления командой проекта и обмена информацией, для обеспечения снижения потерь и увеличения рентабельности
3. системное применение политики, процедур и методов управления целями проекта, анализа, оценки, обработки, мониторинга информацией, для обеспечения снижения потерь и увеличения рентабельности
4. системное применение политики, процедур и методов управления к задачам определения ситуации, мониторинга риска и обмена информацией, для обеспечения снижения потерь.

**Правильны ответ 3**

1. Выберите технологию потоковой обработки событий в режиме реального времени
2. Apache [Hadoop](https://bigdataschool.ru/wiki/hadoop)
3. Apache [Kafka](https://bigdataschool.ru/wiki/kafka)
4. [Spark Streaming](https://bigdataschool.ru/wiki/spark-streaming)
5. [MapReduce](https://bigdataschool.ru/wiki/mapreduce)

**Правильны ответ 2**

5. Автоматизировать запуск пакетных задач в рамках конвейера обработки больших данных по расписанию можно с помощью

1. Apache Kafka
2. Apache Hadoop
3. [Apache Hive](https://bigdataschool.ru/wiki/hive-wiki)
4. Apache [AirFlow](https://bigdataschool.ru/wiki/airflow)

**Правильный ответ 4**

6.Apache [NiFi](https://bigdataschool.ru/wiki/nifi) используется для

1. эффективного хранения больших данных
2. визуализации результатов аналитики
3. маршрутизации потоков Big Data и построения ETL-конвейеров
4. оптимизации SQL-запросов к DWH

**Правильный ответ 3**

**7.** Для реализации микросервисной архитектуры и интеграции разрозненных систем подходит

1. Apache Kafka
2. Apache [Spark](https://bigdataschool.ru/wiki/spark)
3. Apache Hadoop
4. Apache AirFlow

**Правильный ответ 1**

**8.** Для распределенного глубокого машинного обучения (Deep Learning) больше подходит фреймворк

1. PyTorch
2. [Scikit-learn](https://bigdataschool.ru/wiki/scikit-learn)
3. Flask
4. TensorFlow

**Правильный ответ 1**

**9.** Анализировать данные, хранящиеся в Apache Hadoop, с помощью стандартного инструментария SQL-запросов

1. Можно
2. Нельзя

**Правильный ответ 1**

**10.** Для машинного обучения подходят данные

1. Числовые типа int
2. Предварительно подготовленные, очищенные от ошибок, пропусков и выбросов, а также нормализованные и представленные в виде числовых векторов
3. Любых форматов в цифровом виде
4. Бинарные

**Правильный ответ 2**

**Для оценки УК-3. Способен организовывать и руководить работой команды, вырабатывая командную стратегию для достижения поставленной цели**

**11.** Вставьте пропущенные слова

Жизненный цикл данных – это \_\_\_\_\_\_, которую конкретная \_\_\_\_\_\_ проходит от начального этапа создания или получения до момента архивации или удаления

**Правильный ответ:** последовательность этапов, порция данных

12. Какие стадии включает жизненный цикл метаданных?

1. оценка требований и анализ контента
2. спецификация системных требований
3. система метаданных
4. сервис и оценка
5. обзор релевантных стандартов и проектов по метаданным
6. идентификация стратегий для схем метаданных
7. оценка качества метаданных

**Правильный ответ 1, 2, 3, 4**

**13.** Какие пять «V» описывают Большие данные?

1. Volume
2. Variety
3. Velocity
4. Variability
5. Veracity
6. V**alue**
7. **View**
8. Vacillation
9. [Vainness](https://wooordhunt.ru/word/vainness)
10. Vantage

**Правильный ответ 1, 2, 3, 4, 5**

**14.** R является:

1. языком программирования для статистической обработки данных и работы с графикой, созданный на основе языка S
2. свободной программной средой вычислений с открытым исходным кодом
3. высокоуровневым языком программирования общего назначения, ориентированным на повышение производительности разработчика и читаемости кода
4. языком программирования с динамической типизацией данных

**Правильный ответ 1**

**15.** Присваивание в языке R можно осуществить с помощью оператора:

1. =
2. <-
3. ->
4. :=

**Правильный ответ 1, 2, 3**

**16.** Присваивание в R нельзя осуществить с помощью оператора:

1. =
2. <-
3. ->
4. :=

**Правильный ответ 4**

**17.** Какие из вариантов присвоения не вызовут ошибки:

1. x -> 3
2. x <- 3
3. 3 <- x
4. 3 -> x

**Правильный ответ 2, 4**

**18.** Как сделать вектор из трех чисел?

1. (7,7,7)
2. [7,7,7]
3. {7,7,7}
4. c(7,7,7)

**Правильный ответ 4**

**19.** В каких из следующих случаях исполнение выдаст ошибку?

1. (7,7,7)
2. [7,7,7]
3. {7,7,7}
4. c(7,7,7)

**Правильный ответ 1, 2, 3**

**20.** Для каких аргументов функция is.finite вернет true?

1. 1
2. NA
3. NaN
4. (+Inf)

**Правильный ответ 1**

**Для оценки ПКС-1.** Способен управлять этапами жизненного цикла методологической и технологической инфраструктуры анализа больших данных

**21.** Каким образом можно сформировать вектор (FALSE, FALSE, TRUE)?

1. c(FALSE, FALSE, TRUE)
2. 2:4>3
3. 2:4<3
4. С(TRUE, TRUE, FALSE)

**Правильный ответ 1, 2**

**22.** К классификации с точки зрения ... относится деление организационных структур на функциональные, матричные и проектные

**Правильный ответ:** Содержания проекта

**23.** Набор последовательных или перекрывающихся фаз проекта это-

**Правильный ответ:** Жизненный цикл

**24.** Может ли Apache Kafka работать без службы Zookeeper?

1. Да, если используется другой механизм обеспечения надежной согласованности состояния кластера и его конфигурации.
2. Нет, ZooKeeper обязателен для любых распределенных систем, т.к. обеспечивает синхронизацию кластерных узлов.
3. Нет, ZooKeeper является встроенным механизмом платформы Apache Kafka.
4. Да, Apache Kafka является самодостаточной отказоустойчивой распределенной системой и не нуждается в дополнительных сервисах.

**Правильный ответ 1**

**25.** С какими форматами данных может работать Apache Kafka?

1. Только JSON и AVRO
2. бинарный формат (binary)
3. только Parquet
4. С любыми благодаря механизму работы с разными схемами данных, реализованному в компоненте Schema Registry.

**Правильный ответ 1**

**26.** Альтернативы Apache Kafka, которые могут заменить эту платформу потоковой обработки событий в высоконагруженных Big Data проектах

1. Только JSON и AVRO
2. бинарный формат (binary)
3. только Parquet
4. Apache Pulsar

**Правильный ответ: 4**

**27.** Назовите способ повышения производительности Apache Kafka -

**Правильный ответ:** Увеличение коэффициента репликации

28. Что понимается под данными, согласно ГОСТ 7.0-992?

1. это информация, обработанная и представленная в формализованном виде для дальнейшей обработки
2. это представление информации в формализованном виде, пригодном для передачи, интерпретации и обработки
3. это информация, особенно факты и числа, собранные для последующего использования при принятии решений
4. это информация в электронной форме, пригодная для хранения и использования компьютером

**Правильный ответ 1**

1. Что понимается под данными, согласно ГОСТ 7.0-992?
2. это информация, обработанная и представленная в формализованном виде для дальнейшей обработки
3. это представление информации в формализованном виде, пригодном для передачи, интерпретации и обработки
4. это информация, особенно факты и числа, собранные для последующего использования при принятии решений
5. это информация в электронной форме, пригодная для хранения и использования компьютером

**Правильный ответ 1**

1. Какой федеральный закон регулирует право свободно искать, получать, передавать, производить и распространять информацию любым законным способом, при применении информационных технологий?
2. ФЗ-42
3. часть 4 статьи 29 Конституции РФ
4. ФЗ-149
5. часть 1 статьи 24 Конституции РФ

**Правильный ответ 3**

**Критерии оценивания уровня сформированности компетенций**

Оценка «отлично» выставляется обучающемуся, если им даны правильные ответы более, чем на 90% вопросов.

Оценка «хорошо» выставляется обучающемуся, если им даны правильные ответы на 76-90% вопросов.

Оценка «удовлетворительно» выставляется обучающемуся, если им даны правильные ответы на 61-75% вопросов.

Оценка «неудовлетворительно» выставляется обучающемуся, если им даны правильные ответы менее чем на 61% вопросов.