

Питання до екзамену з дисципліни «Big Data та хмарні технології»

1. Дайте визначення поняттю «великі дані», які структури великих масивів даних віднесені до Big Data.
2. Наведіть основні методи та технології обробки та аналізу великих масивів даних
3. Основні принципи роботи з великими масивами даних (масштабованість, локальність, розподільність та відмовостійкість).
4. Інфраструктура збереження та обробки великих масивів даних.
5. Регресійний аналіз даних. Приклади задач, які можуть бути вирішені за допомогою регресійного аналізу.
6. Які ймовірнісні методи використовуються для аналізу великих даних, наведіть приклади ймовірнісних розподілів та виконайте порівняння.
7. Метод опорних векторів. Приклади задач, які можуть бути вирішені за допомогою регресійного аналізу.
8. Інфраструктура збереження і обробки великих об'ємів даних Hadoop.
9. Наведіть методи, які використовуються для класифікації великих даних та прогнозування.
10. Метод опорних векторів. Застосування, переваги та недоліки методу, що використовується для аналізу великих даних.
11. Методи найближчого сусіда KNN, WKNN. Порівняння та вибір коефіцієнтів.
12. Байєсовський метод аналізу великих масивів даних.
13. Порівняння методів аналізу великих даних, методи найближчого сусіда KNN та WKNN, Байєсовський метод. Переваги та недоліки.
14. Сервіси AWS для роботи з масивами великих даних BigData. Amazon EMR, Amazon Glue, Amazon StepFunctions, Amazon Textract, Amazon Lex, Amazon Polly, Amazon Rekognition.
15. Основні визначення щодо «хмарних технологій», види хмар та їхня класифікація.
16. Методи розгортання (приватна, публічна та гібридна).
17. Порівняння та приклади реалізації кожного виду хмарних інфраструктур.
18. Класифікація хмарних систем IaaS, SaaS, PaaS.
19. Актуальні проблеми та перспективи розвитку хмарних технологій.
20. Google та Amazon хмарні сервіси. Конфіденційність та безпека хмарних технологій та сервісів на їх основі.
21. Хмарні сховища. Хмарні додатки. Порівняння та приклади реалізації кожного виду хмарних інфраструктур для використання в задачах дослідження соціальних, економічних та бізнес процесів.
22. Хмарні обчислення Cloud Computing Service.
23. Статистичні методи для аналізу великих масивів даних в задачах аналізу гіпотез економічних та бізнес рішень, шляхом виявлення закономірностей і кореляцій.
24. Основні підходи щодо масштабування, розміщення та резервного копіювання.
25. Використання сплайн- та вейвлет- апроксимації для обробки великих масивів даних.