

[Dashboard](#) / [My courses](#) / [ATA 2020/2021 | 2-FIKTI | Komputasi Big Data | TEAM TEACHING](#) / [Pertemuan 1 : Big Data Computing Overview](#) / [Kuis M1 : Big Data Computing Overview](#)

Started on	Monday, 8 March 2021, 11:31 AM
State	Finished
Completed on	Monday, 8 March 2021, 11:33 AM
Time taken	1 min 46 secs
Grade	95.00 out of 100.00

Question 1

Correct

Mark 5.00 out of 5.00

Beragamnya jenis data di dalam Big Data yang meliputi data terstruktur, semi terstruktur dan bahkan data tidak terstruktur, merupakan penjelasan dari dimensi Big Data dalam hal:

Select one:

- ☐ Velocity
- ☐ Volume
- ☒ Variety ✓
- ☐ Veracity

The correct answer is: Variety

Question 2

Correct

Mark 5.00 out of 5.00

Contoh solusi Big Data, kecuali:

Select one:

- ☐ Mesin rekomendasi
- ☒ Input Nilai Mahasiswa ✓
- ☐ Semua benar
- ☐ Deteksi Penipuan

The correct answer is: Input Nilai Mahasiswa

Question 3

Correct

Mark 5.00 out of 5.00

Dalam arsitektur Big Data, solusi berupa rekomendasi, berada pada posisi sebagai:

Select one:

- ☐ Data Computing & Analysis
- ☐ Data Storage Layer
- ☐ Data Format
- ☒ End Result ✓

The correct answer is: End Result



Question **4**

Correct

Mark 5.00 out of 5.00

Dalam ekosistem Big Data, proses pre-prosesing data dapat dilakukan menggunakan:

Select one:

- ☐ HDFS
- ☒ Sqoop ✓
- ☐ Facebook
- ☐ Tableau

The correct answer is: Sqoop

Question **5**

Incorrect

Mark 0.00 out of 5.00

Data Warehouse dibuat dengan menerapkan serangkaian langkah pemrosesan terhadap data yang berasal dari beberapa basis data. Berikut ini adalah yang termasuk dalam pemrosesan data tersebut:

Select one:

- ☐ Pembersihan data
- ☐ Semua benar
- ☒ Integrasi data ✗
- ☐ Transformasi data

The correct answer is: Semua benar

Question **6**

Correct

Mark 5.00 out of 5.00

Dataset dibentuk dari:

Select one:

- ☐ Semua salah
- ☐ Sekumpulan fitur yang memiliki nilai yang sama
- ☐ Sekumpulan fitur yang memiliki objek yang sama
- ☒ Sekumpulan objek yang memiliki fitur yang sama ✓

The correct answer is: Sekumpulan objek yang memiliki fitur yang sama

Question **7**

Correct

Mark 5.00 out of 5.00

Komponen Komputasi Big Data di lapisan Middleware terdiri dari:

Select one:

- ☒ SDK Layer ✓
- ☐ Semua benar
- ☐ Resource Layer
- ☐ Interface Layer

The correct answer is: SDK Layer

Question **8**

Correct

Mark 5.00 out of 5.00

Kumpulan data yang disusun dalam tabel berformat persegi panjang yang terdiri dari baris dan kolom disebut sebagai:

Select one:

- ☐ Semua salah
- ☐ Data Warehouse
- ☐ Database
- ☒ Flat files ✓

The correct answer is: Flat files

Question **9**

Correct

Mark 5.00 out of 5.00

Manakah yang benar dari pernyataan berikut ini:

Select one:

- ☒ Memperkirakan puncak dari kasus positif Covid-19, adalah contoh analisis prediktif. ✓
- ☐ Prediktif, menilai tindakan
- ☐ Deskriptif, perkiraan berdasarkan data yang tersedia
- ☐ Preskriptif, memodelkan perilaku masa lalu

The correct answer is: Memperkirakan puncak dari kasus positif Covid-19, adalah contoh analisis prediktif.

Question **10**

Correct

Mark 5.00 out of 5.00

Model basis data yang berupa deret waktu (time series):

Select one:

- ☒ Temporal database ✓
- ☐ Text database
- ☐ Semua salah
- ☐ Spatial database

The correct answer is: Temporal database

Question **11**

Correct

Mark 5.00 out of 5.00

Objek disebut juga:

Select one:

- ☒ Semua benar ✓
- ☐ Unit
- ☐ Data points
- ☐ Record

The correct answer is: Semua benar

Question **12**

Correct

Mark 5.00 out of 5.00

Salah satu keunggulan DGX A100:

Select one:

- ☐ Konsumsi daya lebih sedikit
- ☒ Semua benar ✓
- ☐ Biaya lebih murah
- ☐ Tidak membutuhkan ruang yang besar

The correct answer is: Semua benar

Question **13**

Correct

Mark 5.00 out of 5.00

Sekumpulan tabel dimana bentuk setiap tabel analog dengan flat file, disebut:

Select one:

- ☐ Semua salah
- ☐ Data Warehouse
- ☐ DBMS
- ☒ Database ✓

The correct answer is: Database

Question **14**

Correct

Mark 5.00 out of 5.00

Semakin banyak data hingga mencapai ukuran Peta Byte atau bahkan lebih, merupakan penjelasan dari dimensi Big Data dalam hal:

Select one:

- ☐ Variety
- ☒ Volume ✓
- ☐ Velocity
- ☐ Veracity

The correct answer is: Volume

Question **15**

Correct

Mark 5.00 out of 5.00

Sumber data yang digunakan dalam aplikasi Big Data dapat berupa:

Select one:

- ☐ Data Asuransi Kesehatan
- ☐ Data Finansial
- ☒ Semua benar ✓
- ☐ Data Media Sosial

The correct answer is: Semua benar

Question **16**

Correct

Mark 5.00 out of 5.00

Tempat penyimpanan data yang dikumpulkan di lokasi berbeda (basis data relasional) dan disimpan menggunakan skema terpadu:

Select one:

- ☒ Data Warehouse ✓
- ☐ Transactional database
- ☐ Semua benar
- ☐ Spatial Database

The correct answer is: Data Warehouse

Question **17**

Correct

Mark 5.00 out of 5.00

Tipe nilai dapat berupa:

Select one:

- ☐ Object
- ☐ Fitur
- ☐ Semua salah
- ☒ Numerik ✓

The correct answer is: Numerik

Question **18**

Correct

Mark 5.00 out of 5.00

Unit informasi tunggal disebut:

Select one:

- ☐ Fitur/atribut
- ☐ Dataset
- ☒ Nilai ✓
- ☐ Objek

The correct answer is: Nilai

Question **19**

Correct

Mark 5.00 out of 5.00

Visualisasi dari hasil analisis Big Data dapat berupa:

Select one:

- ☒ Semua benar ✓
- ☐ Tabel
- ☐ Grafik 3D
- ☐ Plot

The correct answer is: Semua benar

Question **20**
Correct
Mark 5.00 out of 5.00

Yang merupakan framework untuk Machine Learning:

- Select one:
- ☐ Azure ML
 - ☐ Amazon ML
 - ☒ Apache Spark MLlib ✓
 - ☐ Semua benar

The correct answer is: Apache Spark MLlib

◀ Materi M1 : Big Data Computing Overview

Jump to...

Kuis M1 : Big Data Computing Overview (Khusus Kelas Malam) ▶

Started on	Monday, 15 March 2021, 11:33 AM
State	Finished
Completed on	Monday, 15 March 2021, 11:36 AM
Time taken	3 mins 56 secs
Grade	100.00 out of 100.00

Question **1**
Correct
Mark 5.00 out of 5.00

Arsitektur Hadoop terdiri dari:

Select one:

- ☐ a. Storage layer, Processing Layer, Application Layer
- ☒ b. Storage layer, Resource Management Layer, Application Layer ✓
- ☐ c. Data layer, Processing Layer, Service Layer
- ☐ d. Storage layer, Service Layer, Application Layer

The correct answer is: Storage layer, Resource Management Layer, Application Layer

Question **2**
Correct
Mark 5.00 out of 5.00

Berikut ini adalah saran untuk memasuki era Analytic 3.0:

Select one:

- ☐ a. Mulailah dengan kemampuan yang ada untuk manajemen data dan analitik
- ☐ b. Pilih target dan sub-target analytics 3.0
- ☒ c. Semua benar ✓
- ☐ d. Tambahkan beberapa data volume besar yang tidak terstruktur

The correct answer is: Semua benar

Question **3**
Correct
Mark 5.00 out of 5.00

Berikut ini yang harus dilakukan oleh perusahaan 1.0, kecuali:

Select one:

- ☒ a. Membuat analitik dan keputusan berbasis fakta elemen kunci dari hubungan pelanggan, produk dan layanan, strategi, dan persaingan ✓
- ☐ b. Temukan pelanggan terbaik, tetapkan harga terbaik kepada mereka
- ☐ c. Buat penawaran yang efektif
- ☐ d. Pahami biaya dan dorong kinerja yang lebih baik

The correct answer is: Membuat analitik dan keputusan berbasis fakta elemen kunci dari hubungan pelanggan, produk dan layanan, strategi, dan persaingan



Question **4**
Correct
Mark 5.00 out of 5.00

Contoh Analytic 3.0:

Select one:

- ☒ a. Semua benar ✓
- ☐ b. Data volume tinggi waktu nyata dari sensor dan mesin
- ☐ c. Fokus pada anomali dan pengecualian
- ☐ d. Analisis untuk tren, pola, prediksi, pengoptimalan, kesadaran situasional

The correct answer is: Semua benar

Question **5**
Correct
Mark 5.00 out of 5.00

Contoh Aplikasi Machine Learning dan Analitik Prediktif:

Select one:

- ☐ a. Prediksi wabah penyakit
- ☒ b. Semua benar ✓
- ☐ c. Penetapan harga dinamis
- ☐ d. Prediksi Cashflow

The correct answer is: Semua benar

Question **6**
Correct
Mark 5.00 out of 5.00

Contoh data tidak terstruktur:

Select one:

- ☒ a. Data media sosial ✓
- ☐ b. Data transaksi penjualan
- ☐ c. Data pegawai
- ☐ d. Semua benar

The correct answer is: Data media sosial

Question **7**
Correct
Mark 5.00 out of 5.00

Contoh implementasi dalam perusahaan:

Select one:

- ☐ a. Sistem Manajemen Inventaris Cerdas pada perusahaan Ford, menyarankan pesanan mingguan untuk dealer dan telah meningkatkan pendapatan sebesar \$ 100 juta / tahun
- ☐ b. Procter & Gamble melakukan Analisis sentimen media sosial real time untuk "Pulsa Konsumen"
- ☒ c. Semua benar ✓
- ☐ d. Monsanto membayar hampir \$ 1 miliar untuk The Climate Company, yang mengumpulkan dan menganalisis data cuaca untuk pertanian

The correct answer is: Semua benar

Question **8**

Correct

Mark 5.00 out of 5.00

Dalam computer vision, contoh penerapan Machine Learning adalah, kecuali:

Select one:

- ☒ a. Analitik Rute ✓
- ☐ b. Pengenalan wajah
- ☐ c. Deteksi suhu tubuh
- ☐ d. Pengenalan objek

The correct answer is: Analitik Rute

Question **9**

Correct

Mark 5.00 out of 5.00

Istilah 'Machine Learning' diciptakan pada tahun <T> oleh <P>:

Select one:

- ☐ a. T = 1969, P = Yann LeCun
- ☒ b. T = 1959, P = Arthur Samuel ✓
- ☐ c. T = 1989, P = Tom M. Mitchell
- ☐ d. T = 1999, P = Bill Gates

The correct answer is: T = 1959, P = Arthur Samuel

Question **10**

Correct

Mark 5.00 out of 5.00

Komponen Hadoop yang berfungsi sebagai data processing:

Select one:

- ☐ a. YARN
- ☐ b. HDFS
- ☐ c. Semua benar
- ☒ d. MapReduce ✓

The correct answer is: MapReduce

Question **11**

Correct

Mark 5.00 out of 5.00

Komponen Hadoop yang berfungsi sebagai data storage:

Select one:

- ☐ a. Semua benar
- ☐ b. MapReduce
- ☐ c. YARN
- ☒ d. HDFS ✓

The correct answer is: HDFS

Question **12**

Correct

Mark 5.00 out of 5.00

Manakah pernyataan yang benar tentang era Analytic 3.0:

Select one:

- ☐ a. Data yang kompleks, besar dan tidak terstruktur
- ☒ b. Perpaduan antara analitik tradisional dengan Big Data ✓
- ☐ c. Dibutuhkan kapabilitas analitis dan komputasi baru
- ☐ d. Sumber data dari internal, relatif kecil dan data terstruktur dari pelanggan

The correct answer is: Perpaduan antara analitik tradisional dengan Big Data

Question **13**

Correct

Mark 5.00 out of 5.00

Manakah pernyataan yang benar tentang Hadoop:

Select one:

- ☐ a. Hadoop merupakan Google File System
- ☒ b. Hadoop merupakan framework untuk bekerja dengan Big Data ✓
- ☐ c. Hadoop mulai dikembangkan pada tahun 2008
- ☐ d. Semua benar

The correct answer is: Hadoop merupakan framework untuk bekerja dengan Big Data

Question **14**

Correct

Mark 5.00 out of 5.00

Netflix adalah contoh perusahaan yang tergolong:

Select one:

- ☐ a. 1.0
- ☐ b. Semua benar
- ☒ c. 2.0 ✓
- ☐ d. 3.0

The correct answer is: 2.0

Question **15**

Correct

Mark 5.00 out of 5.00

Perbandingan SQL vs NoSQL:

Select one:

- ☐ a. Structured Query Language vs Un-structured Query Language
- ☐ b. Relational vs Non-Relational
- ☒ c. Semua benar ✓
- ☐ d. Static vs Dynamic

The correct answer is: Semua benar

Question **16**

Correct

Mark 5.00 out of 5.00

Perbandingan Traditional vs Big Data Analytics:

Select one:

- ☐ a. Process centric vs Information centric
- ☐ b. Mahal vs Terjangkau
- ☐ c. Hard to scale vs Scales out forever
- ☒ d. Semua benar ✓

The correct answer is: Semua benar

Question **17**

Correct

Mark 5.00 out of 5.00

Salah satu bentuk data tidak terstruktur:

Select one:

- ☐ a. Document
- ☐ b. Graph
- ☐ c. Key-Value
- ☒ d. Semua benar ✓

The correct answer is: Semua benar

Question **18**

Correct

Mark 5.00 out of 5.00

Salah satu fakta yang menunjukkan aspek Velocity dari Big Data:

Select one:

- ☐ a. 27% responden dalam satu survei tidak yakin dengan banyaknya data mereka yang tidak akurat
- ☐ b. Semua benar
- ☒ c. Mobil modern memiliki 100 sensor yang memantau berbagai item pada mobil ✓
- ☐ d. Pertumbuhan data tidak terstruktur mencapai 60% - 80% per tahun

The correct answer is: Mobil modern memiliki 100 sensor yang memantau berbagai item pada mobil

Question **19**

Correct

Mark 5.00 out of 5.00

Salah satu tonggak penting dalam Big Data Analytic dalam aspek pengadaan infrastruktur:

Select one:

- ☐ a. Analisis data untuk setiap kasus bisnis spesifik Anda dan visualisasikan dengan cara yang bermakna
- ☒ b. Dapatkan infrastruktur yang tepat yang dibutuhkan berdasarkan arsitektur solusi Anda termasuk penyimpanan data dan unit pemrosesan, platform, dan aplikasi ✓
- ☐ c. Adaptasi dan ubah organisasi Anda menuju teknologi big data dan implementasi analitik baru
- ☐ d. Kembangkan arsitektur Solusi untuk kasus bisnis Anda termasuk proses pengumpulan data, arsitektur teknologi dan metode analitik serta visualisasi data

The correct answer is: Dapatkan infrastruktur yang tepat yang dibutuhkan berdasarkan arsitektur solusi Anda termasuk penyimpanan data dan unit pemrosesan, platform, dan aplikasi

Question **20**

Correct

Mark 5.00 out of 5.00

Teknologi yang dibutuhkan pada era Analytic 3.0:

Select one:

- ☐ a. Keduanya benar
- ☐ b. Berpusat pada Data Warehouse
- ☒ c. Ketergantungan yang tinggi pada Machine Learning ✓
- ☐ d. Keduanya salah

The correct answer is: Ketergantungan yang tinggi pada Machine Learning

◀ Materi M2 : [SEGMENT 3] Hadoop & Big Data - Analytics Journey

Jump to...

Kuis M2 : Big Data Computing, Applications and Technologies (khusus kelas Malam) ▶

[Dashboard](#) / [My courses](#) / [ATA 2020/2021 | 2-FIKTI | Komputasi Big Data | TEAM TEACHING](#) / [Pertemuan 3 : Introduction to Apache Hadoop](#) / [Kuis M3 : Introduction to Hadoop](#)

Started on	Monday, 22 March 2021, 11:30 AM
State	Finished
Completed on	Monday, 22 March 2021, 11:34 AM
Time taken	3 mins 51 secs
Grade	95.00 out of 100.00

Question **1**
Correct
Mark 5.00 out of 5.00

Jenis database NoSQL apa yang membuat dan menjalankan aplikasi yang berjalan dengan dataset yang selalu terhubung menjadi lebih mudah dan sering digunakan untuk jaringan media sosial.

Select one:

- ☐ a. Primary key
- ☒ b. Grafik ✓
- ☐ c. Dokumen
- ☐ d. Pencarian

The correct answer is: Grafik

Question **2**
Correct
Mark 5.00 out of 5.00

Bagian pekerjaan utama MapReduce adalah

Select one:

- ☐ a. map dan combine
- ☐ b. combine dan partition
- ☒ c. map dan reduce ✓
- ☐ d. shuffle dan sort

The correct answer is: map dan reduce

Question **3**
Correct
Mark 5.00 out of 5.00

NoSQL database adalah singkatan dari

Select one:

- ☒ a. Not-only Structured Query Language database ✓
- ☐ b. No Like Structured Query Language database
- ☐ c. Numbered of Structured Query Language database
- ☐ d. Numbered of sequel database

The correct answer is: Not-only Structured Query Language database



Online

Question **4**

Incorrect

Mark 0.00 out of 5.00

Database NoSQL menggunakan beragam model data untuk mengakses dan mengelola data. Jenis database ini dioptimalkan secara khusus untuk aplikasi yang memerlukan..., yang dicapai dengan mengurangi pembatasan konsistensi data dari database lainnya.

Select one:

- ☒ a. volume data besar, latensi rendah, dan model data fleksibel ✖
- ☐ b. volume data besar, latensi rendah, dan model data fleksibel
- ☐ c. volume data besar, latensi rendah, dan model data fleksibel
- ☐ d. volume data besar, latensi rendah, dan model data fleksibel

The correct answer is: volume data besar, latensi rendah, dan model data fleksibel

Question **5**

Correct

Mark 5.00 out of 5.00

Type data berikut yang dapat ditangani oleh Hadoop adalah

Select one:

- ☐ a. Tidak terstruktur
- ☐ b. Semi – terstruktur
- ☐ c. Terstruktur
- ☒ d. Semua benar ✔

The correct answer is: Semua benar

Question **6**

Correct

Mark 5.00 out of 5.00

Pernyataan yang benar mengenai MapReduce, adalah:

Select one:

- ☐ a. memproses data besar secara parallel dalam cluster pada commodity hardware
- ☒ b. semua benar ✔
- ☐ c. implementasi dari framework pemrosesan batch
- ☐ d. membagi masalah besar menjadi masalah kecil yang dapat diselesaikan dengan cepat

The correct answer is: semua benar

Question **7**

Correct

Mark 5.00 out of 5.00

Sistem penyimpanan data cluster adalah

Select one:

- ☐ a. Menyimpan data dengan mempartisi secara horizontal dan pecahan horizontal tersebut di simpan secara terpisah pada node server.
- ☐ b. Data disimpan dengan cara menyalin dataset dan menyimpan Salinan tersebut pada beberapa node
- ☐ c. Penyimpanan dan pengaturan data pada perangkat penyimpanan harddisk atau flashdisk, dan menyediakan logical view dari data yang disimpan dalam struktur hirakis direktori file.
- ☒ d. Kumpulan server atau node digabungkan secara erat, dengan spesifikasi perangkat keras yang sama. Serta terhubung melalui jaringan dan bekerja sebagai satu kesatuan ✔

The correct answer is: Kumpulan server atau node digabungkan secara erat, dengan spesifikasi perangkat keras yang sama. Serta terhubung melalui jaringan dan bekerja sebagai satu kesatuan

Question **8**

Correct

Mark 5.00 out of 5.00

Terdapat 2 (dua) teknologi Penyimpanan Big Data, yaitu

Select one:

- ☐ a. Flashdisk dan harddisk
- ☐ b. Reliable dan Fault tolerance
- ☐ c. SQL dan No-SQL
- ☒ d. On Disk Storage Devices dan In Memory Storage Devices ✓

The correct answer is: On Disk Storage Devices dan In Memory Storage Devices

Question **9**

Correct

Mark 5.00 out of 5.00

Berikut yang dimaksud dengan menganalisa data dalam konteks big data adalah:

Select one:

- ☒ a. Descriptive analytics, diagnostic analytics, predictive analytics, dan prescriptive analytics ✓
- ☐ b. mean statistic analytics, deviation standard analytics, dan percentil analytics
- ☐ c. Menentukan jumlah data dan atribut data
- ☐ d. Classification analytics, regression analytics, clustering analytics, dan hadooping analytics

The correct answer is: Descriptive analytics, diagnostic analytics, predictive analytics, dan prescriptive analytics

Question **10**

Correct

Mark 5.00 out of 5.00

Berikut ini yang bukan merupakan Hadoop operation mode, adalah

Select one:

- ☒ a. Globally distributed mode ✓
- ☐ b. Stand alone mode
- ☐ c. Fully-Distributed mode
- ☐ d. Pseudo distributed mode

The correct answer is: Globally distributed mode

Question **11**

Correct

Mark 5.00 out of 5.00

Properti Big data yang penting, salah satunya adalah scalability, yaitu:

Select one:

- ☒ a. Kemampuan menangani penambahan beban yang diberikan tanpa penurunan kinerja ✓
- ☐ b. Ketersediaan layanan secara terus menerus
- ☐ c. Handal
- ☐ d. Kemampuan untuk terus beroperasi walaupun terjadi kegagalan pada beberapa komponen

The correct answer is: Kemampuan menangani penambahan beban yang diberikan tanpa penurunan kinerja

Question **12**

Correct

Mark 5.00 out of 5.00

Hadoop adalah framework yang

Select one:

- ☐ a. Melibatkan satu atau lebih cluster Komputer
- ☒ b. Semua benar ✓
- ☐ c. Menerapkan programming model yang sederhana
- ☐ d. Memungkinkan pemrosesan secara terdistribusi terhadap data yang berukuran besar

The correct answer is: Semua benar

Question **13**

Correct

Mark 5.00 out of 5.00

Hadoop adalah platform pemrosesan big data yaitu data yang mempunyai karakteristik kecuali:

Select one:

- ☒ a. Jumlah fitur yang tertentu ✓
- ☐ b. Terakumulasi dengan cepat (high speed velocity)
- ☐ c. Inconsistent dan uncertaint
- ☐ d. Volume besar, struktur data yang beragam (variety), dan nilai data yang diperlukan (useful value)

The correct answer is: Jumlah fitur yang tertentu

Question **14**

Correct

Mark 5.00 out of 5.00

Properti Big data yang penting, salah satunya adalah reliability, yaitu:

Select one:

- ☐ a. Kemampuan untuk terus beroperasi walaupun terjadi kegagalan pada beberapa komponen
- ☐ b. Kemampuan menangani penambahan beban yang diberikan tanpa penurunan kinerja
- ☒ c. Handal ✓
- ☐ d. Ketersediaan layanan secara terus menerus

The correct answer is: Handal

Question **15**

Correct

Mark 5.00 out of 5.00

Properti Big data yang penting, salah satunya adalah fault tolerance, yaitu:

Select one:

- ☒ a. Kemampuan untuk terus beroperasi walaupun terjadi kegagalan pada beberapa komponen ✓
- ☐ b. Kemampuan menangani penambahan beban yang diberikan tanpa penurunan kinerja
- ☐ c. Handal
- ☐ d. Ketersediaan layanan secara terus menerus

The correct answer is: Kemampuan untuk terus beroperasi walaupun terjadi kegagalan pada beberapa komponen

Question **16**

Correct

Mark 5.00 out of 5.00

YARN singkatan dari

Select one:

- ☐ a. Yahoo's another resource name
- ☒ b. Yet another resource negotiator ✓
- ☐ c. Yet another resource need
- ☐ d. Yahoo's archived Resource names

The correct answer is: Yet another resource negotiator

Question **17**

Correct

Mark 5.00 out of 5.00

Hadoop adalah framework yang bekerja bersama tools lain yang beragam. Kelompok yang umum bekerja bersama Hadoop adalah

Select one:

- ☐ a. MapReduce, Hummer and Iguana
- ☐ b. MapReduce, Heron and Trumpet
- ☒ c. MapReduce, Hive and Hbase ✓
- ☐ d. MapReduce, MySQL and Google Apps

The correct answer is: MapReduce, Hive and Hbase

Question **18**

Correct

Mark 5.00 out of 5.00

Komponen layanan inti Hadoop adalah

Select one:

- ☐ a. Speed, consistency, dan volume
- ☒ b. MapReduce, HDFS, dan YARN ✓
- ☐ c. SQL, NoSQL, Cluster
- ☐ d. ClickStream, Sensor, dan Sentiment

The correct answer is: MapReduce, HDFS, dan YARN

Question **19**

Correct

Mark 5.00 out of 5.00

Pemrosesan data terdistribusi adalah pemrosesan

Select one:

- ☐ a. yang membagi pemrosesan dalam beberapa batch dan dilakukan secara offline
- ☒ b. yang menerapkan divide and conquer dalam fisik terpisah, namun terhunung Bersama dalam satu cluster ✓
- ☐ c. yang melibatkan pelaksanaan sub-tugas secara bersamaan secara kolektif membentuk tugas yang besar melalui beberapa processor
- ☐ d. yang menggunakan pendekatan pendekatan interaktif tanpa penundaan

The correct answer is: yang menerapkan divide and conquer dalam fisik terpisah, namun terhunung Bersama dalam satu cluster

Question **20**

Correct

Mark 5.00 out of 5.00

Properti Big data yang penting, salah satunya adalah availability, yaitu:

Select one:

- ☐ a. Kemampuan untuk terus beroperasi walaupun terjadi kegagalan pada beberapa komponen
- ☐ b. Handal
- ☒ c. Ketersediaan layanan secara terus menerus ✓
- ☐ d. Kemampuan menangani penambahan beban yang diberikan tanpa penurunan kinerja

The correct answer is: Ketersediaan layanan secara terus menerus

[◀ Materi M3 : Introduction to Hadoop](#)[Kuis M3 : Introduction to Hadoop \(khusus kelas MALAM\)\) ▶](#)

Started on	Monday, 29 March 2021, 11:35 AM
State	Finished
Completed on	Monday, 29 March 2021, 11:46 AM
Time taken	10 mins 56 secs
Grade	100.00 out of 100.00

Question **1**
Correct
Mark 5.00 out of 5.00

Bagaimana menangani missing value dengan tepat?

Select one:

☒ A,B,dan C benar ✓

☐ Imputasi nilai manual

☐ Dibiarkan

☐ Eliminasi data object

Question **2**
Correct
Mark 5.00 out of 5.00

Berikut adalah nama lain dari data objek, kecuali

Select one:

☒ Atribut ✓

☐ Record

☐ Instance

☐ Sample

Question **3**
Correct
Mark 5.00 out of 5.00

Cara menangani noisy data, kecuali

Select one:

☐ Regression

☐ Combined Computer

☒ Ignore ✓

☐ Clustering

Question **4**
Correct
Mark 5.00 out of 5.00

Data berupa image, video bahkan suara termasuk jenis data

Select one:

☐ Record

☐ Ordered

☐ Graf

☒ Multimedia ✓

Online

https://v-class.gunadarma.ac.id/mod/quiz/review.php?attempt=2645316&cmid=502729

1/4

Question **5**

Correct

Mark 5.00 out of 5.00

data cleaning pada data preprocessing dilakukan..

Select one:

- ☐ Identifikasi dan integrasi dari multiple database
- ☒ Imputasi, smoothing dan identifikasi ✓
- ☐ Normalisasi dan diskritasi data
- ☐ Dimensionality reduction dan data compression

Question **6**

Correct

Mark 5.00 out of 5.00

Data Cleansing merupakan tahapan untuk membersihkan data kotor sehingga menjadi dataset yang siap diolah. Data kotor tersebut berupa

Select one:

- ☐ missing value
- ☐ duplicate data
- ☐ outlier
- ☒ A,B,C benar ✓

Question **7**

Correct

Mark 5.00 out of 5.00

Data-data yang saling berhubungan, yang biasanya ditemukan pada jaringan internet termasuk dalam jenis data

Select one:

- ☐ Multimedia
- ☒ Graf ✓
- ☐ Record
- ☐ Ordered

Question **8**

Correct

Mark 5.00 out of 5.00

Imputasi nilai secara otomatis saat menangani incomplete data tidak dapat dilakukan menggunakan

Select one:

- ☐ interpolasi
- ☒ Nilai tertinggi ✓
- ☐ Rata-rata
- ☐ Median

Question **9**

Correct

Mark 5.00 out of 5.00

Jenis kelamin, merupakan salah satu atribut data type...

Select one:

- ☐ Numerik
- ☐ Ordinal
- ☒ Nominal ✓
- ☐ Rasio

Question **10**

Correct

Mark 5.00 out of 5.00

Karakteristik data objek merupakan representasi dari

Select one:

- ☐ Dataset
- ☐ Data objek
- ☐ Data sample
- ☒ Atribut ✓

Question **11**

Correct

Mark 5.00 out of 5.00

Nilai yang merepresentasikan urutan seperti nilai mata kuliah termasuk dalam atribut data type..

Select one:

- ☐ Rasio
- ☒ Ordinal ✓
- ☐ Numerik
- ☐ Nominal

Question **12**

Correct

Mark 5.00 out of 5.00

Normalisasi data dilakukan pada tahap data preprocessing ..

Select one:

- ☒ data transformation ✓
- ☐ data Integration
- ☐ data cleaning
- ☐ data reduction

Question **13**

Correct

Mark 5.00 out of 5.00

Penanganan missing value dapat dilakukan menggunakan nilai estimasi dari probabilitas bersyarat menggunakan metode

Select one:

- ☐ Eliminasi
- ☐ Mean
- ☒ Bayessian ✓
- ☐ Regresi

Question **14**

Correct

Mark 5.00 out of 5.00

Salah satu jenis data record yang diekstrak menjadi kumpulan term dan ditulis berdasarkan frekuensi kemunculan data, termasuk pada jenis data

Select one:

- ☒ Dokumen ✓
- ☐ Matriks
- ☐ Sosial network
- ☐ Transaksi

Question **15**

Correct

Mark 5.00 out of 5.00

Seberapa mudah data dapat dipahami merupakan kualitas data dari..

Select one:

- ☒ Interpretability ✓
- ☐ Believability
- ☐ Completeness
- ☐ Consistency

Question **16**

Correct

Mark 5.00 out of 5.00

Statistik dasar untuk mengukur sebaran data salah satunya adalah

Select one:

- ☒ Variance ✓
- ☐ Median
- ☐ Mean
- ☐ Modus

Question **17**
Correct
Mark 5.00 out of 5.00

Studi tentang bagaimana menganalisis data dan membersihkan data merupakan salah satu tujuan dari

Select one:

- ☐ Data testing
- ☒ Data cleansing ✓
- ☐ Data validation
- ☐ Data training

Question **18**
Correct
Mark 5.00 out of 5.00

Tahapan pertama data preprocessing adalah

Select one:

- ☐ data Integration
- ☐ data transformation
- ☐ data reduction
- ☒ data cleaning ✓

Question **19**
Correct
Mark 5.00 out of 5.00

Tanggal penjualan, nama pelanggan, nama barang, jumlah penjualan merupakan contoh dari

Select one:

- ☐ Atribut penjualan
- ☒ Dataset penjualan ✓
- ☐ Data sample penjualan
- ☐ Data objek penjualan

Question **20**
Correct
Mark 5.00 out of 5.00

Tinggi badan, berat badan, usia, dan jenis kelamin merupakan contoh representasi karakteristik dari

Select one:

- ☐ Data sample
- ☐ Atribut
- ☐ Dataset
- ☒ Data objek ✓

◀ Materi M4 : Data Preprocessing
(update 29/03/2021)

Jump to...

Kuis M4 : Data Preprocessing (khusus kelas MALAM)) ▶

[Dashboard](#) / [My courses](#) / [ATA 2020/2021](#) | [2-FIKTI](#) | [Komputasi Big Data](#) | [TEAM TEACHING](#) / [Pertemuan 5 : Introduction to Data Mining](#)
/ [Kuis M5 : Introduction Data Mining](#)

Started on	Monday, 5 April 2021, 11:57 AM
State	Finished
Completed on	Monday, 5 April 2021, 12:20 PM
Time taken	22 mins 53 secs

Question **1**

Correct

Marked out of 5.00

Metode untuk mempresentasikan hasil yang ditemukan dengan cara yang "bagus" secara visual

- Select one:
- ☐ a. Asosiasi
 - ☒ b. Visualisasi ✓
 - ☐ c. Klasterisasi
 - ☐ d. Klasifikasi

Question **2**

Correct

Marked out of 5.00

Untuk menghindari set pengujian yang tumpang tindih, data dibagi menjadi k himpunan bagian dengan ukuran yang sama kemudiansetiap subset secara bergantian digunakan untuk pengujian dan sisanya untuk pelatihan

- Select one:
- ☒ a. k-fold cross-validation ✓
 - ☐ b. Model Sorted list
 - ☐ c. Commulative PCT Hits
 - ☐ d. Direct Marketing Paradigm

Question **3**

Correct

Marked out of 5.00

Metode untuk mengelompokkan secara "alami" dari instance yang diberi data tanpa label

- Select one:
- ☐ a. Asosiasi
 - ☒ b. Klasterisasi ✓
 - ☐ c. Klasifikasi
 - ☐ d. Visualisasi

Question **4**

Correct

Marked out of 5.00

Pernyataan yang kurang sesuai dengan metode peringkasan

- Select one:
- ☐ a. Menjelaskan fitur dari grup yang dipilih
 - ☒ b. Mengelompokkan data tak berlabel ✓
 - ☐ c. Menggunakan bahasa dan grafik alami
 - ☐ d. Biasanya dalam Kombinasi dengan deteksi Deviasi atau metode lain



Question **5**

Correct

Marked out of
5.00

Acuan yang tidak tepat ketika melakukan pengaturan parameter

Select one:

- ☒ a. Data pengujian harus digunakan untuk penyetelan parameter ✓
- ☐ b. Data pengujian tidak dapat digunakan untuk penyetelan parameter
- ☐ c. Data pengujian tidak digunakan dengan cara apa pun untuk membuat pengklasifikasi
- ☐ d. Skema pembelajaran beroperasi dalam dua tahap yaitu membangun struktur dasar dan mengoptimalkan pengaturan parameter

Question **6**

Correct

Marked out of
5.00

Jenis tugas data mining yang memfasilitasi penerimaan / penelusuran informasi bagi pengguna

Select one:

- ☐ a. Asosiasi
- ☐ b. Klasterisasi
- ☐ c. Klasifikasi
- ☒ d. Visualisasi ✓

Question **7**

Correct

Marked out of
5.00

Karakteristik Big Data yang terus bertambah dengan cepat secara real-time dinamakan

Select one:

- ☒ a. Volume ✓
- ☐ b. variety
- ☐ c. Velocity
- ☐ d. Value

Question **8**

Correct

Marked out of
5.00

Fungsi dari data mining adalah salah satunya dapat mengetahui pola yang sering (atau kumpulan item yang sering) misalnya item apa yang sering dibeli bersama di Walmart Anda. Ini dikenal dengan istilah

Select one:

- ☐ a. Klasterisasi
- ☐ b. Outlier Analysis
- ☐ c. Klasifikasi
- ☒ d. Asosiasi dan Analisis Korelasi ✓

Question **9**

Correct

Marked out of
5.00

Proses pencarian konten baru dan konten yang perlu diperbaharui dengan mengirimkan crawler ke seluruh konten yang ditemukan dari tautan halaman merupakan Salah satu konsep data mining pada Search Engine yang dinamakan

Select one:

- ☐ a. Index
- ☒ b. Crawl ✓
- ☐ c. Rank
- ☐ d. Searching

Question **10**

Correct

Marked out of
5.00

Metode untuk memprediksi kelas instance dari instance yang diberi label sebelumnya

Select one:

- ☐ a. Visualisasi
- ☐ b. Klasterisasi
- ☐ c. Asosiasi
- ☒ d. Klasifikasi ✓

Question **11**

Correct

Marked out of
5.00

Melihat suatu objek data yang tidak sesuai dengan perilaku umum data, sering digunakan dalam deteksi penipuan atau analisis kejadian langka

Select one:

- ☒ a. Analisis Outlier ✓
- ☐ b. Klasifikasi
- ☐ c. Klasterisasi
- ☐ d. Asosiasi dan Analisis Korelasi

Question **12**

Correct

Marked out of
5.00

Seberapa mudah data dapat dipahami merupakan kualitas data dari..

Select one:

- ☐ a. Consistency
- ☐ b. Believability
- ☐ c. Completeness
- ☒ d. Interpretability ✓

Question **13**

Correct

Marked out of
5.00

Regression, Decision Trees, Bayesian, Neural Networks, Naïve Bayes Rules, Support Vector Machines, Genetic Algorithms adalah pendekatan yang digunakan untuk

Select one:

- ☐ a. Asosiasi
- ☒ b. Klasifikasi ✓
- ☐ c. Visualisasi
- ☐ d. Klasterisasi

Question **14**

Correct

Marked out of
5.00

Urutan langkah melakukan klasifikasi

Select one:

- ☐ a. Buat model pada set pelatihan, Evaluasi pada set pengujian (Latih ulang?), Pisahkan data menjadi set pelatihan dan pengujian
- ☒ b. Pisahkan data menjadi set pelatihan dan pengujian, Buat model pada set pelatihan, Evaluasi pada set pengujian ✓
- ☐ c. Buat model pada set pelatihan, pilih secara acak jumlah instance kelas minoritas yang diinginkan, Evaluasi pada set pengujian
- ☐ d. Evaluasi pada set pengujian, Buat model pada set pelatihan, Pisahkan data menjadi set pelatihan dan pengujian

Question **15**

Correct

Marked out of
5.00

Pernyataan yang tepat terkait data mining and knowledge discovery

Select one:

- ☐ a. lebih heuristik
- ☐ b. lebih fokus pada pengujian hipotesis
- ☐ c. lebih teori
- ☒ d. mengintegrasikan teori dan heuristik ✓

Question **16**

Correct

Marked out of
5.00

Jenis tugas data mining yang melakukan prediksi kelas

Select one:

- ☐ a. Klasterisasi
- ☐ b. Visualisasi
- ☒ c. Klasifikasi ✓
- ☐ d. Asosiasi

Question **17**

Correct

Marked out of
5.00

Pernyataan yang tepat terkait data mining and knowledge discovery

Select one:

- ☐ a. lebih fokus pada pengujian hipotesis
- ☒ b. fokus pada seluruh proses penemuan pengetahuan ✓
- ☐ c. fokus pada peningkatan kinerja agen pembelajaran
- ☐ d. melihat pembelajaran real-time dan robotika

Question **18**

Correct

Marked out of
5.00

Attributes yang berisi informasi seperti item_ID, name, brand, category, supplier, place_made and price, merupakan salah satu contoh jenis database

Select one:

- ☐ a. Social media
- ☒ b. Penjualan ✓
- ☐ c. Penyewaan
- ☐ d. Marketing

Question **19**

Correct

Marked out of
5.00

Metode klasifikasi pada fungsi data mining yang menggunakan nilai estimasi dari probabilitas bersyarat dinamakan metode

Select one:

- ☒ a. Naive Bayessian ✓
- ☐ b. Mean
- ☐ c. Regresi
- ☐ d. Eliminasi

Question **20**

Correct

Marked out of
5.00

Yang bukan Metode clusterisasi

Select one:

- ☐ a. Exclusive vs. overlapping
- ☒ b. Decision Tree ✓
- ☐ c. Deterministic vs. probabilistic
- ☐ d. Hierarchical vs. flat

◀ Video M5 : Introduction Data Mining

Jump to...

Kuis M5 : Introduction Data Mining
(KHUSUS KELAS MALAM)) ▶

Started on	Monday, 12 April 2021, 12:46 PM
State	Finished
Completed on	Monday, 12 April 2021, 12:53 PM
Time taken	7 mins 37 secs

Question **1**

Correct

Marked out of 5.00

Model Pembelajaran yang mempelajari bagaimana sistem dapat menyimpulkan suatu fungsi untuk mendeskripsikan struktur tersembunyi dari data yang tidak berlabel. Sistem ini tidak memprediksi keluaran yang benar, tetapi mengeksplorasi data dan dapat menarik kesimpulan dari kumpulan data untuk mendeskripsikan struktur tersembunyi dari data yang tidak berlabel.

Select one:

- ☐ a. Reinforcement Learning
- ☐ b. Supervised Learning
- ☐ c. Semi Supervised Learning
- ☒ d. Unsupervised Learning ✓

Question **2**

Correct

Marked out of 5.00

Tahapan yang tepat ketika melakukan testing

Select one:

- ☐ a. Image Features - Training With Labeled Data – Prediction
- ☒ b. Image Features - Learned Model - Prediction ✓
- ☐ c. Image Features - Training With Labeled Data - Learned Model
- ☐ d. Image Features - Training With Unlabel Data -Learned Model - Prediction

Question **3**

Correct

Marked out of 5.00

Jika variabel keluaran adalah berupa kategori misalnya “apel”, “jeruk” termasuk pada machine learning bertipe

Select one:

- ☐ a. Unsupervised
- ☐ b. Supervised -Regresion
- ☐ c. Reinforcement
- ☒ d. Supervised – Classification ✓

Question **4**

Correct

Marked out of 5.00

```
If tear production rate = reduced
then recommendation = none
Otherwise, if age = young and astigmatic = no
then recommendation = soft
```

Gambar pada representasi di atas adalah Pola-pola struktur yang ada pada suatu koleksi data (knowledge) dalam bentuk

Select one:

- ☐ a. Pohon
- ☒ b. Aturan ✓
- ☐ c. Tabel
- ☐ d. Instance



Question **5**

Correct

Marked out of
5.00

Saat Anda ingin mengungkap pengelompokan yang melekat dalam data, seperti mengelompokkan hewan berdasarkan beberapa karakteristik / fitur, mis. jumlah kaki maka model pembelajaran yang sesuai

Select one:

- ☐ a. Reinforcement Learning
- ☐ b. Supervised Learning - Regression
- ☐ c. Supervised Learning - Classification
- ☒ d. Unsupervised Learning - Clustering ✓

Question **6**

Correct

Marked out of
5.00

Pola-pola struktur yang ada pada suatu koleksi data (knowledge) dapat direpresentasi dalam bentuk if then rules :

Select one:

- ☐ a. Tabel
- ☐ b. Instance
- ☐ c. Pohon
- ☒ d. Aturan ✓

Question **7**

Correct

Marked out of
5.00

Model pembelajaran yang menggunakan data berlabel dan tidak berlabel untuk pelatihan.

Select one:

- ☒ a. Semi Supervised Learning ✓
- ☐ b. Unsupervised Learning
- ☐ c. Supervised Learning
- ☐ d. Reinforcement Learning

Question **8**

Correct

Marked out of
5.00

Pembelajaran yang bertujuan memprediksi output dengan akurat untuk data baru dan output ti bernilai kontinu (riil), contoh untuk Prediksi kinerja CPU

Select one:

- ☐ a. Klasifikasi
- ☐ b. Klasterisasi
- ☐ c. Generalisasi
- ☒ d. Regressi ✓

Question **9**

Correct

Marked out of
5.00

Tahapan yang tepat ketika melakukan training

Select one:

- ☐ a. Image Features - Training With Labeled Data – Prediction
- ☐ b. Image Features - Training With UnLabeled Data - Prediction
- ☒ c. Image Features - Training With Labeled Data - Learned Model ✓
- ☐ d. Image Features - Learned Model - Prediction

Question **10**
Correct
Marked out of 5.00

Salah satu teknik yang digunakan untuk mengontrol fenomena over-fitting adalah dengan cara menambah finalti ke fungsi error.

Select one:

- ☒ a. regularization ✓
- ☐ b. generalization
- ☐ c. model selection
- ☐ d. prediction

Question **11**
Correct
Marked out of 5.00

Tujuan pembelajaran unsupervised adalah membangun model yang dapat menemukan komponen / variabel / fitur tersembunyi pada data pelatihan, yang dapat digunakan untuk hal berikut, kecuali

Select one:

- ☐ a. Reduksi dimensi (dimension reduction)
- ☒ b. Pengklasifikasian (classification) ✓
- ☐ c. Pengelompokan (clustering)
- ☐ d. Rekomendasi

Question **12**
Correct
Marked out of 5.00

Age	Spectacle prescription	Astigmatism	Tear production rate	Recommended lenses
Young	Myope	No	Reduced	None
Young	Hypermetrope	No	Normal	Soft
Pre-presbyopic	Hypermetrope	No	Reduced	None
Presbyopic	Myope	Yes	Normal	Hard
...

Gambar pada representasi di atas adalah Pola-pola struktur yang ada pada suatu koleksi data (knowledge) dalam bentuk

Select one:

- ☐ a. Aturan
- ☐ b. Instance
- ☐ c. Pohon
- ☒ d. Tabel ✓

Question **13**
Correct
Marked out of 5.00

Model ini belajar dari kumpulan data berlabel dan kemudian digunakan untuk memprediksi peristiwa di masa depan.

Select one:

- ☐ a. Semi Supervised Learning
- ☐ b. Reinfocement Learning
- ☐ c. Unsupervised Learning
- ☒ d. Supervised Learning ✓

Question **14**
Correct
Marked out of 5.00

Pembelajaran yang menghasilkan Nilai output ti bernilai diskrit (kelas)

Select one:

- ☐ a. Generalisasi
- ☐ b. Klasterisasi
- ☐ c. Regressi
- ☒ d. Klasifikasi ✓

Question **15**

Correct

Marked out of 5.00

Yang bukan tujuan supervised learning untuk membangun model yang dapat menghasilkan output yang benar untuk suatu data input dapat digunakan untuk

Select one:

- ☐ a. Pengklasifikasian (classification)
- ☐ b. Ranking
- ☒ c. Klasterisasi (clustering) ✓
- ☐ d. Regresi (regression)

Question **16**

Correct

Marked out of 5.00

Yang tidak digunakan pada Model Supervised Learning

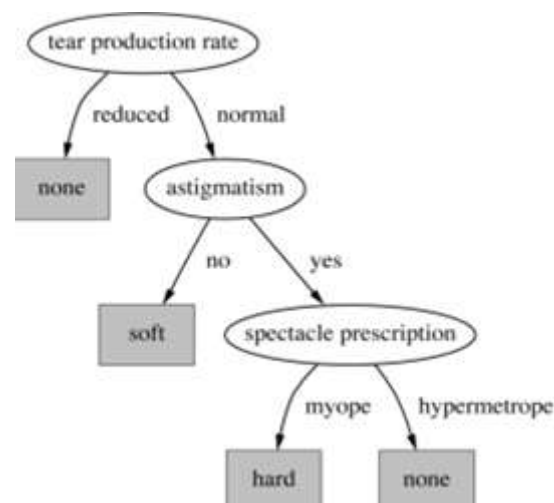
Select one:

- ☐ a. Neural Networks
- ☐ b. Linear Model
- ☒ c. Latent Semantic Analysis ✓
- ☐ d. Support Vector Machine

Question **17**

Correct

Marked out of 5.00



Gambar pada representasi di atas adalah Pola-pola struktur yang ada pada suatu koleksi data (knowledge) dalam bentuk

Select one:

- ☐ a. Aturan
- ☒ b. Pohon ✓
- ☐ c. Tabel
- ☐ d. Instance

Question **18**

Correct

Marked out of 5.00

Kumpulan data yg sudah dipastikan kebenarannya, dr sisi label, data awal, komponen2 data yg akan dijadikan kunci,

Select one:

- ☐ a. Feature Extraction
- ☐ b. Instance
- ☐ c. Feature/ atribut
- ☒ d. Training Set ✓

Question **19**

Correct

Marked out of 5.00

Sebuah proses yg mempersiapkan data sehingga data tsb bisa kita extract cirinya menjadi suatu vektor

Select one:

- ☐ a. Training Set
- ☐ b. Instance
- ☒ c. Feature Extraction ✓
- ☐ d. Feature/ atribut


Question **20**

Correct

Marked out of 5.00

Ciri-ciri dari masing-masing sample data misalnya Smartphone. ciri: memiliki ...

Select one:

- ☐ a. Training Set
- ☒ b. Feature/ atribut 
- ☐ c. Instance
- ☐ d. Feature Extraction

◀ Materi M6 : Introduction to Machine Learning-Supervised

Jump to...

Kuis M6 : Introduction to Machine Learning (Supervised) (KHUSUS KELAS MALAM)) ▶



Started on	Monday, 19 April 2021, 11:40 AM	Online
State	Finished	
Completed on	Monday, 19 April 2021, 11:45 AM	
Time taken	4 mins 38 secs	

Question **1**

Correct

Marked out of 5.00

Pengelompokan yang sebagai "pendekatan bottom-up". Di mana titik datanya diisolasi sebagai pengelompokan terpisah pada awalnya, dan kemudian digabungkan bersama secara iteratif atas dasar kesamaan hingga satu cluster tercapai.

- Select one:
- ☐ a. Pendekatan Overlapping
 - ☒ b. Pendekatan Aglomeratif ✓
 - ☐ c. Pendekatan Divisive
 - ☐ d. Pendekatan Probabilistic

Question **2**

Correct

Marked out of 5.00

Teknik yang digunakan ketika jumlah fitur, atau dimensi, dalam kumpulan data tertentu terlalu tinggi. Teknik ini mengurangi jumlah input data ke ukuran yang dapat dikelola sambil juga menjaga integritas set data sebanyak mungkin, biasanya digunakan dalam tahap data praproses

- Select one:
- ☒ a. Dimensionality reduction ✓
 - ☐ b. Association
 - ☐ c. clustering
 - ☐ d. Regression

Question **3**

Incorrect

Marked out of 5.00

Algoritme yang menggunakan pohon hash untuk menghitung kumpulan item, menavigasi kumpulan data dengan cara yang pertama. Algoritma ini telah dipopulerkan melalui analisis keranjang pasar, yang mengarah ke mesin rekomendasi yang berbeda untuk platform musik dan pengecer online.Misalnya, jika saya memutar radio Black Sabbath di Spotify, dimulai dengan lagu mereka "Anggrek", salah satu lagu lain di saluran ini kemungkinan besar adalah lagu Led Zeppelin, seperti "Over the Hills dan Far Away." Ini didasarkan pada kebiasaan mendengarkan saya sebelumnya serta kebiasaan orang lain.

- Select one:
- ☐ a. Eclat
 - ☒ b. Apriori ✗
 - ☐ c. Apriori
 - ☐ d. FP-Growth

Question **4**

Correct

Marked out of
5.00

Algoritme yang menggunakan pohon hash untuk menghitung kumpulan item, menavigasi kumpulan data dengan cara yang pertama. Algoritma ini telah dipopulerkan melalui analisis keranjang pasar, yang mengarah ke mesin rekomendasi yang berbeda untuk platform musik dan pengecer online. Misalnya, jika saya memutar radio Black Sabbath di Spotify, dimulai dengan lagu mereka "Anggrek", salah satu lagu lain di saluran ini kemungkinan besar adalah lagu Led Zeppelin, seperti "Over the Hills dan Far Away." Ini didasarkan pada kebiasaan mendengarkan saya sebelumnya serta kebiasaan orang lain. Ini termasuk task

Select one:

- ☐ a. Dimensionality reduction
- ☐ b. Regression
- ☐ c. clustering
- ☒ d. Association ✓

Question **5**

Correct

Marked out of
5.00

Algoritme yang berusaha menemukan pola tersembunyi atau pengelompokan data tanpa perlu campur tangan manusia.

Select one:

- ☐ a. Supervised Learning
- ☐ b. Deep Learning
- ☒ c. Unsupervised Learning ✓
- ☐ d. Semi Supervised Learning

Question **6**

Correct

Marked out of
5.00

Metode berbasis aturan untuk menemukan hubungan antara variabel dalam kumpulan data tertentu. Metode ini sering digunakan untuk analisis keranjang pasar, memungkinkan perusahaan untuk lebih memahami hubungan antara produk yang berbeda. Memahami kebiasaan konsumsi pelanggan memungkinkan bisnis untuk mengembangkan strategi penjualan silang dan mesin rekomendasi yang lebih baik.

Select one:

- ☒ a. Association ✓
- ☐ b. clustering
- ☐ c. Regression
- ☐ d. Dimensionality reduction

Question **7**

Correct

Marked out of
5.00

Jenis algoritme pengurangan dimensi yang digunakan untuk mengurangi redundansi dan mengompresi kumpulan data melalui ekstraksi fitur. Metode ini menggunakan transformasi linier untuk membuat representasi data baru, menghasilkan sekumpulan "komponen utama". Komponen utama pertama adalah arah yang memaksimalkan varians dari dataset. Sementara komponen utama kedua juga menemukan varians maksimum dalam data, itu sama sekali tidak berkorelasi dengan komponen utama pertama, menghasilkan arah yang tegak lurus, atau ortogonal, ke komponen pertama. Proses ini berulang berdasarkan jumlah dimensi, dimana komponen utama berikutnya adalah arah ortogonal ke komponen sebelumnya dengan varians terbanyak.

Select one:

- ☐ a. Singular value decomposition
- ☐ b. FP-Growth
- ☐ c. Autoencoders
- ☒ d. Principal component analysis ✓

Question **8**

Correct

Marked out of
5.00

Pendekatan reduksi dimensi lain yang memfaktorkan matriks, A , menjadi tiga matriks peringkat rendah. SVD dilambangkan dengan rumus, $A = USVT$, di mana U dan V adalah matriks ortogonal. S adalah matriks diagonal, dan nilai S dianggap sebagai nilai tunggal matriks A . Principal component analysis

Select one:

- ☐ a. Singular value decomposition
- ☐ b. FP-Growth
- ☒ c. Autoencoders ✓

Question **9**

Correct

Marked out of
5.00

Pengelompokan yang mengambil pendekatan "top-down". Dalam hal ini, satu cluster data dibagi berdasarkan perbedaan antara titik data. Pengelompokan dengan cara ini tidak umum digunakan, tetapi masih perlu diperhatikan dalam konteks pengelompokan hierarkis. Proses pengelompokan ini biasanya divisualisasikan menggunakan dendrogram, diagram mirip pohon yang mendokumentasikan penggabungan atau pemisahan titik data pada setiap iterasi.

Select one:

- ☐ a. Pendekatan Aglomeratif
- ☐ b. Pendekatan Probabilistic
- ☐ c. Pendekatan Overlapping
- ☒ d. Pendekatan Divisive ✓

Question **10**

Correct

Marked out of
5.00

Algoritma yang paling banyak digunakan untuk task Association rules

Select one:

- ☐ a. FP-Growth
- ☐ b. Apriori
- ☒ c. Apriori ✓
- ☐ d. Eclat

Question **11**

Correct

Marked out of
5.00

Salah satu bentuk pengelompokan yang memungkinkan titik data menjadi bagian dari beberapa kluster dengan derajat keanggotaan terpisah. Ini disebut juga Pengelompokan "lembut" atau fuzzy k-means

Select one:

- ☐ a. Memecah belah (Divisive)
- ☐ b. Clustering Eksklusif (specifically exclusive)
- ☒ c. Tumpang tindih (overlapping) ✓
- ☐ d. Hierarkis (hierarchical)

Question **12**

Correct

Marked out of
5.00

Yang bukan tugas dari pembelajaran tanpa pengawasan

Select one:

- ☐ a. Association
- ☐ b. Dimensionality reduction
- ☒ c. Regression ✓
- ☐ d. clustering

Question **13**

Correct

Marked out of 5.00

Salah satu bentuk pengelompokan yang menetapkan bahwa suatu titik data hanya dapat ada dalam satu cluster. Ini juga bisa disebut sebagai pengelompokan "keras". Algoritma yang digunakan salah satunya adalah K-means clustering

Select one:

- ☐ a. Memecah belah (Divisive)
- ☒ b. Clustering Eksklusif (specifically exclusive) ✓
- ☐ c. Hierarkis (hierarchical)
- ☐ d. Tumpang tindih (overlapping)

Question **14**

Correct

Marked out of 5.00

Pendekatan Unsupervised Learning adalah

Select one:

- ☐ a. Mempelajari bagaimana cara kerja pembelajaran dengan pengawasan
- ☐ b. Mempelajari bagaimana cara kerja pembelajaran tanpa pengawasan dan dengan pengawasan
- ☒ c. Mempelajari bagaimana cara kerja pembelajaran tanpa pengawasan ✓
- ☐ d. Mempelajari bagaimana cara kerja pembelajaran yang mendalam

Question **15**

Correct

Marked out of 5.00

Teknik tanpa pengawasan yang membantu menyelesaikan estimasi kepadatan atau masalah pengelompokan "lunak". Dalam hal ini, titik data dikelompokkan berdasarkan kemungkinannya termasuk dalam distribusi tertentu. Teknik yang paling umum digunakan adalah Gaussian Mixture Model (GMM)

Select one:

- ☒ a. Pendekatan Probabilistic ✓
- ☐ b. Pendekatan Aglomeratif
- ☐ c. Pendekatan Overlapping
- ☐ d. Pendekatan Divisive

Question **16**

Correct

Marked out of 5.00

Pembelajaran mesin tanpa pengawasan, menggunakan algoritme pembelajaran mesin untuk menganalisis dan mengelompokkan set data tak berlabel

Select one:

- ☐ a. Supervised Learning
- ☐ b. Deep Learning
- ☒ c. Unsupervised Learning ✓
- ☐ d. Semi Supervised Learning

Question **17**

Correct

Marked out of 5.00

Kemampuannya untuk menemukan persamaan dan perbedaan informasi menjadikannya unsupervised learning sebagai solusi ideal tugas tugas di bawah ini, kecuali

Select one:

- ☐ a. strategi penjualan silang
- ☒ b. Klasifikasi buah jeruk dan buah apel ✓
- ☐ c. Analisis data eksplorasi
- ☐ d. Segmentasi pelanggan

Question **18**

Correct

Marked out of
5.00

Algoritme yang bukan digunakan untuk pengelompokan (Clustering)

Select one:

- ☐ a. Clustering eksklusif (specifically exclusive)
- ☐ b. Tumpang tindih (overlapping)
- ☐ c. Hierarkis (hierarchical)
- ☒ d. Auto Encoder ✓

Question **19**

Correct

Marked out of
5.00

Teknik yang mengelompokkan data yang tidak berlabel berdasarkan persamaan atau perbedaannya. Algoritma ini digunakan untuk memproses objek data mentah dan tidak terklasifikasi menjadi grup yang diwakili oleh struktur atau pola dalam informasi.

Select one:

- ☒ a. Clustering ✓
- ☐ b. Dimensionality reduction
- ☐ c. Regression
- ☐ d. Association

Question **20**

Correct

Marked out of
5.00

Teknik yang memanfaatkan jaringan neural untuk mengompresi data dan kemudian membuat ulang representasi baru dari input data asli. Terdapat lapisan tersembunyi secara khusus bertindak sebagai penghambat untuk memampatkan lapisan masukan sebelum merekonstruksi di dalam lapisan keluaran.

Select one:

- ☐ a. Singular value decomposition
- ☐ b. Principal component analysis
- ☒ c. Autoencoders ✓
- ☐ d. FP-Growth

[◀ Materi M7 : Unsupervised Learning](#)

Jump to...

[Kuis M7 : Unsupervised Learning
\(KHUSUS KELAS MALAM\) ▶](#)

[Dashboard](#) / [My courses](#) / [ATA 2020/2021 | 2-FIKTI | Komputasi Big Data | TEAM TEACHING](#)
/ [Pertemuan 8 : Introduction to Machine Learning \(Reinforcement\)](#) / [Kuis M8 : Reinforcement](#)

Started on	Monday, 26 April 2021, 12:01 PM
State	Finished
Completed on	Monday, 26 April 2021, 12:05 PM
Time taken	4 mins 43 secs

Question **1**
Correct
Marked out of 5.00

Reinforcement learning tepat digunakan untuk menyelesaikan permasalahan utama dalam:

- Select one:
- ☐ a. Segmentasi pelanggan
 - ☒ b. Permainan catur ✓
 - ☐ c. Peramalan cuaca
 - ☐ d. Pengenalan objek

Question **2**
Correct
Marked out of 5.00

Berikut ini merupakan contoh permasalahan yang dapat diselesaikan dengan reinforcement learning:

- Select one:
- ☐ a. Sistem percakapan
 - ☐ b. Self-driving car
 - ☒ c. Semua benar ✓
 - ☐ d. Bermain Game

Question **3**
Correct
Marked out of 5.00

Markov Decision Process dapat didefinisikan kedalam n-tuple, dimana bernilai:

- Select one:
- ☐ a. 4
 - ☐ b. 3
 - ☐ c. 6
 - ☒ d. 5 ✓

Question **4**
Correct
Marked out of 5.00

Efek dari suatu aksi yang dipilih dari suatu state yang bergantung hanya pada state tersebut saja bukan pada state-state sebelumnya dikenal sebagai sifat:

- Select one:
- ☒ a. Markov ✓
 - ☐ b. Bellman
 - ☐ c. Transisi
 - ☐ d. Stokastik



Question **5**

Correct

Marked out of
5.00

Pernyataan yang benar terkait reinforcement learning:

Select one:

- ☐ a. Semua benar
- ☐ b. Untuk aksi yang bersifat stokastik harus memenuhi sifat Markov
- ☐ c. Input yang diterima oleh agen berupa state, action dan reward
- ☒ d. Aksi yang dapat dipilih pada setiap state tidak unik ✓

Question **6**

Correct

Marked out of
5.00

Solusi dari Markov Decision Process disebut sebagai:

Select one:

- ☒ a. Policy ✓
- ☐ b. Reward
- ☐ c. Model
- ☐ d. Action

Question **7**

Correct

Marked out of
5.00

Fungsi reward dalam Markov Decision Process dapat bernilai bilangan:

Select one:

- ☒ a. Semua benar ✓
- ☐ b. Riil
- ☐ c. Bulat
- ☐ d. Asli

Question **8**

Correct

Marked out of
5.00

Sifat Markov dalam Markov Decision Process diperlukan langsung dalam menentukan:

Select one:

- ☐ a. Fungsi reward
- ☐ b. Aksi yang dipilih
- ☐ c. Fungsi nilai state-action
- ☒ d. Matriks probabilitas transisi ✓

Question **9**

Correct

Marked out of
5.00

Reinforcement learning dapat diterapkan dalam permasalahan:

Select one:

- ☐ a. Penentuan jalur terpendek
- ☐ b. Supervised learning
- ☒ c. Semua benar ✓
- ☐ d. Navigasi robot

Question **10**

Correct

Marked out of
5.00

Pemilihan suatu aksi pada suatu state tertentu didalam Markov Decision Process ditentukan oleh:

Select one:

- ☐ a. Model
- ☐ b. Reward
- ☐ c. Value
- ☒ d. Policy ✓

Question **11**

Correct

Marked out of 5.00

Berikut ini merupakan pernyataan yang salah terkait Markov Decision Process:

Select one:

- ☐ a. Fungsi nilai state-action dapat dinyatakan dalam persamaan ekspektasi Bellman
- ☐ b. Elemen matriks probabilitas transisi bernilai antara 0 dan 1
- ☐ c. Fungsi reward dapat dinyatakan sebagai suatu fungsi ekspektasi
- ☒ d. Total policy pada suatu state untuk semua reward adalah 1 ✓

Question **12**

Correct

Marked out of 5.00

Reinforcement learning merupakan bagian dari:

Select one:

- ☐ a. Dynamic Programming
- ☐ b. Markov Decision Process
- ☐ c. Semua benar
- ☒ d. Machine learning ✓

Question **13**

Correct

Marked out of 5.00

Berikut ini merupakan beberapa algoritma reinforcement learning, kecuali:

Select one:

- ☒ a. Hill-Climbing Bagged Ensemble Selection (HCES-Bag) ✓
- ☐ b. State-Action-Reward-State-Action (SARSA)
- ☐ c. Deep Deterministic Policy Gradient (DDPG)
- ☐ d. Q-learning

Question **14**

Correct

Marked out of 5.00

Efek dari suatu pilihan aksi pada suatu state ditentukan oleh:

Select one:

- ☐ a. Semua salah
- ☒ b. Model transisi ✓
- ☐ c. Policy
- ☐ d. Fungsi reward

Question **15**

Correct

Marked out of 5.00

Reward dalam permasalahan shortest path yang dimodelkan dengan Markov Decision Process ditandai dengan:

Select one:

- ☐ a. Total nilai bobot simpul dan ruas
- ☐ b. Nilai bobot pada simpul
- ☒ c. Nilai bobot pada ruas ✓
- ☐ d. Semua salah

Question **16**

Correct

Marked out of 5.00

Reinforcement learning dapat diilustrasikan sebagai interaksi antara:

Select one:

- ☐ a. Semua benar
- ☐ b. Reward dengan state berikutnya
- ☒ c. Agen dengan lingkungannya ✓
- ☐ d. State dengan aksi yang dipilih

Question **17**

Correct

Marked out of
5.00

Berikut ini merupakan pernyataan yang kurang tepat terkait reinforcement learning (RL):

Select one:

- ☐ a. RL dapat digunakan pada robotik untuk otomasi industri
- ☒ b. RL yang diselesaikan dengan Markov Decision Process memerlukan data latih dalam menentukan probabilitas transisinya ✓
- ☐ c. RL dapat digunakan dalam menciptakan sistem pelatihan untuk siswa yang bersifat kustom
- ☐ d. RL memerlukan lingkungan untuk berinteraksi secara repetitif

Question **18**

Correct

Marked out of
5.00

Pada Markov Decision Process, pemilihan ruas jalan dalam permasalahan shortest path merupakan bagian dari:

Select one:

- ☐ a. Value
- ☐ b. Reward
- ☐ c. State
- ☒ d. Action ✓

Question **19**

Correct

Marked out of
5.00

Dalam memilih aksi pada suatu state perlu diperhatikan hal-hal berikut ini:

Select one:

- ☐ a. Nilai state-action pada suatu waktu berlaku untuk sembarang policy
- ☒ b. Memilih aksi yang memberikan state dengan nilai tertinggi ✓
- ☐ c. Semua benar
- ☐ d. Memaksimumkan reward pada masa mendatang

Question **20**

Correct

Marked out of
5.00

Sekumpulan token yang mewakili setiap kemungkinan keadaan dari agen disebut sebagai:

Select one:

- ☒ a. State ✓
- ☐ b. Action
- ☐ c. Semua salah
- ☐ d. Reward

[◀ Materi M8 : Reinforcement Learning](#)[Kuis M8 : Reinforcement \(KHUSUS KELAS MALAM\) ▶](#)

M10 Team Teaching

1. Salah satu library data science yang digunakan untuk pemrosesan sinyal, gambar, dan rumus statistik adalah.. **Semua Benar**
2. Beragamnya jenis data di dalam Big Data yang meliputi data terstruktur, semi terstruktur dan tidak terstruktur, merupakan dimensi Big Data dalam hal: **Variety**
3. Pilihan yang dapat menciptakan hasil yang paling optimal merupakan pendekatan analitik dari.. **Preskriptif**
4. Sumber data yang digunakan dalam aplikasi Big Data dapat berupa: **Semua Benar**
5. Keuntungan menggunakan Mobile Positioning data (MPD) adalah... **Semua Benar**
6. Metode untuk memprediksi kelas instance dari instance yang diberi label sebelumnya... **Klasifikasi**
7. Visualisasi pada hasil analisis Big Data dapat disajikan dalam bentuk.. **Semua Benar**
8. Data yang dikumpulkan secara aktif dari sensor, misalnya melalui penginderaan jauh dan gambar satelit merupakan sumber big data dalam hal.... **Sensing data**
9. Contoh solusi Big Data, kecuali: **Semua Benar**
10. Contoh Aplikasi Machine Learning dan Analitik Prediktif: **Semua Benar**
11. Semakin banyak data hingga dapat melebihi terabyte merupakan penjelasan dari dimensi Big Data dalam hal: **Volume**
12. Karakteristik Big Data yang cepat sekali berubah baik dari sisi variabel maupun tipe data dinamakan... **Velocity**
13. Data Warehouse dibuat dengan menerapkan serangkaian langkah pemrosesan terhadap data yang berasal dari beberapa basis data. Berikut ini adalah yang termasuk dalam pemrosesan data tersebut: **Semua Benar**
14. Manakah yang benar dari pernyataan berikut ini: **Deskriptif, perkiraan berdasarkan data yang tersedia**
15. Data Mobile Positioning data (MPD) berupa... **Semua Benar**
16. Dataset dibentuk dari: **Sekumpulan fitur yang memiliki objek yang sama**
17. Tools Bahasa Pemrograman yang dapat digunakan dalam ilmu data science adalah... **Semua Benar**
18. Social Media termasuk dalam sumber data.. **Digital Content**
19. Cara yang dapat dilakukan untuk meningkatkan analisis bigdata genome bioinformatics adalah.. **Semua Benar**
20. Proses menganalisis tulisan untuk menentukan nada emosional penulisnya merupakan salah satu implementasi machine learning dalam... **Principal Component Analysis**

Team Teaching Big Data M9

1. Berikut ini adalah efek dari mengecilnya nilai stride dalam Algoritma Convolutional Neural Network

Informasi yang diperoleh semakin detail

2. Hasil dari proses flatten pada Algoritma Convolutional Neural Network menjadi masukan dalam

Fully-connected layer

3. Pooling layer bertujuan untuk

Mengurangi dimensi feature map

4. Berikut ini merupakan loss function dalam algoritma berbasis Neural Network

Cross entropy

5. Algoritma Deep Learning berbasis

Semua benar (Neural network, Machine learning, Artificial Intelligence)

6. Berikut ini adalah teknik yang dapat digunakan pada pooling layer dalam Algoritma Convolutional Neural Network

Semua benar (Average pooling, Min pooling, Max pooling)

7. Dalam Algoritma Convolutional Neural Network, ReLU seringkali digunakan pada

Hidden layer

8. Sebuah citra berdimensi 5x5 dengan filter 3x3 dan stride sebesar 2 akan menghasilkan feature map berdimensi

4x4

9. Pooling layer pada Algoritma Convolutional Neural Network terdapat pada

Hidden layer

10. Berikut ini merupakan fungsi aktivasi dalam Algoritma Convolutional Neural Network

Semua benar (Sigmoid, Tanh, ReLU)

11. Tahapan feature extraction pada Algoritma Convolutional Neural Network terdiri dari beberapa lapisan (layer) berikut ini, kecuali

Fully-connected layer

12. Untuk memperbanyak feature yang berhasil diekstrak dalam Algoritma Convolutional Neural Network digunakan

Padding

13. Berikut ini merupakan contoh Algoritma Deep Learning

Semua benar

Algoritma Recurrent Neural Network

Algoritma Long Short Term Memory

Algoritma Convolutional Neural Network

14. Algoritma Deep Learning dapat menyelesaikan permasalahan berikut ini

Semua benar (Deteksi objek, Natural Language Processing, Pengenalan suara)

15. Untuk menyelesaikan permasalahan klasifikasi biner menggunakan algoritma berbasis Neural Network lebih tepat digunakan fungsi aktivasi

ReLU

16. Salah satu ukuran dari confusion matrix yang dapat digunakan dalam mengukur tingkat keberhasilan sistem dalam menemukan kembali sebuah informasi adalah

Recall

17. Untuk mencegah terjadinya overfitting dalam Algoritma Convolutional Neural Network dilakukan

Dropout regularization

18. Dimensi dari feature map dalam Algoritma Convolutional Neural Network bergantung pada

Semua benar

Stride

Dimensi filter dan input

Padding

19. Parameter dalam Algoritma Convolutional Neural Network yang menentukan berapa jumlah pergeseran filter disebut sebagai

Stride

20. Berikut ini merupakan library Python untuk Algoritma Deep Learning

Semua benar (Tensorflow, Keras, PyTorch)

Team Teaching Big Data M11

1. Berikut ini yang bukan merupakan karakteristik dari Apache Spark adalah

Tidak dilengkapi dengan sistem file bawaan

2. Berapa banyak Spark Context yang dapat aktif per Java Virtual Machine

Hanya 1

3. Berikut ini yang bukan merupakan fungsi dari Spark Context adalah

Untuk mengatur konfigurasi

4. Toleransi kesalahan pada Resilient Distributed Data (RDD) dicapai menggunakan

Kemampuan alami dari RDD untuk tidak dapat diubah

5. Berikut ini adalah beberapa manfaat penggunaan Spark, kecuali

Otomatis dalam mengoptimalkan kode program

6. Berikut ini yang bukan operasi transformasi pada Resilient Distributed Data adalah

Flatmap

7. Apache Spark pertama kali diluncurkan open-source pada tahun

2010

8. Berikut ini yang bukan operasi action pada Resilient Distributed Data adalah

Top()

9. Berikut ini adalah pernyataan yang salah mengenai Resilient Distributed Data (RDD), kecuali

RDD merupakan basis data

10. Spark ditulis dalam bahasa pemrograman

Scala

11. Modul untuk memproses data terstruktur dalam Spark adalah

Spark SQL

12. Manajemen memori dalam Spark dikerjakan oleh

Spark Core

13. Cluster manager yang dapat digunakan pada Spark adalah

Semua benar (Hadoop YARN, Mesos cluster, Standalone scheduler)

14. Spark lebih cepat dibandingkan dengan MapReduce karena

Direct Acyclic Graph mengeksekusi mesin dan komputasi didalam memori

15. Api Machine Learning yang utama untuk Spark saat ini adalah API berbasis

DataFrame

16. Apache Spark berpotensi ... lebih cepat untuk menjalankan program pemrosesan batch didalam memori jika dibandingkan dengan MapReduce

100 kali

17. Komponen berikut ini yang tidak terdapat diatas Spark Core adalah

Semua salah (Spark SQL, Spark MLib, Spark RDD)

18. Spark SQL menerjemahkan perintah kedalam bentuk kode yang diproses oleh

Semua salah (Cluster manager, Driver nodes, Executor nodes)

19. Pernyataan berikut ini yang salah mengenai Spark dan Hadoop adalah

Keduanya merupakan mesin komputasi cluster

20. Spark dapat menjalankan ... tugas pada setiap partisi

Satu

Team Teaching Big Data M12

1. Algoritma untuk menentukan jumlah segitiga yang melewati setiap simpul dalam graf

Triangle Count

2. Urutan edge yang memungkinkan Anda berpindah dari vertex A ke vertex B disebut

Path

3. Algoritma untuk mengukur seberapa pentingnya halaman situs web

Page Rank

4. Untuk menampilkan pola di balik hubungan antara entitas data

Visualisasi Grafik

5. Manakah urutan yang benar untuk menjelaskan cara kerja Spark Streaming:

1. Input data stream diterima oleh Spark Streaming,
2. Batch input data / RDD tersebut diproses oleh Spark Engine menggunakan operasi RDD.,
3. Hasilnya adalah batch data yang telah diproses didorong keluar ke sistem eksternal,
4. Oleh Spark Streaming data dibagi menjadi kumpulan / batch input data (diperlakukan sebagai RDD).

1 - 4 - 2 - 3

6. Fitur Spark Streaming

Pemulihan cepat dari kegagalan dan kekeliruan

7. Langkah pipa pemrosesan data grafik :

1. Pembuatan grafik,
2. Pengolahan pasca,
3. Pra-pemrosesan data (yang mencakup pemuatan, transformasi, dan pemfilteran),
4. Analisis

3 - 1 - 4 - 2

8. Algoritma untuk menemukan komponen terhubung dari suatu graf

Connected Components

9. Contoh database grafik

Neo4j

10. Contoh penggunaan Spark Streaming

Netflix

11. Operasi Jendela membutuhkan parameter

window length

12. Framework yang dapat digunakan untuk memproses data grafik dan menjalankan analisis prediktif pada data

Spark GraphX

13. Contoh use case penggunaan Spark GraphX

Semua Benar (Film baru yang paling direkomendasikan orang-orang 1 minggu terakhir, Orang-orang yang sama-sama mensitasi suatu artikel, Teman toni di facebook yang paling banyak pengikutnya)

14. Pernyataan yang benar tentang Graph

Semua benar (Graph terdiri dari simpul dan sisi, Semua yang ada di Facebook adalah contoh penggunaan struktur data graph, Edge menunjukkan relasi antar simpul)

15. Fitur Spark GraphX

Semua benar (Kecepatan, Pustaka Algoritma, Fleksibilitas)

16. Apa yang menyebabkan Spark Streaming memiliki toleransi kesalahan?

Kumpulan data input direplikasi dalam memori

17. Yang diperlukan untuk menangani data besar dalam bentuk grafik

Semua benar (Analisis Data Grafik, Visualisasi data grafik, Database grafik)

18. Aliran data yang berkelanjutan dalam Spark Streaming disebut

DStream

19. Pernyataan yang benar tentang Apache Spark, kecuali

Abstraksi memori terdistribusi untuk komputasi dalam memori pada kluster besar yang toleran terhadap kesalahan

20. Pernyataan yang benar Spark Streaming

Semua benar (Memproses menggunakan algoritma kompleks dengan fungsi seperti map, reduce, join, window, Dapat menyerap data dari berbagai sumber, Data yang diproses dapat dikirim ke sistem file, database, dan live dashboard)

Team Teaching Big Data M13

1. CUDA yang diciptakan oleh NVIDIA adalah singkatan dari

Compute Unified Device Architecture

2. Tiga tahap alur pemrosesan dalam eksekusi program CUDA

Langkah ke-1 Salin input data dari CPU memory ke GPU memory

Langkah-2 Muat program GPU dan eksekusi, caching (simpan) data pada chip untuk kinerja

Langkah-3 Salin hasil dari GPU memory ke CPU memory

3. Arsitektur CUDA menerapkan pendekatan berupa kumpulan streaming multiprocessors (SM) yang

Mengeksekusi sejumlah instruksi yang sama pada beberapa thread pada berbagai wilayah data, sehingga dikenal dengan istilah single instruction, multiple threads (SIMT)

4. Komputasi paralel adalah komputasi yang

Melibatkan banyak unit computer untuk memecahkan masalah yang berbeda-beda untuk masing-masing unit computer, namun dalam waktu yang sama

5. Tujuan utama dari pemrograman paralel adalah untuk meningkatkan performa komputasi. Performa dalam pemrograman paralel diukur dari

Berapa banyak peningkatan kecepatan (speed up) yang diperoleh dalam menggunakan tehnik paralel

6. CUDA C++ adalah perluasan C++ sehingga programmer dapat mendefinisikan fungsi C++ yang ketika dipanggil, akan dieksekusi sebanyak N kali secara paralel dan N CUDA thread yang berbeda. Fungsi ini dikenal dengan istilah

Kernel

7. Kernel pada pemrograman CUDA agar dieksekusi pada device, dideklarasikan menggunakan

`__global__`

8. Sistem terdistribusi (distributed computing) adalah teknik komputasi paralel dimana

Komputer yang digunakan secara bersamaan dalam melakukan komputasi dilakukan oleh computer-komputer terpisah yang terhubung dalam suatu jaringan komputer

9. Perintah kernel `< <N, M> >(...)` menunjukkan device untuk

Mempersiapkan block sebanyak N dengan thread sebanyak M di tiap block

10. Istilah thread pada GPU adalah

Satuan pemrosesan sekuensial yang dikumpulkan dalam satu grup yang disebut sebagai thread block

11. Dua kata istilah penting yang digunakan dalam CUDA, yaitu: host dan device. Istilah host pada CUDA mengacu pada

CPU

12. Dua kata istilah penting yang digunakan dalam CUDA, yaitu: host dan device. Istilah device pada CUDA mengacu

GPU

13. GPU dalam teknologi komputasi adalah singkatan dari

Graphics Processing Unit

14. Pengelolaan memori pada device pada dasarnya menggunakan

CudaMalloc(), cudaMemcpy(), cudaFree()

15. Pernyataan yang benar mengenai thread block pada GPU

Beberapa block dihimpun dalam satu thread, yang kemudian beberapa thread dihimpun dalam satu kesatuan grid

16. CUDA adalah platform komputasi paralel dengan arsitektur komputasi masuk dalam

Heterogenous computing, karena melibatkan CPU dan GPU

17. Pernyataan yang benar mengenai pemrograman paralel pada GPU

Semua pernyataan benar

18. Pernyataan berikut yang benar mengenai CUDA

CUDA adalah suatu platform untuk komputasi paralel, dan memanfaatkan GPU untuk komputasi

19. CPU dari suatu unit computer adalah singkatan dari

Central Processing Unit

20. Perintah kernel < <N, M> >(...) menunjukkan device untuk

Mempersiapkan block sebanyak N dengan thread sebanyak M di tiap block

Team Teaching Big Data M14

1. Bahasa pemrograman Python tergolong cepat dikarenakan

Semua benar

Bahasa pemrograman tingkat tinggi

Dapat diinterpretasikan

Dapat dikodekan secara dinamis

2. Bahasa pemrograman Python tergolong lambat untuk eksekusi

Tugas yang berulang dan tergolong tingkat rendah

3. Modul kompilasi tepat waktu pada Python untuk CPU adalah

Semua benar (Numba, PySpark, CUDA)

4. Berikut ini adalah urutan proses yang dijalankan oleh Numba

Rewrite IR, Lowering, LLVM IR, LLVM/NVVM JIT

5. Banyaknya operasi-operasi kecil yang berulang merupakan salah satu penyebab bahasa pemrograman Python

Tergolong lambat

6. Berikut ini yang pernyataan yang tepat tentang Python adalah

Semua jawaban benar

Python tergolong cepat untuk pengembangan

Python tergolong lambat untuk eksekusi kode program

7. Elemen dalam RDD dikelompokkan kedalam beberapa partisi dan hanya dapat disimpan pada sebuah node yang berbeda merupakan konsep RDD dalam hal

Distributed

8. Numba dapat menangani

Semua jawaban benar (Penyalinan data dari dan ke host jika diperlukan, Alokasi data dari dan ke host jika diperlukan)

9. Komputasi pada GPU cluster dapat menggunakan

Semua jawaban benar (Numba, PySpark)

10. Berikut ini yang salah terkait dengan Numba adalah

Semua salah

11. Bahasa pemrograman Python tergolong cepat digunakan untuk

Semua benar (Mengembangkan kode program, menguji kode program, menulis kode program)

12. Apache Spark bukan merupakan

Semua jawaban salah

13. Spark diimplementasikan dalam

Semua jawaban benar (Scala dan Java)

14. Berikut ini adalah komputasi primitive yang tersedia pada Spark untuk melakukan paralelisme dan meminimumkan komunikasi antar workers

Semua benar (Map, Filter, Reduce)

15. Berikut ini yang benar terkait dengan Numba, kecuali

Open-source

16. Modul kompilasi tepat waktu pada Python untuk GPU adalah

Semua benar (PySpark, Numba, CUDA)

17. API Spark memiliki dukungan yang terbatas pada bahasa pemrograman

R

18. Permasalahan alokasi memori dalam komputasi pada GPU terdapat pada tugas yang berskala

Kecil

19. Type inference pada proses Numba memerlukan

Semua jawaban benar (Functions Arguments, Numba IR)

20. Pemrograman terdistribusi dalam Python adalah

Semua benar (CUDA, Numba, PySpark)