TUGAS

 KONSEP DAN APLIKASI DATA MINING

Oleh:

Agnes Martha Muwa (17.51.0007)

STIMIK PPKIA PRADNYA PARAMITA  
MALANG

2020

1. **Sebutkan Karakteristik Big Data?**
2. Volume yaitu besaran data yang harus dikelolah dan berukuran super besar
3. Velocity berkaitan dengan kecepatan pemrosesan data yang harus mengimbangi pesatnya pertumbuhan jumlah data (update informasi tentang data sangat cepat)
4. Variety yaitu karakteristik sumber data yang sangat beragam,baik itu yang berasal dari basis data yang terstruktu maupun juga dari data-data yang tidak terstruktu.
5. **Sebutkan Sumber Data Big Data?**

Sumber data untuk big data bisa berupa basis data yang terstruktur ataupun data-data yang tidak terstruktur.

* Data Terstruktur = Kelompok data yang memiliki tipe data ,format dan struktur yang telah terdefenisi.Sumber datanya berupa transaksional,OLAP dat,tradisional RDBMS.
* Data Tidak Terstruktu = Kelompok data tekstual dengan format tidak menentu atau tidak memiliki struktur melekat serhingga untuk menjadikannya data terstruktur membutuhkan usaha,tools, dan waktu yang lebih.Data ini dihasilkan oleh aplikasi-aplikasi internet, seperti data URL log,social media,e-mail ,blog,video dan audio

1. **Sebutkan 4 tahapan dalam Big data?**
2. Acquire berhubungan dengan sumber dan cara mendapatkan data.
3. Accessed berhubungan dengan daya akses data; data yang sudah dikumpulkan memerlukan tata kelola,integrasi,storage dan computing agar dapat dikelolah untuk tahap berikutnya.Perangkat untuk pemrosesan(processing tools) menggunakan Hadoop, Nvidia CUDA, Twitter Strom, dan GraphLab.Sedangkan untuk manajemen penyimpanan data (storage tools) menggunakan Neo4J, Titan, dan HDFS (Hadoop File System)
4. Analytic berhubungan dengan informasi yang akan didapatkan, hasil pengelolah data yang telah diproses.Analitik yang dilakukan dapat berupoa descriptive (penggambaran data), diagnostic (mencari sebab akibat berdasarkan data), predictive (memprediksi kejadian dimasa depan) maupun prescriptive analytics (merekomendasikan pilihan dan implikasi dari setiap opsi).Tools untuk tahap analitik menggunakan MLPACK dan Mahout.
5. Aplication,terkait visualisasi dan reporting hasil dari analitik.Tools untuk tahap ini menggunakan RStudio.
6. **Apa manfaat Teknologi Big Data?**

* **Pemanfaatan big data pada sektor bisnis**

1. Mengetahui respon masyarakat terhadap produk-produk yang dikeluarkan melalui analisis sentiment di media social
2. Membantu perusahaan mengambil keputusan secara lebih tepat dan akurat berdasarkan data
3. Membantu meningkata citra perusahaan di mata pelanggaran
4. Perencanaan usaha,dengan mengetahui perilaku pelanggan seperti pada perusahaan telekomunikasi dan perbankan
5. Mengetahui trend pasar dan keinginan konsumen

* **Pemanfaatan big data pada sektor layanan public**

1. Mendapatkan feedback dan respon masyarakat sebagai dasar penyusunan kebijakan dan perbaikan pelayanan public.Feedback tersebut dapat diperoleh dari system informasi layanan pemerintahan maupun dari media social
2. Membuat layanan terpadu dengan segmen khusus sehingga layanan bisa lebih efektif dan efisien
3. Menemukan solusi atas permasalahan yang ada, berdasarkan data.Sebagai contoh : menganalisis informasi cuaca dan informasi pertanian terkait data tingkat kesuburan tanah,pemerintah dapat menetapkan atau menghimbau jenis varietas tanaman yang ditanam oleh petani pada daerah dan waktu tertentu.