



Ujian Tengah Semester

Big Data

Dosen Pengajar Sirojul Munir, S.Si., M.Kom.

MUHAMMAD AZHAR RASYAD

0110217029

TEKNIK INFORMATIKA – 2017

DEPOK

APRIL 2020

1. Apa yang dimaksud dengan Big Data ?

Big Data merupakan kumpulan data yang sangat besar serta kompleks dan tidak dapat dikelola dengan *software* pemrograman database konvensional. Dalam Big Data tidak cukup dengan SQL saja akan tetapi butuh teknologi tambahan yaitu NoSQL serta didalamnya tidak hanya berisi data berstruktur tapi juga data yang tidak berstruktur [1].

2. Mengapa dengan Structured Query Language (Relational Database Management System) tidak cukup menangani Big Data ?

Karena SQL tidak bisa menghandle data yang tidak berstruktur seperti big data dan tidak dirancang untuk sistem yang terdistribusi [2].

3. Sebutkan 5 karakteristik Big Data (5V) ? dan jelaskan selengkapanya

Berikut 5 karakteristik Big Data yaitu [1] :

- *Volume* yaitu ukuran data yang sangat besar dari sisi jumlah hingga mencapai milyaran terra byte atau bahkan trilyunan GB.
- *Velocity* yaitu kecepatan data yang sangat cepat dari sisi kemunculan dan perubahan.
- *Variety* yaitu variasi jenis data yang sangat banyak dari yang *unstructured* dan *multi-structured*.
- *Veracity* yaitu kebenaran dan keakuratan informasi yang tidak mudah dipastikan.
- *Value* yaitu nilai yang dihasilkan sangat besar dari sisi manfaat dalam bentuk uang maupun non-uang.

4. Jika STT-NF ingin menerapkan teknologi Big Data, jelaskan kebutuhan BIG DATA apa yang bisa atau akan ditangani di kampus STT-NF ?

Menurut saya jika STT-NF ingin menerapkan teknologi Big Data maka hal yang dibutuhkan yaitu [3] :

- Menggunakan *Virtualization* supaya dapat menjalankan komputer guest pada komputer host.
- Membuat *Clustering* supaya menjadikan beberapa komputer digabung dalam jaringan menjadi satu komputer.
- Membuat *Distributed & Paralel Computing* supaya menjalankan sebuah proses pada banyak komputer dalam suatu kluster.

Hal di atas merupakan kebutuhan yang akan ditangani kedepannya oleh kampus STT-NF, adapun kebutuhan saat ini yaitu civitas kampus meningkatkan ilmu dan skill Big Data agar dapat menerapkan teknologi Big Data kedepannya [1].

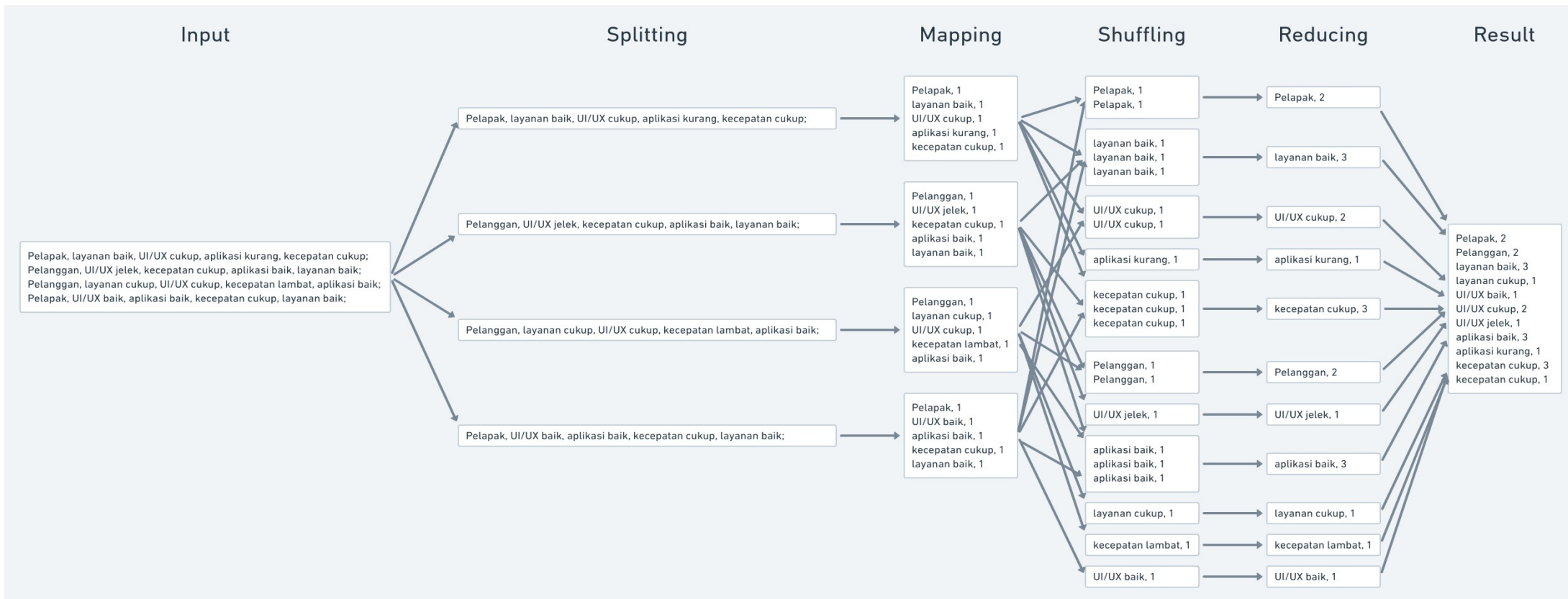
5. Jelaskan apa yang dimaksud dengan MAP REDUCE ?

Gambarkan proses algoritma map reduce jika diterapkan pada sebuah aplikasi e-commerce misal bukalapak.com, untuk analisa data questioner pengguna terhadap aplikasi e-commerce yang diterjemahkan menjadi text data berikut ini:

Pelapak, layanan baik, UI/UX cukup, aplikasi kurang, kecepatan cukup;
Pelanggan, UI/UX jelek, kecepatan cukup, aplikasi baik, layanan baik;
Pelanggan, layanan cukup, UI/UX cukup, kecepatan lambat, aplikasi baik;
Pelapak, UI/UX baik, aplikasi baik, kecepatan cukup, layanan baik;

Map reduce merupakan model algoritma yang digunakan untuk membagi pekerjaan menjadi beberapa bagian pekerjaan dan dikerjakan oleh beberapa komputer [5].

Berikut merupakan analisa dari data questioner diatas terhadap aplikasi e-commerce



6. Jelaskan apa yang dimaksud dengan CAP Theorema, jika diterapkan database MySQL, MongoDB maka kriteira CAP apa yang memenuhi nya, jelaskan!

CAP atau *Consistency Availability Partition Tolerance Theorem* merupakan teori yang digunakan sebagai tool untuk membuat perancangan sistem agar *aware* akan *trade-off* saat merancang *networked shared-data system*, akan tetapi keterpenuhan CAP hanya 2 yang mungkin terpenuhi yaitu CA, CP, atau AP [2].

- Jika diterapkan pada database MySQL maka kriteria CAP yang memenuhi yaitu *Consistency* dan *Availability*, karena database MySQL mempunyai sifat yaitu *client* selalu memiliki tampilan yang sama pada datanya (*Consistency*) dan setiap *client* dapat selalu *read* dan *write* (*Availability*), akan tetapi tidak memiliki *Partition Tolerance* [2].
- Jika diterapkan pada MongoDB maka kriteria CAP yang memenuhi yaitu *Consistency* dan *Partition Tolerance*, karena MongoDB juga mempunyai sifat yaitu *client* selalu memiliki tampilan yang sama pada datanya (*Consistency*) dan sistem berfungsi dengan baik di seluruh partisi pada jaringan fisik (*Partition Tolerance*), akan tetapi tidak memiliki *Availability* [2].

7. Apa yang dimaksud dengan sentiment analisis ? dan apa kegunaannya

Sentiment analysis merupakan proses analisis untuk menentukan suatu sentimen ke dalam bentuk klasifikasi pada orientasi yang terdiri dari positif, netral, atau negatif. Kegunaan dari *sentiment analysis* ini supaya dokumen atau kalimat dapat diklasifikasikan ke dalam 3 orientasi tersebut apakah mengandung hal positif, netral, atau negatif [4].

8. Perhatikan data hasil feedback pelanggan sebuah restoran cepat saji berikut ini

- Q1: makanan semua enak, terutama sop iganya lezat, namun pelayannya sepertinya kurang senyum.
- Q2: makan di resto ini pokoknya makyus, rasanya juara, tapi kok ngak ada wifi nya !
- Q3: baru pertama kali makan di sini, sop iganya enak banget, saya akan rekomendasikan ke temen2 saya deh, tapi tidak ada wifinya kayaknya ngak asik

Rekonstruksi data diatas menjadi sebuah sentimen analisis ? apakah sentimen analisis untuk restoran diatas ? POSITIF atau NEGATIF, berapa skornya !!

Feedback	Kalimat	Klasifikasi	Skor	Total Skor
Q1	makanan semua enak,	Positif	1	1
	terutama sop iganya lezat,	Positif	1	
	namun pelayannya sepertinya kurang senyum.	Negatif	-1	
Q2	makan di resto ini pokoknya makyus,	Positif	1	1
	rasanya juara,	Positif	1	
	tapi kok gak ada wifi nya !	Negatif	-1	
Q3	baru pertama kali makan di sini,	Netral	0	1
	sop iganya enak banget,	Positif	1	
	saya akan rekomendasikan ke temen2 saya deh,	Positif	1	
	tapi tidak ada wifinya kayaknya gak asik	Negatif	-1	
Total Sentimen				3
Overall Sentimen				Positif

Daftar Pustaka

- [1] <https://github.com/mazharrasyad/STTNF/blob/master/Semester%206/Big%20Data/Materi/1.%20Pengantar%20Big%20Data.pdf>
- [2] <https://github.com/mazharrasyad/STTNF/blob/master/Semester%206/Big%20Data/Materi/10.%20NoSQL.pdf>
- [3] <https://github.com/mazharrasyad/STTNF/blob/master/Semester%206/Big%20Data/Materi/3.%20Big%20Data%20Infrastructure.pdf>
- [4] <https://github.com/mazharrasyad/STTNF/blob/master/Semester%206/Big%20Data/Materi/4.%20Big%20Data%20Processing.pdf>
- [5] <https://github.com/mazharrasyad/STTNF/blob/master/Semester%206/Big%20Data/Materi/8.%20Map%20Reduce.pdf>