

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Jumlah data telah meningkat pesat dalam satu dekade terakhir. Tetapi masih minimnya penggunaan dan pemanfaatan mengenai metode pengolahan data selain SQL di Indonesia, dengan adanya perkembangan teknologi yang sangat cepat diperlukannya metode untuk mengatasi masalah dalam menangani pengolahan data yang berukuran sangat besar atau biasa juga disebut dengan *big data*. *big data* dapat ditemukan didalam instansi pemerintah, perusahaan swasta, bahkan di sekolah. Salah satunya adalah sekolah Smk Telkom Jakarta dimana sudah memiliki alumni lebih dari 1000 siswa, dan data alumni ini bisa dimanfaatkan seperti mencari jumlah siswa pada masing masing jurusan dengan perbandingan tahun 2017 dan 2018.[1][2]

SMK Telkom Jakarta merupakan sekolah swasta yang sangat populer dikawasan Jakarta Barat dan sekitarnya, sekolah ini sudah didirikan dari tahun 1992 dengan empat jurusan utama yaitu RPL (Rekayasa Perangkat Lunak), TKJ (Teknik Komputer Jaringan), Transmisi, dan Akses dan sekarang memiliki lulusan dari 26 Angkatan, yang berarti sudah dari 1000 siswa/siswi yang telah menempuh Pendidikan di Smk Telkom Jakarta, dengan adanya data alumni sebanyak ini menjadi ide dasar untuk melakukan implementasi terhadap data besar dengan metode apache Hadoop.[1]

Implementasi ini akan diarahkan dengan menggunakan query atau permasalahan yang digunakan dalam mengolah data alumni, permasalahan ini adalah memberikan sebuah data statistik yaitu berupa data alumni pada tahun 2017 dan 2018 dengan menentukan peminatan jurusan yang paling banyak pada antara tahun tersebut dan data akan bisa digunakan acuan oleh sekolah kedepannya dalam kegiatan mengajar atau dalam kegiatan Kerjasama antara perusahaan dengan sekolah.

Dengan permasalahan diatas, pada penelitian ini dibahas tentang rancang bangun *Big Data* menggunakan Apache Hadoop. Apache Hadoop adalah kerangka kerja yang menyimpan dan memproses data besar secara terdistribusi dan paralel, dan di dalam Hadoop sendiri pun banyak sekali lapisan Penyimpanan Hadoop. Berdasarkan hal tersebut, pada penelitian ini penulis mengambil judul “**RANCANG BANGUN BIG DATA MENGGUNAKAN APACHE HADOOP STUDY CASE DATA ALUMNI SMK TELKOM JAKARTA**”.

## 1.2 Rumusan Masalah

1. Bagaimana cara menggunakan sistem Apache Hadoop?
2. Bagaimana merancang big data menggunakan Apache Hadoop?
3. Bagaimana mengolah data yang terhubung dengan Apache Hadoop?

## 1.3 Batasan Masalah

Dalam tugas ini terdapat batasan – batasan, antara lain:

1. Penelitian ini di fokuskan pada data Alumni SMK Telkom Jakarta
2. Penelitian ini menggunakan *Waterfall Model* untuk metode pengembangan sistem Simulasi menggunakan *website* pada *localhost*
3. Data yang diujikan hanya 2 angkatan yang terdiri dari 769 siswa

## 1.4 Tujuan Penelitian

Tujuan dari penelitian tugas ini adalah sebagai berikut:

1. Dapat membuat big data lebih efektif dalam mengolah data
2. Dapat menghubungkan data dengan sistem Hadoop
3. Menganalisa prinsip kerja antara Hadoop dengan data yang ada.

## 1.5 Manfaat Penelitian

Manfaat dari penelitian tugas ini adalah sebagai berikut:

1. Menghasilkan data yang lebih efisien dalam pengolahan.
2. Membantu peneliti lain untuk mengembangkan Apache Hadoop di tempat lain.
3. Mengetahui hasil dari pengujian data yang ada dengan menggunakan sistem Apache Hadoop.

## 1.6 Metodologi Penelitian

Tahapan – tahapan yang akan dilakukan dalam penelitian ini adalah sebagai berikut :

### Teknik pengumpulan data

#### a. Observasi

Teknik dalam pengumpulan data ini dilakukan dengan cara mengamati langsung dan terjun langsung ke lapangan untuk mengumpulkan data alumni yang berkaitan dengan masalah penelitian. Penelitian berada di Smk Telkom Jakarta demi mendapatkan bukti-bukti yang *valid* dan data alumni dalam laporan yang akan diajukan. Observasi adalah metode untuk melakukan pengumpulan data dimana peneliti mencatat informasi sebagaimana data yang ada dilapangan.

b. Wawancara

Wawancara merupakan percakapan demi mencapai informasi tertentu. Percakapan ini dilakukan oleh pewawancara yaitu peneliti sendiri yang mengajukan pertanyaan dan terwawancara yaitu bagian Hubungan Industri (Hubin) yang memberikan jawaban. Melalui kegiatan wawancara inilah peneliti menggali data lebih dalam berkaitan dengan data, informasi, dan keterangan-keterangan lainnya yang menunjang penelitian.

c. Survei

Penelitian survei merupakan metode penelitian yang digunakan untuk mendapatkan data alumni yang terjadi pada masa lampau. Dengan melakukan survei peneliti bisa mendapatkan data data yang lebih akurat dalam penelitian. Survei dilakukan berada di Smk Telkom Jakarta.

## 1.7 Sistematika Penulisan

Secara umum sistematika penulisan tugas ini terdiri dari bab-bab dengan metode penyampaian sebagai berikut:

### **BAB I PENDAHULUAN**

Berisi latar belakang masalah, tujuan, manfaat, metodologi penelitian, dan sistematika penulisan.

### **BAB II DASAR TEORI**

Berisi teori-teori yang mendukung tugas ini, seperti membahas tentang *Database*, *Data Warehouse*, *Big Data*, skema, dan *software* pendukung lainnya.

### **BAB III PERANCANGAN DAN IMPLEMENTASI**

Membahas masalah perancangan *big data* yang terhubung Hadoop untuk mengelola data Alumni Smk Telkom Jakarta.

### **BAB IV ANALISA HASIL PERANCANGAN**

Pada bab ini membahas hasil dari pembuatan *website* berdasarkan parameter – parameter yang telah ditentukan.

### **BAB V PENUTUP**

Pada bab ini berisi kesimpulan dan saran-saran yang mendukung untuk kesempurnaan tugas ini.