

BAB I

PENDAHULUAN

1.1. Latar Belakang Penelitian

Di masa ini, manusia dituntut mempunyai keahlian dan keterampilan yang mumpuni. Apalagi saat ini dunia mengalami perkembangan teknologi yang masif dan cepat. Banyak hal dan cara yang dapat dilakukan agar mempunyai keahlian dan keterampilan yang dibutuhkan dunia. Salah satunya memperoleh pendidikan dari institusi yang ada. Di Indonesia, pendidikan tinggi menjadi salah satu penyedia sarana untuk memperoleh keahlian dan keterampilan. Selain itu, keahlian dan keterampilan yang didapat akan diarahkan untuk kepentingan masyarakat maupun di dunia kerja nantinya.

Pendidikan tinggi merupakan tingkatan pendidikan terakhir dan biasanya menjadi salah satu tujuan dari para siswa setelah menamatkan pendidikan menengah. Di Indonesia, penyelenggara dari pendidikan tinggi diemban oleh perguruan tinggi. Bentuk perguruan tinggi di Indonesia antara lain, universitas, institut, sekolah tinggi, politeknik, dan akademi. Tidak hanya sebagai lembaga penyelenggara pendidikan, institusi-institusi ini memiliki peran penting lainnya. Menjadi garda terdepan dalam pengembangan dan riset ilmu pengetahuan untuk kepentingan masyarakat luas menjadi salah satu fungsi penting lainnya bagi sebuah perguruan tinggi.

Selain pengembangan dan riset, perguruan tinggi juga mempunyai peran penting dalam menggerakkan roda perekonomian dan sosial sebuah negara. Peran

ini bermaksud bahwa perguruan tinggi harus menghasilkan sumber daya manusia berkualitas untuk negara. Tidak hanya dalam kegiatan akademik, perguruan tinggi juga diharapkan menjadi pusat budaya dan aktivitas sosial yang membentuk identitas, karakter, dan watak masyarakat.

Dalam perjalanannya, perguruan tinggi juga menghadapi berbagai rintangan dan dinamika yang ada. Dinamika yang terjadi seperti kebijakan pendidikan yang terus berubah, kemajuan teknologi, dan peminat yang terus bertambah. Hal ini harus bisa dihadapi dengan bijaksana sehingga standar yang ada pada perguruan tinggi tersebut dapat terus dipertahankan. Pengambilan kebijakan yang tepat menghadapi perubahan zaman yang cepat tentu saja menjadi salah satu langkah yang penting. Universitas Komputer Indonesia (UNIKOM) sebagai salah satu perguruan tinggi yang ada di Indonesia tentu saja juga harus menghadapi dinamika dan rintangan yang ada.

Saat ini Universitas Komputer Indonesia adalah perguruan tinggi swasta yang terdapat di Kota Bandung. Universitas ini berdiri sekitar 23 tahun yang lalu. Perguruan tinggi ini memiliki 7 fakultas sarjana dan 1 fakultas pascasarjana. Tujuh fakultas sarjana dengan 25 program studi untuk S1 dan D3. Satu fakultas pascasarjana dengan 4 program studi untuk S2 dan S3. Pada tahun 2015 jumlah mahasiswa UNIKOM sebanyak 15.000 orang yang berasal dari seluruh tanah air. Tentu saja dalam perjalanannya UNIKOM pasti menghadapi berbagai dinamika yang ada. Untuk menanggulangi dinamika yang akan muncul di masa depan kita bisa melakukan tindakan preventif. Salah satunya melakukan sebuah segmentasi terhadap mahasiswa baru yang masuk di UNIKOM.

Salah satu permasalahan yang biasa dihadapi sebuah universitas adalah kurangnya persiapan dan manajemen yang baik saat memulai tahun ajaran baru. Lalu, belum adanya sebuah sistem yang dapat melakukan pengelompokan mahasiswa berdasarkan alasan mereka memilih Universitas Komputer Indonesia. Di mana saat ini hanya ada pengelompokan mahasiswa berdasarkan hal yang kasat mata seperti kota asal, jenis kelamin, dll. Sehingga, untuk mengetahui lebih dalam alasan tersebut bisa dilakukan penerapan *data mining* di dalamnya.

Segmentasi ini dapat memberikan sebuah gambaran mengenai alasan kenapa para mahasiswa baru memilih Universitas Komputer Indonesia. Sehingga, kebutuhan yang diperlukan untuk dipersiapkan menghadapi mahasiswa baru tiap tahunnya dapat dipenuhi maupun hal yang dirasa kurang menurut pandangan mahasiswa dapat diperbaiki oleh pihak kampus.

Selain itu, sistem pengambilan keputusan yang dilakukan berdasarkan segmentasi profil mahasiswa belum dilakukan. Pengambilan keputusan yang berdasarkan profil mahasiswa ini dapat meningkatkan kepuasan mahasiswa. Keuntungan lainnya adalah dapat meningkatkan nama Universitas Komputer Indonesia di kalangan mahasiswa baru.

Hal ini bisa kita dapatkan setelah mengetahui alasan mereka yang didapatkan dari sebuah pertanyaan (kuesioner) yang diberikan kepada mahasiswa baru. Pertanyaan dalam kuesioner ini akan dibungkus melalui sebuah pertanyaan mengenai strategi pemasaran yaitu *7P marketing mix*.

7P *marketing mix* adalah kegiatan pemasaran yang dilakukan secara terpadu. Artinya kegiatan ini dilakukan secara bersama (gabungan) di antara elemen-elemen yang ada dalam *marketing mix* itu sendiri. Karena dalam konsepnya setiap elemen pemasaran tidak dapat berjalan sendiri tanpa didukung oleh elemen lainnya. [2]

Data mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam *database*. *Data mining* adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai *database* besar. [3] Hasil dari *data mining* ini adalah yang nanti akan digunakan untuk pengambilan keputusan oleh *stakeholder*. Algoritma yang akan digunakan adalah algoritma berjenis *unsupervised learning* yaitu *K-Means*.

K-Means merupakan algoritma yang paling banyak digunakan dalam berbagai aplikasi kecil mulai dari tingkat sederhana sampai dengan tingkat menengah karena kemudahan implementasinya. Ide dasar algoritma ini adalah meminimalkan *Sum of Square Error* (SSE) antara obyek data dengan sejumlah *k centroid*. Algoritma ini bekerja dengan 4 langkah, yaitu himpunan data yang akan dikelompokkan dipilih sejumlah *k* obyek secara acak sebagai *centroid* awal, setiap obyek yang bukan *centroid* akan dimasukkan ke dalam kluster terdekat berdasarkan ukuran jarak tertentu, setiap *centroid* diperbaiki berdasarkan rata-rata dari obyek yang ada dalam setiap klasternya dan dilakukan berulang-ulang sampai dengan semua *centroid* stabil atau konvergen. [4]

Berdasarkan penjelasan pada paragraf sebelumnya dapat dinyatakan bahwa dengan adanya segmentasi mahasiswa baru dapat memberikan keuntungan bagi Universitas Komputer Indonesia. Dengan kata lain Universitas Komputer Indonesia dapat memahami profil mahasiswa baru mereka, dapat mengembangkan program atau kebijakan yang lebih efektif, dan mampu mengembangkan strategi pemasaran yang dimiliki agar menjadi lebih baik lagi. Oleh karena itu, penulis bermaksud untuk menyusun skripsi dengan judul **“Implementasi Data Mining Menggunakan Algoritma K-Means Untuk Segmentasi Mahasiswa Baru”**.

1.2. Identifikasi dan Rumusan Masalah

Identifikasi dan rumusan masalah yang ada pada Universitas Komputer Indonesia adalah sebagai berikut.

1.2.1. Identifikasi Masalah

Berdasarkan uraian latar belakang yang diuraikan, maka identifikasi masalah pada penelitian ini adalah:

1. Sistem segmentasi (pengelompokan) mahasiswa baru hanya sebatas atribut yang tampak seperti kota asal dan faktor-faktor penentu lainnya belum dilibatkan.
2. Pengambilan keputusan kampus belum berdasarkan segmentasi yang terdapat di mahasiswa.

3. Penyebaran alasan mahasiswa baru ketika memilih Universitas Komputer Indonesia belum diteliti menggunakan *data mining* untuk mencari informasi yang dapat membantu perkembangan kampus

1.2.2. Rumusan Masalah

Berdasarkan uraian latar belakang, maka dapat dirumuskan permasalahan sebagai berikut:

1. Bagaimana *preprocessing data* dalam segmentasi mahasiswa baru di Universitas Komputer Indonesia
2. Bagaimana *modelling* dalam segmentasi mahasiswa baru di Universitas Komputer Indonesia
3. Bagaimana evaluasi hasil dari model dalam segmentasi mahasiswa baru di Universitas Komputer Indonesia
4. Bagaimana implementasi model web dalam segmentasi mahasiswa baru di Universitas Komputer Indonesia

1.3. Maksud dan Tujuan Penelitian

1.3.1. Maksud Penelitian

Maksud dari penelitian ini adalah untuk mengaplikasikan teknik *data mining* dengan algoritma *k-means* dalam segmentasi mahasiswa baru di Universitas Komputer Indonesia sehingga dapat membantu pengambilan kebijakan kampus.

1.3.2. Tujuan Penelitian

Adapun tujuan dari penelitian ini yaitu:

1. Mengetahui *data preparation* dalam segmentasi mahasiswa baru di Universitas Komputer Indonesia
2. Mengetahui *modelling* algoritma *k-means* dalam segmentasi mahasiswa baru di Universitas Komputer Indonesia
3. Mengetahui evaluasi hasil dari model algoritma *k-means* dalam segmentasi mahasiswa baru di Universitas Komputer Indonesia
4. Mengetahui implementasi model web dalam segmentasi jumlah mahasiswa baru di Universitas Komputer Indonesia

1.4. Kegunaan Penelitian

Penelitian ini diharapkan dapat bermanfaat bagi Universitas Komputer Indonesia untuk mempersiapkan kebijakan dan hal lainnya setiap tahun dalam rangka menerima mahasiswa baru. Hal ini juga akan bermanfaat bagi peneliti lain dalam melakukan skripsi atau tugas akhir sebagai referensi dalam proses penulisan. Bagi penulis, penelitian ini tentu saja menambah wawasan dan ilmu baru. Ilmu baru yang didapatkan juga diharapkan dapat diimplementasikan di masa depan untuk kepentingan masyarakat luas.

1.5. Batasan Masalah

Batasan masalah dalam penelitian ini bertujuan untuk memberi fokus ruang lingkup dalam permasalahan yang akan diteliti. Berikut adalah batasan masalah pada penelitian ini:

1. Data yang dipakai adalah data yang diperoleh dari hasil penerimaan mahasiswa baru (PMB) di Universitas Komputer Indonesia dari tahun

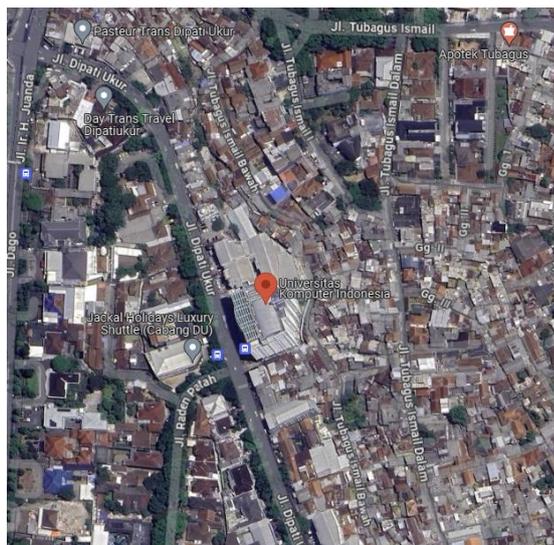
2020 sampai tahun 2022 dengan 5839 jumlah data dari semua jurusan yang ada.

2. Data mahasiswa yang diolah adalah mahasiswa program studi sistem informasi dari tahun 2020 sampai tahun 2022.
3. Metode yang akan digunakan yaitu algoritma k-means untuk menemukan segmentasi mahasiswa baru.
4. Penelitian ini menggunakan *tools Google Colab, Streamlit, dan Visual Studio Code.*

1.6. Lokasi dan Waktu Penelitian

1.6.1. Lokasi Penelitian

Lokasi penelitian ini dilakukan di Universitas Komputer Indonesia yang beralamat di Jl. Dipati Ukur No. 112-116, Lebakgede, Kecamatan Coblong, Kota Bandung, Jawa Barat 40132.



Gambar 1. 1 Lokasi Penelitian

(Sumber: <https://www.google.com/maps>)

1.6.2. Waktu Penelitian

Penelitian ini dilaksanakan selama 4 bulan, dimulai pada bulan April hingga Agustus 2024.

Tabel 1. 1 Waktu Penelitian

No.	Nama Kegiatan	Waktu Kegiatan																
		April				Mei				Juni				Juli				Agustus
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1
1.	Identifikasi Kebutuhan																	
a.	Studi Literatur	■	■															
b.	Observasi		■	■														
c.	Wawancara				■	■												
d.	Pengumpulan Data				■	■												
2.	Perancangan dan Pembangunan																	
a.	Perancangan <i>Preprocessing Data</i>					■	■											
b.	<i>Splitting Data</i>							■	■									
3.	Pengujian dan Implementasi																	
a.	<i>Modelling Data</i>									■	■	■						
b.	<i>Training Data</i>												■	■	■			
c.	<i>Deployment Website</i>															■	■	

1.7. Sistematika Penulisan

Sistematika penulisan penelitian ini terdiri dari beberapa bab yang akan mencakup pokok dari materi yang dibahas pada laporan skripsi ini. Adapun sistematika penulisan laporan ini terdiri dari:

BAB I PENDAHULUAN

Bab ini terdiri atas latar belakang, identifikasi dan rumusan masalah, maksud dan tujuan, kegunaan penelitian, jadwal dan lokasi penelitian, hingga sistematika penulisan.

BAB II LANDASAN TEORI

Bab ini terdiri dari penjelasan atas referensi yang digunakan dalam penulisan skripsi ini. Serta penjelasan dari teori yang digunakan untuk memperjelas penelitian dan topik yang diangkat dalam penelitian ini.

BAB III OBJEK DAN METODE PENELITIAN

Bab ini berisi tentang objek dan metode penelitian yang akan diangkat. Terdiri atas sejarah singkat perusahaan, visi dan misi perusahaan, struktur organisasi perusahaan, deskripsi tugas, tahap penelitian, dan analisis sistem yang berjalan.

BAB IV HASIL DAN PEMBAHASAN

Pada bab ini penulis membahas hasil dari penelitian yang menggunakan algoritma regresi linear, perhitungan manualnya, dan implementasi *website* sederhana untuk segmentasi mahasiswa baru.

BAB V KESIMPULAN DAN SARAN

Pada bab ini penulis memaparkan hasil dari penelitian yang telah dilaksanakan untuk menjawab permasalahan dan tujuan yang sudah dijabarkan

sebelumnya. Selain itu, penulis memberikan saran mengenai penelitian ini yang dapat membantu pengembangan penelitian ke depannya.