

## **BAB II**

### **TINJUAN PUSTAKA**

#### **2.1 Tinjauan Penelitian Terdahulu**

Penelitian terdahulu terkait topik pembahasan yang di angkat sangat penting untuk dijadikan dasar serta sumber data. Hal ini bertujuan untuk mendukung penelitian yang sedang dilakukan. Selain itu untuk mengajukan gagasan atau argumen diperlukan data yang relevan dengan penelitian sehingga dapat dikatakan valid. Tujuan lain dilakukannya studi terhadap penelitian terdahulu adalah untuk menggali kekurangan dan kelebihan suatu metode yang akan digunakan. Dengan begitu peneliti dapat menghindari atau mengatasi permasalahan terkait.

Oleh sebab itu sebelum memulai penelitian perlu dilakukan eksplorasi dan studi pustaka baik melalui internet maupun jurnal – jurnal yang relevan dengan topik Segmentasi Mahasiswa Dengan Unsupervised Data Untuk Prediksi Penerimaan Mahasiswa Baru. Berikut adalah beberapa penelitian yang dijadikan tinjauan pustaka. Pada beberapa halaman selanjutnya disediakan tabel perbandingan antar-penelitian yang dapat dilihat pada Tabel 2.1.

Penelitian pertama dilakukan oleh [3] dengan judul “Peran *Bussines Intelligence* Dalam Menentukan Strategi Promosi Penerimaan Mahasiswa Baru” penelelitian ini menggunakan *Bussines Intelligence* sehingga membantu organisasi atau perusahaan mendapatkan informasi yang jelas dan tepat

mengenai berbagai faktor-faktor yang dapat mempengaruhi kinerja organisasi sehingga dapat memudahkan organisasi tersebut dalam pengambilan keputusan serta sekaligus meningkatkan keunggulannya (*competitive advantage*). Dengan kegiatan BI tersebut maka sebuah organisasi atau perusahaan akan dengan mudah dalam mengambil keputusan secara cepat dan tepat. Dalam penelitian ini BI digunakan untuk mengelompokkan data mahasiswa dengan menggunakan proses data mining, atau di sebut juga metode data *non-hierarchical clustering* yang dapat mengelompokkan data mahasiswa ke dalam beberapa cluster berdasarkan kemiripan data tersebut, sehingga data mahasiswa yang memiliki karakteristik yang sama dikelompokkan dalam satu cluster dan yang memiliki karakteristik berbeda dikelompokkan dalam cluster yang lain. Atribut yang digunakan dalam menentukan pengelompokkan wilayah untuk melakukan promosi di ISTA adalah nama mahasiswa, NIM, kota asal, nilai IPK dan program studi

*Implementasi data mining untuk menentukan persediaan stok burger menggunakan metode K-Means clustering.*

Penelitian yang dilakukan oleh [4] yang digunakan untuk skripsi di Universitas Bina Sarana Informatika pada tahun 2016 membahas tentang mengelompokkan data dari hasil transaksi pada perusahaan burger untuk mengetahui produk burger yang diminati pelanggan digunakan sebagai tingkat persediaan stok burger dengan tujuan membantu perusahaan untuk merancang strategi pemasaran dan pengendalian persediaan yang baik.

Penelitian kedua dilakukan oleh [5] dengan judul penelitian “Metode Microsoft Clustering untuk Meningkatkan Potensi Penerimaan Mahasiswa Baru” pada penelitian kali ini bertujuan untuk menentukan daerah yang berpotensi memberikan sumbangan jumlah calon mahasiswa baru dengan menggunakan metode clustering terhadap data pendaftar. Pada penelitian ini penulis menggunakan tools Microsoft Tools SQL Analysis Service, Microsoft Office Excel 2010 yang sudah dilengkapi dengan Data Mining AddIn for Microsoft Excel 2007. Dan penulis menggunakan data pendaftar tahun 2009 sampai dengan tahun 2011 yang berjumlah 15.000 record.

Penelitian ke tiga dilakukan oleh [6] dengan judul “A dynamic K-means clustering for data mining” dalam paper ini dijelaskan tentang Pengembangan data adalah proses menemukan struktur data dari set data yang besar. Dengan proses ini, para pembuat keputusan dapat membuat keputusan khusus untuk pengembangan lebih lanjut dari masalah-masalah dunia nyata. Beberapa teknik pengelompokan data digunakan dalam penggalian data untuk menemukan pola data tertentu. Metode K-means adalah salah satu teknik pengelompokan yang akrab untuk pengelompokan kumpulan data besar. Metode pengelompokan K-means mempartisi data yang ditetapkan berdasarkan asumsi bahwa jumlah cluster ditetapkan. Masalah utama dari metode ini adalah bahwa jika jumlah cluster harus dipilih kecil maka ada kemungkinan lebih tinggi untuk menambahkan item yang berbeda ke dalam kelompok yang sama. Di sisi lain, jika jumlah cluster dipilih menjadi tinggi, maka ada kemungkinan lebih tinggi untuk menambahkan item serupa di kelompok yang berbeda. Dalam makalah ini, kami mengatasi masalah

ini dengan mengusulkan algoritma pengelompokan K-Means baru. Metode yang diusulkan melakukan pengelompokan data secara dinamis. Metode yang diusulkan awalnya menghitung nilai ambang sebagai centroid dari KMeans dan berdasarkan nilai ini jumlah cluster terbentuk. Pada setiap iterasi K-Means, jika jarak Euclidian antara dua titik kurang dari atau sama dengan nilai ambang batas, maka kedua titik data ini akan berada dalam kelompok yang sama. Jika tidak, metode yang diusulkan akan membuat cluster baru dengan titik data yang berbeda. Hasil menunjukkan bahwa metode yang diusulkan mengunggulkan metode K-Means.

Penelitian ke empat dilakukan oleh [7] peneliti melakukan penelitian yang berjudul “Implementasi Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing President University”. Dalam penelitian ini, melakukan perhitungan dataset mahasiswa yang telah lulus dari President University menggunakan algoritma k-means clustering. Pengelompokkan data mahasiswa ke dalam beberapa kelompok berdasarkan karakteristik dari dataset untuk menemukan informasi yang tersembunyi dari dataset mahasiswa yang telah lulus dari President University. Data atribut yang digunakan dalam penelitian ini adalah asal kota, program studi dan IPK. Tujuan dari penelitian ini adalah untuk membantu departemen pemasaran President University dalam memprediksi strategi promosi yang dilakukan di kota-kota di Indonesia. Informasi yang diperoleh dalam penelitian ini dapat digunakan sebagai referensi dalam menentukan strategi yang tepat untuk tim pemasaran dalam kegiatan promosi di kota-kota di Indonesia sehingga promosi akan lebih efektif dan efisien.

Tabel 2.1 Penelitian Terdahulu

No	Peneliti	Judul Penelitian	Perumusan Masalah	Hasil/Temuan
1.	[3]	<p>Peran <i>Bussines Intelligence</i></p> <p>Dalam Menentukan Strategi Promosi Penerimaan Mahasiswa Baru</p>	<p>Bagaimana cara untuk melakukan pengelompokan data mahasiswa ke dalam beberapa cluster berdasarkan kemiripan dari data tersebut.</p>	<p>Berdasarkan penelitian yang telah dilakukan pengelompokan data mahasiswa melalui persebaran wilayah kota asal, peminatan program studi dan potensi akademik menggunakan K-Means clustering dan diolah dengan Rapid Miner, terbentuk 3 (tiga) cluster yaitu :</p> <p>a. Cluster 1 dengan jumlah 10 mahasiswa berasal dari wilayah kota asal DKI Jakarta dan Jawa Barat dengan peminatan program studi Farmasi dan rata-</p>

				<p>rata IPK 3,26.</p> <p>b. Cluster 2 dengan jumlah 147 mahasiswa berasal dari wilayah kota asal Jawa Barat dan DKI Jakarta dengan peminatan program studi Farmasi dan Teknik Kimia dan rata-rata IPK 3,27. c. Cluster 3 dengan jumlah 175 mahasiswa berasal dari wilayah kota asal Jawa Barat dan DKI Jakarta, peminatan program studi Farmasi, dengan rata-rata IPK 3,26..</p>
2	[5]	Metode Microsoft Clustering untuk Meningkatkan	Bagaimana menentukan daerah yang berpotensi memberikan	Berdasarkan rangkuman hasil cluster serta analisis, dapat direkomendasikan

		<p>Potensi Penerimaan Mahasiswa Baru</p>	<p>sumbangan jumlah calon mahasiswa baru dengan menggunakan metode klatering terhadap data pendaftar</p>	<p>strategi promosi sebagai berikut :</p> <ol style="list-style-type: none"> <li>1. Untuk efektifitas biaya, kegiatan promosi difokuskan di sekolah-sekolah yang sudah terbukti memberikan kontribusi, dimana informasi terkait didapat dari data tahun-tahun sebelumnya.</li> <li>2. Cluster dimana pilihan pertama PGSD dominan, maka pilihan 2 kosong juga dominan. Hal ini menandakan calon mahasiswa yang mendaftar di progdi PGSD tidak memilih</li> </ol>
--	--	--	--	--

				<p>pilihan progdi lain, dan menunjukkan tingkat keseriusan yang tinggi terhadap pilihan progdi, serta juga meningkatkan kepastian untuk melakukan daftar ulang. Progdi PGSD sebaiknya melakukan pemerataan daerah, sehingga tidak terpusat pada Jawa Tengah bagian utara dan Kalimantan Barat saja.</p> <p>3. Sekolah-sekolah Islam, sebaiknya diberikan perhatian dalam kegiatan promosi. Berdasarkan data untuk wilayah Jawa Tengah pada</p>
--	--	--	--	--

				<p>tahun 2011, terdapat 259 orang pendaftar dari sekolah Islam, sementara promotion kit yang dikirimkan untuk sekolah-sekolah Islam hanya untuk 4 sekolah. Jika semua sekolah-sekolah Islam dijadikan tujuan promosi, maka akan lebih banyak jumlah calon pendaftar.</p>
3	[6]	<p>A dynamic K-means clustering for data mining</p>	<p>jumlah cluster harus dipilih kecil maka ada kemungkinan lebih tinggi untuk menambahkan item yang berbeda ke</p>	<p>Berdasarkan hasil paper, penulis telah mensimulasikan metode yang kami usulkan menggunakan MATLAB, Java, MapReduce, dan C ++ di komputer pribadi. Spesifikasi komputer</p>

			<p>dalam kelompok yang sama.</p>	<p>pribadi adalah prosesor 4GB RAM 2,4 GHz Corei5. Pada awalnya, metode yang kami usulkan dikembangkan dalam C ++ dan kemudian dikonversi ke java MapReduce. Hasil dari metode yang kami usulkan diterapkan dalam MATLAB untuk melihat posisi cluster dan data point. Kemudian kami telah mengembangkan algoritma K-Means umum untuk membandingkan dengan metode yang kami usulkan.</p>
4	[7]	Implementasi	Bagaimana	Dari data hasil

		<p>Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing President University</p>	<p>menentukan strategi promosi di University President yang berlokasi di Jababeka.</p>	<p>clustering yang telah dilakukan di atas, maka dapat dibuat beberapa strategi promosi yang dapat dilakukan oleh pihak marketing President University agar promosi yang dilakukan lebih efektif dan efisien, yaitu:</p> <ol style="list-style-type: none"> <li>1. Promosi Dengan Mengirim Tim Marketing yang Sesuai dengan Jurusan yang Paling Banyak Diminati. Strategi promosi pertama yang dapat dilakukan pihak marketing President University berdasarkan hasil clustering adalah</li> </ol>
--	--	--	--	--

				<p>melakukan strategi promosi pada kota-kota di Indonesia berdasarkan jurusan yang paling banyak diminati. Jadi, pihak marketing dapat mengirim tim yang memiliki pengetahuan dan potensi lebih dari suatu jurusan untuk melakukan promosi pada kota yang memiliki minat lebih besar pada jurusan tersebut, sehingga promosi yang dilakukan akan lebih efektif dan efisien karena dengan melakukan strategi ini pihak marketing dapat membagi tim-tim</p>
--	--	--	--	---

				<p>marketing untuk melakukan promosi di kota-kota di Indonesia sehingga sumber daya marketing yang dibutuhkan disetiap kota tidak terlalu banyak</p> <p>2. Promosi Pada Kota Berdasarkan Tingkat Akademik dari Calon Mahasiswa. Strategi promosi kedua yang dapat dilakukan pihak marketing President University, yaitu apabila pihak marketing President University ingin mendapat calon mahasiswa</p>
--	--	--	--	---

## 2.2 Teori Dasar Yang Digunakan

Pada penelitian ini terdapat beberapa teori dasar yang dipergunakan sebagai landasan teori dan referensi terkait pembahasan mengenai Segmentasi Mahasiswa dengan *unsupervised learning* data buntut memprediksi penerimaan mahasiswa baru. Sumber didapatkan dari jurnal dan paper yang relevan dengan penelitian ini serta studi pustaka melalui buku dan internet.

### 2.2.1 Strategi Promosi

Strategi promosi adalah langkah yang harus diperhitungkan dalam rangkaian aktivitas manajemen pemasaran [8]. Kegiatan pada setiap Strategi Promosi merupakan cara di bidang pemasaran berupa interaksi atau komunikasi yang diterapkan perusahaan terhadap konsumen atau pelanggan. Interaksi yang ditentukan pada strategi promosi pemasaran berupa pemberian informasi, mengajak, dan memberi pengaruh mengenai suatu produk.

Strategi Promosi bisa dibagi menjadi dua yaitu strategi promosi jasa dan strategi promosi barang. Keduanya memiliki sifat yang berbeda. Walaupun demikian kedua strategi promosi tersebut bertujuan dalam peningkatan kapasitas penjualan dengan cara memberikan penawaran terbaik ke konsumen atau pelanggan khususnya pada penentuan keputusan pembelian.

### 2.2.2 Data mining

Data Mining merupakan proses pencarian pola tersembunyi dalam sejumlah data dengan tujuan untuk melakukan klasifikasi, prediksi, klusterisasi untuk

mengetahui informasi yang dapat digunakan sebagai strategi bisnis dalam sebuah perusahaan[9].

Kemampuan data mining untuk mencari informasi bisnis yang berharga dari bisnis data yang sangat besar, dapat diartikan dengan pembangunan kota dengan pertumbuhan penduduk yang sangat pesat, teknologi ini dipakai untuk :

1. Prediksi trend dan sifat – sifat bisnis, dimana data mining mengotomatisasi proses pencarian informasi dalam memprediksi data base yang sangat besar.
2. Data mining berguna untuk membuat keputusan yang tepat dalam strategi.
3. Penemuan pola – pola yang belum diketahui sebelumnya, dimana sebelumnya data mining menghitung semua basis data, kemudian mengidentifikasi pola pola yang sebelumnya tersembunyi dalam suatu data.

Berikut ini beberapa definisi data mining dari beberapa sumber :

1. Data mining diartikan sebagai suatu proses ekstrasi yang berguna dan berpotensi terhadap sekumpulan data yang terdapat secara implisit dalam suatu basis data.
2. Data mining adalah analisis pengamatan database untuk menemukan hubungan yang tidak terduga dan menyimpulkan data dengan dengan cara metode baru yang dapat dimengerti, dipahami dan bermanfaat bagi pemilik data,
3. Data mining adalah proses menemukan sesuatu yang bermakna dari suatu korelasi, tren, dan pola yang baru dengan cara memilah milah data berukuran

besar yang disimpan didalam repositori, menggunakan teknologi pengenalan pola serta teknik matematis dan statistis.

4. Data mining merupakan bidang ilmu interdisipliner yang menyatukan teknik pembelajaran dari mesin (*mechine learning*), pengenalan pola, ststistik, database, dan visualilsai untuk mengatasi masalah data basis data yang besar.

### **Fungsi Data Mining**

Data mining mempunyai fungsi yang penting untuk membantu mendapatkan informasi yang berguna serta meningkatkan ilmu pengetahuan bagi pengguna. Pada dasarnya, data mining mempunyai empat (4) fungsi datar yaitu :

- Fungsi prediksi (*prediction*). Proses ini diutamakan untuk menemukan pola data dengan menggunakan beberapavariabel untuk memprediksi variabel lainnya yang belum dapat diketahui jenis dan nilainya.
- Fungsi klasifikasi (*classification*). Klasifikasi merupakan suatu proses untuk menemukan model atau fungsi untuk menggambarkan *class* atau konsep suatu data. Proses yang digunakan untuk mendeskripsikan data yang penting serta dapat memperkirakan kecenderungan data pada masa depan.
- Fungsi deskripsi (*deskription*) proses untuk menemukan suatu karakteristik yang belum ditemukan dalam suatu basis data.
- Fungsi asosiasi proses ini digunakan untuk menemukan suatu hubungan yang terdapat pada nilai atribut dari sekumpulan data.

### 2.2.3 K-means

Metode ini dipilih sebagai alternatif, untuk melihat bagaimana data tangisan suara tersebut dikelompokkan secara natural oleh K-means. Metode ini dapat mengelompokkan data berdasarkan kemiripan dari karakteristik data dengan cara menentukan di awal dahulu berapa kelompok data yang ingin dibentuk. Dalam penelitian ini ditentukan kelompok data sebanyak 5 sesuai dengan kategori Dunstan Baby Language. Sehingga akan terbentuk kelompok atau cluster data yang belum diketahui nama-nama kategori datanya. Diterapkannya metode K-means pada penelitian ini untuk mengetahui seberapa cocok label cluster dengan label aktual pada data tangisan suara bayi tersebut. Biasanya, kategori yang dihasilkan oleh label cluster milik K-means akan memiliki performa akurasi yang baik dibanding label aktualnya. Namun kekurangannya adalah dengan digunakannya label cluster sebagai model data, maka akan menyalahi aturan data label aktualnya yang sudah ditentukan di awal.

Data mining dapat diterapkan untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual. Terdapat beberapa teknik yang digunakan dalam data mining, salah satu teknik data mining adalah clustering. Terdapat dua jenis metode clustering yang digunakan dalam pengelompokkan data, yaitu hierarchical clustering dan non-hierarchical clustering

K-means clustering sebagai salah satu metode data clustering non-hirarki mempartisi data yang ada ke dalam bentuk satu atau lebih cluster atau kelompok, sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam

satu cluster yang sama dan data yang mempunyai karakteristik yang berbeda dikelompokkan ke dalam kelompok yang lain. Kelompok atau cluster yang didapat merupakan pengetahuan atau informasi yang bermanfaat bagi pengguna kebijakan dalam proses pengambilan keputusan [10]. K-means clustering merupakan salah satu metode cluster analysis non hirarki yang berusaha untuk mempartisi objek yang ada kedalam satu atau lebih cluster atau kelompok objek berdasarkan karakteristiknya, sehingga objek yang mempunyai karakteristik yang sama dikelompokkan dalam satu cluster yang sama dan objek yang mempunyai karakteristik yang berbeda dikelompokkan kedalam cluster yang lain.

#### **2.2.4 Universitas Narotama**

Pada tanggal 20 Februari 1981 didirikanlah sebuah Yayasan Pawiyatan Gita Patria sebagai badan hukum pembina Perguruan Tinggi Swasta, Yayasan ini selanjutnya mendirikan Perguruan Tinggi Swata yang diberi nama Universita Narotama. Nama Narotama diambil dari nama seorang tokoh tokoh sejarah Majapahit dari Prabu Airlangga yang sekaligus juga sebagai guru ilmu kenegaraan serta guru agama dan ilmu kedigdayaan.

Jadi tepat kiranya Universitas Narotama mengambil nama guru Prabu Airlangga karena para pendiri dan pengelolanyamempunyai tujuan dan cita-cita yang sama dengan pendiri Universitas sebelumnya. Universitas narotama dalam perjalanan waktu hingga.