

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Metode text mining adalah evolusi dari data mining yang berfokus pada analisis teks. Ini adalah proses otomatis yang dilakukan oleh komputer untuk mengekstraksi informasi berharga dari kumpulan teks dalam sebuah dokumen. Langkah utama dalam metode ini melibatkan identifikasi kata-kata yang mencerminkan isi dokumen, kemudian menganalisis hubungan antar dokumen menggunakan metode statistik seperti teknik clustering.

Terdapat berbagai algoritma yang diajukan untuk melakukan proses clustering pada dataset berukuran besar. Dalam penelitian ini, peneliti akan menerapkan algoritma K-Means untuk menentukan jumlah cluster terbaik. Algoritma K-Means sering digunakan karena keefektifan dan efisiensinya. Kelebihan K-Means adalah mudah dipelajari dan memiliki waktu proses komputasi yang relatif singkat. K-Means adalah salah satu metode clustering yang digunakan dalam data mining untuk mengelompokkan data menjadi beberapa kelompok berdasarkan titik pusat (centroid) [1].

Penentuan hasil cluster didasarkan pada jarak terdekat antara objek data. Sebelum menghitung jarak terdekat ini, jumlah cluster centroid harus ditentukan terlebih dahulu [2]. Selain itu K-Means memiliki ketepatan yang cukup baik dalam mengukur ukuran objek, sehingga algoritma ini cenderung lebih akurat dan efisien untuk memproses sejumlah besar objek [3]. Kelemahan utama dari algoritma K-Means adalah dalam menganalisis dan menentukan jumlah cluster (k) yang optimal untuk mengelompokkan data dalam sebuah dataset. Identifikasi jumlah cluster (k)

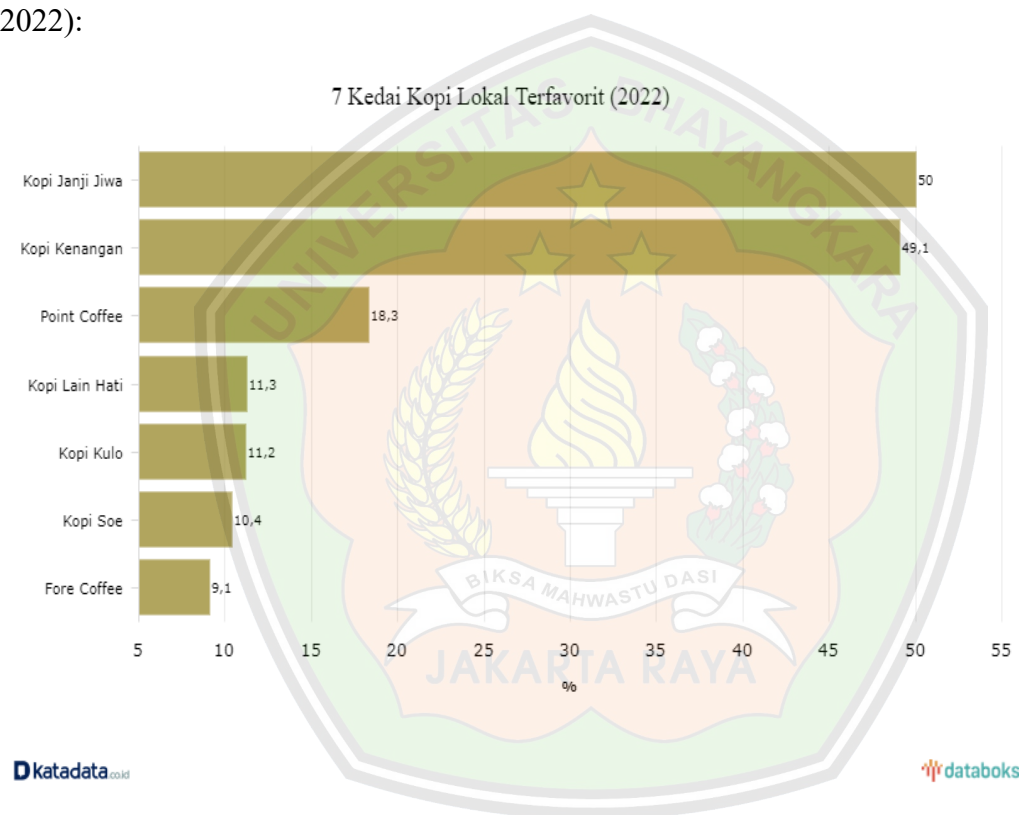
ini sangat penting dalam proses clustering dengan algoritma K-Means, karena hasil clustering sangat bergantung pada jumlah cluster yang ditentukan di awal. Jika jumlah cluster yang dipilih tidak tepat, maka hasil clustering tidak akan sesuai dengan harapan dan tidak akan menghasilkan informasi yang diperlukan oleh pengguna.

Penelitian sejenis yang juga menggunakan metode K-Means untuk pengelompokan data menjelaskan bahwa proses ini dimulai dengan menentukan jumlah cluster. Setelah itu jarak terdekat dengan pusat cluster dihitung, dimana data yang memiliki jarak terdekat dengan pusat cluster dianggap sebagai anggota cluster tersebut. Proses ini diulang hingga data tidak lagi berpindah antara cluster dengan tujuan meminimalkan fungsi objektif [4]. Penelitian berikutnya menunjukkan bahwa algoritma K-Means berhasil digunakan untuk mengidentifikasi pelanggan potensial. Berdasarkan hasil clustering dengan K-Means pelanggan potensial teridentifikasi pada cluster pertama (C1) yang memiliki rata-rata transaksi dan pengeluaran tertinggi. Dengan adanya peningkatan dalam pelayanan dan penjualan, termasuk respon yang cepat serta dukungan sumber daya yang baik, layanan terhadap pelanggan potensial dapat meningkat secara signifikan [5].

Dalam mendukung pertumbuhan dan kelangsungan suatu Perusahaan hubungan antara Perusahaan dengan pelanggan menjadi factor penting dalam mewujudkannya. Manajemen perusahaan harus mampu memahami pelanggan untuk memilih menu yang tepat dan meningkatkan pemahaman tentang kebutuhan masing-masing pelanggan, sehingga dapat menjaga loyalitas mereka. Salah satu hal yang penting dalam pengelolaan pelanggan adalah upaya untuk mempertahankan pelanggan yang ada. Bagi pelaku usaha di sektor *Food & Beverage*, menjaga pelanggan menjadi sangat penting karena banyaknya pesaing di industri ini. Untuk mencegah pelanggan berpindah karena alasan terkait menu, perusahaan perlu mengidentifikasi kelompok pelanggan yang berpotensi untuk membeli kembali. Dengan cara ini, perusahaan dapat melindungi pelanggan potensial melalui pelayanan yang unggul,

pemberian bonus, dan diskon khusus. Potongan harga yang diberikan kepada pelanggan bisa berupa voucher, atau *buy 1 get 1* dan sebagainya.

Hasil implementasi dari metode K-Means dapat digunakan pada kasus penelitian data mining untuk melihat pelanggan potensial yang bisa dilakukan sebagai inovasi bagi para pelaku usaha *Food & Beverage*. Untuk tetap bisa bersaing pada bidang *Food & Beverage* setiap kedai kopi harus selalu berusaha untuk memberikan nilai lebih terhadap pelanggannya. Berikut ini disediakan beberapa kedai kopi lokal terfavorit Masyarakat Indonesia menurut katadata.co.id (2022):



Gambar 1.1

Kedai Kopi Lokal Terfavorit Masyarakat Indonesia [6]

Menurut survei Jakpat, Kopi Janji Jiwa adalah kedai kopi lokal yang paling populer di kalangan masyarakat, dengan proporsi mencapai 50%. Mengusung slogan “kopi dari hati,” Kopi Janji Jiwa berhasil menarik perhatian kaum milenial berkat berbagai menu uniknya, seperti Black Pink, Es Susu Hojicha, dan Es Teh Cincau Pandan. Kedai kopi lokal favorit kedua adalah Kopi Kenangan, dengan proporsi 49,1%. Menu andalan dari kedai ini adalah Kopi

Kenangan Mantan, yaitu campuran kopi susu dengan rasa manis dari gula aren. Selain itu, kedai kopi lokal lainnya yang juga disukai masyarakat meliputi Point Coffee (18,3%), Kopi Lain Hati (11,3%), Kopi Kulo (11,2%), Kopi Soe (10,4%), dan Fore Coffee (9,1%).

TitikSetara Coffee Bekasi sebagai kedai kopi lokal yang saat ini masih bertahan harus memiliki kekuatan dari segi penjualannya. Untuk itu perlu dilakukan strategi penetapan menu yang berpotensi besar agar dapat mempertahankan serta mendapatkan pelanggan baru. Dalam penelitian ini data penjualan akan diolah untuk mencari atau menemukan menu yang potensial untuk direkomendasikan kepada pelanggan. Berikut table mengenai penjualan pada TitikSetara Coffee pada bulan Oktober 2024:

Tabel 1.1
Hasil Penjualan TitikSetara Coffee pada Bulan Oktober

Kode	Nama Menu	Harga Menu	Jumlah Transaksi	Total Pendapatan
A1	Mix Platter	25000	77	1.925.000
A2	Indomie Setara	15000	88	1.320.000
A3	Indomie Spesial Setara	20000	89	1.780.000
A4	Cireng	20000	16	320.000
A5	Friench Fries	20000	111	2.220.000
A6	Taichan	25000	49	1.225.000
A7	Kebab Mini	25000	56	1.400.000
A8	Spaghetti Bolognaise	28000	30	840.000
A9	Burger	25000	50	1.250.000
A10	Snack Bucket	30000	63	1.890.000
A11	Kocok Setara	23000	115	2.645.000
A12	Setara Aren	23000	72	1.656.000
A13	Kopi Setara	21000	188	3.948.000
A14	Titik Semangat	23000	234	5.382.000
A15	Coffee Latte	20000	114	2.280.000
A16	Cappucino	23000	176	4.048.000
A17	Lemon tea	20000	194	3.880.000
A18	On The Rock	16000	31	496.000
....				
A33	Black Tea	15000	56	840.000
Total			2759	62.191.000

Sumber: Data Penelitian (2024)

Pelaku usaha harus bisa melakukan inovasi atau menyediakan produk sesuai dengan apa yang diinginkan oleh pelanggannya seperti meningkatkan pelayanannya, promosi, citra

perusahaan, variasi menu dan segala yang dibutuhkan oleh pelanggan. Namun dalam menjalankan sebuah usaha pastinya akan selalu ada yang namanya hambatan atau masalah. Masalah yang kerap terjadi pada bidang usaha kedai kopi adalah cara mengetahui pelanggan terkait menu mana yang paling memiliki potensi untuk para pelanggan agar membeli kembali di sebuah kedai kopi. Untuk itu, teknik yang akan diterapkan dalam karya ilmiah ini adalah data mining untuk pengelompokan dengan algoritma K-Means Clustering. Metode K-Means Clustering digunakan untuk mengelompokkan cabang yang produktif berdasarkan kriteria yang telah ditentukan. Clustering dalam data mining seperti yang diketahui berfungsi untuk menganalisis objek berdasarkan kesamaan sifat antar objek tersebut. Penelitian ini akan mencakup perancangan perangkat lunak menggunakan Microsoft Excel dan RapidMiner yang diharapkan dapat menyediakan solusi untuk masalah yang diidentifikasi.

Berdasarkan permasalahan di atas, maka peneliti tertarik untuk meneliti lebih lanjut dalam skripsi atau penelitian yang berjudul **“Penentuan Menu Potensial Menggunakan Algoritma K-Means Pada Titiksetara Coffee Bekasi”**

1.2 Identifikasi Masalah

Berdasarkan latar belakang di atas dapat diidentifikasi masalah pada penelitian ini yaitu sebagai berikut:

1. Sulitnya melakukan penentuan pada suatu Menu Potensial
2. Keterbatasan Pengetahuan Keinginan Pelanggan sehingga sulit untuk mempertahankan loyalitas pelanggan dan mengantisipasi perpindahan pelanggan ke pesaing.
3. Pemahaman yang kurang mendalam terhadap preferensi pelanggan terkait menu potensial.
4. Sulitnya mengukur tingkat kepuasan pelanggan terkait menu mana yang paling diminati atau berpotensi untuk dibeli kembali.

1.3 Rumusan Masalah

Berdasarkan latar belakang di atas, dapat dirumuskan beberapa masalah, yaitu:

1. Bagaimana hasil kelompok data dapat diinterpretasikan, dengan mempertimbangkan adanya ketidakakuratan dalam hasil clustering?
2. Bagaimana perhitungan algoritma K-Means dapat memberikan hasil yang akurat untuk menentukan menu potensial?
3. Bagaimana hasil analisis menu potensial dapat diinterpretasikan dengan mempertimbangkan adanya potensi ketidakakuratan dalam data?

1.4 Tujuan dan Manfaat

Adapun beberapa tujuan yang diharapkan pada pembuatan skripsi ini adalah sebagai berikut:

1. Mengelompokkan data menu yang bertransaksi dengan algoritma *K-Means*.
2. Mendapatkan data menu yang potensial dengan menggunakan algoritma *K-Means*.
3. Mengaplikasikan Microsoft Excel dan RapidMiner untuk mendapatkan menu potensial dengan menggunakan algoritma *K-Means*.

Adapun manfaat dari penelitian ini yaitu sebagai berikut:

1. Menyederhanakan proses analisis data besar.
2. Memanfaatkan aplikasi Microsoft Excel dan RapidMiner untuk mengidentifikasi menu yang berpotensi.
3. Diharapkan dapat mendukung pengelola dalam menentukan menu yang efektif, sehingga dapat lebih tepat sasaran dalam mempertahankan pasar di tengah persaingan yang semakin intens.
4. Menyediakan laporan mengenai menu potensial secara cepat dan akurat.

1.5 Batasan Masalah

Pada penyusunan skripsi ini permasalahannya dibatasi pada :

1. Data yang dianalisis adalah sampel transaksi konsumen dari bulan Oktober 2023 hingga Januari 2024
2. Pengelompokan dilakukan dengan menggunakan algoritma K-Means.
3. Hanya data mengenai menu-menu yang dianggap potensial yang akan ditampilkan.
4. Data yang diproses dalam data mining disimpan dalam format Excel dan dianalisis menggunakan aplikasi RapidMiner.
5. Penulisan ini tidak mencakup pembahasan mengenai sistem pendukung keputusan.

1.6 Sistematika Penulisan

Penulisan sistematika bertujuan untuk mempermudah pembahasan dalam suatu penelitian. Sistematika penulisan dalam penelitian ini adalah sebagai berikut:

BAB I PENDAHULUAN

Bab ini mencakup pembahasan mengenai latar belakang masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, serta sistematika penulisan yang merinci struktur bab-bab dalam skripsi ini.

BAB II LANDASAN TEORI

Bab ini membahas landasan teori terkait dengan data mining (K-Means Clustering), menu potensial, Excel, RapidMiner, dan Davies-Bouldin Index. Selain itu, bab ini juga mencakup kerangka pemikiran dan hipotesis yang secara ringkas menjelaskan hubungan antara variabel independen dan variabel dependen yang akan diteliti.

BAB III METODE PENELITIAN

Bab ini akan membahas desain penelitian, tahapan penelitian, jenis dan metode pengambilan data, serta teknik analisis dan proses pengumpulan data yang dilakukan oleh peneliti.

BAB IV HASIL DAN PEMBAHASAN

Bab ini akan menguraikan hasil klasifikasi dalam penentuan menu potensial di Titiksetara Coffee, berdasarkan penelitian yang telah dilakukan menggunakan metode yang telah ditetapkan.

BAB V PENUTUP

Bab ini berisi kesimpulan dari penelitian berdasarkan rumusan masalah dan tujuan penelitian terkait penentuan menu potensial menggunakan algoritma K-Means serta berisi saran yang dapat digunakan untuk pengembangan penelitian lainnya.

