

## **Bab II**

### **Tinjauan Pustaka**

#### **2.1 Usaha Mikro, Kecil, dan Menengah (UMKM)**

UMKM merupakan usaha yang didirikan oleh perorangan dan memiliki karakteristik yang berbeda dengan bidang usaha lainnya. Karakteristik UMKM antara lain struktur organisasi bisnis yang sederhana, milik warga negara Indonesia, aktivitas banyak berupa kegiatan non-formal, memiliki manajemen bisnis yang fleksibel, dan teknologi yang digunakan sederhana. UMKM dinilai memiliki potensi mengentas kemiskinan karena banyaknya peluang penyerapan tenaga kerja dan membantu pemerataan ekonomi rakyat setempat. UMKM bertujuan untuk menguatkan perekonomian nasional, menciptakan lapangan kerja, pemerataan pendapatan masyarakat, pemerataan pembangunan daerah, menciptakan produk unggulan desa, dapat mengentas kemiskinan, meningkatkan kesejahteraan masyarakat setempat dan mendukung pertumbuhan ekonomi. Hal tersebut sesuai dengan pendapat Azzahra dan Wibawa (2021) yang menyebutkan bahwa UMKM memiliki kontribusi dalam perekonomian di Indonesia.

Banyaknya jumlah UMKM dapat memperbesar kesempatan penyerapan tenaga kerja dan penggerak ekonomi di pedesaan. Pertumbuhan ekonomi pedesaan meliputi peningkatan sumber daya manusia, adanya kelembagaan masyarakat setempat, adanya fasilitas produksi, dan tersedianya modal. Selain itu peran UMKM menjadi salah satu sumber devisa ekspor negara serta meningkatkan pendapatan masyarakat. Pertumbuhan ekonomi yang baik dapat diukur dengan tingkat penyerapan kerja dan menurunnya angka pengangguran dalam suatu negara. Noide (2009) menyatakan bahwa UMKM terbukti memiliki keterlibatan besar dalam meningkatkan kesejahteraan masyarakat. Pertumbuhan ekonomi negara dipengaruhi oleh berbagai faktor misalnya masalah kemiskinan, pengangguran, sumber daya alam, ketimpangan pendapatan, arus urbanisasi, modal, sumber daya manusia, jumlah penduduk, aspek sosial budaya, teknologi, dan kekayaan alam.

Pembedaan antara usaha mikro, usaha kecil, dan usaha menengah berdasarkan nilai aset awal, omset rata-rata per tahun, dan jumlah pekerja tetap. Menurut Lilis (2016) manajemen sumber daya manusia bertujuan untuk meningkatkan efektivitas bagi semua karyawan dengan perencanaan, pengembangan, perbaikan, serta melakukan evaluasi terhadap kinerja karyawan. Manajemen sumber daya manusia adalah faktor penting yang menjadi konsentrasi pemilik perusahaan dan UMKM. Hal tersebut karena manajemen sumber daya manusia memiliki peran penting dalam memajukan suatu organisasi, dapat mengoptimalkan kinerja dan efisiensi karyawan untuk mencapai tujuan organisasi. Manajemen sumber daya manusia berhubungan dengan semua aktivitas ketenagakerjaan misalnya pendidikan karyawan, kesejahteraan karyawan, perekrutan karyawan, administrasi dan pelatihan karyawan. Sumber daya manusia di UMKM biasanya mengambil dari daerah sekitar tempat berdirinya usaha, sehingga kualitas sumber daya manusia pada UMKM masih perlu ditingkatkan. Kualitas sumber daya manusia dapat dilihat dengan kualitas produk yang dihasilkan karyawan. Peningkatan kinerja sumber daya manusia bisa dengan mengadakan pelatihan bagi karyawan untuk memahami proses kerja yang seharusnya dijalankan dan menambah pemahaman terhadap bidang kerja.

Manajemen sumber daya manusia dapat diukur dengan pendidikan formal, pengetahuan, dan keterampilan karyawan. Sumber daya manusia menjadi salah satu faktor yang dapat mempengaruhi perkembangan dan memperbesar peluang keberhasilan bisnis UMKM. Adriana dkk (2010) menyatakan bahwa kompetensi SDM mempunyai hubungan yang signifikan dengan kinerja usaha kecil menengah yang dianalisis dengan menggunakan analisis korelasi dan determinasi. Manajemen sumber daya manusia harus dipantau terus menerus karena perilaku dan sifat manusia yang terus mengalami perubahan.

## **2.2 Peran UMKM di Pedesaan**

Pedesaan memiliki alam yang masih terjaga dan sangat subur sehingga potensi untuk mendapatkan produk dengan kualitas yang baik sangatlah tinggi. Apalagi jika pengembangan UMKM desa yang dilakukan dengan cara penerapan inovasi yang kreatif tentunya akan menciptakan ekonomi keberlanjutan. UMKM

merupakan salah satu program di bawah naungan langsung Kementerian Koperasi dan UKM. Beberapa jenis bidang UMKM desa antara lain peternakan, pertanian dan perkebunan serta kerajinan tangan. UMKM desa selain berkontribusi dalam peningkatan PDB, UMKM desa juga memiliki potensi yang bagus di masa mendatang. Bahkan pemerintah memberikan dukungan kepada UMKM di pedesaan dengan berbagai program pembinaan dan pengembangan yang diharapkan UMKM desa dapat bersaing dengan jenis usaha lain hingga kancah internasional.

Namun tidak menutup kenyataan pada saat ini perekonomian mengalami situasi yang kurang menyenangkan karena adanya pandemi yang melanda seluruh dunia. Tidak hanya pemulihan ekonomi yang harus dihadapi oleh banyak orang namun juga berbagai tantangan baru dengan penyesuaian yang baru pula. Diperlukan pengkajian ulang keefektifan dan efisiensi pemanfaatan sumber daya agar terhindar dari kerugian di masa-masa yang sulit. Pemanfaatan sumber daya harus jelas dan terukur supaya ketersediaan sumber daya dapat dikelola dengan optimal.

### **2.3 Desa Petungsewu**

Menurut Badan Pusat Statistik (BPS) Desa Petungsewu Kecamatan Dau merupakan salah satu desa di Kabupaten Malang. Jumlah penduduk yang ada Desa Petungsewu adalah 3.501 jiwa dengan kepadatan penduduk yaitu 1.006,03 per km<sup>2</sup>. Jumlah penduduk Desa Petungsewu cenderung meningkat hal ini dikarenakan tingkat kelahiran yang lebih besar daripada kematian serta penduduk yang masuk desa lebih besar dari penduduk yang keluar desa. Jarak dari pusat kota menuju Desa Petungsewu yakni kurang lebih 8 km menggunakan kendaraan bermotor. Sebagian besar penduduk Desa Petungsewu bekerja sebagai petani. Sedangkan untuk komoditi pertanian dari wilayah ini yaitu cabai, jeruk dan kacang.

### **2.4 *Data mining***

*Data mining* merupakan proses pengolahan data berskala besar untuk menganalisis karakteristik dalam *database* sehingga dapat dipahami manusia. *Data mining* merupakan proses semi otomatis yang menggunakan teknik statistik,

matematika, kecerdasan buatan, dan *machine learning* untuk mengesktrasi dan mengidentifikasi informasi. *Data mining* juga dapat diartikan sebagai metode analisis yang menggunakan satu atau lebih teknik *machine learning* dengan tujuan untuk mengambil informasi dengan cepat dan otomatis. Informasi tersebut dapat dimanfaatkan untuk berbagai kepentingan misalnya optimalisasi, prediksi, mendeskripsikan suatu *database* dan masih banyak lagi. Menurut Tyas dkk (2015) *data mining* merupakan kegiatan yang meliputi pengumpulan dan penggunaan data terdahulu untuk mencari pola dan hubungan dalam data yang besar. *Data mining* dapat menjawab pertanyaan-pertanyaan bisnis dengan cepat dan lebih murah jika dibandingkan dengan cara manual. Larose (2005) membagi *data mining* menjadi beberapa kelompok berdasarkan kegunaannya antara lain:

1. Deskripsi

Deskripsi (*description*) digunakan untuk menggambarkan pola dan kecenderungan dalam data. Hal tersebut karena analisis dilakukan secara sederhana. Deskripsi pola kecenderungan memberikan sebuah penjabaran untuk suatu kecenderungan.

2. Estimasi

Estimasi (*estimation*) bertujuan untuk menemukan kesamaan karakteristik dalam suatu kelas yang berbentuk numerik. Pembuatan model menggunakan rekaman lengkap dari *variable target* sebagai nilai prediksi. Metode dalam estimasi yaitu: *Support Vector Machine*, *Linear Regression*, *Deep Learning*, *Neural Network*, dan *Generalized Liner Model*.

3. Prediksi

Prediksi (*forecasting*) bertujuan untuk memprediksi nilai yang akan muncul dimasa mendatang. Metode dalam prediksi dapat menggunakan metode yang ada dalam klasifikasi dan estimasi. Metode tersebut antara lain dalam prediksi seperti *Support Vector Machine*, *Linear Regression*, *Deep Learning*, *Neural Network*, dan *Generalized Liner Model*.

#### 4. Klasifikasi

Klasifikasi (*classification*) bertujuan untuk menemukan kesamaan karakteristik dalam suatu kelas. Metode ini sering digunakan untuk mengestimasi atau memprediksi kelas dari suatu objek yang belum diketahui labelnya. Metode klasifikasi antara lain *Linear Discriminant Analysis*, *Decision Tree (ID3, C4.5)*, *Naive Bayes*, dan *Logistic Regression*.

#### 5. Clustering

*Clustering* bertujuan untuk mengelompokkan dan membentuk kelas objek yang memiliki kedekatan. *Clustering* hampir sama dengan klasifikasi, namun dalam *clustering* belum ada target kelas yang ditentukan sebelum pengelompokan. Metode *clustering* yaitu *K-Means* dan *K-Medoids*.

#### 6. Asosiasi

Asosiasi (*association*) bertujuan untuk mencari atribut yang muncul bersamaan dalam satu waktu. Asosiasi biasanya sering disebut *market basket analysis* karena sering digunakan untuk menggabungkan suatu *item* yang sering dibeli. Metode asosiasi antara lain Algoritma apriori, *Chi Square*, dan *FP-Growth*.

### 2.5 Asosiasi

Asosiasi adalah salah satu metode dalam *data mining*. Asosiasi bertujuan untuk melakukan perhitungan guna mencari keterkaitan antar data yang muncul dalam waktu yang bersamaan. Asosiasi menganalisis banyaknya frekuensi dan analisisnya menggunakan pola *if-then*. Menurut Yanto dan Khoiriah (2015) aturan asosiasi berhubungan dengan dua indikator yaitu *support* dan *confidence* yang didapat melalui perhitungan hubungan antar item. *Support* merupakan nilai pendukung yang menunjukkan besarnya tingkat dominasi suatu item set dari keseluruhan data. *Confidence* atau nilai keyakinan merupakan nilai yang menggambarkan hubungan antar 2 item.

Asosiasi menggunakan algoritma apriori menghasilkan aturan asosiasi. Aturan asosiasi yang dapat bermanfaat untuk mengetahui item yang dibeli secara bersamaan dan membantu pengambilan keputusan bisnis atau kepentingan

promosi, karena dengan mengetahui keterkaitan antar item dapat memudahkan bagian promosi untuk membuat kebijakan promosi yang akan dilakukan misalnya dengan menggabungkan beberapa item guna meningkatkan penjualan. Perhitungan nilai *support* dan *confidence* dilakukan dengan rumus berikut ini:

$$Support = \frac{\Sigma \text{ item yang dibeli sekaligus}}{\Sigma \text{ seluruh transaksi}} \times 100\% \quad \dots\dots\dots (2.1)$$

$$Confidence = \frac{\Sigma \text{ item yang dibeli sekaligus}}{\Sigma \text{ transaksi pada antecedent}} \times 100\% \quad \dots\dots\dots (2.2)$$

Menurut Kusrini (2009) menyebutkan bahwa dasar analisis aturan asosiasi dibagi menjadi dua tahapan yaitu:

1. Analisis Pola Frekuensi Tinggi

Tahap awal ini dilakukan pencarian kombinasi item yang memenuhi syarat minimum nilai *support* dalam suatu *data base*. Nilai *support* merupakan penunjang persentase dari item yang terdapat pada *data base*. Berikut merupakan rumus nilai *support* yang hanya memiliki 1 jenis item:

$$Support(A) = \frac{\Sigma \text{ Jumlah Transaksi Mengandung A}}{\Sigma \text{ Total Transaksi}} \quad (2.3)$$

Rumus nilai *support* untuk 2 item adalah sebagai berikut:

$$Support(A, B) = P(A \cap B)$$

$$Support(AB) = \frac{\Sigma \text{ Jumlah Transaksi Mengandung A dan B}}{\Sigma \text{ Transaksi}} \quad (2.4)$$

2. Pembentukan Aturan Asosiasi

Setelah pola frekuensi tinggi diketahui hasilnya, selanjutnya dilakukan pencarian aturan asosiasi yang memenuhi syarat minimal *confidence* aturan asosiasi A U B. Rumus nilai *confidence* adalah sebagai berikut:

$$Confidence P(B|A) = \frac{\Sigma \text{ Jumlah Transaksi Mengandung A dan B}}{\Sigma \text{ Transaksi Mengandung A}} \quad (2.5)$$

## 2.6 Clustering

*Clustering* merupakan proses untuk mengelompokkan objek pada suatu kelompok tertentu. *Clustering* bertujuan untuk mendeskripsikan data dan menyelesaikan permasalahan dalam mengelompokkan data serta dapat memudahkan dalam menganalisis persamaan karakteristik setiap data. Menurut Vuldari dalam Pratomo dkk (2019) *clustering* digunakan untuk mengidentifikasi kelompok berdasarkan kemiripan atribut. *Clustering* dalam *data mining* memiliki dua jenis metode untuk mengelompokkan data yaitu *clustering* hierarki dan *partitioning* atau non-hierarki. Ukuran kemiripan data dihitung dengan jarak. Jarak yang baik dalam satu klaster yang memiliki jarak paling dekat dan sebaliknya jarak dengan klaster lain sebaiknya lebih jauh. Dalam penelitian ini menggunakan *clustering* non-hierarki dengan metode *k-means*.

Menurut Waworuntu dan Amin (2018) *k-means* adalah metode yang paling banyak digunakan untuk mengklasterkan data karena mudah diimplementasikan, sederhana, dan mampu mengklasterkan data dengan ukuran yang besar. *K-Means* merupakan algoritma yang mengelompokkan data set ke dalam beberapa kelompok sesuai dengan *k cluster* yang telah ditentukan. Metode *k-means* memiliki basis titik pusat atau *centroid*. Algoritma *k-means* dapat meminimal variasi dalam suatu kelompok serta memaksimalkan variasi dengan kelompok lainnya. Hasil *clustering* data kemudian dilakukan analisis menggunakan visualisasi data untuk melihat kemiripan atau kedekatan suatu objek sehingga dapat mengambil kebijakan untuk tiap klasternya. Untuk menghitung jarak kedekatan dengan *centroid* menggunakan rumus berikut:

$$De = \sqrt{(x1 - s1)^2 + (y1-t1)^2} \dots\dots\dots (2.6)$$

Keterangan:

x,y = koordinat objek

s,t = koordinat *centroid*

i = banyaknya objek

## 2.7 Penggunaan *Software Rapid Miner*

*Software Rapid Miner* memiliki berbagai fungsi khususnya dalam data mining. Dalam penelitian ini *software Rapid Miner* digunakan untuk membuat

kelompok masalah yang dialami UMKM di Desa Petungsewu. Berikut operator dalam *software* Rapid Miner yang digunakan dalam penelitian ini:

### **2.7.1 Read CSV**

Penggunaan fitur ini bertujuan untuk membaca *input* data berupa *file* dengan bentuk *csv*. Operator ini merupakan awal dimulainya suatu model dalam *clustering*. Jika tidak menggunakan fitur ini maka data *input* tidak dapat dibaca dan tidak dapat melanjutkan pada proses selanjutnya.

### **2.7.2 Nominal to Text**

Penggunaan operator *nominal to text* bertujuan untuk mengubah jenis atribut berupa nominal menjadi bentuk teks atau nilai string yang sesuai. Fitur ini memberikan beberapa opsi filter yang dapat digunakan pada fitur ini. Misalnya filter all untuk mengubah semua inputan data menjadi *text*.

### **2.7.3 Process Document from Data**

Tujuan penggunaan operator ini karena data berupa data *text* masalah UMKM. Pemrosesan data tidak dapat dilakukan jika tidak menggunakan fitur ini. *Output* dari fitur ini berupa TF-IDF. Selanjutnya hasil TF-IDF akan dilanjutkan pada proses selanjutnya.

### **2.7.4 Tokenize**

Operator ini merupakan proses pemisahan data menjadi tiap-tiap item set permasalahan UMKM. Fitur ini digunakan agar memudahkan proses *clustering* untuk mengelompokkan data berdasarkan hasil pemisahan setiap permasalahan UMKM.

### **2.7.5 Filter Stopword**

*Filter stopwords* digunakan untuk mengurangi kata yang tidak penting dan kata sambung. Hal tersebut dilakukan untuk menunjang kelancaran proses *clustering*. Karena jika kata sambung tidak dibuang akan mempengaruhi hasil *cluster*.



### **2.7.6 Multiply**

Operator *multiply* dapat membuat salinan objek-objek sebelumnya. Sehingga dapat melakukan proses secara bersamaan tanpa harus membuat objek satu persatu. Operator ini berfungsi untuk menyalin dan menjalankan model sebelumnya pada operator selanjutnya. Operator *multiply* dapat mempercepat pembuatan model.

### **2.7.7 Clustering**

Penggunaan fitur *clustering* bertujuan untuk mengelompokkan data. Operator ini membuat *input* data menjadi kelas-kelas sesuai dengan karakteristik kedekatan data. *Clustering* dalam penelitian ini menggunakan nilai  $k = 2$ ,  $k = 3$ ,  $k=4$ ,  $k= 5$ , dan  $k= 6$ .

### **2.7.8 Cluster Distance Performance**

Fitur ini digunakan untuk menentukan nilai  $k$  yang paling optimal dan sekaligus digunakan sebagai tahap evaluasi *cluster*. Evaluasi *cluster* dilakukan untuk mengetahui kualitas hasil *clustering*. Pada penelitian ini evaluasi hasil *cluster* dilakukan dengan menggunakan metode *Davies Bouldin Index*. Evaluasi dapat dilakukan dengan memilih *Davies Bouldin* pada *main criterion*.

## **2.8 Penelitian Terdahulu**

Penelitian dilakukan oleh Badrul (2016) dengan judul “Algoritma Asosiasi dengan Algoritma Apriori Untuk Analisa Data Penjualan”. Objek amatan dalam penelitian tersebut adalah data transaksi yang telah dihitung nilai *support* dan *confidentnya*. hasil dari penelitian tersebut menunjukkan bahwa algoritma apriori dapat diterapkan untuk mendukung strategi marketing sehingga diharapkan efektivitas pelaksanaan strategi akan dapat semakin ditingkatkan. Jika dibandingkan hasil dari perangkat lunak sama dengan perhitungan manual.

Penelitian berikutnya dilaksanakan oleh Sibuea dan Sapta (2017) dengan judul “Pemetaan Siswa Berprestasi Menggunakan Metode K-Means *Clustering*”. Objek amatan dalam penelitian tersebut adalah sampel rapot peserta didik SMK Yapim Simpang Kawat kelas X AK2 tahun ajaran 2016/2017. Hasil dari penelitian

tersebut menunjukkan bahwa metode k-means dapat mengelompokkan prestasi siswa tinggi, menengah, dan cukup.

Penelitian selanjutnya dilakukan oleh Milfa, Suardi, dan Elpandri (2018). Judul penelitian tersebut yaitu “Data Mining Untuk Analisis Pola Pemilihan Menu Pada Penang Corner Cafe Dan Resto Menggunakan Algoritma Apriori”. Tujuan penelitian tersebut untuk menentukan strategi yang tepat guna meningkatkan penjualan. Hasil penelitian tersebut dapat diketahui kombinasi menu yang paling sering dibeli oleh konsumen. Sehingga menu yang paling sering dibeli dijadikan satu menu paket untuk meningkatkan promosi *cafe*.

Penelitian berikutnya dilaksanakan oleh Jumadi (2018) dengan judul “Peningkatan Hasil Evaluasi *Clustering* Davies-Bouldin Index dengan Penentuan Titik Pusat *Cluster* Awal Algoritma *K-Means*”. Tujuan dilakukannya penelitian tersebut untuk menentukan *centroid* guna meningkatkan hasil evaluasi *cluster* berdasarkan nilai yang paling kecil pada SSE dan membandingkan nilai DBI. Hasil penelitian tersebut yaitu dapat diketahui nilai rata-rata DBI untuk *dataset* 1, *dataset* 2, dan *dataset* 3. Sehingga dapat disimpulkan bahwa untuk mendapatkan nilai SSE yang paling optimal dapat dilakukan setelah mengetahui hasil *clustering*.

Berdasarkan beberapa penelitian terdahulu pada paragraf sebelumnya dengan penelitian ini memiliki persamaan metode penelitian yaitu menggunakan algoritma *k-means* dan asosiasi *data mining*. Dalam penelitian ini penggunaan *data mining* untuk menganalisis dan menggali informasi dari kumpulan masalah yang dialami UMKM di Desa Petungsewu.