

BAB II

TINJAUAN PUSTAKA

II.1 Penelitian Terdahulu

Dalam melakukan penelitian ini penulis menggunakan beberapa contoh referensi dari hasil penelitian-penelitian sebelumnya yang berkaitan dengan latar belakang masalah pada penelitian ini, adapun penelitian yang berhubungan dengan skripsi ini yaitu:

1. Penelitian yang dilakukan oleh Asroni, Adrian Ronald, Mei 2015. “Penerapan Metode K-Means Untuk *Clustering* Mahasiswa Berdasarkan Nilai Akademik Dengan Weka *Interface* Studi Kasus Pada Jurusan Teknik Informatika UMM Magelang”. Jurnal Ilmiah Semesta Teknik Vol. 18, No. 1. Penelitian ini dilakukan untuk menguji data yang telah ada di *data warehouse* Universitas Muhammadiyah Magelang untuk memudahkan dalam mencari 5 orang mahasiswa pada jurusan Teknik Informatika dalam melakukan penyeleksian untuk mengikuti lomba yaitu kompetisi event Cyberjawara yang diselenggarakan oleh Indonesia *Security Incident Response Team on Internet Infrastructure* (ID SIRTII) Kementerian Komunikasi dan Informatika RI.
2. Penelitian selanjutnya dilakukan oleh Surmayanti, dkk, 2015 “Penerapan Analisis *Clustering* Pada Penjualan Komputer Dengan Perancangan Aplikasi *Data Mining* Menggunakan *Algoritma K-Means* (Studi Kasus Toko Tri Buana Komputer Kota Solok)”. Prosiding seminar Ilmiah

Nasional Teknologi Komputer (SENATKOM), ISSN : 2460-2690, Vol.1. Penelitian ini bertujuan untuk membangun sebuah perangkat lunak untuk memprediksi presentase suatu produk laptop, dimana perangkat lunak ini dapat mudah dioperasikan dan dipahami oleh pihak toko Tri Buwana Komputer dalam menganalisa dan mengambil keputusan serta dapat mengoptimalkan dan mengefektifkan kinerja toko tersebut.

3. Selanjutnya penelitian yang dilakukan oleh Gustientiedina, dkk, 2018 “Penerapan *Algoritma K-Means* Untuk *Clustering* Data Obat-Obatan Pada RSUD Pekanbaru”. Jurnal Nasional Teknologi dan Sistem Informasi, Vol. 05, No. 01. Penelitian yang akan dilakukan bertujuan untuk mengetahui *clusterisasi* data obat dengan teknik data *mining*, dimana pengolahan data menggunakan *Algoritma K-Means*. Batasan masalah dari penelitian yang dilakukan adalah data yang akan digunakan yaitu data obat-obatan dari Rumah Sakit Umum Daerah Kota Pekanbaru. Analisis dilakukan dengan menggunakan *Algoritma K-Means*. *Output* yang dihasilkan adalah kelompok obat-obatan yang pemakaiannya rendah, sedang, dan tinggi.
4. Penelitian selanjutnya adalah dilakukan oleh Larasati Sibuea Fitri, dkk 2017 tentang “Pemetaan Siswa Berprestasi Menggunakan Metode *K-Means Clustering*”. Jurnal Teknologi dan Informasi (JURTEKSI) ISSN : 2550-0201 Vol.IV, No.1. Penelitian ini bertujuan untuk dapat mengelompokkan data terkait nilai siswa menurut prestasi masing-masing siswa tersebut apakah prestasi tinggi, menengah, atau cukup. Penelitian yang dilakukan dengan menggunakan metode *Algoritma K-Means* dengan

adanya *software Weka* dalam penelitian ini maka keakuratan data akan cukup baik terhadap permasalahan yang terjadi terkait dengan prestasi siswa dengan hasil 70% dapat mengenali data pada data yang digunakan sebagai sampel.

Penelitian diatas hanyalah sebagai contoh yang menjadi acuan penulis dalam melakukan penelitian, hanya saja topik yang menjadi pembahasan penulis yaitu menerapkan metode *K-Means clustering* dalam pengelompokan pemakaian laptop dikalangan mahasiswa Universitas Potensi Utama Medan. dalam hal ini penulis menerapkan metode tersebut menggunakan *software Weka* dan dilakukan perhitungan dalam *Microsoft Excel 2010*.

II.2. Landasan Teori

II.2.1. Analisa

Analisa berasal dari kata Yunani kuno "*analisis*" yang berarti melepaskan. *Analisis* terdiri dari dua suku kata yaitu "*ana*" yang berarti kembali dan "*luein*" berarti melepas. Sehingga pengertian Analisa yaitu sekumpulan kegiatan, aktivitas dan proses yang saling berkaitan untuk memecahkan masalah atau memecahkan komponen menjadi lebih detail dan digabungkan kembali lalu ditarik kesimpulan. Analisa juga dapat diartikan sebagai sebuah penyelidikan terhadap suatu peristiwa dengan tujuan mengetahui keadaan yang sebenarnya terjadi.

Bentuk dari kegiatan analisa salah satunya yaitu merangkum data mentah menjadi sebuah informasi yang bisa disampaikan ke khalayak. Segala macam

bentuk analisis menggambarkan pola-pola konsisten di dalam data, sehingga hasil analisa dapat dipelajari dan diterjemahkan dengan singkat dan penuh makna. Analisa data biasanya dilakukan meliputi kegiatan seperti mengorganisasikan data, mengelompokkan data, mengklasifikasi data, memaparkan data dan menarik kesimpulan dari keseluruhan data tersebut. Mengorganisasikan data yaitu mengatur data-data yang telah diperoleh peneliti selama kegiatan penelitian sedang berlangsung, sedangkan pengelompokan data yaitu mengelompokkan data mana yang hendak dipakai dan data mana yang tidak dipakai. Mengklasifikasi data juga berarti mengelas-ngelaskan data sesuai kebutuhan. Memaparkan data yaitu menyampaikan hasil proses analisis data dan menarik kesimpulan atas informasi dari data yang telah disampaikan.

II.2.2. Data Mining

Data Mining merupakan suatu proses ekstraksi atau penggalian data dan informasi yang besar, yang belum diketahui sebelumnya, namun dapat dipahami dan berguna dari *database* yang besar serta digunakan untuk membuat keputusan bisnis yang sangat penting. Teknik data mining digunakan untuk memeriksa basis data berukuran besar sebagai cara untuk menemukan pola baru dan berguna.

Terdapat beberapa istilah dalam proses *Data Mining* yaitu:

1. *Data Selection*

Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam *knowledge discovery in database* (KDD) dimulai. Data hasil seleksi yang digunakan untuk proses

data mining, disimpan dalam suatu berkas, terpisah dari basis data operasional.

2. *Pre-processing/cleaning*

Sebelum proses data mining dapat dilaksanakan, perlu dilakukan proses *cleaning* pada data yang menjadi fokus KDD. Proses *cleaning* mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data.

3. *Data Transformation*

Dimana data berubah atau bersatu menjadi bentuk yang tepat untuk menambang dengan ringkasan performa atau operasi agresif. *Coding* adalah proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses data mining. Proses *coding* dalam KDD merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data.

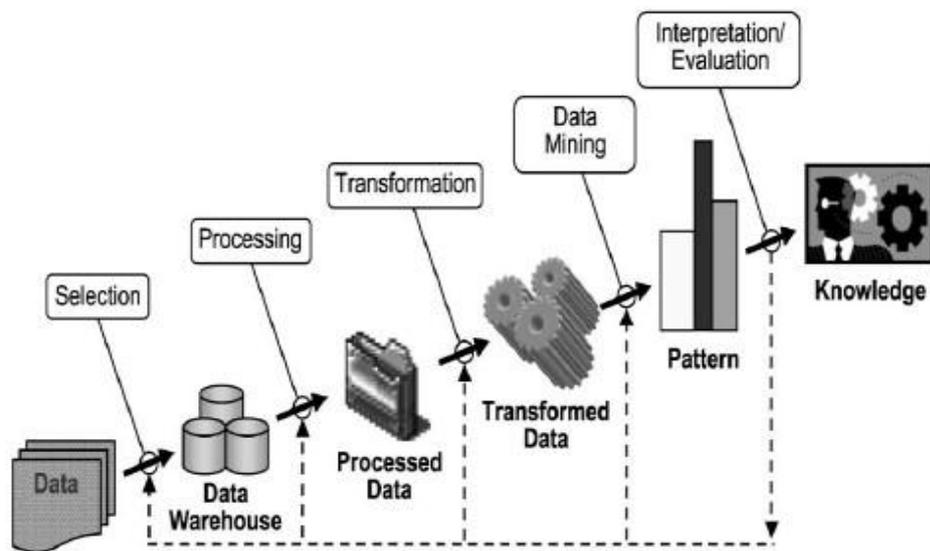
4. *Data Mining*

Data mining adalah proses mencari pola atau informasi yang menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma dalam data mining sangat bervariasi. Pemilihan algoritma atau metode yang tepat sangat bergantung pada tujuan dan proses KDD secara keseluruhan.

5. *Interpretation / evaluation*

Untuk mengidentifikasi pola yang benar-benar menarik yang mewakili pengetahuan berdasarkan atas beberapa tindakan yang menarik. Pola

informasi yang dihasilkan dari proses data mining perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini merupakan bagian dari proses KDD yang disebut *interpretation*. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesis yang ada sebelumnya.



Gambar II.1 Tahapan proses dalam *Data Mining*

Data mining mempunyai beberapa fungsi yang penting untuk mendapatkan informasi yang berguna serta meningkatkan pengetahuan bagi pengguna. Berikut ini adalah beberapa fungsi dari *data mining* :

1. *Classification*, yaitu menyimpulkan definisi-definisi karakteristik sebuah grup.
2. *Clustering*, yaitu mengelompokkan data-data yang mempunyai karakteristik yang sama dan mengelompokkan data yang lainnya yang memiliki karakteristik yang berbeda.

3. *Association*, yaitu mengidentifikasi hubungan antara kejadian-kejadian yang terjadi pada suatu waktu.
4. *Sequencing*, hampir sama dengan *association*, *sequencing* mengidentifikasi hubungan-hubungan yang berbeda pada suatu periode waktu tertentu, seperti pelanggan yang mengunjungi supermarket yang berulang-ulang.
5. *Forecasting*, memperkirakan nilai pada masa yang akan datang berdasarkan pola-pola dengan sekumpulan data yang besar, seperti peramalan permintaan pasar.

II.2.3. *K-Means Clustering*

K-Means merupakan salah satu algoritma *clustering*. Tujuan algoritma ini yaitu untuk membagi data menjadi beberapa kelompok. Sedangkan *clustering* merupakan pengelompokan *record*, pengamatan atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan.

Metode ***K-Means Clustering*** merupakan salah satu metode data *clustering non-hirarki* berusaha mengelompokkan data dalam bentuk satu atau lebih *cluster*/kelompok. Data-data yang memiliki karakteristik yang berbeda dikelompokkan dengan *cluster*/kelompok yang lain sehingga data yang berada dalam satu *cluster*/kelompok memiliki tingkat variasi yang kecil.

Langkah-langkah dalam melakukan *clustering* dengan metode *k-means* adalah sebagai berikut:

1. Menentukan jumlah *cluster* k ,

2. Tentukan *centroid* awal,

Centroid (titik pusat) awal ditentukan secara acak yang juga merupakan titik pusat *cluster* pertama.

3. Menghitung jarak setiap data yang ada terhadap setiap pusat *cluster* menggunakan persamaan *Euclidean Distance Space*.

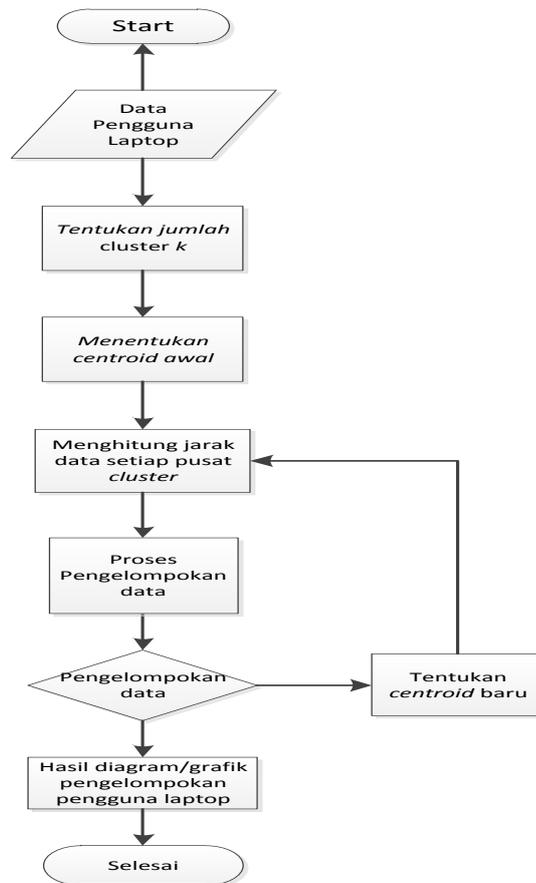
Rumus *Euclidean distance* yaitu:

$$d = |x-y| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

x = Pusat *cluster*

y = data

4. Melakukan pengelompokan data
5. Melakukan penghitungan ulang kembali pusat *cluster* yang baru berdasarkan rata-rata anggota yang ada pada *cluster*.



Gambar II.2 Kerangka algoritma k-means clustering

II.2.4. Weka

Weka merupakan perangkat lunak yang bersifat terbuka (*open source*) yang merupakan sebuah solusi untuk melakukan analisis terhadap data *mining*, *text mining*, dan analisis prediksi. *Weka* menggunakan berbagai teknik deskriptif dan prediksi dalam memberikan wawasan kepada pengguna sehingga dapat membuat keputusan yang paling baik.

II.2.5. Microsoft Excel 2010

Microsoft Excel atau *Microsoft office excel* adalah sebuah program aplikasi lembar kerja yang dibuat dan didistribusikan oleh *Microsoft Corporation* yang dapat dijalankan pada *Microsoft Windows* dan *Mac OS*. Aplikasi ini merupakan bagian dari *Microsoft Office System*. Aplikasi ini memiliki fitur kalkulasi dan pembuatan grafik yang, dengan menggunakan strategi *marketing Microsoft* yang agresif, menjadikan *Microsoft Excel* sebagai salah satu program komputer yang populer digunakan di dalam komputer mikro hingga saat ini. Pada penelitian ini penulis menggunakan *Microsoft Excel 2010* sebagai tempat penyimpanan data-data yang telah diperoleh dari hasil penelitian.

II.2.6. Laptop

Laptop atau komputer jinjing adalah komputer bergerak yang berukuran relatif kecil dan ringan, beratnya berkisar dari 1–6 kg, tergantung pada ukuran, bahan, dan spesifikasi laptop tersebut. Sumber daya laptop berasal dari baterai atau adaptor yang dapat digunakan untuk mengisi ulang baterai dan menyalakan laptop itu sendiri. Baterai laptop pada umumnya dapat bertahan sekitar 2 hingga 6 jam sebelum akhirnya habis, tergantung dari cara pemakaian, spesifikasi, dan ukuran baterai. Laptop terkadang disebut juga dengan komputer *notebook* atau *notebook* saja.

Sebagai komputer pribadi, laptop memiliki fungsi yang sama dengan komputer *desktop (desktop computers)* pada umumnya. Komponen yang terdapat

di dalamnya sama persis dengan komponen pada *desktop*, hanya saja ukurannya diperkecil, dijadikan lebih ringan, lebih tidak panas, dan lebih hemat daya. Berbeda dengan komputer desktop, laptop memiliki komponen pendukung yang didesain secara khusus untuk mengakomodasi sifat komputer jinjing yang portabel. Sifat utama yang dimiliki oleh komponen penyusun laptop adalah ukuran yang kecil, hemat konsumsi energi, dan efisien.

Laptop atau dengan istilah lain disebut dengan komputer jinjing juga memiliki banyak jenis, berikut ini adalah beberapa jenis dari laptop :

1. *Notebook*

Notebook merupakan sebuah perangkat komputer *portable* yang biasanya berukuran layar kurang lebih 13.1 *inchi* hingga 18 *inchi*. Dan *notebook* adalah komputer *portable* mini hasil perkembangan dari laptop. Sedangkan laptop umumnya minimal memiliki ukuran layar 18 *inchi* ke atas. Ukuran ternyata merupakan hal utama yang membedakan antara laptop dan *notebook*.

2. *Netbook*

Netbook termasuk ke dalam kategori perangkat kecil, ringan dan murah. *Netbook* cenderung masih menggunakan sistem operasi *Windows XP* dan secara spesifikasi juga tidak terlalu unggul, karena umumnya hanya dibekali prosesor kelas menengah ke bawah, *RAM* tidak lebih dari 2GB dan tentu saja kualitas grafis ala kadarnya. Maka demikian, *netbook* tidak dapat menjalankan program atau aplikasi yang berat dan mungkin perangkat ini lebih sering digunakan untuk keperluan secukupnya saja, seperti mengolah kata menggunakan *Microsoft Word* dan lain sebagainya yang sejenis.

1. *Tablet PC*

Tablet PC adalah komputer *portable* yang berbentuk buku. Memiliki layar sentuh atau teknologi *tablet* digital yang memungkinkan pengguna komputer mempergunakan *stylus* atau pulpen digital selain *keyboard* ataupun *mouse* komputer, didukung dengan teknologi *Wifi* dan HSDPA untuk koneksi internet. Komputer *tablet* ini merupakan komputer *mobile*, lebih besar daripada ponsel atau *personal digital assistant* (PDA) yang diintegrasikan ke dalam layar sentuh datar dan terutama dioperasikan dengan menyentuh layar daripada menggunakan *keyboard* fisik.