

## BAB II

### LANDASAN TEORI

#### 2.1. Dasar Teori

##### 2.1.1. *Knowledge Discovery in Database (KDD)*

Data mining sering disebut juga *Knowledge Discovery in Database (KDD)* merupakan kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar. Keluaran dari data mining ini bisa dipakai untuk memperbaiki pengambilan keputusan di masa depan. Data mining merupakan metode pengolahan data berskala besar oleh karena itu data mining ini mempunyai peranan penting dalam bidang industri, keuangan, suaca, ilmu, dan teknologi (Santosa, 2007).

Menurut Fayyad istilah data mining dan *Knowledge Discovery in Database (KDD)* sering kali digunakan secara bergantian untuk menjelaskan proses penggalian informasi yang tersembunyi dalam suatu basis data yang besar. Sebenarnya, kedua istilah tersebut memiliki konsep yang berbeda, tetapi berkaitan satu sama lain. Proses KDD secara garis besar dapat dijelaskan sebagai berikut: (dalam Kusrini, 2009)

##### 1. *Data selection*

Pemilihan(seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai. Data hasil seleksi yang akan 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, seperti kesalahan cetak (*tipografi*). Juga dilakukan proses

*enrichment*, yaitu proses “memperkaya” data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk KDD, seperti data atau informasi eksternal.

### 3. *Transformation*

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

### 5. *Interpretation / evaluation*

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.

Menurut (Larose, 2005) Data mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan (dalam Kusriani, 2009) yaitu :

#### 1. Deskripsi

Terkadang peneliti dan analisis secara sederhana ingin mencoba mencari cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data. Sebagai contoh, petugas pengumpulan suara mungkin tidak dapat menemukan keterangan atau fakta bahwa siapa yang tidak cukup profesional akan sedikit disukung dalam

pemilihan presiden. Deskripsi dari pola dan kecenderungan sering memberikan kemungkinan penjelasan untuk suatu pola atau kecenderungan.

## 2. Estimasi

Estimasi hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih ke arah numerik daripada ke arah kategori. Model dibangun menggunakan *record* lengkap yang menyediakan nilai dari variabel target sebagai nilai prediksi. Sebagai contoh, akan dilakukan estimasi tekanan darah sistolik pada pasien rumah sakit berdasarkan umur pasien, jenis kelamin, indeks berat badan, dan *level* sodium darah. Hubungan antara tekanan darah *sistolik* dan nilai variabel prediksi dalam proses pembelajaran menghasilkan model estimasi. Model estimasi yang dihasilkan dapat digunakan untuk kasus baru lainnya.

Contoh lain yaitu nilai estimasi nilai indeks prestasi kumulatif mahasiswa program pascasarjana dengan melihat nilai indeks prestasi mahasiswa tersebut pada saat mengikuti program sarjana.

## 3. Prediksi

Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada di masa mendatang.

Contoh prediksi dalam bisnis dan penelitian adalah :

- a) Prediksi harga beras dalam tiga bulan yang akan datang.
- b) Prediksi presentase kenaikan kecelakaan lalu lintas tahun depan jika batas bawah kecepatan dinaikkan.

Beberapa metode dan teknik yang digunakan dalam klasifikasi dan estimasi dapat pula digunakan (untuk keadaan yang tepat) untuk prediksi.

## 4. Klasifikasi

Dalam klasifikasi, terdapat target variabel kategori. Sebagai contoh, penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu pendapatan tinggi, pendapatan sedang, pendapatan rendah.

Contoh lain klasifikasi dalam bisnis dan penelitian adalah :

- a) Menentukan apakah suatu transaksi kartu kredit merupakan transaksi yang curang atau bukan.
- b) Memperkirakan apakah suatu pengajuan *hipotek* oleh nasabah merupakan suatu kredit yang baik atau buruk.
- c) Mendiagnosis penyakit seorang pasien untuk mendapatkan termasuk kategori apa.

#### 5. Pengklusteran

Pengklusteran merupakan pengelompokan *record*, pengamatan, atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan. Kluster adalah kumpulan *record* yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidakmiripan dengan *record-record* dalam kluster lain.

Pengklusteran beda dengan klasifikasi yaitu tidak adanya variabel target dalam pengklusteran. Pengklusteran tidak mencoba untuk melakukan klasifikasi, mengestimasi, atau memprediksi nilai dari target. Akan tetapi, algoritma pengklusteran mencoba untuk melakukan pembagian terhadap keseluruhan data menjadi kelompok-kelompok yang mempunyai kemiripan (*homogen*), yang mana kemiripan *record* dalam suatu kelompok akan bernilai maksimal, sedang kemiripan dengan *record* dalam kelompok lain akan bernilai minimal.

Contoh pengklusteran dalam bisnis dan penelitian adalah :

- a) Mendapatkan kelompok-kelompok konsumen untuk target pemasaran dari suatu produk bagi perusahaan yang tidak memiliki dana pemasaran yang besar.
- b) Untuk tujuan *audit* akuntansi, yaitu melakukan pemisahan terhadap perilaku finansial dalam baik dan mencurigakan.
- c) Melakukan pengklusteran terhadap ekspresi dari gen, untuk mendapatkan kemiripan perilaku dari gen jumlah besar.

#### 6. Asosiasi

Tugas *asosiasi* dalam data mining adalah menemukan atribut yang muncul dalam satu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja.

Contoh *asosiasi* dalam bisnis dan penelitian adalah :

- a) Meneliti jumlah pelanggan dari perusahaan telekomunikasi seluler yang diharapkan untuk memberikan respons positif terhadap penawaran upgrade layanan yang diberikan.
- b) Menemukan barang dalam supermarket yang dibeli secara bersamaan dan barang yang tidak pernah dibeli secara bersamaan.

### 2.1.2. *Clustering*

*Clustering* adalah mengelompokkan item data kedalam sejumlah kecil grub sedemikian sehingga masing-masing grub mempunyai suatu persamaan yang esensial (Garcia-Molina et al, 2002). Pengelompokan merupakan pemisahan/pemecahan/segmentasi data ke dalam sejumlah kelompok (*cluster*) menurut karakteristik tertentu yang diinginkan, dalam pekerjaan pengelompokan, label dari setiap data belum diketahui, dan dengan pengelompokan diharapkan dapat diketahui kelompok data untuk kemudian diberi label sesuai keinginan (Prasetyo, 2012). (Santosa, 2007) Pada dasarnya *clustering* merupakan suatu metode untuk mencari dan mengelompokkan data yang memiliki kemiripan data (*similarity*) antara satu data dengan data yang lain. *Cluster* adalah salah satu metode data mining yang bersifat tanpa arahan (*unsupervised*), maksudnya metode ini diterapkan tanpa adanya latihan (*training*) dan tanpa ada guru (*teacher*) serta tidak memerlukan target *output*. Guru disini adalah label dari data.

(William) membagi algoritma *clustering* kedalam kelompok besar seperti berikut :

1. *Partitioning algorithms*: algoritma dalam kelompok ini membentuk bermacam partisi dan kemudian mengevaluasinya dengan berdasarkan beberapa kriteria.
2. *Hierarchy algorithms*: pembentukan dikomposisi hirarki dari sekumpulan data menggunakan beberapa kriteria.

3. *Density-based*: pembentukan cluster berdasarkan pada koneksi dan fungsi densitas.
4. *Grid-based*: pembentukan cluster berdasarkan pada struktur *multiple-level granularity*.
5. *Model-based*: sebuah model dianggap sebagai hipotesa untuk masing-masing cluster dan model yang baik dipilih diantara model hipotesa tersebut

### 2.1.3. K-Means Clustering

Metode K-Means merupakan salah satu metode *non-hierarchy clustering* yang umum digunakan. Metode ini mempartisi data ke dalam *cluster*/kelompok sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam *cluster* yang sama dan data yang mempunyai karakteristik berbeda dikelompokkan ke dalam kelompok yang lain. Adapun tujuan dari *clustering* ini adalah untuk meminimalisasikan *objective function* yang diset dalam proses *clustering*, yang pada umumnya berusaha meminimalisasikan variasi di dalam suatu *cluster* dan memaksimalkan variasi antar *cluster* (Agusta, 2007). Karakteristik K-Means dapat diringkas sebagai berikut : (prasetyo, 2012)

1. K-means merupakan metode pengelompokan yang sederhana dan dapat digunakan dengan mudah.
2. Pada jenis set data tertentu, K-Means tidak dapat melakukan segmentasi data dengan baik dimana hasil segmentasinya tidak memberikan pola kelompok yang mewakili karakteristik bentuk alami data.
3. K-Means bisa mengalami masalah ketika pengelompokan data yang mengandung *outlier*.

Algoritma K-means *clustering* menurut (Prasetyo, 2012,2014) adalah sebagai berikut :

1. Tentukan jumlah kelompok.
2. Alokasikan data kedalam kelompok secara acak.
3. Hitung pusat kelompok (*centorid*/rata-rata) dari data yang ada di

masing-masing kelompok.

Rumus untuk menghitung *centroid* adalah sebagai berikut :

$$C = \frac{1}{M} \sum_{j=1}^M x_j \dots \dots \dots (2.1)$$

M= Menyatakan jumlah data dalam sebuah kelompok

i = Menyatakan fitur ke-i dalam sebuah kelompok

4. Alokasikan masing-masing data ke *centroid*/rata-rata terdekat dengan matrik jarak yang sudah ditentukan

Rumus menghitung jarak euclidean :

$$D(x_2, x_1) = \|x_2 - x_1\| = \sqrt{\sum_{j=1}^P |x_{2j} - x_{1j}|^2} \dots \dots \dots (2.2)$$

D = Adalah jarak antara data  $x_2$  dan  $x_1$ , dan  $|\cdot|$  adalah nilai mutlak

P = menyatakan dimensi data

5. Hitung kembali *centroid* C berdasarkan data yang mengikuti *cluster* masing-masing.
6. Kembali ke langkah 3 dan 4, apabila : (1) masih ada data yang berpindah kelompok, (2) apabila ada perubahan nilai centroid di atas nilai ambang yang ditentukan, (3) apabila perubahan nilai pada fungsi objektif yang digunakan masih diatas nilai ambang yang telah ditentukan.

#### 2.1.4. Penyedia / Vendor

Penyedia barang/jasa adalah badan usaha, badan hukum, perorangan yang kegiatan usahanya menyediakan barang/jasa yang sah secara hukum dan mampu menyediakan barang/jasa dengan baik (Peraturan Perusahaan No. 02 th 2007). Menurut Prabowo, kinerja vendor menjadi salah satu faktor penting dalam *supply chain* karena merupakan salah satu strategi perusahaan untuk dapat bersaing dengan perusahaan lain dalam hal kepuasan konsumen dan juga untuk meningkatkan dan mempertahankan *service level* perusahaan tersebut dalam memenuhi permintaan konsumen (Prabowo, 2012). *Supplychain* disini

adalah sebagai manajemen rantai suplai yang dilakukan oleh beberapa perusahaan produksi Indonesia.

### **2.1.5. Pengadaan**

Pengadaan (*procurement*) adalah proses antar organisasi yang dilakukan dua peran : pembeli dan penjual dengan tujuan si pembeli membeli barang / jasa dari penjual. Hal ini dikerjakan dengan adanya pertukaran berbagai dokumen, diikuti dengan pertukaran barang dan uang. Terdapat perbedaan yang cukup jelas antara pembelian dan pengadaan, dimana pembelian hanyalah semata-mata aktifitas pembayaran barang dan jasa. Sementara pengadaan, selain proses pembelian juga merupakan sebuah siklus. Siklus tersebut merupakan serangkaian peristiwa permintaan barang, persetujuan, hingga proses pembayarannya. (Ritchi )

Dalam Peraturan Perusahaan PDAM di jelaskan proses pengadaan terdapat beberapa metode yaitu :

#### **1. Pelelangan umum**

Yaitu pengadaan barang/jasa dengan menciptakan persaingan yang sehat di antara penyedia barang/jasa yang setara dan memenuhi syarat, berdasarkan metode dan tata cara tertentu, yang telah ditetapkan dan di ikuti oleh pihak-pihak terkait secara taat sehingga terpilih penyedia barang/jasa terbaik.

#### **2. Pelelangan terbatas**

Adalah metode pemilihan barang/jasa tertentu dengan jumlah penyedia yang mampu melaksanakan diyakini terbatas dilakukan dan di umumkan secara terbuka.

#### **3. Pemilihan langsung**

Metode pemilihan penyedia barang/jasa dalam hal pelelangan umum atau pelelangan terbatas dinilai tidak efisien dalam segi biaya pelelangan, yaitu pemilihan penyedia barang/jasa dilakukan dengan mengundang sekurang-kurangnya tiga penyedia barang/jasa yang memilik SKT, serta dilakukan negosiasi baik secara teknis maupun biaya dan di umumkan secara resmi di papan pengumuman



perusahaan. Proses pelelangan pemilihan langsung yang apabila gagal, akan dilanjutkan dengan penunjukan langsung. Prosedur / tata cara pemilihan langsung adalah sebagai berikut :

- a) Dilakukan oleh urusan Pengadaan
- b) Di umumkan melalui papan pengumuman perusahaan dan/atau melalui *website* perusahaan.
- c) Minimal yang mendaftar adalah tiga (3) penyedia barang/jasa.
- d) Pada prinsipnya pemilihan calon penyedia barang/jasa didasarkan kepada kinerja sebelumnya.
- e) Calon penyedia barang/jasa yang dipilih, diprioritaskan yang sudah terdaftar di perusahaan.
- f) Apabila yang memasukkan penawaran kurang dari dua (2), pelelangan dinyatakan gagal.

#### 4. Penunjukan langsung

Metode pemilihan penyedia barang/jasa dalam hal pelelangan umum atau pelelangan terbatas dinilai tidak efisien dalam segi biaya pelelangan. Pelaksanaan pengadaan barang/jasa dilakukan satu penyedia barang/jasa tanpa melalui pelelangan atau pemilihan langsung dan perlu melakukan klarifikasi / negosiasi baik teknis maupun harga.

#### 5. Pengadaan langsung

Yaitu pengadaan barang / jasa langsung kepada penyedia barang/jasa, tanpa melalui pelelangan/seleksi/penunjukan langsung dan dapat dilakukan terhadap pengadaan barang/pekerjaan konstruksi/jasa lainnya yang bernilai paling tinggi Rp. 100.000.000 (dalam draft perpres no 54 tahun 2010)

## **2.2. Profil Perusahaan**

PDAM atau Perusahaan Daerah Air Minum Surya Sembada Kota Surabaya merupakan salah Satu Badan Usaha Milik Daerah Surabaya yang bergerak dalam bidang pendistribusian air bersih untuk masyarakat luas, khususnya untuk

masyarakat Surabaya. Karena PDAM merupakan Badan usaha Milik Daerah, maka kinerja perusahaan dimonitor dan diawasi langsung oleh Pemerintah Kota Surabaya.

Berdirinya PDAM Surya Sembada Kota Surabaya merupakan peninggalan jaman belanda, dibentuk berdasarkan (website resmi PDAM Surabaya) :

- 1) Peraturan Daerah No. 7 tahun 1976, tanggal 30 maret 1976.
- 2) Dikukuhkan dan disahkan dengan Surat Keputusan Gubernur Kepala Daerah Tingkat I Jawa Timur No. 11/155/76, tanggal 6 November 1976.
- 3) Diundangkan dalam Lembaran Daerah Kota Madya Daerah Tingkat II Surabaya Seri C No. 4C, tanggal 23 November 1976. (Company Profile)

### **2.2.1. Sejarah perkembangan PDAM Surya Sembada kota Surabaya**

Karena PDAM Surabaya merupakan perusahaan peninggalan Belanda Berikut penjelasan mengenai sejarah PDAM Surabaya Sembada Kota Surabaya (*Website* resmi PDAM Surabaya):

Tahun 1890 : Penyediaan pertama air minum untuk Kota Surabaya pertama yang diperoleh dari sumber mata air desa Purut di Kabupaten Pasuruan dan diangkut menggunakan kereta api.

Tahun 1901 : Pembangunan sistem penyediaan air minum mata air Pandaan oleh Carel Willem Weijjs.

Penyelesaian pekerjaan membutuhkan waktu 2.5 tahun Pekerjaan terdiri dari :

- a) Pembangunan sumber mata air Toyo Arang (107 SHVP) dengan kapasitas 62-73 liter per detik.
- b) Pembangunan sumber mata air pintahan (264 SHVP) dengan kapasitas 102-125 liter/detik.
- c) Pembangunan reservoir/tandon tamanan (103 SHVP).
- d) Pemasangan pipa transmisi diameter 450 mm dengan panjang pipa 38.318km.

- e) Pemasangan 133km jaringan pipa distribusi (22 km pipa dengan diameter 20-50 mm, 111 km pipa sirkulasi dengan diameter 60-150 mm).
  - f) Selain itu juga pemasangan 16 km pipa ke daerah militer / laut.
  - g) 1.000 hidran dan 150 air mancur jalan
- Tahun 1903 : 8 Oktober 1903 peresmian pekerjaan sistem penyediaan air minum sumber mata air Pandaan.  
Dan Perusahaan Air Minum didirikan dibawah pemerintahan kolonial Belanda.
- Tahun 1906 : Jumlah pelanggan kurang lebih 1.500 sambungan.
- Tahun 1922 : IPAM Ngagel dibangun dengan kapasitas 60 liter / detik.
- Tahun 1932 : Pembangunan sistem penyediaan air Umbulan untuk memenuhi kebutuhan air minum Kota Surabaya, pekerjaan meliputi pembangunan rumah pompa baru beserta aksesorisnya.
- Tahun 1942 : Peningkatan kapasitas IPAM Ngagel I menjadi 180 liter/detik.
- Tahun 1950 : Perusahaan Air minum diserahkan pada Pemerintah Republik Indonesia (Kota Praja Surabaya).
- Tahun 1954 : Peningkatan kapasitas IPAM Ngagel I menjadi 350 Liter/detik.
- Tahun 1959 : Pembangunan IPAM Ngagel II kapasitas 1.000 liter / detik, didesain & dilaksanakn oleh F.A Degremont (Prancis).
- Tahun 1976 : Perusahaan Air Minum Ddisahkan sebagai Perusahaan Daerah dengan Perda No.7 tanggal 30 Maret 1976.
- Tahun 1977 : Peningkatan Kapasitas IPAM Ngagel I menjadi 500 liter/detik.
- Tahun 1978 : Pengalihan sattus menjadi Perusahaan Daerah Air

Minum dari Dinas Air Minum berdasarkan SK Walikotamadya Dati II Surabaya No. 657/WK/77 tanggal 30 Desember 1977.

- Tahun 1980 : Peningkatan kapasitas IPAM Ngagel I menjadi 1.000 liter/detik.
- Tahun 1982 : Pembangunan IPAM Ngagel III kapasitas 1.000 liter/detik dengan lisensi dari Neptune Microfloc (Amerika Serikat).
- Tahun 1990 : Pembangunan IPAM Karangpilang I dengan kapasitas 1.000 liter/detik dengan dana Loan IBRD No.2632 IND.
- Tahun 1991 : Pembangunan gedung kantor PDAM yang terletak di Mayjend Prof.DR. Moestopo NO.02 SURabaya yang dibiayai dana MDAM murni.
- Tahun 1994 : Peningkatan kapasitas IPAM Ngagel I menjadi 1.500 liter/detik.
- Tahun 1996 : Peningkatan kapasitas IPAM Ngagel I menjadi 1.800 liter/detik, peningkatan kapasitas IPAM Karangpilang I menjadi 1.200 liter/detik dan dimulainya pembangunan IPAM KARangpilang II dengan kapasitas 2.000 liter/detik yang didanai Loan IBRD No. 3726 IND.
- Tahun 1997 : Peningkatan kapasitas IPAM Ngagel III menjadi 1.500 liter/detik. Dan produksi awal 500 liter/detik IPAM Karangpilang II didistribusikan ke pelanggan.
- Tahun 1999 : Pembangunan IPAM Karangpilang II dengan kapasitas 2.000 liter/detik telah selesai.
- Tahun 2001 : Pekerjaan peningkatan kapasitas IPAM Karangpilang II menjadi 2.500 liter/detik.
- Tahun 2005 : Peningkatan kapasitas IPAM Ngagel III menjadi 1.750 liter/detik.
- Tahun 2006 : Peningkatan kapasitas IPAM Karangpilang I menjadi

1.450 liter/detik dan Peningkatan kapasitas IPAM Karangpilang II menjadi 2.750 liter /detik.

Tahun 2009 : Pembangunan IPAM Karangpilang III dengan kapasitas 2.000 liter/detik.

### **2.2.2. Visi PDAM Surya Sembada Kota Surabaya**

Manjadi Perusahaan Air Minum yang mandiri, berwawasan global dan terbaik dikelasnya.

### **2.2.3. Misi Pdam Surya Sembada Kota Surabaya**

Berikut ini misi dari PDAM Surya Sembada Kota Surabaya :

1. Menyediakan air minum yang memenuhi kualitas, kuantitas, dan kontinuitas yang dapat dipertanggungjawabkan melalui pelayanan prima dalam rangka menciptakan nilai tambah bagi masyarakat Kota Surabaya.
2. Melakukan pengolahan usaha secara profesional dengan teknologitepat guna dan prinsip-prinsip manajemen yang berwawasan global sehingga mampu memberikan kontribusi pada peningkatan Pendapatan Asli Daerah.
3. Mengembangkan lingkungan kerja yang kondusif sehingga menjadi pilihan utama karyawan bekerja dan berkarir secara profesional.
4. Turut yang berpartisipasi dalam mengemban tanggung jawab sosial secara proposional melalui aktifitas *Corporate Social Responsibility (CSR)*.

### **2.2.4. Budaya Kerja PDAM Surya Sembada Kota Surabaya**

Budaya kerja yang diterapkan di PDAM Surya Sembada Kota Surabaya yaitu S.M.I.L.E :

*Saticfaction* : Mengutamakan kepuasan dalam melayani

- pelanggan.
- Morale* : Memiliki semangat juang yang gigih dan pantang menyerah dalam berupaya mencapai kesuksesan.
- Integrity* : Memahami komitmen untuk mewujudkan loyalitas.
- Leadership* : Berjiwa kepemimpinan sebagai teladan dalam sikap, kompetensi, dan jati diri.
- Enterpreneurship* : Memiliki keberanian dalam mengambil resiko dengan perhitungan yang masuk akal serta terkendali.

### **2.3. Penelitian terdahulu**

Pada kepustakaan penelitian ini akan dibahas mengenai penelitian-penelitian yang pernah dilakukan sebelumnya. Dari beberapa penelitian terdahulu, untuk lebih jelasnya dapat dilihat pada tabel *road map* berikut :

**Tabel 2.1 Tabel *Road Map* penelitian terdahulu**

No	Peneliti	Fokus Penelitian	Metode Penelitian	Hasil Penelitian
1	Ivan Rosyadi	Pengelompokkan Prestasi Akademik Siswa MTsN Gresik	K-Means	Hasil uji system dengan evaluasi klaster internal Davides-Bouldin Index yang paling bagus terdapat pada percobaan ke-9 pada cluster 1 dengan nilai DBI 7,06. Dengan jumlah prosentase 17,46% untuk jumlah 62 siswa dengan prestasi akademik baik, 40,56% untuk jumlah 144 siswa dengan prestasi akademik sedang, dan 41,97% dengan jumlah 149 siswa dengan prestasi akademik buruk.
2	Banatus Sa'adah	Pengelompokkan Potensi Akademik Siswa RA Tarbiyatul Aulad	K-Means	Pada penelitian ini pengelompokan menggunakan parameter mata pelajaran sekolah yaitu : akhlak, hafalan, agama, bahasa, dan motorik. Hasil akhir dari perhitungan menunjukkan bahwa hasil nilai siswa yang berubah maka di akhir perhitungan k-mean juga ikut berubah dimana nilai ujian siswa akan berpengaruh.

3	Nizam Al-Anshori	Pengelompokan kompetensi database mahasiswa UMG	K-Harmonic Means	<p>Pada penelitian ini dilakukan percobaan menggunakan 2, 3, 5, 7, 9 cluster. Dan pelabelan cluster yang digunakan yaitu K=3 yaitu mahasiswa dengan kompetensi tinggi, sedang, dan rendah.</p> <p>Penentuan cluster menggunakan voting dan prosentase kesesuaian hasil pengelompokan kompetensi mahasiswa menggunakan metode k-harmonic means dengan perhitungan penilaian dosen sebesar 83,53%.</p>
---	---------------------	---	---------------------	--