

BAB II

LANDASAN TEORI

2.1 Pembangunan Daerah

Adapun pengertian dari pembangunan daerah ialah sebagai berikut.

2.1.1 Pengertian Pembangunan Daerah

Pembangunan daerah merupakan suatu proses perubahan terencana yang dilakukan oleh pemerintah daerah bersama masyarakat untuk meningkatkan kesejahteraan, kualitas hidup, serta kemandirian wilayah melalui pemanfaatan potensi lokal yang dimiliki daerah tersebut. Pembangunan tidak hanya berpotensi pada pertumbuhan ekonomi, tetapi juga pada pemerataan antarwilayah, peningkatan sumber daya manusia, serta penguatan kelembagaan lokal.

Implementasi pembangunan daerah menunjukkan bahwa pembangunan daerah adalah proses yang melibatkan masyarakat lokal dan pemerintah daerah dalam merancang, melaksanakan, dan mengevaluasi program pembangunan (Zaili Rusli, R. P. S 2020).

Di Labuhanbatu Utara, pembangunan daerah melibatkan penguatan sektor-sektor strategis seperti perkebunan rakyat, plasma, pertanian, dan peningkatan kualitas sumber daya manusia. Pemerintah daerah memiliki peran penting dalam mengatur dan mendorong pembangunan ekonomi masyarakat melalui program-program pemberdayaan dan kebijakan pembangunan perkebunan (Oktavinai et al., 2024).

2.1.2 Tujuan dan Indikator Pembangunan Daerah

Untuk memahami tujuan dan indikator pembangunan daerah, berikut akan dijelaskan apa saja tujuan dan indikator pembangunan daerah.

2.1.2.1 Tujuan Pembangunan Daerah

Pembangunan daerah memiliki tujuan strategis yang tidak hanya berfokus pada peningkatan ekonomi, tetapi juga pemerataan pelayanan publik, peningkatan kualitas hidup, dan pengurangan ketimpangan antarwilayah. Berikut ini adalah tujuan-tujuan pembangunan daerah :

1. Meningkatkan Kesejahteraan Masyarakat

Tujuan utama pembangunan daerah adalah meningkatkan kesejahteraan penduduk, hal ini mencakup tingkat pendapatan yang lebih baik, kualitas pendidikan dan kesehatan yang meningkat, kesempatan kerja yang lebih luas, dan meningkatnya akses terhadap layanan publik. Dengan meningkatkan kesejahteraan, masyarakat dapat hidup lebih layak, produktif, dan mandiri.

2. Mengurangi Ketimpangan Antarwilayah

Pembangunan daerah bertujuan untuk menurunkan kesenjangan pembangunan antar-kecamatan atau antarwilayah, kesenjangan bisa berupa perbedaan akses pendidikan, kualitas infrastruktur, peluang pekerjaan, dan tingkat ekonomi. Pengurangan ketimpangan penting agar setiap wilayah mampu berkembang secara optimal.

3. Mengoptimalkan Pemanfaatan Sumber Daya Daerah

Setiap wilayah memiliki potensi yang berbeda seperti sumber daya alam, sumber daya manusia, potensi wisata, potensi industri lokal. Pembangunan daerah bertujuan memaksimalkan potensi tersebut sehingga memberikan kontribusi terhadap pendapatan daerah dan kesejahteraan masyarakat.

4. Mendorong Kemandirian Ekonomi Daerah

Pembangunan diarahkan agar daerah tidak bergantung penuh pada pemerintah pusat. Hal ini dicapai melalui peningkatan PAD (Pendapatan Asli Daerah), pengembangan sektor unggulan daerah, investasi daerah, pengembangan UMKM. Dengan kemandirian ekonomi, daerah dapat menyusun kebijakan pembangunan yang lebih tepat dan berkelanjutan.

5. Meningkatkan Kualitas Layanan Publik

Pembangunan daerah berorientasi pada peningkatan kualitas publik seperti pendidikan, kesehatan, transportasi, air bersih dan sanitasi, administrasi pemerintahan. Layanan publik yang baik membuat masyarakat lebih produktif dan nyaman.

6. Memperkuat Daya Saing Daerah

Daerah yang memiliki keunggulan kompetitif mampu menarik investasi, meningkatkan ekspor, dan mendorong pertumbuhan ekonomi. Pembangunan diarahkan pada peningkatan infrastruktur, pengembangan sumber daya manusia, teknologi informasi, sistem inovasi daerah.

7. Mendukung Tata Kelola Pemerintahan yang Baik

Pembangunan daerah juga bertujuan memperbaiki tata kelola publik, diantaranya yaitu transparansi, akuntabilitas, efisiensi birokrasi, pelayanan berbasis digital (*e-government*).

2.1.2.2 Indikator Pembangunan Daerah

Agar pembangunan dapat diukur dan dievaluasi secara objektif, diperlukan Indikator. Berikut indikator pembangunan daerah yang umum dipakai dalam penelitian dan perencanaan pembangunan :

1. Indikator Ekonomi

Indikator ekonomi digunakan untuk mengukur sejauh mana suatu daerah mampu menghasilkan kegiatan ekonomi dan meningkatkan kesejahteraan masyarakat.

2. Indikator Sosial

Indikator sosial menunjukkan tingkat kualitas hidup masyarakat yang mencakup pendidikan, kesehatan, dan kesejahteraan manusia.

3. Indikator Infrastruktur

Infrastruktur berfungsi sebagai penopang utama kegiatan ekonomi, sosial, dan mobilitas masyarakat.

4. Indikator Lingkungan

Pembangunan berkelanjutan menuntut pengelolaan sumber daya alam dan lingkungan yang baik.

5. Indikator Pemerintahan

Indikator pemerintahan mengukur kualitas tata kelola dan pelayanan publik.

2.1.3 Pembangunan Daerah di Labuhanbatu Utara

Kabupaten Labuhanbatu Utara merupakan hasil pemekaran dari Kabupaten Labuhanbatu. Pemekaran ini dilakukan agar pemerintahan lokal lebih efektif dalam mengelola sumber daya, meningkatkan pelayanan publik, dan mempercepat pembangunan di wilayah yang sebelumnya terpinggirkan. Sejak dimekarkan, Labuhanbatu Utara berfokus pada pembangunan infrastruktur, pemberdayaan masyarakat, dan peningkatan kesejahteraan sosial dan ekonomi masyarakat (Pohan, Z. S 2015).

Pembangunan di Kabupaten Labuhanbatu Utara dilakukan melalui perencanaan yang bersifat jangka panjang, jangka menengah, dan tahunan. Tujuan dari perencanaan ini adalah untuk memastikan pembangunan daerah berjalan terarah, berkelanjutan, dan sesuai dengan kebutuhan masyarakat.

1. Rencana Pembangunan Jangka Menengah Daerah (RPJMD) 2025-2029

RPJMD merupakan dokumen perencanaan pembangunan daerah untuk jangka waktu lima tahun. Dokumen ini menjadi pedoman bagi pemerintah daerah dalam menetapkan program dan kegiatan prioritas. RPJMD Kabupaten Labuhanbatu Utara 2025-2029 menitikberatkan pembangunan pada empat aspek utama yaitu, Infrastruktur, pendidikan, kesehatan, dan pemberdayaan masyarakat (Kanwil Kemenhumham Sumut 2025).

2. Rencana Pembangunan Jangka Panjang Daerah (RPJPD) 2025-2045

RPJPD merupakan dokumen perencanaan pembangunan daerah untuk jangka waktu dua puluh tahun. Dokumen ini berfungsi sebagai pedoman bagi seluruh perencanaan pembangunan jangka menengah dan tahunan di

daerah. RPJPD menekankan pembangunan berkelanjutan, yaitu pembangunan yang memperhatikan keseimbangan ekonomi, kesejahteraan sosial, dan kelestarian lingkungan.

3. Rencana Kera Pemerintah Daerah (RKPD) 2025

RKPD adalah dokumen perencanaan pembangunan tahunan yang disusun berdasarkan RPJMD dan RPJPD. Dokumen ini menekankan program dan kegiatan prioritas yang harus dilaksanakan pada tahun bersangkutan. Prioritas pembangunan Kabupaten Labuhanbatu Utara dalam RKPD 2025 meliputi, pembangunan dan pemberdayaan desa, peningkatan infrastruktur transportasi, dan pengembangan ekonomi lokal melalui pemberdayaan sektor pertanian, perkebunan, UMKM, dan perdagangan lokal (Perbup Labura Nomor 6 Tahun 2024).

2.1.4 Peran Dinas PUTR Kab. Labuhanbatu Utara

Dinas Pekerjaan Umum dan Tata Ruang (PUTR) Kabupaten Labuhanbatu Utara memegang peran strategis dalam mendukung pembangunan daerah, terutama pada sektor infrastruktur dan pengelolaan tata ruang. Peran ini mencakup perencanaan, pelaksanaan, dan pemeliharaan infrastruktur, serta penyusunan tata ruang wilayah agar pembangunan berjalan terarah dan berkelanjutan (Pemerintah Kabupaten Labuhanbatu Utara, 2024).

Dalam hal perencanaan, Dinas PUTR bertanggung jawab menyusun rencana pembangunan infrastruktur, termasuk jalan, jembatan, irigasi, dan fasilitas publik lainnya, serta merancang tata ruang wilayah yang mendukung pemanfaatan lahan sesuai peruntukannya, baik untuk pemukiman, pertanian, industri, maupun

kawasan konversi. Tata ruang yang baik menjadi dasar pembangunan berkelanjutan dan pengendalian penggunaan lahan agar tidak merusak lingkungan (Badan Perencanaan Pembangunan Daerah Kabupaten Labuhanbatu Utara, 2023).

Selain perencanaan, Dinas PUTR melaksanakan pembangunan fisik sesuai program pemerintah daerah. Pekerjaan ini meliputi konstruksi dan pemeliharaan jalan serta jembatan untuk meningkatkan konektivitas antarmasyarakat, pembangunan sistem drainase dan irigasi guna mendukung sektor pertanian, serta pembangunan fasilitas publik seperti pasar dan terminal yang mendukung kesejahteraan masyarakat. Pelaksanaan pembangunan ini juga disertai pengawasan mutu dan anggaran proyek untuk memastikan infrastruktur dibangun sesuai standar dan tepat waktu (Kemen PUPR RI, 2022).

Setelah pembangunan selesai, Dinas PUPR bertanggung jawab dalam pemeliharaan infrastruktur agar tetap layak digunakan. Pemeliharaan mencakup perbaikan jalan, jembatan, irigasi, dan fasilitas publik lainnya. Infrastruktur yang terawat tidak hanya menunjang mobilitas masyarakat, tetapi juga mendukung kegiatan ekonomi lokal, distribusi hasil pertanian, dan mendorong investasi di wilayah tersebut (Pemkab Labuhanbatu Utara, 2024).

Selain itu Dinas PUTR berperan dalam mendukung pembangunan berkelanjutan dengan memastikan pembangunan infrastruktur dan tata ruang ramah lingkungan. Hal ini dilakukan melalui penerapan standar konstruksi yang efisien energi, mitigasi risiko bencana, dan pengelolaan penggunaan lahan yang adil dan berkelanjutan. Dengan demikian, Dinas PUTR tidak hanya berfokus pada pengembangan fisik, tetapi juga pada peningkatan kualitas hidup masyarakat dan

keseimbangan lingkungan hidup di Kabupaten Labuhanbatu Utara (Kementrian PUPR RI, 2022).

2.2 Data Mining

Untuk memahami apa itu *Data Mining*, berikut akan di jelaskan apa pengertian dari *Data Mining*.

2.2.1 Pengertian Data Mining

Data mining adalah proses menemukan pola, hubungan, atau informasi berguna dari sekumpulan data yang besar menggunakan teknik statistik, algoritma, dan *machine learning*. Tujuan utama *data mining* adalah mengubah data menjadi informasi yang dapat digunakan untuk pengambilan keputusan.

Data mining sering disebut sebagai bagian dari *Knowledge Discovery in Databases* (KDD), yaitu proses penemuan pengetahuan dari data yang mencakup pemilihan data, pra-pemrosesan, transformasi, penerapan algoritma, evaluasi, dan interpretasi hasil. Menurut Edhy Tri Cahyono (2020), *data mining* bisa diartikan sebagai proses menggunakan teknik statistik, matematika, AI, dan *machine learning* untuk mengidentifikasi informasi penting dan pengetahuan dari bisnis data yang besar.

Dalam konteks pemerintahan dan pembangunan daerah, *data mining* menjadi alat yang efektif dalam mengolah data besar untuk kebutuhan perencanaan kebijakan (Siahaan, 2022).

Melalui *data mining*, pola dan kondisi pembangunan antarwilayah dapat dianalisis secara objektif sehingga menghasilkan rekomendasi yang tepat sasaran.

2.2.2 Proses *Data Mining* (KDD)

Proses data mining tidak dilakukan secara langsung, tetapi mengikuti tahapan sistematis yang dikenal sebagai *Knowledge Discovery in Database* (KDD). *Knowledge Discovery in Database* merupakan serangkaian langkah yang bertujuan mengubah data mentah menjadi informasi atau pengetahuan yang dapat digunakan. Secara umum, tahapan dalam proses *Knowledge Discovery in Database* terdiri dari lima langkah utama sebagai berikut :

1. *Data Selection* (Pemilihan Data)

Pada tahap ini, peneliti memilih data yang relevan dengan tujuan analisis dari berbagai sumber yang tersedia. Sumber data dapat berupa basis data instansi pemerintah, data survei, data statistik, maupun data hasil observasi. *Data Preprocessing atau Cleaning* (Pembersih Data)

Pembersihan data merupakan langkah untuk memastikan bahwa data yang digunakan memiliki kualitas baik. Pada tahap ini, data diperiksa dan diperbaiki dari berbagai masalah seperti, nilai kosong, data ganda, kesalahan input, dan ketidakkonsistenan antarvariabel. Tahap pembersihan data sangat menentukan kualitas hasil analisis.

2. *Data Transformation* (Transformasi Data)

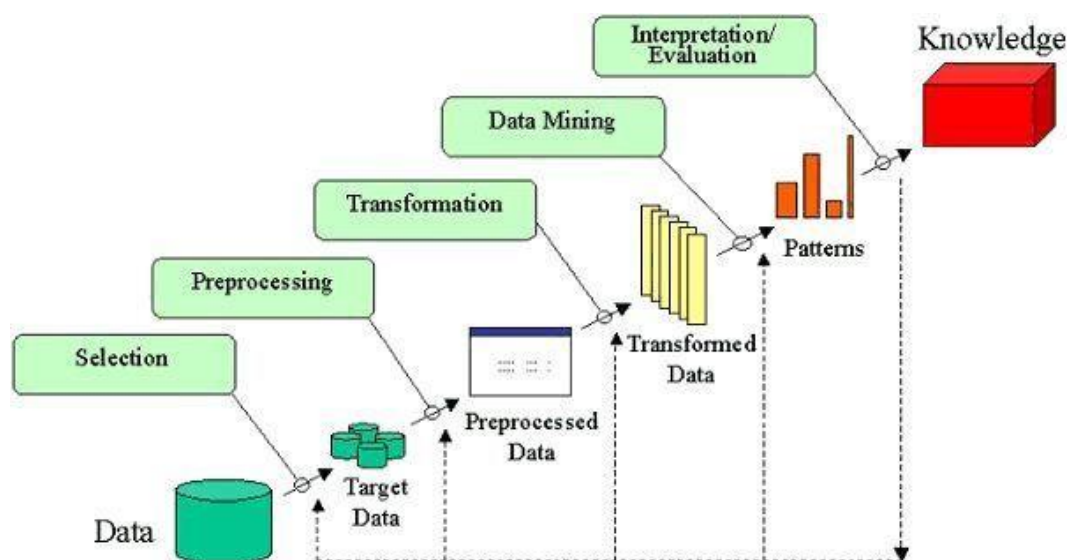
Setelah data dibersihkan, langkah berikutnya adalah menyesuaikan bentuk data agar cocok digunakan oleh algoritma *data mining*. Transformasi data dapat dilakukan dengan berbagai cara seperti, normalisasi, standardisasi, pengelompokan ulang variabel, dan pengkodean data. proses transformasi

bertujuan untuk mempermudah algoritma membaca pola dalam data dan memastikan bahwa semua variabel memiliki bobot yang seimbang.

3. *Data Mining* (Penerapan Algoritma)

Tahap ini merupakan inti dari seluruh proses *Knowledge Discovery in Database* (KDD). Pada tahap *data mining*, peneliti menerapkan metode atau algoritma tertentu untuk menemukan pola dari data yang telah dipersiapkan. Metode yang digunakan dapat berupa klasifikasi, regresi, asosiasi, atau *clustering*. *Evaluation and Interpretation* (Evaluasi dan Interpretasi)

Setelah proses *data mining* selesai, hasil yang diperoleh perlu dievaluasi dan ditafsirkan. Evaluasi dilakukan untuk memastikan bahwa pola atau *cluster* yang terbentuk benar-benar menggambarkan kondisi sebenarnya dan sesuai dengan tujuan analisis. Berikut ini adalah ilustrasi KDD :



Gambar 2.1 *Knowledge Discovery in Database*

2.3 Metode *K-Means*

Berikut ini akan di paparkan mengenai pengertian *K-Means*, konsep dasar *K-Means*, proses *K-Means*, dan penerapan *K-Means*

2.3.1 Pengertian *K-Means*

K-Means adalah salah satu algoritma *clustering* dalam *data mining* yang digunakan untuk mengelompokkan sekumpulan data mejadi beberapa *cluster* berdasarkan kemiripan data. *K-Means* termasuk metode *unsupervised learning*, yaitu metode pembelajaran tanpa label kelas. Dengan kata lain, algoritma ini tidak memerlukan informasi sebelumnya mengenai kategori data tetapi mampu menemukan pola atau struktur tersembunyi dalam dataset.

Tujuan utama *K-Means* adalah meminimalkan variasi atau jarak antar anggota dalam satu *cluster* dan memaksimalkan jarak antar *cluster* berbeda.

Dalam konteks penelitian ini, *K-Means* digunakan untuk menganalisis data pembangunan Kabupaten Labuhanbatu Utara, dengan sumber data dari Dinas Pekerjaan Umum dan Tata Ruang (PUTR). Data yang dianalisis meliputi data pembangunan proyek pada tahun 2024, sehingga dapat dikelompokkan menjadi *cluster* berdasarkan data pembangunan proyek terbanyak, sedang, dan rendah.

2.3.2 Konsep Dasar *K-Means*

K-Means bekerja dengan menentukan *centroid* atau titik pusat untuk setiap *cluster*, kemudian mengelompokkan data berdasarkan jarak ke *centroid*. Proses ini berulang hingga posisi *centroid* stabil dan tidak berubah secara signifikan. Konsep dasarnya meliputi :

- *Cluster*, yaitu sekumpulan data yang memiliki karakteristik serupa.

- *Centroid*, yaitu titik pusat *cluster* yang dihitung berdasarkan rata-rata data dalam *cluster*.
- Jarak yaitu ukuran kedekatan data dengan *centroid*, biasanya menggunakan *Euclidean Distance*.

2.3.3 Proses *K-Means*

Proses *K-Means* terdiri dari beberapa langkah yang dilakukan secara berulang untuk menghasilkan pengelompokan data yang optimal. Tahap pertama yang dilakukan dalam metode ini adalah menentukan jumlah *cluster* atau kelompok yang akan dibentuk, yang dilambangkan dengan k . Penentuan jumlah *cluster* ini sangat penting karena berpengaruh langsung terhadap hasil pengelompokan, dan biasanya ditetapkan berdasarkan kebutuhan penelitian, karakteristik objek yang dianalisis, atau menggunakan pendekatan tertentu seperti *Elbow Method* untuk memperoleh jumlah *cluster* yang paling representatif. Setelah nilai k ditentukan, algoritma melanjutkan proses dengan memilih *centroid* awal secara acak dari sekumpulan data, sehingga pemilihan *centroid* awal ini sangat memengaruhi kualitas hasil akhir karena *K-Means* sensitif terhadap inisialisasi awal.

Setelah *centroid* awal ditetapkan, langkah berikutnya adalah menghitung jarak masing-masing data terhadap setiap *centroid*. Pada tahap ini, algoritma umumnya menggunakan ukuran jarak *Euclidean* karena kesederhanaan dan kemampuan metrik ini menggambarkan kedekatan geometris antar data. perhitungan jarak ini dilakukan agar setiap objek dapat ditentukan kedekatannya terhadap pusat *cluster*. Berdasarkan jarak yang dihasilkan, algoritma kemudian memasukkan setiap data

ke dalam *cluster* yang memiliki *centroid* paling dekat dengannya. Proses ini disebut sebagai tahap *assignment*, yakni penempatan data ke dalam kelompok yang paling sesuai berdasarkan jarak minimum. Dengan demikian, data-data yang memiliki kemiripan karakteristik akan kumpul dalam *cluster* yang sama.

Setelah semua data dikelompokkan, algoritma melanjutkan ke tahap pembaruan *centroid*. Pada tahap ini, *centroid* baru dihitung ulang dengan mengambil nilai rata-rata dari semua data yang berada dalam setiap *cluster*. Proses ini berlangsung untuk memastikan bahwa posisi *centroid* merepresentasikan pusat sesungguhnya dari kelompok data setelah pengelompokan awal dilakukan. Selanjutnya, algoritma memeriksa apakah terdapat perubahan posisi *centroid* dibandingkan dengan *centroid* pada iterasi sebelumnya. Jika terjadi perubahan yang cukup signifikan, maka algoritma akan kembali melakukan perhitungan jarak dan pengelompokan ulang hingga *centroid* stabil. Namun, apabila tidak ada perubahan atau perubahan sangat kecil, maka kondisi ini disebut sebagai konvergensi, yang menandakan bahwa proses iteratif *K-Means* sudah mencapai titik akhir.

Secara keseluruhan, proses *K-Means* berjalan secara iteratif hingga tercapai keadaan stabil, dan hasil akhirnya berupa pembagian data ke dalam beberapa *cluster* yang memiliki karakteristik serupa berdasarkan kedekatannya dengan *centroid* masing-masing. Hasil ini juga mencakup posisi akhir *centroid* serta pola pengelompokan data yang dapat digunakan sebagai dasar analisis lebih lanjut. Dengan mekanisme tersebut, *K-Means* menjadi salah satu algoritma *clustering* yang sederhana namun efektif dalam menggambarkan struktur alami dari data yang dianalisis.

2.3.4 Penerapan *K-Means* dalam Analisis Data Pembangunan Daerah

Penerapan metode *K-Means* dalam analisis data proyek pembangunan daerah Labuhanbatu Utara dilakukan untuk pengelompokan berbagai kegiatan pembangunan yang dikelola oleh Dinas Pekerjaan Umum dan Tata Ruang (PUTR) pada tahun 2024 berdasarkan karakteristik tertentu yang menggambarkan kondisi, prioritas, serta tingkat kebutuhan setiap wilayah. Dinas PUTR Kab. Labuhanbatu Utara memiliki berbagai jenis proyek, seperti pembangunan infrastruktur jalan, rehabilitasi jembatan, peningkatan kualitas drainase, pembangunan fasilitas umum, serta penyediaan sarana dasar lainnya. Banyaknya jenis proyek dengan nilai anggaran, lokasi, dan tingkat penyelesaian yang berbeda-beda membuat proses analisis secara manual menjadi kurang efektif. Oleh karena itu, *K-Means* digunakan sebagai alat bantu untuk mengelompokkan data sehingga pada pembangunan dapat terlihat lebih jelas dan terstruktur (Yulianti & Nurfadilah, 2020; Mulyati & Surbakti, 2021).

Dalam penerapannya, data proyek pembangunan yang diperoleh dari Dinas Pekerjaan Umum dan Tata Ruang (PUTR) terlebih dahulu dipersiapkan melalui proses *preprocessing*, seperti pembersihan data, pemilihan atribut yang relevan, serta normalisasi data agar setiap variabel memiliki skala yang sebanding. Atribut yang umum digunakan dalam analisis ini meliputi nilai anggaran proyek, presentase progres pekerjaan, tingkat urgensi kebutuhan masyarakat, kondisi eksisting infrastruktur, serta sebaran lokasi proyek di berbagai kecamatan. Setelah data diproses, metode *K-Means* diterapkan dengan menentukan jumlah *cluster* yang ingin dibentuk, misalnya tiga *cluster* yang mewakili kategori data

proyek banyak, sedang, rendah. Selanjutnya algoritma *K-Means* akan mengelompokkan proyek berdasarkan kedekatan karakteristik antar data sehingga proyek-proyek dengan pola serupa akan tergabung dalam *cluster* yang sama, sesuai dengan prinsip metode *K-Means* yang telah banyak digunakan untuk pengelompokan infrastruktur daerah (Kusumadewi & Saputra, 2019).

Hasil penerapan *K-Means* pada data pemabangunan daerah Labuhanbatu Utara memeberikan gambaran yang lebih terstruktur mengenai kondisi pembangunan di wilayah tersebut. Melalui *cluster* yang terbentuk, Dinas PUTR dapat mengetahui kelompok proyek yang membutuhkan perhatian khusus, seperti proyek dengan progres rensah tetapi urgensi tinggi, proyek dengan anggaran besar yang membutuhkan pengawasan intensif, atau proyek dengan dampak besar terhadap masyarakat. Selain itu *cluster* yang dihasilkan juga dapat menjadi dasar dalam pengambilan keputusan strategis, seperti penyusunan prioritas anggaran, perencanaan pembangunan tahun berikutnya, serta evaluasi kinerja pelaksanaan proyek di berbagai kecamatan. Temuan ini sejalan dengan penelitian sebelumnya yang menunjukkan bahwa *clustering* mampu memeberikan wawasan penting untuk mendukung pengambilan keputusan pembangunan daerah (Astuti & Rahmadani, 2022; Simangunsong & Siregar, 2020).

2.3.5 Rumus Perhitungan Jarak (*Euclidean Distance*)

Untuk menentukan kedekatan antara data dengan *centroid*, metode *K-Means* umumnya menggunakan *Euclidean Distance*. Rumusnya adalah sebagai berikut :

$$d(x_i, c_j) = \sqrt{\sum_{k=0}^n (x_{ik} - c_{jk})^2}$$

Keterangan :

- $d(x_i, c_j)$: jarak data ke – i dengan *centroid* ke- j
- x_{ik} : nilai atribut ke- k data ke- i
- c_{jk} : nilai atribut ke- k dari *centroid* ke- j
- n : jumlah atribut

Data akan dimasukkan ke dalam cluster yang memiliki jarak paling kecil terhadap *centroid*.

2.3.6 Langkah-langkah Penentuan dan Perhitungan *Centroid*

Langkah-langkah dalam menentukan *centroid* pada metode *K-Means* adalah sebagai berikut :

1. Menentukan jumlah *cluster* (k)

Peneliti menentukan jumlah *cluster* yang akan digunakan sesuai dengan kebutuhan penelitian, 3 *cluster* (banyak, sedang, rendah).

2. Menentukan *centroid* awal

Centroid awal ditentukan secara acak dari data yang tersedia atau berdasarkan nilai tertentu sebagai titik awal perhitungan.

3. Menghitung jarak data ke *centroid*

Setiap data dihitung jaraknya terhadap seluruh *centroid* menggunakan rumus *Euclidean Distance*.

4. Pengelompokan data

Data dimasukkan ke dalam cluster yang memiliki jarak terdekat dengan *centroid*.

5. Menghitung *centroid* baru

Setelah semua data dikelompokkan, *centroid* baru dihitung dengan mencari nilai rata-rata dari seluruh data dalam masing-masing *cluster*.

6. Iterasi

Langkah perhitungan jarak dan pembaruan *centroid* diulang sampai tidak terjadi perubahan cluster atau nilai *centroid* sudah stabil.

2.4 **RapidMiner**

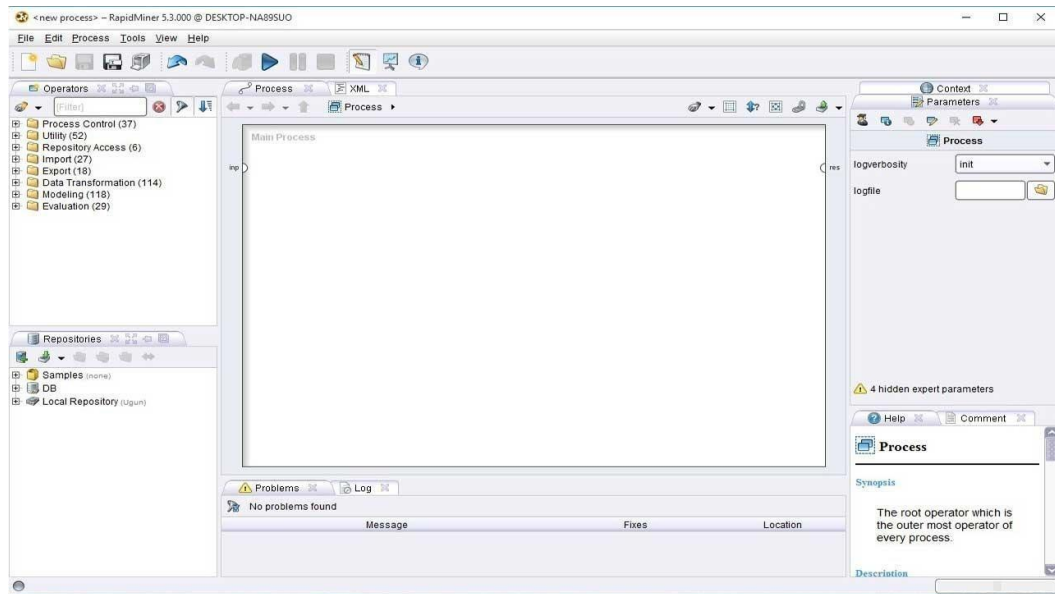
RapidMiner adalah perangkat lunak *data science* dan *data mining* berbasis *open-source* yang dirancang untuk memfasilitasi proses analisis data, pembelajaran mesin, dan prediksi secara visual (Witten, Frank, Hall, & Pal, 2016). *RapidMiner* menyediakan antarmuka *drag-and-drop*, sehingga pengguna dapat membuat alur kerja (*workflow*) analisis tanpa harus menulis kode pemrograman secara manual. Platform ini mendukung berbagai tahapan analisis data, mulai dari *preprocessing*, transformasi data, pemodelan, evaluasi, hingga visualisasi hasil (Tan, Steinbach, & Kumar, 2019). Dengan lebih dari 500 operator yang tersedia, *RapidMiner* memungkinkan integrasi berbagai metode *data mining* seperti klasifikasi, regresi, *clustering*, asosiasi, dan prediksi.

Dalam konteks *clustering*, khususnya metode *K-Means*, *RapidMiner* mempermudah proses pengelompokan data dengan menyediakan operator khusus yang dapat langsung diterapkan pada dataset (Sari, Sudewa, Lestari, & Jaya, 2020). Pengguna dapat mengimpor data pembangunan daerah, misalnya indikator

infrastruktur, anggaran, dan capaian proyek, dari berbagai format seperti CSV, *Excel*, atau *database*. Sebelum dilakukan *clustering*, data dapat diolah melalui *preprocessing*, seperti normalisasi, penghapusan *missing value*, dan transformasi variabel, agar hasil *clustering* menjadi lebih akurat (Nur & Fatah, 2024). Operator *K-Means* di *RapidMiner* dapat diatur parameter seperti jumlah *cluster* (*k*), metode jarak, jumlah iterasi maksimal, dan pengaturan *centroid* awal. Setelah proses *clustering* selesai, *RapidMiner* menyediakan visualisasi hasil berupa grafik *cluster*, distribusi data, dan statistik *centroid* yang mempermudah interpretasi hasil (Idham, Rosika, & Yuliadi, 2022)

Kelebihan *RapidMiner* dibandingkan alat analisis lainnya adalah *user-friendly*, *visualisasi workflow* yang jelas, dan mudah direproduksi, sehingga peneliti maupun praktisi dapat meninjau ulang analisis dengan mudah. Beberapa penelitian telah membuktikan efektivitas *RapidMiner* dalam pengelompokan data, misalnya untuk segmentasi pelanggan, evaluasi kinerja sekolah, analisis data kesehatan, maupun analisis pembangunan daerah (Sari et al., 2020; Nur & Fatah, 2024; Idham et al., 2022). Dengan *RapidMiner*, Dinas Pekerjaan Umum dan Tata Ruang dapat melakukan analisis data pembangunan secara sistematis, menghasilkan *cluster* data pembangunan proyek berdasarkan karakteristik pembangunan, dan memberikan dasar rekomendasi kebijakan yang lebih tepat sasaran. Secara keseluruhan, *RapidMiner* merupakan alat bantu yang tepat untuk penelitian ini karena memadukan kemudahan penggunaan, kemampuan analisis yang lengkap, serta visualisasi yang mendukung pengambilan keputusan berbasis data (Witten et al., 2016; Tan et al., 2019). Penerapan *RapidMiner* dalam penelitian

ini diharapkan dapat membantu mengelompokkan wilayah Labuhanbatu Utara berdasarkan tingkat pembangunan, sehingga Dinas Pekerjaan Umum dan Tata Ruang dapat menyusun strategi pembangunan yang lebih efektif dan efisien.



Gambar 2.2 Tampilan Awal RapidMiner

2.5 Penelitian Terdahulu

Tabel 2.1 Penelitian Terdahulu

Referensi Penelitian	1
Judul	"Penerapan Algoritma <i>K-Means Clustering</i> untuk Pengelompokan Pembangunan Jalan pada Dinas Pekerjaan Umum dan Penataan Ruang"
Nama Penulis	Dede Kurniadi, Yoga Handoko Agustin, Hari Ilham Nur Akbar & Ida Farida
Tahun	2023

Hasil	Hasil penelitian menunjukkan algoritma <i>K-Means</i> , ketika diterapkan pada data pembangunan jalan milik Dinas Pekerjaan Umum dan Tata Ruang Kabupaten Garut, mampu menghasilkan pengelompokan (<i>klaster</i>) jalan dengan sangat baik berdasarkan atribut tipe jalan, panjang jalan, dan biaya pembangunan. Peneliti menggunakan pendekatan CRISP-DM (<i>Cross-Industry Standart Process for Data Mining</i>) untuk proses <i>data mining</i> secara sistematis.
Referensi Penelitian	2
Judul	" <i>Data Mining</i> Strategi Pembangunan Infrastruktur Menggunakan Algoritma <i>K-Means</i> "
Nama Penulis	Mangapul Siahaan
Tahun	2022
Hasil	Penelitian ini menganalisis data pembangunan infrastruktur dari beberapa wilayah menggunakan algoritma <i>K-Means</i> untuk memetakan tingkat prioritas pembangunan. Variabel yang digunakan meliputi panjang jalan, jumlah fasilitas umum, anggaran pembangunan, dan kondisi infrastruktur. Hasil penelitian menunjukkan bahwa metode <i>K-Means</i> mampu mengelompokkan wilayah menjadi tiga klaster utama yaitu, klaster tinggi, klaster sedang, dan klaster rendah. Penelitian ini menyimpulkan bahwa <i>K-Means</i> dapat membantu pemerintah menemukan prioritas pembangunan infrastruktur secara objektif berdasarkan data.
Referensi Penelitian	3

Judul	<i>"K-Means Cluster Algoritm for Grouping Inequality in Regional Developmeant"</i>
Nama Penulis	Tb Ai Munandar & Dwipa Handayani
Tahun	2023
Hasil	Penelitian ini menganalisis ketimpangan pembangunan regional menggunakan algoritma <i>K-Means</i>. Penelitian ini memanfaatkan indikator pembangunan yang meliputi infrastruktur dasar, akses jalan, kondisi ekonomi dan indsutri, angka pendidikan, serta layanan kesehatan. Hasil klasterisasi menunjukkan terbentuknya tiga kelompok ketimpangan pembangunan, yaitu klaster ketimpangan tinggi, yang mencakup daerah tertinggal dengan infrastruktur yang sangat terbatas, akses transportasi yang buruk, serta banyak desa yang belum memiliki layanan kesehatan memadai. Klaster sedang, terdiri dari wilayah yang beberapa fasilitasnya sudah tersedia namun belum merata, dengan tingkat pendidikan menengah dan kondisi ekonomi yang belum stabil. Kalster ketimpangan rendah, meliputi daerah dengan pembangunan relatif merata, infrastruktur lengkap, dan laju pertumbuhan ekonomi yang baik. Penelitian ini menunjukkan bahwa algoritma <i>K-Means</i> dapat menjadi alat yang efektif dalam mengidentifikasi dan menganalisis kesenjangan pembangunan antarwilayah, sehingga dapat membantu pemerintah dalam merencanakan kebijakan pembangunan yang lebih tepat sasaran.

Referensi Penelitian	4
Judul	"Pengelompokan Kabupaten/Kota di Provinsi Sumatera Utara Berdasarkan Indikator Indeks Pembangunan Manusia dengan Menggunakan Analisis <i>Cluster</i> Metode <i>K-Means</i> ",
Nama Penulis	Ebdy Rafles Marbun & Chairunnisa
Tahun	2024
Hasil	Penelitian ini menerapkan <i>K-Means</i> untuk mengelompokkan Kabupaten/Kota Sumatera Utara berdasarkan indikator pembangunan manusia, termasuk harapan hidup, rata-rata lama sekolah, harapan lama sekolah, dan pengeluaran perkapita. Penelitian ini menemukan tiga kluster pembangunan, yaitu rendah, sedang, tinggi, di mana kabupaten dengan kluster rendah memiliki keterbatasan akses pendidikan dan kesehatan, sedangkan kluster tinggi mencerminkan kabupaten/kota dengan kualitas hidup dan pendapatan yang lebih baik.
Referensi Penelitian	5
Judul	" <i>Application of the K-Means Clustering Algorithm in Grouping Regencies/Cities in North Sumatra Province Based on Human Development Index Indicators</i> ",
Nama Penulis	Sonia Sihombing, Rahmat W. Sembiring & Eka Irawan
Tahun	2022

Hasil	Penelitian ini menggunakan <i>K-Means</i> untuk mengelompokkan Kabupaten/Kota di Sumatera Utara berdasarkan IPM dan komponennya. Hasil analisis menunjukkan terbentuknya tiga klaster utama, yaitu klaster tinggi, menengah, dan rendah, yang menggambarkan tingkat pembangunan manusia di masing-masing Kabupaten/Kota. Penelitian ini memberikan dasar bagi pemerintah provinsi untuk mengidentifikasi wilayah yang membutuhkan perhatian khusus agar pembangunan lebih merata.
-------	--

Berdasarkan kajian penelitian terdahulu, algoritma *K-Means* terbukti efektif dalam pengelompokan wilayah atau proyek pembangunan berdasarkan indikator pembangunan, kesejahteraan masyarakat, dan infrastruktur. Hasil klasterisasi dari penelitian-penelitian sebelumnya menunjukkan adanya perbedaan karakteristik pembangunan antarwilayah, mulai dari tinggi, sedang, hingga rendah, sehingga dapat menjadi dasar pengambilan keputusan dan perencanaan pembangunan yang lebih tepat sasaran. Penelitian terdahulu juga menekankan pentingnya pemilihan variabel yang relevan, seperti IPM, kemiskinan, pendapatan daerah, serta akses pendidikan dan kesehatan, untuk menghasilkan klaster yang akurat. Oleh karena itu, penelitian ini akan menerapkan algoritma *K-Means* pada data pembangunan daerah Labuhanbatu Utara dengan tujuan memperoleh pengelompokan data proyek pembangunan daerah sehingga dapat mendukung perumusan strategi pembangunan yang efektif dan berbasis data.