DATA MINING PENGOLAHAN DATA CALON PEKERJA MIGRAN INDONESIA (PMI) DENGAN PENERAPAN METODE KLUSTERING K-MEANS DAN METODE KLASIFIKASI K-NEAREST NEIGHBOR (KNN): STUDI KASUS PT. SAM

Dedy[1], Anis Cherid[2]
Jurusan Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana[1,2]
Jl. Raya Meruya Selatan, Kembangan, Jakarta
E-mail: 41518110182@student.mercubuana.ac.id [1], anis.cherid@mercubuana.ac.id[2]

Abstrak

Informasi adalah suatu hal yang sangat penting bagi perusahaan untuk melaksanakan proses bisnisnya secara efektif dan efisien. Informasi dapat diperoleh dari hasil pengolahan data, salah satunya adalah dengan proses Data Mining. Data Mining dapat menggali dan mengolah data menjadi suatu informasi yang sangat penting dan berguna yang mungkin belum diketahui sebelumnya. Clustering menganalisis objek data dimana label kelas tidak diketahui dan dapat digunakan untuk menentukan label kelas tidak diketahui dengan cara mengelompokkan data untuk membentuk kelas baru. Klasifikasi adalah proses menemukan model (fungsi) yang menjelaskan dan membedakan kelas-kelas atau konsep, dengan tujuan agar model yang diperoleh dapat digunakan untuk memprediksikan kelas atau objek yang memiliki label kelas tidak diketahui. Metode clustering diterapkan dalam penelitian ini untuk menghasilkan kelompok (kluster) data yang dapat menggambarkan pola kemiripan karakteristik data atribut penilaian kualitatif penentu dan data atribut lamanya waktu Calon Pekerja Migran Indonesia (CPMI) tersebut dari perekrutan sampai dengan berangkat ke luar negeri untuk bekerja (perhitungan waktu dari tanggal masuk Balai Latihan Kerja (BLK) dan tanggal keberangkatan). Sedangkan Metode klasifikasi KNN diterapkan untuk mengolah dataset hasil pengolahan K-Means sebelumnya dengan tujuan untuk menghasilkan pola klasifikasi data dalam memprediksi klasifikasi nilai atribut data pendukung CPMI baru.
Kata Kunci: Clustering, Classification, CPMI, K-Means, KNN

Abstract

Information is something that is very important for a company to carry out its business processes effectively and efficiently. Information can be obtained from the results of data processing, one of which is the Data Mining process. Data mining can dig and process data into very important and useful information that may not have been known before. Clustering analyzes data objects where the class label is unknown and can be used to determine the unknown class label by grouping the data to form a new class. Classification is the process of finding a model (function) that describes and differentiates classes or concepts, with the aim that the model obtained can be used to predict classes or objects that have an unknown class label. The clustering method is applied in this study to produce data clusters that can describe the pattern of similarity characteristics of the determinant qualitative assessment attribute data and attribute data of the length of time the Prospective Indonesian Migrant Worker (CPMI) from recruitment to going abroad to work (time calculation from the date of entry of the Work Training Center (BLK) and the date of departure). Meanwhile, the KNN classification method is applied to process the dataset of previous K-Means processing results with the aim of producing a data classification pattern in predicting the value classification of the new CPMI supporting data attributes.
Key words: Clustering, Classification, CPMI, K-Means, KNN

I. PENDAHULUAN

Informasi adalah suatu hal yang sangat penting bagi perusahaan atau organisasi untuk melaksanakan proses bisnisnya secara efektif dan efisien dalam rangka mencapai target dan tujuannya. Informasi dapat diperoleh dari hasil pengolahan data, salah satunya adalah dengan proses Data Mining. Data Mining dapat menggali dan mengolah data menjadi suatu informasi yang sangat penting dan berguna yang mungkin belum diketahui sebelumnya.

Begitu pula dengan yang terjadi pada PT. SAM, suatu perusahaan yang bergerak di bidang perekrutan dan penempatan Calon Pekerja Migran Indonesia (CPMI) di luar negeri. Banyak data CPMI yang telah dihasilkan dan disimpan dalam basis data PT. SAM, namun belum pernah dilakukan proses pengolahan data untuk menghasilkan informasi yang penting dan berguna dalam mendukung proses bisnis perusahaan.

Kondisi saat ini, PT. SAM belum memiliki landasan yang bersifat ilmiah berdasarkan data yang ada untuk menggambarkan faktor atau atribut apa saja dari biodata CPMI yang berpengaruh terhadap minat ketertarikan pemberi kerja di luar negeri untuk mempekerjakan mereka di sektor rumah tangga (house maid). Ruang lingkup house maid dapat berupa: asuh balita, asuh anak, tata laksana rumah tangga, dan rawat lansia.

Faktor atau atribut biodata CPMI yang akan diteliti yaitu: asal kota CPMI (tren atau sifat karakter CPMI secara umum), umur (kemampuan produktivitas), tinggi dan berat badan (kondisi fisik), pengalaman kerja (jenis pekerjaan dan lama bekerja sebelumnya), personality (karakter kepribadian), facial expression (ekspresi wajah), tidines (kerapian), house maid experience (penilaian pengalaman kerja), kemampuan bahasa kantonis (bahasa penduduk Hongkong), kemampuan bahasa Mandarin (sebagian besar penduduk Hongkong menguasai bahasa Mandarin), kemampuan bahasa Inggris (sebagian besar penduduk Hongkong dapat berbahasa Inggris) dan lamanya waktu CPMI dari masuk pelatihan sampai dengan keberangkatannya.

Kondisi yang diharapkan dengan adanya penelitian ini, dapat dihasilkannya informasi yang memberikan gambaran tentang pola kemiripan dan klasifikasi dari nilai atribut biodata CPMI sehingga diperoleh informasi yang berarti tentang nilai atribut biodata tersebut dan dapat berguna untuk melakukan prediksi klasifikasi data CPMI bagi keperluan pemasaran PT. SAM. Terkait dengan kondisi tersebut, penelitian ini ditujukan untuk melakukan proses Data Mining terhadap data CPMI dengan penerapan metode klustering K-Means dan metode klasifikasi K-Nearest Neighbor (KNN).

Metode klustering K-Means diterapkan dalam penelitian ini untuk menghasilkan kelompok (kluster) data yang dapat menggambarkan pola kemiripan karakteristik data atribut penilaian kualitatif penentu (misalnya: usia, tinggi badan, berat badan, kemampuan bahasa, pengalaman kerja, keterampilan kerja, perilaku, kepribadian dan lain sebagainya) dan lamanya waktu CPMI tersebut dari perekrutan sampai dengan berangkat ke luar negeri untuk bekerja (perhitungan waktu dari tanggal masuk Balai Latihan Kerja (BLK) dan tanggal keberangkatan).

Dari pengolahan data K-Means sebelumnya, akan dihasilkan dataset dengan atribut data baru berupa nama cluster (misal: Cluster 1, Cluster 2, dst). Dalam proses klasifikasi KNN, atribut data nama cluster tersebut akan menjadi atribut yang akan dipengaruhi (dependent variable) oleh atribut data lain (independent variable) yang sudah ada sebelumnya pada pengolahan data K-Means (misalnya: usia, tinggi badan, berat badan, kemampuan bahasa, pengalaman kerja, keterampilan kerja, perilaku, kepribadian dan lain sebagainya).

Metode klasifikasi KNN diterapkan untuk mengolah dataset hasil pengolahan K-Means sebelumnya dengan tujuan untuk menghasilkan pola klasifikasi data dalam memprediksi klasifikasi nilai atribut data pendukung CPMI baru. Sehingga setelah proses KNN, maka akan diperoleh jumlah tetangga terdekat (nearest neighbor) sesuai klasifikasi masing-masing yang akan digunakan sebagai prediksi atau proyeksi klasifikasi dari suatu nilai independent variable yang baru. Contoh: jika ada data CPMI baru dengan nilai data: Usia = 30 tahun, Tinggi Badan = 160 cm, Berat Badan = 60 kg, Pengalaman Kerja di Taiwan = 2 tahun, maka melalui proses klasifikasi KNN, data tersebut akan diolah untuk diperoleh hasil prediksi klasifikasinya, misalnya data tersebut masuk dalam klasifikasi kelompok 1.

Perbedaan dengan penelitian terkait sebelumnya adalah penelitian ini akan dilakukan dalam 2 (dua) tahap "mining" yaitu metode klustering dan klasifikasi. Perbedaan selanjutnya adalah dari sumber data yang akan digunakan dalam penelitian yaitu penelitian ini menggunakan data CPMI pada PT. SAM dengan negara tujuan Hongkong.

Terdapat penelitian lain terkait dengan Data Mining metode klustering K-Means dan klasifikasi KNN yaitu:

1. Analisa data kecelakaan dengan K-Means untuk mengetahui faktor-faktor penyebab kecelakaan lalu lintas yang terjadi pada tahun 2016 s.d. 2017 di Kecamatan Pelaihari Kabupaten Tanah Laut Provinsi Kalimantan Selatan yang menggunakan dataset berupa data lokasi dan data lainnya yang terkait dengan terjadinya kecelakaan lalu lintas [1].
2. Analisa Data Mining *Customer Lifetime Value* (CLV - digunakan dalam menghitung nilai profitabilitas pelanggan) pada setiap segmen pelanggan dengan algoritma K-Means untuk klusterisasi pelanggan. Penelitian ini menggunakan dataset PT. HPAI Cabang Pekanbaru pada periode analisis Februari 2017 sampai September 2017 [2].
3. Analisa *big data E-commerce* menggunakan algoritma clustering K-Means dengan dataset berupa catatan transaksi yang dikumpulkan dari UCI Machine Learning Repository dengan pengelompokan berdasarkan CountryID dan Customer ID [3].
4. Analisa model KNN-LSTM dalam meningkatkan akurasi prediksi arus lalu lintas spatiotemporal. KNN digunakan untuk memilih sebagian besar stasiun tetangga yang terkait dengan stasiun uji dan menangkap fitur spasial dari arus lalu lintas. LSTM digunakan untuk menambang variabilitas temporal dari arus lalu lintas, dan jaringan LSTM dua lapis diterapkan untuk memprediksi aliran lalu lintas masing-masing di stasiun yang dipilih. Hasil prediksi akhir diperoleh dengan fusi tingkat hasil dengan metode pembobotan pangkat-eksponen.

Kinerja prediksi dievaluasi dengan data arus lalu lintas waktu nyata yang disediakan oleh Transport Data Research Lab (TDRL) di Pusat Data University of Minnesota Duluth (UMD) [4].

5. Analisa klasifikasi KNN untuk melakukan klasifikasi Lao Text dalam proses kategorisasi teks. Kategorisasi teks adalah skenario aplikasi umum di bidang NLP (NLP digunakan untuk menganalisis teks agar mesin memahami bahasa manusia). Saat ini, hanya ada beberapa klasifikasi untuk teks Lao. Algoritma KNN menghitung bahwa sebagian besar tetangga terdekat dalam ruang fitur termasuk kategori tertentu, dan sampel juga termasuk dalam kategori ini. Algoritma ini melibatkan beberapa faktor utama: pengukuran jarak, pemilihan nilai k dan sebagainya [5].
6. Prediksi pola kelulusan mahasiswa menggunakan teknik *data mining classification emerging pattern*. Untuk memprediksi pola dan menganalisa tingkat kelulusan mahasiswa peneliti menggunakan data mining proses klasifikasi menggunakan algoritma *emerging pattern*. Dalam penelitian ini data yang digunakan berasal dari data gabungan antara data induk mahasiswa dan data kelulusan. Hasil pengujian data yang dilakukan oleh peneliti pada aplikasi data mining ini menghasilkan pola kelulusan dengan berbagai variasi sesuai dengan atribut learning yang digunakan yaitu jenis kelamin, angkatan, program studi, system kuliah dan IPK mahasiswa. [6].

Topik penelitian ini adalah pengolahan data Calon Pekerja Migran Indonesia (CPMI) dengan proses Data Mining metode klustering K-Means dan klasifikasi KNN.

Rumusan Masalah penelitian ini adalah:

1. bagaimana hasil pengelompokan (klustering) K-Means data CPMI dapat memberikan informasi tentang pola kemiripan karakteristik data atribut penilaian penentu dan atribut lamanya waktu keberangkatan CPMI;
2. bagaimana hasil klasifikasi KNN data CPMI dapat memberikan prediksi klasifikasi dari data CPMI baru.

Tujuan dilaksanakannya penelitian ini adalah untuk melakukan proses Data Mining dengan penerapan metode K-Means dan KNN yang hasilnya dapat memberikan informasi kepada PT. SAM berupa:

1. pengelompokan data berdasarkan pola keimiripan karakteristik data atribut penilaian dan lamanya waktu keberangkatan CPMI;
2. klasifikasi data CPMI yang dapat melakukan prediksi klasifikasi CPMI baru.

Manfaat penelitian ini bagi:

1. ilmu pengetahuan, dapat digunakan sebagai referensi untuk studi pengolahan data Data Mining dengan penerapan K-Means dan KNN;
2. masyarakat khususnya PT. SAM, memberikan informasi tentang gambaran faktor penilaian penentu CPMI yang berpengaruh dalam minat pemberi kerja di luar negeri.

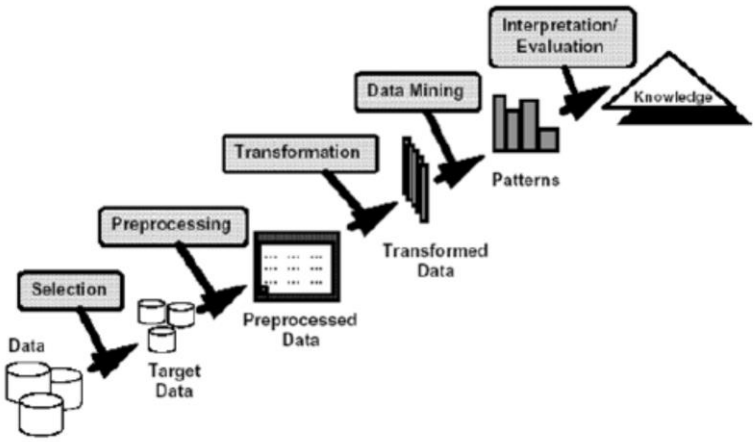
Batasan masalah penelitian ini adalah hanya untuk pengolahan data CPMI pada PT. SAM untuk penempatan negara tujuan Hongkong, dengan metode Data Mining penerapan klustering K-Means dan klasifikasi KNN dengan rencana jumlah dataset kurang lebih 353 (tiga ratus lima puluh tiga) data.

II. METODOLOGI PENELITIAN

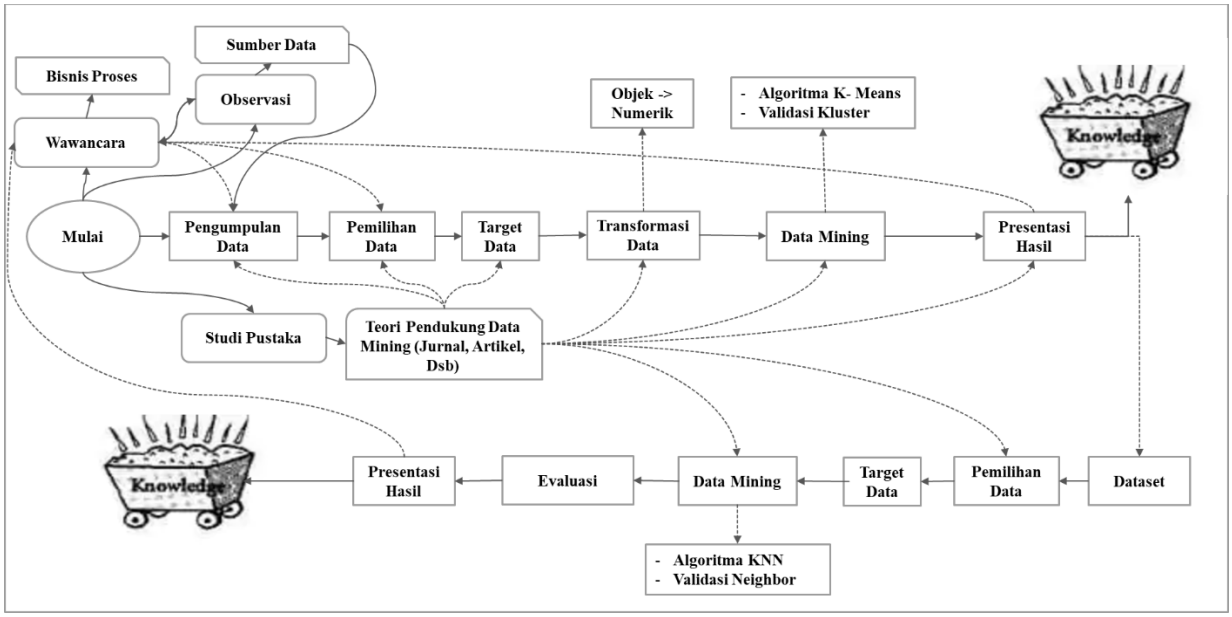
Berdasarkan tingkat eksplanasinya (tingkat kejelasan), penelitian ini dapat digolongkan sebagai penelitian yang bersifat asosiatif karena bertujuan untuk mengetahui pengaruh ataupun juga hubungan antara dua variabel atau lebih yang membangun suatu teori yang dapat berfungsi untuk menjelaskan, memprediksi dan mengontrol suatu gejala.

Menurut Hermawati dalam jurnal (Nurdin, Dewi Astika, 2015) tahapan proses dalam penggunaan data mining yang merupakan proses Knowledge Discovery in Database (KDD) dapat diuraikan sebagai berikut [7]:

a. memahami domain aplikasi untuk mengetahui dan menggali pengetahuan awal serta apa sasaran pengguna;
b. membuat target data-set yang meliputi pemilihan data dan fokus pada subset data.
c. pembersihan dan transformasi data meliputi eliminasi derau, outliers, missing value, serta pemilihan fitur dan reduksi dimensi.
d. penggunaan algoritma data mining yang terdiri dari asosiasi, sekuensial, klasifikasi, klasterisasi, dll.
e. Evaluasi dan visualisasi pola untuk melihat apakah ada sesuatu yang baru dan menarik dan dilakukan iterasi jika diperlukan.



Gambar 1. Proses KDD (Hermawati, 2015)



Gambar 2. Metodologi Penelitian

Kerangka proses data mining tersusun atas tiga tahapan, yaitu pengumpulan data (data collection), transformasi data (data transformation), dan analisis data (data analysis). Proses tersebut diawali dengan preprocessing yang terdiri atas pengumpulan data untuk menghasilkan data mentah (raw data) yang dibutuhkan oleh data mining, yang kemudian dilanjutkan dengan transformasi data untuk mengubah data mentah menjadi format yang dapat diproses oleh data mining, misalnya melalui filtrasi atau agregasi. Hasil transformasi data akan digunakan oleh analisis data untuk membangkitkan pengetahuan dengan menggunakan teknik seperti analisis statistik, machine learning, dan visualisasi informasi [8].

Penjelasan Metodologi Penelitian (Gambar 2.):

a. Metode pengumpulan data yang akan dilakukan dalam penelitian ini adalah:
   1. Studi literatur;
   2. Wawancara, dilakukan dengan karyawan perwakilan PT. SAM diperoleh hasil berupa informasi tentang profil perusahaan dan bisnis proses CPMI dari awal perekrutan sampai dengan penempatannya di luar negeri. Dari wawancara tersebut juga diperoleh informasi bahwa PT. SAM telah memiliki sertifikasi ISO 9001-2015 Sistem Manajemen pada tahun 2016 berlaku s.d. 2019 (BSI) dan tahun 2019 s.d. 2022 (IMS).
   3. Observasi langsung ke kantor PT. SAM dengan kegiatan yang dilakukan berupa pengumpulan informasi tentang sumber data yang akan digunakan untuk penelitian. Diperoleh hasil berupa sampel sumber data yang akan digunakan berupa: Form Biodata CPMI, Data Registrasi CPMI pada Sistem SISKOTKLN (BP2MI) dan Buku Laporan Keberangkatan CPMI.
b. Metode Data Mining dengan K-Means
   1. Pengumpulan Data

a) tahap awal proses penelitian ini adalah tahap pengumpulan data dengan metode wawancara dan observasi untuk memperoleh data primer, serta dengan studi pustaka untuk memperoleh teori pendukung tentang mekanisme proses pengumpulan data dan tahap selanjutnya Data Mining;
b) pada tahap pengumpulan data ini diperoleh informasi tentang sumber data yang akan digunakan dalam penelitian yaitu:
   1) Data dari Sistem SISKOTKLN Badan Perlindungan Pekerja Migran Indonesia (BP2MI) dengan akun PT. SAM. Pada form ini terdapat sekitar 40 (empat puluh) atribut data antara lain: Nomor ID, Biodata CPMI, Biodata orang tua CPMI, Tanggal Masuk BLK, Tanggal Selesai BLK, data lainnya);
   2) Data dari Form Biodata CPMI PT. SAM, pada form ini terdapat sekitar 20 (dua puluh) atribut data antara lain: biodata CPMI, Tinggi dan Berat Badan, Pengalaman Kerja, Penilaian Kepribadian, Kemampuan Bahasa, dan lainnya;
   3) Data dari Buku Laporan Keberangkatan CPMI dari Marketing Hongkong PT. SAM, pada form ini terdapat atribut data: Kode CPMI, Asal Kabupaten/Kota CPMI, Tempat/Tanggal Lahir, Nama Pemberi Kerja, Nama Agensi di Hongkong, Tanggal Penerbangan, dan lainnya.

2. Pemilihan Data
   Tahap selanjutnya adalah Pemilihan Data, pada tahapan ini adalah proses pemilihan atribut data sesuai dengan kebutuhan penelitian. Penelitian ini berfokus pada atribut penilaian pendukung CPMI yang menjadi faktor minat pemberi kerja dan lamanya waktu dari CPMI masuk pelatihan BLK.
   Faktor atau atribut biodata CPMI yang akan diteliti yaitu: asal kota CPMI (tren atau sifat karakter CPMI secara umum), usia (kemampuan produktivitas), tinggi dan berat badan (kondisi fisik), pengalaman kerja (jenis pekerjaan dan lama bekerja sebelumnya), personality (karakter kepribadian), facial expression (ekspresi wajah), tidines (kerapian), house maid experience (penilaian pengalaman kerja), kemampuan bahasa kantonis (bahasa penduduk Hongkong), kemampuan bahasa Mandarin (sebagian besar penduduk Hongkong menguasai bahasa Mandarin), kemampuan bahasa Inggris (sebagian besar penduduk Hongkong dapat berbahasa Inggris) serta lamanya hari CPMI dari masuk pelatihan sampai dengan keberangkatannya.
3. Transformasi dan Target Data
   Tahapan berikutnya adalah persiapan Target Data dan Transformasi Data (perubahan nilai data objek menjadi numerik).
4. Penerapan Algoritma K-Means
   Proses berikutnya adalah Tahapan Data Mining dengan metode klustering K-Means. Pada tahapan ini, dataset yang telah ditransformasikan akan diolah menggunakan algoritma K-Means untuk memperoleh hasil pengelompokan data berdasarkan pola kemiripan karakteristik data pada dataset dengan atribut baru berupa label kelompok (misal: cluster 1, cluster 2, dst). Melakukan penelitian validasi jumlah kluster yang tepat sesuai dengan karakteristik dari dataset yang ada.
   Metode K-Means adalah metode yang dikategorikan ke dalam metode clustering partisi (Kandeil, Saad and Youssef, 2014). Algoritma ini paling banyak digunakan dalam CRM dan pemasaran (Kandeil, Saad and Youssef, 2014). Hal ini disebabkan karena algoritma ini memiliki kesederhanaan dan kemudahan dalam menggunakannya serta pengguna bisa menentukan sendiri jumlah cluster nya. Tahapan dalam metode K-means adalah sebagai berikut:
   1. tentukan jumlah kluster;
   2. pilih centroid awal secara acak sesuai jumlah kluster;
   3. hitung jarak data ke centroid dengan rumus Euclidean Distance:
      dxy =$\sqrt{\Sigma}(xi-yi)2ni=1$
   4. perbaharui centroid dengan menghitung nilai rata-rata nilai pada masing-masing kluster;
   5. kembali ke tahapan ke 3 jika masih terdapat data yang berpindah kluster atau perubahan nilai centroid.

   Pada tahap 1 dan 2 dari tahapan K-Means, jumlah kluster dapat ditentukan dengan metode dunn index, silhoutte coefficient, atau Elbow yang merupakan metode validasi dalam menentukan jumlah kluster yang terbaik.
5. Evaluasi dan Presentasi Hasil
   - Evaluasi adalah fase lanjutan terhadap tujuan data mining. Evaluasi dilakukan secara mendalam dengan tujuan agar hasil pada tahap pemodelan sesuai dengan sasaran yang ingin dicapai dalam tahap *business understanding.*
   - Dalam tahap ini hasil dari teknik data mining berupa pola-pola yang khas dievaluasi untuk menilai apakah hipotesa yang ada memang tercapai. Bila ternyata hasil yang diperoleh tidak sesuai hipotesa ada beberapa alternatif yang dapat diambil seperti: menjadikannya umpan balik untuk memperbaiki proses data mining, mencoba teknik data mining lain yang lebih sesuai, atau menerima hasil ini sebagai suatu hasil yang di luar dugaan yang mungkin bermanfaat.
   - Tahap terakhir dari proses data mining adalah bagaimana memformulasikan keputusan atau aksi dari hasil analisa yang didapat. Ada kalanya hal ini harus melibatkan orang-orang yang tidak memahami data mining. Karenanya presentasi hasil data mining dalam bentuk pengetahuan yang bisa dipahami

semua orang adalah satu tahapan yang diperlukan dalam proses data mining. Dalam presentasi ini, visualisasi juga bisa membantu mengkomunikasikan hasil data mining.

c. Data Mining dengan KNN

1. Pengumpulan Data
   Dataset yang akan digunakan adalah data hasil pengolahan K-Means sebelumnya ditambah satu atribut baru "Nama Cluster".
2. Pemilihan Data
   Melakukan pemilihan data yang sesuai dengan fokus penelitian dan presentasi hasil K-Means sebelumnya yang sudah dapat menentukan atribut utama yang menjadi minat pemberi kerja.
3. Target Data
   Menyusun dataset ke dalam bentuk yang siap diolah untuk tahapan berikutnya.
4. Penerapan Algoritma KNN
   - Metode klasifikasi KNN diterapkan untuk mengolah dataset hasil pengolahan K-Means sebelumnya dengan tujuan untuk menghasilkan pola klasifikasi data dalam memprediksi klasifikasi nilai atribut data pendukung CPMI baru. Dari pengolahan data K-Means sebelumnya, akan dihasilkan dataset dengan atribut data baru yaitu "Label Cluster" dengan nilai tertentu (Cluster 0, Cluster 1, dst).
   - Dalam proses klasifikasi KNN, atribut data Label Cluster tersebut akan menjadi atribut yang akan dipengaruhi (dependent variable) oleh atribut data lain (independent variable) yang sudah ada sebelumnya pada pengolahan data K-Means. Sehingga setelah proses KNN, maka akan diperoleh jumlah tetangga terdekat (nearest neighbor) sesuai klasifikasi masing-masing yang akan digunakan sebagai prediksi atau proyeksi klasifikasi dari suatu nilai independent variable yang baru. Melakukan validasi jumlah k (tetangga terdekat) data latih dan data test untuk meningkatkan akurasi hasil.
   - K-Nearest Neighbor (KNN) adalah sebuah metode supervised yang berarti membutuhkan data training untuk mengklasifikasikan objek yang jaraknya paling dekat. Prinsip kerja K-Nearest Neighbor adalah mencari jarak terdekat antara data yang akan dievaluasi dengan k tetangga (neighbor) dalam data pelatihan (Whidhiasih et al., 2013). Pada proses pelatihan, data dikelompokkan secara manual sesuai dengan kategori yang telah ditentukan. Setelah itu data tersebut akan melalui tahapan preprocessing yang akan menghasilkan bobot untuk setiap kata yang ada di semua dokumen latih. Selanjutnya menghitung kemiripan vektor data uji dengan setiap data latih yang telah diklasifikasikan. Untuk mengetahui kemiripan data digunakan metode cosine similarity (Ridok dan Indriati, 2015).
   - Metode ini dapat digunakan untuk menginterpretasikan jarak tiap data berdasarkan kemiripan data (Rivki dan Bachtiar, 2017). Perhitungan jarak dengan metode cosine similarity dapat dilihat pada Persamaan:
     $(i, k) = \sum(didk)\ \sqrt{\sum}\ dik\ 2\ k\ \sqrt{\sum}\ djk\ 2\ k$ (5)
     dimana:
     $\sum(didk)$: vector dari produk i dan k $\sqrt{\sum}\ dik\ 2\ k$: panjang dari vector i $\sqrt{\sum}\ djk\ 2\ k$: panjang dari vector j i : data uji ke-i j : data latih ke-j
   - Selanjutnya yaitu mengurutkan jarak tersebut berdasarkan nilai terkecil (terdekat) hingga yang terbesar (terjauh). Kemudian menentukan jumlah tetangga (nilai k) yang ingin digunakan sebagai acuan untuk proses klasifikasi. Dari nilai k inilah dapat ditentukan kategori data berdasarkan nilai jarak terdekat.
5. Evaluasi dan Presentasi Hasil
   - Evaluasi adalah fase lanjutan terhadap tujuan data mining. Evaluasi dilakukan secara mendalam dengan tujuan agar hasil pada tahap pemodelan sesuai dengan sasaran yang ingin dicapai dalam tahap *business understanding.*
   - Dalam tahap ini hasil dari teknik data mining berupa pola-pola yang khas dievaluasi untuk menilai apakah hipotesa yang ada memang tercapai. Bila ternyata hasil yang diperoleh tidak sesuai hipotesa ada beberapa alternatif yang dapat diambil seperti: menjadikannya umpan balik untuk memperbaiki proses data mining, mencoba teknik data mining lain yang lebih sesuai, atau menerima hasil ini sebagai suatu hasil yang di luar dugaan yang mungkin bermanfaat.
   - Tahapan akhir dari proses ini adalah presentasi hasil KNN dengan menggunakan beberapa data CPMI baru yang akan diprediksi klasifikasinya, sehingga dapat memberikan informasi terkait gambaran klasifikasi CPMI sekarang yang akan diproses penempatan bekerjanya di Hongkong untuk mendukung strategi perusahaan.

III. HASIL DAN PEMBAHASAN

A. Hasil Data Mining K-Means

a. Pengumpulan Data Dataset dari Sumber Data

Pada tahapan ini, penulis memilih atribut yang relevan dengan tujuan penelitian ini. Berdasarkan hasil wawancara dan observasi, atribut yang relevan tersebut adalah berupa atribut yang terkait dengan identitas

dan nilai dari faktor pendukung CPMI untuk diminati oleh pemberi kerja di Hongkong. Atribut-atribut tersebut yaitu:

1. No (nomor urut berupa angka yang unik mengidentifikasikan terhadap satu Kode CPMI);
2. Kode CPMI (data unik yang berupa huruf dan angka yang diberikan perusahaan untuk mengidentifikasikan satu CPMI, data ini diperoleh dari sumber data Buku Laporan Keberangkatan CPMI;
3. Kabupaten/Kota Asal CPMI (data nama kabupaten/Kota asal CPMI berdasarkan dokumen Kartu Tanda Penduduk (KTP) yang bersangkutan, data ini diperoleh dari dari sumber data Buku Laporan Keberangkatan CPMI);
4. Tanggal Lahir (data berupa tanggal lahir CPMI berdasarkan data KTP yang bersangkutan, data ini diperoleh dari dari sumber data Buku Laporan Keberangkatan CPMI);
5. Umur (perhitungan umur CPMI yang bersangkutan saat masuk pelatihan, data ini diperoleh dari perhitungan menggunakan formula Ms. Excel dengan pengurangan tanggal masuk pelatihan dengan tanggal lahir yang bersangkutan);
6. Tinggi Badan (tinggi badan CPMI berdasarkan hasil pengukuran yang dilakukan oleh petugas wawancara perusahaan saat CPMI akan masuk pelatihan, data ini diperoleh dari sumber data Form Biodata CPMI);
7. Berat Badan (berat badan CPMI berdasarkan hasil pengukuran oleh petugas wawancara perusahaan saat CPMI akan masuk pelatihan, data ini diperoleh dari Form Biodata CPMI);
8. Negara/Daerah Pengalaman Kerja Terakhir (nama negara atau daerah tempat terakhir CPMI bekerja berdasarkan hasil wawancara oleh petugas wawancara saat CPMI akan masuk pelatihan, data ini diperoleh dari Form Biodata CPMI);
9. Lama Pengalaman Kerja Terakhir (jumlah tahun bekerja terakhir CPMI sebelum masuk pelatihan berdasarkan hasil wawancara oleh petugas wawancara saat CPMI akan masuk pelatihan, data ini diperoleh dari Form Biodata CPMI);
10. Jenis Pekerjaan Pengalaman Kerja Terakhir (data berupa jenis pekerjaan pengalaman kerja terakhir CPMI sebelum masuk pelatihan berdasarkan hasil wawancara oleh petugas wawancara saat CPMI akan masuk pelatihan, data ini diperoleh dari Form Biodata CPMI). Data ini bernilai: Asuh Balita, Asuh Anak, Rawat Lansia, dan Tata Laksana Rumah Tangga;
11. Personality (data penilaian dari hasil wawancara oleh petugas wawancara pada saat CPMI akan masuk pelatihan terkait dengan kepribadian dari CPMI, data ini diperoleh dari Form Biodata CPMI). Data ini bernilai: Good, Fair, dan Poor;
12. Facial Expression (data penilaian dari hasil wawancara oleh petugas wawancara pada saat CPMI akan masuk pelatihan terkait dengan ekspresi wajah dari CPMI, data ini diperoleh dari Form Biodata CPMI). Data ini bernilai: Good, Fair, dan Poor;
13. Tidines (data penilaian dari hasil wawancara oleh petugas wawancara pada saat CPMI akan masuk pelatihan terkait dengan kerapihan dari CPMI, data ini diperoleh dari Form Biodata CPMI). Data ini bernilai: Good, Fair, dan Poor;
14. Housemaid Experience (data penilaian dari hasil wawancara oleh petugas wawancara pada saat CPMI akan masuk pelatihan terkait dengan kemampuan tata laksana rumah tangga dari CPMI, data ini diperoleh dari Form Biodata CPMI). Data ini bernilai: Good, Fair, dan Poor;
15. Kemampuan Bahasa Kantonis (data penilaian dari hasil wawancara oleh petugas wawancara pada saat CPMI akan masuk pelatihan terkait dengan kemampuan bahasa Kantonis dari CPMI, data ini diperoleh dari Form Biodata CPMI). Data ini bernilai: Proses Belajar dan Berkomunikasi;
16. Kemampuan Bahasa Mandarin (data penilaian dari hasil wawancara oleh petugas wawancara pada saat CPMI akan masuk pelatihan terkait dengan kemampuan bahasa Mandarin dari CPMI, data ini diperoleh dari Form Biodata CPMI). Data ini bernilai: Good, Fair dan Poor;
17. Kemampuan Bahasa Inggris (data penilaian dari hasil wawancara oleh petugas wawancara pada saat CPMI akan masuk pelatihan terkait dengan kemampuan bahasa Inggris dari CPMI, data ini diperoleh dari Form Biodata CPMI). Data ini bernilai: Good, Fair dan Poor;
18. Tanggal Masuk Pelatihan (data berupa tanggal masuk pelatihan CPMI, data diperoleh dari List PMI pada Siskotkln BP2MI menggunakan user login PT. SAM);
19. Tanggal Berangkat (data berupa tanggal keberangkatan CPMI ke Hongkong untuk bekerja, data diperoleh dari Buku Laporan Keberangkatan CPMI);
20. Total Hari CPMI Dari Masuk Pelatihan Sampai Dengan Berangkat (data berupa jumlah hari hasil perhitungan pengurangan dengan formula Ms. Excel dari tanggal berangkat dengan tanggal masuk pelatihan).

NO	KABUPATEN/KOTA ASAL	TANGGAL LAHIR	UMUR (TAHUN)	TINGGI BADAN (CM)	BERAT BADAN (KG)	PENGALAMAN KERJA TERAKHIR	LAMA PENGALAMAN KERJA (TAHUN)	PENGALAMAN KERJA (JENIS PEKERJAAN)	PERSONALITY	FACIAL EXPRESSION	TIDINES	HOUSEMAID EXPERIENCE	KEMAMPUAN BAHASA KANTONIS	KEMAMPUAN BAHASA MANDARIN SAAT MASUK PELATIHAN	KEMAMPUAN BAHASA INGGRIS SAAT MASUK PELATIHAN	TGL MASUK PELATIHAN	TGL BERANGKAT	TOTAL HARI (MASUK PELATIHAN S.D BERANGKAT)
1	INDRAMAYU	10/05/1988	28	150	46	SINGAPORE	2	ASUH ANAK	GOOD	GOOD	GOOD	GOOD	PROSES BELAJAR	POOR	GOOD	08/06/2016	05/09/2016	89
2	PASURUAN	11/06/1994	22	153	63	SURABAYA	2	ASUH ANAK	GOOD	GOOD	GOOD	GOOD	PROSES BELAJAR	POOR	POOR	11/06/2016	08/09/2016	89
3	JEMBER	17/02/1989	27	157	46	TAIWAN	2	RAWAT LANSIA	GOOD	GOOD	GOOD	GOOD	PROSES BELAJAR	GOOD	POOR	13/06/2016	08/09/2016	87
4	INDRAMAYU	25/01/1984	32	157	45	TAIWAN	3	RAWAT LANSIA	GOOD	GOOD	GOOD	GOOD	PROSES BELAJAR	GOOD	POOR	11/06/2016	09/09/2016	90
5	BLITAR	03/05/1986	30	156	55	MALAYSIA	2	ASUH BALITA	GOOD	GOOD	GOOD	GOOD	PROSES BELAJAR	POOR	POOR	07/06/2016	28/09/2016	113
6	MALANG	26/03/1990	26	151	47	MALAYSIA	2	ASUH BALITA	GOOD	GOOD	GOOD	GOOD	PROSES BELAJAR	POOR	POOR	19/05/2016	28/09/2016	132
7	SUBANG	07/01/1994	22	155	43	JAKARTA	3	ASUH ANAK	GOOD	GOOD	GOOD	GOOD	PROSES BELAJAR	POOR	POOR	03/05/2016	04/10/2016	154
8	INDRAMAYU	16/05/1983	33	145	48	ARAB SAUDI	3	ASUH ANAK	GOOD	GOOD	GOOD	GOOD	PROSES BELAJAR	POOR	POOR	02/08/2016	11/10/2016	70
9	SUBANG	09/06/1976	40	150	49	HONGKONG	1	RAWAT LANSIA	GOOD	GOOD	GOOD	GOOD	BERKOMUNIKASI	POOR	POOR	08/06/2016	24/10/2016	138
10	BLITAR	03/09/1994	22	153	57	BLITAR	2	RAWAT LANSIA	GOOD	GOOD	GOOD	GOOD	PROSES BELAJAR	POOR	POOR	20/07/2016	12/10/2016	84
11	SUBANG	27/12/1982	34	155	50	TAIWAN	1	ASUH ANAK	GOOD	GOOD	GOOD	GOOD	PROSES BELAJAR	GOOD	POOR	22/07/2016	15/11/2016	116
12	INDRAMAYU	10/01/1992	25	156	40	MALAYSIA	3	TATA LAKSANA RT	GOOD	GOOD	GOOD	GOOD	PROSES BELAJAR	POOR	POOR	02/08/2016	15/11/2016	105
13	SUBANG	14/06/1994	22	155	53	JAKARTA	2	ASUH BALITA	GOOD	GOOD	GOOD	GOOD	PROSES BELAJAR	POOR	POOR	09/09/2016	15/11/2016	67
14	INDRAMAYU	21/07/1984	32	155	75	HONGKONG	1	RAWAT LANSIA	GOOD	GOOD	GOOD	GOOD	PROSES BELAJAR	POOR	GOOD	26/08/2016	15/11/2016	81
15	CIREBON	18/06/1979	37	155	56	HONGKONG	2	ASUH ANAK	GOOD	GOOD	GOOD	GOOD	BERKOMUNIKASI	FAIR	GOOD	29/07/2016	29/11/2016	123
16	SUBANG	31/12/1982	34	148	51	ARAB SAUDI	3	ASUH BALITA	GOOD	GOOD	GOOD	GOOD	PROSES BELAJAR	POOR	POOR	02/08/2016	29/11/2016	119
17	SUBANG	05/05/1981	35	156	69	TAIWAN	2	RAWAT LANSIA	GOOD	GOOD	GOOD	GOOD	PROSES BELAJAR	FAIR	POOR	02/08/2016	07/12/2016	127
18	SUBANG	18/03/1983	34	153	53	TAIWAN	3	RAWAT LANSIA	GOOD	GOOD	GOOD	GOOD	PROSES BELAJAR	GOOD	POOR	09/09/2016	07/12/2016	89
19	SUBANG	06/07/1994	22	163	52	JAKARTA	2	ASUH ANAK	GOOD	GOOD	GOOD	GOOD	PROSES BELAJAR	POOR	POOR	24/08/2016	16/01/2017	145
20	SUBANG	29/06/1977	39	146	47	MALAYSIA	2	ASUH BALITA	GOOD	GOOD	GOOD	GOOD	PROSES BELAJAR	POOR	POOR	20/05/2016	24/01/2017	249

Gambar 3. Contoh Tabel Dataset Awal dari 353 data

b. Data Cleansing dan Integration
Setelah dilakukan proses Data Cleansing dan Integration, diperoleh data untuk dataset awal adalah sebanyak 353 (tiga ratus lima puluh tiga) record data.

c. Tahapan Transformasi Data
Tahapan ini dilakukan untuk memudahkan pengolahan data dengan penerapan algoritma data mining yang akan dilakukan. Tahapan ini berupa perubahan nilai data nominal menjadi numerik.

KABUPATEN/KOTA ASAL	NUMERIK	JUMLAH
BIMA	1	7
BLITAR	2	2
BOJONEGORO	3	1
CIANJUR	4	1
CIREBON	5	53
DEPOK	6	2
DOMPU	7	1
INDRAMAYU	8	66
JAKARTA PUSAT	9	2
JAKARTA UTARA	10	3
JEMBER	11	1
KARAWANG	12	8
LAMPUNG TENGAH	13	13
LAMPUNG TIMUR	14	2
LAMPUNG UTARA	15	1
MALANG	16	3
METRO	17	1
PASURUAN	18	1
PROBOLINGGO	19	1
SUBANG	20	163
SUKABUMI	21	3
SUMBAWA	22	15
TANGGAMUS	23	1
TULUNG AGUNG	24	1
SUMEDANG	25	1
		353

PENGALAMAN KERJA TERAKHIR	KODE	JUMLAH
ARAB SAUDI	1	40
BANDUNG	2	5
BEKASI	3	1
BLITAR	4	1
BOGOR	5	1
BRUNAI DARUSSALAM	6	4
HONGKONG	7	36
JAKARTA	8	106
LAMPUNG	9	4
MALAYSIA	10	28
SUMBAWA	11	1
SINGAPORE	12	38
SURABAYA	13	4
TAIWAN	14	82
MALANG	15	2
		353

PENGALAMAN KERJA (JENIS PEKERJAAN)	KODE	JUMLAH
ASUH BALITA	1	127
ASUH ANAK	2	74
RAWAT LANSIA	3	113
TATA LAKSANA RT	4	39
		353

PERSONALITY	KODE	JUMLAH		FACIAL EXPRESSION	KODE	JUMLAH		TIDINES	KODE	JUMLAH		HOUSEMAID EXPERIENCE	KODE	JUMLAH
GOOD	1	283		GOOD	1	283		GOOD	1	283		GOOD	1	259
FAIR	2	70		FAIR	2	70		FAIR	2	70		FAIR	2	87
POOR	3	0		POOR	3	0		POOR	3	0		POOR	3	7
		353				353				353				353

KEMAMPUAN BAHASA KANTONIS	KODE	JUMLAH		KEMAMPUAN BAHASA MANDARIN SAAT MASUK PELATIHAN	KODE	JUMLAH		KEMAMPUAN BAHASA INGGRIS SAAT MASUK PELATIHAN	KODE	JUMLAH
PROSES BELAJAR	1	322		GOOD	1	42		GOOD	1	17
BERKOMUNIKASI	2	31		FAIR	2	46		FAIR	2	41
		353		POOR	3	265		POOR	3	295
						353				353

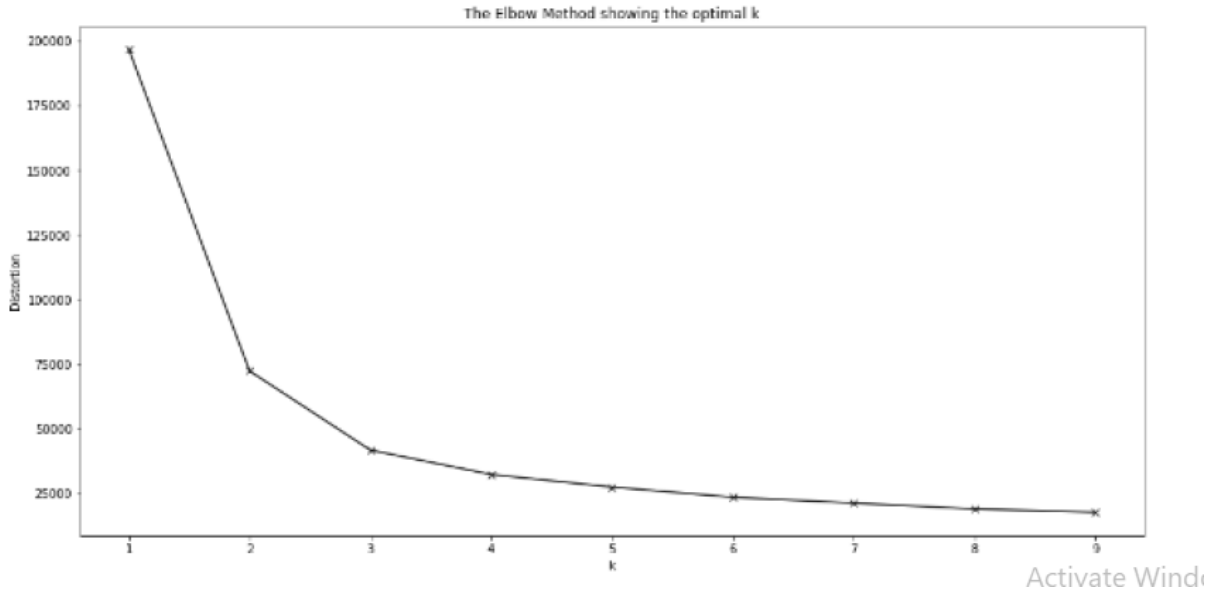
Gambar 4. Tabel Transformasi Dataset

d. Target Data
   - Pada tahapan ini, dataset awal yang terdiri dari 20 (dua puluh) atribut, dikurangi menjadi 16 (enam belas) atribut dengan menghapus atribut tanggal lahir, tanggal masuk pelatihan, tanggal keberangkatan, dan Kode CPMI;
   - Atribut tanggal lahir, tanggal masuk pelatihan, dan tanggal keberangkatan dihapus karena tujuan awal penggunaan atribut-atribut tersebut adalah untuk menghitung atribut umur dan lamanya hari CPMI dari tanggal masuk pelatihan sampai dengan keberangkatan;
   - Atribut Kode CPMI juga dihapus, cukup diwakilkan dengan atribut No yang unik dan identik untuk masing-masing Kode CPMI;
   - Target Data yang akan diolah terdiri dari atribut:
     1. No (simbol kolom “A”);
     2. Kabupaten/Kota Asal CPMI (simbol kolom “B”);
     3. Umur (simbol kolom “C”);
     4. Tinggi Badan (simbol kolom “D”);
     5. Berat Badan (simbol kolom “E”);
     6. Negara Pengalaman Kerja Terakhir (simbol kolom “F”);
     7. Lama Pengalaman Kerja Terakhir (simbol kolom “G”);
     8. Jenis Pekerjaan Pengalaman Kerja Terakhir (simbol kolom “H”);
     9. Personality (simbol kolom “I”);
     10. Facial Expression (simbol kolom “J”);
     11. Tidines (simbol kolom “K”);
     12. Housemaid Experience (simbol kolom “L”);
     13. Kemampuan Bahasa Kantonis (simbol kolom “M”);
     14. Kemampuan Bahasa Mandarin (simbol kolom “N”);
     15. Kemampuan Bahasa Inggris (simbol kolom “O”);
     16. Total Hari CPMI Dari Masuk Pelatihan Sampai Dengan Berangkat (simbol kolom “P”).

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	8	28	150	46	12	2	2	1	1	1	1	1	3	1	89
2	18	22	153	63	13	2	2	1	1	1	1	1	3	3	89
3	11	27	157	46	14	2	3	1	1	1	1	1	1	3	87
4	8	32	157	45	14	3	3	1	1	1	1	1	1	3	90
5	2	30	156	55	10	2	1	1	1	1	1	1	3	3	113
6	16	26	151	47	10	2	1	1	1	1	1	1	3	3	132
7	20	22	155	43	8	3	2	1	1	1	1	1	3	3	154
8	8	33	145	48	1	3	2	1	1	1	1	1	3	3	70
9	20	40	150	49	7	1	3	1	1	1	1	2	3	3	138
10	2	22	153	57	4	2	3	1	1	1	1	1	3	3	84
11	20	34	155	50	14	1	2	1	1	1	1	1	1	3	116
12	8	25	156	40	10	3	4	1	1	1	1	1	3	3	105
13	20	22	155	53	8	2	1	1	1	1	1	1	3	3	67
14	8	32	155	75	7	1	3	1	1	1	1	1	3	1	81
15	5	37	155	56	7	2	2	1	1	1	1	2	2	1	123
16	20	34	148	51	1	3	1	1	1	1	1	1	3	3	119
17	20	35	156	69	14	2	3	1	1	1	1	1	2	3	127
18	20	34	153	53	14	3	3	1	1	1	1	1	1	3	89
19	20	22	163	52	8	2	2	1	1	1	1	1	3	3	145
20	20	39	146	47	10	2	1	1	1	1	1	1	3	3	249
21	20	33	155	58	14	3	2	1	1	1	1	1	1	3	123
22	8	24	150	42	8	2	2	1	1	1	1	1	3	3	182
23	8	30	155	48	8	5	1	1	1	1	1	1	3	3	193
24	20	34	158	68	14	6	3	1	1	1	1	1	1	3	101
25	20	33	152	53	14	2	3	1	1	1	1	1	1	3	108
26	20	24	150	48	10	2	2	1	1	1	1	1	3	2	79

Gambar 5. Contoh Tabel Target Data dari 353 data

e. Metode Elbow, dari hasil algoritma Elbow dan hasil grafik Elbow diperoleh jumlah cluster yang optimal = 3 cluster.



Gambar 6. Contoh Tabel Target Data dari 353 data

f. Hasil impor dataset ke dalam bentuk Dataframe Python

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.cluster import KMeans
from sklearn.preprocessing import MinMaxScaler

datakmeans=pd.read_excel(r'C:\Users\User\Documents\TA UMB\KONSEP TA\KONSEP TA\KONSEP SERIUS\DATASET\DATASET KMEANS\DATAS
datakmeans1=pd.DataFrame(datakmeans, columns=['B','C','D','E','F','G','H','I','J','K','L','M','N','O','P'])
```

```
print (datakmeans1)
```

```
      B          C    D   E   F  G  H  I  J  K  L  M  N  O    P
0     8  28.098630  150  46  12  2  2  1  1  1  1  1  3  1   89
1    18  22.016438  153  63  13  2  2  1  1  1  1  1  3  3   89
2    11  27.336986  157  46  14  2  3  1  1  1  1  1  1  3   87
3     8  32.400000  157  45  14  3  3  1  1  1  1  1  1  3   90
4     2  30.117808  156  55  10  2  1  1  1  1  1  1  3  3  113
..   ..        ...  ...  ..  .. .. .. .. .. .. .. .. .. ..  ...
348   5  34.057534  158  66  14  2  3  1  1  1  1  1  2  3   50
349   8  36.572603  150  50   1  5  1  2  2  2  2  1  3  3   49
350   8  30.402740  154  60  12  1  3  1  1  1  1  1  3  3  156
351   3  44.791781  150  58   7  3  3  1  1  1  1  2  3  3  105
352  24  28.515068  155  67  13  2  1  1  1  1  1  1  3  3   78

[353 rows x 15 columns]
```

g. Penerapan algoritma model K-Means dengan Python, jumlah Cluster = 3

```
kmeans = KMeans(n_clusters = 3, random_state=123)
kmeans.fit(datakmeans1_scaled)
```

```
KMeans(algorithm='auto', copy_x=True, init='k-means++', max_iter=300,
       n_clusters=3, n_init=10, n_jobs=None, precompute_distances='auto',
       random_state=123, tol=0.0001, verbose=0)
```

h. Hasil cluster model K-Means Python

```
print (kmeans.labels_)
```

```
[0 0 2 2 0 0 0 0 0 0 2 0 0 0 0 0 2 2 0 0 2 0 0 2 2 0 0 0 0 2 0 0 2 0 0 2 0
 0 0 0 0 0 0 2 2 0 0 0 0 0 0 0 0 0 2 0 2 0 0 0 0 0 0 0 0 0 0 2 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 2 0 0 0 0 0 0
 2 0 0 2 0 0 0 0 0 0 0 0 0 0 2 0 2 0 2 0 0 0 0 2 2 0 0 2 0 0 0 0 0 0 0 0 0 2 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 2 2 2 0 2 2 0 0 0 1 0 0 0 0 0 2 2 0
 2 2 1 0 2 0 0 2 0 0 1 0 0 1 0 0 0 0 1 0 1 0 1 0 0 1 0 0 0 0 2 0 0 0 1 0 0
 1 1 1 2 0 1 1 1 1 0 0 1 0 1 0 1 2 0 1 1 1 1 1 1 0 1 1 2 1 0 1 0 1 1 2 1 1
 0 1 1 0 1 1 2 1 1 1 1 1 1 1 0 2 2 1 2 0 1 1 0 2 1 0 0 0 1 0 1 0 2 0 1 0 1
 1 1 0 2 2 1 0 0 0 1 0 0 1 0 0 0 0 1 0 2 1 1 1 2 2 0 0 0 0 0 0 0 0 2 1 1 0
 0 2 0 0 0 1 1 1 0 0 0 0 0 0 1 2 2 1 0 0 0]
```

i. Menambahkan atribut baru kolom Cluster pada dataset K-Means

```
datakmeans1["cluster"] = kmeans.labels_
print (datakmeans1)
```

```
      B          C    D   E   F  G  H  I  J  K  L  M  N  O    P  cluster
0     8  28.098630  150  46  12  2  2  1  1  1  1  1  3  1   89        0
1    18  22.016438  153  63  13  2  2  1  1  1  1  1  3  3   89        0
2    11  27.336986  157  46  14  2  3  1  1  1  1  1  1  3   87        2
3     8  32.400000  157  45  14  3  3  1  1  1  1  1  1  3   90        2
4     2  30.117808  156  55  10  2  1  1  1  1  1  1  3  3  113        0
..   ..        ...  ...  ..  .. .. .. .. .. .. .. .. .. ..  ...      ...
348   5  34.057534  158  66  14  2  3  1  1  1  1  1  2  3   50        2
349   8  36.572603  150  50   1  5  1  2  2  2  2  1  3  3   49        1
350   8  30.402740  154  60  12  1  3  1  1  1  1  1  3  3  156        0
351   3  44.791781  150  58   7  3  3  1  1  1  1  2  3  3  105        0
352  24  28.515068  155  67  13  2  1  1  1  1  1  1  3  3   78        0

[353 rows x 16 columns]
```

j. Hasil keseluruhan data hasil cluster

```
datakmeans1["cluster"] = kmeans.labels_
print (datakmeans1)
```

```
      B          C    D   E   F  G  H  I  J  K  L  M  N  O    P  cluster
0     8  28.098630  150  46  12  2  2  1  1  1  1  1  3  1   89        0
1    18  22.016438  153  63  13  2  2  1  1  1  1  1  3  3   89        0
2    11  27.336986  157  46  14  2  3  1  1  1  1  1  1  3   87        2
3     8  32.400000  157  45  14  3  3  1  1  1  1  1  1  3   90        2
4     2  30.117808  156  55  10  2  1  1  1  1  1  1  3  3  113        0
..   ..        ...  ...  ..  .. .. .. .. .. .. .. .. .. ..  ...      ...
348   5  34.057534  158  66  14  2  3  1  1  1  1  1  2  3   50        2
349   8  36.572603  150  50   1  5  1  2  2  2  2  1  3  3   49        1
350   8  30.402740  154  60  12  1  3  1  1  1  1  1  3  3  156        0
351   3  44.791781  150  58   7  3  3  1  1  1  1  2  3  3  105        0
352  24  28.515068  155  67  13  2  1  1  1  1  1  1  3  3   78        0

[353 rows x 16 columns]
```

k. Presentasi Hasil K-Means

1. Hasil Clustering Atribut Kondisi Fisik CPMI (Umur, Tinggi Badan, dan Berat Badan) dan Lama Waktu:

CLUSTER	JUMLAH	UMUR			TINGGI BADAN			BERAT BADAN			LAMA WAKTU		
		MEAN	MAX	MIN	MEAN	MAX	MIN	MEAN	MAX	MIN	MEAN	MAX	MIN
0	225	28	45	20	154	170	142	54	78	35	115	266	30
1	70	32	46	21	154	167	147	57	75	43	82	182	41
2	58	33	44	26	154	169	145	59	78	43	99	197	36
TOTAL	**353**												

2. Hasil Clustering Asal Kota CPMI:

NO	NAMA KOTA ASAL	CLUSTER 0	CLUSTER 1	CLUSTER 2	JUMLAH
1	BIMA	7	0	0	7
2	BLITAR	2	0	0	2
3	BOJONEGORO	1	0	0	1
4	CIANJUR	1	0	0	1
5	CIREBON	42	3	8	53
6	DEPOK	2	0	0	2
7	DOMPU	1	0	0	1
8	INDRAMAYU	33	19	14	66
9	JAKARTA PUSAT	2	0	0	2
10	JAKARTA UTARA	3	0	0	3
11	JEMBER	0	0	1	1
12	KARAWANG	4	4	0	8
13	LAMPUNG TENGAH	11	0	2	13
14	LAMPUNG TIMUR	1	0	1	2
15	LAMPUNG UTARA	1	0	0	1
16	MALANG	3	0	0	3
17	METRO	1	0	0	1
18	PASURUAN	1	0	0	1
19	PROBOLINGGO	0	0	1	1
20	SUBANG	93	43	27	163
21	SUKABUMI	3	0	0	3
22	SUMBAWA	11	1	3	15
23	TANGGAMUS	0	0	1	1
24	TULUNG AGUNG	1	0	0	1
25	SUMEDANG	1	0	0	1
TOTAL		**225**	**70**	**58**	**353**

3. Hasil Clustering Kota/Negara Pengalaman Kerja

NO	NEGARA/KOTA PENGALAMAN KERJA	CLUSTER 0	CLUSTER 1	CLUSTER 2	JUMLAH
1	ARAB SAUDI	31	9	0	40
2	BANDUNG	5	0	0	5
3	BEKASI	1	0	0	1
4	BLITAR	1	0	0	1
5	BOGOR	0	1	0	1
6	BRUNAI DARUSSALAM	3	1	0	4
7	HONGKONG	21	12	3	36
8	JAKARTA	93	13	0	106
9	LAMPUNG	4	0	0	4
10	MALAYSIA	24	3	1	28
11	SUMBAWA	1	0	0	1
12	SINGAPORE	29	8	1	38
13	SURABAYA	4	0	0	4
14	TAIWAN	6	23	53	82
15	MALANG	2	0	0	2
TOTAL		**225**	**70**	**58**	**353**

4. Hasil Clustering Jenis dan Lama Pengalaman Kerja

CLUSTER	JUMLAH	JENIS PENGALAMAN PEKERJAAN				HOUSE MAID EXPERIENCE			LAMA BEKERJA		
		ASUH BALITA	ASUH ANAK	RAWAT LANSIA	TATA LAKSANA RT	GOOD	FAIR	POOR	MEAN	MAX	MIN
0	225	103	56	35	31	202	19	4	3	13	1
1	70	19	14	30	7	2	66	2	3	9	1
2	58	5	4	48	1	55	2	1	3	9	1
TOTAL	**353**	**127**	**74**	**113**	**39**						

5. Hasil Clustering Penilaian Pribadi

CLUSTER	JUMLAH	PERSONALITY			FACIAL EXPRESSION			TIDINES		
		GOOD	FAIR	POOR	GOOD	FAIR	POOR	GOOD	FAIR	POOR
0	225	225	0	0	225	0	0	225	0	0
1	70	0	70	0	0	70	0	0	70	0
2	58	58	0	0	58	0	0	58	0	0
TOTAL	**353**	**283**	**70**	**0**	**283**	**70**	**0**	**283**	**70**	**0**

6. Hasil Clustering Kemampuan Bahasa

CLUSTER	JUMLAH	KANTONIS		MANDARIN			INGGRIS		
		BELAJAR	KOMUNIKASI	GOOD	FAIR	POOR	GOOD	FAIR	POOR
0	225	207	18	0	3	222	16	18	191
1	70	60	10	2	25	43	0	17	53
2	58	55	3	40	18	0	1	6	51
TOTAL	**353**	**322**	**31**	**42**	**46**	**265**	**17**	**41**	**295**

B. Hasil Eksperimen Klasifikasi KNN

1. Import Dataset ke Dataframe Python

```
import pandas as pd
import numpy as np
knncpmi1=pd.read_excel(r'C:\Users\User\Documents\TA UMB\KONSEP TA\KONSEP TA\KONSEP SERIUS\DATASET\DATASETKNN\DATASETUTAM
knncpmi2=pd.DataFrame(knncpmi1, columns=['B','C','D','E','F','G','H','I','J','K','L','M','N','O','P','CLUSTER','NAMA_CLU
print (knncpmi2)
```

```
     B         C    D   E   F  G  H  I  J  K  L  M  N  O    P  CLUSTER  \
0    8  28098630  150  46  12  2  2  1  1  1  1  1  3  1   89        0
1   18  22016438  153  63  13  2  2  1  1  1  1  1  3  3   89        0
2   11  27336986  157  46  14  2  3  1  1  1  1  1  1  3   87        2
3    8  32400000  157  45  14  3  3  1  1  1  1  1  1  3   90        2
4    2  30117808  156  55  10  2  1  1  1  1  1  1  3  3  113        0
..  ..       ...  ...  ..  .. .. .. .. .. .. .. .. .. .. ...      ...
348  5  34057534  158  66  14  2  3  1  1  1  1  1  2  3   50        2
349  8  36572603  150  50   1  5  1  2  2  2  2  1  3  3   49        1
350  8  30402740  154  60  12  1  3  1  1  1  1  1  3  3  156        0
351  3  44791781  150  58   7  3  3  1  1  1  1  2  3  3  105        0
352 24  28515068  155  67  13  2  1  1  1  1  1  1  3  3   78        0

    NAMA_CLUSTER
0      CLUSTER 1
1      CLUSTER 1
2      CLUSTER 3
3      CLUSTER 3
4      CLUSTER 1
..           ...
348    CLUSTER 3
349    CLUSTER 2
350    CLUSTER 1
351    CLUSTER 1
352    CLUSTER 1

[353 rows x 17 columns]
```

2. Menentukan Variabel Independent (explanatory variable)

```
x=knncpmi2[["B","C","D","E","F","G","H","I","J","K","L","M","N","O"]]
```

```
print (x)
```

```
      B         C    D   E   F  G  H  I  J  K  L  M  N  O
0     8  28098630  150  46  12  2  2  1  1  1  1  1  3  1
1    18  22016438  153  63  13  2  2  1  1  1  1  1  3  3
2    11  27336986  157  46  14  2  3  1  1  1  1  1  1  3
3     8  32400000  157  45  14  3  3  1  1  1  1  1  1  3
4     2  30117808  156  55  10  2  1  1  1  1  1  3  3
..   ..       ...  ...  ..  .. .. .. .. .. .. .. .. .. ..
348   5  34057534  158  66  14  2  3  1  1  1  1  1  2  3
349   8  36572603  150  50   1  5  1  2  2  2  2  1  3  3
350   8  30402740  154  60  12  1  3  1  1  1  1  1  3  3
351   3  44791781  150  58   7  3  3  1  1  1  1  2  3  3
352  24  28515068  155  67  13  2  1  1  1  1  1  1  3  3

[353 rows x 14 columns]
```


3. Menentukan Variabel Dependent (Target Value)

```
y=knncpmi2["CLUSTER"]
```

```
print (y)
```

```
0      0
1      0
2      2
3      2
4      0
      ..
348    2
349    1
350    0
351    0
352    0
Name: CLUSTER, Length: 353, dtype: int64
```

4. Membagi Dataset menjadi Data Training dan Data Testing

```
from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test = train_test_split(x,y, test_size=0.20, random_state=0)
```

```
print ('x_train=', x_train.shape)
print ('x_test=', x_test.shape)
print ('y_train=', y_train.shape)
print ('y_test=', y_test.shape)
```

```
x_train= (282, 14)
x_test= (71, 14)
y_train= (282,)
y_test= (71,)
```

5. Mengaktifkan package untuk klasifikasi KNN (n_neighbors = 5)

```
from sklearn.neighbors import KNeighborsClassifier
knn = KNeighborsClassifier (n_neighbors = 5)
```

6. Input Data Training ke model KNN

```
knn.fit(x_train, y_train)
```

```
KNeighborsClassifier(algorithm='auto', leaf_size=30, metric='minkowski',
                     metric_params=None, n_jobs=None, n_neighbors=5, p=2,
                     weights='uniform')
```

7. Skor Akurasi

```
knn.score(x_test, y_test)
```

```
1.0
```

8. Prediksi atau Peramalan dan Probabilitasnya

```
y_pred = knn.predict (x_test)
y_pred
```

```
array([0, 0, 2, 0, 0, 0, 2, 0, 0, 1, 0, 1, 0, 1, 2, 1, 0, 1, 2, 1, 0, 1,
       0, 0, 1, 1, 0, 0, 0, 0, 1, 2, 0, 2, 0, 2, 0, 0, 1, 0, 2, 0, 1, 0,
       2, 1, 0, 0, 0, 0, 0, 0, 2, 2, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 1,
       1, 0, 0, 0, 0], dtype=int64)
```

```
knn.predict_proba (x_test)
```

```
array([[1. , 0. , 0. ],
       [1. , 0. , 0. ],
       [0. , 0. , 1. ],
       [1. , 0. , 0. ],
       [1. , 0. , 0. ],
       [1. , 0. , 0. ],
       [0. , 0. , 1. ],
       [1. , 0. , 0. ],
       [1. , 0. , 0. ],
       [0. , 1. , 0. ],
       [1. , 0. , 0. ],
       [0. , 1. , 0. ],
       [1. , 0. , 0. ],
       [0. , 1. , 0. ],
       [0.2, 0. , 0.8],
       [0. , 1. , 0. ],
       [1. , 0. , 0. ],
       [0. , 1. , 0. ],
       [0. , 0. , 1. ],
       [0. , 1. , 0. ],
       [0.8, 0. , 0.2],
```


9. Data hasil prediksi dengan data aktualnya

```
from sklearn.metrics import classification_report, confusion_matrix
```

```
print (confusion_matrix (y_test, y_pred))
```

```
[[43  0  0]
 [ 0 17  0]
 [ 0  0 11]]
```

```
print (classification_report(y_test, y_pred))
```

```
              precision    recall  f1-score   support

           0       1.00      1.00      1.00        43
           1       1.00      1.00      1.00        17
           2       1.00      1.00      1.00        11

    accuracy                           1.00        71
   macro avg       1.00      1.00      1.00        71
weighted avg       1.00      1.00      1.00        71
```

10. Prediksi Klasifikasi Data Baru

Melihat nilai akurasi klasifikasi dan prediksi dan melakukan uji coba Prediksi Klasifikasi Data CPMI Baru, misal:

1. Asal Kota: Subang (B = 20);
2. Umur (C) = 35 tahun;
3. Tinggi Badan (D) = 155 cm;
4. Berat Badan (E) = 45 kg;
5. Kota/Negara Pengalaman Kerja: Singapore (F = 12);
6. Pengalaman Kerja (G) = 2 tahun;
7. Jenis Pengalaman Kerja (H): Asuh Balita = 1;
8. Personality (I): Good = 1;
9. Facial Expression (J): Fair = 2;
10. Tidines (K): Fair = 2;
11. House Maid Experience (L): Fair = 2;
12. Kemampuan Bahasa Kantonis (M): Belajar = 1;
13. Kemampuan Bahasa Mandarin (N): Good = 1
14. Kemampuan Bahasa Inggris (O): Good = 1.

```
lookup_cluster_name = dict(zip(knncpmi2.CLUSTER.unique(), knncpmi2.NAMA_CLUSTER.unique()))
print(lookup_cluster_name)
```

```
{0: 'CLUSTER 1', 2: 'CLUSTER 3', 1: 'CLUSTER 2'}
```

```
cluster_prediction = knn.predict([[20, 35, 155, 45, 12, 2, 1, 1, 2, 2, 2, 1, 1, 1]])
lookup_cluster_name[cluster_prediction[0]]
```

```
'CLUSTER 1'
```

IV. KESIMPULAN

1. Fokus penelitian ini adalah mencari gambaran informasi terkait faktor apa saja (fisik, asal kota, pengalaman kerja sebelumnya, kepribadian, dan kemampuan bahasa) pendukung penilaian CPMI yang berpengaruh pada minat pemberi kerja di Hongkong dengan mempertimbangkan nilai atribut lamanya waktu CPMI dari masuk pelatihan (BLK) sampai dengan keberangkatannya.
2. Dari hasil pengolahan data mining penerapan algoritma clustering K-Means diperoleh data sebagai berikut:
   a. Urutan cluster dengan waktu rata-rata keberangkatan tercepat adalah:
      1) Cluster 1, dengan mean = 82 hari, maksimal = 182 hari, dan minimal = 30 hari.
         Jumlah data pada cluster ini adalah 70 orang dengan rata-rata: umur = 28 tahun, tinggi badan = 154 cm, dan berat badan = 54 kg. Mayoritas data: kota asal CPMI: Kota Subang = 43 orang, Negara Pengalaman Kerja = Taiwan (23 orang), Jenis Pengalaman Kerja = Rawat Lansia (30 orang), House Maid Experience = Fair (66 orang), Mean Lama Pengalaman Kerja = 3 tahun, Personality = Fair (70 orang/semua data), Facial Expression = Fair (70 orang/semua data), Tidines = Fair (70 orang/semua data), Bahasa Kantonis = Belajar (60 orang), Bahasa Mandarin = Poor (43 orang), dan Bahasa Inggris = Poor (53 orang);
      2) Cluster 2, dengan mean = 99 hari, maksimal = 197 hari, dan minimal = 36 hari.
         Jumlah data pada cluster ini adalah 58 orang dengan rata-rata: umur = 32 tahun, tinggi badan = 154 cm, dan berat badan = 57 kg. Mayoritas data: kota asal CPMI: Kota Subang = 27 orang, Negara Pengalaman Kerja = Taiwan (53) orang, Jenis Pengalaman Kerja = Rawat Lansia (48 orang), House Maid Experience = Good (55 orang), Mean Lama Pengalaman Kerja = 3 tahun, Personality = Good (58 orang/semua data), Facial Expression = Good (58 orang/semua data), Tidines = Good (58 orang/semua data), Bahasa Kantonis = Belajar (55 orang), Bahasa Mandarin = Good (40 orang), dan Bahasa Inggris = Poor (51 orang);
      3) Cluster 0, dengan mean = 115 hari, maksimal = 266 hari, dan minimal = 30 hari.
         Jumlah data pada cluster ini adalah 225 orang dengan rata-rata: umur = 33 tahun, tinggi badan 154 cm, dan berat badan = 59 kg . Mayoritas data: kota asal CPMI: Kota Subang = 93 orang, Kota Pengalaman Kerja = Jakarta (93 orang), Jenis Pengalaman Kerja = Asuh Balita (103 orang), House Maid Experience = Good (202 orang), Mean Lama Pengalaman Kerja = 3 tahun, Personality = Good (225 orang/semua data), Facial Expression = Good (225 orang/semua data), Tidines = Good (225 orang/semua data), Bahasa Kantonis = Belajar (207 orang), Bahasa Mandarin = Poor (222 orang), dan Bahasa Inggris = Poor (191 orang);
   b. Dari data di atas dapat diasumsikan bahwa:
      1) faktor fisik (umur, tinggi badan, dan berat badan) tidak menunjukkan perbedaan yang jauh antar semua cluster, sehingga dapat disimpulkan faktor fisik tidak berpengaruh signifikan terhadap minat pemberi kerja;
      2) faktor asal kota CPMI di semua cluster didominasi oleh kota Subang dengan jumlah keseluruhan sebanyak = 163 orang, hal ini dimungkinkan karena populasi CPMI asal kota subang adalah 46,18% dari keseluruhan populasi data penelitian (353 data);
      3) faktor kota atau negara tempat pengalaman kerja sebelumnya, pada cluster 1 dan cluster 2 didominasi oleh negara Taiwan dengan jumlah = 76, persentase total = 92% dari populasi Taiwan keseluruhan (82 orang). Memberikan gambaran informasi bahwa negara tempat pengalaman kerja sebelumnya yang cenderung diminati pemberi kerja adalah negara Taiwan;
      4) faktor jenis pengalaman pekerjaan sebelumnya pada cluster 1 dan cluster 2 didominasi jenis pekerjaan rawat lansia. Memberikan gambaran informasi bahwa jenis pekerjaan sebelumnya yang cenderung diminati pemberi kerja adalah rawat lansia;
      5) faktor-faktor lain (kepribadian dan kemampuan bahasa) tidak menunjukkan pengaruh yang signifikan terhadap lamanya waktu, karena pada cluster 1 nilai keseluruhan kepribadian adalah: Fair dan bahasa didominasi proses belajar dan nilai: Poor. Sehingga dapat diperoleh gambaran informasi bahwa pemberi kerja tidak terlalu mempertimbangkan kemampuan bahasa karena mayoritas seluruh CPMI dalam proses belajar bahasa Kantonis.
   c. Berdasarkan informasi tersebut di atas, penulis memperoleh kesimpulan bahwa faktor yang sangat berpengaruh terhadap minat pemberi kerja terhadap CPMI (dilihat dari rata-rata lamanya waktu CPMI berangkat) adalah faktor negara dan jenis pengalaman pekerjaan CPMI sebelumnya.
3. Kesimpulan Hasil Eksperimen Klasifikasi KNN
   a. Skor akurasi dari penerapan algoritma klasifikasi KNN pada dataset cluster data CPMI adalah = 1.0 (100%), hal ini berarti bahwa penerapan model KNN tersebut telah berhasil mengklasifikasikan dengan tepat cluster data test (20%) terhadap data train (80%) yang telah ditentukan (353 data) dalam proses algoritma KNN;
   b. Hasil evaluasi klasifikasi KNN dengan penerapan metode confusion matrix menyatakan bahwa prediksi klasifikasi telah tepat terhadap semua data masing-masing cluster;
   c. Nilai Precision cluster 0 = 1.0, cluster 1 = 1.0 , dan cluster 2 = 1.0, hal ini menyatakan bahwa rasio prediksi benar positif dibandingkan dengan keseluruhan hasil yang diprediksi positif 100% tepat:

d. Nilai Recall cluster 0 = 1.0, cluster 1 = 1.0, dan cluster 2 = 1.0, hal ini menyatakan bahwa rasio prediksi benar positif dibandingkan dengan keseluruhan data yang benar positif 100% tepat;
e. F1 Score (perbandingan rata-rata presisi dan recall yang dibobotkan), cluster 0 = 1.0, cluster 1 = 1.0, dan cluster 2 = 1.0;
f. Berdasarkan nilai akurasi klasifikasi dan prediksi di atas, dapat dilakukan proses prediksi data CPMI baru dengan asumsi bahwa prediksi klasifikasi cluster terhadap data CPMI baru tersebut tepat pada masing-masing clusternya.
g. Sesuai dengan hasil tersebut, dilakukan ujicoba prediksi klasifikasi cluster untuk mengetahui rata-rata, maksimal, dan minimal lama waktu dari data CPMI baru berdasarkan hasil clusternya;
h. Ujicoba prediksi klasifikasi dilakukan dengan menginput nilai atribut asal kota, umur, tinggi badan, berat badan, pengalaman kerja sebelumnya, kepribadian, dan kemampuan bahasanya dan diperoleh hasil data tersebut masuk ke dalam cluster mana.

V. DAFTAR PUSTAKA

[1] Winda Aprianti & Jaka Permadi (2018). *K-Means Clustering* Untuk Data Kecelakaan Lalu Lintas Jalan Raya Di Kecamatan Pelaihari. Jurnal: Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK), 5(5), 613-620.
[2] Siti Monalisa (2018). Klasterisasi *Customer Lifetime Value* Dengan Model LRFM Menggunakan Algoritma K-Means. Jurnal: Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK), 5(2), 247-252.
[3] Indivar Shaik, Swapna Suhasini Nittela, Trayabak Hiwarkar & Srinivas Nalla (2019). K-means Clustering Algorithm Based on E-Commerce Big Data. Jurnal: International Journal of Innovative Technology and Exploring Engineering (IJITEE), 8(11), September 2019.
[4] Xianglong Luo, Danyang Li, Yu Yang, and Shengrui Zhang (2019). Spatiotemporal Traffic Flow Prediction with KNN and LSTM. Jurnal: Journal of Advanced Transportation, Volume 2019, Article ID 4145353, 10 pages.
[5] Zhuo Chen, Lan Jiang Zhou, Xuan Da Li, Jia Nan Zhang, and Wen Jie Huo (2019). The Lao Text Classification Method Based on KNN. Jurnal: 3rd International Conference on Mechatronics and Intelligent Robotics (ICMIR-2019).
[6] Ida Farida dan Spits Warnars Harco Leslie Hendric (2019). Prediksi Pola Kelulusan Mahasiswa Menggunakan Teknik Data Mining Classification Emerging Pattern. Jurnal: Jurnal Petir Vol. 12, No. 1, Maret 2019, P-ISSN 1978-9262, E-ISSN 2655-5018.
[7] Pangestu, Daniel Harry (2019) *Penerapan Data Mining Untuk Klasifikasi Penjurusan Sekolah Menengah Atas Pada Sman 5 Kota Jambi Dengan Menggunakan Algoritma Naãïve Bayes Classifier.* Skripsi Thesis, STIKOM Dinamika Bangsa Jambi.
[8] Diky Firdaus (2017). Penggunaan Data Mining dalam Kegiatan Sistem Pembelajaran Berbantuan Komputer. Jurnal: Jurnal Format Volume 6 Nomor 2 Tahun 2017 :: ISSN : 2089 -5615.