

## PENERAPAN DATA MINING PADA SUKU BUNGA INVESTASI DEPOSITO DI INDONESIA MENGGUNAKAN METODE K-MEANS CLUSTERING UNTUK PENGELOMPOKAN PROFITABILITAS (Application of Data Mining on Deposit Investment Rate in Indonesia Using K-Means Clustering Method for Profitability Clustering)

Arief Wibowo<sup>1</sup>, Raden Sasongko<sup>2</sup>

<sup>12</sup>Universitas Budi Luhur

Email: [arief.wibowo@budiluhur.ac.id](mailto:arief.wibowo@budiluhur.ac.id)

*Abstract: Investment in Indonesia has several types, one of which is deposits through banks, namely individual or collective customers lending excess funds to banks and channeling them in the form of credit. This research is about clustering on several Banks, those are Regional, Private, Foreign, Persero and General Banks which have different interest rates for investments. The method used in this study is the Knowledge Discovery using Database (KDD) method, using RapidMiner tools. The algorithm used to perform clustering is the K-Means algorithm. The data that was used is data on investment credit interest rates from several banks in Indonesia obtained from BPS (Central Statistics Agency) Indonesia. This data is taken from 2009-2020. The clustering results obtained are 3 clusters, where cluster 1 is Regional Bank, which turns out to have stable and high loan interest rates, then cluster 2 is Foreign Bank, which also has the lowest and unstable interest rate. While cluster 3 is a private bank and a state-owned company, these two banks have similar interest rates and levels of stability. In addition, these two banks are in the middle for investment options.*

*Keywords; Clustering, , Investment Loan Interest Rate, K-Means, Knowledge Discovery using Database, RapidMiner.*

### PENDAHULUAN

Instrumen investasi terdiri dari beberapa jenis mulai dari pasar modal, deposito, obligasi dan lainnya. Investasi yang sehat merupakan investasi yang menghasilkan profit atau keuntungan yang bertumbuh seiring dengan berjalannya waktu, salah satu investasi yang tetap bertumbuh adalah investasi dalam bentuk deposito. Deposito merupakan suatu instrumen investasi yang meminjamkan dana berlebih kepada pihak perbankan dan kemudian dana tersebut akan disalurkan kepada nasabah atau pengusaha yang membutuhkan tambahan modal dalam bentuk kredit. Peran perbankan dalam hal ini sangat dibutuhkan untuk membantu pengalokasian agar alokasi dana dapat efisien. Dalam menjalankan investasi dalam bentuk deposito keuntungan yang diperoleh oleh investor dalam bentuk persentase dari dana yang dipinjamkan atau lebih dikenal dengan suku bunga.

Suku bunga investasi yang berasal dari deposito biasanya berbeda untuk setiap bank dan berdasarkan waktu, menurut Yunita (2016) nilai suku bunga tergantung pada masing-masing bank menentukan kebijakannya. Dengan demikian tingkat suku bunga mencerminkan tingkat kelangkaan atau kecukupan dana di masyarakat, sehingga tingkat suku bunga mempunyai kaitan yang cukup erat dengan berbagai indikator ekonomi lainnya. Di sisi internal tingkat suku bunga berkaitan dengan inflasi, permintaan dalam negeri dan nilai tukar rupiah. Dalam lingkup eksternal tingkat suku bunga sangat berperan terhadap arus modal masuk dan keluar.

Keuntungan dalam bentuk suku bunga (profitable) deposito dapat digolongkan menjadi beberapa bagian yaitu bunga rendah, bunga sedang dan tinggi, hal ini bisa dikaji melalui clustering atau pengelompokan suku bunga secara otomatis berdasarkan nilai pada masing-masing bank dan waktu. Analisis ini sangat berguna untuk mendapatkan tingkat

keuntungan atau profitabel yang tinggi dengan melihat kategori suku bunga sehingga para investor lebih tertarik untuk meminjamkan dananya kepada bank tertentu.

Nurdiawan (2018) melakukan penelitian yang berhubungan dengan data mining menggunakan metode Naive bayes untuk proses klasifikasi yang bertujuan untuk mengoptimalkan strategi pemasaran, dimana menggunakan aplikasi RapidMiner untuk mendapatkan hasil dengan cepat dan tepat sehingga penelitian ini didapatkan hasil optimasi sebesar 65% dan akurasi untuk metode klasifikasi sebesar 82,08% sehingga penelitian ini terbilang sangat akurat menggunakan aplikasi RapidMiner untuk mendapatkan hasil dengan akurasi yang tinggi sehingga pada penelitian ini aplikasi RapidMiner digunakan sebagai tool utama untuk melakukan pengolahan data.

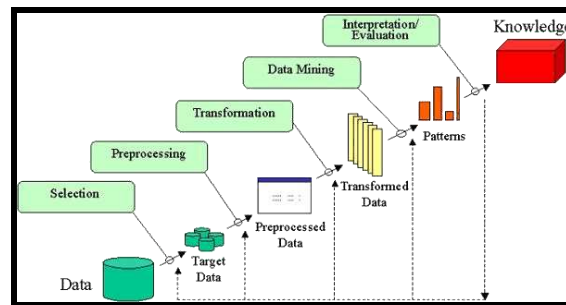
Dwiasuti (2020) Penelitian ini mengangkat tema terkait dengan pengaruh sektor perbankan terhadap proses pertumbuhan ekonomi indonesia, dimana studi kasus dilakukan di daerah Kalimantan Barat dengan melihat indikator kredit modal kerja, kredit investasi dan kredit konsumsi dengan metode regresi analisis berganda, dan hasil penelitian menghasilkan bahwa kredit investasi dan kredit konsumsi berpengaruh positif dan signifikan terhadap pertumbuhan ekonomi di kabupaten Kalimantan Barat, sehingga dari penelitian ini didapatkan kredit investasi bisa menjadikan sebuah acuan untuk pertumbuhan ekonomi yang menandakan perputaran dana di sektor perbankan berjalan lancar.

Penelitian yang dilakukan oleh Harianto (2016) implementasi dari K-means *clustering* dimana pada penelitian tersebut berfokus kepada membangun sebuah system berbasis web untuk mengolah data dan mengelompokkan berdasarkan cluster, penelitian ini menggunakan data yang bersumber dari data keuangan perusahaan pada sector pertambangan dan industry barang konsumsi dengan tujuan untuk mendapatkan rasio profitabilitas di setiap perusahaan. Pada penelitian ini yang membedakan adalah data yang digunakan tetapi metode yang digunakan sama yaitu K-means clustering untuk mengelompokkan berbagai bank yang memberikan kredit investasi sehingga mengetahui profitabilitas untuk setiap kelompok tersebut.

Maka dari itu penelitian ini akan melakukan analisa terhadap suku bunga deposito investasi dari berbagai jenis bank dengan menggunakan metode k-means clustering untuk mendapatkan kelompok perbankan berdasarkan nilai suku bunga dengan judul “Penerapan Data mining pada Suku Bunga Investasi Deposito di Indonesia Menggunakan Metode K-means Clustering untuk Pengelompokan Profitabilitas”. Yang bertujuan untuk meningkatkan minat daripada investor untuk melakukan kegiatan investasi pada kelompok perbankan yang memiliki nilai profitabilitas tinggi dan mendorong perbankan yang memiliki profitabilitas rendah untuk lebih bersaing meningkatkan nilai suku bunga sehingga perputaran ekonomi Indonesia lebih besar.

## METODE

Penelitian ini menggunakan *tools Rapid Mining* dalam mengaplikasikan metode KDD. Dalam melakukan *clustering* Metode penelitian yang digunakan adalah metode analisis deskriptif dengan pendekatan kuantitatif artinya penelitian yang dilakukan adalah menekankan analisisnya pada data-data numeric (angka), yang bertujuan untuk mendapatkan gambaran yang jelas mengenai suatu keadaan berdasarkan data yang diperoleh dengan cara menyajikan, mengumpulkan dan menganalisis data tersebut sehingga menjadi informasi baru yang dapat digunakan untuk menganalisa mengenai masalah yang sedang diteliti. Adapun tahapan yang dilakukan berdasarkan kepada Knowledge Discovery Data (KDD) adalah sebagai berikut;



Gambar 1 Metode Knowledge Discovery using Database

### 2.1 Data Selection

Data Selection ini bertujuan untuk membuat himpunan data target, pemilihan himpunan data, atau memfokuskan pada subset variabel atau sampel data, dimana penemuan (discovery) akan dilakukan. 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.

### 2.2 Data Preprocessing

Dalam proses preprocessing data, dilakukan pembersihan data dari *missing value* dan *outlier*. Selain itu proses lainnya adalah *cleaning* mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak (tipografi). 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.

### 2.3 Data Transformation

Untuk menentukan variabel target dan peubah dilakukan pada tahap ini. Pencarian fitur-fitur yang berguna untuk mempresentasikan data bergantung kepada goal yang ingin dicapai. Merupakan proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses data mining. Proses ini merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data.

### 2.4 Data Mining

Proses dari data mining adalah merupakan proses dari proses KDD itu sendiri. Proses dalam discovery knowledge bisa dilakukan dengan cara misalnya klasifikasi, regresi, clustering, dll. Pemilihan algoritma data mining untuk pencarian (searching), Proses Data mining yaitu 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.

### 2.5 Data Interpretation

Penerjemahan pola-pola yang dihasilkan dari data mining. 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 mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada sebelumnya.

## HASIL DAN PEMBAHASAN

### 3.1 Kajian Teori

#### 3.1.1 Bank

Kata bank berasal dari bahasa Italia “banca” yang berarti “bence” yaitu suatu bangku tempat duduk. Sebab, pada zaman pertengahan, pihak bankir Italia yang memberikan pinjaman-pinjaman melakukan usahanya tersebut dengan duduk dibangku-bangku di halaman pasar. Istilah bank dimaksudkan sebagai suatu jenis pranata finansial yang melaksanakan jasa-jasa keuangan yang cukup beraneka ragam, seperti pinjaman, memberi pinjaman, mengedarkan mata uang, mengadakan pengawasan terhadap mata uang, bertindak sebagai tempat penyimpanan untuk benda-benda berharga, membiayai usaha-usaha perusahaan.

Berdasarkan SK Menteri Keuangan RI No. 792 Tahun 1990, pengertian bank adalah Bank merupakan suatu badan yang kegiatannya dibidang keuangan melakukan penghimpunan dana dan penyaluran dana kepada masyarakat terutama guna membiayai investasi perusahaan.

Dari pengertian di atas dapat disimpulkan bahwa bank adalah lembaga keuangan yang menghimpun serta menyalurkan dari masyarakat kepada masyarakat melalui tabungan, pinjaman dalam bentuk kredit yang terus diputar.

#### 3.1.2 Investasi

Investasi adalah komitmen atas sejumlah dana atau sumber daya lainnya yang dilakukan pada saat ini, dengan tujuan memperoleh sejumlah keuntungan dimasa datang, Istilah investasi bisa berkaitan dengan berbagai macam aktivitas. Menginvestasikan dana pada sektor riil (tanah, emas, mesin atau bangunan) maupun aset finansial (deposito, saham atau obligasi), merupakan aktivitas yang umum dilakukan.

Menurut Menurut Suindyah (2017) kegiatan investasi yang dilakukan oleh masyarakat secara terus menerus akan meningkatkan kegiatan ekonomi dan kesempatan kerja, meningkatkan pendapatan nasional dan meningkatkan taraf kemakmuran masyarakat. Peranan ini bersumber dari tiga fungsi penting dari kegiatan investasi, yakni (1) investasi merupakan salah satu komponen dari pengeluaran agregat, sehingga kenaikan investasi akan meningkatkan permintaan agregat, pendapatan nasional serta kesempatan kerja; (2) penambahan barang modal sebagai akibat investasi akan menambah kapasitas produksi; (3) investasi selalu diikuti oleh perkembangan teknologi.

#### 3.1.3 Deposito

Deposito adalah simpanan yang penarikannya hanya dapat dilakukan pada waktu tertentu berdasarkan perjanjian nasabah penyimpan dengan bank (Undang-Undang No. 10 Tahun 1998). Deposito merupakan salah satu tempat bagi nasabah untuk melakukan investasi dalam bentuk surat-surat berharga. Pemilik deposito disebut dengan deposan. Bunga yang diberikan kepada para deposan merupakan bunga tertinggi jika dibandingkan dengan simpanan giro atau tabungan. Keuntungan bagi bank dengan menghimpun dana lewat deposito adalah uang yang tersimpan relatif lama, mengingat deposito memiliki jangka waktu yang relatif panjang dan frekuensi penarikan yang juga jarang. Jangka waktu deposito adalah 1 bulan, 3 bulan, 6 bulan, 12 bulan, atau 24 bulan. Semakin lama deposito, tingkat suku bunganya seharusnya akan semakin besar pula.

#### 3.1.4 Suku bunga

Tingkat suku bunga menjadi acuan seberapa besar imbalan yang akan diperoleh atas sejumlah dana yang dipercayakan oleh investor atau nasabah kepada bank yang bersangkutan. Sehingga, penting bagi pihak perbankan untuk menganalisis sejumlah faktor untuk menetapkan suku bunga yang ditargetkan sesuai dengan kondisi perbankan dan keadaan perekonomian. Menurut Yunita (2016) tingkat bunga ditentukan oleh permintaan dan penawaran akan uang (ditentukan dalam pasar uang). Perubahan tingkat suku bunga

selanjutnya akan mempengaruhi keinginan untuk mengadakan investasi, misalnya pada surat berharga, dimana harga dapat naik atau turun tergantung pada tingkat bunga (bila tingkat bunga naik maka surat berharga turun dan sebaliknya), sehingga ada kemungkinan pemegang surat berharga akan menderita capital loss atau gain.

### 3.1.5 K-Means Clustering

K-Means merupakan salah satu algoritma dalam data mining yang bisa digunakan untuk melakukan pengelompokan/clustering suatu data. Ada banyak pendekatan untuk membuat cluster, diantaranya adalah membuat aturan yang mendikte keanggotaan dalam group yang sama berdasarkan tingkat persamaan diantara anggota-anggotanya. Pendekatan lainnya adalah dengan membuat sekumpulan fungsi yang mengukur beberapa properti dari pengelompokan tersebut sebagai fungsi dari beberapa parameter dari sebuah clustering. Metode K-Means adalah metode yang termasuk dalam algoritma clustering berbasis jarak yang membagi data ke dalam sejumlah cluster dan algoritma ini hanya bekerja pada atribut numerik. Proses Pengelompokan data dengan metode K-Means adalah sebagai berikut (Windha, 2015):

1. Menentukan berapa jumlah cluster
2. Menempatkan data ke dalam cluster secara acak
3. Hitung centroid rata-rata dari data yang ada di masing-masing cluster
4. Menempatkan masing-masing data ke centroid/rata-rata terdekat
5. Kembali ke langkah 3, apabila masih ada data yang berpindah cluster atau apabila perubahan nilai centroid, ada yang di atas nilai threshold yang ditentukan atau apabila perubahan nilai pada object function yang digunakan di atas nilai threshold yang ditentukan.

### 3.1.6 RappidMiner

RapidMiner merupakan perangkat lunak yang bersifat terbuka (open source). RapidMiner adalah sebuah solusi untuk melakukan analisis terhadap data mining, text mining dan analisis prediksi. RapidMiner menggunakan berbagai teknik deskriptif dan prediksi dalam memberikan wawasan kepada pengguna sehingga dapat membuat keputusan yang paling baik. RapidMiner memiliki kurang lebih 500 operator data mining, termasuk operator untuk input, output, data preprocessing dan visualisasi (Aprilia, 2013).

Beberapa fitur dari Rapidminer adalah (Uska, dkk, 2020):

1. Open source
2. Multi-Platform Language (Python, Java, dll)
3. Memiliki interface yang memudahkan penggunaannya
4. Memiliki fitur automodel untuk aplikasi machine learning dan deep learning.

## 3.2 Pembahasan

### 3.2.1 Deskripsi Data

Data suku bunga kredit yang digunakan dalam penelitian ini diperoleh dari situs Badan Pusat Statistik Indonesia yang bebas untuk diakses. Data yang tersedia merupakan rekaman data suku bunga kredit per bulan untuk tiap tahun, yaitu dari tahun 2008 sampai dengan

2021. Dalam penelitian ini digunakan data rekaman 2008 sampai dengan 2020. Untuk gambaran data bisa dilihat pada gambar tabel berikut;

Kelompok Bank	Suku Bunga Kredit Rupiah Menurut Ke							
	2020							
	Januari	Februari	Maret	April	Mei	Juni	Juli	Agustus
Bank Persero - Modal Kerja	9,77	9,75	9,70	9,38	9,18	9,14	9,09	9,06
Bank Persero - Investasi	9,70	9,66	9,57	9,26	9,12	9,12	9,08	9,06
Bank Persero - Konsumsi	11,04	11,02	10,99	10,92	10,78	10,75	10,73	10,70
Bank Pemerintah Daerah - Modal Kerja	10,76	10,77	10,75	10,75	10,64	10,58	10,55	10,54
Bank Pemerintah Daerah - Investasi	11,02	10,98	10,90	10,87	10,73	10,66	10,62	10,57
Bank Pemerintah Daerah - Konsumsi	11,73	11,72	11,69	11,67	11,72	11,69	11,68	11,66
Bank Swasta Nasional - Modal Kerja	10,59	10,48	10,36	10,14	10,04	9,85	9,87	9,85
Bank Swasta Nasional - Investasi	10,02	9,98	9,81	9,57	9,43	9,42	9,28	9,20
Bank Swasta Nasional - Konsumsi	11,10	11,06	11,01	10,91	10,86	10,83	10,78	10,77
Bank Asing dan Bank	7,79	7,72	7,57	7,64	7,59	7,51	7,44	7,30

Gambar 2 Data suku bunga kredit tahun 2020 dari situs BPS

Dari gambar 2 diatas diperlihatkan bahwa ada 5 tipe bank dengan 3 jenis suku kredit bunga dengan nilai masing-masing dalam tiap bulan. Sehingga dalam satu file, data tersebut terdapat 13 kolom dan 16 baris. Kemudian data tiap tahun tersebut diunduh dan diperoleh 12 file dalam format xlsx (*Microsoft Excel*).

### 3.2.2 Data Pre-processing

Untuk pre-processing data, tidak ditemukan *missing values* pada data yang sudah diunduh. Selanjutnya dilakukan penggabungan 12 file suku bunga kredit menjadi satu file dan dilakukan transformasi untuk mendapatkan format data menjadi data *Time Series*. Sehingga dalam file tersebut terdapat 16 kolom dan 157 baris. Dengan kolom pertama merupakan *Date Time*. Sedangkan 15 kolom lainnya merupakan 5 tipe bank dengan 3 jenis suku bunga kredit. Untuk hasil transformasi data bisa dilihat pada tabel berikut:

Tabel 1 . Sample hasil preprosesing data

time	Bank Persero - Modal Kerja	Bank Persero - Investasi	Bank Persero - Konsumsi	Bank Pemerintah Daerah - Modal Kerja	Bank Pemerintah Daerah - Investasi	Bank Pemerintah Daerah - Konsumsi	Bank Swasta Nasional - Modal Kerja	Bank Swasta Nasional - Investasi	Bank Swasta Nasional - Konsumsi	Bank Asing dan Bank Campuran - Modal Kerja	Bank Asing dan Bank Campuran - Investasi	Bank Asing dan Bank Campuran - Konsumsi	Bank Umum - Modal Kerja	Bank Umum - Investasi	Bank Umum - Konsumsi
1/1/2008	12.52	11.47	12.64	13.56	12.25	13.40	13.31	13.02	13.45	10.39	10.59	28.46	12.76	12.29	13.62
2/1/2008	12.55	11.45	12.68	13.38	12.13	13.47	13.29	13.03	13.51	10.26	10.51	28.59	12.74	12.27	13.68
3/3/2008	12.65	11.49	12.67	13.71	12.37	13.49	13.36	13.06	13.52	10.26	10.47	28.62	12.82	12.32	13.68
4/2/2008	12.64	11.45	12.72	13.71	12.36	13.48	13.25	13.06	13.58	10.23	10.77	28.79	12.75	12.32	13.73
5/2/2008	12.61	11.45	12.74	13.72	12.38	13.49	13.20	13.02	13.62	10.21	10.75	28.97	12.72	12.30	13.76
6/1/2008	12.60	11.46	12.84	13.74	12.39	13.49	13.17	13.02	13.67	10.18	10.60	29.13	12.70	12.29	13.82
7/2/2008	12.54	11.46	12.91	13.56	12.18	13.30	13.14	12.98	13.65	10.28	10.68	28.40	12.65	12.26	13.73
8/1/2008	12.59	11.45	12.99	13.47	12.45	13.50	13.03	12.87	13.56	10.30	10.67	28.95	12.63	12.21	13.82
9/1/2008	12.48	11.44	13.03	13.47	12.52	13.50	13.00	12.83	13.61	10.51	10.78	29.05	12.58	12.19	13.85
10/2/2008	12.43	11.42	13.06	13.45	12.25	13.30	12.98	12.85	13.69	10.72	10.95	28.94	12.58	12.19	13.85
11/1/2008	12.36	11.40	13.08	13.66	12.44	13.49	12.92	12.77	13.68	10.79	11.13	28.87	12.55	12.14	13.89
12/2/2008	12.30	11.35	13.08	13.50	12.19	13.49	12.82	12.77	13.66	10.79	11.25	28.93	12.46	12.12	13.88

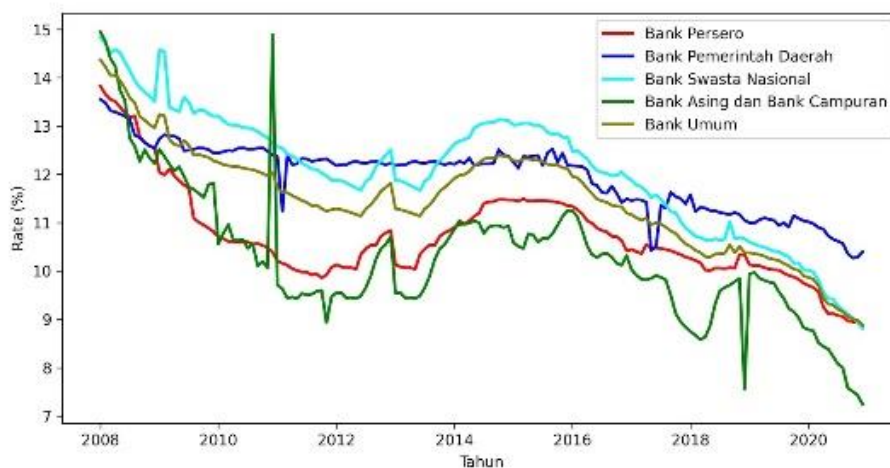
### 3.2.3 Data Transformation

Sesuai dengan tujuan dari penelitian ini adalah untuk melakukan *clustering* pada data suku bunga kredit investasi untuk 5 jenis bank. Maka dilakukan filtering untuk memilih 5 bank dengan suku bunga kredit investasi. Sehingga dataset akan memiliki 13 kolom dengan 5 baris. Untuk gambaran data dapat dilihat pada gambar tabel berikut:

Tabel 2 . Hasil filter data berdasarkan suku bunga kredit investasi

time	Bank Persero	Bank Pemerintah Daerah	Bank Swasta Nasional	Bank Asing dan Bank Campuran	Bank Umum
1/1/2008	13.83	13.55	14.83	14.95	14.37
2/1/2008	13.66	13.48	14.72	14.75	14.23
3/3/2008	13.55	13.32	14.52	14.40	14.05
4/2/2008	13.51	13.29	14.58	14.24	14.05
5/2/2008	13.39	13.25	14.52	13.74	13.94
6/1/2008	13.28	13.20	14.33	13.53	13.78
7/2/2008	13.18	13.13	14.14	12.74	13.58
8/1/2008	13.21	12.81	13.94	12.58	13.48
9/1/2008	12.78	12.76	13.80	12.27	13.20
10/2/2008	12.66	12.66	13.70	12.51	13.12
11/1/2008	12.58	12.59	13.62	12.34	13.03
12/2/2008	12.56	12.54	13.51	12.22	12.96
1/1/2009	12.04	12.74	14.58	12.52	13.24
2/1/2009	11.99	12.82	14.54	12.37	13.21
3/4/2009	12.11	12.79	13.38	12.21	12.72
4/3/2009	11.97	12.79	13.34	12.07	12.62
5/3/2009	11.87	12.74	13.30	12.16	12.59
6/2/2009	11.78	12.49	13.60	11.96	12.70
7/2/2009	11.71	12.50	13.49	11.74	12.60
8/1/2009	11.09	12.52	13.26	11.69	12.40
9/1/2009	11.01	12.55	13.33	11.61	12.41
10/2/2009	10.95	12.54	13.33	11.52	12.38

Dari gambar 4 diperlihatkan bahwa data suku bunga kredit untuk 5 tipe bank yaitu; Bank Pemerintah Daerah, Persero, Swasta Nasional, Asing dan Umum. Kemudian dari dataset tersebut bisa dilihat perbandingan sementara tarif suku bunga kredit investasi. Bisa dilihat pada grafik berikut:



Gambar 3 Perbandingan 5 tipe bank berdasarkan tarif suku bunga kredit investasi

Dari grafik tersebut bisa dilihat bahwa bank Swasta Nasional memiliki tarif suku bunga kredit tertinggi pada tahun 2008-2011. Sedangkan yang terendah pada tahun tersebut adalah bank Persero dan Umum. Selain itu, kedua bank tersebut memiliki tarif suku kredit bunga investasi terendah dari 3 tipe bank lainnya dari tahun 2008 sampai

dengan tahun 2020. Untuk bank Pemerintah Daerah memiliki tarif yang stabil dari tahun 2010 sampai dengan 2017 dan selanjutnya menjadi yang tertinggi di 2 tahun terakhir.

### 3.2.4 Data Mining

Dalam proses KDD (*Knowledge Discovery using Database*), digunakan *tools rapid mining*. Selanjutnya dipilih fitur clustering menggunakan K-Means, untuk prosesnya adalah sebagai berikut;

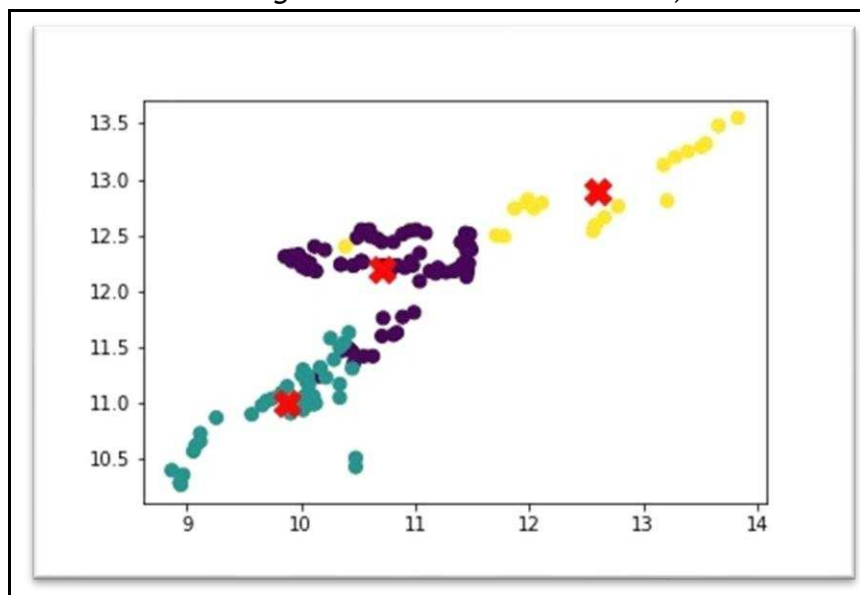
1. Buka tools RapidMiner Studio
2. Pada jendela pertama pilih Auto Model, akan muncul perintah untuk mengimport data
3. Pilih data yang akan digunakan
4. Ketika data sudah masuk pada jendela kerja, pilih semua kolom data dan klik next
5. Simpan data sementara sebagai backup dengan format RM HDF5 TABLE agar langkah 1-4 tidak perlu dilakukan.
6. Selanjutnya pilih task yang dipakai, terdapat tiga pilihan yaitu, predict, clustering dan outliers. Predict adalah merupakan fitur klasifikasi dan outlier adalah fitur untuk melakukan mendeteksi outlier pada data atau abnormal data.
7. Pilih clustering, secara otomatis akan dihitung korelasi antar kolom pada data.
8. Terdapat tiga warna indikator korelasi, warna merah adalah tidak memiliki korelasi, warna biru korelasi sedang dan warna hijau korelasi tinggi. Pada data ini, kelima kolom bank memiliki korelasi tinggi, sehingga bisa dilanjutkan.
9. Akan muncul pilihan model clustering yaitu K-Means dengan default jumlah cluster adalah 2. Selain itu ada model pembanding yaitu X-Means yaitu, memilih jumlah cluster maksimal yang mungkin bisa dihasilkan dari data.
10. Selain itu, pilih extract information from date, karena kita memiliki data time series
11. Maka clustering menggunakan algoritma K-Means menggunakan rapidminer sudah selesai, save semua hasil agar bisa diakses untuk interpretasi data nantinya.

### 3.2.5 Data Interpretation

Dari hasil *K-Means Clustering*, diperoleh beberapa insight:

#### 3.2.5.1 Hasil *clustering* dengan tidak melakukan ekstraksi informasi terhadap waktu

Pemilihan ekstraksi informasi waktu merupakan salah satu fitur dalam rapidminer dalam melakukan *K-Means clustering*. Dalam eksperimen ini, pertama tidak dilakukan pemilihan fitur ini, selanjutnya diperoleh hasil clustering menjadi tiga klaster dan hasilnya bisa dilihat dalam gambar 6. Ketiga klaster terpisah berdasarkan centroidnya, berdasarkan tabel 1. Terlihat bahwa *clustering* berdasarkan nilai maksimum, medium dan minimum.



Gambar 4 Pembagian Cluster terhadap korelasi antar tarif suku bunga



Untuk anggota masing-masing klaster dapat dilihat pada gambar 5 berikut, klaster 1 terdapat 43 anggota, dipilih berdasarkan persentase data dengan suku bunga terbesar yaitu Bank Persero 64.60%, Bank Asing dan Campuran 53.68% dan Bank Umum 47.97%. Sedangkan Klaster 2 dengan 41 anggota, memiliki Persentase suku bunga yang stabil berkisar antara 54-57% untuk Bank Pemerintah Daerah, Swasta Nasional dan Bank Umum. Untuk klaster terakhir yaitu, klaster 3 dengan 72 anggota dengan Bank Persero 11.70%, Bank Pemerintah Daerah 9.70% dan Bank Asing 6.93%.



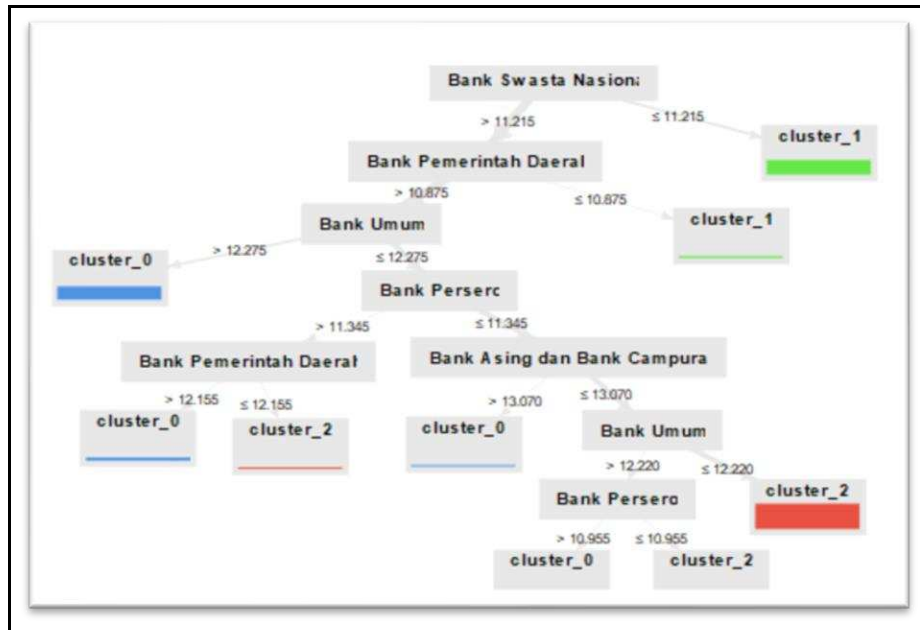
Gambar 5 K-Means Summary dari proses clustering pada tools rapidminer

Selanjutnya untuk melihat perbandingan centroid pada masing-masing klaster, bisa dilihat pada gambar 8 dibawah. Dari gambar dilihat bahwa centroid untuk klaster 1 memiliki nilai yang tinggi-tinggi sehingga bisa dikategorikan bahwa klaster satu memiliki potensi yang menguntungkan untuk melakukan investasi. Sedangkan klaster 2 memiliki nilai centroid yang rendah, artinya nilai suku bunga berada dikisaran nilai centroidnya. Selanjutnya klaster 3 menjadi pilihan kedua untuk melakukan investasi karena nilai centroidnya lebih besar dari klaster 2.

Cluster ↑	Bank Asing dan Ban...	Bank Pemerintah Daerah	Bank Persero	Bank Swasta Nasional	Bank Umum
Cluster 0	11.946	12.613	11.920	13.472	12.752
Cluster 1	8.951	10.985	9.848	10.287	10.070
Cluster 2	10.090	12.122	10.506	12.237	11.560

Gambar 6 Perbandingan centroid ketiga cluster

Selanjutnya dalam proses pengelompokan dengan menggunakan *K-Means*, ilustrasinya bisa dijelaskan dalam diagram pohon *K-Means clustering* berikut;

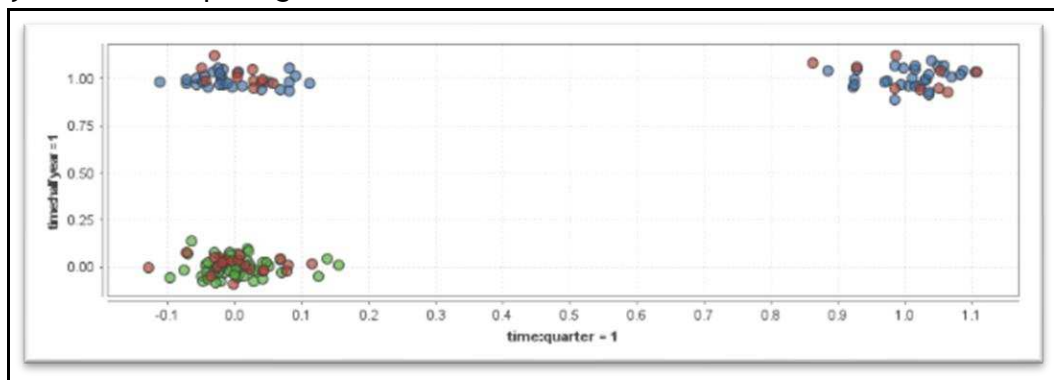


Gambar 7 K-Means Clustering Tree, diagram alir pengelompokan

Dari gambar 8 dapat diketahui bahwa, pengelompokan dilakukan berdasarkan kemiripan dengan masing-masing rentang suku bunga. Dimana suku bunga dengan besar lebih dari 12 akan masuk menjadi klaster 1, kemudian klaster 2 memilih suku bunga paling kecil. Selanjutnya klaster 3 memilih anggotanya dengan suku bunga yang berada di antara klaster 1 dan 2.

### 3.2.5.2 Hasil clustering dengan melakukan ekstraksi informasi terhadap waktu:

Pada pemilihan fitur ini, data akan dibagi menjadi beberapa bagian berdasarkan waktu. Proses ini akan membuat kolom baru yang berisi hari, minggu, bulan dan tahun. Dengan menambah kolom waktu ini pola data memungkinkan untuk mendapatkan informasi. Untuk hasilnya bisa dilihat pada grafik 3 dibawah.



Gambar 8 Hasil Cluster dengan fitur ekstraksi informasi waktu

Dari gambar 9 didapatkan bahwa hasil clustering tidak cukup baik, karena klaster masih berkumpul dalam satu titik. Padahal terdapat tiga buah kumpulan data. Hal ini disebabkan kemungkinan terdapat *overfitting*. *Overfitting* yaitu, penambahan variabel baru menyebabkan data menjadi tidak berkorelasi. Sehingga fitur ekstraksi informasi waktu tidak cocok dalam penelitian ini.

Hasil yang diperoleh dari kedua percobaan untuk mendapatkan knowledge yaitu berdasarkan klustering atau pengelompokan suku bunga investasi dan klustering dengan ekstraksi informasi berdasarkan waktu implementasi dari hasil ini bisa di jadikan sebuah acuan bagi nasabah untuk memilih tempat invesatasi yang mendatangkan keuntungan tidak hanya besar tetapi juga setabil karena ada informasi berdasarkan waktu yang diperoleh.

## KESIMPULAN

Dari hasil K-Means Clustering pada 5 jenis bank, yaitu Bank Persero, Pemerintah Daerah, Swasta Nasional, Asing dan Umum dengan masing-masing tarif suku bunga investasinya. Sesuai dengan goal dari hasil clustering ini diharapkan diperoleh sebuah *knowledge*. Knowledge ini berguna sebagai decision making dalam menentukan bank yang tepat untuk berinvestasi. Setiap bank memiliki karakteristik tarif suku bunga investasi yang berbeda-beda terhadap waktu. Dari hasil clustering diperoleh bahwa bank dengan suku bunga terbesar berada dalam klaster 1 yaitu Bank Persero, Asing dan Umum. Akan tetapi klaster 1 memiliki tarif suku bunga yang kurang stabil, karena hanya memiliki anggota yang sedikit. Untuk lebih stabil dan memiliki tarif yang hampir sama dengan klaster 1, maka klaster 3 adalah pilihan yang paling aman karena memiliki 73 anggota. Bank yang masuk dalam klaster 3 adalah Pemerintah Daerah, Bank Persero dan Bank Asing. Sedangkan klaster 2, tidak direkomendasikan untuk melakukan investasi, karena memiliki tarif suku bunga yang rendah dan tidak stabil.

Untuk saran dari penelitian ini adalah, perlu dilakukan perbandingan hasil clustering dengan metode similarity dengan distance measure menggunakan Hamming Distance, Euclidean Distance, Manhattan Distance dan Minkowski Distance. Untuk mendapatkan pembandingan sebagai pengukur tingkat akurasi penelitian ini.

## BIBLIOGRAFI

- Aprilla Dennis. (2013). Belajar Data Mining dengan RapidMiner. Innovation and Knowledge Management in Business Globalization: Theory & Practice, Vols 1 and 2, 5(4), 1-5. <https://doi.org/10.1007/s13398-014-0173-7.2>
- Dwiasuti, Ninuk (2020). Pengaruh Kredit Perbankan Terhadap Pertumbuhan Ekonomi dan Hubungannya Dengan Kesejahteraan Masyarakat Kabupaten/Kota di Provinsi Kalimantan Barat. Prosiding Seminar Akademik Tahunan Ilmu Ekonomi dan Studi Pembangunan 2020. Retrieved from <http://feb.untan.ac.id/wp-content/uploads/2020/12/Ninuk-Dwiasuti.pdf>
- Hariato, Kristanto N. & Christopher A.L (2016) Implementasi K-Means Clustering untuk Pengelompokan Analisis Rasio Profitabilitas dalam Working Capital. JUISI, Vol. 02, No. 01, Februari 2016. <https://journal.uc.ac.id/index.php/JUISI/article/view/111>
- Nurdiawan, O., & Salim, N. (2018). Penerapan Data Mining Pada Penjualan Barang Menggunakan Metode Metode Naive Bayes Classifier Untuk Optimasi Strategi Pemasaran. Jurnal Teknologi Informasi Dan Komunikasi, 11(1), 84-95. Retrieved from <https://jurnalstmiksubang.ac.id/index.php/jtik/article/view/124>
- Suindyah D, S. (2017). Pengaruh Investasi, Tenaga Kerja Dan Pengeluaran Pemerintah Terhadap Pertumbuhan Ekonomi Di Provinsi Jawa Timur. Ekuitas (Jurnal Ekonomi Dan Keuangan), 15(4), 477. <https://doi.org/10.24034/j25485024.y2011.v15.i4.2312>
- Uska, M. Z., Usuluddin, U., & Rasyid Hardi Wirasasmita, Usuluddin, Baiq Desi Dwi Arianti . (2020). Evaluation of Rapidminer-Application in Data Mining Learning using PerSIVA Model. Edumatic: Jurnal Pendidikan Informatika. DOI:10.29408/edumatic.v4i2.2688.
- Windha Mega P. D. (2015). Clustering menggunakan K-Means untuk Menentukan Status Gizi Balita. Jurnal Informatika Vol 15.
- Yunita, R. (2016). Faktor-Faktor Yang Mempengaruhi Tingkat Profitabilitas Perbankan Syariah Di Indonesia (Studi Kasus pada Bank Umum Syariah di Indonesia Tahun 2009 -2012). Jurnal Akuntansi Indonesia, 3(2), 143. <https://doi.org/10.30659/jai.3.2.143-160>

## COPYRIGHT

Copyright (c) 2021 Arief Wibowo, Raden Sasongko



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/).