

PENGELOMPOKAN BIDANG LAJU PERTUMBUHAN EKONOMI INDONESIA MENGUNAKAN ALGORITMA K-MEANS

Amril Mutoi Siregar

Universitas Buana Perjuangan Karawang

amrilmutoi@ubpkarawang.ac.id

ABSTRACT

Indonesian is one of countries with economic development in the very good category. Economic growth is seen from several supporting fields, Indonesia has a lot of excess natural resources, which can support the economy compared to other countries. But the problem faced is the lack of maximum management of the economy, Indonesia has economic support categorized into 17 fields. Among the fields not in the same development because they are still stuck in one area, it turns out that Indonesia has all the potential to improve all fields. To increase the growth of all fields, the government must have correct, accurate and relevant data to group these fields. In this study using the Decision Tree algorithm to classify fields supporting economic growth automatically. The grouping results into three classes, namely high, medium, low. After the research was conducted the results were that the high category group was Mining and Excavation, Construction, transportation and warehousing, Provision of accommodation and food Drinking, Information and Communication, Financial Services and Insurance, Real Estate, Educational Services, Health Services and Social Activities, medium groups were Procurement of Electricity and Gas, Company Services and low-income groups are in the fields of Agriculture, Forestry, and Fisheries, Processing Industry, water supply, waste management, Waste and Recycling, large Trade and retail, car and motorcycle repair, Government Administration, Defense and Compulsory Social Security, Other Services

Keywords: *Clustering, Datamining, K-Means Algorithm, Economics.*

ABSTRAK

Indonesia adalah salah satu negara yang perkembangan ekonomi dalam kategori sangat baik. Pertumbuhan ekonomi dilihat dari beberapa bidang yang menunjang, Indonesia memiliki banyak kelebihan sumber daya alam, yang dapat menunjang perekonomian dibandingkan negara lain. Tapi permasalahan yang dihadapi adalah kurang maksimalnya pengelolaan bidang ekonomi, Indonesia memiliki penunjang ekonomi dikategorikan ke dalam 17 bidang. Ternyata Indonesia memiliki semua potensi untuk meningkatkan semua bidang. Untuk meningkatkan pertumbuhan semua bidang, pemerintah harus memiliki data yang benar, akurat dan relevan untuk mengelompokkan bidang tersebut. Dalam penelitian ini menggunakan algoritma Decision tree untuk mengklasifikasikan bidang penunjang pertumbuhan ekonomi secara otomatis. Hasil pengelompokan kedalam tiga kelas yaitu: tinggi, sedang, rendah. Setelah dilakukan penelitian dengan hasilnya yaitu kelompok kategori tinggi adalah pertambangan dan penggalian, konstruksi, transportasi dan pergudangan, penyediaan akomodasi dan minum makan, informasi dan komunikasi, jasa keuangan dan asuransi, real estate, jasa pendidikan, jasa kesehatan dan kegiatan sosial, kelompok sedang adalah pengadaan listrik dan gas, jasa perusahaan dan kelompok rendah adalah bidang pertanian, kehutanan, dan perikanan, industri pengolahan, pengadaan air, pengelolaan sampah, limbah dan daur ulang, perdagangan besar dan eceran, reparasi mobil dan sepeda motor, administrasi pemerintahan, pertahanan dan jaminan sosial wajib, dan jasa lainnya.

Kata kunci: *Clustering, Datamining, Algoritma K-Means, Ekonomi.*

1. PENDAHULUAN

1.1 Latar Belakang Masalah

Pertumbuhan ekonomi Indonesia dalam 10 tahun terakhir ini mencapai 5–5.5%, (Republika.co.id) ini didukung dari berbagai bidang yang menjadi variabel pertumbuhan ekonomi. Proses perkembangan ekonomi suatu Negara secara berkelanjutan disebut pertumbuhan ekonomi, atau proses perkembangan ekonomi dari rendah menjadi lebih tinggi selama dalam kurun tertentu. Peningkatan proses produksi suatu perekonomian yang diwujudkan dalam peningkatan pendapatan nasional disebut pertumbuhan ekonomi. Peningkatan ekonomi adalah indikator keberhasilan pembangunan ekonomi dalam kehidupan bermasyarakat dan bernegara.

Perkembangan ekonomi pada sebuah Negara dapat di ukur dengan cara membandingkan, contoh indikator nasional, pendapatan nasional bruto, pada tahun yang sedang berjalan dengan tahun sebelumnya. Teori dibuat berlandaskan pengalaman empiris, sehingga teori yang dapat kita jadikan sebagai dasar untuk prediksi membuat suatu kebijakan. Beberapa teori yang menerangkan pertumbuhan ekonomi. Teori diatas sering dipakai landasan kebijakan. Pemerintahan presiden Jokowi dan Jusuf Kalla, ekonomi Indonesia tercatat sebagai Negara kemajuan ekonominya paling cepat. Termasuk Negara yang naiknya luar biasa berusaha dari bank dunia. Pada Tahun 2007, peringkat Indonesia meningkat dari posisi 106 menjadi posisi 91. Perbaikan terus berkelanjutan, hasil eodb yang tahun 2018 terbaru, menunjukkan tercapainya ke peringkat 72.

Lihat kondisi saat ini, presiden Jokowi memanfaatkan momentum ekonomi yang sudah membaik. Terlebih pemerintah menargetkan pertumbuhan ekonomi Indonesia pada level 5.5 pada tahun 2018. Dari peningkatan 5,02 persen, sedangkan target tahun ini 5,2 persen. Menurut badan pusat statistik mengumumkan pertumbuhan ekonomi Indonesia mencapai 5,07 persen pada tahun 2017, meleset dari target pemerintah 5,2 persen. Meski hasilnya seperti itu, menteri koordinator bidang perekonomian Darmin Nasution optimis target pertumbuhan ekonomi tahun ini yang sebesar 5,4 persen bakal tercapai.

Lebih lanjut, Darimin mengatakan tanda tanda ekonomi membaik juga tercermin dari kinerja bisnis diberbagai sektor, yang menarik untuk diperdebatkan adalah kinerja sektor perdagangan tapi perdagangan akomodasi dan restaurant, pergudangan dan transportasi telekomunikasi semua naik. Jadi bisa dikatakan perdagangan dikatakan lesu/tidak bergairah. Salah satu sektor yang paling potensial yang dapat mendongkrak pertumbuhan laju ekonomi Indonesia dari 5,17 persen menjadi 7 persen adalah sektor ekonomi pembangunan ekonomi maritim. Seperti yang dilansir pengamat kemaritiman Prof.Dr.Rokmin Dahuri kepada tim Indonesia jumat (09/02/2019).

Rokmin mengatakan Indonesia sebagai Negara banyak pulau pulau dengan jumlah 17.504, pulau dan terdiri 108.000 km garis pantai seluruh Indonesia atau sekitar $\frac{3}{4}$ wilayah berupa laut sebagai ekonomi strategis bidang maritim, menawarkan sangat banyak potensi, sektor maritim dianggap dapat meningkatkan pertumbuhan ekonomi yang tinggi berkualitas dan inklusif secara kontinu. Rokmin mengatakan bahwa dari sektor ekonomi Indonesia meliputi 11 sektor, dimana potensi nilai ekonomi dari ekonomi yang dari 11 sektor mencapai 1,5 triliun dollar atau setara 1,5 PDB dengan potensi lapangan kerja mencapai 45 juta orang.

Dengan permasalahan diatas bahwa sangat penting informasi yang dibutuhkan untuk bisa mengelompokkan bidang apa saja yang dalam pertumbuhan ekonomi suatu bangsa, agar dapat mengambil keputusan yang baik, karena di antara bidang masing terdapat perkembangan tidak signifikan, ada juga mengalami penurunan. Dalam ilmu *datamining* adalah salah yang cocok untuk mengeksplorasi data tersebut. Karena *datamining* memiliki banyak memiliki teknik (Larose, DT, 2014), seperti asosiasi (Tan, et al, 2006), klasifikasi (Ransi N, 2014), *clustering*, yang banyak digunakan untuk mengolah data menjadi informasi (Yunanto W, et al, 2012).

1.2 Penelitian Sebelumnya

Penelitian ini memiliki banyak keterkaitan dengan penelitian sebelumnya, *prosiding* (Siregar AM, 2018), didalam artikelnya yang berjudul penerapan *clustering* algoritma dengan SVD dalam meningkatkan produksi tanaman jagung. Hasil penelitiannya adalah memberikan data yang akurat, relevan dalam mengambil keputusan oleh pemerintah dalam rangka meningkatkan produksi tanaman jagung.

Penelitian Eko Wicaksono Pambudi dan Miyasto (2013) yang berjudul analisis pertumbuhan ekonomi dan faktor-faktor yang mempengaruhi Kabupaten/Kota di Provinsi Jawa Tengah, hasil penelitiannya menunjukkan bahwa aglomerasi berpengaruh negatif tetapi tidak signifikan terhadap pertumbuhan ekonomi, investasi berpengaruh positif dan signifikan terhadap pertumbuhan ekonomi, variabel angkatan kerja berpengaruh positif dan signifikan terhadap pertumbuhan ekonomi, dan variabel modal manusia berpengaruh positif. tetapi tidak signifikan terhadap pertumbuhan ekonomi.

Penelitian Susanti H, et al. (2017) yang berjudul analisis pengaruh pertumbuhan ekonomi dan pengeluaran Pemerintah Aceh terhadap pendapatan asli daerah Provinsi Aceh setelah tsunami. Hasilnya penelitiannya adalah estimasi model menunjukkan bahwa jika semua variabel nol maka provinsi PAD Aceh setelah tsunami (Y) berjumlah 1.442 miliar rupiah; jika pertumbuhan ekonomi (X1) tumbuh sebesar 1 persen akan meningkat sebesar 0,016 persen y; jika pengeluaran untuk Gubernur Aceh (X2) meningkat sebesar 0,388 persen Y. Pemerintah Aceh perlu meningkatkan laju pertumbuhan ekonomi di Provinsi Aceh dengan memberikan perhatian khusus pada sektor-sektor unggulan dan potensi untuk dikembangkan di Provinsi Aceh adalah diharapkan dapat meningkatkan jumlah pendapatan Provinsi Aceh.

Penelitian Zumaidah L,N et al.(2018) yang berjudul pengaruh total aset, dana pihak ketiga, dan kredit pada bank umum terhadap pertumbuhan ekonomi Provinsi-provinsi di Indonesia pada tahun 2013-2016. Hasil penelitian menunjukkan total aset (X1) berpengaruh positif signifikan pertumbuhan ekonomi dengan nilai koefisien 0.175711 dan nilai probabilitas 0.0075. Nilai Dana Pihak Ketiga (X2) berpengaruh positif dan signifikan terhadap pertumbuhan ekonomi dengan nilai koefisien 0.285815 dan nilai probabilitas 0.0000. Nilai Kredit (X3) berpengaruh positif signifikan terhadap pertumbuhan ekonomi dengan nilai koefisien 0.088468 dan nilai probabilitas 0.0023.

1.3 Tinjauan Pustaka

1.3.1 Datamining

Datamining berhubungan dengan bidang ilmu lainnya misalnya *artificial intelligence*, *Database System*, *Statistic*, *Information Retrieval*, *Machine Learning*, dan Komputasi dan lain lain. Selain itu, datamining di *support* dalam beberapa bidang contoh NN (*neural network*), *algoritma K-Means*, *Decision tree*, *algoritma C.4.5*, *KNN Algorithm*, *SVM*, dan *Naïve Bayes algorithm* dan sebagainya. Pengolahan gambar dengan menggunakan pola, bagaian data analisis, proses sinyal, dan beberapa survey tentang pemodelan dan metode yang dapat diambil kesimpulan, bahwa datamining diperlukan sebagai penunjuk yang mana datamining menampilkan hasil atas sejarahnya, ditapsirkan untuk petunjuk mengenai dalam proses model datamining (mariscal, et al, 2010).

Menurut Connolly, at al. (2010), *Datamining* yaitu sebuah penambangan atau penggalan pola dalam data yang belum pernah didapatkan sebelumnya, dipahami yang berguna diambil dari database yang skala besar untuk digunakan membantu sebuah hasil yang berguna. Menurut Segall, et al. (2008) *Datamining* disebut juga dengan “Data atau knowledge discovery (KDD)” yaitu yang tersembunyi pola pada sebuah datanya. Menurut Han & Kamber (2016) *Datamining* adalah proses ekstraksi atau menamban pengetahuan yang dibutuhkan dalam data skala besar. Jadi *Datamining* adalah analisis data dari sudut pandang yang berbeda, yang dapat disimpulkan ke dalam informasi yang berguna. Pada dasarnya datamining untuk mengekstrak

hasil yang berharga dalam menganalisis dengan model/ pola/cara atau relasi, keterhubungan dan data yang skala besar, (Siregar AM, et al, 2017).

Datamining adalah sebuah ekstraksi informasi dari sekumpulan data, dan model yang penting dan penting dari dalam data yang berukuran besar. Beberapa jurnal ilmiah, nama lain dari *datamining* adalah penemuan pengetahuan dalam database atau istilah KDD. Adanya *datamining* dilatarbelakangi dengan kompleksnya permasalahan data yang membludak yang dialami terakhir, dimana banyak instansi, kelompok, perusahaan telah mengumpulkan data sekian tahun lamanya, contoh data penjualan, pembelian, nasabah dan transaksi dan lain lain.

1.3.2 Clustering/Pengelompokan

Pengelompokan yaitu sebuah kumpulan objek data yang hampir sama dengan yang lain, didalam *cluster* yang kesamaan dan perbedaan dengan objek data dan pada kelompok lainnya. Padahal, *clustering* atau analisa *cluster* ialah suatu pengelompokan satu set benda benda fisik dan yang ilustrasi kedalam kelas kesamaan objek. Tujuannya ialah untuk mendapatkan kelompok objek yang sama atau mirip satu sama lain dalam kelompoknya. Semakin tinggi tingkat kesamaan objek dalam satu kelompok dan semakin tinggi tingkat beda pada tiap kelompok baik analisa kelompok semakin baik *performance*.

Clustering ini termasuk hirarki *clustering* pada data dari *cluster* yang besar hingga ke *cluster* kecil, adalah masing masing jarak dari *cluster* data tersebut. Sebaliknya *agglomerative* hirarki *clustering* dimulai dengan mengelompokan data yang kecil ke *cluster* data yang besar (Qudri at al, 2010).

Clustering dapat mengenali *cluster* alami dari dalam data yang berdasarkan kemiripan atribut yang disebut dengan segmentasi. *Unsupervised learning method* adalah proses ekstraksi tanpa dapat diprediksi hasilnya dan tidak satupun atribut yang digunakan untuk dapat membantu proses belajar yang tidak berlabel yaitu seperti metode, *fuzzy c-means*, *k-means*, *k-medoid*, hirarki dan lain lain.

1.3.3 Algoritma K-Means

Dalam *Datamining* untuk mengekstrak hasil pengelompokan dengan menggunakan algoritma *k-means* merupakan algoritma *clustering* paling baik *performance* berdasarkan hasil peneliti sebelumnya. Algoritma *k-means* adalah model *clustering* yang tidak hirarki, mengelompokan data kedalam salah satu bentuk satu, dua, dan lebih *cluster*. Data mempunyai karakteristik yang mirip di *cluster* menjadi satu kelompok. Data yang memiliki karakteristik yang berbeda akan dikelompokan ke dalam satu *cluster*/kelompok lainnya. Sampai data kedalam 1 *cluster* mempunyai tingkat beda yang sangat kecil (Agusta, 2007). Berikut langkah-langkah algoritma *k-means* sesuai menurut (Tan, et al, 2006).

1. Menentukan nilai k, sebagai data *centroid*, k merupakan jumlah *cluster* digunakan para peneliti.
2. Setiap data titik akan dicari *centroid* terdekat.
3. Tiap himpunan data titik untuk dijadikan *centroid* titik dinamakan kelompok.
4. Menghitung ulang *centroid* titik pada tiap kelompok.
5. Kembali ulangi yang sama 1-4 sampai *centroid* tidak berubah atau sampai nilainya sama.

Model *clustering* yang digunakan algoritma *k-means*, akan mengukur kedekatan data yang kita hitung menggunakan *euclidean distance*. Algoritma *k-means* untuk bertujuan meminimalkan jarak *euclidean distance*, pada tiap titik X_i dan kelompok terdekat yaitu X_j . Menurut Wu, at al. (2009) *Euclidean Distance* dihitung dengan menggunakan rumus 1 sebagai berikut:

$$d_{ij} = \sqrt{\sum_{k=0}^n \{x_{ik} - x_{jk}\}^2} \quad (1)$$

Dimana:

- D_{ij} = Jarak point data ke i ke data j
- n = banyaknya data
- X_{ik} = titik data ke i dalam dimensi k
- X_{jk} = titik data ke j dalam dimensi k

Formula *Euclidean distance* yaitu rumus dalam menghitung jarak antar jarak, satu titik ke titik yang lain, atau dua titik dalam satu, dua, dan tiga dimensi secara sekuen, rumusnya dapat dilihat pada persamaan 2, 3, 4, sebagai berikut :

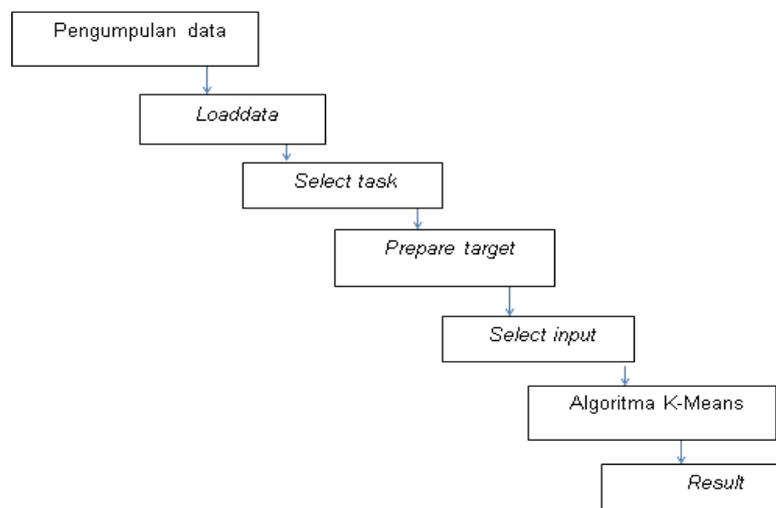
$$\sqrt{(x - y)^2} = |x - y| \quad (2)$$

$$d(p, q) = \sqrt{(p1 - q1)^2 + (p2 - q2)^2} \quad (3)$$

$$d(p, q) = \sqrt{(p1 - q1)^2 + (p2 - q2)^2 + (p3 - q3)^2} \quad (4)$$

2. METODOLOGI

Metode penelitian ini, peneliti mengajukan metode yang sesuai dengan masalah yaitu *clustering*, menggunakan tiga kategori 3 *cluster* yaitu tinggi, sedang dan rendah. Dataset yang akan diperlukan saat penelitian adalah data perkembangan ekonomi dalam berbagai sektor atau bidang, untuk mengelompokkannya menggunakan *K-means Algorithm*. Dalam penelitian peneliti menggunakan *tools rapidminer* visual 9.2, berikut tahapan penelitiannya:



Sumber: Rapidminer Studio 9.2 (2019)

Gambar 1. Metode Penelitian yang digunakan

2.1 Pengumpulan Dataset

Data yang digunakan dalam penelitian ini adalah laju pertumbuhan ekonomi Indonesia yang diambil <https://data.go.id>. Dataset terdiri dari beberapa atribut yaitu bidang, 2010, 2011, 2012, 2013, 2014, 2015, dan 2016. Data yang diambil adalah laju pertumbuhan ekonomi selama 7 tahun dari 2010 sampai 2016, dan terdiri dari 15 bidang yang digunakan adalah sebagai berikut :

1. Pertanian, Kehutanan, dan Perikanan
2. Pertambangan dan Pengalihan

3. Industri Pengolahan
4. Pengadaan Listrik dan Gas
5. Pengadaan Air, Pengelolaan Sampah, Limbah dan Daur Ulang
6. Konstruksi
7. Perdagangan eceran dan besar, reparasi mobil dan sepeda motor
8. Pergudangan dan transportasi
9. Penyediaan makan minum dan akomodasi
10. Komunikasi dan informasi
11. Jasa asuransi dan keuangan
12. Real estate
13. Jasa perusahaan
14. Pertahanan dan jaminan social wajib dan administrasi pemerintahan
15. Jasa pendidikan
16. Jasa kesehatan dan social
17. Jasa lainnya

Berikut adalah dataset yang digunakan.

Tabel 1. Dataset yang digunakan

Bidang	Tahun						
	2010	2011	2012	2013	2014	2015	2016
Pertanian, kehutanan dan perikanan	29.42	3.53	3.26	4.57	2.95	3.3	3.76
Pertambangan dan pengalihan	0.08	8.9	8.8	9.79	8.87	9.9	15.4
Industri pengolahan	14.92	4.75	5.59	6.4	5.47	7.92	3.98
Pengadaan listrik dan gas	0.05	10.04	11.61	10.51	25.22	14.44	8.11
Pengadaan air, pengelolaan sampah, limbah dan daur ulang	0.06	6.11	4.88	2.37	4.63	1.86	2.42
Konstruksi	11.66	8.72	8.87	6.29	6.01	2.59	7.48
Perdagangan besar dan eceran, reparasi mobil dan sepeda motor	15.07	6.94	7.16	6.38	6.36	3.47	5.77
Transportasi dan pergudangan	3.63	8.83	9.9	8.18	8.43	12.81	6.51
Penyediaan akomodasi dan makan minum	1.91	7.29	8.97	9.96	6.59	11.21	4.97
Informasi dan komunikasi	4.49	11.94	10.91	8.24	7.37	9.27	6.74
Jasa keuangan dan asuransi	3.87	8.46	9.06	8.67	8.01	2.75	5.89
Real estate	3.75	7.84	8.3	9.09	8.5	6.24	4.94
Jasa perusahaan	0.18	12.31	13.32	13.99	13.38	7.59	3.93
Administrasi pemerintahan, pertahanan dan jaminan sosial wajib	4.04	4.44	5.06	5.73	6.58	5.13	1.52
Jasa pendidikan	4.46	9.19	9.08	9.28	9.41	6.29	5.5
Jasa kesehatan dan kegiatan sosial	1.35	6.4	7.02	8.04	8.67	7.8	6.06
Jasa lainnya	1.07	5.59	5.15	4.93	5.77	9.22	4.8

Sumber: <https://data.go.id> Dataset Penelitian (2019)

2.2 Load Data

Dalam penelitian ini setelah dataset sudah seleksi atribut kemudian dilakukan *load data*, yaitu proses upload load dataset untuk sebuah pengujian dalam penelitian ini. Pertama-tama yang dilakukan adalah dengan memilih kumpulan dataset yang akan digunakan dalam penelitian ini. Jadi *load data* dari *repositori*, setelah langkah ini lanjut ke berikutnya.

2.3 Select Task

Dalam proses load data selesai, selanjutnya memilih permasalahan apa yang kita uji, Pengujian dalam penelitian ini untuk clustering. Task yang digunakan menentukan kedalaman Cluster. Tujuannya untuk menemukan set jarak data yang berdekatan dengan data yang lain.

2.4 Prepare Target/Select Input

Prepare data dan *select input* dilakukan secara otomatis oleh *tools rapidminer* untuk memilih target/tujuan. Tahapan untuk mendapatkan pola atau model, dengan tidak melihat macam macam data, dan banyak cara untuk melihat data yang tidak digunakan. Untuk *prepare* tujuan dan klik masukan dengan banyak dapat dipertimbangan yaitu kolom tabel yang terdekat dengan column tujuan, value dalam column diharuskan berbeda dengan nilai-nilai column tidak boleh ada yang kosong.

2.5 Algoritma K-Means

Penelitian sekarang dilakukan dengan menggunakan metode atau *K-means algorithm* untuk kelompok masuk kedalam bukan hirarki *clustering* dalam bidang, dengan memilih 3 *cluster*.

3. HASIL DAN PEMAHASAN

Tahapan ini peneliti mendapatkan hasil yang dimunculkan secara *automatic*. Pada hasil tergantung pada *dataset* yang kita pilih. Untuk melihat dan melakukan proses data yang digunakan untuk optimasi pola. Sebagai model digunakan titik awal untuk mengoptimalkan. Berikut langkah langkah *clustering* dengan metode *Euclidean distance* adalah:

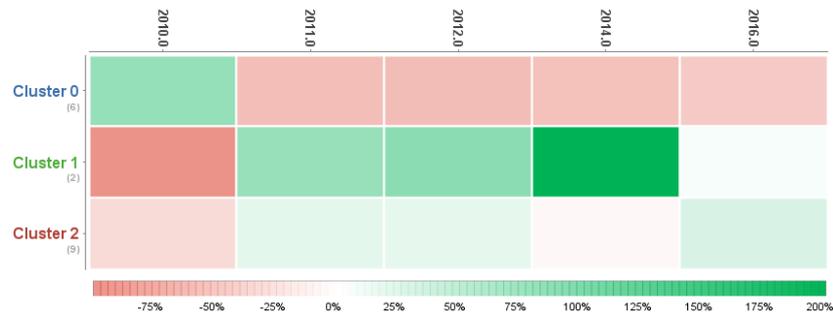
1. *Heatmap* ialah identifikasi atribut yang penting untuk setiap kelompok.
2. *Clustertree* yaitu untuk menghasilkan *decision tree* untuk menggambarkan perbedaan dalam kelompok.
3. *Centroidchart* yaitu menampilkan *values cluster centroid* dalam bentuk grafik.
4. *Centroidtable* yaitu untuk menampilkan *values cluster centroid* dan bentuk tabel.
5. *Clusterview* yaitu menunjukkan ilustrasi data kedalam cluster.
6. *Dataclustered* yaitu menunjukkan semua tabel data dan termasuk label kelompok pada data untuk sebuah titik.
7. Hasil Jumlah per cluster yaitu jumlah bidang dalam cluster.

3.1 Hasil Penelitian

Setiap penelitian harus di akhiri dengan hasil, untuk penelitian ini didapat beberapa informasi dalam *clustering*. Menggunakan *rapidminer* adalah *tools* dalam pengujian, dan mengolah data. Menggunakan prinsip algoritma dalam *datamining*, *rapidminer* dapat mengekstrak berbagai macam model atau pola dari dataset yang besar dengan melibatkan kecerdasan buatan, *database*, statistika. *Tools* ini sangat mudah digunakan oleh peneliti untuk pengujian data, dan difasiltasi dengan banyak menggunakan *operator*.

Hasilnya dapat ditampilkan secara visual berupa grafik, tabel, gambar, persentasi, sehingga mudah dipahami oleh pengguna dalam menganalisa data laju pertumbuhan ekonomi Indoensia, dengan menggunakan pengelompokan data, kedalam 3 kelompok/*cluster*, terdiri 17 bidang ekonomi dengan rentang tahun 2010 sampai 2016 dengan 7 atribut yaitu bidang, 2010, 2011, 2012, 2013, 2014, 2015, 2016. Hasil penelitian seperti ini berupa: *Heatmap*, *Clustertree*, *Centroidchart*, *Centroidtable*, *Clusterview*, *Dataclustered*. *K-Means Heat Map*, menampilkan untuk mengenali atribut yang penting dalam kelompok.

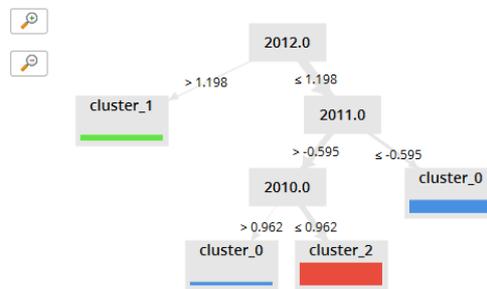
k-Means - Heat Map



Sumber: Hasil Penelitian (2019)
Gambar 2. Heatmap K-means

Cluster Tree pada K-Means untuk mempermudah membaca keputusan atau kesimpulan, berikut hasil pohon keputusan, menggambarkan perbedaan utama antar cluster.

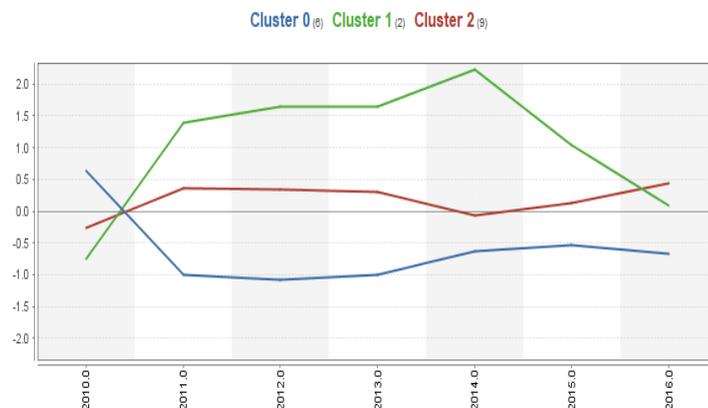
k-Means - Cluster Tree



Sumber: Hasil Penelitian (2019)
Gambar 3 K-Means Cluster Tree

Centroid Chart dalam K-means, melihat pergerakan grafik pada nilai-nilai untuk cluster centroid dalam grafik paralel.

k-Means - Centroid Chart



Sumber: Hasil Penelitian (2019)
Gambar 4 K-Means Centroid Chart

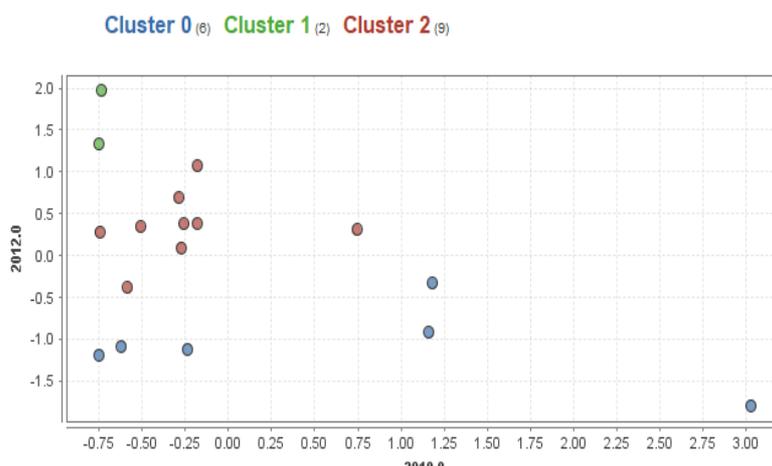
Centroid Table dalam K-means adalah hasil nilai-nilai dalam centroid cluster dalam gambar.

k-Means - Centroid Table

Cluster	2010.0	2011.0	2012.0	2013.0	2014.0	2015.0	2016.0
Cluster 0	0.627	-1.011	-1.073	-1.004	-0.625	-0.541	-0.678
Cluster 1	-0.742	1.399	1.648	1.643	2.226	1.035	0.088
Cluster 2	-0.253	0.363	0.349	0.304	-0.078	0.131	0.432

Sumber: Hasil Penelitian (2019)
Gambar 5 K-Means Centroid Table

Clustering dalam *K means* adalah mengilustrasikan titik-titik data dalam sebuah cluster.



Sumber: Hasil Penelitian (2019)
Gambar 6 Clustering dalam K-Means

Data Clustered dalam clustering, berikut hasil yang ditampilkan pada tabel untuk semua data, termasuk label *cluster* untuk setiap titik data.

Tabel 2. Hasil Clustering Penelitian

Bidang	Tahun							Cluster
	2010	2011	2012	2013	2014	2015	2016	
Pertanian, Kehutanan, dan Perikanan	3.03	(1.70)	(1.79)	(1.19)	(1.10)	(1.04)	(0.66)	cluster_0
Industri Pengolahan	1.16	(1.20)	(0.92)	(0.51)	(0.59)	0.20	(0.59)	cluster_0
Pengadaan Air, Pengelolaan Sampah, Limbah dan Daur Ulang	(0.75)	(0.65)	(1.19)	(2.00)	(0.76)	(1.43)	(1.10)	cluster_0
Perdagangan Besar dan Eceran; Reparasi Mobil dan Sepeda Motor	1.18	(0.32)	(0.33)	(0.52)	(0.41)	(0.99)	0.01	cluster_0

Administrasi Pemerintahan, Pertahanan dan Jaminan Sosial Wajib	(0.24)	(1.33)	(1.12)	(0.76)	(0.36)	(0.55)	(1.40)	cluster_0
Jasa Lainnya	(0.62)	(0.86)	(1.09)	(1.05)	(0.53)	0.55	(0.32)	cluster_0
Pengadaan Listrik dan Gas	(0.75)	0.94	1.33	1.00	3.43	1.96	0.78	cluster_1
Jasa Perusahaan	(0.73)	1.86	1.97	2.28	1.02	0.11	(0.60)	cluster_1
Pertambangan dan Penggalian	(0.75)	0.48	0.28	0.74	0.10	0.74	3.21	cluster_2
Konstruksi	0.74	0.40	0.30	(0.55)	(0.48)	(1.23)	0.57	cluster_2
Transportasi dan Pergudangan	(0.29)	0.45	0.69	0.14	0.01	1.52	0.25	cluster_2
Penyediaan Akomodasi dan Makan Minum	(0.51)	(0.18)	0.34	0.80	(0.36)	1.09	(0.26)	cluster_2
Informasi dan Komunikasi	(0.18)	1.71	1.07	0.17	(0.20)	0.57	0.33	cluster_2
Jasa Keuangan dan Asuransi	(0.26)	0.30	0.38	0.32	(0.07)	(1.19)	0.05	cluster_2
Real Estate	(0.27)	0.05	0.09	0.48	0.03	(0.25)	(0.27)	cluster_2
Jasa Pendidikan	(0.18)	0.59	0.38	0.55	0.21	(0.24)	(0.08)	cluster_2
Jasa Kesehatan dan Kegiatan Sosial	(0.58)	(0.54)	(0.39)	0.09	0.06	0.17	0.10	cluster_2

Sumber: Hasil Penelitian (2019)

3.2 Pembahasan Penelitian

Berikut hasil penelitian setelah melakukan pengujian, hasil ditampilkan dalam bentuk tabel, hasil *clustering* dalam 17 bidang yang di jadikan variabel pertumbuhan ekonomi Indoensia dengan menggunakan 3 *cluster* sebagai berikut :

Tabel 3. Hasil *Clustering*

Cluster	Kategori Cluster		
	Cluster rendah (0)	Cluster Sedang (1)	Cluster Tinggi (2)
Jumlah Cluster	6	2	9

Sumber: Hasil Penelitian (2019)

Hasil pengelompokan dari penelitian ini adalah memiliki 3 cluster yaitu tinggi, sedang dan rendah. Hasil ini bertujuan untuk mendapatkan data bidang yang dikelompokan sesuai dengan perkembangan dari tahun 2010 sampai 2016. Perhitungan disesuaikan dengan tingkat kedekatan yang terdekat, cara ini sangat baik akurasi dalam pengelompokan berikut hasil yang didapatkan:

1. **Cluster rendah;** meliputi: bidang kehutanan, pertanian, industri pengolahan, perikanan, pengelolaan sampah, limbah dan daur ulang, pengelolaan air, perdagangan eceran dan besar, reparasi mobil dan sepeda motor, administrasi pemerintahan, pertahanan dan Jaminan social dan jasa lainnya.
2. **Cluster sedang;** meliputi: bidang pengadaan listrik dan gas, jasa perusahaan.

3. **Custer tinggi**; meliputi: bidang pertambangan dan penggalian, konstruksi, transportasi dan pergudangan, penyediaan akomodasi dan minum makan, jasa pendidikan, komunikasi dan informasi, jasa keuangan dan asuransi, real estate, jasa kesehatan dan kegiatan sosial.

Perlu dipahami dalam penelitian ini, data dari tahun 2010 sampai 2016 sangat fluktuatif, contoh : bidang industri pengolahan tahun 2010 = 14.92, tahun 2011 = 4.75, tahun 2012 = 5.59, tahun 2013 = 6.4, tahun 2014 = 5.47, tahun 2015 = 7.92 tahun 2016 = 3.98. ketika diperhatikan datanya di awal tahun 2010 sangat tinggi perkembangan turun terus jadi selama 5 tahun terakhir praktis tidak perkembangan yang baik ternyata bidang ini masuk *cluster rendah*.

Bidang Pengadaan Listrik dan Gas tahun 2010 = 0.05, tahun 2011 = 10.04, tahun 2012 = 11.61, tahun 2013 = 10.51, tahun 2014 = 25.22, tahun 2015 = 14.44, tahun 2016 = 8.11 perhatikan datanya di awal sangat kecil dan tahun berikut lebih stabil perkembangannya, secara angka fluktuatif, bidang ini masuk *cluster sedang*.

Bidang Pertambangan dan Penggalian tahun 2010 = 0.08, tahun 2011 = 8.9, tahun 2012 = 8.8, tahun 2013 = 9.79, tahun 2014 = 8.87, tahun 2015 = 9.9, tahun 2016 = 15.44 perhatikan datanya di awal sangat kecil dan tahun berikut lebih stabil perkembangannya, secara angka tidak terlalu jauh fluktuatif, bidang ini masuk *cluster tinggi*

4. PENUTUP

4.1 Kesimpulan

Tahapan terakhir penelitian ini adalah tahap pengambilan kesimpulan dengan *dataset* sebanyak 17 bidang laju pertumbuhan ekonomi dengan rentang waktu tahun 2010 sampai 2016 yang diambil data pemerintah berikut kesimpulan yang didapat:

1. Hasil *clustering* dengan algoritma *K-means* menunjukkan hasil *cluster* rendah adalah Pertanian, Kehutanan, dan Perikanan, Industri Pengolahan, pengadaan air, pengelolaan sampah, limbah dan daur ulang, perdagangan besar dan eceran, reparasi mobil dan sepeda motor, administrasi pemerintahan, pertahanan dan jaminan social wajib, Jasa Lainnya. Dimana peningkatan kurang baik selama 6 tahun yang diuji.
2. Hasil *cluster* rendah data tidak stabil dari tahun ke tahun, untuk *cluster* sedang data meningkat tapi tidak terlalu tinggi peningkatannya dari tahun ke tahun, dan *cluster* tinggi selama 6 tahun yang uji jauh lebih stabil perkembangannya.

4.2 Saran

Berdasarkan hasil penelitian ini, peneliti memberikan masukan kepada pemangku kebijakan sebagai data akurat, informasi yang relevan untuk menunjang keputusan dalam rangka peningkatan bidang pertumbuhan ekonomi. Agar kebijakan pemerintah lebih maksimal, diharapkan oleh peneliti selanjutnya dapat memberikan sebagai referensi ilmiah untuk menganalisa data yang sejenis.

DAFTAR PUSTAKA

- Agusta. (2017). *K-means Penerapan Permasalahan di STMIK STIKOM Bali*. Denpasar Bali.
- Connoly, et al.(2005). *A Practical Approach to Design Implementation and Management*, 4th. Harlow Addison Wesley.
- Larose, DT. (2014). *Discovering Knowledge in Data:an Introduction to Data Mining*. Jhon Wiley and Sons.
- Pambudi, Eko Wicaksono dan Miyasto. (2013). *Analisis Pertumbuhan Ekonomi dan Faktor-faktor yang Mempengaruhi Kabupaten/Kota di Provinsi Jawa Tengah*. <http://ejournal-s1.undip.ac.id/index.php/jme>. Volume 2, Nomor 2. Semarang.

- Qudri, et al. (2010). *Drop Out Feature of Student Data for Academic Performance Using Decision Tree Techniques*. Global Journal of Computer Science and Technology,2-4.
- Rahmat B, et al. (2017). *Implementasi K-means Clustering dengan Rapidminer untuk Analisis Daerah Rawan Kecelakaan*.
- Siregar AM. (2018). *Implementasi Algoritma Clustering dengan Singular Vector Decomposition untuk Menunjang Keputusan dalam Meningkatkan Produksi Tanaman Jagung*. Institut Pertanian Bogor.
- Siregar AM,Et al. (2017). *Data Mining Pengolahan data Menjadi Informasi dengan Rapidminer*. Kekata Publisher : Solo.
- Susanti H, et al. (2017). *Analisis Pengaruh Pertumbuhan Ekonomi dan Pengeluaran Pemerintah Aceh terhadap Pendapatan Asli Daerah Provinsi Aceh Setelah Tsunami*. Volume 4 Nomor 1, Mei 2017. E-ISSN. 2549-8355.
- Tan, et al. (2006). *Introduction to Data Mining*. Boston: Pearsong education, ltd.
- Wu, et al.(2009). *The Top Ten Algoritihms in Data Mining*. Boca Raton:CRC Press.
- Zumaidah L,N et al. (2018). *Pengaruh Total Aset, Dana Pihak Ketiga, dan Kredit pada Bank Umum terhadap Pertumbuhan Ekonomi Provinsi-provinsi di Indonesia pada Tahun 2013-2016*. Jurnal Ilmu Ekonomi Vol 2 Jilid 2/Tahun 2018 Hal. 251 – 263.