UNES Journal of Information System

Volume 3, Issue 1, June 2018

P-ISSN: 2528-3502

E-ISSN: 2528-5955

Open Access at: <http://lppm.ojs.unespadang.ac.id/index.php/UJIS>

ALGORITMA K-MEANS CLUSTERING UNTUK ANALISA DATA SERVICE BERDASARKAN PENGADUAN PELANGGAN

K-MEANS CLUSTERING ALGORITHM FOR SERVICE DATA ANALYSIS BASED ON CUSTOMERS COMBINATION

Zulhendra¹, Gunadi Widi Nurcahyo², Julius Santony³

¹ Mahasiswa Pascasarjana Ilmu Komputer Univ. Putra Indonesia "YPTK" Padang

E-mail: ltikiindocomputer@gmail.com

² Dosen Univ. Putra Indonesia "YPTK" Padang

³ Dosen Univ. Putra Indonesia "YPTK" Padang

INFO ARTIKEL

Koresponden:

Zulhendra

ltikiindocomputer@gmail.com

Kata kunci:

data mining, k-means clustering, software weka, algoritma, student data

hal: 1-7

ABSTRAK

Dalam penelitian ini menggunakan Data Mining yaitu K-Means Clustering. Data Mining dapat digunakan dalam mencari sebuah analisa data yang cukup besar yang bertujuan agar Indocomputer mengetahui dan pengelompokan data servis berdasarkan keluhan pelanggan menggunakan Software Weka. Dalam penelitian ini menggunakan Algoritma K-Means Clustering untuk memprediksi atau pengelompokan data komplain kerusakan hardware pada Indocomputer Payakumbuh. Dan dapat mengetahui data merek Laptop paling banyak melakukan service di Indocomputer Payakumbuh sebagai salah satu rekomendasi kepada konsumen untuk pemilihan Laptop.

Copyright © 2018 UJIS. All rights reserved

ARTICLE INFO

Correspondent:

Zulhendra

ltikiindocomputer@gmail.com

Keywords:

*data mining, k-means
clustering, software
weka, algoritma,
student data.*

page: 1-7

ABSTRACT

In this study using Data Mining, namely K-Means Clustering. Data Mining can be used in searching for a large enough data analysis that aims to enable Indocomputer to know and classify service data based on customer complaints using Weka Software. In this study using the algorithm K-Means Clustering to predict or classify complaints about hardware damage on Payakumbuh Indocomputer. And can find out the data of Laptop brands most do service on Indocomputer Payakumbuh as one of the recommendations to consumers for the selection of Laptops.

Copyright © 2018 UJIS. All rights reserved

PENDAHULUAN

Dalam berbagai bidang kehidupan sehari-hari saat ini, banyak sekali data yang dihasilkan oleh teknologi informasi yang semakin canggih. Mulai dari bidang industri, ekonomi, pendidikan, ilmu dan teknologi serta berbagai bidang ilmu pengetahuan lainnya yang menghasilkan data yang sangat berlimpah. Hal ini pun juga terjadi di toko *service* dan *reparasi* komputer, data-data *service* dan pengaduan pelanggan bisa menjadi acuan serta tolak ukur sebagai bahan perencanaan dan strategi di tempat *service* komputer tersebut.

Algoritma *K-Means* mendasari metode operasi berdasarkan namanya yaitu „mean“ *rerata*. Di mana akan dilakukan pengamatan ke kelompok „*k*“, di mana „*k*“ diberikan sebagai parameter masukan kemudian dilakukan pengamatan pola di setiap kelompok berdasarkan kedekatan pada *rerata* jarak *cluster* adalah titik tengah. Selanjutnya titik pusat atau *centroid* dan proses akan dimulai kembali. *K-Means* adalah teknik yang dikategorikan *greedy*, komputasi dilakukan dengan teknik yang efisien, menjadi salah satu algoritma *Clustering* yang paling banyak digunakan untuk klasifikasi (Sulianta, 2014).

METODE PENELITIAN

1. Kerangka Penelitian

Pada bagian ini diuraikan kerangka kerja dalam penelitian yang akan memandu pekerjaan yang akan dilakukan agar senantiasa fokus pada tujuan penelitian.

Gambar 1. Kerangka Kerja Penelitian melalui

2. Data Mining

Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma dalam data mining sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses KDD secara keseluruhan (Fina Nasari and Surya Darma, 2015). *Data Mining* juga sering disebut sebagai Knowledge Discovery in Databases (KDD). KDD terdiri dari tiga proses utama yaitu:

1. Preprocessing

Preprocessing dilakukan terhadap data sebelum algoritma *Data Mining* diaplikasikan. Proses ini meliputi *Data Cleaning*, integrasi, seleksi dan transformasi.

2. Data Mining

Proses utama dalam KDD adalah proses *Data Mining*, dalam proses ini algoritma-algoritma *Data Mining* diaplikasikan untuk mendapatkan pengetahuan dari sumber data.

3. Post processing

Hasil yang diperoleh dari proses *Data Mining* selanjutnya akan dievaluasi pada post processing.

Gambar 2. Data Mining Sebagai Sebuah Langkah Dalam Proses Knowledge Discovery

3. Clustering (Pengelompokan)

Salah satu teknik yang dikenal dalam Data Mining adalah teknik Clustering. Teknik Clustering merupakan salah satu metode Data Mining yang melakukan pemisahan/pemecahan/segmentasi data kedalam sejumlah kelompok (Cluster) menurut karakteristik tertentu yang diinginkan.

Gambar 3. Pengelompokan Cluster

4. K-Means Clustering

K-Means Clustering adalah suatu metode penganalisaan data atau metode Data Mining yang melakukan proses pemodelan tanpa supervisi (*unsupervised*) dan merupakan salah satu metode yang melakukan pengelompokan data dengan sistem partisi. Metode K-Means berusaha mengelompokkan data yang ada ke dalam beberapa kelompok, di mana data dalam satu kelompok mempunyai karakteristik yang sama satu sama lainnya dan mempunyai karakteristik yang berbeda dengan data yang ada di dalam kelompok yang lain (Fina Nasari and Surya Darma, 2015).

Clusterer	Sum of Squared Errors	Silhouette
2	10.000	0.000
3	8.000	0.000
4	6.000	0.000
5	4.000	0.000
6	2.000	0.000
7	1.000	0.000
8	0.500	0.000
9	0.250	0.000
10	0.125	0.000

Gambar 4. Pengklusteran dengan Algoritma K-Means

HASIL DAN PEMBAHASAN

Tampil hasil dari pengklasteran yang dilakukan dengan teknik data mining metode *k-means clustering*. Pada tampilan hasil *cluster* diperoleh jumlah rata-rata dari setiap *cluster* sama dengan proses yang dilakukan pada sebelumnya.

Cluster	Center	Assignment
1	1.000	1.000
2	2.000	2.000
3	3.000	3.000
4	4.000	4.000
5	5.000	5.000
6	6.000	6.000
7	7.000	7.000
8	8.000	8.000
9	9.000	9.000
10	10.000	10.000

Gambar 5. Hasil Pengklasteran

Pada gambar diatas menampilkan hasil interaksi yang didapatkan dari nilai masing cluster yang terbentuk. Setelah dilakukan penyimpanan output maka kita dapat melihat hasil dari proses yang kita lakukan, untuk hasilnya dapat kita lihat berbentuk textpad seperti Gambar 6.

Cluster	Center	Assignment
1	1.000	1.000
2	2.000	2.000
3	3.000	3.000
4	4.000	4.000
5	5.000	5.000
6	6.000	6.000
7	7.000	7.000
8	8.000	8.000
9	9.000	9.000
10	10.000	10.000

Gambar 6. Tampilan Output Untuk 7Record

Dari dua hasil pengujian yang telah dilakukan yaitu proses secara manual dan menggunakan *software* aplikasi WEKA 3.8.0 dapat diambil sebuah kesimpulan bahwa

hasil pengujian yang diperoleh sangat baik karena *rule* yang dihasilkan adalah sama antara proses dengan cara manual dan dengan menggunakan aplikasi. Informasi yang dihasilkan secara umum dapat dilihat bahwa jumlah Merek baru yang direkap selama 3 tahun terakhir yang memiliki jumlah *service* besar berasal dari merek Asus.

Sehingga dapat diambil suatu keputusan bahwa tujuan *service* yang paling utama dilakukan pada merek Samsung. Selanjutnya pada merek yang berada pada *cluster* 2 yaitu HP. Sedangkan merek pada *cluster* 3 yaitu Asus akan dijadikan pilihan akhir dalam melakukan *service*. Hal ini disebabkan karena jumlah Merek yang berada pada *cluster* 3 tidak terlalu signifikan.

SIMPULAN DAN SARAN

Simpulan

1. Metode *Clustering* yang diproses dengan Weka lebih efektif dan fleksibel jika digunakan pada proses penentuan jenis merek laptop yang rusak.
2. Pemilihan variable (atribut kondisi dan atribut keputusan) yang akan digunakan dalam memprediksi juga sangat mempengaruhi *rule* atau *knowledge* yang dihasilkan.
3. Algoritma *K-Means Clustering* dapat memberikan informasi eksekutif dan digunakan untuk menggambarkan proses yang terkait dengan penentuan tujuan jenis merek laptop yang rusak berdasarkan merk laptop yang di- *service* oleh pelanggan.
4. Sistem pengelompokan data *service* berdasarkan merk menggunakan algoritma *K-Means Clustering* dapat digunakan dalam pengambilan keputusan.

Saran

1. Pada penelitian ini, penulis mencoba salah satu teknik yang digunakan dalam memprediksi yaitu teknik *data mining* dengan metode *K-Means Clustering*. Untuk mendapatkan hasil prediksi yang baik dapat digunakan beberapa atau penggabungan beberapa teknik prediksi. Oleh sebab itu perlunya dilakukan penelitian lanjut untuk membandingkan hasil prediksi dan menentukan teori mana yang menghasilkan prediksi yang baik.
2. Untuk mendapatkan hasil prediksi yang baik diperlukan sumber data yang lengkap. Oleh sebab itu diharapkan pada proses penyimpanan data keluhan pelanggan sudah dilakukan secara komputerisasi sehingga akan memudahkan dalam pencarian data yang dibutuhkan untuk melakukan proses penemuan *knowledge* seperti untuk memprediksi merek laptop yang rusak yang akan dilakukan selanjutnya.

DAFTAR PUSTAKA

- Sulianti. 2014. *Customer Profiling Pada Supermarket Menggunakan Algoritma K- Means Dalam Memilih Produk Berdasarkan Selera Konsumen Dengan Daya Beli Maksimum*. Jurnal Ilmiah Teknologi Informasi Terapan (ISSN.2407-3911), Hal: 41-45.
- Park, Park and Kim. 2015. *Kim, Kwang Beak Appendix Analysis from Ultrasonography with Cubic Spline Interpolation and Means Clustering*. International Journal of Bio- Science and Bio-Technology Vol.7, No.1, Hal: 1-10.

- Liu, Fang, Wu, Xu and Dai. 2015. *Improved K-means Algorithm with the Pretreatment of PCA Dimension Reduction*. International Journal of Hybrid Information Technology, Hal: 195-204.
- Pamulaparty, Rao and Rao, 2016. *Cluster Analysis of Medical Research Data using R* ISSN: 0975-4172 & Print ISSN: 0975-4350, Hal: 1-8.
- Fina Nasari and Surya Darma. 2016. *Penerapan K-Means Clustering Pada Data Penerimaan Mahasiswa Baru*. Seminar Nasional Teknologi Informasi dan Multimedia 2015, Hal: 6-8.
- Goldie Gunadi and Indra Sensuse, 2012. *Penerapan Metode Data Mining Market Basket Analysis Terhadap Data Penjualan Produk Buku Dengan Menggunakan Algoritma Apriori Dan Frequent Pattern Growth*. Jurnal Telematika Mkom Vol.4. No.1, Maret 2012 ISSN: 2085-725X, Hal: 118-132.
- Tampubolon, et al, 2013. *Implementasi Data Mining Algoritma Apriori Pada Sistem Persediaan Alat-Alat Kesehatan*. Informasi dan Teknologi Ilmiah (INTI) ISSN: 2339-210X, Hal: 93-106.
- Seragih, at al, 2013. *Implementasi Data Mining Algoritma Apriori Pada Sistem Persediaan Alat-Alat Kesehatan*. Informasi dan Teknologi Ilmiah (INTI) ISSN: 2339-210X, Hal: 93-106.
- Johan Oscar Ong. 2013. *Implementasi Algoritma K-means clustering untuk menentukan strategi marketing president university*. Jurnal Ilmiah Teknik Industri, Vol. 12, No. 1, Juni 2013 ISSN 1412-6869. Hal: 10-20.
- Izzudin A, 2015. *Optimasi Cluster pada Algoritma K-Means dengan Reduksi Dimensi Dataset Menggunakan Principal Component Analysis untuk Pemetaan Kinerja Dosen*. ISSN: 2088-4591 Vol. 5 No. 2 Hal: 41-46.

=====