

PERBANDINGAN TINGKAT PERFORMA METODE *K-MEANS* DAN *HIERACHICAL CLUSTERING* PADA SISTEM REKOMENDASI PEMILIHAN KOST

(COMPARISON OF *K-MEANS* AND *HIERACHICAL CLUSTERING* METHODS PERFORMANCE IN SYSTEM BOARDING COSTS SELECTION RECOMMENDATIONS)

Aiman Ayadi¹⁾, Kusri²⁾, dan Eko Pramono³⁾

^{1, 2, 3)} Magister Teknik Informatika Universitas Amikom Yogyakarta

Jl. Ring Road Utara, Ngringin, Condongcatur, Kec. Depok, Kabupaten Sleman, Daerah Istimewa Yogyakarta

e-mail: aiman.ayadi@students.amikom.ac.id¹⁾, kusri@amikom.ac.id²⁾, eko.p@amikom.ac.id³⁾

ABSTRAK

Perkembangan teknologi saat ini semakin maju dan berkembang semakin cepat, khususnya dibidang teknologi informasi. Keberadaan kost di sekitar tempat fasilitas publik sangat diharapkan dan menguntungkan bagi pendatang di suatu daerah baru. Salah satu fasilitas publik yang biasanya terdapat kost adalah sebuah universitas atau perguruan tinggi. Mahasiswa yang menuntut ilmu di suatu perguruan tinggi tidak hanya berasal dari dalam kota namun juga ada yang berasal dari luar daerah.. Namun kebanyakan dari mahasiswa luar daerah kesulitan dalam memilih kost atau kontrakan dikarenakan keterbatasan informasi. Informasi yang terbatas mengenai fasilitas dan keberadaan kost yang kurang akurat menjadi kesulitan pada awal bagi mereka yang menjadi mahasiswa baru. Dengan adanya kemajuan teknologi informatika dapat menjawab kebutuhan mencari tempat kost, dan akan sangat membantu baik dari sisi pemilik maupun penyewa. Untuk mengatasi permasalahan yang kompleks dapat menggunakan model algoritma *K-Means Clustering* dan *Hierarchical Clustering* yang dioptimalkan dengan *naïve bayes*. Hasil akhir dari penelitian ini adalah nilai akurasi *K-Means* dan *naïve bayes* lebih tinggi dengan akurasi 90,82%, presisi 90,56%, recall 90,68% dan waktu lebih lama yaitu 10 detik, sedangkan untuk nilai *hierachical* dan *naïve bayes* mendapatkan nilai akurasi 88,02%, presisi 87,82%, recall 88,00% dan waktu lebih cepat 7,6 detik.

Kata Kunci: Kost, Mahasiswa, Hierachical, *K-Means*, *Naïve Bayes*.

ABSTRACT

The development of technology is currently getting more advanced and growing faster, especially in the field of information technology. The existence of houses in the vicinity of public facilities is very desirable and beneficial for newcomers to a new area. One of the public facilities that usually have boarding is a university or college. Students who study at a university do not only come from within the city but also come from outside the region. However, most students from outside the region have difficulty choosing boarding or rented houses due to limited information. Limited information about the facilities and the inaccuracy of boarding houses becomes a difficulty at the beginning for those who are new students. With the advancement of information technology, it can answer the need for finding a boarding house, and will be very helpful both from the side of the owner and tenant. To solve complex problems, you can use the *K-Means Clustering* and *Hierarchical Clustering* algorithm models that are optimized with *naïve Bayes*. The final result of this study is that the *K-Means* and *naïve bayes* accuracy values are higher with 90.82% accuracy, 90.56% precision, 90.68% recall and longer time that is 10 seconds, while for *hierachical* and *naïve bayes* values. *Bayes* got 88.02% accuracy, 87.82% precision, 88.00% recall and 7.6 seconds faster time.

Keywords: boarding, students, hierachical, *K-Means*, *Naïve Bayes*.

I. PENDAHULUAN

Perkembangan teknologi saat ini semakin maju dan berkembang semakin cepat, khususnya dibidang teknologi informasi. Pemanfaatannya yang luas dalam berbagai bidang kehidupan yang menyebabkan manusia berusaha membuat sesuatu untuk mempermudah segala aktifitasnya. Hal ini berkaitan dengan penggunaan perangkat

komputer, program aplikasi pendukung, perangkat komunikasi dan internet sebagai sarana pengelolaan informasi. Implementasi komputer sudah meliputi berbagai bidang salah satunya untuk akses mencari informasi pencarian kost.

Keberadaan kost di sekitar tempat fasilitas publik sangat diharapkan dan menguntungkan bagi pendatang di suatu daerah baru. Salah satu fasilitas publik yang biasanya terdapat kost adalah sebuah

Universitas atau perguruan tinggi. Mahasiswa yang menuntut ilmu di suatu perguruan tinggi tidak hanya berasal dari dalam kota namun juga ada yang berasal dari luar daerah. Mahasiswa dari luar daerah yang akan menuntut ilmu di perguruan tinggi mencari tempat tinggal sementara yang berada di dekat tempat kuliah mereka. Para mahasiswa itu ada yang memilih untuk tinggal di rumah kontrakan, tetapi tidak sedikit pula yang lebih memilih untuk tinggal di kost.

Namun kebanyakan dari mahasiswa luar daerah kesulitan dalam memilih kost atau kontrakan dikarenakan keterbatasan informasi. Tidak jarang dari mereka mendapatkan tempat kost tidak sesuai keinginan. Informasi yang terbatas mengenai fasilitas dan keberadaan kost yang kurang akurat menjadi kesulitan pada awal bagi mereka yang menjadi mahasiswa baru. Informasi tentang tempat kost lebih sering beredar lewat mulut ke mulut, sehingga tidak semua mahasiswa mendapat informasi tersebut.

Dengan adanya kemajuan teknologi informatika dapat menjawab kebutuhan mencari tempat kost, dan akan sangat membantu baik dari sisi pemilik maupun penyewa. Dengan adanya teknologi internet, memungkinkan para pemilik tempat kost mempromosikan kostnya lewat media ini dengan biaya yang relatif murah dan dapat dilihat langsung oleh calon penyewa. Dan bukan hanya itu, dengan teknologi internet ini, para calon penyewa juga dapat memilih tempat kost yang sesuai dengan keinginan mereka.

Perencanaan suatu sistem yang dapat membantu dalam penentuan hunian sementara sangat diperlukan oleh mahasiswa dalam mencari hunian yang sesuai dengan kriteria yang diinginkan. Dengan dibuatnya sebuah sistem rekomendasi Pemilihan Tempat Kost diharapkan bisa membantu menyelesaikan masalah alternatif tempat tinggal. Oleh karena itu dibutuhkan program aplikasi sistem rekomendasi untuk menentukan tempat tinggal sementara atau kost. Sistem rekomendasi merupakan suatu sistem berbasis komputer yang bertujuan untuk membantu pengambilan keputusan dengan memanfaatkan data dan model untuk menyelesaikan masalah-masalah yang tidak terstruktur pada data yang kompleks.

Adanya data yang banyak menyebabkan para pengambil keputusan membutuhkan tools yang membantu dalam rekomendasi pencarian kost secara cepat dalam pengambilan keputusan. Untuk menjalankan proses penilaian dengan banyak data, maka diperlukan sebuah sistem yang mampu

melakukan rekomendasi guna meningkatkan efektivitas pengambilan keputusan serta mengurangi subyektivitas dalam proses pengambilan keputusan (Friedyadi. 2016). Ketika keputusan yang akan diambil bersifat kompleks dengan resiko yang besar seperti perumusan kebijakan, pengambilan keputusan memerlukan alat bantu analisis yang bersifat ilmiah, logis, dan terstruktur/konsisten (norhikmah at all, 2013). Salah satu alat analisis tersebut berupa clustering model (model klaster) yang memungkinkan membuat prediksi secara cepat untuk masalah yang bersifat kompleks (Saefulloh at all, 2016).

Untuk mengatasi permasalahan yang kompleks dapat menggunakan model algoritma *K-Means Clustering* seperti penelitian yang telah dilakukan oleh Kusdarnowo dkk (2017) dengan judul “Pencarian Lokasi Perumahan Berdekatan Dengan Fasilitas Kesehatan Dan Belanja Menggunakan Algoritma *K-Means*” dan penelitian agus (2017) dengan judul “Penerapan *Data Mining* Pada Ekspor Buah-Buahan Menurut Negara Tujuan Menggunakan *K-Means Clustering*”. Ataupun dengan Menggunakan model algoritma *Hierarchical Clustering* seperti yang penelitian dilakukan oleh Rizal dan Miftahul (2016) dengan judul “Pencarian Kemiripan Judul Tugas Akhir Mahasiswa Dengan Menggunakan *Metode Single Linkage Hierarchical*” serta Intan at al (2018) dengan judul “Analisis Dan Implementasi Algoritma *Agglomerative Hierarchical Clustering* Untuk Deteksi Komunitas Pada Media Sosial Facebook”.

Sedangkan untuk meningkatkan perbandingan hasil performa yang lebih baik dapat menggunakan perbandingan dua metode seperti saran pada peneltian Rendy et al (2014) dengan peneltian yang berjudul “Perbandingan Metode *Clustering* Menggunakan *Metode Single Linkage dan K - Means* Pada Pengelompokan Dokumen” serta kesimpulan yang temukan oleh Asep (2017) dengan judul peneltian “Analisa Metode *Hierarchical Clustering Dan K-Mean* Dengan Model *Lrfmp* Pada Segmentasi Pelanggan”.

Dalam pengambilan keputusan diperlukana suatu metode klasifidikasi untuk menempatkan objek yang ditugaskan hanya ke salah satu kategori yang disebut keputusan (Han et al, 2012). Salah satu metode klasifikasi yang sering digunakan adalah *Naïve Bayes Classifier* yang pertama kali dikemukakan oleh Revered Thomas Bayes. Penggunaan *Naïve Bayes Classifier* sudah dikenal sejak tahun 1702-1761. *Naïve Bayes Classifier* merupakan pendekatan yang sangat sederhana dan

sangat efektif untuk pelatihan klasifikasi. Sedangkan Kononenko dan Langley menyimpulkan bahwa *Naïve Bayes Classifier* merupakan kemungkinan label kelas data atau bisa diasumsikan sebagai atribut kelas yang diberi label (Kusumadewi, 2003).

Penelitian yang akan dilakukan peneliti adalah melakukan perbandingan model algoritma *K-Means* dan *Hierarchical Clustering* untuk melakukan klustering data serta *naïve bayes* untuk menentukan klasifikasi atau keputusan pada sistem rekomendasi pemilihan kost pada daerah Yogyakarta untuk menghitung nilai performa kedua model algoritma tersebut..

II. STUDI PUSTAKA

Dalam melakukan penelitian perlu diberikan rujukan atau tinjauan pustaka sebagai acuan dasar penelitian. Pada penelitian pertama yang dirujuk adalah penelitian yang dilakukan oleh Fandy et al (2017) dengan judul “Analisis Sentimen untuk Komentar pada Sistem Pencarian Kost Menggunakan Metode *Support Vector Machine (SVM)*”. Pada penelitian ini membahas Analisis Sentimen untuk Komentar pada Sistem Pencarian Kost Menggunakan *Metode Support Vector Machine (SVM)*. Dari pengujian SVM pada 5 tempat kost berbeda dengan menggunakan parameter $C = 2.5$ dan $P = 0.3$ didapat rating tertinggi sebesar 4.615 pada kost Kemiri 2 dan yang terkecil sebesar 3.076 pada kost Turen II dan untuk akurasi didapat pada tempat kost Turen II menghasilkan akurasi 92.30% , Dipo 88 mendapat akurasi 81.81%, Kemiri 2 mendapat 92.30%, Graha Widya mendapat 54.54%, dan Wisma Mawar mendapatkan akurasi 90.90%. Terdapat akurasi yang bernilai kecil yaitu 54.54% dikarenakan sistem masih belum bisa menentukan polaritas sebuah komentar yang berisi perbandingan [1].

Penelitian kedua yang menjadi referensi adalah penelitian yang dilakukan oleh Agus (2017) dengan judul “Penerapan Data Mining Pada Ekspor Buah-Buahan Menurut Negara Tujuan Menggunakan *K-Means Clustering*”. Penelitian ini membahas Penerapan *Data Mining* Pada Ekspor Buah-Buahan Menurut Negara Tujuan Menggunakan *K-Means Clustering*. Untuk melakukan penilaian terhadap hasil ekspor buah-buahan berdasarkan negara tujuan dapat menerapkan metode *clustering K-Means*. Data diolah untuk memperoleh nilai dari produksi ekspor buah-

buahan berdasarkan negara tujuan. Data tersebut diolah menggunakan *Rapidminer* untuk ditentukan nilai centroid dalam 3 cluster yaitu cluster tingkat ekspor tinggi, cluster tingkat ekspor sedang dan cluster tingkat ekspor rendah. Centroid data untuk cluster tingkat ekspor tinggi 904.276,5, Centroid data untuk cluster tingkat ekspor sedang 265.501 dan Centroid data untuk cluster tingkat ekspor rendah 34.280,1 [2].

Penelitian ketiga yang menjadi referensi adalah penelitian yang dilakukan oleh Kusdarnowo et al (2020) dengan judul “Pencarian Lokasi Perumahan Berdekatan Dengan Fasilitas Kesehatan Dan Belanja Menggunakan *Algoritma K-Means*”. Penelitian ini membahas Pencarian Lokasi Perumahan Berdekatan Dengan Fasilitas Kesehatan Dan Belanja Menggunakan *Algoritma K-Means*. *Metode k-means clustering* yang digunakan pada penelitian ini juga membantu dalam pengelompokan perumahan yang berdekatan dengan supermarket dan rumah sakit, Untuk merancang aplikasi pencarian lokasi perumahan disertai informasi supermarket dan rumah sakit terdekat berbasis web dapat dilakukan menggunakan metode pengembangan sistem yaitu Model Prototipe. Desain sistem menggunakan *flowmap* dan *Unified Modeling Language (UML)*, dan pengujian aplikasi menggunakan pengujian *blackbox* [3].

Penelitian keempat yang menjadi referensi adalah penelitian yang dilakukan oleh Rizal dan Miftahul (2018) dengan judul “Pencarian Kemiripan Judul Tugas Akhir Mahasiswa Dengan Menggunakan *Metode Single Linkage Hierarchical*”. Penelitian membahas Pencarian kemiripan judul tugas akhir mahasiswa dengan menggunakan metode *Single Linkage Hierarchical*. Hasil dari percobaan 60 data uji abstrak persentase keberhasilan kecocokan pada kategori multimedia sebesar 100%. pengujian kategori pengolahan citra sebesar 87,5%. Pengujian kategori pengenalan pola sebesar 11,11%. Pengujian pada kategori jaringan dan data mining tidak ditemukan kecocokan. Dan pada halaman user, hasil berupa judul tugas akhir yang ada pada database berdasarkan dengan tema judul tugas akhir yang di-inputkan oleh user [4].

Penelitian kelima yang menjadi referensi adalah penelitian yang dilakukan oleh Rendy et all (2016) dengan judul “Perbandingan Metode *Clustering* Menggunakan *Metode Single Linkage Dan K - Means* Pada Pengelompokan Dokumen”. Penelitian membahas Perbandingan Metode *Clustering* Menggunakan *Metode Single Linkage dan K*

- *Means* Pada Pengelompokan Dokumen. Metode pengujian yang digunakan untuk mengukur kualitas cluster adalah *Silhouette Coefficient* dan *Purity*. Berdasarkan hasil pengujian yang dilakukan, dapat disimpulkan, bahwa metode *Single Linkage* memiliki performansi yang lebih baik dibandingkan dengan metode *K-means*. Nilai *Silhouette Coefficient Single Linkage* selalu lebih unggul dibandingkan dengan *K-Means*. Pertambahan jumlah dokumen membuat nilai *Silhouette Coefficient single linkage* semakin kecil sedangkan *K-means* terkadang menghasilkan nilai yang negatif. Untuk nilai *Purity*, *Single Linkage* selalu bernilai 1 sedangkan *K-Means* tidak pernah bernilai 1. Hasil pertambahan jumlah cluster dan jumlah dokumen memberikan pengaruh terhadap nilai *Silhouette Coefficient* dan *Purity*. Hal ini berarti *single linkage* selalu menghasilkan dokumen yang sama, sedangkan *K-means* masih bercampur dengan dokumen yang lain [5].

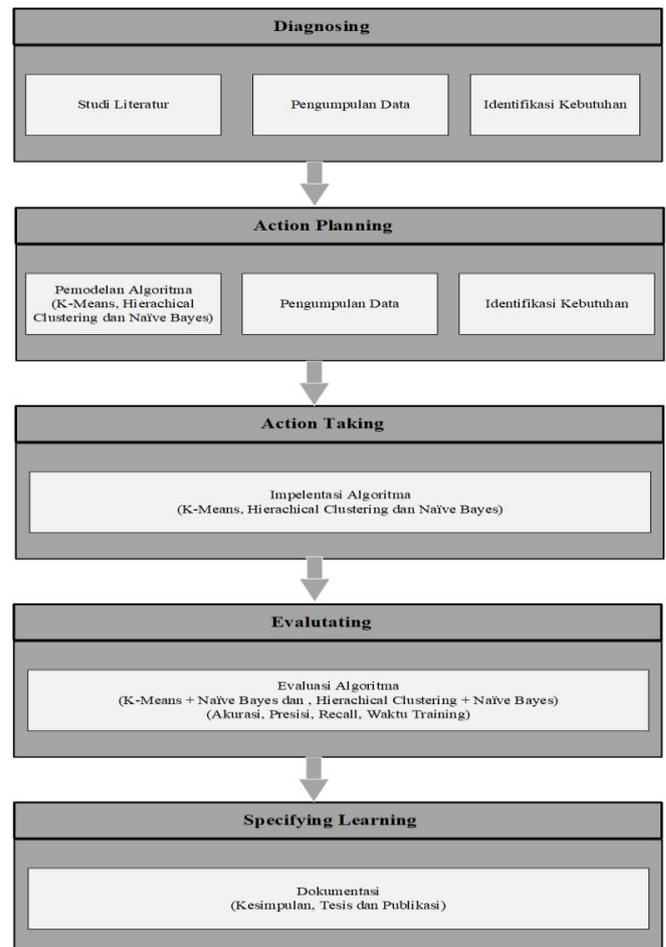
Penelitian keenam yang menjadi referensi adalah penelitian yang dilakukan oleh Intan (2018) dengan judul “Analisis Dan Implementasi *Algoritma Agglomerative Hierarchical Clustering* Untuk Deteksi Komunitas Pada Media Sosial Facebook”. Penelitian membahas analisis dan implementasi *Algoritma Agglomerative Hierarchical Clustering* untuk deteksi komunitas pada media sosial facebook. Pembentukan komunitas menggunakan algoritma *Agglomerative Hierarchical Clustering (AHC)* berbasis *Enhanced Similarity* pada data Facebook, yang mana merupakan algoritma yang secara berturut-turut memodelkan pengelompokan yang menggunakan nilai threshold sebagai alat pengelompokkan, dua simpul yang memiliki nilai similarity lebih besar dibandingkan nilai threshold maka akan dijadikan satu komunitas dengan memperhitungkan hubungan yang dimiliki simpul. Dalam pengujiannya, peneliti menggunakan *modularity* sebagai alat penentu kualitas cluster, yang mana kualitas cluster dikatakan baik jika memiliki hubungan yang pada di dalam satu komunitasnya dibandingkan hubungan di luar komunitasnya. Hasil dari Algoritma AHC mampu menunjukkan bahwa AHC dapat mendeteksi komunitas dengan kualitas cluster (*modularity*) mencapai 0.4453 pada jaringan nyata [6].

Penelitian ketujuh yang menjadi referensi adalah penelitian yang dilakukan oleh Asep (2017) dengan judul “Analisa Metode *Hierarchical Clustering* dan *K-Mean* Dengan Model *Lrfmp* Pada Segmentasi Pelanggan”. Penelitian membahas

analisa metode *Hierarchical Clustering* dan *K-Mean* dengan model *Lrfmp* pada segmentasi pelanggan. Metode *Hierarchical Clustering Single* adalah metode yang terbaik untuk penentuan nilai K awal I dibandingkan dengan metode *Single Ward*, *Complete Linkage*, *Average Linkage* dan *Centroid*. Penambahan parameter P (payment) mampu membuat nilai DBI lebih baik. Penambahan bobot dapat meningkatkan nilai DBI cluster. Nilai rasio akurasi model LRFM lebih baik dari model LRFMP. Jumlah kluster (K) hasil metode *Hierarchical Clustering*, bisa diterapkan untuk nilai K awal k-mean. Tetapi dengan jumlah kluster yang sama, DBI metode *single linkage* lebih baik dari pada metode k-mean [7].

III. METODE PENELITIAN

Pada bagian ini berisi diagram alur langkah penelitian secara lengkap dan terinci termasuk di dalamnya mencakup algoritma, rute, pemodelan-pemodelan, desain yang terkait dengan perancangan sistem. Alur penelitian pada sistem pendukung keputusan rekomendasi pencarian kost ditunjukkan pada Gambar 1 di bawah ini.



Gambar 1. Metode Penelitian

Adapun secara garis besar urutan penelitian adalah sebagai berikut:

A. Diagnosing

Pada tahap ini peneliti melakukan studi literatur dengan membaca buku, jurnal, makalah dan laporan penelitian yang terkait dengan topik penelitian. Kemudian melakukan pengumpulan data dan informasi seperti melakukan wawancara, dokumentasi dan observasi untuk mengambil kebutuhan data. Selanjutnya dilakukan proses definisi kebutuhan dengan melakukan identifikasi data yang dibutuhkan, melihat prosedur yang sedang berjalan, menganalisis sistem yang sedang berjalan serta membuat hasil evaluasi sistem tersebut.

B. Action Planning

Pada tahap ini berisikan proses desain sistem dan software seperti pemodelan metode *K-Means Clustering* dan *Hierarchical Clustering* sebagai metode untuk menentukan kelas atau klastering serta metode *Naive Bayes* untuk pengambilan keputusan, pemodelan proses dengan *Unified Modeling Language (UML)*, pemodelan data dengan *Entity Relationship Diagram (ERD)* serta perancangan *User Interface*.

C. Action Taking

Pada tahap ini dilakukannya implementasi kedua algoritma yaitu *K-Means Clustering* dan *Hierarchical Clustering* untuk klastering pada dataset serta *Naive Bayes* untuk rekomendasi pencarian kost.

D. Evaluating

Pada tahap ini dilakukan perbandingan antara dua algoritma untuk didapatkan hasil penelitian dengan pengukuran performa akurasi, presisi, recall dan waktu training berupa nilai angka, tabel dan diagram.

E. Specifying Learning

Pada tahap ini dilakukan proses dokumentasi dan publikasi thesis berisi hasil penelitian yang telah dilakukan.

Naive Bayes memiliki presisi yang lebih baik, selanjutnya untuk nilai rata-rata *recall* algoritma *K-Means + Naive Bayes* juga memiliki nilai *recall* yang lebih baik akan tetapi Dalam hal eksekusi waktu training model algoritma algoritma *Hierarchical + Naive Bayes* lebih unggul karena memiliki nilai angka lebih cepat. Penggabungan *Hierarchical + Naive Bayes* terlihat lebih unggul dikarenakan pada saat penentuan bobot probabilitas *Hierarchical + Naive Bayes*, nilai probabilitas tersebut dioptimasi atau ditambahkan dengan algoritma Min Max sehingga nilai bobot probabilitas baru tersebut lebih optimal jika dibandingkan dengan menggunakan *Hierarchical + Naive Bayes* saja akan tetapi proses waktu training menjadi lebih lama, hal ini berbanding terbalik dengan penerapan algoritma tunggal *Hierarchical + Naive Bayes* saja yang memiliki pemrosesan waktu training lebih cepat akan tetapi memiliki nilai akurasi yang lebih lama. Hasil perbandingan performa dapat dilihat pada Tabel 1 sampai Tabel 4 dan Gambar 2.

Tabel 1. Performa Akurasi.

Algoritma	Akurasi (%)
Hierarchical + Naive Bayes	88,02
K-Means + Naive Bayes	90,82

Tabel 2. Performa Presisi.

Algoritma	Presisi (%)
Hierarchical + Naive Bayes	87,82
K-Means + Naive Bayes	90,56

Tabel 3. Performa Recall.

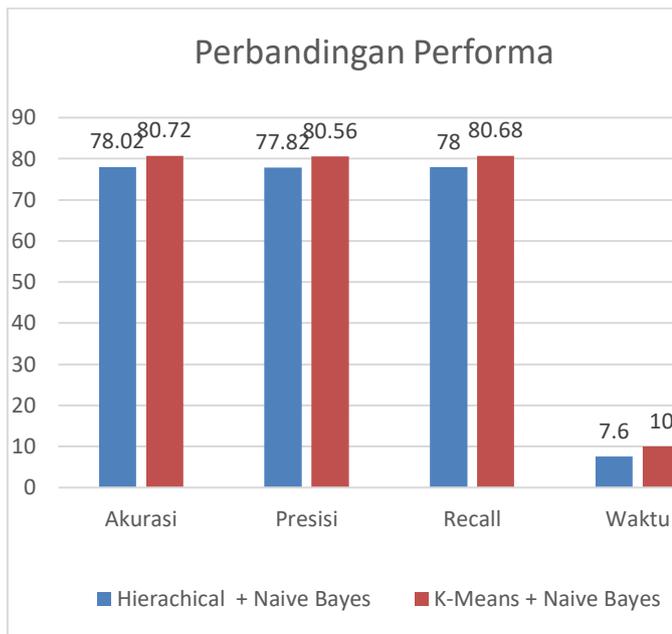
Algoritma	Recall (%)
Hierarchical + Naive Bayes	88,00
K-Means + Naive Bayes	90,68

Tabel 3. Performa Waktu.

Algoritma	Waktu (detik)
Hierarchical + Naive Bayes	7,6
K-Means + Naive Bayes	10

IV. HASIL DAN PEMBAHASAN

Kesimpulan hasil perbandingan tingkat performa pada penelitian ini adalah model algoritma *K-Means + Naive Bayes* memiliki nilai rata-rata akurasi yang tertinggi, kemudian nilai rata-rata presisi model algoritma algoritma *K-Means +*



Gambar 2 Perbandingan Hasil Performa

V. KESIMPULAN

Setelah melalui tahap pengujian tingkat performa dengan parameter akurasi, presisi, *recall* dan Waktu Training pada model *algoritma Hierachical + Naive Bayes* serta *Hierachical + Naive Bayes* dan Optimalisasi untuk prediksi mahasiswa pada rekomendasi pemilihan kost, maka dapat diambil beberapa kesimpulan antara lain:

1. Model *algoritma Hierachical + Naive Bayes* memiliki nilai tingkat performa akurasi dengan nilai 88,02%, performa presisi dengan nilai 87,82%, performa *recall* dengan nilai 88% serta performa waktu training dengan nilai 7,6 detik. Sedangkan pada model *algoritma Hierachical + Naive Bayes* dan Optimalisasi memiliki nilai tingkat performa akurasi dengan nilai 90,82%, performa presisi dengan nilai 90,56%, performa *recall* dengan nilai 90,68% serta performa waktu training dengan nilai 10 detik.
2. Pengaruh jumlah *records* terhadap model algoritma yang digunakan memiliki pengaruh baik terhadap nilai akurasi, presisi dan *recall* dimana nilai akurasi, presisi dan *recall* akan meningkat jika jumlah *records* semakin bertambah, akan tetapi berdampak buruk pada kecepatan waktu training dimana waktu training akan melambat jika jumlah *records* semakin bertambah.

DAFTAR PUSTAKA

- [1] Fandy. O. S., Banu. W. Y., Saptadi. N., (2017), *Analisis Sentimen untuk Komentar pada Sistem Pencarian Kost Menggunakan Metode Support Vector Machine (SVM)*, Techno Jurnal Ilmiah Elektroteknika Vol. 16 No. 1 April 2017 Hal 41 – 47.
- [2] Agus. P. W., (2017), *Penerapan Data Mining Pada Ekspor Buah-Buahan Menurut Negara Tujuan Menggunakan K-Means Clustering*, Techno.COM, Vol. 16, No. 4, November 2017 : 348-3577.
- [3] Kusdarnowo. H., Andi. C., Siti. A., (2020): *Pencarian Lokasi Perumahan Berdekatan Dengan Fasilitas Kesehatan Dan Belanja Menggunakan Algoritma K-Means*, Jurnal IKRA-ITH Informatika Vol 4 No 1, Maret 2020.
- [4] Rizal. T. A, Miftahul. J., (2018): *Pencarian Kemiripan Judul Tugas Akhir Mahasiswa Dengan Menggunakan Metode Single Linkage Hierarchical*, urnal SAINTEKOM Volume 8 Nomor 1, Maret 2018.
- [5] R Handoyo, R Mangkudjaja, SM Nasution (2014), *Perbandingan Metode Clustering Menggunakan Metode Single Linkage Dan K - Means Pada Pengelompokan Dokumen*, Jurnal Sifo Mikroskil, Vol 15, No 2, Oktober 2014.
- [6] Intan. W. H., Imelda. A., Anisa H., (2018). *Analisis Dan Implementasi Algoritma Agglomerative Hierarchical Clustering Untuk Deteksi Komunitas Pada Media Sosial Facebook*. e-Proceeding of Engineering : Vol.5, No.1 Maret 2018 | Page 1460.
- [7] Asep. M., (2017), " *Analisa Metode Hierarchical Clustering Dan K-Mean Dengan Model Lrfmp Pada Segmentasi Pelanggan*. SIGMA Jurnal Teknologi Pelita Bangsa Volume 7 Nomor 1, September 2017 ISSN : 2407-3903.