

BAB I

PENDAHULUAN

1.1. Latar Belakang

Aktivitas *search* sangat bervariasi dan berkembang dari hari ke hari sehingga fitur *search* terus dikembangkan untuk menambah daya tarik pengguna. Proses *search query* terutama pada *e-commerce* bergantung kepada konteks waktu dan tempat dengan dimotivasi oleh rasa keingintahuan, *time killing*, dan pembelajaran (Dan Wu, Man Zhu, dan Aihua Ran, 2016). *Query* pengguna terus mengalami perubahan sesuai dengan tren dan informasi dari hasil *query* (Lin YC., Datta, A., dan Fabrizio, G.D., 2018). Perubahan mengharuskan optimisasi lebih mendalam terhadap fitur *search* sehingga hasil *query* mampu mencerminkan perilaku dari penggunaannya, seperti *query* yang sudah diketahui targetnya, *query* untuk tujuan pengambilan keputusan, ataupun *query* untuk tujuan eksplorasi saja (Sondhi, P. et al., 2018 dan Su, N. et al., 2018). Preferensi pengguna *e-commerce* atas *query* yang dilakukan sangat bervariasi, tetapi ketika preferensi tersebut dapat terealisasi melalui *query* maka implikasinya adalah minat beli (Guo, Y. et al., 2018). Informasi seperti atribut-atribut dari hasil *query* yang disimpan di *e-commerce* dibentuk menjadi *cluster* untuk menghasilkan model dasar suatu pengetahuan bisnis (Singh, S. et al., 2018 dan Lin, YC., Datta, A., dan Fabrizio, G.D., 2018).

Algoritma *k-means* menghasilkan taksonomi untuk menjelaskan proses pengambilan keputusan dalam melakukan transaksi pembelian di *e-commerce*. Taksonomi tersebut menjadi dasar peningkatan fitur *search* yang kategorinya terdiri dari *shallow exploration queries*, *targeted purchase queries*, *major-item shopping queries*, *minor-item shopping queries*, dan *hard-choice shopping queries* (Sondhi, P. et al., 2018). Pengguna melakukan *search query* tidak hanya berdasarkan nama produk, tetapi juga kategori dari produk sehingga *cluster* dari klasifikasi *query* harus mengutamakan sifat informatif pada keluarannya (Kathuria, A. et al., 2010 dan Lin, YC., Datta, A., dan Fabrizio, G.D., 2018).

Berbeda dengan *k-means*, *adaptive DBSCAN* bekerja berdasarkan tingkat kepadatan (*density*) data dan tidak perlu mendefinisikan nilai Eps dan nilai MinPts (Khan, M.M.R. et al., 2018). *Adaptive DBSCAN* juga memiliki kemampuan dalam menghasilkan *cluster* dari bentuk yang sembarang tetapi yang diteliti peneliti sebelumnya hanya dalam konteks analisis waktu pemrosesan (Chakraborty, S., Dey, L., dan Nagwani, N.K., 2011, Kanagala, H.K. dan Krishnaiah, J.R., 2016). Berdasarkan uraian di atas, maka peneliti termotivasi untuk melakukan komparasi *search query clustering* pada *e-commerce* menggunakan *k-Means* dan *Adaptive DBSCAN*.

1.2. Masalah Penelitian

Masalah penelitian yang akan dibahas dari penulisan tesis ini dibagi menjadi 2 (dua) yaitu identifikasi dan rumusan masalah.

1.2.1. Identifikasi Masalah

Berikut ini akan dilakukan identifikasi masalah berdasarkan latar belakang yang telah diuraikan di atas yaitu :

1. Hasil *clustering* dari *search query* selain dikelompokkan menjadi taksonomi, juga dapat dimanfaatkan sebagai model dasar pengetahuan bisnis *e-commerce* dalam meningkatkan fitur *search*.
2. Algoritma *k-means* menyelesaikan masalah *clustering* terhadap *search query* tetapi belum pernah diselesaikan menggunakan *adaptive DBSCAN*.
3. Belum ada penelitian sebelumnya yang membandingkan keakuratan proses *clustering* menggunakan algoritma *k-means* dan *adaptive DBSCAN*.

1.2.2. Rumusan Masalah

Adapun rumusan masalah dalam penelitian tesis ini adalah bagaimana akurasi dari proses *search query clustering* menggunakan algoritma *k-means* dibandingkan dengan *adaptive DBSCAN*.

1.3. Tujuan dan Manfaat Penelitian

Tujuan dari penelitian adalah untuk membandingkan keakuratan dari hasil *search query cluster* pada *e-commerce* menggunakan *k-means* dan *adaptive DBSCAN*.

Manfaat dari penelitian adalah :

1. Menjadi dasar dari peningkatan fitur *search* sehingga hasil *query* mampu mencerminkan perilaku penggunaannya.
2. Memberikan model dasar pengetahuan bisnis bagi pengguna *e-commerce*.
3. Membandingkan tingkat keakuratan dari algoritma *k-means* dan *adaptive DBSCAN* dalam menghasilkan *search query cluster*.

1.4. Ruang Lingkup

Berikut akan dijelaskan ruang lingkup penelitian dalam penulisan tesis :

1. *Dataset* yang digunakan dalam penelitian diambil dari PT Royal Utama Makmur dari tanggal 1 Juli 2018 s.d. 30 Juni 2019.
2. *Fields dataset* yang digunakan dalam penelitian adalah *browse time (minutes)*, *page view time (minutes)*, *add-to-cart*, *order*, *direct order*, *search depth*, *click per query*, dan *token*.
3. *Dataset* yang digunakan adalah *search query* yang telah memiliki transaksi.
4. Banyaknya *cluster* yang digunakan dalam komparasi adalah 5 (lima), sesuai dengan banyaknya kategori taksonomi.
5. Pengujian hasil komparasi *k-means* dan *adaptive DBSCAN* menggunakan *confusion matrix*.

1.5. Metodologi Penelitian

Adapun langkah-langkah dalam penelitian ini adalah sebagai berikut :

1. Studi literatur

Pada tahap ini dilakukan pendalaman mengenai bagaimana cara kerja algoritma *k-means* dan *adaptive DBSCAN* dalam menghasilkan *search query cluster*.

2. Analisis

Pada tahap ini dilakukan proses untuk mengidentifikasi data yang dibutuhkan, masalah dan tantangan yang harus diselesaikan dan menjelaskan solusi yang diusulkan untuk menyelesaikan masalah dan tantangan yang ada dengan menggunakan *flowchart*.

3. Implementasi

Implementasi dimulai dengan membuat program menggunakan bahasa pemrograman Java untuk membandingkan algoritma *k-means* dan *adaptive DBSCAN* dalam menghasilkan *search query cluster*.

4. Pengujian

Melakukan pengujian keakuratan implementasi menggunakan *confusion matrix*.

5. Evaluasi

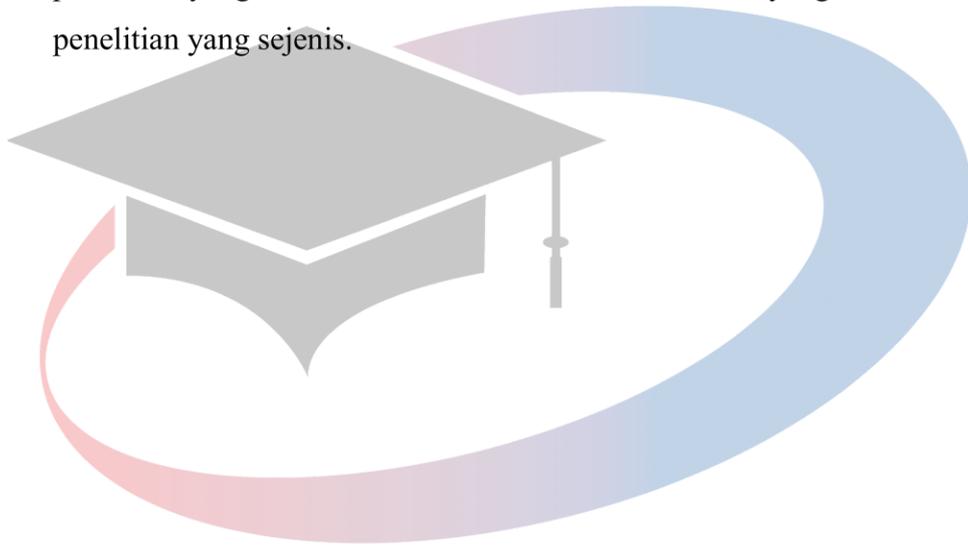
Pada tahap ini, dilakukan evaluasi terhadap hasil pengujian yang sudah dilakukan untuk mengambil kesimpulan dan saran.

1.6. Sistematika Penulisan

Sistematika penulisan laporan penelitian ini terdiri dari 5 bab, dimana secara garis besar masing-masing bab membahas hal-hal berikut :

1. Bab 1 Pendahuluan, berisi penjelasan umum, masalah dan solusi yang akan dilakukan penelitian.
2. Bab 2 Kajian Literatur, berisi tinjauan pustaka dan kerangka konsep terkait masalah dan metode yang berhubungan dengan penelitian yang akan dilakukan.

3. Bab 3 Metodologi Penelitian, berisi identifikasi masalah, langkah-langkah dari metode yang diusulkan, data yang digunakan, alat-alat penelitian dan teknik analisis.
4. Bab 4 Hasil dan Pembahasan, berisi hasil dari sistem yang ditingkatkan dan berisi hasil pengujian.
5. Bab 5 Penutup, berisi kesimpulan yang diperoleh dari hasil dan pengujian penelitian yang sudah dilakukan dan saran untuk hasil yang lebih baik dalam penelitian yang sejenis.



UNIVERSITAS MIKROSKIL