

PENENTUAN TRENDING TOPIC MENGUNAKAN PENDEKATAN FUZZY LOGIC

UTS PENCARIAN INFORMASI MEDIA ONLINE

**Diajukan untuk Menyelesaikan UTS Pencarian Informasi Media
Online**

Oleh

Ryan Nathaniel (1872019)

Kelvin William (1872020)

Edwin Desemsky (1872051)

Salim Arizi (1872054)



PROGRAM STUDI S1 TEKNIK INFORMATIKA

FAKULTAS TEKNOLOGI INFORMASI

UNIVERSITAS KRISTEN MARANATHA

BANDUNG

2021

Commented [RT1]: Diganti dengan tahun Tugas Akhir/
Kerja Praktik dibuat dan dikumpulkan

DAFTAR PUSTAKA

BAB 1 PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	1
1.3 Tujuan Pembahasan	2
1.4 Ruang Lingkup.....	2
1.5 Sumber Data.....	2
1.6 Sistematika Penyajian	2
BAB 2 KAJIAN TEORI	4
2.1 <i>Data Streaming</i>	4
2.2 <i>Clustering</i>	4
2.3 <i>Fuzzy Logic</i>	4
BAB 3 Analisis dan desain	5
3.1 Deskripsi Sistem	5
3.2 Langkah Kerja.....	5
3.2.1 <i>Data Streaming</i>	5
3.2.2 <i>Clustering by Topic</i>	6
3.2.3 <i>Determining Trending Topic (Most Searches)</i>	6
3.2.3.1 Penerapan <i>Fuzzy Logic</i>	6
3.2.3.2 <i>Time Schedule</i>	6
3.3 Desain UML.....	7
3.3.1 <i>Use Case</i>	7
3.3.2 <i>Use Case Description</i>	7
3.3.2.1 <i>Use Case Description Melihat Trending Topic</i>	7
3.3.2.2 <i>Use Case Description Mencari Topik</i>	8

3.3.3 Activity Diagram	8
3.3.3.1 Activity Diagram Melihat Trending Topic.....	8
3.3.3.2 Activity Diagram Mencari Topik	9
3.3.4 GUI (<i>Graphical User Interface</i>)	9
3.3.5 ERD (<i>Entity Relationship Diagram</i>).....	11
DAFTAR PUSTAKA	1

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Di era yang serba cepat dan instan ini, segala sesuatu terasa terjadi dengan hanya sekejap mata, hal ini juga didukung dengan perkembangan teknologi yang semakin canggih. Orang-orang dapat membicarakan sesuatu terhadap sesamanya secara praktis melalui gawai, media menyebarkan informasi dalam jumlah yang besar dengan waktu yang singkat. Perbincangan satu topik akan beralih dan berganti ke topik perbincangan lainnya secara cepat dan tak terkendali. Hal ini tentunya mengakibatkan penyebaran informasi dan keakuratan informasi menjadi tidak efektif dan efisien.

Untuk mengatasi hal tersebut, penelitian ini akan berfokus pada pembelajaran dan penerapan *trending topic*. *Trending topic* sendiri ialah parameter kondisi di mana suatu topik perbincangan sedang banyak diperbincangkan atau dipertanyakan oleh banyak orang dalam kurun waktu tertentu. *Trending topic* dapat memberikan manfaat dalam hal pengaksesan informasi yang lebih tertata dan akurat terhadap suatu topik tertentu, di mana *trending topic* hanya akan mengumpulkan informasi-informasi yang hanya berkaitan dengan suatu topik tertentu yang juga berdampak positif pada keakuratan informasi tersebut. *Trending topic* juga dapat mendorong sekumpulan orang untuk membicarakan hal tertentu.

Penerapan *trending topic* ini akan dilakukan kepada Philoit.id sebagai subjek daripada penelitian ini. Dengan diterapkannya *trending topic* pada Philoit.id maka hal-hal yang telah disebutkan sebelumnya tidak akan terjadi mengingat bahwa Philoit.id belum memiliki fitur *trending topic* sampai penelitian ini selesai. *Trending topic* yang diterapkan pada Philoit.id ini bertujuan untuk memberikan pengalaman baru yang dapat bermanfaat bagi penggunaanya.

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah yang telah dibahas, maka dapat diambil kesimpulan rumusan masalah yaitu :

1. Bagaimana cara menentukan trending topic pada setiap topik yang terdapat pada Philoit?
2. Bagaimana algoritma yang efektif dan efisien untuk mencari trending topic?

1.3 Tujuan Pembahasan

Tujuan pembahasan dari penelitian ini adalah untuk menentukan *trending topic* yang terdapat pada Philoit.id berdasarkan topik yang sering dibicarakan dan dicari oleh pengguna Philoit.id.

1.4 Ruang Lingkup

Ruang lingkup yang digunakan pada pembuatan laporan ini adalah sebagai berikut:

1. Membangun *framework* yang dapat menentukan *trending topic* berdasarkan topik yang sering dibicarakan dan dicari.
2. Mengimplementasikan *framework* ke dalam Philoit.id untuk memberikan manfaat kepada pengguna dalam hal mendapatkan informasi.

1.5 Sumber Data

Sumber data yang digunakan dalam penelitian ini adalah sebagai berikut:

1. Sumber data primer

Data diperoleh melalui data-data yang telah tersimpan dalam basis data Philoit.id terkait pengguna, topik, dan hal lainnya yang berhubungan dengan penerapan *trending topic*.

2. Sumber data sekunder

Sumber data berasal buku-buku teks, situs internet, dan referensi jurnal atau artikel ilmiah yang memuat materi yang memiliki hubungan dengan penelitian.

1.6 Sistematika Penyajian

Sistematika penyajian ini dituliskan dengan tujuan agar alur pembahasan laporan ini dapat lebih terarah dan terstruktur. Sistematika penyajian ini direncanakan sebagai berikut:

BAB 1 PENDAHULUAN

Pada bagian ini akan membahas tentang latar belakang, rumusan masalah, tujuan pembahasan, ruang lingkup, sumber data, dan sistematika penyajian dari penelitian ini.

BAB 2 PENGKAJIAN TEORI

Pada bagian ini akan membahas tentang teori-teori yang akan diterapkan dalam proses penyelesaian penelitian ini.

BAB 3 ANALISIS DAN RANCANGAN SISTEM

Pada bagian ini akan menguraikan analisis dan desain sistem yang akan dibangun pada penelitian ini.

BAB 4 IMPLEMENTASI

Pada bagian ini akan membahas tentang implementasi *framework* yang akan digunakan dalam proses konversi teks dalam rekaman pembelajaran.

BAB 5 PENGUJIAN

Pada bagian ini akan berisi uraian hasil pengujian terhadap *framework* yang diimplementasi dan evaluasi dari sistem yang telah dibangun.

BAB 6 KESIMPULAN DAN SARAN

Pada bagian ini akan berisi tentang kesimpulan dan saran dari seluruh rangkaian penelitian yang telah dilakukan.

BAB 2

KAJIAN TEORI

2.1 Data Streaming

Data Streaming merupakan suatu metode yang digunakan untuk memantau aliran data dalam suatu sistem. *Data Streaming* harus memenuhi beberapa batasan yang terkait dengan respons *real-time*, memori yang terbatas, *single-pass*, dan pendeteksian concept-drift. Masalah utama dari hal ini ialah mengidentifikasi pola dan pengetahuan tersembunyi untuk memahami konteks untuk mengidentifikasi tren dan aliran data yang berlanjut dalam sistem [1].

2.2 Clustering

Clustering adalah suatu proses pengelompokan data ke dalam beberapa kelompok di mana data yang berada dalam satu kelompok cenderung memiliki tingkat kemiripan yang tinggi, sedangkan yang tidak berada dalam satu kelompok cenderung memiliki tingkat kemiripan yang rendah. Untuk meningkatkan kualitas clustering dapat menggunakan model topik dan fusi. Pertama, metode clustering diaplikasikan ke dalam sebuah set dokumen, lalu topik diekstraksi dari masing-masing gugus. Metode ini terdiri dari empat langkah. Langkah pertama, dokumen serupa dikelompokkan dengan metode clustering. Langkah kedua, algoritma LDA diterapkan ke setiap kelompok untuk mengidentifikasikan topik kelompok. Langkah ketiga, frekuensi global berasal dari topik yang diekstraksi dan kata-kata yang mirip secara semantik. Pada langkah terakhir, kalimat yang mengandung istilah dan kata-kata yang mirip akan dipilih dan diringkaskan [2].

2.3 Fuzzy Logic

Fuzzy Logic merupakan salah satu algoritma yang dapat menentukan persentase tingkat kemiripan satu hal terhadap hal yang lainnya. Pada penerapan *trending topic*, inferensi *fuzzy* dan pengukuran *fuzzy* digunakan untuk menemukan kalimat paling signifikan [3].

BAB 3

ANALISIS DAN DESAIN

3.1 Deskripsi Sistem

Trending topic merupakan suatu subjek yang mengalami lonjakan popularitas dalam kurun waktu tertentu. Dalam menentukan subjek mana saja yang dapat menjadi *trending* maka diperlukan kumpulan data dari data Q&A dan *data writing* dari seluruh pengguna philoit.id. Data tersebut akan diolah dengan memanfaatkan *clustering* berdasarkan topik yang kemudian tiap kata kunci dari data akan dipilih dan ditentukan untuk menjadi *trending*.

3.2 Langkah Kerja

Adapun langkah kerja yang akan dilakukan dalam penelitian ini adalah sebagai berikut :

- *Data Streaming*
- *Clustering by Topic*
- *Determining Trending Topic (Ranking and Counting Comments and Upvotes)*
- *Determining Trending Topic (Most Searches)*

3.2.1 Data Streaming

- Untuk bisa menentukan *trending topic* secara aktual dan akurat, kita perlu data sebagai salah satu kebutuhan.
- Data yang diperoleh dan telah tersimpan ke dalam *database* akan diawasi dan dianalisis untuk mendapatkan hasil yang diinginkan.
- Data yang ada di Philoit.id akan diawasi dan dianalisis dengan menggunakan pendekatan algoritma *query* (mengingat data tersimpan dalam bentuk basis data). Namun, tidak menutup kemungkinan untuk menggunakan metode pendekatan lainnya untuk mendapatkan hasil serupa.
- *Query* sederhana yang digunakan adalah perintah *SELECT*, untuk menampilkan data sesuai dengan kebutuhan

3.2.2 Clustering by Topic

- Setelah data-data tersebut selesai dianalisis, maka selanjutnya kita akan melakukan pengelompokan data atau *data clustering*. dengan begitu, proses penentuan *trending topic* tidak akan terlalu sulit sebat data-data tersebut telah dikelompokkan ke dalam masing-masing topik dan tidak bersifat rancu.
- Pengelompokan atau *clustering* sendiri dapat ditentukan dari banyak faktor, karakteristik data akan menjadi salah satu ciri khas dalam penentuan tersebut, contoh daripada karakteristik tersebut bisa berupa kesamaan kata dalam data atau mungkin hal lainnya.
- Dalam *clustering* ini sendiri, kita dapat menggunakan perintah *GROUP_BY* untuk pengelompokan datanya sendiri.

3.2.3 Determining Trending Topic (Most Searches)

- Selain menggunakan *comments* dan *upvotes*, kita juga bisa menentukan *trending* topik berdasarkan kata kunci yang paling sering dicari oleh pengguna.
- Selanjutnya, kata kunci ini akan kita sesuaikan (mencari kesamaan) dengan topik, hal ini akan dilakukan dengan menggunakan metode *FUZZY LOGIC* yang mana kita mencari persentase kemiripan kata kunci tersebut.
- Setelah menentukan kata kunci yang paling banyak dicari sesuai dengan topik yang ada maka topik tersebut juga dapat disebut sebagai *trending topic*.

3.2.3.1 Penerapan Fuzzy Logic

Fuzzy Logic merupakan pendekatan untuk memperoleh tingkat persentase antara satu hal dengan hal yang lain. dalam penelitian ini pendekatan ini berarti untuk memperoleh persentase antara satu kata kunci dengan kata kunci yang lain dan menentukan seberapa besar persentasenya yang mana jika persentasenya tinggi maka akan dapat diasumsikan bahwa topik tersebut memang sedang dicari-cari oleh sejumlah pengguna.

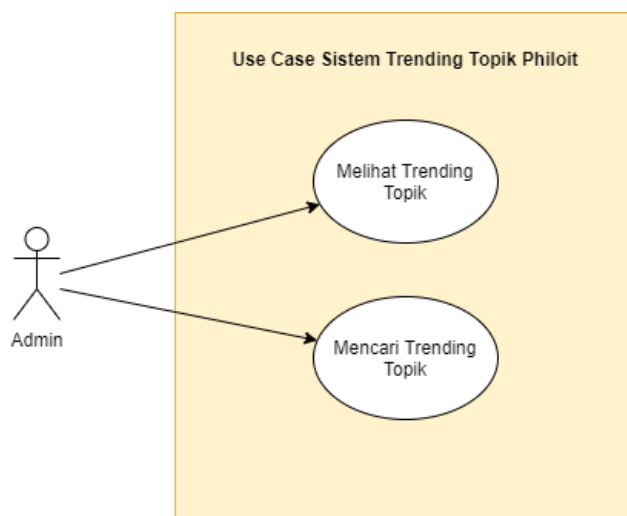
3.2.3.2 Time Schedule

Week	Execution
5-6	Melakukan analisis data dengan menggunakan <i>query</i> .
7-8	Melakukan pemodelan dengan <i>Fuzzy Logic</i> .

9-12	Melakukan <i>training data</i>
13-14	Melakukan <i>clustering</i> atau pengelompokan data
15-16	Membuat laporan hasil pekerjaan

3.3 Desain UML

3.3.1 Use Case



3.3.2 Use Case Description

3.3.2.1 Use Case Description Melihat Trending Topic

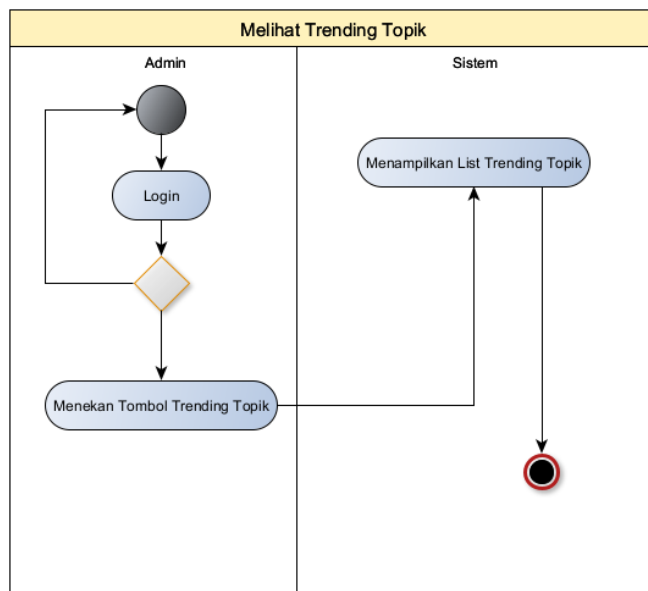
Name	Melihat Trending Topic
Actor	Admin
Description	Proses ini merupakan proses melihat <i>trending topic</i> yang dilakukan oleh pengguna
Input Condition	Pengguna login kedalam philoit sebagai member dan menekan tombol trending
Output Condition	Sistem akan menampilkan daftar topik yang sedang trending pada saat ini.

3.3.2.2 Use Case Description Mencari Topik

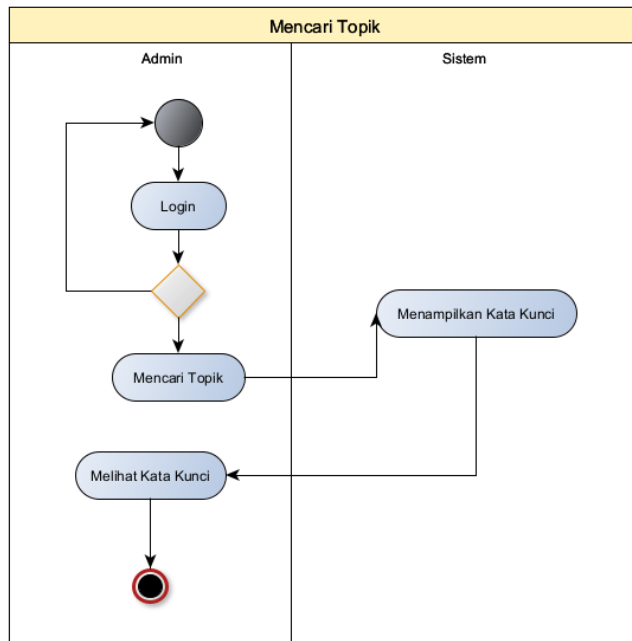
Name	Mencari Topik
Actor	Admin
Description	Proses ini merupakan proses mencari trending topik yang dilakukan oleh pengguna
Input Condition	Pengguna login kedalam philoit sebagai member dan mencari topik pada kotak pencarian

3.3.3 Activity Diagram

3.3.3.1 Activity Diagram Melihat Trending Topic



3.3.3.2 Activity Diagram Mencari Topik



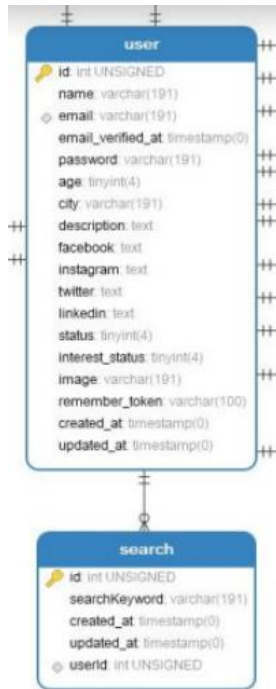
3.3.4 GUI (Graphical User Interface)

Tampilan *trending* dapat diakses melalui halaman utama philoit.id. Halaman *trending* juga dapat diakses melalui *text trending* yang ada pada halaman utama untuk dapat melihat informasi jumlah pencarian pada kurun waktu tertentu yang menjadikan suatu subjek pencarian menjadi *trending*.



TRENDING	
🔥 Puasa	392 kali dicari
🔥 Menu Buka	216 kali dicari
🔥 Ramadhan	205 kali dicari
🔥 Vaksin	186 kali dicari
🔥 Investasi	174 kali dicari
🔥 Liburan Mudik	105 kali dicari
🔥 Corona	96 kali dicari

3.3.5 ERD (Entity Relationship Diagram)



DAFTAR PUSTAKA

- [1] V. S. M. O. J. (. X. H. Z. Ashwin Lall, "Data Streaming Algorithms for Estimating Entropy," 2020.
- [2] G. Y. Iskandarli, "Applying Clustering and Topic Modeling to Automatic Analysis of Citizens' Comments in E-Government," 2020.
- [3] H. C. Darshna Patel, "Fuzzy logic-based single document summarisation with improved sentence scoring technique," *Int. J. Knowledge Engineering and Data Mining*, 2018.