

PENDAHULUAN

Latar Belakang Masalah

Dengan kemajuan teknologi informasi dewasa ini, kebutuhan akan informasi yang akurat sangat dibutuhkan dalam kehidupan sehari-hari, sehingga informasi akan menjadi suatu elemen penting dalam perkembangan masyarakat saat ini dan waktu mendatang. Namun kebutuhan informasi yang tinggi kadang tidak diimbangi dengan penyajian informasi yang memadai, sering kali informasi tersebut masih harus digali ulang dari data yang jumlahnya sangat besar. Perkembangan teknologi ini semakin memudahkan pengguna mendapatkan informasi yang diinginkan. Salah satu sumber yang mudah digunakan untuk mendapatkan informasi. Dalam Al-Quran disebutkan, *“Dan Sesungguhnya kamu telah mengetahui penciptaan yang pertama, maka mengapakah kamu tidak mengambil pelajaran (untuk penciptaan yang kedua)? (QS. Al-Waaqi’ah : 62)”*. Ayat di atas menuntut kita untuk selalu mengambil pelajaran dari setiap peristiwa atau kejadian, termasuk mengambil pelajaran berupa informasi yang dapat digali dari sekumpulan besar data adalah salah satu implementasinya.

pencarian berita dalam media online merupakan suatu hal yang umum dilakukan oleh pengguna internet. Seiring dengan terus bertambahnya pengguna internet diseluruh dunia maka semakin banyak pula orang mencari berita yang tersedia disetiap akun penyedia berita. Ketika melakukan pencarian berita pada akun

tertentu, maka kita dapat melihat data berita pada akun penyedia berita dengan tampilan yang mudah untuk dibaca, pada salah satu akun penyedia berita yang menjadi studi kasus pada skripsi ini yaitu akun The Jakarta Post jejaring sosial Twitter, ketika membuka akun tersebut data berita yang ditampilkan tidak tersusun sesuai dengan topik pembahasan dari setiap berita..

Dalam skripsi ini digunakan algoritma Lingo yang dapat menemukan dan mendeskripsikan kelompok dari dokumen yang saling berhubungan. Lingo mengkombinasikan beberapa metode yang menarik seperti metode Stop Words removal, Stemming, Singular Value Decomposition (SVD), dan sebagainya dengan tujuan dapat menarik perhatian user akan membaca berita pada akun tersebut sesuai dengan topik dari setiap berita.

Information Retrieval (IR) sebagai tindakan, metode dan prosedur untuk menemukan kembali data yang tersimpan, kemudian menyediakan informasi mengenai subyek yang dibutuhkan. Tindakan tersebut mencakup text indexing, inquiry analysis, dan relevance analysis. Data mencakup teks, tabel, gambar, ucapan, dan video. Informasi termasuk pengetahuan terkait yang dibutuhkan untuk mendukung penyelesaian masalah dan akuisisi pengetahuan. Tujuan dari sistem IR adalah memenuhi kebutuhan informasi pengguna dengan me-retrieve semua dokumen yang mungkin relevan, pada waktu yang sama me-retrieve sesedikit mungkin dokumen yang tak-relevan. Sistem ini menggunakan fungsi heuristik untuk mendapatkan dokumen-dokumen yang relevan dengan query

pengguna. Sistem IR yang baik memungkinkan pengguna menentukan secara cepat dan akurat apakah isi dari dokumen yang diterima memenuhi kebutuhannya. Agar representasi dokumen lebih baik, dokumen-dokumen dengan topik atau isi yang mirip dikelompokkan bersama-sama

Text mining memiliki definisi menambang data yang berupa teks dimana sumber data biasanya didapatkan dari dokumen, dan tujuannya adalah mencari kata-kata yang dapat mewakili isi dari dokumen sehingga dapat dilakukan analisa keterhubungan antar dokumen. Tujuan dari text mining adalah untuk mendapatkan informasi yang berguna dari sekumpulan dokumen. Jadi, sumber data yang digunakan pada text mining adalah kumpulan teks yang memiliki format yang tidak terstruktur atau minimal semi terstruktur. Adapun tugas khusus dari text mining antara lain yaitu pengkategorisasian teks (text categorization) dan pengelompokan text clustering.

Clustering adalah pengelompokan objek yang berdasarkan kemiripan yang dimiliki ke dalam suatu grup. Clustering melakukan pengelompokan dokumen tanpa ditentukan terlebih dahulu kelompok-kelompok atau kelas-kelas penampungnya. Clustering sangat berguna dalam pengelompokan dokumen, karena dapat mengekstrak topik dari dokumen secara otomatis dan tanpa perlu ditentukan terlebih dahulu cluster apa saja yang harus ada untuk menampungnya. Clustering sangat berguna untuk pengelompokan dokumen hasil pencarian online karena hasil pencarian yang ditampilkan bisa sangat bermacam-macam walaupun

dalam satu kata kunci. Diharapkan pengguna dapat melakukan seleksi dan menemukan informasi lebih cepat dari hasil penelusuran yang telah dikelompokkan. Sistem pengelompokan ini, dinilai penting untuk mempermudah pengguna untuk mendapatkan informasi yang dibutuhkan. Hal ini dapat diimplementasikan dalam sebuah aplikasi pencarian berita, yang diharapkan mampu mempermudah pengguna dalam pencarian berita.

Berdasarkan uraian diatas, maka akan dilakukan penelitian dan penyusunan tugas akhir yang berjudul “Pengelompokan (Clustering)Berita pada Titter Menggunakan Algoritma Lingo”.

Perumusan Masalah

Berdasarkan latar belakang masalah di atas, maka perumusan masalah yang dibahas adalah Bagaimana menerapkan Algoritma Lingo untuk mengelompokkan berita pada akun jejaring sosial penyedia berita.

Tujuan

Mengimplementasikan Algoritma Lingo untuk mengcluster berita pada jejaring sosial.

Batasan Masalah

Jejaring sosial yang digunakan adalah Twitter.

Nama akun yang digunakan adalah @jakpost.

Algoritma yang digunakan untuk mengcluster adalah Lingo.

Bahasa pemrograman yang digunakan adalah bahasa pemrograman C#.

Metodologi Penelitian

Penelitian ini menggunakan metodologi penelitian sebagai berikut:

Metode pengumpulan data

Metode pengumpulan data dalam tugas akhir ini adalah sebagai berikut:

a. Studi Literatur

Pengumpulan data dilakukan dengan cara mempelajari literatur tentang penulisan dan mengenai hal-hal yang mendukung program aplikasi serta mempelajari dari sumber data yang lain seperti dari internet dan lain-lain.

Wawancara

Teknik pengumpulan data dengan mengadakan tanya jawab secara langsung dan bimbingan dari pembimbing.

Metode pengembangan perangkat lunak

Dalam proses pengembangan perangkat lunak, metode yang digunakan adalah metode pengembangan prototype. Prototype tersebut dievaluasi oleh pelanggan atau pemakai dan dipakai untuk menyaring kebutuhan pengembangan perangkat lunak. Iterasi terjadi pada saat disetel untuk memenuhi kebutuhan pelanggan dan pada saat yang sama memungkinkan pengembang untuk secara lebih baik memahami apa yang harus dilakukan. (Pressman,2002).

Tahapan-tahapan dalam Prototyping sebagai berikut:

Pengumpulan kebutuhan

Mendefinisikan format seluruh perangkat lunak, mengidentifikasi semua kebutuhan, dan garis besar sistem yang akan dibuat.

Membangun prototyping

Membangun prototyping dengan membuat perancangan sementara yang berfokus pada penyajian kepada user (misalnya dengan membuat input dan format output).

Evaluasi prototyping

Evaluasi ini dilakukan oleh user apakah prototyping yang sudah dibangun sudah sesuai dengan keinginan user. Jika sudah sesuai maka langkah 4 akan diambil. Jika tidak prototyping direvisi dengan mengulangi langkah 1, 2, dan 3.

Mengkodekan sistem

Dalam tahap ini prototyping yang sudah disepakati diterjemahkan ke dalam bahasa pemrograman yang sesuai

Menguji sistem

Setelah sistem sudah menjadi sebuah perangkat lunak yang siap pakai, maka dilakukan langkah 3 dahulu sebelum digunakan. Pengujian ini dilakukan dengan Black Box Testing.

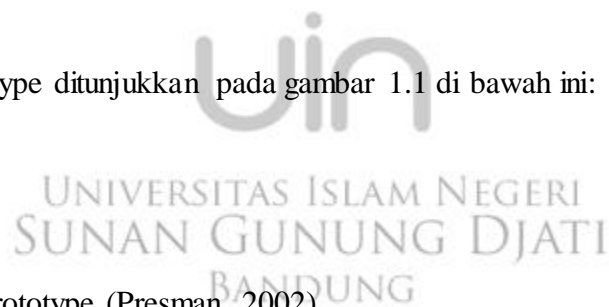
Evaluasi Sistem

User mengevaluasi apakah sistem yang sudah jadi sudah sesuai dengan yang diharapkan . Jika ya, langkah 7 dilakukan; jika tidak, ulangi langkah 4 dan 5.

Menggunakan sistem

Perangkat lunak yang telah diuji dan diterima user siap untuk digunakan.

Tahapan prototype ditunjukkan pada gambar 1.1 di bawah ini:



Gambar 1.1 Prototype (Presman, 2002)

Sistematika Penulisan

Untuk menjelaskan gambaran secara umum mengenai isi dari pembahasan ini, berikut akan dikemukakan mengenai sistematika penulisan tugas akhir ini yaitu:

BAB I PENDAHULUAN

Bab ini menjelaskan tentang latar belakang masalah, rumusan masalah, tujuan, batasan masalah, dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Bab ini menjelaskan tentang teori-teori apa saja yang berkaitan dengan topik penelitian yang telah dibuat berdasarkan hasil penelitian dan hal-hal yang berguna dalam proses penulisan tugas akhir ini.

BAB III ANALISIS DAN PERANCANGAN SISTEM

Bab ini berisi tentang analisis dan perancangan dari penerapan metode Lingo pada pengelompokan berita pada akun jejaring sosial.

BAB IV IMPLEMENTASI DAN PENGUJIAN SISTEM

Bab ini berisi tentang implementasi dan pengujian penerapan metode lingo pada pengelompokan berita pada akun jejaring sosial.

BAB V KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan dari hasil pembahasan dari bab – bab sebelumnya, serta saran yang merupakan tindak lanjut dari kesimpulan.

