

BAB I

PENDAHULUAN

Dalam bab ini akan dibahas latar belakang dilaksanakannya penelitian, identifikasi masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

Latar Belakang

Seiring dengan pesatnya perkembangan zaman, kebutuhan informasi di tengah masyarakat juga mulai berkembang. Masyarakat tidak lagi mengkonsumsi informasi sebagai kebutuhan sesaat, namun telah menjadikan informasi tersebut sebagai kebutuhan kontinyu dan rutin. Berbagai kebutuhan informasi dimulai dari informasi dunia politik, hukum, ekonomi, bisnis, sampai hiburan telah menjadi konsumsi vital bagi masyarakat. Semakin tingginya kebutuhan akan informasi juga akan berdampak pada tingkat peredaran informasi tersebut di tengah masyarakat.

Perkembangan lalu lintas informasi yang semakin cepat di tengah masyarakat tentunya tidak terlepas dari peranan media informasi. Dengan adanya media informasi online banyak membantu masyarakat untuk mengetahui informasi yang beredar dengan mudah. Informasi ini ditampung dan ditampilkan dalam suatu *website* berita untuk kemudian diakses oleh pengguna secara *online*. Namun dengan karakteristik dari *website* berita yang hanya menampilkan kumpulan berita secara tidak terorganisir akan menyulitkan pengguna untuk mengetahui berita yang diinginkan pada suatu kategori tertentu dan akan sangat tidak praktis jika pengguna harus membaca semua berita untuk mendapatkan berita yang diinginkan.

News aggregator dibentuk untuk mengurangi usaha dan waktu dari pengguna untuk mengetahui berita yang diinginkannya. *News aggregator* ini mengumpulkan berita dari berbagai sumber atau *website* berita ke dalam satu halaman *website* dan mengelompokkannya ke dalam beberapa kategori berita.

Sudah terdapat banyak *news aggregator* yang dapat diakses saat ini, diantaranya adalah Google News yang dapat diakses pada tautan <https://news.google.com/> dan TechMeme yang dapat diakses pada tautan <https://techmeme.com/>. Google News mengumpulkan berita-berita terbaru yang berasal dari berbagai *website* berita dan mengelompokkannya ke dalam kategori berita tertentu. Sedangkan TechMeme lebih terpusat pada topik mengenai teknologi dan mengumpulkan berita teknologi pada satu *website* portal TechMeme.

Sumber berita tentunya tidak hanya terfokus pada berita yang telah diliput oleh wartawan dan ditampilkan pada *website* berita. Banyak informasi berguna yang tersebar secara cepat di media sosial. Wicaksono, seorang jurnalis beritagar.id pada *meet up class* “Peran *Social Media* dalam Diseminasi Berita” di Social Media Week (SMW) di Pacific Place, Jakarta pada Selasa (24 Februari 2015) mengatakan bahwa bahwa 8 dari 10 wartawan Indonesia mendapatkan ide berita dari media sosial. Berita yang diperoleh wartawan dari media sosial itulah yang kemudian dipublikasikan pada media informasi. Salah satu media sosial yang memegang peranan penting dalam penyebaran informasi adalah Twitter.

Twitter merupakan salah satu media sosial yang populer dan banyak digunakan saat ini. Twitter memungkinkan pengguna untuk terhubung dengan orang-orang di seluruh dunia dan saling bertukar pesan maupun buah pemikiran. Twitter memungkinkan pengguna untuk mengirim dan membaca pesan dengan panjang sebanyak 140 karakter dan ditampilkan pada halaman profil pengguna. Pesan ini dinamakan *tweet* atau kicauan. *Tweet* dapat berupa pesan teks, gambar, video, maupun pesan atau media yang berasal dari *link* eksternal lainnya.

Twitter dapat diakses menggunakan perangkat komputer atau *mobile* dengan memanfaatkan layanan internet. Pengguna dapat mengakses Twitter dengan membuka langsung *website* resmi Twitter pada tautan berikut ini <https://twitter.com/> dan langsung memanfaatkan semua fitur umum Twitter yang tersedia. Selain itu pengguna juga dapat mengakses Twitter melalui aplikasi *third-party twitter client* seperti TweetDeck, Hootsuite, Tweetbot, dan lain sebagainya. *Third-party twitter client* berisi semua fitur umum Twitter serta fitur tambahan

Rany Kasman, 2015

PEMBANGUNAN TWEET AGGREGATOR DENGAN MENGGUNAKAN METODE NAÏVE BAYES

Universitas Pendidikan Indonesia | [\ .upi.edu](http://upi.edu) perpustakaan.upi.edu

seperti *multiple account*, penjadwalan pengiriman *tweet*, pengaturan *user interface*, pengaturan navigasi yang lebih mudah, dan lainnya. Twitter juga memungkinkan pengguna untuk mengirim tweet melalui *Short Message Service (SMS)*.

Kemudahan yang ditawarkan oleh Twitter menarik banyak pengguna dari berbagai kalangan dan jumlah penggunanya bertambah setiap harinya. Hal ini tentunya akan berbanding lurus dengan tingkat penyebaran informasi di masyarakat. Dengan banyaknya informasi yang beredar melalui Twitter maka perlu diperhatikan tingkat relevan dari informasi tersebut. Untuk mem-*filter* informasi yang beredar ini maka diperlukan sumber-sumber terpercaya yang berkecimpung langsung di bidang berita tersebut. Seperti contohnya adalah berita berbau politik dan hukum berasal dari tokoh-tokoh berkompeten yang terjun langsung di bidang tersebut sehingga informasi yang dihasilkan lebih relevan, mendalam, dan menarik. Informasi ini cenderung menyebar lebih cepat jika dibandingkan dengan informasi yang berasal dari website berita lainnya.

Untuk mengumpulkan informasi yang tersebar melalui Twitter maka dibentuk suatu *website* portal seperti *news aggregator* berbasis *tweet* yang disebut juga dengan *tweet aggregator*. Pada *tweet aggregator*, *tweet* yang berisi informasi berguna dari orang-orang yang berkompeten dikumpulkan dan dikelompokkan ke dalam beberapa kategori seperti kategori politik dan hukum, hiburan, ekonomi, kesehatan, bisnis, teknologi, olahraga, dan lainnya. Pengelompokkan ini dilakukan dengan menggunakan teknik klasifikasi teks dengan pembelajaran mesin.

Penelitian tentang klasifikasi teks dengan menggunakan sumber data *tweet* sebelumnya pernah dilakukan oleh Sriram pada tahun 2010. Data yang digunakan pada penelitian ini adalah *tweet* berbahasa Inggris dan mengklasifikasikannya ke dalam kategori topik pembicaraan yang umum dibahas pada Twitter seperti *news*, *opinions*, *deals*, *events*, dan *private messages*. Pengklasifikasian ini dilakukan berdasarkan informasi profil *user* dan fitur spesifik yang diekstrak dari *tweet* seperti pemendekkan kata, kata *slang*, frasa berbasis waktu, kata opini, tekanan kata, kemunculan *@username* pada awal *tweet*, dan kemunculan *@username* di

dalam *tweet*. Sriram dkk menggunakan metode *Naïve Bayes*, *Decision Tree*, dan *Sequential Minimal Optimization* (SMO). Hasil pengklasifikasian kemudian dievaluasi dengan menggunakan teknik *5-fold cross validation* dan menghasilkan akurasi tertinggi pada metode *Naïve Bayes*. (Sriram, 2010)

Penelitian lainnya dilakukan oleh Go dkk pada tahun 2009 dengan menerapkan klasifikasi teks pada analisis sentimen dengan menggunakan sumber data *tweet*. Dataset *tweet* yang telah dikumpulkan mengandung emotikon sebagai *noisy* label. Klasifikasi dilakukan dengan menggunakan metode *Naïve Bayes*, *Maximum Entropy*, dan *Support Vector Machine* (SVM) menghasilkan akurasi di atas 80% saat dilatih menggunakan data emotikon. Data *tweet* sebelumnya di praproses dengan menghilangkan *username*, *link*, dan perulangan huruf. Namun dari banyaknya penelitian yang telah ditemukan, penelitian klasifikasi yang terfokus pada informasi berita berbahasa Indonesia yang tersebar pada Twitter cukup jarang ditemukan. Jika ada, penelitian yang ditemukan adalah klasifikasi dengan menggunakan data *tweet* berbahasa Inggris. Oleh karena itu dilakukan pembangunan *Tweet Aggregator* dengan menerapkan klasifikasi informasi berita berbahasa Indonesia yang tersebar melalui Twitter. (Go, Bhayani, & Huang, 2009)

Tweet aggregator ini dibangun dengan menggunakan salah satu metode pengklasifikasian teks yang memiliki karakteristik sederhana, mudah diaplikasikan, dan memiliki akurasi tinggi, yaitu algoritma *Naïve Bayes* untuk mengelompokkan *tweet* berdasarkan kategori yang ada.

Identifikasi Masalah

Identifikasi masalah yang akan dibahas dalam “Pembangunan *Tweet Aggregator* dengan Menggunakan Metode *Naïve Bayes*” adalah:

1. Bagaimana pengklasifikasian *tweets* berdasarkan kategori tertentu dengan menggunakan metode *Naïve Bayes*?
2. Bagaimana tingkat kepuasan pengguna terhadap aplikasi pengelompokan *tweet* berita?

Batasan Masalah

Adapun batasan masalah dalam penelitian ini, diantaranya :

- a. Metode yang digunakan adalah algoritma *Naïve Bayes* untuk pengklasifikasian teks berdasarkan kategori berita tertentu.
- b. Data yang digunakan yaitu data *tweets* yang berasal dari akun tokoh yang memiliki kompetensi pada bidang berita. Daftar tokoh berkompeten didapatkan dari Ensiklopedi Tokoh Indonesia Online pada tautan <http://www.tokohindonesia.com/> dan daftar wartawan terkenal Indonesia.
- c. *Tweets* yang digunakan yaitu *tweets* berbahasa Indonesia.
- d. Pelabelan *tweet* berita dilakukan atas pemahaman penulis sebagai *supervisor* dalam pengklasifikasian *tweet*.

Tujuan Penelitian

Tujuan diadakannya penelitian ini adalah:

1. Untuk dapat mengklasifikasikan *tweets* berdasarkan kategori tertentu dengan menggunakan metode *Naïve Bayes*.
2. Untuk dapat membuktikan sejauh mana tingkat kepuasan pengguna terhadap aplikasi pengelompokkan *tweet* berita.

Manfaat Penelitian

Dengan adanya penelitian ini, diharapkan dapat memudahkan pengguna untuk mendapatkan berita terbaru dari berbagai macam kategori berita secara cepat langsung dari orang-orang yang berkompeten di dalamnya.

Sistematika Penulisan

Sistematika penulisan dalam proposal ini sebagai berikut :

BAB I PENDAHULUAN

Bab ini berisi pembahasan masalah secara umum, terdiri dari latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Bab ini berisi dasar teori yang digunakan dalam penelitian ini. Adapun yang dibahas pada bab ini adalah teori yang berkaitan dengan

pembangunan *tweet aggregator* dengan menggunakan metode *Naïve Bayes*.

BAB III METODOLOGI PENELITIAN

Bab ini berisi penjelasan langkah-langkah yang akan dilakukan dalam penelitian.

BAB IV HASIL PENELITIAN

Bab ini berisi uraian tentang hasil penelitian dan pembahasan terhadap hasil penelitian yang dilakukan

BAB V KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan dari keseluruhan penelitian yang telah dilakukan serta saran dari penulis untuk kegiatan penelitian selanjutnya terkait dengan topik yang sedang dibahas.