

BAB 1

Latar Belakang

Jejaring sosial menjadi fenomena dunia yang marak dibicarakan di kalangan masyarakat. Lebih dari 80% pengguna Internet adalah pengguna jejaring sosial [1]. Salah satu jejaring sosial yang banyak digunakan adalah jejaring sosial Twitter. Twitter memiliki pengguna yang sangat besar, terdiri dari jutaan pengguna [2]. Ini terbukti dari tahun ke tahun, pengguna Twitter kian menambah, ditambah dengan banyaknya pengguna aktif Twitter yang tercatat. Dari awal tahun 2010 sampai akhir tahun 2014 menunjukkan bahwa jumlah pengguna Twitter aktif di seluruh dunia rata-rata ada 288 juta pengguna [3].

Twitter memiliki fitur yang dapat digunakan untuk mem-*posting tweet*, dimana dari *tweet* tersebut dapat memberikan informasi ke pengguna lain, salah satunya opini terkait tempat makan yang dikunjungi. Informasi tersebut sebenarnya dapat dicari melalui akun Twitter yang membahas informasi seputar makanan. Contoh akun yang membahas tempat makan di Bandung adalah @infobdgkuliner, @Kulinier_Bandung, @KulinierBDG, dll. Selain akun tersebut, pengguna biasanya memberikan informasi tempat makan yang dikunjungi dengan memberikan *hashtag* atau langsung *mention* ke akun Twitter tempat makan tersebut. Akan tetapi, pengguna Twitter yang memberikan opini tentang tempat makan tersebut tidak hanya satu dan tidak saling terintegrasi, sehingga informasi tersebut menjadi kurang bermanfaat. Pada penelitian terdahulu mengkombinasikan metode SentiStrenght, Hybrid TF-IDF dan Cosine Similarity untuk mengekstraksi ringkasan sentiment positif dan negatif terhadap data Twitter [4].

Dalam penelitian ini penulis memanfaatkan ilmu data mining, yakni menemukan pola yang menarik dari data yang didapat dari *Twitter*. Data tersebut merupakan *tweet* yang mengandung opini terhadap tempat makan yang diambil menggunakan NodeXL. Data Twitter yang didapat dari NodeXL dilakukan *filtering* manual untuk mendeteksi apakah opini yang didapat sesuai dengan penelitian penulis, yaitu opini terhadap tempat makan, dan mengubah kata-kata yang tidak baku menjadi kata baku. Kemudian data tersebut dilakukan *preprocessing* dengan *cleaning*/menghilangkan *noise* (menghilangkan url, *retweet*, *mention*), *case folding*, *tokenizing*, *stopword removal*, dan *stemming*. Hasil dari *preprocessing* kemudian diklasifikasikan menggunakan metode klasifikasi *Naïve Bayesian Classification*. Pendekatan ini merupakan pendekatan mengacu pada teorema Bayes yang merupakan prinsip peluang statistika untuk mengkombinasikan pengetahuan sebelumnya dengan pengetahuan baru. Prinsip ini kemudian digunakan untuk memecahkan masalah klasifikasi [5] [6]. Hasil yang didapat dari klasifikasi data twitter menggunakan *Naïve Bayesian Classification* berupa dua kelas, yakni kelas positif dan negatif. Dan dari hasil tersebut dapat disimpulkan apakah tempat makan rekomen atau tidak.

Topik dan Batasannya

Berdasarkan latar belakang diatas, permasalahan yang ada adalah bagaimana menerapkan dan menganalisis penggunaan *Naïve Bayesian Classification* dalam mengklasifikasikan *tweet* yang berisi informasi opini pengguna Twitter terkait tempat makan di Bandung beserta tingkat akurasi yang didapatkan. Adapun batasan masalah dari penelitian ini adalah data yang digunakan adalah data Twitter dan *tweet* berbahasa Indonesia, data yang diambil hanya 3 tempat makan yang ada di Bandung, hasil pengklasifikasian adalah kelas positif dan negatif, implementasi dilakukan menggunakan bahasa pemrograman PHP dan sistem berjalan secara *offline*.

Tujuan

Tujuan dari penelitian ini adalah dapat menerapkan dan menganalisis *Naïve Bayesian Classification* dalam mengklasifikasikan *tweet* yang berisi informasi opini pengguna *Twitter* terkait tempat makan di Bandung dan dapat menentukan tingkat akurasi.

Organisasi Tulisan

Penulisan penelitian ini disusun dalam beberapa bab yang dari setiap bab berisi: Bab 1 – Pendahuluan, Bab 2 – Studi Terkait, Bab 3 – Sistem yang Dibangun, Bab 4 – Evaluasi, Bab 5 – Kesimpulan.