

BAB I

PENDAHULUAN

1.1 Latar Belakang

DC Universe merupakan sebuah alam semesta fiksi dimana kumpulan pahlawan super dan penjahat super berdasarkan pada karakter yang muncul dalam buku-buku komik oleh DC Comics ada didalamnya [1]. DC Comics sendiri adalah penerbit buku komik terbesar dan tertua di amerika milik DC Entertainment yang memproduksi dan menampilkan pahlawan super dan penjahat super [2].

Twitter merupakan sosial media yang sampai saat ini selalu dipilih untuk mencari informasi dan memberi informasi. Jumlah pengguna *twitter* 2019 mencapai 152 juta, naik hampir 21 persen dari jumlah 126 juta yang tercatat di periode yang sama tahun sebelumnya (tekno.kompas.com, 2020). Informasi yang tersebar dalam *twitter* juga sangat banyak mulai dari berita, opini, komentar, kritik dan semua informasi tersebut ada yang bersifat positif, negatif, dan netral.

Klasifikasi opini atau analisis sentimen adalah salah satu cara untuk mengetahui sentimen seseorang atau sekelompok orang terhadap produk, layanan, isu atau golongan tertentu dari berbagai platform media sosial dan internet. Klasifikasi sentimen dapat dilakukan salah satunya dengan mengumpulkan data melalui *twitter* dengan topik tertentu, salah satunya dapat diperoleh melalui *tweet* yang ada pada *twitter*. Pada penelitian ini kasus topik yang di angkat adalah *tweet* yang berhubungan dengan kata kunci DC Universe. *Tweet* akan dicari dan dihimpun dengan menggunakan tools yaitu API *Twitter*. [3]

Data yang didapat dari proses penghimpunan akan diolah, selanjutnya data tersebut akan dilakukan klasifikasi opini atau sentimen dengan menggunakan algoritma *naïve bayes classifier*, selanjutnya proses klasifikasi akan menghasilkan kelas sentimen positif, negatif dan netral. [3]

Untuk memulai sebuah bisnis bertema pahlawan super, ada beberapa contoh bisnis yang bisa jadi referensi dan juga dapat dimanfaatkan untuk meraup keuntungan dari bisnis tersebut yaitu, persewaan komik, menjual *merchandise*, membuat *clothing line*, membuat kostum *cosplay*, membuat makanan bertema pahlawan super, menjual *action figure*. [4]

Dalam upaya untuk memulai bisnis bertema pahlawan super khususnya bertema DC Universe, perlu memperhatikan / mendengarkan konsumen DC Universe di Indonesia. Dengan mengetahui sentimen masyarakat Indonesia dan hal-hal apa saja yang seringkali muncul / dibicarakan di media sosial terhadap DC Universe. Dengan mengetahui sentimen, pendapat, keinginan masyarakat Indonesia terhadap DC Universe, maka hal itu dapat digunakan sebagai acuan untuk memulai bisnis bertema pahlawan super di Indonesia. Oleh karena itu penulis merasa penting untuk mengajukan penelitian dengan judul Analisis Opini Terhadap DC Universe Pada Media Sosial Twitter Menggunakan Metode Naïve Bayes.

1.2 Rumusan Masalah

Rumusan masalah pada penelitian ini adalah :

1. Berapa persentase sentimen masyarakat di *twitter* terkait DC Universe ?

2. Seberapa tepat algoritma *naïve bayes classifier* dalam mengklasifikasi sentimen terkait DC Universe ?
3. Bagaimana mendapatkan informasi yang bermanfaat untuk memulai bisnis bertema pahlawan super dalam proses analisis sentimen terkait DC Universe ?

1.3 Batasan Masalah

Batasan Masalah pada penelitian ini adalah :

1. Algoritma yang digunakan dalam pengklasifikasian ini adalah *naïve bayes classifier*.
2. Menggunakan metode pembobotan TF-IDF.
3. *Tweet* yang digunakan mengandung kata kunci DC Universe.
4. *Tweet* yang digunakan berbahasa indonesia.
5. Menggunakan bahasa pemrograman PHP.

1.4 Maksud dan Tujuan Penelitian

Maksud dan tujuan dari penelitian ini adalah :

1. Mengimplementasikan algoritma *naïve bayes classifier* untuk mengklasifikasi sentimen masyarakat di *twitter* terkait DC Universe.
2. Mengetahui persentase sentimen masyarakat di *twitter* terkait DC Universe.
3. Mengetahui seberapa tepat algoritma *naïve bayes classifier* dalam mengklasifikasi sentimen terkait DC Universe.
4. Mendapatkan informasi bermanfaat dalam proses analisis sentimen terkait DC Universe sebagai acuan untuk memulai bisnis bertema pahlawan super.

1.5 Manfaat Penelitian

Manfaat yang diharapkan peneliti pada penelitian ini adalah :

1. Memberi pengetahuan tambahan bagaimana cara mengimplementasikan algoritma *naïve bayes classifier* untuk permasalahan berupa *text mining* di Twitter.
2. Sistem yang dibuat dapat digunakan untuk mengklasifikasi sentimen masyarakat terhadap suatu topik di twitter.
3. Sebagai panduan untuk penelitian dengan permasalahan berupa *text mining* di Twitter.

1.6 Metode Penelitian

Metode penelitian ini sendiri merupakan cara utama yang digunakan peneliti untuk mencapai tujuan dan menjawab masalah penelitian, adapun beberapa metode-metode penelitian yang akan digunakan antara lain :

1.6.1 Metode Pengumpulan Data

Beberapa metode pengumpulan *data* yang dilakukan peneliti untuk mencari *data* dan informasi terkait penelitian, Metode tersebut antara lain :

a. Studi kepustakaan

Metode studi kepustakaan ini dilakukan dengan mengumpulkan jurnal, *literature*, paper, makalah, buku, maupun situs *internet* sebagai sumber pustaka yang berkaitan dengan penelitian.

b. Pengumpulan data *Twitter*

Data *tweet* dengan keyword DC Universe dikumpulkan dengan memanfaatkan *search API Twitter* yang telah disediakan oleh *twitter*.

1.6.2 Metode Analisis Data

Metode Analisis data dilakukan peneliti untuk mengolah data yang telah dikumpulkan, yaitu :

a. *Preprocessing*

Preprocessing ini dilakukan untuk mengatur data yang tidak teratur menjadi data yang teratur dan lebih terstruktur, serta menghindari data yang tidak perlu.

b. TF-IDF

Data hasil dari *preprocessing* akan dihitung bobot tiap-tiap katanya menggunakan metode TF-IDF.

c. *Naïve Bayes Classifier*

Data akan dihitung dan di klasifikasi menggunakan *Naïve Bayes Classifier* untuk ditentukan apakah data tersebut berkategori positif, negatif, atau netral.

d. *Confusion Matrix*

Confusion Matrix dilakukan untuk evaluasi akurasi hasil dari klasifikasi.

1.6.3 Metode Perancangan

Metode perancangan dilakukan peneliti untuk merancang sistem terkait penelitian, yaitu :

- a. Merancang sebuah aplikasi analisis sentimen *twitter* berbasis web.
- b. Merancang *database* sebagai penyimpanan data dan hasil analisis.

1.7 Sistematika Penulisan

Penulisan penelitian ini disusun dengan sistematika penulisan yang dibagi kedalam lima bab, yaitu :

BAB I Pendahuluan

Bab ini berisi Latar belakang, rumusan masalah, batasan masalah, maksud dan tujuan penelitian, manfaat penelitian, metode penelitian, dan sistematika penulisan.

BAB II Landasan Teori

Bab ini berisi penjelasan tentang teori-teori dan konsep algoritma yang berkaitan dan digunakan dalam penelitian ini.

BAB III Metode Penelitian

Bab ini berisi penjelasan langkah-langkah dan metodologi yang dilakukan dalam penelitian ini.

BAB IV Hasil dan Pembahasan

Bab ini berisi tentang hasil penelitian dan pembahasan terkait penelitian yang telah dilakukan.

BAB V Penutup

Bab ini berisi kesimpulan dari seluruh penelitian yang telah dilakukan, serta saran dari peneliti untuk penelitian selanjutnya terkait dengan topik yang sedang dibahas.