

BAB I

PENDAHULUAN

Bab ini membahas tentang latar belakang, rumusan masalah, tujuan, manfaat, batasan masalah, dan sistematika penulisan penelitian ini.

1.1 Latar Belakang

Perkembangan pasar game dunia pada tahun 2019 telah mencapai nilai US\$152 miliar atau sekitar Rp.2,15 kuadriliun dan telah mengalami peningkatan sebesar 9,6% dibandingkan dengan tahun sebelumnya. Pada wilayah Asia sendiri, Indonesia merupakan salah satu pasar *video game* yang mencapai nilai US\$941 juta atau sekitar Rp.13 triliun. Berdasarkan angka yang terlihat, maka bukanlah hal yang mengherankan jika industri dunia *e-sport* mengalami perkembangan yang pesat. *E-sport* sendiri merupakan permainan kompetitif yang berbasis *video game* [1]. Perkembangan industri *e-sport* yang ada di Indonesia telah mendapatkan respon yang positif dari pemerintah. Indonesia Esports Association atau biasa disebut juga dengan IeSPA merupakan wadah resmi sebagai asosiasi untuk menaungi komunitas ini di Indonesia. IeSPA juga membantu para pemain agar bisa mengembangkan bakat dan kemampuan dalam dunia game kompetitif. Selain itu juga, pemerintah menyelenggarakan kejuaraan seperti IEC University Series 2019, Youth National eSport Championship, dan Piala Presiden [2].

Indonesia merupakan pengguna internet terbesar ketiga yang ada di Asia, dimana terdapat 143,26 juta jiwa dan diperkirakan bisa mencapai 269,54 juta jiwa yang artinya adalah 53% dari total populasi yang ada [3]. Twitter merupakan salah satu situs jejaring sosial yang dapat membantu manusia untuk saling berhubungan dan berbagi informasi satu sama lain. Twitter juga dikenal sebagai layanan *micro-blogging*, dimana penggunanya diberikan fasilitas untuk menulis, membaca, serta mengirimkan informasi pada forum umum. Twitter mampu untuk mengembangkan fasilitas agar perusahaan dan pengguna dapat saling berinteraksi walau tidak secara langsung. Perusahaan bisa menyebarkan informasi terbaru kepada pengguna, dan

memperoleh *feedback* dari pengguna dengan waktu yang singkat, dengan demikian perusahaan mampu mengevaluasi layanan-layanan yang telah diberikan.

Dengan menggunakan analisis sentiment, teks dapat diklasifikasikan ke dalam polaritas (positif, netral, dan negatif) [4]. Analisis sentimen juga dijadikan sebagai pendukung keputusan dalam menilai suatu keadaan berdasarkan kata yang terkandung atau dalam bentuk informasi yang membantu memprediksi dan memberikan pengaruh dalam pengambilan kebijakan. Banyak metode analisis yang dapat digunakan untuk menganalisis pendapat masyarakat berdasarkan informasi yang ada pada media sosial semacam Twitter. Metode analisis sentiment adalah salah satu di antaranya. Metode analisis sentimen merupakan salah satu metode untuk menganalisis data yang didapatkan dari media sosial atau internet sehingga dapat diketahui polaritas dari data tersebut. Polaritas dari tanggapan yang ada dapat dikumpulkan, sehingga akan dapat digunakan untuk memprediksi suasana publik atau gambaran perasaan netizen bersifat negatif atau positif dengan menggunakan analisis sentiment [5]. Penggunaan analisis sentimen dapat di terapkan dalam mengategorikan atau mengklasifikasikan tweets di Twitter terhadap opini masyarakat. Namun, untuk melakukan analisis sentimen harus ada suatu model atau metode perhitungan probabilitas yang bisa membantu dalam proses pengklasifikasian sebuah tulisan. Beberapa model algoritma pengklasifikasian data yang sering digunakan antara lain algoritma *Naïve Bayes*, *Decision Tree* C4.5, dan *k-Nearest Neighbors* (KNN) [6].

Penelitian ini akan menggunakan model algoritma *Naïve Bayes* sebagai algoritma pengklasifikasian sentimen pada tweet media sosial Twitter, dan merupakan suatu algoritma yang terbukti efektif pada berbagai bidang klasifikasi teks. TF-IDF merupakan metode ekstraksi fitur dan pembobotan kata. TF-IDF merupakan sebuah ukuran statistik yang digunakan untuk mengevaluasi suatu kata dalam dokumen atau dalam sekumpulan kata [7].

Algoritma *Naïve Bayes* digunakan dalam penelitian ini didasarkan pada banyaknya data yang berupa tweet dari twitter yang akan ditentukan polaritinya. *Naïve Bayes* akan mengklasifikasikan objek (tweet) ke dalam salah satu kelas yang telah ada dalam data sampel yang sebelumnya telah ditetapkan [8]. Penelitian sebelumnya yang membandingkan *Naïve Bayes* dan *Holistic Lexicon Based*

memperoleh hasil akurasi terbaik sebesar 80% dan nilai presisi sebesar 75% pada algoritma *Naive Bayes*, sedangkan pada algoritma *Holistic Lexicon Based* mendapatkan nilai presisi sebesar 70% dan akurasi sebesar 70% [9]. Hal ini menandakan bahwa metode *Naive Bayes* masih lebih baik jika dibandingkan dengan metode *Holistic Lexicon Based* pada tingkat presisi dan akurasi. Dengan adanya penelitian ini diharapkan dapat memberikan suatu informasi yang akurat dan bermanfaat bagi masyarakat mengenai sentimen pengguna Twitter tentang *e-sport* di Indonesia.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang di atas, maka dapat dipaparkan perumusan masalah sebagai berikut:

1.2.1 Permasalahan

Perkembangan industri *e-sport* yang ada di Indonesia telah mendapatkan respon positif dari pemerintah. IeSPA sebagai wadah untuk menaungi komunitas dan para pemain agar dapat mengembangkan bakat dan kemampuan dalam dunia game yang kompetitif. Dunia *e-sport* cukup menarik banyak perhatian masyarakat khususnya bagi para pecinta video game. Hal ini menimbulkan opini-opini dari masyarakat tentang dunia *e-sport* yang telah menjadi trend baru masa kini. Twitter menjadi sarana untuk mempublikasikan opini tersebut dalam hal ini berbentuk tweet.

1.2.2 Usulan Solusi

Solusi yang ditawarkan adalah dengan mengumpulkan sentimen atau opini dari masyarakat mengenai dunia *e-sport* di Indonesia. Data dikumpulkan dari media sosial Twitter berupa tweet. Data lalu diolah dan diberikan nilai di setiap kata. Pembobotan kata dilakukan dengan menggunakan metode TF-IDF yang berperan sebagai ekstraksi fitur. Pengklasifikasian dilakukan dengan menggunakan *Naive Bayes* serta menggunakan *K-Fold Cross Validation* sebagai metode untuk melakukan evaluasi model.

1.2.3 Pertanyaan Penelitian

Berapa tingkat akurasi dari klasifikasi untuk data Twitter dengan topik E-Sport dengan menggunakan *Naive Bayes* dan TF-IDF?

1.3 Tujuan dan Manfaat

Dalam penelitian ini akan menghasilkan tujuan dan manfaat sebagai berikut:

1.3.1 Tujuan

1. Melakukan analisis serta melakukan evaluasi terhadap sentiment masyarakat tentang perkembangan dunia *e-sport* di Indonesia pada Twitter.
2. Menerapkan model algoritma *Naïve Bayes* dan TF-IDF untuk mengklasifikasikan data tweet tentang dunia *e-sport* di Indonesia.

1.3.2 Manfaat

1. Memperoleh informasi tentang bagaimana opini masyarakat pengguna Twitter terhadap dunia *e-sport* di Indonesia.
2. Terklasifikasinya opini masyarakat sehingga dapat menjadi suatu media pendukung untuk mendapatkan *feedback* dari masyarakat dalam proses pengambilan keputusan.

1.4 Batasan Masalah

Batasan masalah pada penelitian ini melingkupi hal berikut yaitu:

1. Data yang digunakan adalah data yang diambil dari Twitter yaitu berupa tweet para pengguna dan akan menjadi target analisis opini.
2. Penelitian ini mengumpulkan data yang hanya menggunakan Bahasa Indonesia.

1.5 Sistematika Proposal

Sistematika penulisan proposal tugas akhir ini adalah sebagai berikut:

BAB I PENDAHULUAN

Bab ini membahas tentang latar belakang dari masalah yang dibahas, usulan solusi yang dapat menyelesaikan permasalahan, tujuan dan manfaat penerapan atas solusi yang diajukan, batasan-batasan, serta penyusunan proposal yang sistematis.

BAB II KAJIAN PUSTAKA

Bab ini memuat tentang penjelasan dari setiap teori yang membantu memuat uraian tentang tinjauan pustaka terdahulu dan kerangka teori relevan dan terkait dengan tema penelitian.

BAB III METODOLOGI PENELITIAN

Pada bab ini akan membahas tentang tahapan penelitian dengan melakukan beberapa tahapan, serta diagram sistem pada penelitian ini.