

BAB I

PENDAHULUAN

1.1. Latar Belakang

Barclays Premier League atau disebut juga dengan Liga Primer Inggris merupakan kompetisi liga sepak bola profesional di Inggris. Berdasarkan data yang dilansir <http://premierleague.com>, Liga Primer Inggris merupakan kompetisi tertinggi di Inggris dengan tingkat popularitas tinggi dan paling banyak ditonton di dunia, disiarkan di 212 wilayah kepada 643 juta pemirsa. Liga Primer Inggris juga merupakan liga paling kompetitif, dengan tidak adanya dominasi yang terlalu kuat sehingga membuat liga ini sangat digemari oleh penggemar sepak bola. Kekuatan tim yang lebih merata dan tidak semua pemain bintang bergabung ke klub-klub besar membuat liga ini selalu diwarnai berbagai kejutan.

Setiap minggu, terdapat 10 pertandingan yang berlangsung dan disiarkan secara langsung ke berbagai negara. Liga Primer Inggris memiliki setidaknya enam klub besar yang masing-masing memiliki suporter fanatik dari berbagai negara. Hal ini membuat banyak penonton berkomentar, berpendapat, dan meluapkan emosi melalui berbagai cara di antaranya melalui media sosial seperti Twitter, Facebook, maupun Instagram. *The Football Association* sebagai penyelenggara Liga Primer Inggris secara resmi mengeluarkan *hashtag* untuk masing-masing tim seperti pada tim Arsenal dengan *hashtag* #ARS, Chelsea dengan *hashtag* #CHE, Manchester United dengan *hashtag* #MUN, dan juga *hashtag* resmi untuk tiap pertandingan *Barclays Premier League* yaitu #BPL.

Twitter merupakan sumber informasi yang terbukti relevan untuk analisis berbagai macam keperluan, seperti klasifikasi terhadap tokoh publik (Hidayatullah, 2014), klasifikasi pating Twitter kemacetan lalu lintas kota Bandung (Rodiyansyah, 2012), dan prediksi hasil pertandingan bola (Kampakis dan Adamides, 2014). Twitter adalah sebuah situs *microblogging* yang bersifat *real-time*, dimana penggunaannya dapat saling berinteraksi melalui sebuah pesan pendek berjumlah 140

karakter yang sering disebut sebagai *tweet*. Twitter dengan mudah berkembang karena keterbukaan, fleksibilitas, dan kemudahan penggunaannya (Smith *et al.*, 2013). Twitter terbukti bisa digunakan untuk mendapatkan informasi yang berharga, karena pada beberapa kasus mampu menyajikan data yang tidak ditemukan pada sumber statistik lain (Kampakis dan Adamides, 2014). Sebagai contoh, pendapat penonton Liga Primer Inggris mengenai pemain bola melalui twitter terkadang berkebalikan dengan statistik performa pemain bola tersebut. Beberapa pemain boleh memiliki statistik yang bagus dalam sebuah pertandingan karena torehan gol, *assist*, dan kontribusi ke tim, tetapi penonton bisa jadi memiliki kebencian yang diekspresikan melalui Twitter karena sang pemain sering kali berbuat curang, memprovokasi, ataupun *diving*.

Twitter bukan satu-satunya media untuk menyalurkan pendapat para penggemar Liga Primer Inggris. *Professional Footballers' Association (PFA)*, organisasi pesepakbola profesional liga Inggris, menyediakan sebuah penghargaan pemain terbaik versi penonton (*Fans' Player of the Year*). Seluruh penonton yang menyetujui persyaratan *voting* pada situs skysports.com berhak melakukan *voting* terhadap pemain yang dianggap memiliki performa impresif dalam semusim terakhir.

Penelitian ini membahas tentang keterkaitan tweet mengenai pemain sepak bola liga Inggris untuk prediksi *Fans' Player of the Year*. Tweet yang dikumpulkan kemudian dikategorisasikan berdasarkan sentimennya. Hasil yang didapatkan kemudian dibandingkan dengan hasil *voting Fans' Player of the Year*.

1.2. Rumusan Masalah

Berdasarkan latar belakang di atas, ada beberapa pokok masalah yang akan dikaji dalam penelitian ini, antara lain :

1. Bagaimana membangun model analisis sentimen untuk mengklasifikasi tweet berdasarkan sentimen terhadap pemain bola liga Inggris dengan metode *Support Vector Machine* menggunakan data dari Twitter.
2. Apakah hasil analisis sentimen berpengaruh dalam menentukan prediksi hasil *Fans' Player of the Year*.

1.3. Batasan Masalah

Beberapa batasan masalah yang penulis gunakan antara lain:

- a) Penelitian ini hanya melakukan analisis terhadap tweet berbahasa Inggris.
- b) Pemain bola yang dinilai dalam penelitian ini adalah pemain yang ada dalam daftar sepuluh besar *Player Performance Index*/indeks performa pemain pada situs *premierleague.com*.
- c) Proses klasifikasi dilakukan pada level kalimat.
- d) Klasifikasi sentimen dibagi menjadi tiga kelas yaitu positif, negatif, dan netral.
- e) Klasifikasi sentimen tidak membedakan tweet berdasarkan sarkasme.
- f) Metode analisis sentimen menggunakan *Support Vector Machine* dengan kernel *linear*, polinomial derajat dua, dan RBF.

1.4. Tujuan Penelitian

Tujuan dari penelitian ini untuk membangun model analisis sentimen untuk mengklasifikasi tweet berdasarkan sentimen terhadap pemain bola liga Inggris sehingga bisa dijadikan prediksi *Fans' Player of the Year*. Selain itu juga untuk mengukur nilai akurasi penggunaan metode *Support Vector Machine*.

1.5. Manfaat Penelitian

Manfaat dari penelitian ini untuk mengetahui tanggapan penonton liga Inggris melalui Twitter terhadap pemain bola sehingga dijadikan acuan untuk *voting Fans' Player of the Year*.

1.6. Metodologi Penelitian

1. Studi Literatur

Melakukan studi dari literatur-literatur yang mendukung penelitian. Literatur dapat berupa buku, paper, jurnal, website resmi, dll. Informasi yang dikumpulkan adalah materi mengenai analisis sentimen dengan menggunakan SVM. Selain itu, juga dilakukan pencarian informasi dari penelitian - penelitian sebelumnya.

2. Analisis dan Perancangan

Mengumpulkan dan mengolah data-data yang berasal dari tweet yang berkaitan dengan Liga Primer Inggris. Perancangan sistem berdasarkan studi literatur yang dilakukan sebelumnya.

3. Implementasi

Mengimplementasikan rancangan sistem yang telah disusun dalam bahasa *Python* sehingga dihasilkan sebuah program untuk proses pengujian selanjutnya. Sistem manajemen basis data yang digunakan adalah MySQL. Library SVM yang digunakan adalah *Scikit-learn*.

4. Pengujian

Pada tahap ini dilakukan penghitungan *precision*, *recall*, dan *F1 score* terhadap model analisis sentimen dan hasil klasifikasi, serta hasil perbandingan metode ekstraksi fitur, pembobotan, dan algoritma klasifikasi yang digunakan.

5. Penulisan Laporan

Hasil penelitian dianalisis dan didokumentasikan ke dalam bentuk laporan sehingga bisa ditarik kesimpulan dari penelitian ini.

1.7. Sistematika Penulisan

Rincian dari penulisan laporan penelitian akan dijabarkan sebagai berikut:

1. BAB I PENDAHULUAN

Berisi tentang latar belakang penelitian dan permasalahan, rumusan masalah, batasan masalah, serta tujuan dan manfaat penelitian.

2. BAB II TINJAUAN PUSTAKA

Berisi tentang uraian sistematis tentang informasi hasil penelitian yang disajikan dalam pustaka dan menghubungkannya dengan masalah penelitian yang sedang diteliti sebagai acuan untuk melakukan penelitian.

3. BAB III LANDASAN TEORI

Berisi tentang teori yang berasal dari pustaka yang digunakan dalam penelitian ini.

4. BAB IV ANALISIS DAN RANCANGAN

Berisi tentang analisis permasalahan dan rancangan data, rancangan algoritma, dan rancangan pengujian.

5. BAB V IMPLEMENTASI

Berisi tentang implementasi rancangan dan demo program hingga menghasilkan output.

6. BAB VI HASIL PENELITIAN DAN PEMBAHASAN

Berisi tentang pembahasan hasil penelitian dan pengujian. Hasil penelitian berupa eksperimen untuk menentukan parameter terbaik,

untuk selanjutnya digunakan untuk melakukan proyeksi. Hasil proyeksi dari program dibandingkan dengan hasil proyeksi dari metode lainnya.

7. BAB VII PENUTUP

Berisi tentang kesimpulan dari penelitian dan saran akan penelitian selanjutnya.