

1. Pendahuluan

Latar Belakang

Anime merupakan suatu animasi yang diproduksi di Jepang dan ditayangkan dalam format film maupun series TV[1]. Pebisnis pada industri ini memerlukan cara untuk mengetahui seberapa bagus suatu anime salah satunya dengan menggunakan *review*. Anime yang bagus bagi penonton merupakan suatu indikator bagi para pebisnis dalam berinvestasi pada anime tersebut. *Review* yang diberikan penonton anime terkadang sangat sulit untuk dibedakan apakah bersifat positif maupun negatif. Analisis sentimen dengan kemampuannya dalam mengklasifikasi suatu teks diharapkan mampu dalam menyelesaikan masalah tersebut.

Sebagai pendukung teori pada penelitian ini diperlukan beberapa penelitian yang telah dilakukan sebelumnya. Salah satunya adalah penelitian yang dilakukan oleh Samah Mohammed AlSulaim dan Ali Mustafa Qamar[1] Penelitian tersebut menggunakan *dataset* yang sama menggunakan *Word2Vec* untuk mengekstrak fitur dan *CNN* sebagai metode untuk mengklasifikasi data *review* anime. Penelitian tersebut memiliki hasil yang sangat memuaskan yaitu *f1-score* sebesar 97%. Penelitian yang dilakukan oleh Manal Mohammed dan Nazila Omar[2] membandingkan penggunaan TF-IDF dan *Word2Vec* dengan hasil *f1-score* dengan menggunakan TF-IDF yaitu 82,6% dan *Word2Vec* sebesar 89,7%.

Untuk membedakan penelitian ini dengan penelitian sebelumnya[1] diperlukan penelitian yang serupa namun menggunakan metode klasifikasi yang berbeda. Salah satunya adalah penelitian yang dilakukan oleh Joan Santoso beserta rekannya[3] yang menggunakan *Word2Vec* namun menggunakan metode klasifikasi berbeda yaitu *Naïve Bayes* dan *dataset* yang berbeda pula. Pada penelitian tersebut menggunakan berita berbahasa Indonesia yang memiliki hasil yang memuaskan yaitu *f1-score* sebesar 94,31%. Penelitian yang dilakukan oleh Muhammad Ridho Fauzi[4] yang melakukan perbandingan antara penggunaan *Naïve Bayes* dan *KNN* dengan hasil akurasi *Naïve Bayes* sebesar 80% dan *KNN* sebesar 67,2%.

Pada penelitian ini penulis akan melakukan penelitian yang hampir sama dengan penelitian sebelumnya[1]. Penulis mempertahankan penggunaan *Word2Vec* dengan alasan pada penelitian[2] *Word2Vec* memang memiliki hasil yang lebih baik dibandingkan dengan TF-IDF. Untuk memberikan perbedaan dengan penelitian[1] penulis menggunakan metode *Naïve Bayes* yang pada penelitian[3] dengan mempertimbangkan hasil penelitian[4] yang menggunakan *Naïve Bayes* sebagai hasil terbaik.

Topik dan Batasannya

Topik pada penelitian ini yaitu mengetahui performa yang dihasilkan oleh sistem yang dibangun berdasarkan scenario yang digunakan pada penelitian ini. Batasan masalah pada penelitian ini yaitu pembuatan sistem analisis sentimen berfokus pada penggunaan *Word2Vec* sebagai metode ekstraksi fitur dan *Naïve Bayes* sebagai metode pengklasifikasian data. Data *review* anime yang digunakan berasal dari *MyAnimeList*. Data ini berisikan 1000 data *review* dengan 500 data berlabel negatif dan 500 data dengan label positif berbahasa Inggris.

Tujuan

Tujuan pada penelitian ini adalah membangun suatu sistem analisis sentimen pada *review* anime berbahasa Inggris dengan menggunakan *Word2Vec* dan *Naïve Bayes*. Model yang telah dibangun akan diukur performansinya untuk mengetahui sebaik apa model tersebut.

Organisasi Tulisan

Bab 2 menjabarkan berbagai penelitian yang telah dilakukan sebelumnya yang dapat menjadi referensi pada penelitian kali ini. Bab 3 menjelaskan bagaimana sistem yang akan dibangun dan juga metode-metode yang digunakan pada penelitian ini. Bab 4 menjabarkan hasil yang didapat pada penelitian ini dan menganalisis hasil penelitian tersebut. Pada bab 5 memberikan kesimpulan penelitian ini berdasarkan analisis yang dilakukan.