

1. Pendahuluan

Latar Belakang

Dengan berkembangnya Revolusi Industri 4.0, media sosial pun juga mengalami perkembangan. Media sosial adalah salah satu alat yang sering digunakan untuk mengekspresikan dan melaporkan berbagai aktivitas. Twitter adalah salah satu media sosial populer, yang memungkinkan *user* untuk membuat cuitan singkat “tweets”, mengikuti pengguna lain “follow”, mencari informasi populer “trending”, dan berkomunikasi dengan singkat “direct message” [1]. Dikarenakan fungsi-fungsi tersebut, banyak *user* Twitter yang hobi menonton film, menggunakannya untuk mencari informasi ataupun memberikan cuplikan singkat seputar film.

Film adalah salah satu media hiburan yang memiliki berbagai judul dan *genre*. Dikarenakan Revolusi Industri 4.0 ini, untuk mengakses berbagai film sangatlah mudah. Banyak perusahaan industri film telah mencapai pasar *direct-to-customer* untuk memanfaatkan kenyamanan hiburan berbasis internet [2]. Salah satu penyedia yang sering digunakan ialah Disney+. Dengan banyaknya judul dan *genre* yang ada, sistem rekomendasi dapat dimanfaatkan untuk membantu *user* dalam memilih film berdasarkan preferensinya. Selain dapat dimanfaatkan di penyedia *movie streaming*, sistem rekomendasi juga dapat diimplementasikan di penyedia *music streaming*, TV, *e-commerce*, dan lainnya [3].

Sistem rekomendasi adalah salah satu alat yang berguna untuk menyortir dan mengatasi informasi yang tidak teratur. Ide utama dibalik penggunaan sistem ini adalah untuk membangun hubungan antara produk, *user*, dan pengambilan keputusan dalam memilih produk, sesuai preferensi *user*. Collaborative Filtering (CF), Content-Based Filtering (CBF), dan Hybrid Filtering adalah tiga teknik populer dalam pengimplementasian sistem rekomendasi. CF merupakan teknik yang berdasarkan kemiripan pengguna dalam hal pemilihan atau pemberian penilaian pada suatu *item*. Sedangkan untuk CBF, berdasarkan kemiripan *item* dalam hal konten atau isi yang disukai oleh *user*. Hybrid Filtering, sesuai namanya, yaitu teknik yang menggabungkan dua atau lebih teknik lainnya [4], [5].

Pendekatan Hybrid Filtering adalah untuk mengatasi berbagai isu pada teknik lainnya, yaitu *cold start*, *overspecialization*, dan *sparsity*. Hybrid Filtering juga diadopsi untuk meningkatkan akurasi dan efisiensi dalam memproses rekomendasi [6].

Walaupun Hybrid Filtering lebih unggul dari teknik lainnya, pada langkah pengklusteran data poin, teknik ini memiliki masalah *gradient issue*. Untuk mengatasi masalah tersebut, penelitian ini akan menggunakan bobot dari CF dan CBF untuk diimplementasikan ke dalam Hybrid Filtering, yang berikutnya akan diklasifikasikan dengan Long Short-Term Memory (LSTM) [7], [8]. LSTM sendiri adalah salah satu varian dari jaringan yang menggunakan konsep sel memori dengan *self-connection* untuk menyimpan informasi sebagai *temporal state* [8], [9].

Pada penelitian ini, berbagai jurnal ilmiah yang terkait dengan topik, akan digunakan sebagai acuan dan batasan untuk metode dan teknik yang digunakan. Geetha dan kolega pada *projectnya* menjelaskan, bahwa perbandingan dan hasil dari sistem rekomendasi CF dan CBF menghadapi isu *content-independent*, dan *portfolio effect*, yaitu sistem akan memberikan rekomendasi *item* yang sudah dimiliki, seperti merekomendasikan film yang telah ditonton, memberikan saran lagu yang sudah disimpan, dan lainnya [10]. Hilmi Arfisko dan kolega juga mengalami isu yang sama pada CF dan CBF, yaitu *cold start* dan *sparsity*. Untuk menghindari isu tersebut, pendekatan Hybrid Filtering digunakan untuk mengatasi masalah tersebut. Teknik ini menghasilkan nilai *hit-rate* tertinggi (0.8042), melebihi CF (0.7904) dan CBF (0.7009). Teknik ini juga dapat menampilkan *item-item* rekomendasi yang relevan, baik, dan tersembunyi dengan nilai MAP 0,3159, dibandingkan dengan CF (0.3054) dan CBF (0.2219) [11]. Berikutnya, penelitian oleh Jerónimo dan kolega membuktikan, bahwa penggunaan Hybrid Filtering dengan pembobotan CF-CBF, dapat menghasilkan rekomendasi yang relevan, sesuai dengan referensi penelitian terdahulu mereka [12].

Pada kasus yang lain, penelitian oleh Chopra dan kolega mengindikasikan, bahwa pengimplementasian LSTM dalam Hybrid Filtering menghasilkan nilai performansi metrik yang tinggi, dengan 98.50% *Precision*, 98.83% *Recall*, 98.67% *F-Measure*, dan 98.66% *Accuracy* [13]. Penelitian oleh Nguyen dan kolega juga memdemonstrasikan, bahwa penggunaan LSTM pada Hybrid Filtering dapat menghasilkan rekomendasi yang tepat. Mereka bereksperimen dengan membandingkan LSTM, RNN, dan K-NN, pada sebuah sistem rekomendasi. Mereka menyimpulkan bahwa jaringan *LSTM* dapat membangun sistem rekomendasi yang spesifik dengan film yang relevan dengan pengguna. *LSTM* juga dapat meninggalkan hal-hal yang tidak berguna di masa yang akan datang, serta metode jaringan ini merupakan metode terbaru dan memiliki optimasi yang baik [14].

Berdasarkan referensi diatas, penggabungan Hybrid Filtering dan klasifikasi jaringan LSTM pada penelitian ini, akan menghasilkan metrik performansi yang tinggi dan *output* yang historis.

Topik dan Batasannya

Penelitian ini akan membuat sistem rekomendasi Hybrid Filtering dengan pembobotan CF dan CBF, yang akan diklasifikasikan dengan metode jaringan LSTM. Metrik performansi akan dihitung dengan Confusion Matrix,

berdasarkan nilai *precision*, *recall*, *accuracy*, dan *F1-Score*. Dataset yang digunakan berasal dari situs *web* Kaggle, yang mencakup *tweet* film terkait *platform* Disney+.

Tujuan

Penelitian ini bertujuan untuk membuat sistem rekomendasi Hybrid Filtering yang didasarkan dari pembobotan nilai CF dan CBF. Hasil rekomendasi akan diklasifikasikan dengan metode jaringan LSTM, dengan berbagai *optimizer*. Berikutnya, nilai performansi akan dihitung menggunakan Confusion Matrix, berdasarkan nilai *precision*, *recall*, *accuracy*, dan *F1-Score*.

Organisasi Tulisan

Berikutnya, jurnal ini berisikan studi terkait (referensi yang digunakan pada jurnal ini), sistem yang dibangun, evaluasi, dan kesimpulan.