

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Pada era digital ini, di mana terdapat banyak sekali pilihan secara daring dibutuhkan adanya penyaringan informasi agar menghindari pengguna internet dari masalah berlebihan informasi. Sistem rekomendasi dapat membantu dalam permasalahan ini dengan membantu menyediakan pengguna dengan konten dan layanan yang dikhususkan untuk pengguna tersebut.

Sistem rekomendasi adalah strategi pengambilan keputusan untuk pengguna dengan lingkungan data yang rumit[1]. Sistem rekomendasi bekerja cukup baik untuk merekomendasikan produk, film, musik dan lainnya. Metode-metode tradisional sistem rekomendasi termasuk *Content Based Recommendation* dan *Collaborative Filtering Recommendation* yang membuat rekomendasi berdasarkan hal yang sudah dibeli pengguna di masa lalu, tetapi kebanyakan model sistem rekomendasi bersifat statis. Namun kala ini orang lebih banyak menggunakan perangkat seluler untuk menelusuri produk secara daring, dan menghasilkan banyak *Behavioural Data* [2].

Contohnya seperti pada penelitian "A SVM-based Personal Recommendation System for TV Programs" [3], dibuatlah sistem rekomendasi personal untuk program televisi menggunakan pendekatan SVM. Pada penelitian ini dihasilkan sistem rekomendasi dengan akurasi tertinggi pada angka 92,5% dan menjamin kemampuan adaptif dinamis yang tinggi [3].

Pada tahun 1992 *Support Vector Machine* atau SVM pertama kali diperkenalkan oleh Boser, Guyon, dan Vapnik. SVM menjadi sangat populer karena dapat memberikan akurasi yang lebih baik dari pada *Neural Network* yang konvensional. SVM adalah algoritma yang biasa digunakan untuk klasifikasi dan regresi. SVM bekerja dengan cara membentuk *HyperPlane* yang optimal. *HyperPlane* sendiri bekerja sebagai alat pemisah untuk proses klasifikasi pada ruang tiga dimensi atau lebih [4].

Proses klasifikasi dengan metode SVM telah digunakan untuk membuat prediksi di berbagai bidang. Seperti pada buku "Cancer Genomics Proteomics", SVM dapat digunakan untuk mendeteksi subtype kanker dari *genomic feature patterns* atau memprediksi hasil dari prognosa, juga memprediksi keuntungan dari obat

yang digunakan [5]. Pada penelitian tersebut, data *Cancer genomics* adalah data berdimensi tinggi, heterogen, dan banyak mengandung *noise*, penerapan *SVM* terbukti akurat dengan nilai error 0% dari 34 data, walaupun penelitian tersebut tidak menjelaskan proses pembentukan kombinasi terbaik (Pemilihan *Kernel function*, nilai *C*, dan nilai *r*) [5].

Berdasarkan penelitian "*Supervised Machine Learning Algorithms: Classification and Comparison*" [6], dengan jumlah *dataset* yang besar maupun kecil *SVM* memiliki nilai akurasi dan presisi tertinggi, selain itu *SVM* membutuhkan waktu lebih sedikit untuk menghasilkan prediksi. Nilai akurasi *SVM* pada jumlah *dataset* yang besar (768 data) adalah 77,34% dengan waktu 0,09 detik, sementara untuk nilai akurasi *SVM* pada jumlah *dataset* yang kecil (384 data) adalah 72,92% dengan waktu 0,04 detik. Tingkat presisi algoritma *SVM* pada jumlah *dataset* yang besar (768 data) adalah 0,74, sementara tingkat presisi algoritma *SVM* pada jumlah *dataset* yang kecil (384 data) adalah 0,711 [6].

Objek penelitian yang dipilih ialah film dikarenakan film memiliki atribut yang dapat digunakan untuk proses klasifikasi. Dengan menggunakan *The Movie Database API* untuk meminta data film yang akan datang agar dapat direkomendasikan kepada pengguna. Maka dari itu, tujuan penelitian ini adalah untuk membuat sistem yang dapat mempelajari kebiasaan atau minat penikmat film sebagai pengguna atau *user*. Dengan mengumpulkan *Input Data* akan film yang diminati *user* dan menggunakan *SVM* untuk membuat model klasifikasi, lalu menerapkan *Hyper Parameter Optimization* untuk mencari *Kernel Function* yang paling cocok pada masalah ini untuk menghasilkan klasifikasi film yang akan di minati oleh *user* sebagai rekomendasi kepada *user* tersebut.

1.2 Rumusan Masalah

Berdasarkan latar belakang permasalahan yang sudah dijelaskan di atas, rumusan masalah yang akan digunakan dalam penelitian ini, yaitu sebagai berikut:

1. Bagaimana menerapkan algoritma *Support Vector Machine* agar sistem rekomendasi film dapat membantu *user* dalam pengambilan keputusan ?
2. Apakah *Kernel Function* yang terbaik dalam permasalahan ini?
3. Berapa akurasi dan *F1-Score* dari sistem rekomendasi ini ?

1.3 Batasan Permasalahan

Batasan masalah pada penelitian ini adalah sebagai berikut:

1. Sistem rekomendasi hanya akan berfokus pada satu objek yaitu film.
2. Sistem rekomendasi hanya berfokus pada film terbaru dalam rentang 1 tahun.
3. Parameter yang digunakan dalam *SVM* adalah sebagai berikut:
 - *Kernel Function* yang diuji adalah *RBF*, *Poly*, dan *Sigmoid* [7].
 - nilai *C* dengan rentang 10^{-3} sampai 10^3 [7].
 - nilai *Gamma* dengan rentang 10^{-3} sampai 10^3 [7].
4. *Attribute* film yang digunakan dari *The Movie Database* adalah sebagai berikut:
 - *adult* menunjukkan apakah film tersebut film dewasa atau bukan.
 - *genre id* menunjukkan genre apa saja yang dimiliki film tersebut.
5. Jumlah film awal yang digunakan untuk *knowledge* sistem adalah 10, 20, 50, 100, 200, dan 250 *Input Data*.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

1. Menerapkan algoritma *Support Vector Machine* agar sistem rekomendasi dapat mengenali karakteristik dari pilihan-pilihan *user*.
2. Menerapkan *Kernel Function* yang terbaik untuk masalah ini.
3. Mengukur akurasi dan *F1-Score* dari sistem rekomendasi telah dibuat.

1.5 Manfaat Penelitian

Melalui penelitian ini diharapkan beberapa manfaat sebagai berikut:

1. Memberikan *user* rekomendasi film bioskop terbaru yang sesuai dengan kebiasaan menonton *user*.
2. Mengetahui *Kernel Function* terbaik dan performa dari *SVM* untuk permasalahan ini.

1.6 Sistematika Penulisan

Sistematika penulisan laporan adalah sebagai berikut:

- Bab 1 PENDAHULUAN
Bab ini berisikan latar belakang masalah dari pemilihan judul Skripsi “Implementasi Algoritma Support Vector Machine Untuk Sistem Rekomendasi berdasarkan kebiasaan pengguna”, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.
- Bab 2 LANDASAN TEORI
Bab ini berisikan beberapa literatur yang mendukung penelitian terkait dengan penelitian yang dilakukan. Beberapa literatur tersebut ialah *Support Vector Machine*, Sistem Rekomendasi, Akurasi dan *F-1 Score*.
- Bab 3 METODOLOGI PENELITIAN
Bab ini berisikan analisis kebutuhan, perancangan sistem, *Flowchart*, skema kerja, struktur tabel, dan *Site Map* yang dijabarkan.
- Bab 4 HASIL DAN DISKUSI
Bab ini berisikan hasil implementasi yang dilakukan dari perancangan Sistem Rekomendasi. Hasil uji coba, hasil evaluasi dan *Heat Map* untuk menjawab rumusan masalah penelitian ini juga dijabarkan pada bab ini.
- Bab 5 SIMPULAN DAN SARAN
Bab ini berisikan simpulan dan hasil penelitian dari tujuan yang ingin dicapai pada penelitian ini, beserta saran untuk pengembangan aplikasi lebih lanjut.

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A