

BAB II

TINJAUAN PUSTAKA

2.1 Sistem Rekomendasi

Sistem rekomendasi adalah sistem yang berfungsi memberikan rekomendasi atau preferensi dari suatu benda kepada pengguna. Menurut (Konstan & Riedl, 2012), Sistem rekomendasi adalah sistem berbasis penyaringan informasi yang dikembangkan untuk mengatasi masalah terkait sejumlah besar informasi yang akan diproses. Contoh sistem rekomendasi bisa Anda temukan di beberapa aplikasi streaming seperti Spotify, Netflix, Prime Video dan YouTube.

Dari penelitian-penelitian tersebut terlihat bahwa sudah banyak penerapan sistem rekomendasi dalam dunia pendidikan dengan jangkauan yang sangat beragam. Sistem rekomendasi tidak hanya ditujukan kepada pelajar dan mahasiswa, tetapi juga digunakan oleh penyedia layanan pendidikan seperti universitas. Sistem rekomendasi memiliki keunggulan dalam pengambilan keputusan secara efisien dan berdasarkan kriteria.

Dalam proses ini, beberapa faktor penting dipertimbangkan, yaitu penghasilan orang tua, yang mempengaruhi kemampuan finansial siswa untuk menempuh program studi tertentu. asal sekolah, yang memberikan informasi mengenai kurikulum yang telah dipelajari dan kesiapan akademik. nilai lulusan, yang mencerminkan prestasi akademik dan kecocokan dengan program studi. Minat siswa, yang diidentifikasi melalui tes atau survei, sangat penting untuk menentukan bidang studi yang paling menarik dan sesuai bagi siswa, memastikan bahwa rekomendasi yang diberikan selaras dengan ketertarikan dan motivasi mereka. Pertimbangan faktor-faktor ini memastikan bahwa rekomendasi program studi disesuaikan dengan kebutuhan dan potensi siswa, memaksimalkan peluang mereka untuk sukses dalam pendidikan tinggi.

2.2 Support vector machine (SVM)

Support vector machine (SVM) adalah metode klasifikasi *machine learning* yang cukup populer digunakan untuk kasus klasifikasi (Cortes & Vapnik, 1995). Ide dari metode SVM adalah menemukan pemisah (*hyperplane*) yang memaksimalkan jarak antar kelas (Kancherla, Bodapati, & Veeranjanyulu, 2019).

Hyperplane adalah batas yang memisahkan dan mengklasifikasikan data ke dalam kelas-kelas yang berbeda, sedangkan *support vector* adalah titik dalam setiap kelas yang jaraknya paling dekat dengan *hyperplane*. Dalam kasus *nonlinier*, *hyperplane* model SVM tidak linier tetapi sesuai dengan distribusi data. SVM umumnya digunakan dalam skenario klasifikasi biner seperti regresi logistik Pendekatan OVO atau OVA dapat digunakan ketika ingin menggunakan SVM dalam skenario *multi-class*.

SVM mendefinisikannya sebagai klasifikasi linear atau klasifikasi nonlinier. SVM hanya dapat didefinisikan sebagai metode prediksi dengan mencari garis tertentu atau batas keputusan yang disebut *hyperplane* yang dengan mudah memisahkan kumpulan data atau kelas untuk menghindari overfit ekstra ke data. Penggunaan kernel bertujuan untuk mentransformasikan data ke ruang berdimensi tinggi, dengan menjadikan data non linier terpisah secara linier. Ada beberapa pilihan fungsi kernel yang dipakai pada sebuah aplikasi untuk mengatasi masalah pada metode Support Vector Machine (SVM) yaitu:

1. Linear Kernel

$$K(x_i, x) = x_i^T x \quad (1)$$

Dengan x_i merupakan data latih (training), x adalah data uji

2. Polynomial Kernel

$$K(x_i, x) = (\gamma(x_i^T x) + r)^p \quad (2)$$

Dengan x merupakan data latih (training), adalah data uji, p adalah derajat polinomial.

3. Radial Basis Function (RBF)

$$K(x_i, x') = \exp \left(-\frac{\|x - x'\|^2}{2\sigma^2} \right) \quad (3)$$

4. Sigmoid Kernel

$$K(x_i, x) = \tanh(\gamma(x_i^T x) + r) \quad (4)$$

Dengan x merupakan data latih (training), adalah data uji, r adalah koefisien.

2.3. Metode Pengembangan Sistem

Pengembangan sistem rekomendasi program studi berbasis machine learning bertujuan untuk membantu siswa SMA memilih program studi yang sesuai dengan kemampuan akademis, minat, dan latar belakang keluarga. Dalam pengembangan ini, metode 4D (Define, Design, Develop, Disseminate) diterapkan untuk memastikan proses berjalan sistematis dan efektif. Tahapan tahapan 4D yaitu:

2.3.1. Define

Pada tahap *Define* (Pendefinisian), dilakukan analisis kebutuhan pengguna, yang dalam hal ini adalah siswa SMA. Analisis ini mencakup pengumpulan informasi penting seperti nilai akademis, penghasilan orang tua, serta minat siswa. Selain itu, tahap ini juga mengidentifikasi masalah yang dihadapi siswa dalam memilih program studi yang sesuai, seperti kebingungan akibat minimnya informasi. Tujuan sistem kemudian ditetapkan, yakni untuk memberikan rekomendasi program studi berdasarkan data tersebut menggunakan algoritma Support Vector Machine (SVM).

2.3.2. Design

Tahap berikutnya adalah Design (Perancangan), yang melibatkan perancangan arsitektur sistem secara keseluruhan. Pada tahap ini, dirancang bagaimana sistem

akan menerima input berupa data siswa, memproses data tersebut, melatih model SVM, dan menghasilkan output berupa rekomendasi program studi. Selain itu, antarmuka pengguna juga dirancang agar mudah digunakan oleh siswa dalam memasukkan data dan menerima hasil rekomendasi.

2.3.3. *Development*

Tahap ketiga adalah Develop (Pengembangan), yang merupakan proses implementasi dari desain yang telah dibuat. Pada tahap ini, data siswa dikumpulkan dan diolah melalui proses normalisasi agar siap digunakan dalam pelatihan model SVM. Pengembangan model SVM dilakukan dengan melatih model menggunakan data yang telah diproses dan melakukan evaluasi performa model berdasarkan metrik seperti akurasi. Setelah model selesai dikembangkan, dilakukan pengujian sistem secara menyeluruh untuk memastikan fungsionalitasnya berjalan baik.

2.3.4. *Disseminate*

Tahap *Disseminate* (Penyebaran), di mana sistem yang telah dikembangkan disebarluaskan kepada pengguna. Sistem ini diluncurkan agar siswa dapat mulai menggunakannya untuk memilih program studi yang sesuai. Pelatihan pengguna juga dilakukan untuk membantu siswa memahami cara menggunakan sistem serta membaca hasil rekomendasi yang diberikan. Evaluasi dilakukan berdasarkan umpan balik dari pengguna, dan hasil evaluasi digunakan untuk perbaikan dan pengembangan sistem lebih lanjut.

2.4. Perangkat Lunak Pendukung

2.4.1. PHP

PHP Hypertext Preprocessor atau yang sering dikenal dengan PHP merupakan sebuah Bahasa pemrograman berorientasi pada objek. PHP ini termasuk kedalam tipe Bahasa pemrograman *Scripting Language* yang dirancang untuk pengembangan aplikasi web dinamis yang kode nya dapat ditanamkan ke dalam HTML. Bahasa ini berjalan pada sisi *Web Server*.

2.4.2. Java

Untuk dapat menjalankan sebuah website dinamis dengan menggunakan Bahasa pemrograman PHP, maka dibutuhkan *web server* yang dapat menjalankan kode PHP. *Apache* merupakan salah satu *web server* yang dapat dijalankan pada banyak sistem operasi (Unix, Linux, Windows, dan BSD). Protokol yang digunakan oleh *web server* umumnya HTTP dan HTTPS. *Apache* sendiri dikembangkan oleh *Apache Software Foundation* dan juga sebagai *Open Source Software*.

2.4.3. MySQL Database

Salah satu basis data yang banyak digunakan dan mudah dijalankan adalah *MySQL*. Basis data sendiri digunakan untuk menyimpan data secara dinamis yang akan ditampilkan pada aplikasi. *MySQL* mendukung perintah database berupa *Create*, *Read*, *Update*, *Delete*. Hal ini memungkinkan pengguna melakukan pengolahan data dengan mudah. *MySQL* termasuk ke dalam *Relational Database Management System* atau RDBMS.

2.4.4. Laravel

Laravel adalah salah satu framework PHP yang populer dan banyak digunakan oleh pengembang web. Framework ini menawarkan berbagai fitur unggulan seperti routing yang elegan, ORM (Eloquent) yang kuat, dan sistem templating (Blade) yang mudah digunakan. Dengan dukungan komunitas yang besar dan dokumentasi yang lengkap, Laravel memudahkan pengembang untuk membangun aplikasi web yang scalable, maintainable, dan efisien. Selain itu, Laravel juga menyediakan berbagai alat bantu seperti Artisan CLI, yang mempercepat proses pengembangan dan deployment aplikasi.

2.5. Penelitian Terkait

Pada pembuatan penelitian ini, terdapat beberapa inspirasi yang berasal dari penelitian yang terkait dengan permasalahan dan dijadikan sebagai referensi pada penelitian ini

Tabel 2. 1. Penelitian Terdahulu

No.	Judul	Penulis	Metode/ Algoritma	Hasil / kesimpulan	Dataset
1	PENERAPAN MACHINE LEARNING UNTUK PENENTUAN MATA KULIAH PILIHAN PADA PROGRAM STUDI INFORMATIKA	Fathorazi Nur Fajri, Abu Thalib, Wiwin Yuliana, (2022)	Neural Network	Pengolah data menggunakan Excel serta Google Collab, model yang dihasilkan mencapai akurasi 79% dalam memprediksi mata kuliah pilihan berdasarkan nilai yang diperoleh..	mengumpulkan data dari 70 mahasiswa Teknik Informatika yang mencakup 35 mata kuliah selama 6 semester
2	Model Klasifikasi Calon Mahasiswa Baru Untuk Sistem Rekomendasi Program Studi Sarjana Berbasis Machine Learning	Pratama, Ahmad R Aryanto, Rio Rizky Pratama, Arif Taufiq M, (2022)	Random Forest	model Random Forest yang dihasilkan mencapai tingkat akurasi 86%, presisi 84%, recall 86%, dan AUC 97%	seleksi database berdasarkan jumlah SKS, nilai IPK, dan status mahasiswa memberikan hasil yang akurat dan dapat diandalkan dalam proses rekomendasi program studi.
3	Sistem Rekomendasi Pemilihan Program MSIB Bagi Mahasiswa Pendidikan Informatika	Elnursa, Dian Budi Nofriana, Vesy Syamsuri, Agus Cahyani, Laili (2023)	content-based filtering	Evaluasi sistem menunjukkan tingkat presisi yang tinggi, dengan skor rata-rata 89,4% dalam pengujian, yang menegaskan keakuratan dan relevansi rekomendasi yang diberikan.	27 data mitra yang dikumpulkan, serta 5 data mahasiswa berdasarkan preferensi profil minat.
4	Sistem Pendukung Keputusan Pemilihan Tempat Kerja Praktek Bagi Mahasiswa Fakultas Teknik Elektro Universitas Telkom	Ibrahim Ikhsan Aditia, Roswan Latuconsina, Ashri Dinimaharawati	Support Vector Machine	Hasil pengujian algoritma Support Vector Machine yaitu mendapatkan nilai akurasi sebesar 50%.	nama mahasiswa, gender, jurusan, peminatan, domisili, nama perusahaan, alamat

		(2021)			
5	Sistem Rekomendasi Program Studi Sarjana Berbasis Machine Learning untuk Model Klasifikasi Calon Mahasiswa Baru	Anato, Ali Akbar, Yogi, Suprayuandi Pratama. (2023)	Random Forest (RF), Multinomial Logistic Regression (MLR), dan Support Vector Machine (SVM)	Akurasi tinggi tidak disebutkan dengan persentase.	Dataset ini mencakup informasi seperti nilai rapor, nilai Ujian Nasional (UN), jurusan SMA, minat, dan preferensi calon mahasiswa
6	Analisis Perbandingan Klasifikasi Topik Skripsi Mahasiswa menggunakan K Nearest Neighbor dan Support Vector Machine (Studi Kasus: Jurusan Sistem Informasi, Fakultas Ilmu Komputer, Universitas Brawijaya)	Fitria Yesisca, Dian Eka Ratnawati, Bayu Rahayudi	K-Nearest Neighbor And Support Vector Machine	Hasil klasifikasi berdasarkan judul serta judul abstrak metode SVM mendapatkan nilai akurasi 97.08%,	3 kategori yakni pengembangan, manajemen data dan informasi,