**Pengembangan Sistem Rekomendasi Buku Berbasis Data Google Books Menggunakan Content-Based Recommendation dan BERT  
  
Penerapan Machine Learning pada Web Rekomendasi Buku Goodreads Menggunakan Collaborative Filtering, Content-Based, dan Clusteringss**

**Latar Belakang**

Di era digital saat ini, jumlah buku yang tersedia secara online terus berkembang pesat. Google Books, sebagai salah satu platform terbesar untuk pencarian buku, menyediakan berbagai informasi tentang buku, rating, dan metadata lainnya. Banyaknya pilihan buku seringkali membuat pengguna kesulitan menemukan buku yang sesuai dengan minat mereka.

Sistem rekomendasi berbasis Content-Based Recommendation dan menggunakan model BERT (Bidirectional Encoder Representations from Transformers) dapat membantu pengguna menemukan buku yang relevan dengan preferensi pribadi mereka. Dengan menggunakan Google Books API, sistem akan mengambil data buku yang berkaitan dengan deskripsi, genre, rating, dan metadata lainnya untuk memberikan rekomendasi berbasis konten.

Batasan Masalah

1. Sumber Data  
   Data yang digunakan berasal dari Google Books API, yang memuat informasi buku, rating, dan deskripsi buku. Data akan diambil secara langsung dari API, dengan informasi berikut:
   * id: Identifikasi buku
   * title: Judul buku
   * authors: Pengarang buku
   * description: Deskripsi buku
   * categories: Kategori buku
   * averageRating: Rating rata-rata
   * ratingsCount: Jumlah rating
   * publisher: Penerbit
   * publishedDate: Tanggal terbit
   * language: Bahasa buku
2. Metode Rekomendasi  
   Sistem akan dibangun menggunakan Content-Based Recommendation, yang akan merekomendasikan buku-buku serupa berdasarkan kemiripan deskripsi, kategori, dan konten lainnya. Untuk memproses teks secara lebih mendalam dan akurat, sistem ini akan menggunakan BERT, model deep learning yang dapat menangkap konteks dan hubungan dalam teks dengan lebih baik.
3. Output Sistem  
   Hasil dari model Machine Learning akan diintegrasikan ke dalam aplikasi web interaktif dengan fitur-fitur berikut:
   * Rekomendasi buku untuk user tertentu: Memberikan rekomendasi buku berdasarkan teks yang relevan dengan buku yang sudah disukai atau dilihat oleh pengguna.
   * Pencarian buku serupa berdasarkan judul tertentu: Mencari buku lain yang mirip berdasarkan judul yang dimasukkan oleh pengguna.
   * Eksplorasi buku berdasarkan kesamaan konten: Menampilkan buku yang memiliki kemiripan konten berdasarkan deskripsi, kategori, dan atribut lainnya.
4. Lingkup Evaluasi  
   Evaluasi model dilakukan menggunakan metrik yang sesuai:
   * Akurasi Cosine Similarity untuk melihat seberapa mirip rekomendasi yang diberikan dengan buku yang diinginkan pengguna berdasarkan deskripsi.
   * Analisis manual terhadap hasil rekomendasi untuk memastikan relevansi buku-buku yang disarankan.
5. Tidak Mencakup
   * Sistem ini tidak mengintegrasikan data real-time dari Google Books.
   * Tidak ada proses login pengguna sesungguhnya; hanya simulasi rekomendasi berdasarkan ID pengguna yang tersedia.
   * Skala sistem hanya untuk tujuan demonstrasi dan penelitian, bukan produksi massal.

**Penerapan Machine Learning Random Forest/logistic regression untuk Prediksi Tingkat Depresi Mahasiswa Secara Online**

**Latar Belakang**

Kesehatan mental menjadi isu penting dalam dunia pendidikan, khususnya di kalangan mahasiswa. Tekanan akademik, tuntutan sosial, dan perubahan lingkungan dapat meningkatkan risiko gangguan kesehatan mental seperti depresi. Namun, deteksi dini terhadap potensi depresi masih belum banyak dilakukan secara efektif.

Dengan kemajuan teknologi, penerapan machine learning memungkinkan pembuatan model prediktif yang dapat membantu mendeteksi gejala depresi secara cepat dan akurat. Salah satu algoritma yang handal dalam klasifikasi adalah Random Forest, yang mampu mengelola data kompleks dan menghasilkan prediksi yang kuat.

Model prediksi ini dikembangkan berdasarkan dataset kesehatan mental mahasiswa dan dirancang untuk diintegrasikan ke dalam aplikasi web, sehingga dapat diakses dengan mudah oleh mahasiswa, konselor, maupun pihak kampus untuk upaya preventif terhadap kasus depresi.

**Batasan Masalah**

* Dataset yang digunakan adalah **"Student Mental Health"** dari Kaggle, dengan fitur seperti usia, jenis kelamin, program studi, dan kondisi kesehatan mental.
* Model machine learning yang digunakan adalah **Random Forest Classifier** tanpa optimasi hyperparameter lanjutan.
* Target prediksi terbatas pada **deteksi tingkat depresi** (apakah mahasiswa mengalami depresi atau tidak).
* Data preprocessing dilakukan dengan metode **Label Encoding** untuk mengubah data kategorikal menjadi numerik.
* Implementasi model diarahkan untuk penggunaan dalam **aplikasi web sederhana**, bukan untuk diagnosis medis profesional.
* Evaluasi model menggunakan metrik **accuracy, precision, recall, F1-score**, dan **confusion matrix**.
* Asumsi bahwa data input user nantinya memiliki format yang sesuai dengan data pelatihan model.

**Pengembangan Model Prediksi Risiko Serangan Jantung Menggunakan Random Forest / SVMClassifier Berbasis Web**

**Latar Belakang**

Serangan jantung merupakan salah satu penyebab kematian terbesar di dunia. Deteksi dini terhadap risiko serangan jantung menjadi sangat penting untuk mencegah komplikasi yang lebih serius.  
Namun, banyak orang yang sulit mengakses layanan kesehatan untuk melakukan pemeriksaan rutin.  
Seiring dengan perkembangan teknologi, penerapan machine learning untuk prediksi kesehatan menjadi solusi yang potensial.

Model prediksi berbasis data medis dapat membantu memberikan gambaran awal terhadap risiko serangan jantung berdasarkan faktor-faktor seperti tekanan darah, riwayat diabetes, kebiasaan merokok, hingga tingkat olahraga.  
Dalam penelitian ini, dikembangkan sebuah model prediksi risiko serangan jantung menggunakan algoritma Random Forest Classifier yang kemudian direncanakan untuk diimplementasikan dalam bentuk aplikasi web agar lebih mudah diakses oleh masyarakat luas.

**Batasan Masalah**

1. Dataset yang digunakan adalah "heart\_attack\_prediction\_dataset.csv" yang telah tersedia dan diolah.
2. Model hanya memprediksi dua kategori: "Berisiko" dan "Tidak Berisiko" mengalami serangan jantung.
3. Algoritma yang digunakan adalah Random Forest Classifier, dengan hyperparameter tuning menggunakan GridSearchCV.
4. Jika ditemukan ketidakseimbangan data pada target, dilakukan penanganan menggunakan SMOTE.
5. Evaluasi model dilakukan menggunakan akurasi, confusion matrix, classification report, cross-validation, dan feature importance.
6. Model akan di-export untuk digunakan dalam aplikasi berbasis web.
7. Model ini tidak menggantikan diagnosis medis profesional, melainkan hanya sebagai alat bantu prediksi awal.