

Prediksi Harga Smartphone Menggunakan K-Nearest Neighbors (KNN)

Identifikasi Persoalan :

Prediksi harga smartphone dengan menggunakan algoritma K-Nearest Neighbors (KNN) melibatkan sejumlah persoalan potensial. Dari kualitas data yang mungkin tidak lengkap atau mengandung outliers, pemilihan fitur yang relevan, hingga sensitivitas KNN terhadap skala data, setiap aspek memerlukan perhatian khusus. Penentuan parameter K yang tepat, upaya menghindari overfitting atau underfitting, serta evaluasi model yang baik melalui validasi, menjadi kunci kesuksesan. Dinamika pasar dan perubahan harga smartphone juga perlu dipertimbangkan, dengan pembaruan model secara berkala. Kendala interpretabilitas KNN dan ketersediaan sumber daya komputasi juga harus diatasi. Terakhir, konsistensi data dan pembersihan data yang teliti juga penting untuk memastikan prediksi yang akurat.

Research Questions :

1. Bagaimana kualitas data yang diperoleh memengaruhi akurasi prediksi harga smartphone dengan KNN, dan bagaimana mengatasi ketidaklengkapan data atau keberadaan outliers?
2. Apa saja fitur-fitur yang paling relevan dalam memprediksi harga smartphone ?
3. Bagaimana pengaruh skala data terhadap performa KNN, dan apa strategi terbaik untuk normalisasi atau standarisasi fitur-fitur yang memiliki rentang nilai yang berbeda?
4. Bagaimana pemilihan nilai K mempengaruhi hasil prediksi, dan bagaimana menentukan nilai K yang optimal??

Teori Terkait :

- **Machine Learning (ML)**

Machine Learning (ML), atau Pembelajaran Mesin dalam bahasa Indonesia, adalah bidang kecerdasan buatan yang berfokus pada pengembangan sistem yang dapat belajar dari data dan mengidentifikasi pola atau tren untuk membuat keputusan atau prediksi tanpa harus secara eksplisit diprogram. Tujuan utama dari machine learning adalah memberikan kemampuan komputer untuk memahami dan meningkatkan kinerjanya seiring berjalannya waktu, tanpa intervensi manusia yang terus-menerus.

- **K-Nearest Neighbors (KNN)**

K-Nearest Neighbors (KNN) adalah salah satu algoritma machine learning yang digunakan untuk klasifikasi dan regresi. Algoritma ini beroperasi dengan prinsip dasar bahwa suatu titik data akan memiliki label atau nilai yang serupa dengan mayoritas kategori tetangga terdekatnya. "K" dalam KNN merujuk pada jumlah tetangga terdekat yang akan diperhitungkan dalam menentukan prediksi untuk suatu data yang tidak diketahui.