

Klasifikasi Harga Smartphone Menggunakan K-Nearest Neighbors (KNN)

Identifikasi Persoalan :

Dalam melakukan klasifikasi harga smartphone dengan menggunakan algoritma K-Nearest Neighbors (KNN), beberapa persoalan yang perlu diidentifikasi melibatkan aspek kualitas data seperti kelengkapan dan label yang akurat, seleksi fitur yang relevan, normalisasi skala data, pemilihan nilai K yang optimal untuk jumlah tetangga terdekat, potensi overfitting atau underfitting, evaluasi model yang tepat, adaptasi terhadap perubahan dinamika pasar, interpretasi model yang kurang jelas, ketersediaan sumber daya komputasi, dan konsistensi data. Menanggulangi kendala-kendala ini termasuk pemilihan fitur yang cermat, normalisasi data, dan pemilihan parameter K yang tepat, serta memastikan pembaruan model yang teratur untuk mencerminkan perubahan harga dan dinamika pasar yang terus berubah. Selain itu, pembersihan data yang teliti dan evaluasi model yang hati-hati perlu dilakukan untuk meningkatkan akurasi dan kehandalan klasifikasi harga smartphone.

Research Questions :

1. Bagaimana kualitas data, termasuk kelengkapan dan ketepatan label, memengaruhi kinerja model KNN dalam klasifikasi harga smartphone ?
2. Apa saja fitur-fitur yang paling relevan dalam memprediksi klasifikasi harga smartphone, dan bagaimana cara memilih fitur yang optimal untuk meningkatkan performa model?
3. Bagaimana pengaruh skala data terhadap hasil klasifikasi harga menggunakan KNN, dan apa strategi terbaik untuk normalisasi atau standarisasi fitur-fitur dengan rentang nilai yang berbeda?
4. Bagaimana pemilihan nilai K dalam algoritma KNN mempengaruhi akurasi klasifikasi harga, dan bagaimana menentukan nilai K yang optimal untuk dataset smartphone yang bersangkutan?

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.