

ABSTRACT

Optimizing marketing strategies, particularly in price prediction, is a common challenge in the retail industry. This study aims to assist retail companies in predicting product prices based on market price trends more accurately through a machine learning approach, thereby avoiding decision-making based solely on estimations. Initially, the Random Forest model demonstrated the best performance with an R^2 value of 80%, MAPE of 25%, and RMSE of 1.07. After applying feature selection and feature importance, the model's performance significantly improved, achieving an R^2 of 92%, MAPE of 13%, and RMSE of 0.84. This indicates a 15.3% increase in prediction accuracy accompanied by a reduction in overall prediction error. In addition, the product recommendation system developed using a content-based filtering approach based on cosine similarity by considering similarities in product name and category attributes successfully achieved the highest similarity score of 0.42%. The results of this study show that machine learning, supported by feature optimization techniques, can significantly enhance decision-making accuracy in pricing and product recommendations within the retail sector, thus serving as a valuable tool for strategic business planning.

Keywords – Content-Based Filtering, Feature Importance, Feature Selection, Price Prediction, Random Forest

ABSTRAK

Pengoptimalan strategi pemasaran, khususnya dalam hal prediksi harga, merupakan tantangan umum di industri ritel. Penelitian ini bertujuan untuk membantu perusahaan ritel dalam memprediksi harga produk berdasarkan tren harga pasar secara lebih akurat melalui pendekatan *machine learning*, sehingga dapat menghindari pengambilan keputusan yang hanya didasarkan pada perkiraan. Pada awalnya, model *Random Forest* menunjukkan performa terbaik dengan nilai R^2 sebesar 80%, MAPE sebesar 25%, dan RMSE sebesar 1,07%. Setelah dilakukan *feature selection* dan *feature importance*, performa model meningkat signifikan dengan R^2 sebesar 92%, MAPE 13%, dan RMSE 0,84%. Hal ini menunjukkan peningkatan akurasi prediksi sebesar 15,3% disertai dengan penurunan tingkat kesalahan prediksi secara keseluruhan. Selain itu, sistem rekomendasi produk yang dikembangkan menggunakan pendekatan *content-based filtering* berbasis *cosine similarity* dengan mempertimbangkan kemiripan pada atribut nama dan kategori produk berhasil mencapai skor kemiripan tertinggi sebesar 0,42%. Hasil penelitian ini menunjukkan bahwa *machine learning*, yang didukung dengan teknik optimasi fitur, dapat secara signifikan meningkatkan ketepatan pengambilan keputusan dalam penetapan harga dan rekomendasi produk di sektor ritel, sehingga menjadi alat yang bernilai dalam perencanaan bisnis strategis.

Kata kunci - *Content-Based Filtering, Feature Importance, Feature Selection, Prediksi Harga, Random Forest*