Penelitian tentang pengaruh teknik normalisasi terhadap algoritma KNN telah banyak dilakukan untuk mengeksplorasi cara meningkatkan kinerja klasifikasi. Salah satu penelitian penting adalah studi oleh Alshdaifat et al. (2021) yang mengevaluasi dampak berbagai teknik normalisasi, termasuk Min-Max Normalization, Z-Score Normalization, dan Decimal Scaling Normalization, terhadap kinerja algoritma klasifikasi seperti Support Vector Machines (SVM) dan Artificial Neural Networks (ANN). Meskipun penelitian ini berfokus pada SVM dan ANN, temuan yang menunjukkan pengaruh signifikan teknik normalisasi terhadap performa algoritma klasifikasi juga relevan untuk algoritma K-Nearest Neighbor (KNN), mengingat KNN sangat sensitif terhadap skala fitur dalam perhitungan jaraknya. Penelitian ini dilakukan pada 18 dataset benchmark dari berbagai domain, termasuk kesehatan, bisnis, dan fisika. Hasilnya menunjukkan bahwa Z-Score Normalization sering kali memberikan hasil terbaik, khususnya karena kemampuannya menangani outlier, sedangkan Decimal Scaling Normalization dianggap kurang efektif dalam meningkatkan akurasi model secara keseluruhan. Temuan ini menjadi dasar penting untuk mengeksplorasi pengaruh normalisasi pada algoritma KNN.

SITASI: The Effect of Preprocessing Techniques, Applied to Numeric Features, on Classification Algorithms’ Performance

Penelitian lainnya dilakukan oleh Peshawa J. Muhammad Ali (2022), yang mengevaluasi dampak Min-Max Normalization terhadap kinerja algoritma K-Nearest Neighbor (KNN) dalam tugas regresi menggunakan delapan ukuran kesamaan, yaitu City Block, Euclidean, Chebychev, Cosine, Correlation, Hamming, Jaccard, dan Mahalanobis. Penelitian ini dilakukan pada lima dataset benchmark dari UCI Machine Learning Repository yang mencakup berbagai domain, seperti fisika, ilmu hayat, energi, teknik sipil, dan bisnis. Hasil penelitian menunjukkan bahwa Min-Max Normalization memiliki pengaruh signifikan pada ukuran kesamaan berbasis jarak, seperti City Block, Euclidean, dan Chebychev, terutama pada dataset dengan perbedaan besar dalam rentang fitur. Namun, teknik ini tidak memengaruhi ukuran kesamaan seperti Hamming, Jaccard, dan Mahalanobis, karena sifat matematis mereka tidak bergantung pada skala fitur. Selain itu, ukuran kesamaan Mahalanobis secara konsisten memberikan hasil terbaik di semua dataset, baik sebelum maupun setelah normalisasi.

SITASI: Investigating the Impact of Min-Max Data Normalization on the Regression Performance of K-Nearest Neighbor with Different Similarity Measurements