

INTISARI

Masalah klasifikasi data yang tidak seimbang, terutama dalam pemeliharaan prediktif, menimbulkan tantangan signifikan bagi prediksi yang akurat dan efisiensi operasional. Tesis ini memperkenalkan pendekatan terintegrasi yang menggabungkan Multi-Granularity Relabeled Undersampling (MGRU), Affinity Propagation (AP), dan Granular Ball Sampling (GBS) (MGRUAPGBS) untuk mengatasi tantangan ini. Algoritma MGRUAPGBS yang diusulkan menyempurnakan proses pengambilan sampel data, meningkatkan deteksi instans kelas minoritas dalam kumpulan data yang tidak seimbang. Pendekatan ini divalidasi menggunakan sembilan belas kumpulan data dari Repositori KEEL, membandingkan kinerjanya dengan metode pengambilan sampel lain melalui skor F1 dan Nilai Prediksi Negatif (NPV).

Hasil eksperimen menunjukkan bahwa metode MGRUAPGBS secara konsisten mengungguli metode pengambilan sampel lain dalam hal NPV, dengan peringkat tertinggi pada empat dari tujuh pengklasifikasi untuk Pohon Keputusan, Pohon Keputusan Gradient Boost, Hutan Acak, dan pengklasifikasi XGBoost. Algoritme ini juga menunjukkan kinerja yang kompetitif dalam hal skor F1, khususnya unggul dalam empat dari tujuh pengklasifikasi untuk K-Nearest Neighbors, Regresi Logistik, Support Vector Machine, dan pengklasifikasi XGBoost.

Selain itu, algoritma ini diterapkan pada studi kasus untuk pemeliharaan prediktif, yang menunjukkan kegunaan praktisnya dalam meningkatkan akurasi dan keandalan prediktif. Hasilnya menyoroti kemanjuran integrasi komputasi granular dan teknik pengelompokan dalam menangani data yang tidak seimbang, yang berkontribusi pada strategi pemeliharaan prediktif yang lebih tangguh dan andal.

Kata kunci: Klasifikasi data tidak seimbang, Pemeliharaan prediktif, Undersampling multi-granularitas yang diberi label ulang, Propagasi afinitas, Pengambilan sampel bola granular, Pembelajaran mesin, Pengambilan sampel data

ABSTRACT

The problem of imbalanced data classification, particularly in predictive maintenance, poses significant challenges for accurate prediction and operational efficiency. This thesis introduces an integrated approach combining Multi-Granularity Relabeled Undersampling (MGRU), Affinity Propagation (AP), and Granular Ball Sampling (GBS) (MGRUAPGBS) to address these challenges. The proposed MGRUAPGBS algorithm refines data sampling processes, enhancing the detection of minority class instances in imbalanced datasets. This approach is validated using nineteen datasets from the KEEL Repository, comparing its performance with other sampling methods through F1 score and Negative Prediction Value (NPV).

The experimental results demonstrate that the MGRUAPGBS method consistently outperforms other sampling methods in terms of NPV, ranking highest in four out of seven classifiers for Decision Tree, Gradient Boost Decision Tree, Random Forest, and XGBoost classifiers. It also shows competitive performance in terms of F1 scores, particularly excelling in four out of seven classifiers for K-Nearest Neighbors, Logistic Regression, Support Vector Machine, and XGBoost classifiers.

Additionally, the algorithm is applied to a case study for predictive maintenance, demonstrating its practical utility in improving predictive accuracy and reliability. The results highlight the efficacy of integrating granular computing and clustering techniques in handling imbalanced data, contributing to more robust and reliable predictive maintenance strategies.

Keywords: Imbalanced data classification, Predictive maintenance, Multi-granularity relabeled undersampling, Affinity propagation, Granular ball sampling, Machine learning, Data sampling.