UDC 004

Roman Kormysh, Master's Student,
Olena Chyzhmotria, Senior Lecturer,
Iryna Dmytrenko, Senior Lecturer
Zhytomyr Polytechnic State University

THE ROBUSTNESS OF THE NAIVE BAYES CLASSIFIER TO DATA IMBALANCES IN FIRST AID DATASETS

The increasing use of machine learning models in critical domains like healthcare and emergency response necessitates a deep understanding of their performance on non-ideal data. First aid datasets often suffer from severe data imbalance, where minor injuries are common but life-threatening events are rare. This skew can compromise a classifier's accuracy, particularly for the critical minority classes. The Naive Bayes classifier is a computationally efficient algorithm, but its core assumption of feature independence can make it vulnerable to data imbalance when the training data disproportionately favors one class over another [1]. This study conducts a comprehensive investigation into the robustness of the Naive Bayes classifier on imbalanced first aid datasets, evaluating its predictive performance and reliability on both majority and minority classes under a range of conditions.

Naive Bayes is rooted in Bayes' Theorem, which computes the posterior probability for a class given a feature vector. The algorithm's "naive" assumption of conditional independence simplifies calculations, making it fast. However, in the presence of severe class imbalance, the algorithm's prior probability estimation can become heavily skewed. A model might learn to classify all instances as the majority class to achieve high overall accuracy, effectively ignoring the critical minority class [2]. This is catastrophic in first aid applications, where misclassifying a rare but critical emergency could lead to delayed or inappropriate treatment. Therefore, our analysis will use more robust performance metrics, including precision, recall, and the F1-score, and a detailed confusion matrix to quantify the model's ability to correctly identify true positives and avoid false negatives, which is paramount for safety-critical systems.

To systematically evaluate the classifier's robustness, this research will use a simulated first aid dataset with varying degrees of class imbalance, from moderate (10:1) to extreme (100:1). The dataset will mimic realistic scenarios, including features like patient vitals, reported symptoms, and external factors. A series of controlled experiments will be conducted to measure the impact of different mitigation strategies, including both data-level and algorithm-level approaches.

Data-level techniques will include re-sampling methods like random undersampling of the majority class and oversampling of the minority class using SMOTE (Synthetic Minority Oversampling Technique) [3]. We will also investigate algorithm-level techniques, specifically cost-sensitive learning, which modifies the algorithm to assign a higher misclassification cost to errors on the minority class. A false negative (failing to identify a critical condition) can be assigned a cost much greater than a false positive. We will conduct several key experiments: first, we will evaluate the Naive Bayes classifier on the raw, imbalanced dataset to establish a baseline performance. Next, we will compare its performance on datasets modified with undersampling and SMOTE. We will then analyze the impact of cost-sensitive learning. Finally, we will explore a hybrid approach by combining the most effective data and algorithm-level methods to determine if they yield superior performance. [4]. By understanding the specific limitations of the Naive Bayes classifier, we can make informed decisions about its suitability for various applications.

In conclusion, the robustness of the Naive Bayes classifier to data imbalances is a critical consideration for its application in real-world first aid and medical datasets. This study provides a comprehensive framework for assessing its performance and evaluating various mitigation techniques. By focusing on metrics beyond overall accuracy, we can gain deeper insights into the algorithm's behavior, particularly concerning the correct identification of rare yet critical events. The outcomes of this research will provide actionable, data-driven insights for building reliable and trustworthy AI systems in the emergency response domain, ensuring that artificial intelligence enhances, rather than compromises, patient outcomes [5]. Future work could extend this study to evaluate the performance of other classifiers on similar imbalanced datasets, offering a broader comparative analysis to guide model selection in critical applications.

References:

1. Wu S. & He S. C., The Impact of Class Imbalance on Machine Learning Classifiers: A Comparative Study. – 2023. – P. 45-50.

2. Li Q. & Zhang Y., Understanding and Mitigating the Effects of Class Imbalance on Naive Bayes Classifiers. – 2022. – P. 112-118.

3. Chawla N. V., Bowyer K. W., Hall L. O. & Kegelmeyer W. P., SMOTE: Synthetic Minority Oversampling Technique. – 2002. – P. 321-331.

4. Brown A., Chen J. & Wang L., Performance of Machine Learning Models in Emergency Medical Systems: A Critical Review. – 2021. – P. 78-85.

5. Miller K. & Jones T., Building Trustworthy AI for Healthcare: Ethical and Technical Considerations. – 2024. – P. 20-25.