Digital Health and Informatics Innovations for Sustainable Health Care Systems
J. Mantas et al. (Eds.)

doi:10.3233/SHTI240532

A Conformal Prediction Approach to Enhance Predictive Accuracy and Confidence in Machine Learning Application in Chronic Diseases

Christina PAPANGELOU[a,1], Thomas CHATZIKONSTANTINOU[a], Kostas STAMATOPOULOS[a], Anastasia CHATZIDIMITRIOU[a] and Evangelia MINGA[a]
[a] *Institute of Applied Biosciences, Center for Research and Technology Hellas, Thessaloniki, Greece*

Abstract. Heterogeneity in chronic malignancies raises an increasing interest for the integration and study of predictive models. This study presents a machine learning model approach to predict outcomes and improve their trustworthiness in multi-factorial diseases with highly heterogeneous outcomes, like Chronic Lymphocytic Leukemia (CLL). We incorporated Conformal Prediction to quantify our models uncertainty, and generate confident personalized prediction outcomes that can be integrated into clinical practice.

Keywords. Chronic Malignancies, Conformal Prediction, Machine Learning

1. Introduction

In a paradigmatic hematological malignancy, CLL, where the clinical course and outcome is highly heterogeneous [1], the importance of accurate and reliable predictions cannot be overstated. Uncertainty quantification techniques, such as Conformal Prediction, could play a crucial role, providing predictions and also measures of confidence or uncertainty associated with those predictions. This can enhance the reliability of predictions to foster a more personalized approach in patient care.

2. Methodology

Based on our medical expertise in hematological diseases and experience in data collection and analysis of biomedical data for research purposes, we utilized an anonymized dataset of 950 CLL patients and used data bootstrapping and perturbation techniques to generate a synthetic dataset of 2500 cases in order to explore the integration of AI techniques in predicting clinical outcomes for CLL. We developed prediction models for survival outcomes and time-to-second treatment. Our methodology included

[1] Corresponding Author: Christina Papangelou, INAB|CERTH. 6th km Charilaou-Thermi Rd, Thessaloniki, Greece; E-mail: papaggeloux@certh.gr.

rigorous pre-processing, feature selection, and algorithm optimization to ensure model robustness. Random Forest Classifier (RF) and Linear Regression model (LR) revealed as the top-performing algorithms for survival outcomes and time-to-second treatment, respectively. In order to enhance the predictive accuracy and to adopt a more patient-centric approach, providing reliable confidence estimates for our predictions, we used Conformal Prediction (CP) [2,3]. Our objective was to generate larger prediction sets for cases difficult-to-predict by AI models [4], aiming to mitigate the risk of erroneous decisions in clinical practice.

3. Results

To predict the survival outcome, the RF classifier was trained and evaluated on 20% of the dataset, achieving an accuracy of 89.53% (53 error cases). By integrating CP for a 95% desire level of confidence and the hinge loss as the non-conformity measure, our model achieved an empirical coverage (the average frequency of the true label falls in prediction region) of 95.4%. CP flagged 98 patient cases as uncertain, demanding classification from a physician. Notably, 33/53 erroneous predictions of RF model flagged as uncertain by CP. The LR model, that was developed to predict the time between first- and second-line treatment, was tested on 20% of the dataset and generated predictions with 0.3 Mean Squared Error. CP integration applied for a 95% desire level of confidence. Non-conformity scores calculated on 20% of our training data, used to calculate the 95% quantile (q), that was equal to 1.17 (95% of the calibration set has a deviation value from the actual time value below the 1.17). Finally, after adding and subtracting the q from the predicted value, the prediction interval for the test set achieved a 95.1% empirical coverage.

4. Discussion and Conclusions

Uncertainty quantification techniques utilizing a generous amount of informative data can fill the gap in trust between AI and medical experts. Our approach supports that conformal prediction improves decision-making in multi-factorial diseases with highly heterogeneous outcomes like CLL, by offering multiple predictions in scenarios where the AI model cannot confidently provide singular predictions. We managed to mitigate the risk of erroneous predictions and identified complex patient cases that require personalized medical care.

References

[1] Kipps TJ, Stevenson FK, Wu CJ, Croce CM, Packham G, Wierda WG, et al. Chronic lymphocytic leukaemia. Nature Reviews Disease Primers [Internet]. 2017 Jan 19;3(16096):16096. doi:10.1038/nrdp.2016.96

[2] Vovk V, Gammerman A, Shafer G. Algorithmic Learning in a Random World. Springer eBooks. Springer Nature; 2005. doi: 10.1007/b106715

[3] ShaferGlenn, & VovkVladimir. (2008). A Tutorial on Conformal Prediction. Journal of Machine Learning Research. https://doi.org/10.5555/1390681.1390693

[4] Angelopoulos, A. N., & Bates, S. (2021). A Gentle Introduction to Conformal Prediction and Distribution-Free Uncertainty Quantification. ArXiv (Cornell University). https://doi.org/10.48550/arxiv.2107.07511