Empowering Glioma Prognosis With Transparent Machine Learning and Interpretative Insights Using Explainable AI

Abstract

Gliomas, particularly Glioblastoma Multiforme (GBM), are aggressive brain tumors requiring precise prognosis methods. This project employs machine learning (ML) and explainable artificial intelligence (XAI) techniques to develop a reliable and interpretable glioma prediction system. Using a dataset of 889 patients, key predictors such as IDH1 mutation and age at diagnosis were identified through Pearson's correlation, mutual information, and principal component analysis (PCA). ML models, including XGBoost, Random Forest, and LightGBM, along with deep learning architectures (ANN, CNN), were evaluated. XAI tools—SHAP, LIME, ELI5, and QLattice—were used for interpretability. The XGBoost model achieved 88% accuracy and an AUC of 92%, providing a transparent tool for clinical decision-making.

1 Background and Motivation

Gliomas are the most common primary brain tumors, with GBM being the deadliest subtype. Accurate and early prognosis is critical for effective treatment planning and patient care. Traditional methods rely heavily on manual interpretation of medical and genetic data, which can be inconsistent and time-consuming. This project aims to leverage machine learning for automated prediction and explainable AI to ensure trust and interpretability in model outputs. By bridging computational power and clinical applicability, this work offers a pathway to personalized treatment planning.

2 Methodology

2.1 Data Description and Preprocessing

The dataset comprises 889 patients with 23 clinical and genetic features. Preprocessing steps included:

- Balancing the dataset.

- Standardization for consistent scaling.

- Feature selection using Pearson's correlation, mutual information, and PCA.

2.2 Machine Learning Models

Multiple ML and deep learning models were employed:

- **ML Models**: Random Forest, XGBoost, LightGBM, CatBoost.

- **Deep Learning Models**: Artificial Neural Networks (ANN) and Convolutional Neural Networks (CNN).

An ensemble stacking method combined individual model outputs for enhanced performance.

2.3 Explainable AI (XAI) Techniques

To interpret the results, four XAI techniques were implemented:

- **SHAP**: Quantified feature contributions using game theory principles.

- **LIME**: Provided local explanations by approximating black-box models with interpretable ones.

- **ELI5**: Simplified complex predictions for user understanding.

- **QLattice**: Explored feature interactions through a probabilistic graphical model.

2.4 Evaluation Metrics

Performance was evaluated using accuracy, precision, recall, F1-score, and AUC.

3 Implementation Details

The models and workflows were implemented using Python and libraries like scikit-learn, TensorFlow, and SHAP. Training utilized an 80:20 train-test split, with hyperparameter tuning conducted through grid search.

4 Results

- The XGBoost model achieved the best performance with 88% accuracy and an AUC of 92%.

- Key features identified for prediction included IDH1, TP53, age at diagnosis, and EGFR.

- XAI methods provided clear explanations, improving trust and usability in clinical settings.

5 Discussion and Conclusion

This project demonstrated the potential of transparent machine learning systems in glioma prognosis. By combining predictive accuracy with interpretability, the developed system bridges the gap between computational techniques and practical medical use. Future work involves expanding the dataset, incorporating advanced imaging features, and validating models in diverse clinical environments.