Classification of Gravitational Wave Events by Source Type

Harshith V. H.[1]

[1]Department of Physics and Computer Science, Kristu Jayanti College, Bengaluru, India

March 5, 2025

Abstract

This study investigates the classification of gravitational wave (GW) events into three distinct categories: Binary Black Hole (BBH), Binary Neutron Star (BNS), and Neutron Star-Black Hole (NSBH) mergers. Leveraging machine learning techniques, we analyze astrophysical parameters sourced from the Gravitational Wave Open Science Center (GWOSC) to develop robust classifiers for distinguishing these event types. Our results indicate that ensemble learning methods outperform other approaches, achieving high accuracy. This work establishes a foundation for future advancements through deep learning and sophisticated sampling techniques.

1 Introduction

Gravitational wave astronomy has provided unprecedented insights into the dynamics of compact object mergers, including BBH, BNS, and NSBH systems. Precise classification of these events is essential for advancing our understanding of astrophysical populations and optimizing detection methodologies. This project employs supervised machine learning to classify GW events based on the properties of the merging objects.

2 Methodology

2.1 Data Preprocessing

- Class Labeling:
  - BBH: Both objects have masses $\geq 3\,M_\odot$.
  - NSBH: One object is a black hole ($\geq 3\,M_\odot$) and the other a neutron star ($< 3\,M_\odot$).
  - BNS: Both objects have masses $< 3\,M_\odot$.
- Handling Missing Data: Missing values (NaN) were imputed using median values.

- Feature Encoding: Class labels were encoded using LabelEncoder:
  - BBH = 0, NSBH = 1, BNS = 2.
- Data Splitting: The dataset was divided into 70% training and 30% testing sets.

2.2 Addressing Class Imbalance

To mitigate significant class imbalance, Random Over-Sampling (ROS) was employed to enhance the representation of minority classes prior to model training.

2.3 Model Training and Evaluation

The following machine learning models were trained:

- AdaBoost Classifier
- Gradient Boosting Classifier
- Logistic Regression
- Random Forest Classifier
- XGBoost (XGB)

Evaluation Metrics:

- Accuracy
- Precision, Recall, and F1-Score
- Confusion Matrix Visualization

3 Results

The Random Forest, XGBoost, and AdaBoost models achieved 100% accuracy, precision, and recall, indicating successful differentiation of event types in the test set. However, these results may reflect overfitting due to the limited test set size (18 instances: 17 BBH, 1 NSBH, 0 BNS). The Gradient Boosting model failed to classify NSBH events, resulting in poor recall and F1-scores. Logistic Regression exhibited a convergence warning, likely due to insufficient iterations; increasing the max$_{i}terparametercouldaddressthisissue$.

4 Discussion and Future Work

Although the results are promising, the small test set constrains generalizability. Future enhancements include:

- Implementing Synthetic Minority Over-sampling Technique (SMOTE) for improved class balancing.
- Advancing feature engineering to capture more nuanced patterns.
- Exploring deep learning architectures to enhance generalization.

These improvements are expected to yield robust performance on larger, more diverse datasets.

5 Conclusion

This study underscores the efficacy of ensemble methods in classifying gravitational wave events. Despite challenges related to data imbalance and sample size, the results demonstrate the potential of machine learning for astrophysical classification tasks.

Acknowledgements

We express our gratitude to the Gravitational Wave Open Science Center (GWOSC) for providing the data and to the developers of open-source machine learning libraries that facilitated this research.