ARRYTHMIA PREDICTION MACHINE LEARNING MODEL

Onwusika Somkenechukwu Ikechukwu,

Revised: 3rd October, 2024.

Abstract:

Cardiac arrhythmias, such as atrial fibrillation (Nesheiwat et al., n.d.), pose significant health risks and are often difficult to detect accurately using conventional diagnostic methods (*ECG Interpretation of Arrhythmias [TUSOM | Pharmwiki]*, 2022). This project aims to develop a machine learning-based application that leverages electrocardiograph (ECG) signals to identify cardiac arrhythmias. By analysing differences between normal sinus rhythm and atrial fibrillation patterns, a detection algorithm is trained using a labelled ECG dataset from the MIT-BIH Arrhythmia Database (Moody & Mark, 2001).

The project utilizes several key Python libraries, including NumPy, Pandas, and SciPy for data manipulation and employs machine learning libraries such as TensorFlow to implement and evaluate models. The methodology involves preprocessing raw ECG data to remove noise, segmenting the data into relevant intervals, and then training a machine learning model to recognize atrial fibrillation patterns. Performance metrics such as accuracy and precision will be used to determine the effectiveness of the model.

This project provides a foundation for creating automated arrhythmia detection systems and will be extended to include other arrhythmias such as ventricular tachycardia (Whitaker et al., n.d.) in future versions. The developed model could be integrated into portable ECG devices for real-time heart health monitoring, potentially improving early diagnosis and patient outcomes. Furthermore, this could ensure greater access to quality healthcare as this could replace other alternative methods currently in place, which are expensive.

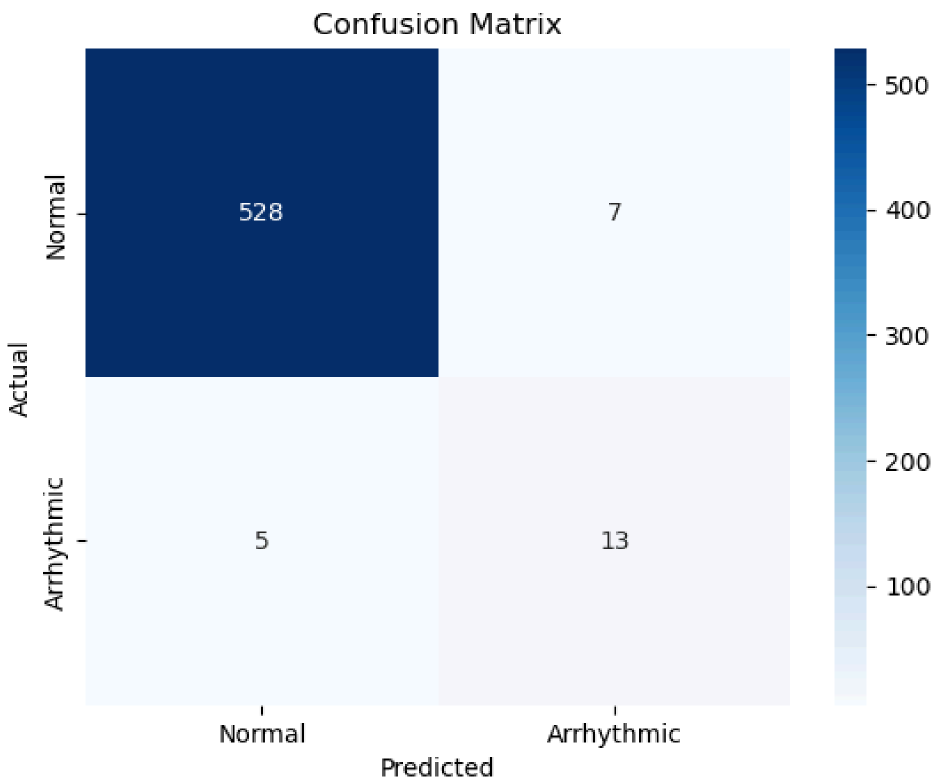
Findings:

The project, a machine learning model made use of the SVC classification library in order to identify arrhythmias based off of existing databases, such as the MIT-BIH Arrhythmia Database (Moody & Mark, 2001) which was used to carry out this research. It used the following key metrics to adjudge its results: Overall Accuracy, Precision, Recall (which was the ratio of true positives to the total actual positives), and the F-1 Score, which is the harmonic mean of the Precision and Recall categories. After multiple attempts at optimization, the machine learning model produced the following results:

	precision	recall	f1-score	support
Class 0 (Normal)	0.99	0.99	0.99	535
Class 1 (Arrhythmic)	0.65	0.72	0.68	18
accuracy			0.98	553
macro avg	0.82	0.85	0.84	553

weighted avg	0.98	0.98	0.98	553

The model produced an overall accuracy of 98%, with it being able to correctly identify regular heartbeats in 99% of all instances, and an accuracy of 65% when identifying arrhythmic heartbeats. In addition to this, a confusion matrix was used to procure a breakdown of the results for ease of viewing, as attached below:



From this matrix, we can identify that the model was able to almost fully predict all five hundred and thirty-five cases of regular heartbeats, only failing to correctly account for seven instances. When predicting the arrhythmic beats, it correctly predicted thirteen out of the eighteen heartbeats. While this correlates to a lower percentage of success when compared to the degree of accuracy with normal heartbeats, the reduced sample size must be taken into account as the model may not have been exposed to enough data to make better informed decisions.

Conclusion:

With this being the initial version of this study, it can be said that the model predicted to a great degree of accuracy, given the bias placed within the training data. In further editions the model would make use of more expansive databases which can supply it with more cases of arrhythmic cycles in order to make fully informed decisions.

Once this model has achieved a degree of accuracy suitable for clinical usage and trials, it can be exported to a network module or application for general public usage.

Machine Learning Model:
Below is a link to the machine learning model, which was made using Jupyter Notebooks:
https://github.com/onwusikasomkenechukwu/Arrhythmia-Detection-ML-Project

References:

ECG Interpretation of Arrhythmias [TUSOM | Pharmwiki]. (2022, April 27). TMedWeb. Retrieved October 2, 2024, from https://tmedweb.tulane.edu/pharmwiki/doku.php/ecg_interpretation_of_arrhythmias

Goldberger, A., Amaral, L., Glass, L., Hausdorff, J., Ivanov, P. C., Mark, R., ... & Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation [Online]. 101 (23), pp. (2005, February 24). PhysioNet. Retrieved October 2, 2024, from https://physionet.org/content/mitdb/1.0.0/

Moody, G. B., & Mark, R. G. (2001). *Moody GB, Mark RG. The impact of the MIT-BIH Arrhythmia Database. IEEE Eng in Med and Biol 20(3):45-50 (May-June 2001). (PMID: 11446209)*. Retrieved 10 2, 2024, from http://ecg.mit.edu/george/publications/mitdb-embs-2001.pdf

Nesheiwat, Z., Goyal, A., & Jagtap, M. (n.d.). *Atrial Fibrillation - StatPearls*. NCBI. Retrieved October 2, 2024, from https://www.ncbi.nlm.nih.gov/books/NBK526072/

Whitaker, J., Wright, M. J., & Tedrow, U. (n.d.). *Diagnosis and Management of Ventricular Tachycardia - PMC*. NCBI. Retrieved October 2, 2024, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10541285/