

Deep learning assisted heart arrhythmia detection

*Johnathan Fernandes¹, Sarthak Chudgar², Harshal Dharap³,
Aneesh Poduval⁴*

¹Department of Instrumentation and Control Engineering, Vishwakarma Institute of Technology Pune, Maharashtra, India

²Department of Instrumentation and Control Engineering, Vishwakarma Institute of Technology Pune, Maharashtra, India

³Department of Instrumentation and Control Engineering, Vishwakarma Institute of Technology Pune, Maharashtra, India

⁴Department of Instrumentation and Control Engineering, Vishwakarma Institute of Technology Pune, Maharashtra, India

E-mail: johnathan.fernandes17@vit.edu

Abstract. Heart arrhythmia, or irregular heart rhythm, is an extremely common heart affliction experienced by a large percent of the world's population every day, mostly going unnoticed. However, if left unchecked for an extended period of time, it poses an inherent risk to human life. Advancements in technology have enabled us to leverage the awesome computational power of graphics processing units in parallel in order to derive solutions to real life medical issues, by analyzing tremendous amounts of data in a relatively short amount of time. Owing to their ability to parse huge amounts of data and quickly perform multiple complex computations in parallel, machine learning algorithms have repeatedly and consistently outperformed humans in tasks such as pattern recognition and data analysis. Through this research project, we seek to contribute to the medical field by implementing deep learning technology along with machine learning algorithms into a system which can detect heart arrhythmias from electrocardiogram (ECG) reports quickly and effectively.

1. Introduction

“Heart Arrhythmia” is an umbrella term referring to a multitude of issues caused by irregular electrical impulses to the heart, causing it to beat too fast, too slow, or irregularly. The most common type of arrhythmia is “Atrial Fibrillation”, estimated to be present in 12.1 million people in the United States by 2030 [18,5]. In 2018 alone, atrial fibrillation was mentioned on 175,326 death certificates and was the underlying cause of death in 14.74% of those cases [3]. Technology today is progressing in ways that make detection of such conditions faster, less invasive and more accessible to ordinary citizens. In a study conducted by Apple [17], out of 400,00 participants, about 2,000 subjects had irregular pulse. A third of these cases were confirmed to have atrial fibrillation, while about 84 percent of them turned out to be atrial fibrillation episodes. This is just one example of the amazing performance in embedded systems brought about by recent advancements in technology, which raises the bar for traditional computational systems as well.

“Deep learning” is a subset of machine learning which primarily focuses on artificial neural networks, which aim to replicate the way the human brain processes data via “representation learning”. Advancements in computing technology on both hardware and software fronts have led to a boom in popularity of deep learning among enthusiasts and professionals alike.

In recent times, deep learning methods such as convolutional neural networks have made enormous progress in the medical field, most notably in cases involving medical imaging. Given enough computational power, they have been shown to even outperform human experts in their relevant areas of expertise [23].

Through this research project, we have implemented a deep learning system which, after suitable feature engineering and hyperparameter tuning, can detect arrhythmic heartbeats with an accuracy of 91.52%.

2. Related Work

We surveyed a multitude of papers to familiarize ourselves with progress in the field. Several papers used a traditional machine learning approach [25, 9, 2, 27, 15, 24, 19, 21, 20], while others utilized a more complex deep learning approach [8, 22, 1, 28, 10, 23, 16, 26, 14, 4, 12]. Some have even applied additional signal processing techniques to enhance model performance [24]. Deep learning methods predictably outperform traditional algorithms in most cases, with the obvious trade-off of computational complexity, especially in the case of convolutional neural networks.

Rajpurkar, Pranav, et al. [23] compared the performance of their convolutional neural network model to that of professional cardiologists, and as such will be taken as the baseline for our project. Mustaqeem, Anam, et al. [21] used a dataset from the University of California Irvine Machine Learning Repository [7]. This dataset was compiled using a 12 lead ECG machine and contains 15 arrhythmia classes. We chose to use this dataset for our project.

3. Methodology

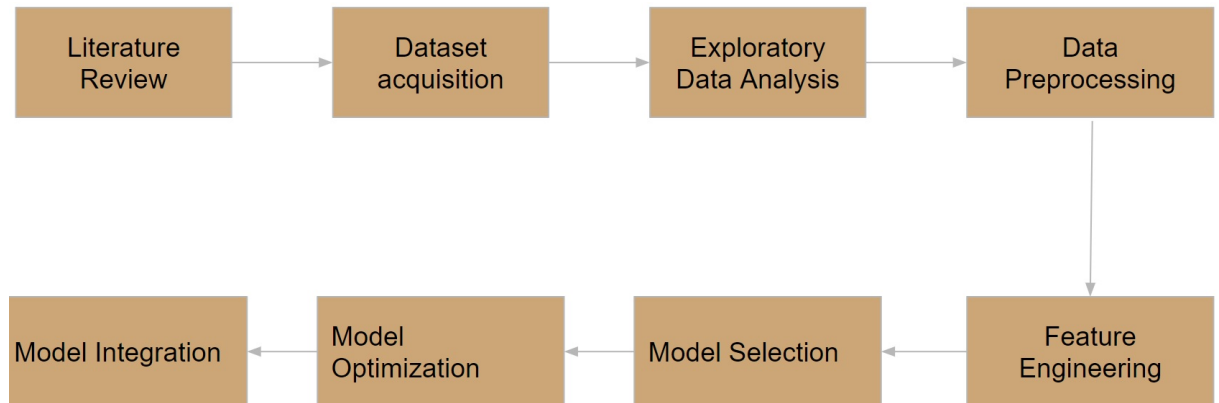


Figure 1. Block diagram of project flow

After acquiring the dataset, we conduct exploratory data analysis and data pre-processing to make it suitable for model ingestion. We then apply multiple machine learning and deep learning models to establish a baseline performance. After evaluating the models, we tune the hyperparameters of the best performing one to improve its performance and then integrate it into a system to generate predictions from new data.

4. Dataset

The dataset was obtained from the University of California Irvine Machine Learning Repository [7]. This dataset was compiled using a 12 lead ECG machine and contains 15 arrhythmia classes.

It contains 452 data points, each with 279 features. Further information about the classes can be inferred from the source website.

4.1. Exploratory Data Analysis

We conduct exploratory data analysis by visualizing the data in Tableau.

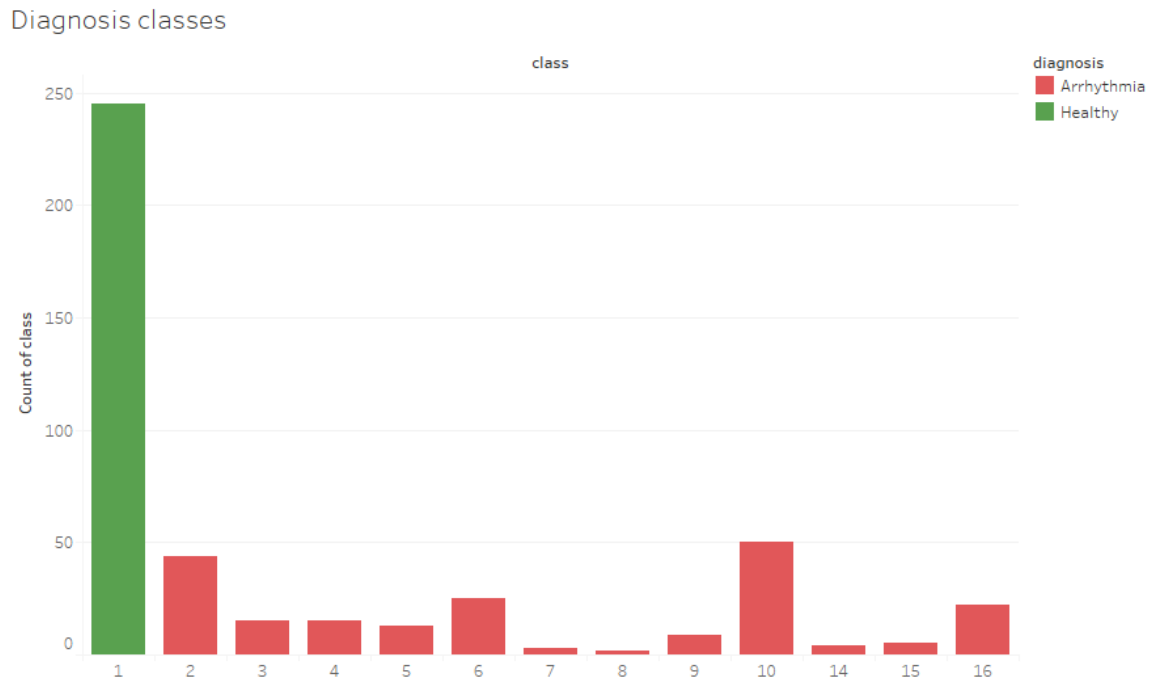


Figure 2. Frequency of classes in dataset

From figure 2, we see that the classes are very imbalanced. Training a model on this dataset will result in an imbalanced model, hyper sensitive to high frequency classes. So, in order to stabilize our predictions, we group the classes into “healthy” and “arrhythmia”.



Figure 3. Frequency of classes in dataset, grouping arrhythmic cases together

As we can see from figure 3, grouping the data leads to a better, more balanced dataset which will be suitable for predictive modelling.

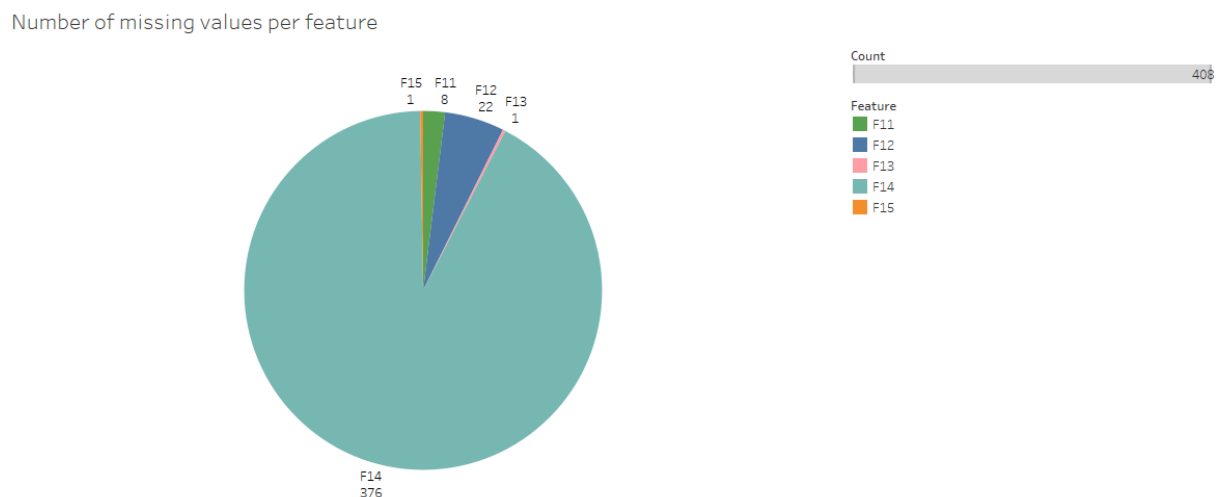


Figure 4. Number of missing values per feature

Figure 4 shows that the dataset contains a large number of missing values, especially in feature 14. Upon further investigation, feature 14 indicated the length of the “J wave”, which is generally not present in normal ECG readings. The presence of this feature might be of

importance to our model, so we cannot simply omit it. We address this issue during the “data pre-processing” stage.

Diagnosis by Age

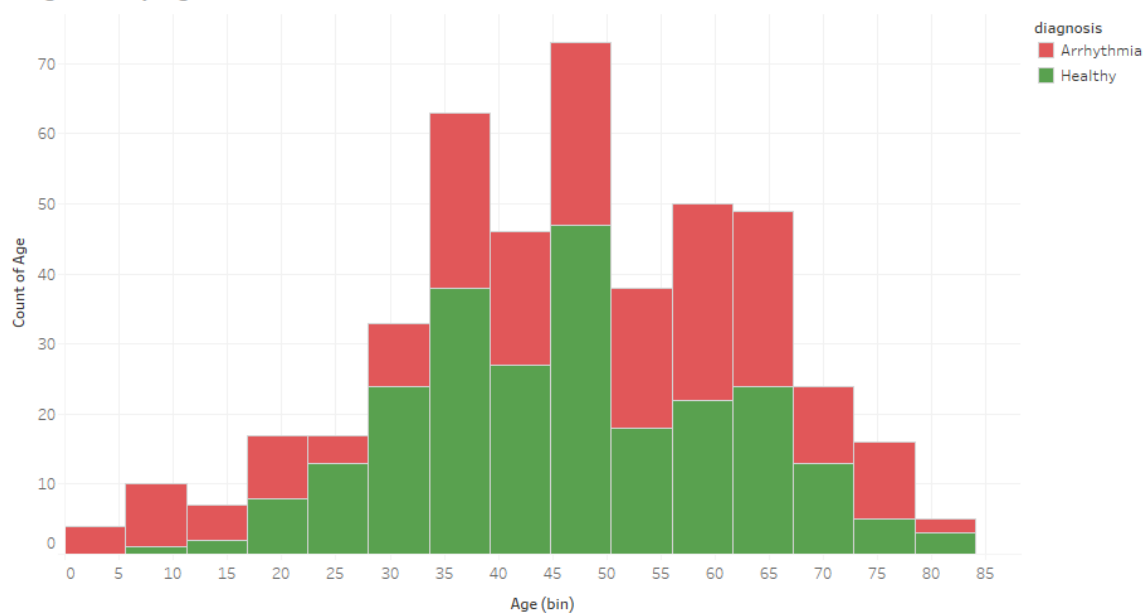


Figure 5. Arrhythmia diagnosis by age

Diagnosis by Sex

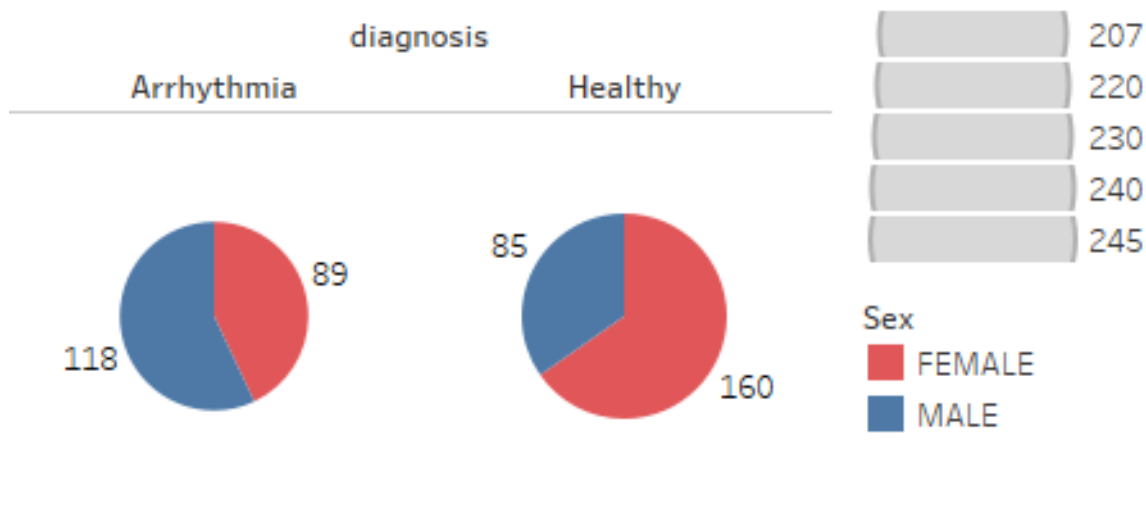


Figure 6. Arrhythmia diagnosis by sex

Figures 5 and 6 are more for general interest than statistical significance. They show that

arrhythmia diagnosis by age follows a Gaussian distribution, with more people in the age range of 35-65 being diagnosed, and arrhythmia also occurs more frequently in males than in females.

5. Experimental Analysis

5.1. Data Pre-processing

Given the size of the dataset, manually labelling each class is unfeasible. Instead, we label each column as “F1, F2, ..., F279”. We then imported the dataset into a MySQL database to simulate an industrial production environment and improve processing speeds. As shown in figure 4, feature 14 contains a considerable amount of NA values (83%) and is converted to a Boolean, indicating the presence of the J wave, instead of the value. We then remove all rows containing missing values, which downsized the dataset to 420 samples. We also create a “diagnosis” column which indicates whether the data point is healthy or arrhythmic.

5.2. Feature Engineering

While multiple techniques exist for selecting the best features for a model, due to the size of the dataset, we chose to eliminate features based on variance. Features with low variance are unlikely to affect the outcome of the model. On examining the features, we observe a steady decrease in the variance of features with a sudden drop at variance of 50. Based on this, we decided to utilize all features with variance above 50. The total number of features in the dataset is now 93.

Principal component analysis

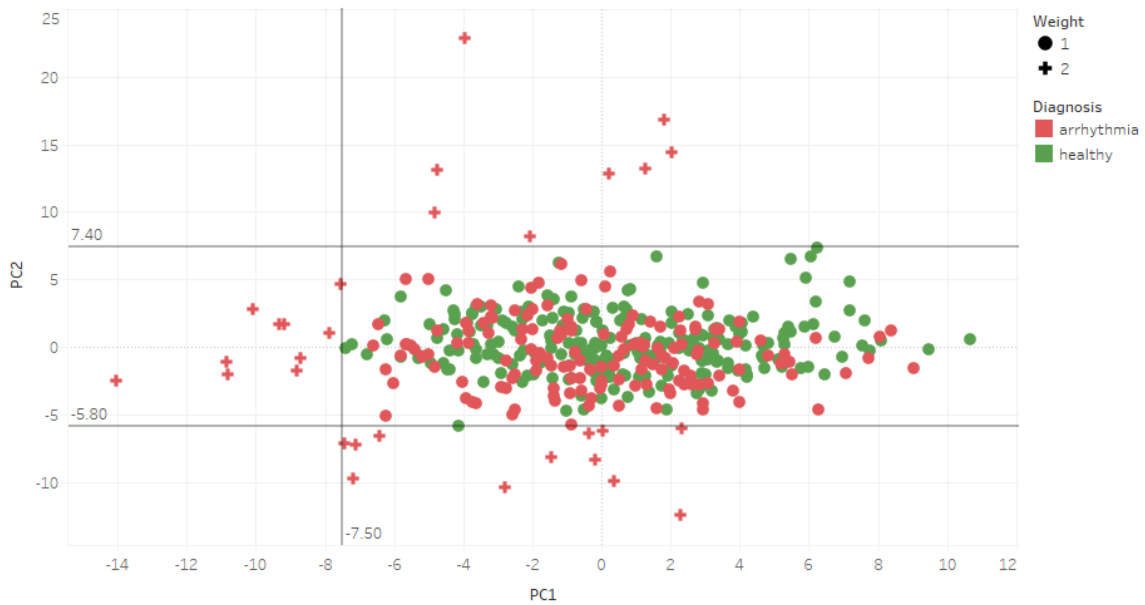


Figure 7. Principal component analysis

In order to initialize the weights for the neural network, we performed a principal component analysis and visualized the results. From figure 6, we can see that there isn’t much difference between the healthy and arrhythmic data points. In order to reinforce the characteristics of the more prominent data points, we assign those points a heavier weight when initializing the neural network.

5.3. Model Selection

We tested multiple classification models using Python (scikit-learn) and R (h2o.ai), and using accuracy as a performance metric, noted down their performance in table 1.

Table 1. Comparison of various classification models

Classification Model	Accuracy%
K-Nearest Neighbors	69%
Random Forest	75%
Support Vector Machine	58%
Artificial Neural Network	78%

As seen in table 1, the deep learning approach yields the best results when judged based on accuracy.

5.4. Model Optimization

In order to maximize the performance of the model, we performed a random grid search, where we iterate through a two-dimensional list of possible hyperparameters, applying random permutations and combinations of various hyperparameters and observing the performance of the models generated using 25-fold cross validation. We use a randomized approach as iterating through each and every listed hyperparameter would be extremely computationally expensive, while randomized grid search usually gives suitable results in a shorter amount of time.

Random grid search

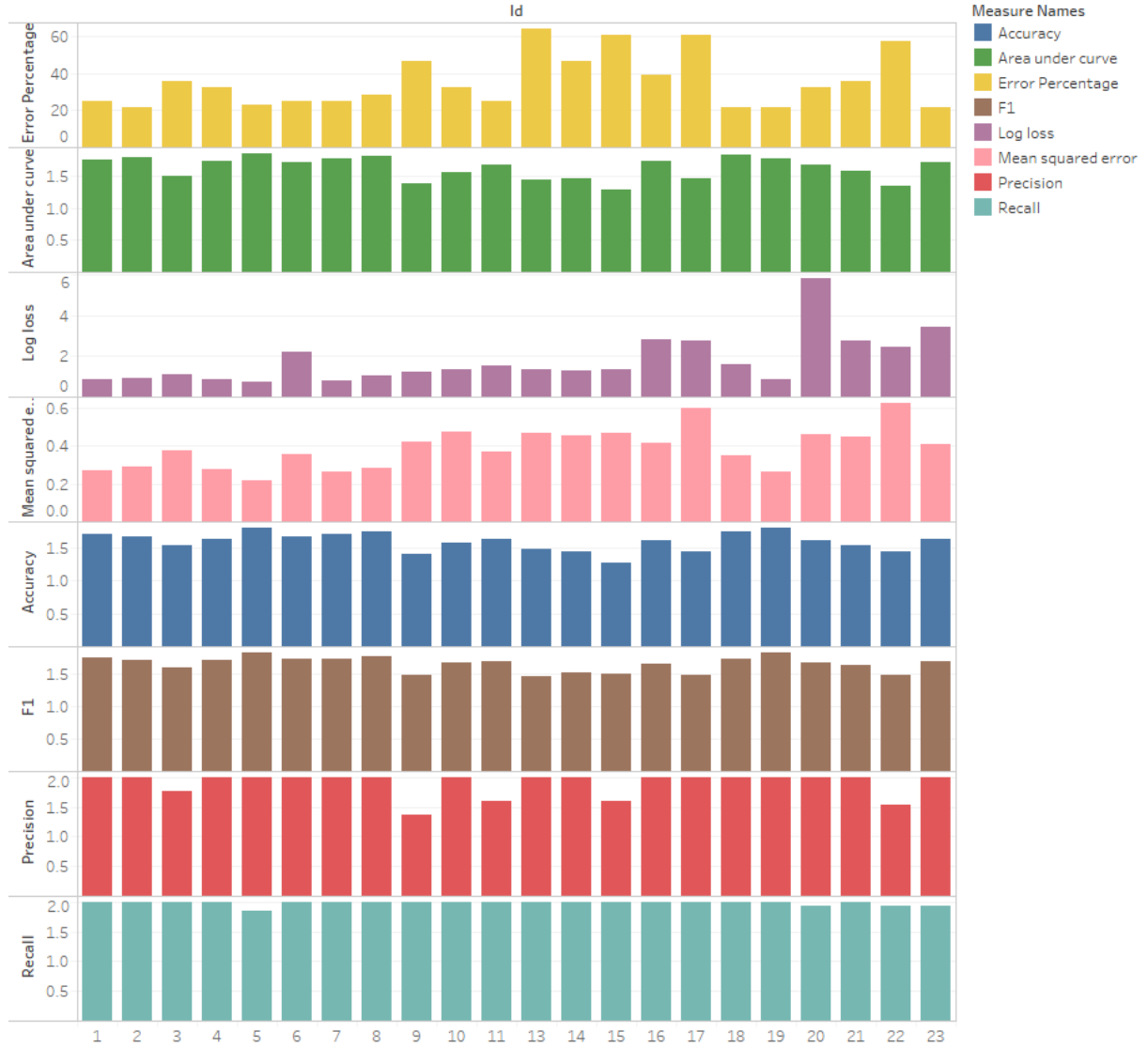


Figure 8. Random grid search model performance comparison

Each run of the program generates twenty to thirty models at a go. In total, we generated about 250 models over the course of this project. Figure 8 shows the output of one such run, with 23 models generated, and their performance metrics visualized.

6. Results

The performance metrics of the best generated model are documented in table 2.

Table 2. Confusion matrix of best model

		Predicted	
		Healthy	Arrhythmia
Actual	Healthy	35	4
	Arrhythmia	1	19

Table 2 displays the confusion matrix, a commonly used method to gauge the performance of classification models. In our testing set, the model correctly classified 54 out of 59 cases. It mis-reported 4 healthy cases as arrhythmic (false positive, type I error) and one arrhythmic case as healthy (false negative, type II error).

Table 3. Performance metrics of the best model

Accuracy	Precision	Recall	F1	MSE
0.915254	0.826087	0.95	0.883721	0.081473
RMSE	Log Loss	Mean Per-Class Error	AUC%	
0.285435	0.2788	0.100845	0.963768	

Table 3 showcases the performance metrics of the model. According to Rajpurkar, Pranav, et al. [23], average cardiologists attained a Sensitivity of 0.6181 and an F1 score of 0.78, which are much lower than our model.

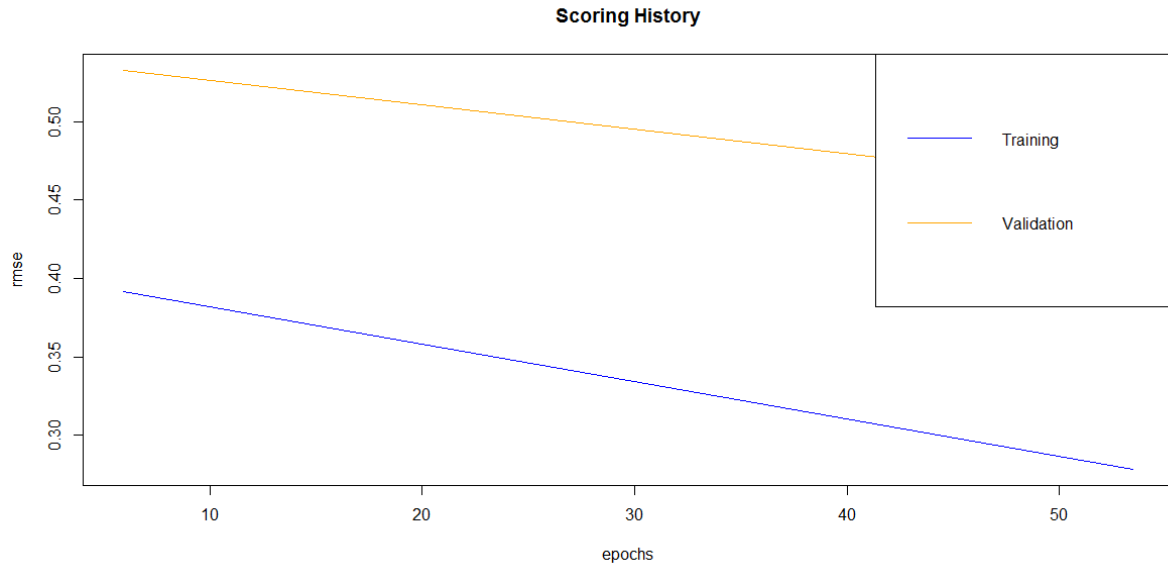


Figure 9. RMSE progression of model

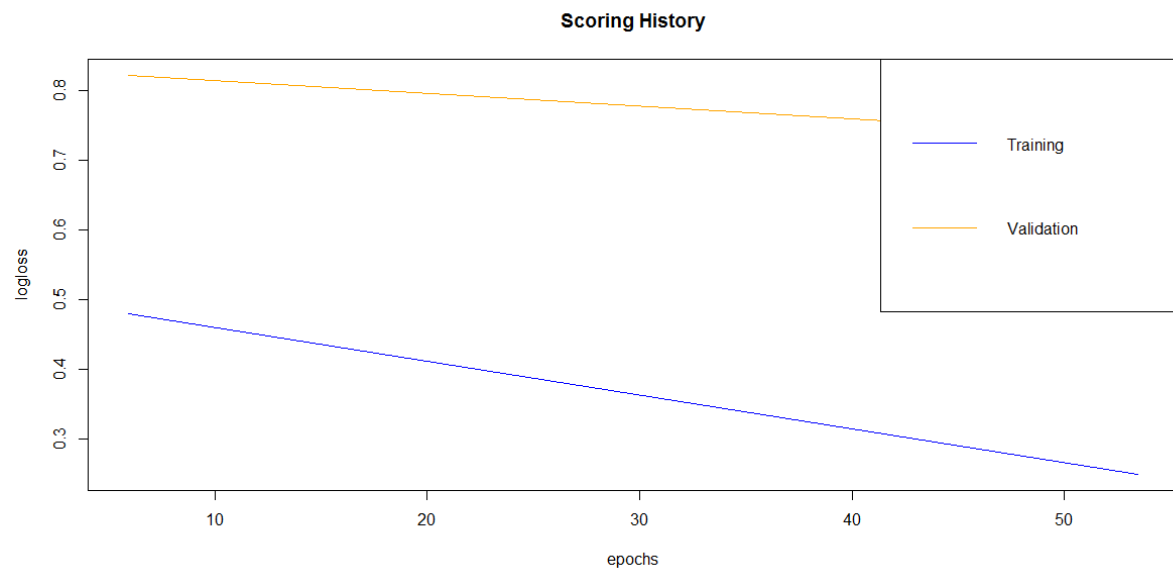


Figure 10. Log loss progression of model

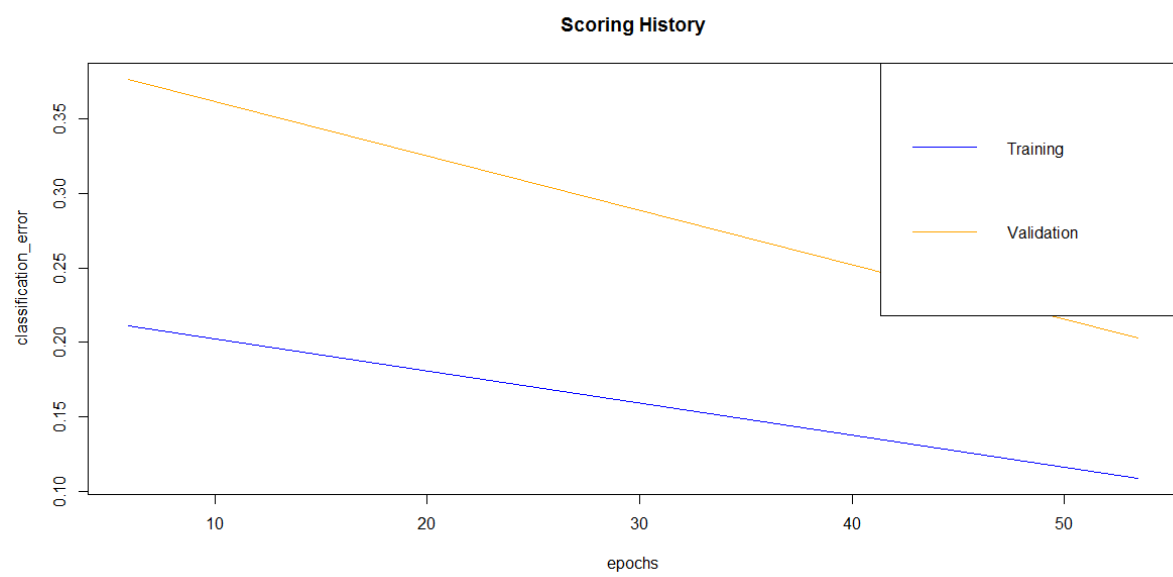


Figure 11. Classification error progression of model

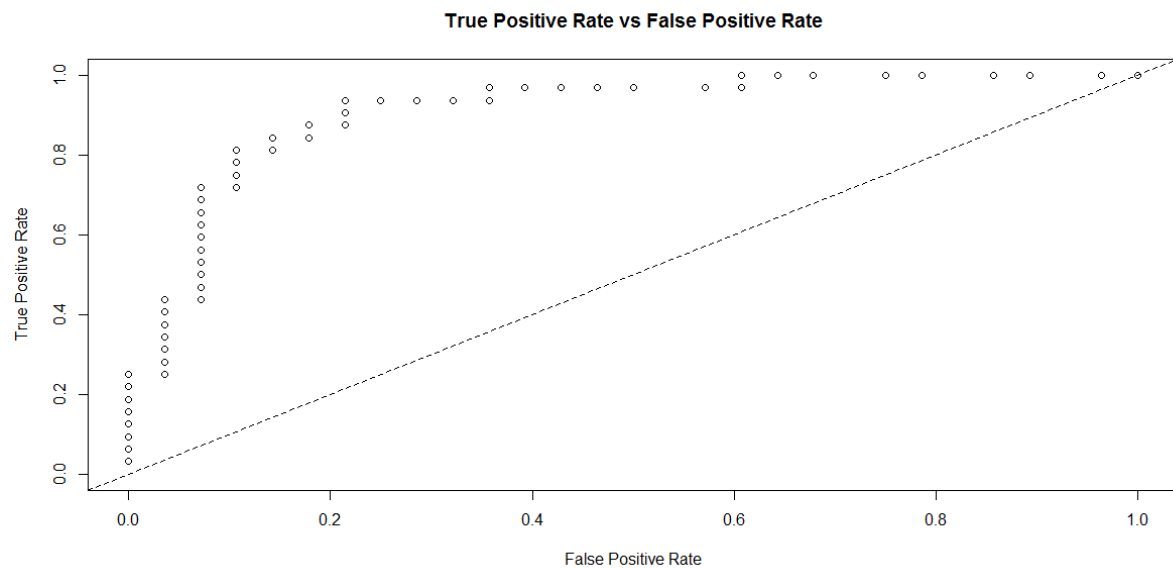


Figure 12. True positive vs False positive rate of model

Figures 9, 10, 11 and 12 show various graphs from the training process. They indicate generally good progression as training was carried out, with error gradually decreasing at each step, while overall model performance increased.

7. System Integration

In order to make the system practical and usable for real life applications, we implemented the model into a Graphical User Interface (GUI) application which can also be deployed as a web app.

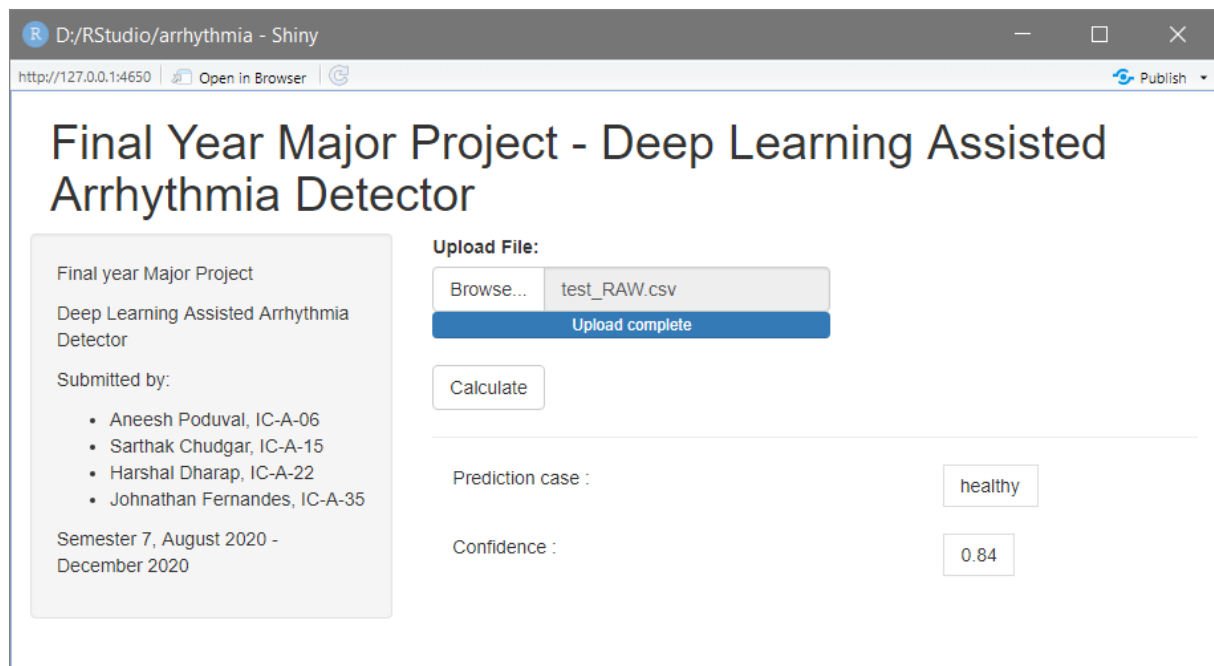


Figure 13. Graphical user interface

As shown in figure 13, the GUI app accepts data in the form of a comma separated values (.csv) file, Initializes the h2o cluster and generates the prediction along with the confidence margin.

8. Future Work

The performance of the model can be improved by using larger, more varied datasets, and fine-tuning the model hyperparameters by implementing a genetic algorithm based optimization technique. Advancements in deep learning and medical data recording are leading to more efficient algorithms and processing techniques which may be utilized to further improve the model.

9. Conclusion

Through this project, we have utilized deep learning algorithms to create a model which can detect heart arrhythmia from ECG signals with an accuracy of 91.52%, and successfully deployed it onto a web application for easy use. While this model outperforms average cardiologists and the results may seem promising, one must keep in mind that the error rate is still too high to fully replace professionals in the medical field. This model, like all machine learning models is not perfect, but is still useful and can serve as a preliminary guide in most situations. This project primarily serves as a demonstration of the progress made by deep learning in the past decade, and it will take several more decades for algorithms to reach the level of medical professionals.

Acknowledgments

The authors of this paper would like to thank Prof. Vijaykumar R. Bhanuse (project guide), Prof. Dr. (Mrs.) Shilpa Y. Sondkar (head of department) and Dr Rajesh Jalnekar (Director) for their constant guidance, and for giving us the opportunity and resources to perform this project.

References

- [1] Acharya, U. Rajendra, et al. "Automated identification of shockable and non-shockable life-threatening ventricular arrhythmias using convolutional neural networks." *Future Generation Computer Systems* 79 (2018): 952-959.
- [2] Batra, Anish, and Vibhu Jawa. "Classification of arrhythmia using conjunction of machine learning algorithms and ECG diagnostic criteria." *Training Journal* (1975).
- [3] Centers for Disease Control and Prevention, National Center for Health Statistics. Underlying Cause of Death, 1999–2018. Accessed March 12, 2020.
- [4] Chen, Min, et al. "Disease prediction by machine learning over big data from healthcare communities." *Ieee Access* 5 (2017): 8869-8879.
- [5] Colilla S, Crow A, Petkun W, Singer DE, Simon T, Liu X. Estimates of current and future incidence and prevalence of atrial fibrillation in the U.S. adult population. *Am J Cardiol.* 2013;112:1142–1147. doi: 10.1016/j.amjcard.2013.05.063.
- [6] Cybenko, G. (1989) "Approximations by superpositions of sigmoidal functions", *Mathematics of Control, Signals, and Systems*, 2 (4), 303-314
- [7] Dua, D. and Graff, C. (2019). UCI Machine Learning Repository [<http://archive.ics.uci.edu/ml>]. Irvine, CA: University of California, School of Information and Computer Science.
- [8] Gao, Dayong, et al. "Arrhythmia identification from ECG signals with a neural network classifier based on a Bayesian framework." *Proceedings of the 24th SGAI International Conference on Innovative Techniques and Applications of Artificial Intelligence*. Citeseer, 2004.
- [9] Ge, Dingfei, Narayanan Srinivasan, and Shankar M. Krishnan. "Cardiac arrhythmia classification using autoregressive modeling." *Biomedical engineering online* 1.1 (2002): 5.
- [10] Gothwal, Himanshu, Silky Kedawat, and Rajesh Kumar. "Cardiac arrhythmias detection in an ECG beat signal using fast fourier transform and artificial neural network." *Journal of Biomedical Science and Engineering* 4.04 (2011): 289.
- [11] Hinton, G. E.; Osindero, S.; Teh, Y. W. (2006). "A Fast Learning Algorithm for Deep Belief Nets" (PDF). *Neural Computation*. 18 (7): 1527–1554
- [12] Kukar, Matjaž, et al. "Analysing and improving the diagnosis of ischaemic heart disease with machine learning." *Artificial intelligence in medicine* 16.1 (1999): 25-50.
- [13] Kurt Hornik (1991) "Approximation Capabilities of Multilayer Feedforward Networks", *Neural Networks*, 4(2), 251–257. doi:10.1016/0893-6080(91)90009-T
- [14] Lee, Susan Ciarrocca. "Using a translation-invariant neural network to diagnose heart arrhythmia." *Advances in neural information processing systems*. 1990.
- [15] Lyon, Aurore, et al. "Distinct ECG phenotypes identified in hypertrophic cardiomyopathy using machine learning associated with arrhythmic risk markers." *Frontiers in physiology* 9 (2018): 213.
- [16] Ma, Fenglong, et al. "Dipole: Diagnosis prediction in healthcare via attention-based bidirectional recurrent neural networks." *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining*. 2017.
- [17] Mishra, M. (2019, March 16). Apple Watch detects irregular heartbeat in large U.S. study. Reuters. Retrieved from <https://www.reuters.com/article/us-health-heart-apple-idUSKCN1QX0EI>
- [18] Miyasaka Y, Barnes ME, Gersh BJ, et al. Secular trends in incidence of atrial fibrillation in Olmsted County, Minnesota, 1980 to 2000, and implications on the projections for future prevalence of external icons. *Circulation.* 2006;114:199–225.
- [19] Mohammadzadeh-Asl, Babak, and Seyed Kamaledin Setarehdan. "Neural network based arrhythmia classification using heart rate variability signal." 2006 14th European Signal Processing Conference. IEEE, 2006.
- [20] Mohan, Senthilkumar, Chandrasegar Thirumalai, and Gautam Srivastava. "Effective heart disease prediction using hybrid machine learning techniques." *IEEE Access* 7 (2019): 81542-81554.
- [21] Mustaqeem, Anam, et al. "Wrapper method for feature selection to classify cardiac arrhythmia." 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2017.
- [22] Oh, Shu Lih, et al. "Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length heart beats." *Computers in biology and medicine* 102 (2018): 278-287.
- [23] Rajpurkar, Pranav, et al. "Cardiologist-level arrhythmia detection with convolutional neural networks." *arXiv preprint arXiv:1707.01836* (2017).
- [24] S. Karimifard, A. Ahmadian, M. Khoshnevisan and M. S. Nambakhsh, "Morphological Heart Arrhythmia Detection Using Hermitian Basis Functions and kNN Classifier," 2006 International Conference of the IEEE Engineering in Medicine and Biology Society, New York, NY, 2006, pp. 1367-1370, doi: 10.1109/IEMBS.2006.260182.

- [25] Sannino, Giovanna, and Giuseppe De Pietro. "A deep learning approach for ECG-based heartbeat classification for arrhythmia detection." *Future Generation Computer Systems* 86 (2018): 446-455.
- [26] Schwob, Michael R., et al. "Robust Multimodal Heartbeat Detection Using Hybrid Neural Networks." *IEEE Access* 8 (2020): 82201-82214.
- [27] Soman, Thara, and Patrick O. Bobbie. "Classification of arrhythmia using machine learning techniques." *WSEAS Transactions on computers* 4.6 (2005): 548-552.
- [28] Tsipouras, Markos G., and Dimitrios I. Fotiadis. "Automatic arrhythmia detection based on time and time-frequency analysis of heart rate variability." *Computer methods and programs in biomedicine* 74.2 (2004): 95-108.
- [29] What users are saying about Apache Spark vs H2O. Retrieved from <https://www.trustradius.com/compare-products/apache-spark-vs-h2o>