

Detecting Myocardial Infarction by Electrocardiogram Machine Learning Models With Greater Accuracy; A Technical Advance Article

M.D.S. Sudaraka

Department of Obstetrics and Gynaecology, Faculty of Medicine, University of Peradeniya

I. Abeyagunawardena

Department of Physiology, Faculty of Medicine, University of Peradeniya

E. S. De Silva

Faculty of Medicine, University of Peradeniya

S Abeyagunawardena (✉ shamali.agwd@yahoo.com)

Department of Medicine, Faculty of Medicine, University of Peradeniya

Research Article

Keywords: Myocardial infarction, Electrocardiogram, Machine learning models for Myocardial infarction detection.

Posted Date: January 29th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-150700/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background

Electrocardiogram (ECG) is a key diagnostic test in cardiac investigation. Interpretation of ECG is based on the understanding of normal electrical patterns produced by the heart and alterations of those patterns in specific disease conditions. With machine learning techniques, it is possible to interpret ECGs with increased accuracy. However, there is a lacuna in machine learning models to detect myocardial infarction (MI) coupled with the affected territories of the heart.

Methods

The dataset was obtained from the University of California, Irvine, Machine Learning Repository. It was filtered to obtain observations categorized as Normal, Ischemic changes, Old Anterior MI and Old Inferior MI. The dataset was randomly split into a training set (70%) and a test set (30%). 73 out of the 270 ECG features were selected based on the changes observed following MI, after excluding predictors that had near zero variance across the observations. Three machine learning classification models (Bootstrap Aggregation Decision Trees, Random Forest, Multi-layer Perceptron) were trained using the training dataset, optimizing for the Kappa statistic and the parameter tuning was achieved with repeated 10-fold cross validation. Accuracy and Kappa of the samples were used to evaluate performance between the models.

Results

The Random Forest model identified old anterior and old inferior MIs with 100% sensitivity and specificity and all 4 categorized observations with an overall accuracy of 0.9167 (95% CI 0.8424 - 0.9633). Both the Bootstrap Aggregation Decision Trees and the Multi-layer Perceptron models identified old anterior MIs with 100% sensitivity and specificity and their overall accuracies for all 4 observations were 0.8958 (95% CI 0.8168 - 0.9489) and 0.8542 (95% CI 0.7674 - 0.9179) respectively.

Conclusion

With a medically informed feature selection we were able to identify old anterior MI with 100% sensitivity and specificity by all three models in this study, and old inferior MI with 100% sensitivity and specificity by Random Forest Model. If the data set can be improved it is possible to utilize these machine learning models in hospital setting to identify cardiac emergencies by incorporating them into cardiac monitors, until trained personnel become available.

Background

Since its development in 1902, Electrocardiogram (ECG) has evolved steadily to the cost effective, simple and non-invasive 12 lead ECG used in the modern age, which is one of the most widely used investigations worldwide (1). It remains a cornerstone investigation in the medical field, utilized for the diagnosis of countless medical conditions including cardiac conditions such as acute coronary syndromes and arrhythmias and non-cardiac conditions such as electrolyte imbalances. A standard 12 lead ECG records information from 6 limb leads namely lead I, II, III, aVR, aVL and aVF and 6 chest leads from V1-V6. These leads look at the heart from different directions and record the electrical signals accordingly and the standard ECG has a p wave, QRS complex and T wave (1).

The interpretation of ECG is based on the understanding of normal electrical patterns produced by the electrical activity of the heart and the variations in specific conditions. This pattern recognition process requires targeted knowledge and depends on the ability and training of the interpreter. With the development of machine learning techniques, it is possible to train a model to recognize these patterns and interpret the ECG tracing.

There have been models reported which aimed to detect arrhythmias by classifying ECGs as normal and abnormal and to narrow down the specific arrhythmia by classifying them into multiple classes (2,3). However, as a review done in 2017 reveals, studies regarding myocardial infarction (MI) and the affected territory are few (4). A vast amount of information can be gathered using the ECG, from the presence of an MI up to the affected area, timing and complications. The ECG changes of an MI depend on the lead and the timing and may include tall T waves in the initial stage, ST segment elevation, left bundle branch block, prolonged QRS

complex, increased heart rate (reduced R-R interval) and q waves in an old MI. According to the affected territory, these changes may be seen in different leads. For example, in leads II, III and aVF in an inferior MI and in the chest leads in an anterior MI (5).

Therefore, with these changes in mind we aimed to create a machine learning model utilizing existing ECG data to detect an MI along with the information regarding the affected territory and timing.

Materials And Methods

The dataset used was obtained from University of California, Irvine (UCI) Machine Learning Repository, which contained 279 predictor variables of 452 observations (6). The variables included age, sex, height, weight and attributes derived from standard 12 lead ECG tracings. Each observation had been assigned one of sixteen possible outcomes.

The original dataset was filtered to obtain observations categorized as Normal (class 1), Ischemic changes (class 2), Old Anterior MI (class 3) and Old Inferior MI (class 4).

The dataset was randomly split into a training set (70%) and a test set (30%). Ninety nine features were selected based on the ECG changes that are observed following a MI. They are heart rate, the amplitudes of q, R and T waves and the durations of the QRS complexes along with the age and the sex of the subject (Table 1).

Table 1 Selection of ECG features based on medical knowledge

Feature		Number of leads	Number of predictors
Age			1
Sex			1
Heart rate			1
Amplitudes	Q wave amplitude	12	12
	R wave amplitude	12	12
	T wave amplitude	12	12
QRS Duration	Q wave duration	12	12
	R wave duration	12	12
	S wave duration	12	12
	R' wave duration	12	12
	S' wave duration	12	12
Total			99

26 predictors (R' and S' waves of all the 12 leads and Q wave amplitudes of V2 and V3 leads) that had near-zero variance across the observations were excluded. The selected 73 features were scaled to have a mean of 0 and standard deviation of 1.

Three machine learning classification models (Bootstrap Aggregation Decision Trees, Random Forest, Artificial Neural Network) were trained using the training dataset, following data up-sampling to overcome class imbalances, optimizing for the Kappa statistic, which is a measure of inter-rater reliability which takes the possibility of percentage agreement occurring by chance, into account.

Parameter tuning of the models was done with 10 fold repeated cross validation. Cross validation folds were sampled identically to make inter-model performance evaluation possible using accuracy and Kappa statistics.

Performance of the final model was evaluated using sensitivity, specificity, positive predictive value, negative predictive value and balanced accuracy for each class for the test dataset.

The evaluation metrics are summarized below.

Accuracy = Number of correct predictions / The total number of predictions

Sensitivity= True positive / (True positive + False negative)

Specificity= True negative / (True negative + False positive)

Positive predictive value = True positive / (True Positive + False positive)

Negative predictive value = True negative / (True negative + False negative)

Sensitivity= True positives / (True positives + False negatives)

Specificity= True negatives / (True negatives + False positives)

Results

The Bootstrap Aggregation classifier had a mean accuracy of 0.9788 and a mean Kappa value of 0.9718 for the 100 cross validation resamples.

The Random Forests model yielded the highest cross validation accuracy (0.9927) and Kappa metric (0.9902) out of the three models, with 3 randomly selected variables for splitting each tree node (Figure 1).

The Feed forward Artificial Neural Network was a Multi-layer Perceptron with one hidden layer (Figure 2). This model yielded the highest cross validation accuracy (0.9772) with 10 hidden units (Figure 3).

The distributions of the evaluation matrices of the cross validation samples and the scatter plot matrix for Kappa statistic of the three models are depicted in following figures (Figure 4, 5, 6).

All the three models had excellent cross validation accuracies.

The previously unseen test data split was used to evaluate the final model performances.

The Random Forests model predicted the classes of the test dataset with an overall accuracy of 0.9167 (95% CI 0.8424 - 0.9633) with 100% sensitivity and specificity for identifying old anterior and old inferior myocardial infarctions.

The overall accuracies of the Bootstrap Aggregation Decision Trees and the Multi-layer Perceptron models were 0.8958 (95% CI 0.8168 - 0.9489) and 0.8542 (95% CI 0.7674 - 0.9179). Both these models could identify old anterior myocardial infarctions with 100% sensitivity and specificity.

The confusion matrices and evaluation matrices of the three models are summarized below (Tables 2-7).

Random Forests model

Table 2 Confusion matrix of Random Forest Model

Original label of the class	Radom forest Model identified as			
	Class 1	Class 2	Class 3	Class 4
Class 1	71	6	0	0
Class 2	2	10	0	0
Class 3	0	0	5	0
Class 4	0	0	0	2

Table 3 Performance of the Random Forest model on the test data

	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value	Balanced Accuracy
Class 1	0.9726027	0.7391304	0.9220779	0.8947368	0.8558666
Class 2	0.6250000	0.9750000	0.8333333	0.9285714	0.8000000
Class 3	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000
Class 4	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000

Bootstrap Aggregation Decision Trees Model

Table 4 Confusion Matrix of Bootstrap Aggregation Decision Trees Model

	Class 1	Class 2	Class 3	Class 4
Class 1	69	7	0	1
Class 2	1	10	0	1
Class 3	0	0	5	0
Class 4	0	0	0	2

Table 5 Performance of the Bootstrap Aggregation Decision Trees Model on the test data

	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value	Balanced Accuracy
Class 1	0.9857143	0.6923077	0.8961039	0.9473684	0.8390110
Class 2	0.5882353	0.9746835	0.8333333	0.9166667	0.7814594
Class 3	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000
Class 4	0.5000000	1.0000000	1.0000000	0.9787234	0.7500000

Multi-layer Perceptron Model

Table 6 Confusion Matrix of Multi-layer Perceptron Model

	Class 1	Class 2	Class 3	Class 4
Class 1	65	8	0	4
Class 2	2	10	0	0
Class 3	0	0	5	0
Class 4	0	0	0	2

Table 7 Performance of Multi-layer Perceptron Model on the test data

	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value	Balanced Accuracy
Class 1	0.9701493	0.5862069	0.8441558	0.8947368	0.7781781
Class 2	0.5555556	0.9743590	0.8333333	0.9047619	0.7649573
Class 3	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000
Class 4	0.3333333	1.0000000	1.0000000	0.9574468	0.6666667

Discussion

The UCI machine learning dataset has been utilized by previous studies for the detection of heart disease, mainly arrhythmias. A modular neural network model proposed by Jadhav et al., aimed to detect arrhythmias by classifying the ECG tracings as 'normal' and 'abnormal'. It was found that the model classified the tracings into the two classes with 82.22% accuracy (2). Batra et al., also utilized multiple machine learning algorithms for the classification of arrhythmias into multiple classes. The maximum accuracy achieved was 84.82% using Support Vector Machine (SVM) algorithm and gradient boosting, in conjunction (3). A model proposed by Mitra et al., utilized neural networks for the classification of arrhythmias into 'normal' and 'arrhythmia'. It was found that an accuracy of 87.71% was achievable using this method (7).

The feature extraction and reduction was carried out by correlation-based feature subset selection. A notable characteristic in the above models is that the feature selection was not performed based on medical knowledge regarding the diseases but rather by mathematical feature reduction techniques. It is possible that this may be responsible for the relatively similar accuracies which do not exceed 90%. Furthermore, this data set has not been utilized specifically for the detection of myocardial infarctions and its subsets.

It is evident by this study that the detection and classification of myocardial infarction with regard to timing and territory can be achieved with a relatively higher accuracy. The better model performances demonstrated in our study may be attributed to the feature selection technique utilized, which was predominantly based on the known electrophysiological changes occurring during a myocardial Infarction. All three models were able to identify old anterior MI with 100% sensitivity and 100 % specificity, with the random forest model being able to identify old inferior MI with similar matrices as well. Therefore, this study highlights the potential real world applications of this technology in the healthcare sector. The availability of an interpretation along with the tracing would be of paramount importance in conditions like myocardial infarctions where time is of the essence. The need to wait for trained personnel for the interpretation of ECG can be avoided and hence delays minimized. Patients requiring emergent care can be identified in settings with a large patient load. In addition, implanted cardioverter defibrillators (ICDs) and pacemakers of patients can also be equipped with this technique, enabling the detection of MI as well. Hence, it is evident that further studies to explore the boundaries of these new upcoming leaps in technology should be performed for the improvement of medical diagnostics.

Conclusion

We used three machine learning models to detect MI along with timing and territory, using data derived from ECG tracings. The overall accuracies obtained by Random Forests, Bootstrap Aggregation Decision Trees and Multi-layer Perceptron were 0.9167, 0.8958 and 0.8542 respectively. Furthermore, all the models could identify old anterior MIs with a specificity and sensitivity of 100%.

This study can be expanded by using a more extensive data set, to include more detail regarding timing and territories, as well as the detection and classification of arrhythmias. Hence, it can be concluded that these models can be utilized in hospital settings to minimize the delays until trained personnel are available for ECG interpretation and to alert the staff of a potential emergency.

List Of Abbreviations

ECG - Electrocardiogram

MI - Myocardial Infarction

ICD - Implanted Cardioverter Defibrillator

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

The dataset analyzed during the current study are available in the University of California, Irvine (UCI) Machine Learning Repository , [<http://archive.ics.uci.edu/ml/datasets/Arrhythmia>]

Competing interests

The authors declare that they have no competing interests

Funding

Study was self-funded

Authors' contributions

MDS extracted the data and designed the models while IA analyzed the data and drafted the manuscript. ESK also contributed to drafting the manuscript and S was in charge of overall supervision, interpretation and revising of the article.

All authors have approved the submitted version and to have agreed both to be personally accountable for the author's own contributions and to ensure that questions related to the accuracy or integrity of any part of the work, even ones in which the author was not personally involved, are appropriately investigated, resolved, and the resolution documented in the literature.

Acknowledgements

Not applicable

References

1. Electrocardiography: Overview, ECG Indications and Contraindications, Preparation [Internet]. Emedicine.medscape.com. 2020 [cited 28 September 2020]. Available from: <https://emedicine.medscape.com/article/1894014-overview>
2. Jadhav S, Nalbalwar S, Ghatol A. ECG arrhythmia classification using modular neural network model. 2010 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES). 2010;.
3. Batra A. Classification of Arrhythmia Using Conjunction of Machine Learning Algorithms and ECG Diagnostic Criteria. [Internet]. 2016 [cited 28 September 2020];. Available from: <https://www.semanticscholar.org/paper/Classification-of-Arrhythmia-Using->

4. Roopa C, Harish B. A Survey on various Machine Learning Approaches for ECG Analysis [Internet]. 2017 [cited 28 September 2020]. Available from: https://www.researchgate.net/publication/317400128_A_Survey_on_various_Machine_Learning_Approaches_for_ECG_Analysis
5. Burns E. Myocardial Ischaemia • LITFL • ECG Library Diagnosis [Internet]. Life in the Fast Lane • LITFL • Medical Blog. 2020 [cited 28 September 2020]. Available from: <https://litfl.com/myocardial-ischaemia-ecg-library/>
6. Altay Guvenir, BurakAcar, GulsenDemiroz, AyhanCekin "A Supervised Machine Learning Algorithm for Arrhythmia Analysis." Proceedings of the Computers in Cardiology Conference, Lund, Sweden, 1997.
7. Mitra M, Samanta R. Cardiac Arrhythmia Classification Using Neural Networks with Selected Features. *Procedia Technology*. 2013;10:76-84.
8. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)*. 2012;22(3):276-282.

Figures

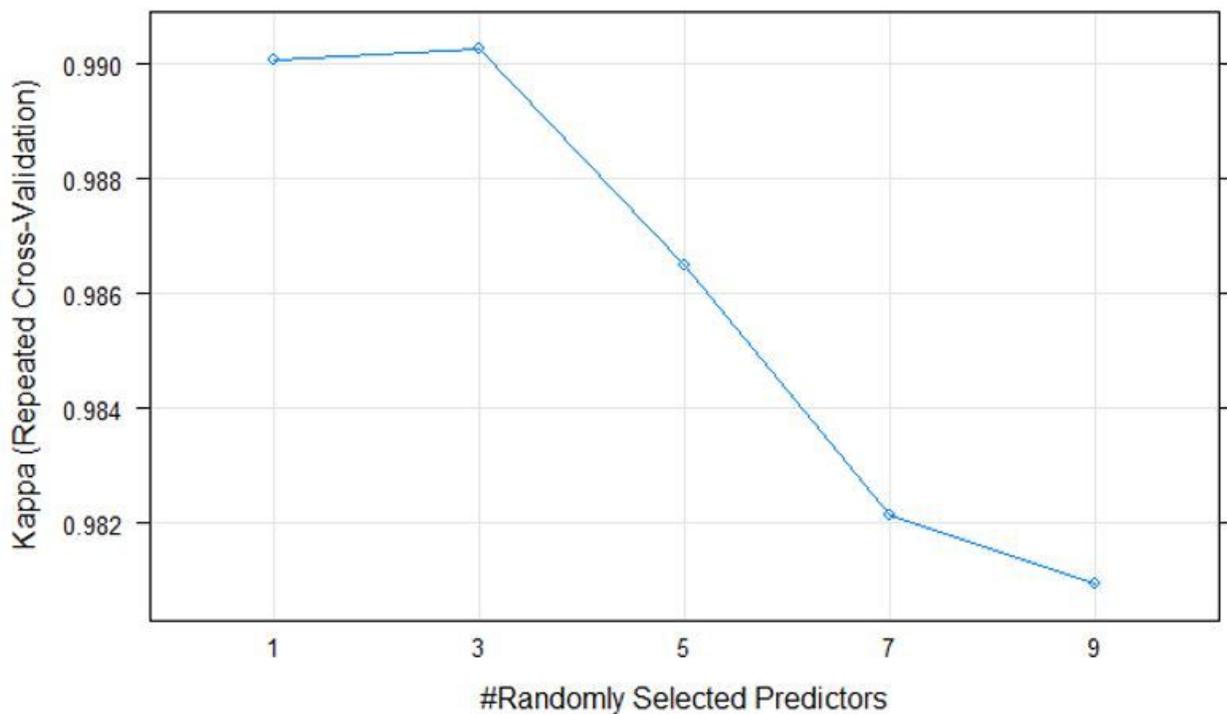


Figure 1

Cross validation Kappa statistics for different numbers of randomly selected predictors of the Random Forests model.

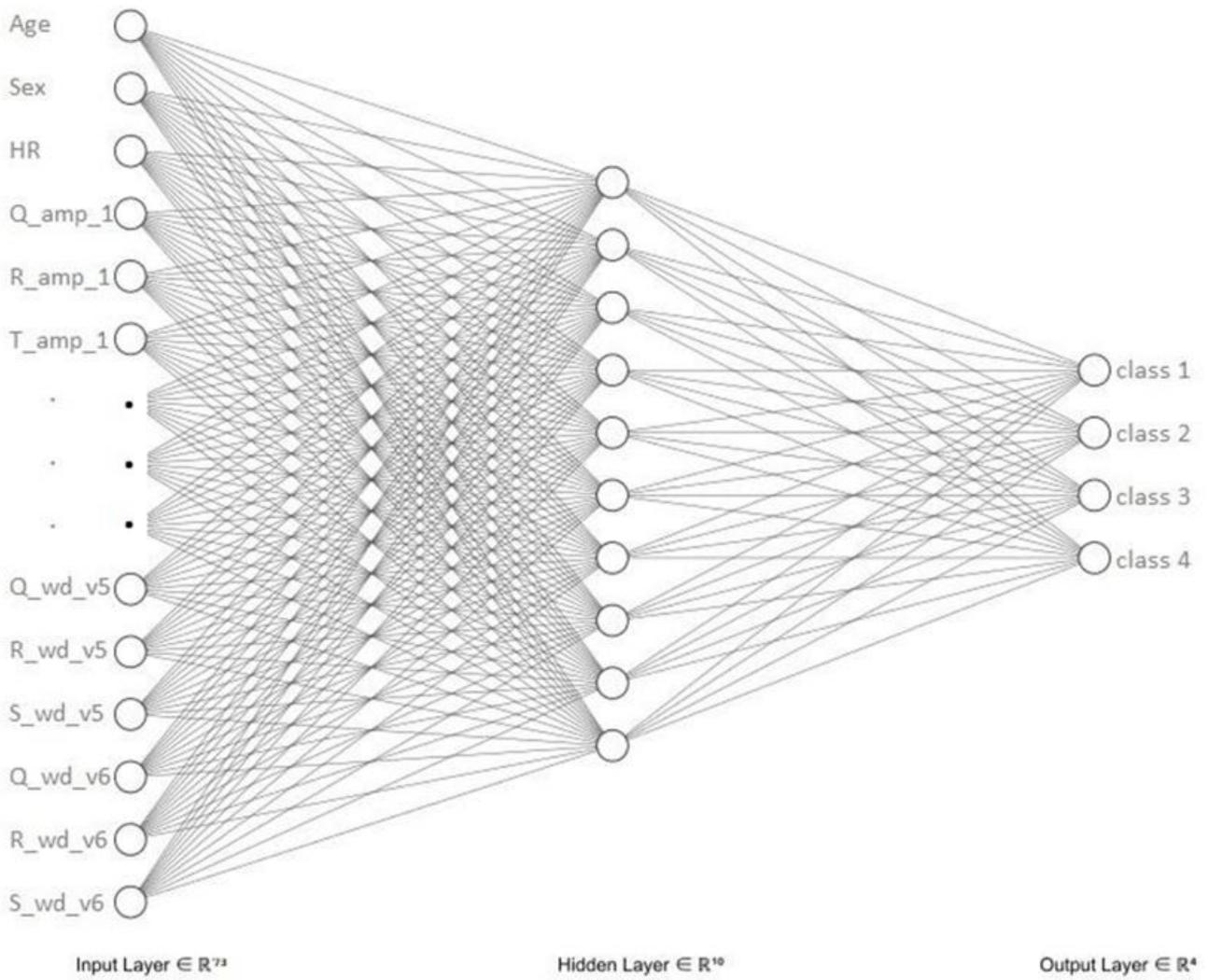


Figure 2

The architecture of the Perceptron

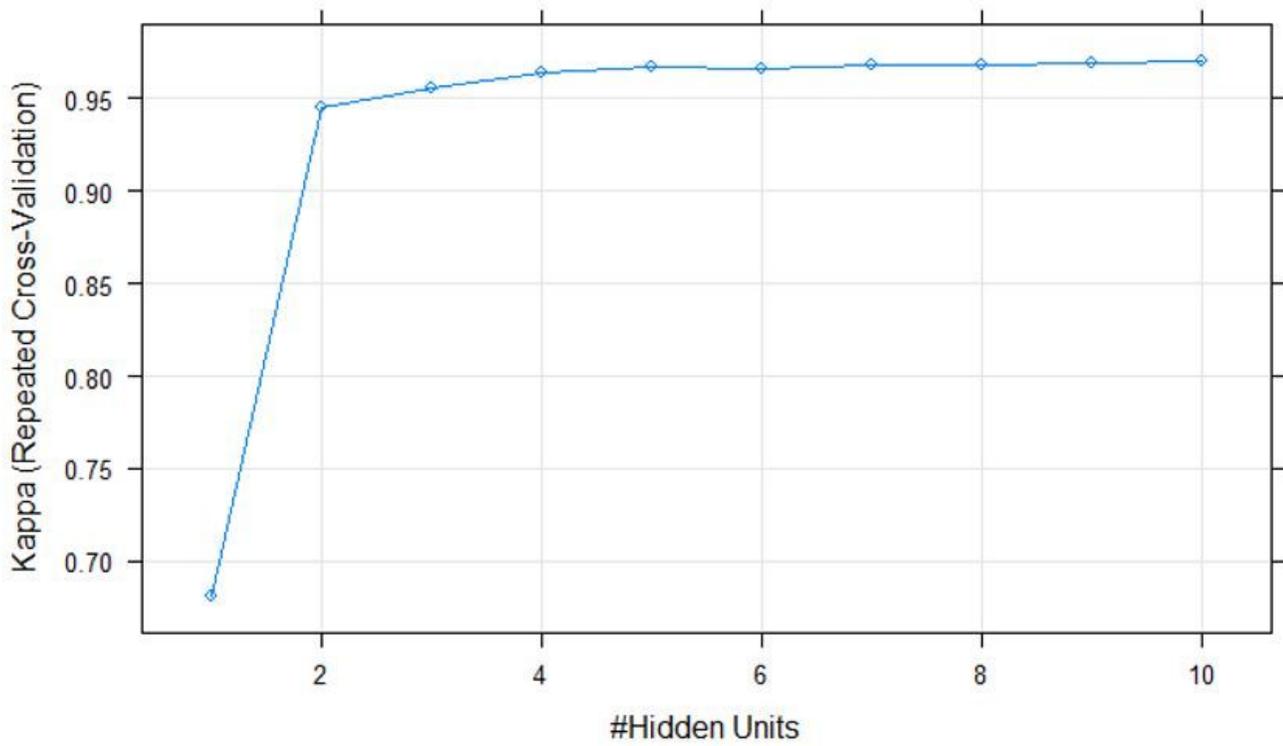


Figure 3

Cross validation Kappa statistics for different numbers hidden units of the neural network.

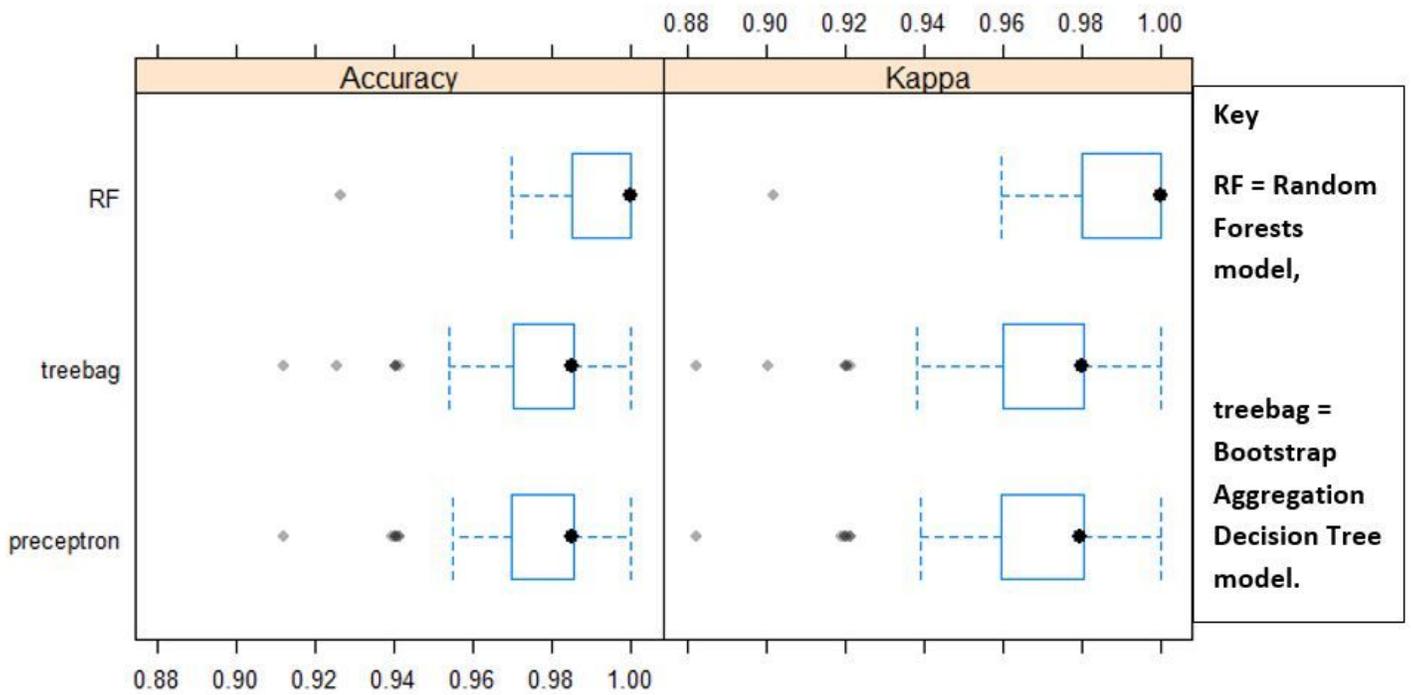


Figure 4

Box-and-whisker plots of the resampling distributions of the accuracy and the Kappa between the models.

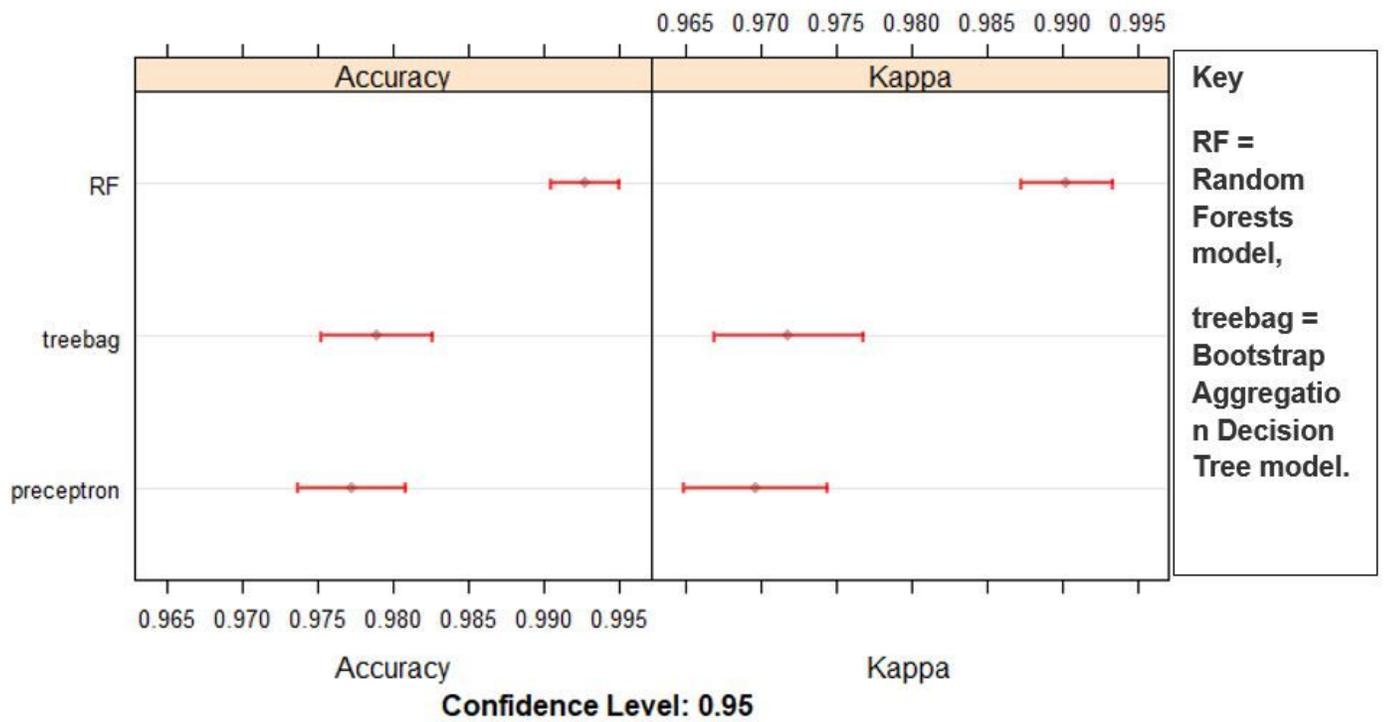


Figure 5

Resampling distributions of the accuracy and the Kappa between the models.

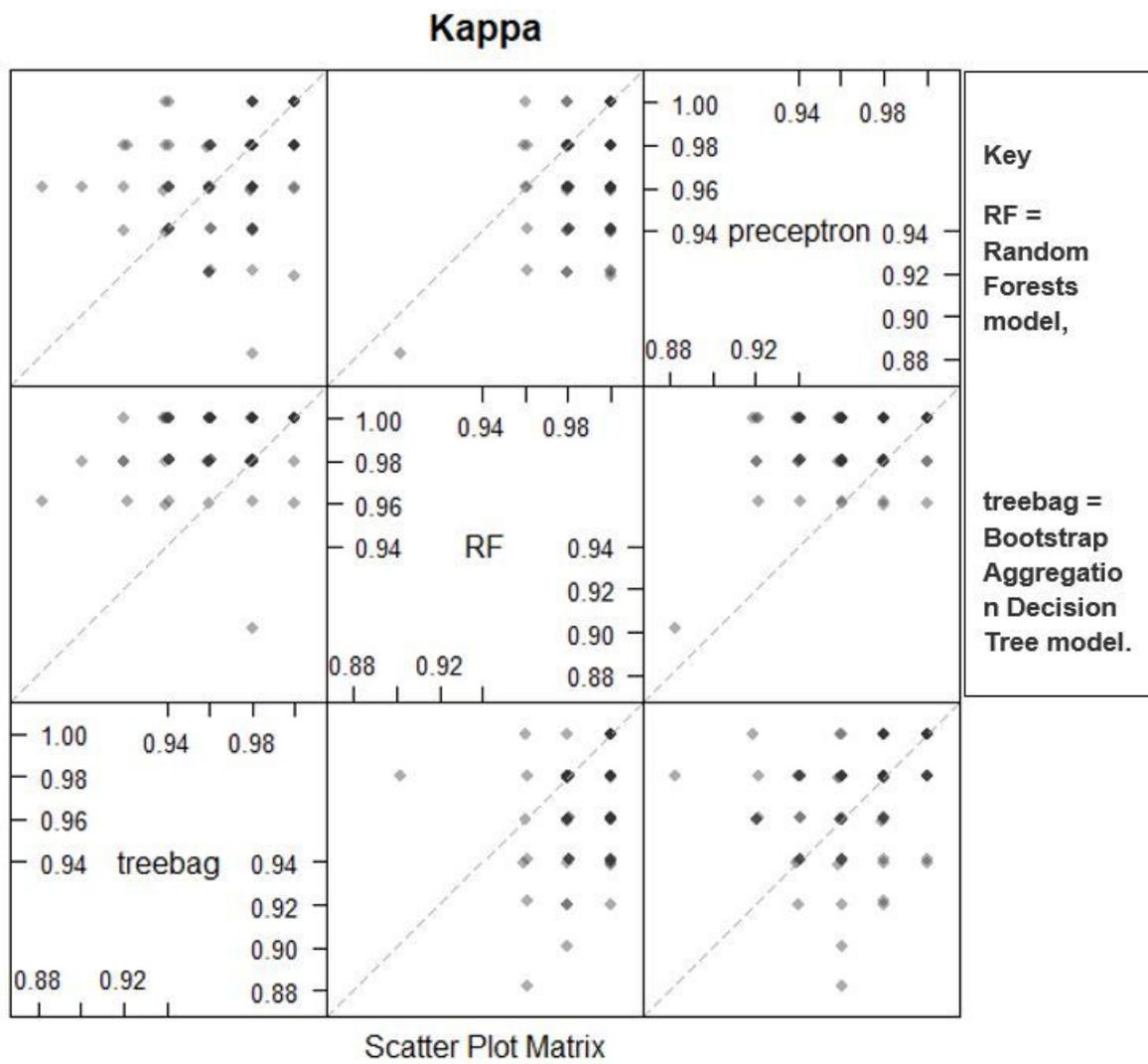


Figure 6

Scatter plot matrix of the resampling distributions of the Kappa between the models.