

## **CSC8003 – Machine Learning Project Part I – Model Building Report**

### **Development of a Machine Learning Model to Calculate a Depth of Anaesthesia Index in Patients Under General Anaesthesia**

Name: Diana Kither

Table of Contents

1. Introduction	3
2. Literature Review	3
3. Data Analysis	4
4. Machine Learning Methodology	5
5. Evaluation Methods	7
6. Parameter Selection	7
7. Index Design	9
8. Performance Evaluation	12
9. Reflection	13
10. Conclusion	14
11. References	15
12. Appendices	17

## 1. Introduction

Efficient monitoring of general anaesthesia is vital for maintaining patient health and welfare (Gu et al 2019). A variety of parameters are monitored by anaesthesiologists to determine the depth of anaesthesia in a patient, including the electroencephalogram (EEG) and the Bispectral Index (A-2000 BIS monitor; Aspect Medical Systems Inc, Newton MA) (Jiang et al 2015). EEG monitors the electrical activity of the brain and can be utilised in a number of areas of medicine including anaesthesia (St Vincent's, 2014). BIS is calculated from the EEG data of patients and is a useful indicator for monitoring depth of anaesthesia (Lee et al, 2019). With the advent of technology, investigations into utilising machine learning techniques with EEG data have been conducted in an effort to determine new parameters with which to gauge depth of anaesthesia in patients (Gu et al 2019; Lee et al, 2019). This project aims to develop a model to determine a depth of anaesthesia (DoA) index using EEG data and to assess the efficacy of this model in comparison to the BIS index.

## 2. Literature Review

General anaesthesia involves placing a patient in an unconscious state using medication to enable health procedures to be conducted (Liu et al 2018). Effective monitoring of anaesthesia is crucial for patient safety, ensuring not only that a patient is adequately anaesthetised but that the plane of patient anaesthesia is not too deep which can lead to potentially serious and life threatening complications (Gu et al 2019; Liu et al 2018). A variety of parameters are used by clinicians to monitor depth of anaesthesia including, heart rate, respiratory rate, oxygen saturation, blood pressure and other measurements (Gu et al 2019; Hernandez-Meza et al 2017). EEG monitors the electrical activity of the brain and in turn gives important insights into how anaesthesia is influencing the nervous system (Gu et al 2019).

The Bispectral Index (BIS) is also widely utilised to monitor depth of anaesthesia (Diykh et al 2020). BIS is calculated using EEG data and is considered to be an effective tool as part of anaesthetic monitoring due to its strong correlation with anaesthetic depth (Gu et al 2019; Lee et al, 2019, Ortolani et al 2002). BIS measurements range from 0 with an isoelectric EEG to a level of 100 in a patient that is awake (Jiang et al 2015). A range of 40-60 is considered adequate anaesthesia (Jiang et al 2015). A number of investigations have been conducted into using machine learning models to predict anaesthetic depth based on EEG data (Hashimoto et al 2020). Given the frequent use of BIS as part of anaesthetic monitoring, this index is often utilised as a comparison to determine the efficacy of modelling techniques (Liu et al 2018).

Amongst the machine learning techniques that have been utilised to predict anaesthetic depth using EEG data, neural networks have been frequently used. Gu et al (2019) used an artificial neural network model to predict a DoA index using EEG data from patients under anaesthesia. The results showed a strong correlation with BIS values (Gu et al 2019). Ortolani et al (2002) also had success in using a neural network model in predicting anaesthetic depth. When the DoA indices calculated by the model developed in the study were compared with patients BIS values, there was a 94% correlation (Ortolani et al 2002). Jiang et al (2015) demonstrated the efficacy of a neural network model when the results of their model were compared to both BIS values and expert opinions of anaesthesiologists when assessing depth of anaesthesia. Sadrawi et al (2015) again used a neural network model to assess depth of anaesthesia using EEG data and doctors scoring of anaesthetic depth. When model results were compared to BIS values it was found the neural network model was more effective at predicting anaesthetic depth and had a lower mean absolute error (Sadrawi et al 2015).

Hernandez-Meza et al (2017) demonstrated the effectiveness of using a support vector machine (SVM) model classifier when compared to both BIS values and mean alveolar concentration (MAC) in assessing anaesthetic depth. Tacke et al (2020) compared several machine learning algorithms for assessing EEG data to determine anaesthetic depth and found the SVM model to have the greatest accuracy. Jahanseir et al (2018) demonstrated that a multioutput least square support vector regression method was effective in predicting anaesthetic depth with no significant statistical difference when compared to BIS values.

Lee et al (2019) showed that decision trees and multiple linear regression could be used to successfully build a predictive model for anaesthetic depth based on EEG data. Liu et al (2017) compared the use of random forests, SVM and neural networks for regression modelling of EEG data to determine depth of anaesthesia. They found that random forests out performed both SVM and neural networks when compared to BIS values (Liu et al 2017). Interestingly SVM and neural networks had similar levels of performance (Liu et al 2017).

Multiple machine learning techniques have shown to be efficacious in predicting depth of anaesthesia from EEG data which may have future applications in medicine. As discussed in the review conducted by Hashimoto et al (2020), artificial intelligence has the advantage of being able to quickly analyse large volumes of data to gain further insights, making it a mechanism which would be of assistance to clinicians in their decision making.

### 3. Data Analysis

Data analysis was conducted on the EEG training patient data sets using the Rapid Miner Studio platform (Version 9.2, RapidMiner 2020). The RapidMiner process for data analysis is shown in Appendix A, Figure 8. The data sets contain the EEG data and BIS values for 10 patients over a period of time under general anaesthesia. Readings are taken every second and are in chronological order.

As the first step in exploratory data analysis, the features of the data set, BIS, x1, x2, x3, x4, and x5 were examined for each patient. The summary statistics for each patient's data are shown in Appendix B, Table 4. When examining BIS values for each patient, the majority of patients had right skewed distribution towards lower BIS values which would be expected given that each patient was under general anaesthesia. Lower values indicate that the patient is at a deeper level of anaesthesia. Patients 1 and 9 however had bimodal distributions of their BIS values (Appendix C, Figure 9). The average BIS value ranged between 37 and 48. The lowest minimum BIS values were patient 4 with a score of 2 and patient 6 with a score of 5. BIS minimums otherwise ranged between 10 and 20. Maximum BIS values were all above 96 as would be expected for awake patients, however patients 6 and 9 had maximum BIS values of 89 and 85 respectively. There was also some range in standard deviations for patients, varying between 10 and 25.

Feature x1 had a very even distribution across all patients. When examining minimum and maximum values, x1 had a very low range between 0 and 0.04 and very low standard deviations. Patient 7 had a standard deviation of 0. Feature x2 also had a narrow range of values across all patients of between 0.003 and 0.006. The highest standard deviation for x2 was 0.019. The graphical distributions of x2 generally had a right skewed distribution with the exception of patients 1 and 9 who both had bimodal distributions (Appendix C, Figure 10). Interestingly these patients also had bimodal distributions for their BIS values.

Feature x3 again had a narrow range of values and low standard deviations across all patients. The distribution of values for x3 was generally right skewed with the exception of patients 2, 6 and 9 who all had relatively normal distributions for this feature (Appendix C, Figure 11). Patients 4, 5 and

10 had highly right skewed distributions (Appendix C, Figure 12). Feature x4 also had very low standard deviations across patients with a left skewed distribution for all patients excluding patient 4 who had a normal distribution for this feature (Appendix C, Figure 13).

Feature x5 had a greater range of values compared to other features and the standard deviation for patients across feature x5 ranged between 0.6 and 1.5. With the exception of patient 4, feature x5 had a right skewed distribution amongst patients in the data set. Patient 4 again had a relatively normal distribution of x5. Box plots were examined of the features and no outliers were found.

The next step in exploratory data analysis was to examine the correlation matrices between the features of each patient. The correlation matrices are shown in Appendix D. All patients had moderate to strong correlations with features x2, x4 and x5 with the label variable BIS. In turn all patients also had weak correlations between x1 and BIS. Patients 5, 7, 8 and 10 had moderate correlations between x3 and BIS whilst the remaining patients had weak correlations. When comparing the correlations between features x1, x2, x3, x4 and x5, feature x1 had a weak correlation with all features amongst all patients. Moderate to high correlations were present between the other features. The highest correlations were found between features x3 and x5 (patients 5, 7, 8 and 10) at greater than 0.86, features x4 and x5 at greater than 0.8 (patients 1, 5, 8, and 10), features x2 and x4 at greater than 0.8 (patients 1 and 8) and finally a correlation of -0.835 between features x3 and x4 (patient 5).

The patient training data sets were then combined into a single data set in Microsoft Excel (Version 16.4, Microsoft 2019) and loaded into the MATLAB platform (Version R2020a, Mathworks 2020) for further analysis. The MATLAB code for this data analysis is presented in Appendix A. The first step when surveying the combined data set was to examine the descriptive statistics and correlation matrix (Appendix D, Table 5 and 16). Similarly to the correlations between features of individual patients, feature x1 had a very low correlation when compared to BIS of just -0.05 in the combined patient data set. X1 also had a low correlation with the other features. Features x2 to x5 had moderate to high correlations with BIS. Features x3 and x5, x4 and x5, x4 and x2 all had moderate correlations with one another. The other feature combinations had low to moderate correlations. When looking at the distribution of BIS values we can see that there is a greater proportion of BIS readings from anaesthetised patients compared to awake patients.

On examination of the scatterplot matrix of the combined patient data set (Appendix C, Figure 14), a degree of linear relationship is evident between the following features x4 and x5, x3 and x5, x3 and x4. A mild linear relationship is also evident between BIS and both x4 and x5 and possibly x2. The other relationships between variables are non-linear.

#### 4. Machine Learning Methodology

##### i) Parameter Selection

The first step in model building is parameter selection. Results from data analysis and from construction of a linear model using the combined patient data set were utilised for parameter selection as part of the model building process. Reducing parameter numbers in turn reduces model complexity, improves understanding of the model and assists in preventing model overtraining.

The first step in parameter selection was to examine the distribution and correlation of features with one another and with the label variable. This was done for the combined patient data set as well as for individual patients. Features with higher correlations with the label variable were considered to

be important in model building and therefore these features were retained in the data set. If two features were very highly correlated with one another then they were considered to be measuring the same characteristics thus one feature can be removed from the data set to avoid redundancy. Any features with even distributions across all values were considered to have less influence on varying BIS values and therefore these features were also removed from the data set prior to building the model.

The next step of parameter selection and model building involved construction of a linear model. On examination of the relationships between BIS values and features in the data set it was evident that there was a degree of linear relationship between some features and BIS. As a result of this a linear model was constructed to investigate the relationships between features of the data set and BIS values which may assist with further parameter selection. The linear model was built with the remaining features in the data set after initial parameter selection as the independent variables and BIS as the dependent variable. Weights and p values for each feature in the linear model were assessed as to the significance of these features in the model. Any features whose coefficients had high p values ( $> 0.05$ ) were considered not to be significant thus these were also removed from the data set prior to the model building process.

### ii) Model Building

To construct the machine learning model to predict the DoA index, the combined patient data set of 10 patients was used for training and testing. Data from patients 1 to 7 were used for model training and data from patients 8 to 10 for model testing. The whole data set was used in effort to maximise the amount of data available for model building so as to improve the quality of the model. The two machine learning techniques selected for building the DoA index model were the linear model and a neural network. This combination was selected based on research conducted in the literature review, characteristics of the patient data set and results from trialling a combination of several other machine learning techniques as part of the model building process.

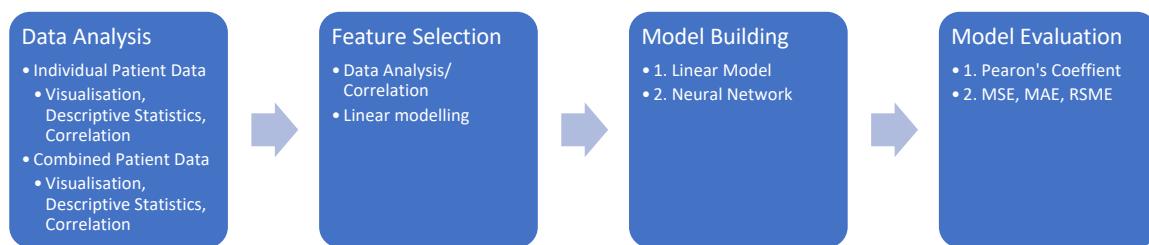
A neural network is a complex supervised learning algorithm that is based on the neuron structure of the human brain (Gu et al 2019; Ognjanovski 2019; Protocol 2017). The neural network is comprised of input, hidden and output layers. (Gu et al 2019; Ognjanovski 2019; Protocol 2017). Data is received by the input layer, the hidden layer applies functions to the data as part of the learning process then the final output is received by the output layer (Ognjanovski 2019; Protocol 2017). The hidden layer is comprised of nodes or neurons (Ognjanovski 2019; Protocol 2017). Inputs are received by the nodes in the hidden layer, weights are calculated for each input and functions are applied (Ognjanovski 2019; Protocol 2017). As the network is trained, the weights of these inputs at each node are recalculated so as to reduce the error of the predicted value of the function. (Ognjanovski 2019; Protocol 2017). One of the most utilised algorithms for learning in the neural network is back propagation (Gu et al 2019; Ognjanovski 2019; Protocol 2017, Sadwari et al 2015). The process of back propagation involves calculating weights of inputs within the network based on the error rate of the previous training loop of the neural network so as to reduce the error rate of the model (Al-Masri, 2019). On review of the literature, the neural network model combined with back propagation was one of the most commonly used models used to predict depth of anaesthesia using EEG and BIS data and achieved favourable results (Gu et al 2019; Jiang et al 2014; Ortolani et al 2002; Sadwari et al 2015). Therefore this model was selected as part of the model building process.

The neural network model was also selected given the characteristics of the patient data set . On examination of the data set in data analysis, it was evident that there is a mix of linear and non-linear relationships in the data set between features and the label variable. Neural networks are

effective with both linear and non-linear relationships in the data. Neural networks are also effective with large amounts of data. Given the volume of data in the data set this algorithm was again an appropriate choice.

To improve the performance of the neural network, an initial machine learning algorithm was applied to the data set prior to data input into the neural network model. The features selected for model building were shown to have a moderate to high correlation with the BIS value. Therefore it was deemed appropriate to use a linear model for this task. The results of the linear model were then used as input data for the neural network model. This reduces the volume of input for the neural network model, in turn reduces model complexity in an effort to improve model performance, reduce training time and avoid overtraining.

A diagram of the process for developing the DoA Index model is shown below.



*Figure 1 : Process of Depth of Anaesthesia Index Model Building*

## 5. Model Evaluation Methods

To evaluate the performance of the new DoA index model, the model was applied to the data sets of patients 8 to 10 and then the calculated DoA index was compared to the BIS values from these patients. The first parameter used for model evaluation is the Pearson coefficient. The Pearson coefficient measures the strength of the linear relationship between two variables. A high Pearson coefficient between the new DoA index and BIS values indicates a strong relationship between the predicted and actual value (Hiregoudar 2020).

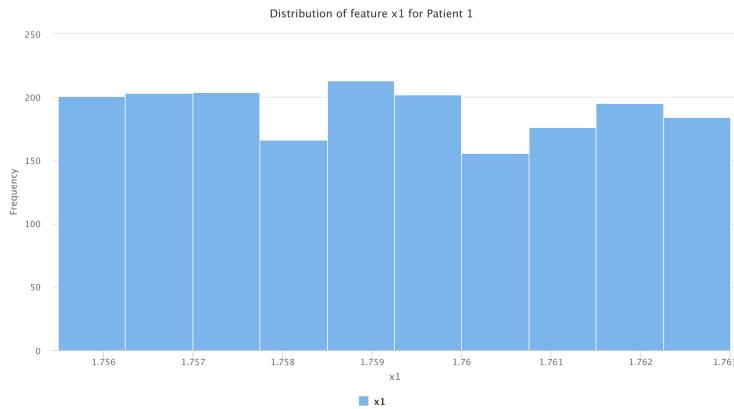
Secondly, the error rates of the model were assessed with mean squared error (MSE), root mean squared error (RMSE) and mean absolute error (MAE). Error rates give a guide to how effective a model is at predicting a value and the goal of model building is to achieve a low error rate (Hiregoudar 2020; Openclassrooms 2020). The MSE and RMSE both give a higher indication of error rates of the model compared to MAE due to their sensitivity to large errors (Openclassrooms 2020).

Finally, the coefficient of determination,  $R^2$  was also calculated to assess the DoA index calculated by the new model in relation to BIS.  $R^2$  is a useful parameter as it indicates the degree of variation in the label variable that can be attributed to the input data of the model (Hiregoudar 2020).

## 6. Parameter Selection

Parameter selection for the DoA index model was conducted in the MATLAB platform (Version R2020a, Mathworks 2020). The code for parameter selection is shown in Appendix E. The first step in parameter selection involves assessment of the feature distribution, descriptive statistics and correlation matrices of the individual and combined patient data sets as determined through the

steps of data analysis. On examination of the distribution of feature  $x_1$  for each patient it was evident that the range of values was very narrow with a low standard deviation (Appendix B, Table) and in the case of patient 7, standard deviation had a value of 0. Furthermore the distribution of  $x_1$  was very even across all values for individual patients as shown in the distribution of feature  $x_1$  for Patient 1 in Figure.



*Figure 2: Distribution of feature  $x_1$  for Patient 1*

Furthermore, feature  $x_1$  had a very low correlation with other features and with BIS as shown in the correlation matrices in Appendix C. Given these results, it was determined that feature  $x_1$  is unlikely to have a significant effect on predicting the BIS result and so it was removed from the data set. When examining correlation between the features of individual patient data sets, some features had high correlations with one another. However when comparing features in the combined patient data set, moderate correlations were evident. Thus it was elected to keep the remaining four parameters at this point.

The second aspect of feature selection involved developing and interpreting the weights and p values of each parameter in a linear model with BIS as the dependent variable. A linear model was built using features  $x_2$  to  $x_5$  and the BIS values for all 10 patients. The equation of the linear model is shown below and the model parameters are shown in Table.

*Equation 1: Linear model for BIS and features  $x_2$  to  $x_5$*

$$BIS = 1034.041 - 441.711(x_2) + 0.798(x_3) - 468.015(x_4) + 3.734(x_5)$$

*Table 1: MATLAB Results of Linear model*

	Estimate	SE	tStat	pValue
<b>Intercept</b>	1034.041	10.077	102.616	0
<b>X2</b>	-441.711	6.510	-67.849	0
<b>X3</b>	0.7984	2.030	0.393	0.69
<b>X4</b>	-468.015	6.30	-74.291	0
<b>X5</b>	3.734	0.087	42.988	0

On examination of the results of the linear model it is evident that features  $x_2$  and  $x_4$  carry the greatest weight in the model with respect to their coefficients. On assessment of the p values,  $x_2$ ,  $x_4$  and  $x_5$  are deemed significant in the model at a significance level of  $<0.05$ .  $X_3$  however is not

significant given the high p value of 0.69. As a result x3 was excluded from model building. Therefore the final features selected for model building are x2, x4 and x5.

## 7. Index Design

The DoA index model was constructed in the MATLAB platform (Version R2020a, Mathworks 2020). The code for model building is presented in Appendix E. The first step in model building for the DoA index model was to divide the data set into training and test sets. Data from patients 1 to 7 was used as the model building and training data set. Data from patients 8 to 10 was used as the test data set to evaluate model performance.

The first element in the DoA index model building process involves constructing a linear model using the selected features as the independent variables and BIS as the dependent variable from the training data set. The calculated results of this linear model are then used as input for the neural network to calculate the DoA index. The equation for the linear model is shown below and its parameters are shown in Table .

*Equation 2: Linear Model for DoA Index Design*

$$BIS = 1034.149 - 440.738(x2) - 467.901(x4) + 3.755 (x5)$$

*Table 2: Linear Model Parameters for DoA Index Design*

	Estimate	SE	t Stat	p value
<b>Intercept</b>	1034.149	10.073	102.666	0
<b>X2</b>	-440.738	6.021	-73.195	0
<b>X4</b>	-467.901	6.293	-74.352	0
<b>X5</b>	3.755	0.068	54.966	0

As is seen in the model parameters, x2 and x4 again have the greatest weight in the model based on model coefficients. All the terms in the model are significant to a value of p <0.05. The linear model was then applied to the patient data to calculate an anaesthesia score. This result was then used as the input for the neural network model.

The second part of the DoA index design involves building the neural network. A neural network model was built with one input layer which utilises the results of the linear model, one hidden layer with 10 nodes and one output layer, the DoA index. Given there is only one input and output layer, only one hidden layer was included in the neural network to reduce complexity of the model. Furthermore, when more than one model layer was trialled in model building it did not improve model performance. As discussed by Gu et al (2019) the formula below was considered whilst trialling numbers of nodes in the hidden layer.

*Equation 3 – Node selection for neural network hidden layer (Gu Y et al 2015)*

$$d = \sqrt{a + b} + c$$

where d= hidden nodes, a = input nodes, b = output nodes, c = regulation constant (value between 1 and 10)

After trialling various numbers of nodes within the hidden layer a final number of 10 nodes were selected as lower numbers of nodes decreased model performance slightly and an increased number of nodes did not improve model performance.

The network training function used for the model was Levenberg-Marquardt optimisation, a backpropagation algorithm. The learning rate was set at 0.05 as lower rates did not result in improved model performance. Epochs were set at a higher level of 1000 given the volume of data and the higher learning rate. The final error goal was set at 0.05. A slightly higher error rate was selected, in an effort to improve model performance but also to decrease computation time. The data set was randomly divided for training validation and testing. The validation checks were left at the default setting of 6 so as to avoid model overfitting. During training of the neural network, model training terminated due to the number of failed validation checks prior to reaching the maximum epoch number. The results of the model building process are shown below.

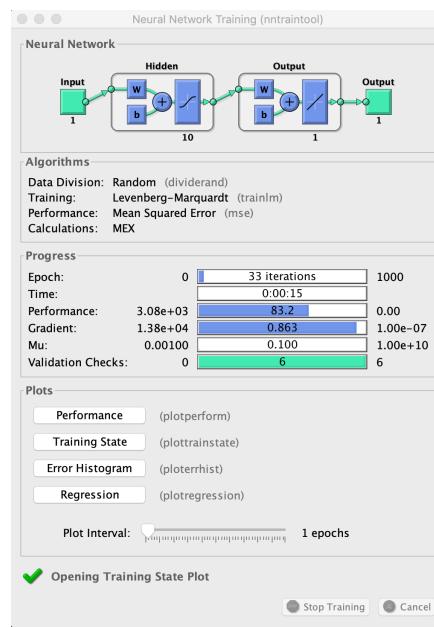


Figure 3: Structure of Neural Network in DoA Index Model

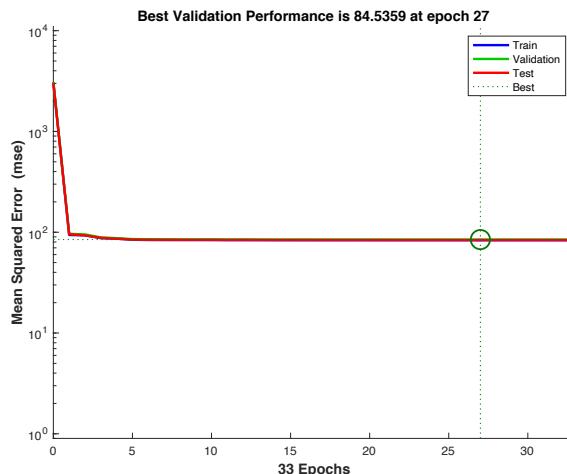
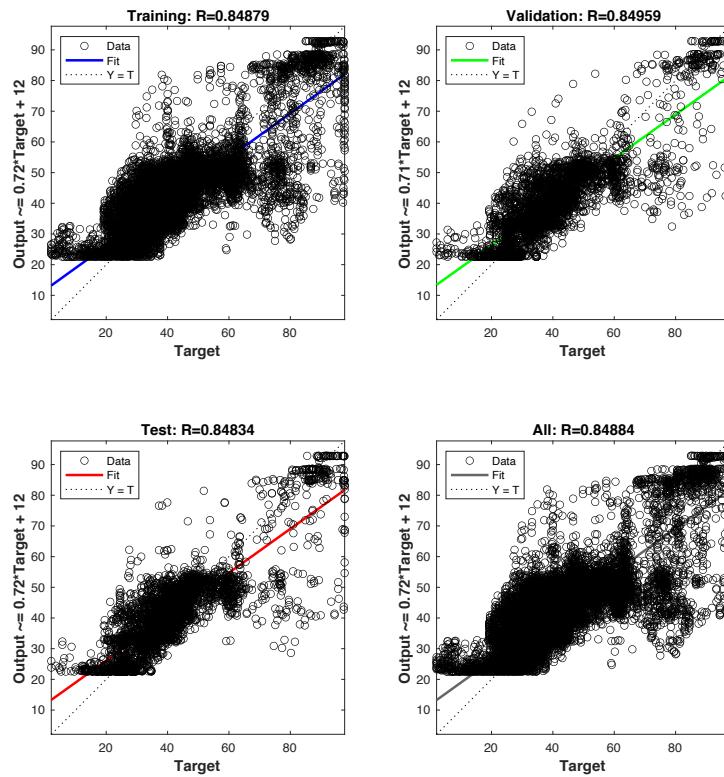


Figure 4: Performance in training of Neural Network in DoA Index Model

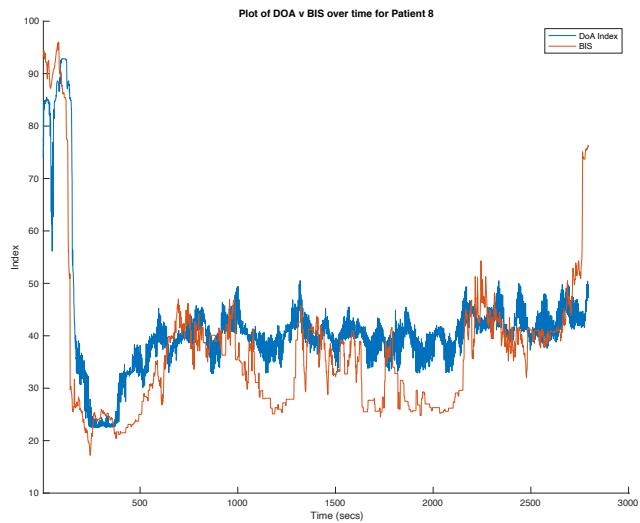


*Figure 5: Regression Plots for neural network in DoA Index Model*

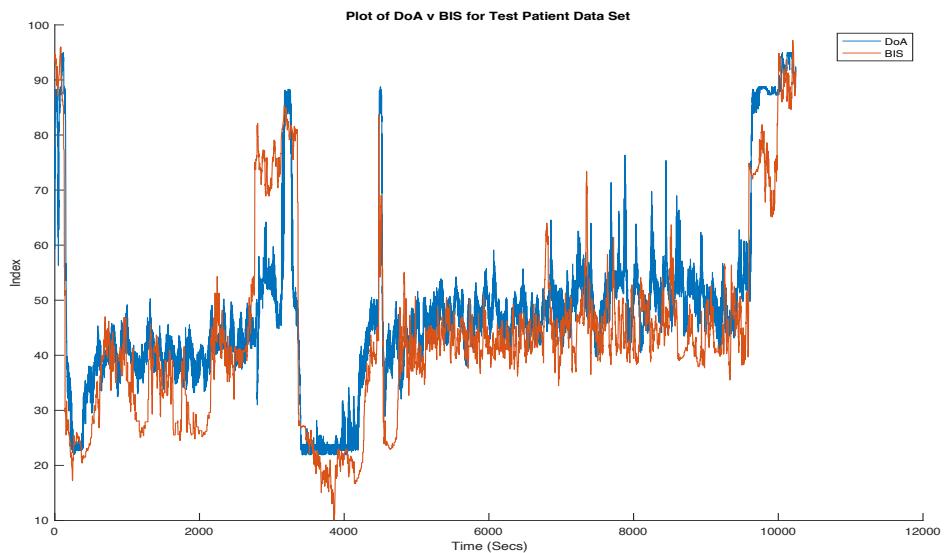
The final results of neural network training were mixed. On examination of the performance plots, there doesn't appear to be any overfitting of the model however the mean squared error of the model is relatively high at 84.5359. On examination of the regression plots the correlation coefficient between target and output results for the model was fairly high at 0.84 for training, test and validation sets however the level of error in the model is also evident when looking at the higher intercept value of 12 in the regression equation.

## 8. Performance Evaluation

To assess performance of the DoA index, the model was applied to the test set of data from patients 8 to 10. The results were then compared to the patients' BIS values through several statistics. The output of the model was also plotted against BIS values to compare results visually. These plots and statistics were generated using the MATLAB platform (Version R2020a, Mathworks 2020) and the code is shown in Appendix E. Below are the plots of the DoA index and BIS scores for patient 8 (Figure 6) and the combined patient data set (Figure 7).



*Figure 6: Plot of DoA Index v BIS over time for Patient 8*



*Figure 7: Plot of DoA Index v BIS for Test Patient Data Set*

As is evident in the distribution of the DoA index and BIS values for patient 8 and also looking at the distribution of the whole test patient data set generally the new index follows BIS relatively well. However there is also some deviation between the new index and the BIS value in some cases this appears to be quite significant. The greatest deviation between results appears to be at higher index and BIS levels or when the patient is conscious.

*Table 3: Performance Statistics for the DoA Index Model*

Pearson Coefficient	R <sup>2</sup>	MAE	RMSE	MSE
0.8365 ± 0.0059	0.6997	7.724	10.1097	102.7266

The Pearson correlation coefficient for the model when it was applied to the test data set was  $0.8365 \pm 0.0059$  indicating a strong relationship between the DoA index of the model and the BIS values for the test patients. However some studies reviewed in the literature did achieve higher correlations with the use of neural networks. Ortolani et al (2002) demonstrated a correlation coefficient of 0.9411 and Gu et al 0.892 (2019). Ortolani et al (2002) had a much larger data set compared to the one used in this case which may have influenced results. The  $R^2$  value of the model is moderate at 0.6997 indicating the 69.9% of the variation in the DoA index is attributed to the input variables in the model. This is a satisfactory result however ideally this value would be higher.

Three measurements of error were calculated for the model. The mean absolute error (MAE) for the model was high at 7.724 and the root mean squared error (RSME) and the mean squared error (MSE) for the model were both very high at 10.1097 and 102.7266 respectively. The RSME and the MSE are much more sensitive to small numbers of larger errors within the data set which are amplified due to the use of the square of the error in calculating the values when compared to MAE (Openclassroom 2020) . As is seen in the distribution plots there are some very large differences in the BIS and DoA curves as is reflected by the MSE and RMSE of the model. Due to their sensitivity to large errors, MSE and RMSE are particularly utilised when large errors are significant in model results (Openclassroom 2020). In this case this is important. Given the scale of our index is for 0 to 100 for consciousness having an error of 10 can have significant implications for how a patient is assessed. Changing model parameters during model building including training methods, number of iterations and error goals did not significantly improve error performance of the model.

On examination of the distribution of BIS values of the whole patient data set, the distribution is skewed to the right indicating a lower number of higher BIS scores or a lesser amount of data for patients that are conscious compared to patients under anaesthesia. On examination of the test patient distribution of the DoA index and BIS values the greatest deviations in the DoA compared to BIS values appear to be moreso in patients with higher BIS scores. Large errors in turn will have the greatest influence on the MSE error of the model. Therefore increasing the amount of data in conscious patients may improve model training. MSE could also be improved through changes in data pre-processing and changes in modelling of neural network input data that may improve model performance.

## 9. Reflection

The first issue encountered with model design was achieving an adequate level for the Pearson coefficient between BIS values and the new DoA index. The initial method trialled for dimension reduction as the first stage of model building was principal components analysis. These results were the used as input for the neural network model similar to final model design. The PCA results were also used as input to an alternative model using support vector regression. Applying PCA to the patient data set resulted in one principal component representing approximately 99% of the variation in the data set thus was effective in simplifying the input for the second machine learning algorithm. However in both cases this yielded a Pearson coefficient of around 0.7 with the index design model comprised of the neural network achieving a slightly higher result. In an attempt to improve this result and alternative model, the linear was applied to the input data. This led to an improvement in the correlation coefficient of approximately 0.1.

The second challenge faced with model building was the size of the error in the model. The first model design using PCA with a neural network or a support vector regression model also resulted in very high values for the mean squared error in both cases of approximately 180. Using the results of the linear model as the input for the second step of model building decreased model mean square

error in the case of the neural network by approximately 100 a significant improvement however the final model error level was still high. When the linear model was combined with a support vector regression model the results were almost improved however the results utilising the neural network were slightly better therefore this combination of machine learning models was selected for index design.

#### 10. Conclusion

The regression model designed to calculate the DoA index had a good correlation with BIS data however did have a large degree of error, more so at higher BIS scores.

## 11. References

1. Al-Masri A, 2019 'How Does Back Propagation in Artificial Neural Networks Work?' *Towards Data Science* weblog post , 30<sup>th</sup> January, viewed 4<sup>th</sup> October 2020, <https://towardsdatascience.com/how-does-back-propagation-in-artificial-neural-networks-work-c7cad873ea7>
2. Battiti R & Brunato M 2017 '*The Lion Way – Machine Learning plus intelligent optimization*', University of Trento, Italy, viewed 13<sup>th</sup> August 2020, <http://intelligent-optimization.org/LIONbook>
3. Diykh M, Miften FS, Abdullah S, Saleh K & Green JH 2020 'Robust Approach to the depth of anaesthesia assessment based on hybrid transform and statistical features' *IET Science Measurement and Technology* Vol 14, No 1, pp 128-136
4. Gu Y, Liang Z & Haghira S, 2019, 'Use of Multiple EEG Features and Artificial Neural Network to Monitor the Depth of Anaesthesia' *Sensors*, Vol 19, No 11:2499 pp 1-12
5. Hashimoto DA, Witkowski E, Gao L, Meireles O & Rosman G 2020 'Artificial Intelligence in Anaesthesiology: Current Techniques, Clinical Applications and Limitations', *Anaesthesiology* Vol 132, No 2, pp 379-394
6. Hernandez-Meza G, Izzetoglu M, Sacan A, Green M & Izzetoglu K (2017) *Neurophotonics*, 4(4), pp 1-10
7. Hiregoudar S 2020 'Ways to evaluate regression models' *Towards Data Science*, weblog post, 5<sup>th</sup> August, viewed 4<sup>th</sup> October 2020, <https://towardsdatascience.com/ways-to-evaluate-regression-models-77a3ff45ba70>
8. Jahanseir M, Setarehdan SK & Momenzadah S 2018 'Estimation of depth of anaesthesia by using a multioutput least square support vector regression' *Turkish Journal of Electrical Engineering and Computer Sciences*, Vol 26, pp 2792-2801
9. Jiang GJA, Fan SZ, Abbot MF, Huang HH, Lan JY, Tsai FF, Chang HC, Yang YW, Chuang FL, Chiu YF, Jen KK & Wu JF 2015 'Ample Entropy Analysis of EEG Signals via Artificial Neural Networks to Model Patients' Consciousness Level Based on Anaesthesiologists Experience' *Biomed Research International* Vol 2015: 343478, pp 1-8
10. Lee HC, Ryu HG, Park Y, Yoon SB, Yang SM, Oh HW & Jung CW, 2019, 'Data Driven Investigation of Bispectral Index Algorithm, *Nature: Scientific Reports*, Vol 9:13769, pp 1-12
11. Liu Q, Ma L, Fan SZ, Abbot MF & Shieh JS 2018, 'Sample Entropy Analysis for the estimating depth of anaesthesia through human EEG signal at different levels of unconsciousness during surgeries *PeerJ*, Vol 6: e4817, pp 1-25
12. Mathworks Inc, 1994-2020 'Support' Mathworks Inc, Natick MA, USA viewed 16<sup>th</sup> August 2020, [https://au.mathworks.com/support.html?s\\_tid=gn\\_supp](https://au.mathworks.com/support.html?s_tid=gn_supp)
13. Ognjanovski G 2019 'Everything you need to know about Neural Networks and Back Propagation – Machine Learning Easy and Fun' *Towards Data Science*, weblog post, 15<sup>th</sup> January, viewed 4<sup>th</sup> October 2020, <https://towardsdatascience.com/everything-you-need-to-know-about-neural-networks-and-backpropagation-machine-learning-made-easy-e5285bc2be3a>

14. Openclassrooms 2020 ‘Evaluating the performance of a regression model’, *Openclassrooms*, weblog, 29<sup>th</sup> June, viewed 4<sup>th</sup> October, <https://openclassrooms.com/en/courses/6401081-improve-the-performance-of-a-machine-learning-model/6519016-evaluate-the-performance-of-a-regression-model>
15. Ortolani O, Conti A, Fillipo D, Adembri C, Moraldi E, Evangelisti A, Maggini M & Roberts SJ 2002, ‘EEG signal processing in anaesthesia. Use of neural network technique for monitoring general anaesthesia’, *British Journal of Anaesthesia*, Vol 88, No 5, pp 644-648
16. Protocol, R 2017, ‘Everything you need to know about Neural Networks’ *Medium*, weblog post, 5<sup>th</sup> December, viewed 4<sup>th</sup> October 2020, <https://medium.com/ravenprotocol/everything-you-need-to-know-about-neural-networks-6fcc7a15cb4>
17. Sadrawi M, Fan SZ, Abbot MF, Jen KK & Sheih JS 2015 ‘Computation Depth of Anaesthesia via Multiple Vital Signs based on Artificial Neural Networks’ *Biomed Research International*, Vol 2015: 536683 pp 1-13
18. St Vincent’s Hospital, 2014 ‘EEG Test’, *St Vincent’s Hospital*, web article, viewed 12<sup>th</sup> August 2020, <https://www.betterhealth.vic.gov.au/health/conditionsandtreatments/eeg-test>
19. Tacke M, Kochs EF, Mueller M, Kramer S, Jordan D & Schneider G (2020) ‘Machine Learning for a combined electroencephalographic anaesthesia index to detect awareness under anaesthesia’ *PLoS ONE*, Vol 15, No 8, pp 1-12

## Appendix A:



Figure 8 : RapidMiner Process for Data Analysis

## MATLAB Code Data Analysis

```

%% Project
%% Clear command window

clc
clear all

%% 
%% Read excel data into matlab
  
```

```

train1 = readtable('Data set (stage 1)', 'Sheet', 'Train1');
train2 = readtable('Data set (stage 1)', 'Sheet', 'Train2');
train3 = readtable('Data set (stage 1)', 'Sheet', 'Train3');
train4 = readtable('Data set (stage 1)', 'Sheet', 'Train4');
train5 = readtable('Data set (stage 1)', 'Sheet', 'Train5');
train6 = readtable('Data set (stage 1)', 'Sheet', 'Train6');
train7 = readtable('Data set (stage 1)', 'Sheet', 'Train7');
train8 = readtable('Data set (stage 1)', 'Sheet', 'Train8');
train9 = readtable('Data set (stage 1)', 'Sheet', 'Train9');
train10 = readtable('Data set (stage 1)', 'Sheet', 'Train10');

eeg = readtable('Data set (stage 1)', 'Sheet', 'Total');

%%
%%%Convert data to arrays and create scatterplot matrix for each patient
and total data set%%
t1 = table2array(train1)
t2 = table2array(train2)
t3 = table2array(train3)
t4 = table2array(train4)
t5 = table2array(train5)
t6 = table2array(train6)
t7 = table2array(train7)
t8 = table2array(train8)
t9 = table2array(train9)
t10 = table2array(train10)

g = table2array(eeg)

%%
%%% Scatter Plot for whole data set
gplotmatrix(g,[]);

%% Correlation matrix for whole data set
coveeg = cov(g)
correeg = corrcoef(coveeg)

%%Descriptive Statistics
Meang = mean(g);
SDg = std(g);
summary (eeg);

```

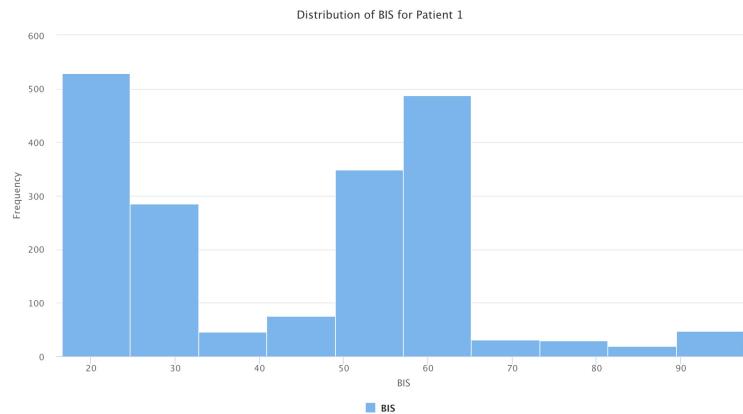
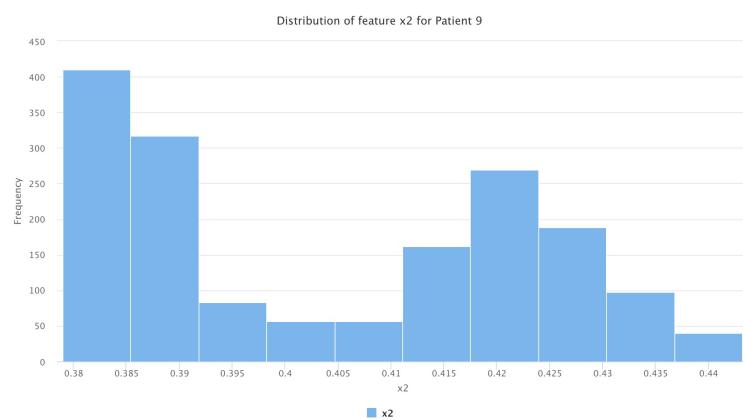
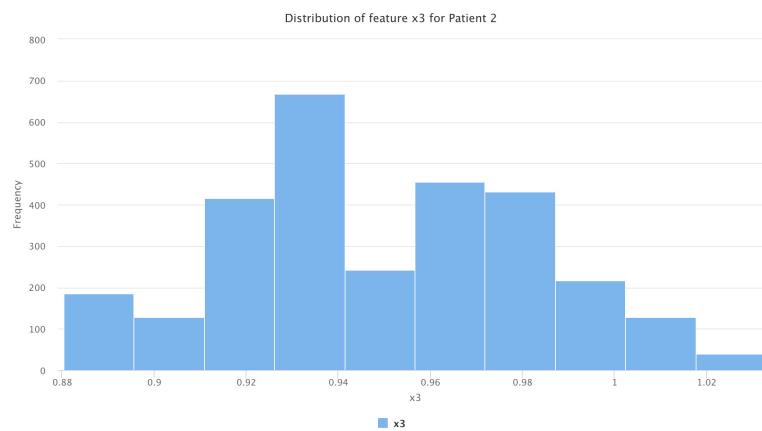
**Appendix B***Table 4 : Descriptive Statistics for Individual Patient Data Sets*

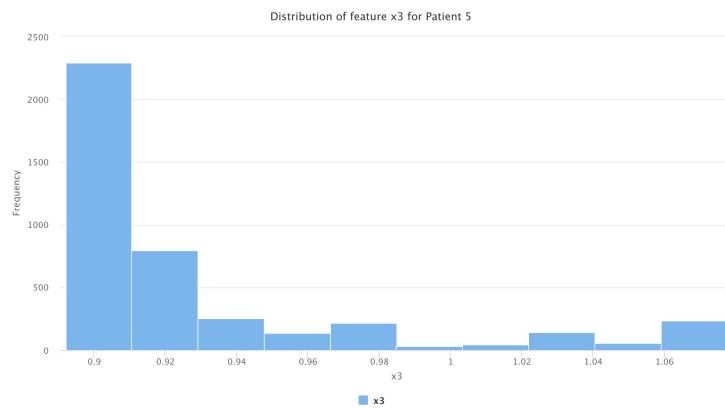
Feature	Min	Max	Mean	Standard Deviation	Range
<b>Patient 1</b>					
BIS	16.6	97.6	43.871	19.934	81
X1	1.755	1.763	1.759	0.002	0.008
X2	0.376	0.438	0.399	0.017	0.062
X3	0.887	1.004	0.922	0.028	0.117
X4	1.722	1.787	1.766	0.013	0.065
X5	1.095	5.976	2.42	0.805	4.881
<b>Patient 2</b>					
BIS	19.4	93.9	36.766	17.705	74.5
X1	1.769	1.78	1.775	0.003	0.011
X2	0.378	0.433	0.401	0.011	0.055
X3	0.88	1.033	0.95	0.032	0.153
X4	1.718	1.789	1.773	0.014	0.071
X5	0.306	6.859	1.606	1.078	6.553
<b>Patient 3</b>					
BIS	20.3	96.1	43.112	14.655	75.8
X1	1.767	1.773	1.77	0.002	0.006
X2	0.376	0.427	0.39	0.008	0.051
X3	0.879	1.025	0.915	0.039	0.146
X4	1.734	1.786	1.766	0.008	0.052
X5	0.266	3.98	1.24	0.681	3.714
<b>Patient 4</b>					
BIS	2.1	97.7	46.039	17.87	95.6
X1	1.72	1.761	1.741	0.012	0.041
X2	0.376	0.438	0.39	0.01	0.062
X3	0.879	1.027	0.908	0.035	0.148
X4	1.732	1.788	1.764	0.01	0.056
X5	0.732	4.893	2.192	0.688	4.161
<b>Patient 5</b>					
BIS	19.2	97.7	48.417	21.015	78.5
X1	1.771	1.778	1.775	0.002	0.007
X2	0.377	0.443	0.391	0.011	0.066
X3	0.892	1.078	0.931	0.048	0.186
X4	1.685	1.786	1.76	0.021	0.101

X5	1.691	8.8	3.305	1.556	7.109
<b>Patient 6</b>					
BIS	5.3	89.5	41.54	16.132	84.2
X1	1.765	1.773	1.769	0.003	0.008
X2	0.378	0.432	0.394	0.011	0.054
X3	0.8	1.053	0.921	0.04	0.253
X4	1.726	1.788	1.768	0.01	0.062
X5	0.707	4.71	1.841	0.685	4.003
<b>Patient 7</b>					
BIS	20.8	97.7	36.14	10.919	76.9
X1	1.765	1.772	1.768	0.002	0.007
X2	0.379	0.44	0.394	0.008	0.061
X3	0.861	1.026	0.912	0.045	0.165
X4	1.711	1.787	1.776	0.008	0.076
X5	0.41	6.357	1.443	1.078	5.947
<b>Patient 8</b>					
BIS	17.2	96	37.201	14.563	78.8
X1	1.761	1.761	1.761	0	0
X2	0.381	0.433	0.394	0.008	0.052
X3	0.886	1.076	0.921	0.036	0.19
X4	1.695	1.787	1.765	0.011	0.092
X5	1.23	9.57	1.882	1.104	8.34
<b>Patient 9</b>					
BIS	10.1	85.4	41.893	25.869	75.3
X1	1.778	1.78	1.779	0.001	0.002
X2	0.379	0.443	0.405	0.019	0.064
X3	0.883	0.988	0.925	0.024	0.105
X4	1.731	1.787	1.77	0.013	0.056
X5	1.222	5.427	2.136	0.737	4.205
<b>Patient 10</b>					
BIS	22.9	97.2	47.646	13.239	74.3
X1	1.76	1.766	1.763	0.002	0.006
X2	0.376	0.413	0.387	0.005	0.037
X3	0.887	1.064	0.92	0.043	0.177
X4	1.68	1.782	1.755	0.019	0.102
X5	1.585	7.921	2.507	1.159	6.336

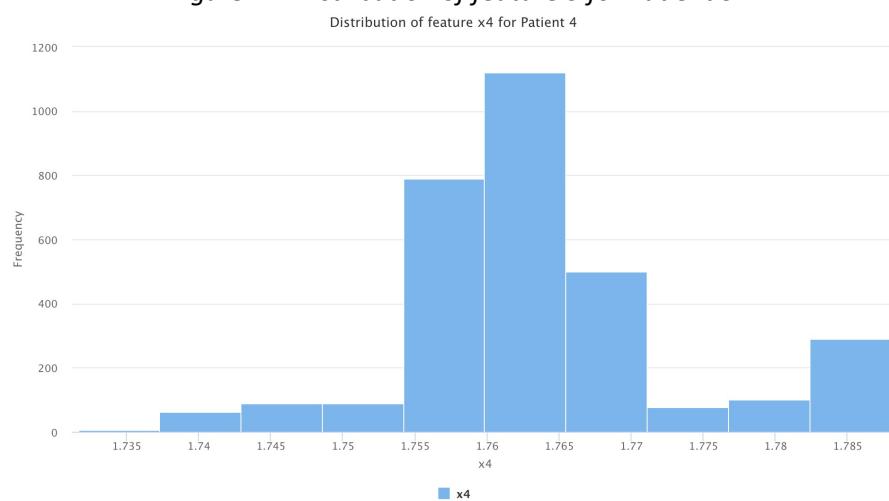
*Table 5: Descriptive Statistics of Combined Patient Data Set*

	Min	Max	Mean	Standard Deviation	Range
BIS	2.1	97.7	42.417	17.210	95.6
X1	1.720	1.780	1.765	0.010	0.06
X2	0.376	0.443	0.393	0.011	0.067
X3	0.800	1.078	0.922	0.041	0.278
X4	1.680	1.790	1.766	0.015	0.11
X5	0.266	9.570	2.07	1.217	9.304

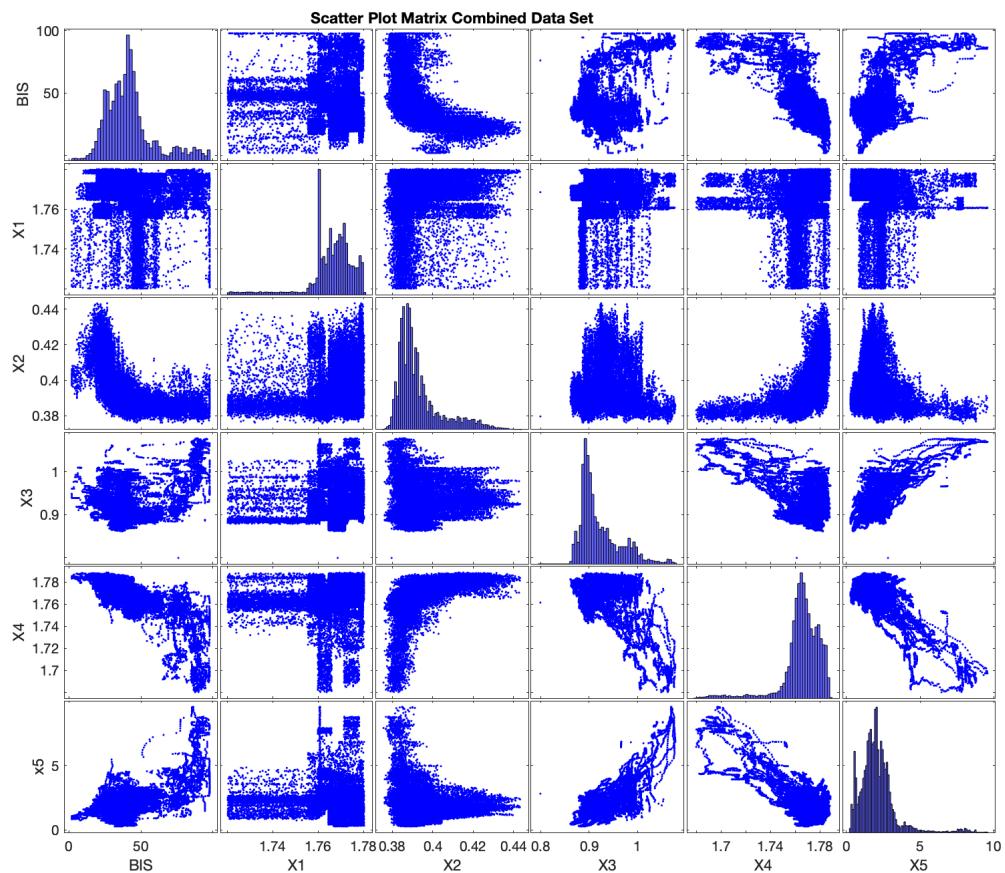
**Appendix C:***Figure 9: Distribution of BIS for Patient 1**Figure 10 : Distribution of feature x2 for Patient 9**Figure 11: Distribution of feature x3 for Patient 2*



*Figure 12: Distribution of feature 5 for Patient 5*



*Figure 13: Distribution of feature x4 for Patient 4*



*Figure 14 : Scatterplot Matrix Combined Patient Data Set*

Appendix D:*Table 6: Correlation matrix Patient 1*

Attribu...	BIS	x1	x2	x3	x4	x5
BIS	1	-0.050	-0.833	0.006	-0.863	0.708
x1	-0.050	1	0.035	-0.035	0.065	-0.054
x2	-0.833	0.035	1	0.263	0.803	-0.539
x3	0.006	-0.035	0.263	1	-0.132	0.501
x4	-0.863	0.065	0.803	-0.132	1	-0.826
x5	0.708	-0.054	-0.539	0.501	-0.826	1

*Table 7: Correlation matrix Patient 2*

Attribu...	BIS	x1	x2	x3	x4	x5
BIS	1	-0.002	-0.595	0.071	-0.841	0.501
x1	-0.002	1	-0.007	-0.010	0.001	0.004
x2	-0.595	-0.007	1	0.408	0.675	-0.017
x3	0.071	-0.010	0.408	1	-0.015	0.655
x4	-0.841	0.001	0.675	-0.015	1	-0.415
x5	0.501	0.004	-0.017	0.655	-0.415	1

*Table 8: Correlation matrix Patient 3*

Attribu...	BIS	x1	x2	x3	x4	x5
BIS	1	-0.018	-0.420	-0.015	-0.713	0.615
x1	-0.018	1	-0.006	-0.028	0.011	-0.039
x2	-0.420	-0.006	1	0.325	0.555	0.030
x3	-0.015	-0.028	0.325	1	0.257	0.569
x4	-0.713	0.011	0.555	0.257	1	-0.374
x5	0.615	-0.039	0.030	0.569	-0.374	1

*Table 9: Correlation matrix Patient 4*

Attribu...	BIS	x1	x2	x3	x4	x5
BIS	1	-0.000	-0.496	0.004	-0.808	0.860
x1	-0.000	1	0.000	-0.003	-0.001	-0.002
x2	-0.496	0.000	1	0.464	0.651	-0.396
x3	0.004	-0.003	0.464	1	0.167	0.165
x4	-0.808	-0.001	0.651	0.167	1	-0.743
x5	0.860	-0.002	-0.396	0.165	-0.743	1

*Table 10: Correlation matrix Patient 5*

Attribu...	BIS	x1	x2	x3	x4	x5
BIS	1	0.007	-0.541	0.668	-0.826	0.754
x1	0.007	1	-0.026	-0.013	0.010	-0.008
x2	-0.541	-0.026	1	-0.111	0.473	-0.328
x3	0.668	-0.013	-0.111	1	-0.835	0.902
x4	-0.826	0.010	0.473	-0.835	1	-0.945
x5	0.754	-0.008	-0.328	0.902	-0.945	1

*Table 11: Correlation matrix Patient 6*

Attribu...	BIS	x1	x2	x3	x4	x5
BIS	1	0.010	-0.652	-0.002	-0.781	0.400
x1	0.010	1	0.011	0.021	-0.003	0.020
x2	-0.652	0.011	1	0.283	0.752	-0.022
x3	-0.002	0.021	0.283	1	0.195	0.749
x4	-0.781	-0.003	0.752	0.195	1	-0.172
x5	0.400	0.020	-0.022	0.749	-0.172	1

*Table 12: Correlation matrix Patient 7*

Attribu...	BIS	x1	x2	x3	x4	x5
BIS	1	-0.009	-0.343	0.231	-0.711	0.464
x1	-0.009	1	0.012	0.003	-0.004	-0.001
x2	-0.343	0.012	1	0.149	0.292	-0.032
x3	0.231	0.003	0.149	1	-0.383	0.867
x4	-0.711	-0.004	0.292	-0.383	1	-0.685
x5	0.464	-0.001	-0.032	0.867	-0.685	1

*Table 13: Correlation matrix Patient 8*

Attribu...	BIS	x1	x2	x3	x4	x5
BIS	1	-0.028	-0.476	0.467	-0.736	0.644
x1	-0.028	1	0.030	0.001	0.016	-0.008
x2	-0.476	0.030	1	0.094	0.515	-0.138
x3	0.467	0.001	0.094	1	-0.524	0.873
x4	-0.736	0.016	0.515	-0.524	1	-0.802
x5	0.644	-0.008	-0.138	0.873	-0.802	1

*Table 14: Correlation matrix Patient 9*

Attribu...	BIS	x1	x2	x3	x4	x5
BIS	1	-0.007	-0.809	-0.070	-0.831	0.714
x1	-0.007	1	-0.005	-0.028	0.014	-0.022
x2	-0.809	-0.005	1	0.408	0.823	-0.496
x3	-0.070	-0.028	0.408	1	0.052	0.428
x4	-0.831	0.014	0.823	0.052	1	-0.811
x5	0.714	-0.022	-0.496	0.428	-0.811	1

*Table 15: Correlation matrix Patient 10*

Attribu...	BIS	x1	x2	x3	x4	x5
BIS	1	0.012	-0.396	0.693	-0.837	0.816
x1	0.012	1	-0.015	0.005	-0.005	0.006
x2	-0.396	-0.015	1	-0.060	0.392	-0.207
x3	0.693	0.005	-0.060	1	-0.628	0.894
x4	-0.837	-0.005	0.392	-0.628	1	-0.790
x5	0.816	0.006	-0.207	0.894	-0.790	1

*Table 16: Correlation Matrix Combined Data Set*

	BIS	X1	X2	X3	X4	X5
BIS	1	-0.0556	-0.5717	0.2598	-0.7641	0.6291
X1	-0.0556	1	0.1601	0.1369	0.1057	-0.0086
X2	-0.5717	0.1601	1	0.1940	0.5387	-0.2172
X3	0.2598	0.1369	0.1940	1	0.3307	0.6693
X4	-0.7641	0.1057	0.5387	0.3307	1	0.7228
X5	0.6291	-0.0086	-0.2172	0.6693	0.7228	1

## Appendix E

### *MATLAB Code for Feature Selection, Model Building and Model Performance Evaluation*

```

%%
%%% Create a linear model of the patient data set
%%%%
X = [eeg.x2, eeg.x3, eeg.x4, eeg.x5];
mdl = fitlm(X, eeg.BIS);

%%%
%%%Parameter selection based on data analysis and weights in linear model
Y = [eeg.x2, eeg.x4, eeg.x5];

%%%
%%%Apply linear model to data set the use these parameters to build NN model

%Calculate inputs from linear model

lin = fitlm(Y, eeg.BIS);

linResult = predict(lin,Y);

TrainInput = transpose(linResult(1:24136, 1));
TestInput = transpose(linResult(24137:34377, 1));

%%%Target values from eeg data

TrainResult = transpose(g(1:24136, 1));
TestResult = transpose(g(24137:34377, 1));

%%%
%%%Build neural Network
netconf = [10]

net = feedforwardnet(netconf, 'trainlm');
net = train(net,TrainInput,TrainResult);

net.trainParam.epochs=1000;
net.trainParam.goal=0.05;
net.trainParam.show=50;
net.trainParam.lr=0.05;
net.trainParam.mc=0.8;
net.divideFcn= 'dividerand';

%%%Calculate model performance
predT = sim(net,TestInput);

[R,P,RL,RU] = corrcoef(predT,TestResult);

e = TestResult-predT;
perf = mae(e); %%MAE

err = mse(net,TestResult,predT) %%MSE

```

```
%%
% % Plot Results of test set

figure; hold on;
plot (predT);
plot(TestResult);

%%
% Apply model to a test patient 8

%Calculate inputs from linear model
Z = [train8.x2, train8.x4, train8.x5]
W = [train8.BIS];

linResult2 = predict(lin,Z);

Train8Input = transpose(linResult2);

pred8 = sim(net,Train8Input);

%%
%%Plot Results of Patient 8
figure; hold on;
plot (pred8);
plot(W);
```