# Using Machine Learning to Predict Decisions on EEG Data

Bachelor's Project Thesis

Sukhleen Kaur, s3157423, s.kaur@student.rug.nl,
Supervisor: Dr. M.K. van Vugt

**Abstract:** Many studies have been conducted to determine how decisions are made in the brain. The Drift Diffusion Model (DDM) proposes an explanation for this. Studies have then tried to apply machine learning classification to determine whether decision making processes exhibit activity as described by the DDM. Athough these studies find classification accuracies above chance, they still tend to be low. Hence, this study investigates whether these results can be improved using different machine learning classification techniques. This study classifies intracranial EEG (iEEG) data collected while participants performed perceptual and memory decision making tasks. Fisher's Linear Discriminant Analysis (LDA), a Support Vector Machine (SVM) with Radial Basis Function (RBF) kernel and a SVM with Polynomial kernel were fit onto this data and their classification was evaluated. It was found that classification was done above chance only for a few participants ($M_{accuracy} > 60\%$). The classifiers did not perform well, their F1 scores only ranged between 0.52 to 0.54 across the tasks. Polynomial SVM ($\mathrm{AUC}_{\mathrm{Poly}} = 0.55, \mathrm{F1} = 0.54$) classified perceptual decisions best and RBF SVM ($\mathrm{AUC}_{\mathrm{RBF}} = 0.53, \mathrm{F1} = 0.53$) classified memory decisions best compared to the other classifiers based on their ROC curves and F1-scores. The reason for a low classification rate is because of low amounts of data and because of inconsistency of the data (difference in electrode placements). In conclusion, this study was not able to classify the decisions with significant accuracy.

## 1 Introduction

Decision making is a cognitive process that chooses a preferred option among a set of alternatives (Wang et al., 2006). People can make decisions over a long period of time or even in a few seconds. The speed at which people make decisions depends on what the decision is being made. People tend to take longer in making decisions regarding, for instance, moving to a different country or choosing between two jobs. People take longer with these decisions since they have a bigger consequences. They tend to take less time in making decisions, for instance, choosing between whether to go right or left. In order to be able to able to collect data in a laboratory and monitor activity in the brain throughout the entire decision making process, i.e., from the presentation of options till the choice is made, it would be important that the decision making is that of a fast type. Hence, this paper looks into the fast decisions that people can make.

The Drift Diffusion Model (DDM) is a popular model that propose an explanation as to how a decision that does not involve a sequence of reasoning is made. That is, DDM proposes an explanation of how a decision is made for simple two-choice decision processes (Ratcliff and McKoon, 2008). This model suggests integrators accumulate information overtime for one choice over the other and it gradually increases activity. This accumulation is noisy. The rate at which this increase occurs is called the drift rate (Ratcliff and McKoon, 2008) and it is suggested that with clearer evidence, the activity increases faster and vice versa. If this activity surpasses a certain decision threshold, the decision is made. (Vasily, 2018). There also exists another parameter in the DDM which specifies the non-decision time. This depicts actions that do not relate to decision making, for instance, the response selection or perception of the situation. This model,

when given a particular response time distribution and accuracy can estimate the model parameters, such as drift rate and non-decision time, through model fitting. As a consequence, a characteristic response time distribution is given which is right skewed, similar to what is observed in human experiments (Ratcliff and McKoon, 2008).

The DDM was initially based on behavioural data but some studies suggest some neural activity in the brain does resemble the DDM. In an random-dot experiment where monkeys had to choose in which direction the dots were moving in, Shadlen and Newsome (1996) found that while the monkey made decisions, the neurons in the lateral intraparietal area carried signals which can be used to predict the decision the monkey was going to make. They found that when many dots were moving in the same direction, the decision was made faster. This finding is consistent with DDM.

Studies with evidence accumulation has not only taken place in monkeys but humans as well. Mulder et al. (2014) found that evidence accumulation is linked with fronto-parietal network and that the decision threshold is linked with fronto-basal ganglia networks. Both of these are found for choice biases. Another study by Ho et al. (2009) also showed evidence for neural activity in cortical areas before the decision is made. Evidence accumulation in humans have also been studied through EEG experiments. In 2012, van Vugt et al. found that EEG oscillations, occurring in the parietal 4-9Hz theta band, show activity as expected with evidence of accumulation where sufficient information is increasingly gathered until one option is favored over another. They also showed that cortical theta oscillations have been linked to perception and the hippocampal theta oscillations have been linked to memory. Hence, in their experiment, they asked their participants to perform perceptual and memory decision making tasks in order to study evidence accumulation.

Other studies, mostly marketing studies, have used different machine learning techniques to predict consumers' choices/preferences. One of these studies (Hakim et al., 2018), recorded the EEG activity of participants while they ranked 6 items from most preferred to least preferred. They then applied a Support Vector Machine, a Logistic Regression, a K-Nearest Neighbour and Decision Tree classifiers onto this data and found that they were able to predict decisions with an accuracy of 67%.

In another investigation done by van Vugt et al. (2018), they applied machine learning to examine whether certain decision making processes exhibit activity as described by the DDM. A Logistic Regression classifier was applied to the data and it was found to have a higher accuracy for the perceptual task than the memory task ($M_{perceptual} = 0.64$, $M_{memory} = 0.61$). Given this, *can the outcome of a decision making process be predicted using EEG?*

Since the accuracy of the logistic regression classifier applied in the study conducted by van Vugt et al. (2018) is quite low, this paper investigates whether other classifiers such as SVMs and LDAs can perform better.

Some of the most popular Brain-Computer Interface (BCI) systems are based on EEG. This makes for a good comparison since BCI systems are becoming increasingly popular and lot of machine learning classifications are applied on data retrieved from BCI systems. Kołodziej et al. (2012) found that Fisher's Linear Discriminant Analysis (LDA) provide good results with BCI systems that analyze EEG signal. Lotte et al. (2007) also suggests that LDAs tend to provide good results with BCI systems. Although LDA and logistic regressions have the same underlying mechanism, logistic regression needs longer time for computation. Logistic regression need a bigger sample size but this restriction does not apply to LDA. Support Vector Machines (SVM) have also shown good results with BCI systems and they tend to have good generalization properties (Lotte et al., 2007). They also require less data than logistic regression models and SVMs tend to have higher accuracy than logistic regression as well. (Verplancke et al., 2008). Given this, LDA and SVM makes for good classifiers to investigate in this study. This classification will then provide more information about decision making in general and also about neural correlates of the DDM.

# 2 Methods

The data-set used for the classification was collected by van Vugt et al. (2018) hence, the subsections preceding the classification descriptions follow the information provided by them.

## 2.1 Participants

A total of 16 participants took part in the experiment. All of the participants were suffering from epilepsy and as a consequence had invasive electrodes placed in their brain within a period of 1-3 weeks, prior to the operation to figure out where their seizures seizure comes from. These electrodes allow for retrieving intracranial EEG (iEEG) data and also allow for better spatial resolution compared to scalp EEG. The participants were undergoing their treatment at the Freiburg University Hospital in Germany.

## 2.2 Tasks

The data-set of these participants consists of data collected through two different types of tasks; a memory task and a perceptual task. To do this, synthetic faces were created using Basel Face Model (Paysan et al., 2009). This allowed for manipulating the similarity between the faces and as a consequence the difficulty of the tasks.

In the perceptual task, participants were presented with an image containing two faces which were shifted outwards (Figure. 2.1a). They then had to determine whether the two faces belonged to the same person. In the memory task, participants were first presented with an image containing two faces for 2000-2075ms and after a delay of 1000-1075ms, they were presented with an image containing one face (Figure. 2.1b). They then had to determine whether the face presented on the second image was one from the image shown previously, which contained the two faces. These tasks were chosen so different kind of decision processes can be compared since in the perception task, participants would need to mentally rotate the image whereas in the memory task they would need to recover the image from their memory of the first image.

## 2.3 Data Recordings

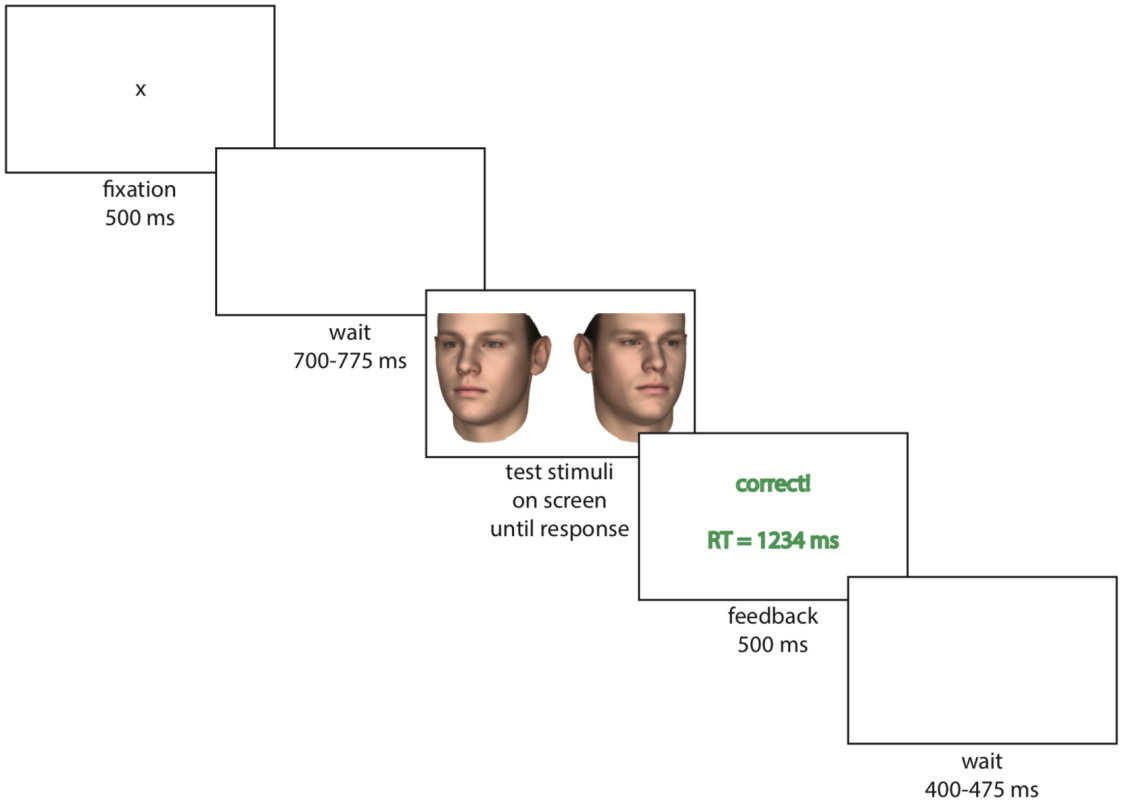
The data were recorded with a 2000Hz sampling rate on a clinical EEG recording system. Each of the trials performed by the participant was then divided into 4000ms duration, starting 200ms before the stimulus was presented to the participants. Then, epileptic spikes were checked for. These epileptic spikes are deviant from the data and these deviations end up at the tails of the data distribution. Hence, kurtosis was used find these epileptic spikes which measures the values at the tails of the distribution. If this value of kurtosis was more than 15, the trial was removed. Also, if the response time was more than 3800ms the trial was removed. Participant 3 was removed because there were epileptic spikes throughout their entire data-set. This data recordings only consisted of the correct trials since supervised learning is applied which require labels containing the correct outputs (Maglogiannis, 2007).

## 2.4 Data Pre-Processsing

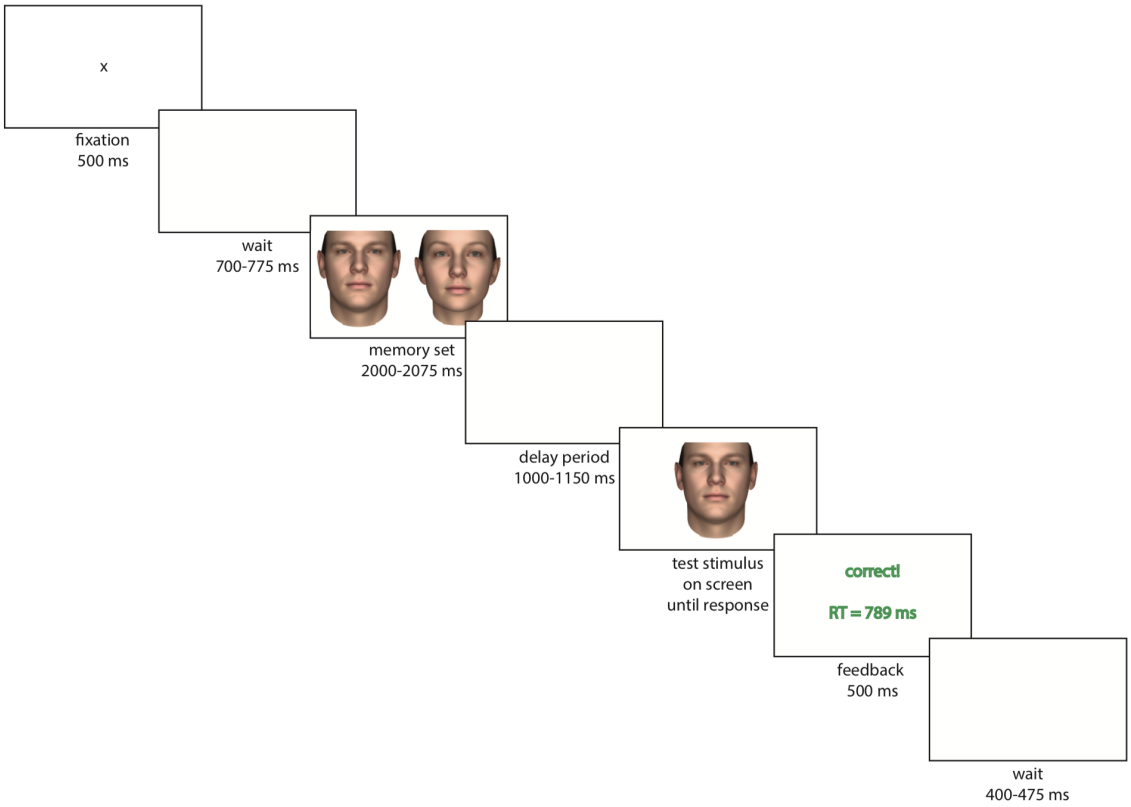
The data were pre-processed using Matlab code based on toolboxes developed by Jelmer Borst and Per Sederberg. This data-set has two classes; match, where participants reported seeing the same face and non-match, where participants reported seeing a different face. The classifiers are trained to distinguish between these two classes. The two classes were then separated such that there were equal number of trials within each class. This was done by randomly removing trials from larger classes. Then, wavelet transformation was performed in the 4-9Hz theta band which is said to be important for decision making according to van Vugt et al. (2012). The data were then divided into equal number of bins between the stimulus and the response. By this, regardless of the response time, the number of bins for each task remained the same. For the perception task 50 bins exist and for the memory, 32 bins. Each of these bins, on average, were of 50ms duration since the focus of the study was primarily in the theta band in the range of 4-9Hz. This short windows also avoids missing any critical information in this theta band.

## 2.5 Classification

In order to determine the neural activity associated with decision making, supervised learning classifiers were fit on the binned data to distinguish between the match and non-match decisions. The data were classified in Python 3 using Scikit-learn. The classifiers were run on the two tasks separately in order to also determine whether there are any




**(a) Perception Task**






**(b) Memory Task**

**Figure 2.1: Example trials of the two tasks.**

differences between the classification of perceptual and memory task.

### 2.5.1 Data Set-Up

As mentioned earlier, the data-set was divided into 50 bins for the perception task and 32 bins in the memory task. Each classification instance was fit onto each of these bins.

For a classifier to perform well and be able to generalize, it needs to have a low bias and low variance. Bias is the difference between the average predicted decisions and the correct decisions that has to be predicted. Having a high bias would lead to over-simplification of the model and hence would lead to errors in classification of the data-sets. Variance is how much the a decision varies, it provides the spread of the data. Having a high variance would lead to over-fitting of the data and hence terrible generalization. Hence, the classifiers were cross validated.

The classifiers were cross validated with 8 folds. Here, the data were divided into 8 subsets where each subset consists of a set of time bins. Each time, one subset is kept apart for testing and the data is trained on the 7 subsets. This is iterated over the number of folds, in this case 8. The error estimates are then averaged over the 8 iterations. 8 fold was chosen since there were not a lot of data samples and so dividing a small sample into large number of folds would mean that the number of iterations that are different will be less. Hence in this case, 8 fold is not too large. Also, it is not too large to become computationally expensive but it is large enough to have less bias compared to having fewer folds. Cross validating helped reduce bias since a lot of the data was used for fitting and also to reduce variance since it was validated on most of it too.

### 2.5.2 Data Classification

Supervised learning is used to classify decisions as labeled data were available. These labeled data consisted of `[-1, 1]` labels where `[-1]` was the non-match decision label and `[1]` was the match decision label. These labels were classified using three different types of classifiers as mentioned earlier in Section 1; Fisher's Linear Discriminant Analysis, Support Vector Machine with a Radial Basis Function kernel and a Polynomial kernel. Different classifiers are used in order to compare them and hence determine which one would be best in determining decisions.

**Fisher's Linear Discriminant Analysis (LDA):** In LDA, the aim is to find the linear function that maximizes the distance between the projected means normalised by the within class variance (Elhabian and Farag, 2009). This means a projection will be found such that examples from the same class (match/ non-match) will be projected close to each other whereas the mean between the two classes will be maximized. This projection is then given as (Quarizmi, 2015):

$$\mathbf{w} = (n_{\mathrm{m}}\Sigma_{\mathrm{m}} + n_{\mathrm{nm}}\Sigma_{\mathrm{nm}}^{-1})(\boldsymbol{\mu}_{\mathrm{m}} - \boldsymbol{\mu}_{\mathrm{nm}})$$

Where:

- $\mathbf{w}$ is the projection of the data
- $n_m, n_{nm}$ is the number of observations in the match and non-match class
- $\Sigma_{\mathrm{m}}, \Sigma_{\mathrm{nm}}$ is the co-variance matrices of the distributions in the match and non-match class after projection
- $\boldsymbol{\mu}_{\mathbf{m}}, \boldsymbol{\mu}_{\mathbf{nm}}$ is the mean vectors of the distributions in the match and non-match class after projection

Linear classifiers are the most popular when it comes to classification of EEG data (Lemm et al., 2011) and so LDA makes for a good classifier for this experiment. According to James et al. (2017), LDA is stable when classes are well separated contrary to linear classifiers that use logistic regression. LDA is also stable when it comes to classifying with small amounts of data, which is the case in this experiment.

**Support Vector Machine (SVM):** The basis of SVMs is to find a hyper-plane that divides the classes the best. This hyper-plane maximizes the distance, $\gamma$, between the hyper-plane and points in the training set (Leskovec et al., 2015). SVMs also use a penalty parameter $C$, which is set in order to state how much misclassification should be avoided.

In this experiment, a Radial Basis Function Kernel and the Polynomial Kernel was used. These

two were chosen since they are non-linear and to investigate if non-linear classifiers can perform better than linear ones. The difference between these two kernels lie in the fact that they describe the hyper-plane differently as described below.

**Radial Basis Function Kernel (RBF):** An RBF kernel is defined as:

$$K(x, y) = \exp(-\gamma||x - y||^2)$$

where $x$ and $y$ are feature vectors in some input space and $\gamma$ is the spread of the kernel. The penalty parameter was $C = 4$ since it was not too big to cause over-fitting and not too small such that the classifier made much misclassifications.

**Polynomial Kernel:** A polynomial kernel is defined as:

$$K(x, y) = (x^T y + a)^d$$

where $x$ and $y$ are feature vectors in some input space, $a$ is a free parameter that trade-off the influence of high-order and low-order terms (Crowley, 2016) and $d$ is the order of the kernel. In this experiment, the penalty parameter was $C = 2$ for the same reason as above. $C = 4$ was not chosen since it did not make a difference and it is better not to have a higher value. Polynomial kernel of degree $d = 2$ was chosen because then the kernel is not linear and also because higher-degrees tend to over-fit.

## 2.6 Classification Evaluation

It is important to not just look at the accuracy (Eq. 2.1), which states the ratio of the correctly predicted decisions and the total number decisions, since this is most useful only when the data has the same number of false positives and false negatives. Hence, the F1 score (Eq. 2.4) is determined which is the weighted average of precision (Eq. 2.2) and recall (Eq. 2.3). Precision is defined to be the ratio between true positives and all of the positive decisions. Recall is defined as the ratio between true positives and all the decisions in the matching class. The F1 score is considered to be good if the values are far above 0.50. Hence, the classifier will be evaluated by looking at the accuracy, precision, recall and therefore also the F1 score.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{2.1}$$

Where:

- TP : True Positives
- TN : True Negatives
- FP : False Positives
- FN : False Negatives

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{2.2}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{PN}} \tag{2.3}$$

$$\text{F1-Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{2.4}$$

The classifiers will also be evaluated by plotting the Receiving Operating Characteristics (ROC) curves. The ROC curves are plotted with True Positive Rate (TPR) against False Positive Rate (FPR). If the classifiers have a higher TPR then they are said to classify well. The Area Under these Curves (AUC) represent how well the match and non-match classes can be distinguished. If the AUC is 1 then the classification is said to be perfect, if the AUC is 0.5 then the classification is said to have happened by chance and the classifier is not considered to have a good performance.

# 3 Results

In order to examine whether decisions can be predicted, first, how the classifiers performed is determined. Then, the best classifier for each task is determined and finally, classification for individual participants is looked at.

## 3.1 Classification Results

### 3.1.1 Accuracy, Precision, Recall and F1 Scores

The accuracy, precision, recall and F1 score for the perception task can be found in Table 3.1. These values were evaluated based on the confusion matrices shown in Figure 3.1. Table 3.1 shows that the classifiers do not perform well. The F1 scores range

from 0.52 to 0.54, close to chance level. This shows that the classifier is not that precise and also is not that robust.

**Table 3.1: Accuracy, precision, recall and F1 score of the different classifiers in the perception task.**

| Classifier | Accuracy | Precision | Recall | F1 Score |
|---|---|---|---|---|
| LDA | 0.53 | 0.52 | 0.52 | 0.52 |
| RBF | 0.54 | 0.54 | 0.54 | 0.54 |
| Poly | 0.55 | 0.55 | 0.55 | 0.55 |

The F1 scores of the memory task are similar to that of the perception task (Table 3.2). These values were evaluated based on the confusion matrices shown in Figure 3.2. The F1 score for the memory task again, does not perform well. They range from 0.52 to 0.53, close to chance level. This again, shows that the classifier is not that precise and also is not that robust.

**Table 3.2: Accuracy, precision, recall and F1 score of the different classifiers in the memory task.**

| Classifier | Accuracy | Precision | Recall | F1 Score |
|---|---|---|---|---|
| LDA | 0.52 | 0.52 | 0.52 | 0.52 |
| RBF | 0.53 | 0.53 | 0.53 | 0.53 |
| Poly | 0.52 | 0.52 | 0.52 | 0.52 |

### 3.1.2 ROC Curve

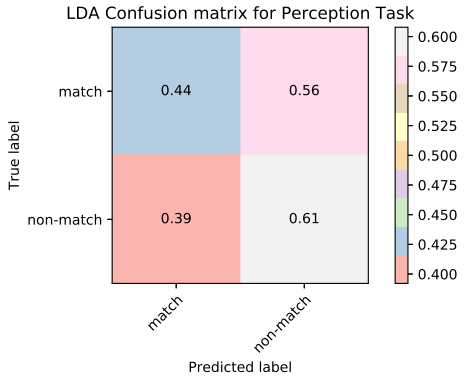
The Receiving Operating Characteristics (ROC) curves can be seen in Figure 3.3. The AUC for each of the classifiers can also be seen on Figure 3.3. In general the classifiers do not perform well ($\text{AUC}_{min} = 0.52$, $\text{AUC}_{max} = 0.55$) as was seen with the F1 scores.

For the perception task (Figure 3.3a), Polynomial SVM performs the best ($\text{AUC}_{\text{Poly}_{\text{Perception}}} = 0.55$) compared to RBF SVM ($\text{AUC}_{\text{RBF}_{\text{Perception}}} = 0.54$) and LDA ($\text{AUC}_{\text{LDA}_{\text{Perception}}} = 0.52$).

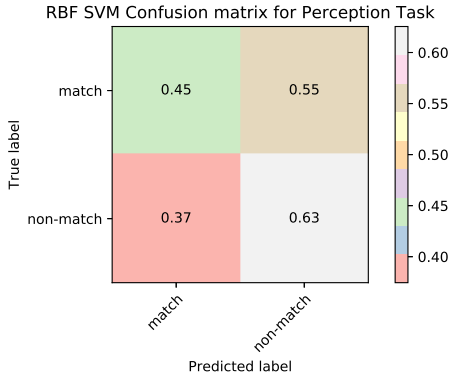
For the memory task (Figure 3.3b), RBF SVM performs the best ($\text{AUC}_{\text{RBF}_{\text{Memory}}} = 0.53$) compared to Polynomial SVM ($\text{AUC}_{\text{Polynomial}_{\text{Memory}}} = 0.52$) and LDA ($\text{AUC}_{\text{LDA}_{\text{Memory}}} = 0.52$).

## 3.2 Task Results

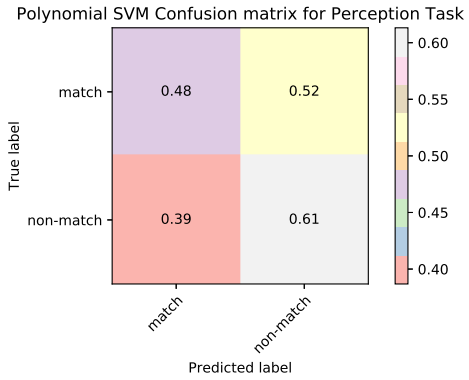
Table 3.3 shows the result of the Bayes Factor Analysis done over the classifiers in between the two con-




**(a) LDA. TP: 0.44, TN: 0.61, FP: 0.39, FN: 0.56**

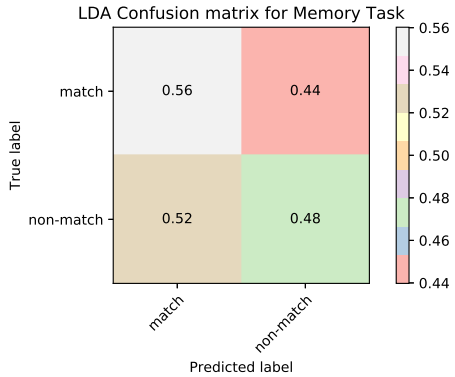



**(b) RBF. TP: 0.45, TN: 0.63, FP: 0.37, FN: 0.55**

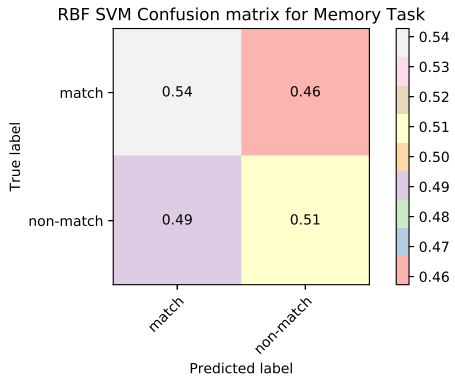



**(c) Poly. TP: 0.48, TN: 0.61, FP: 0.39, FN: 0.52**

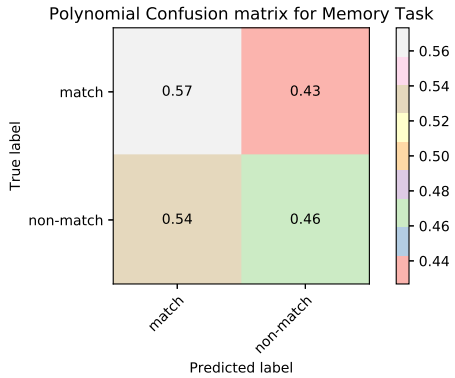
**Figure 3.1: Normalized confusion matrices of all the classifiers in the perception task.**




**(a) LDA. TP: 0.56, TN: 0.48, FP: 0.52, FN: 0.44**




**(b) RBF. TP: 0.54, TN: 0.51, FP: 0.49, FN: 0.46**




**(c) Poly. TP: 0.57, TN: 0.46, FP: 0.54, FN: 0.43**

**Figure 3.2: Normalized confusion matrices of all the classifiers in the memory task.**

ditions. The Bayes Factor ranges between 0.5 and 1.7, showing that there is no statistical difference between classification of decisions between the two conditions.

**Table 3.3: Results of Bayes Factor Analysis between the two conditions, across all classifiers.**

| Classifier | Bayes Factor |
|---|---|
| LDA | 0.50 |
| RBF | 1.71 |
| Poly | 0.50 |

## 3.3 Participant Results

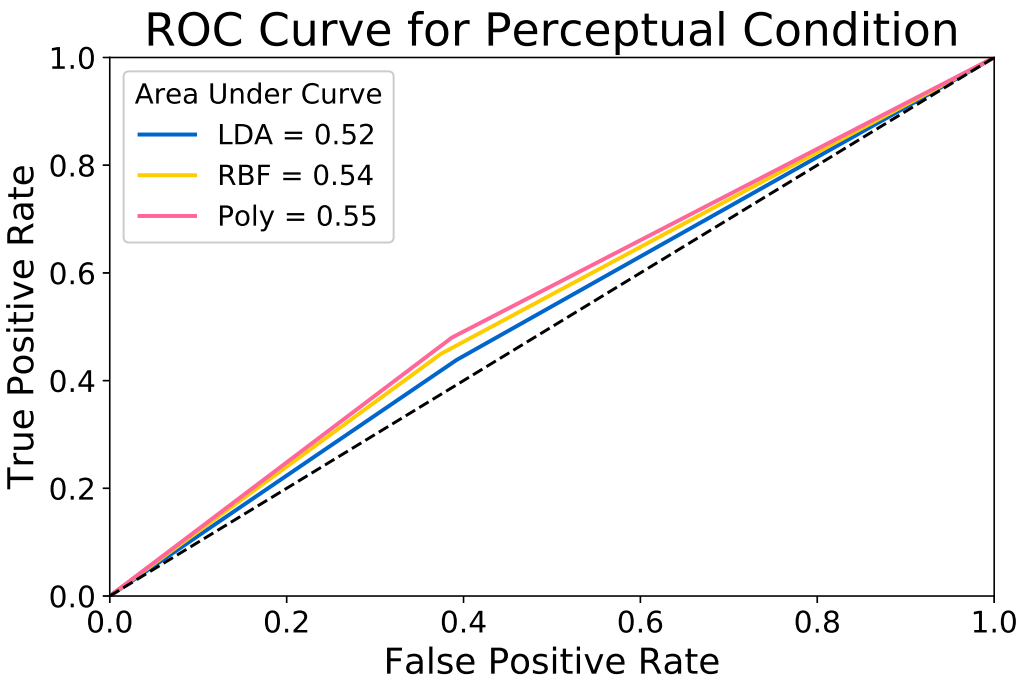
After running the classification for each task using LDA, RBF-SVM and Polynomial SVM, the mean accuracy of over the bins were plotted for each participant and each classifier (Figure 3.4).

Looking at Figure 3.4a, the classifiers seem to classify best for participant 10 and participant 14. A binomial proportion test was performed to determine whether the classification for these participants were significant. The results of this test can be seen Table 3.4 as well as the mean accuracy of the different classifiers for these participants. This shows that for participant 10, RBF ($M_{\text{RBF}_{\text{Participant10}}} = 66\%$) and Polynomial SVMs ($M_{\text{Poly}_{\text{Participant10}}} = 66\%$) classify the best and that the results of the mean accuracy are significant ($p < 0.05$) across all the classifiers. Whereas, for participant 14, the Polynomial SVM ($M_{\text{Poly}_{\text{Participant14}}} = 69\%$) classify best and the results of the mean accuracy are significant only for Polynomial SVM and RBF SVM ($p < 0.05$).

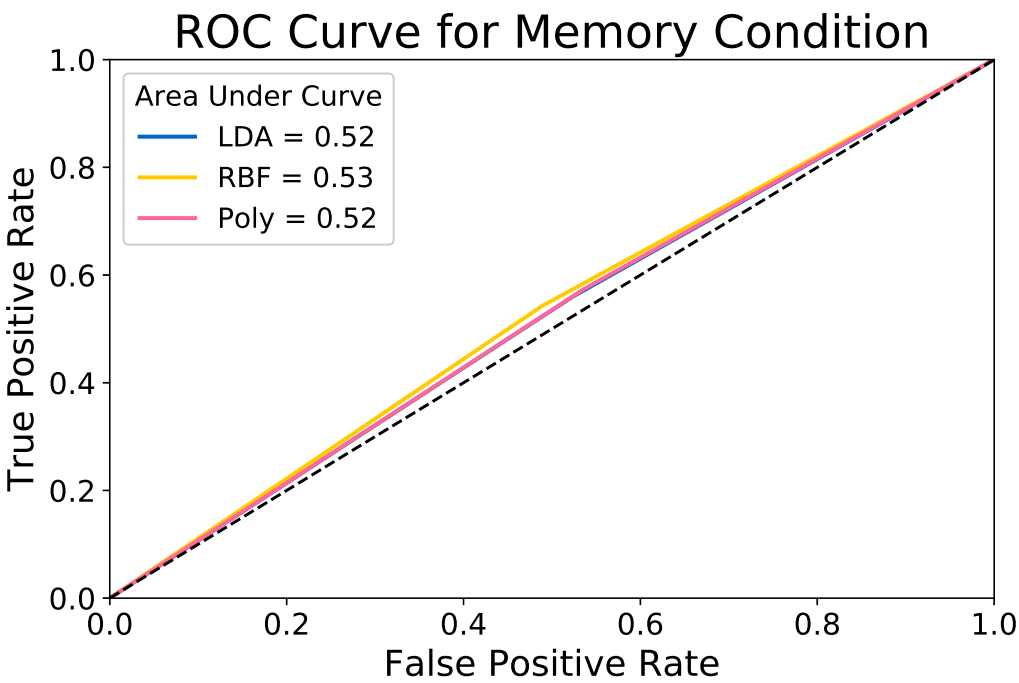
**Table 3.4: Summary of the mean accuracy and p-values for significance for participant 10 and 14 in the perceptual task (perception task).**

| | Participant 10 | | Participant 14 | |
|---|---|---|---|---|
| Classifier | Mean Accuracy | p-value | Mean Accuracy | p-value |
| LDA | 59% | 0.002 | 54% | 0.20 |
| RBF | 66% | <0.001 | 65% | <0.001 |
| Poly | 66% | <0.001 | 69% | <0.001 |

Looking at Figure 3.4b, the classifiers seem to classify best for participant 12. Again, a binomial proportion test was performed to determine whether classification for this participant was significant. The result of this test can be seen in Table 3.5 as well as the mean accuracy of the differ-

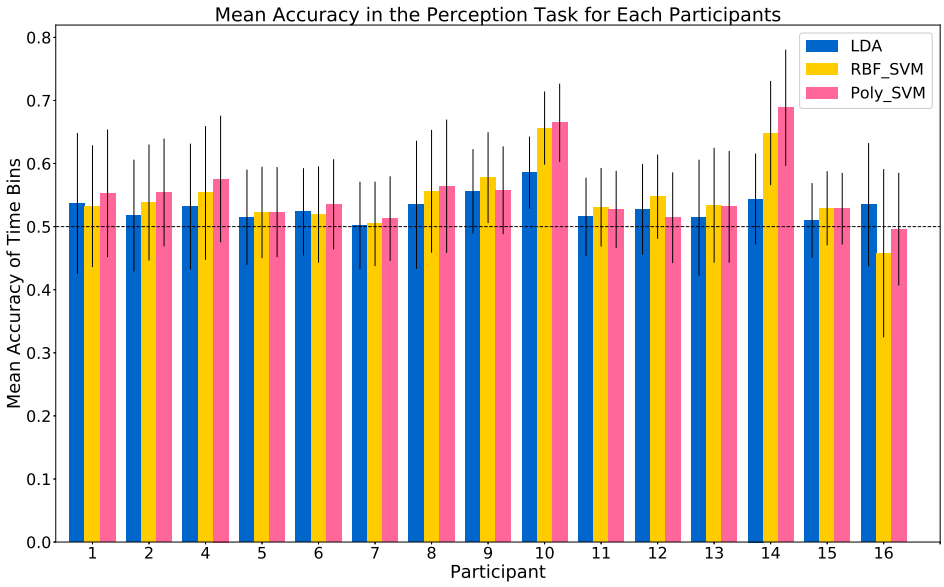



(a) Perceptual condition.

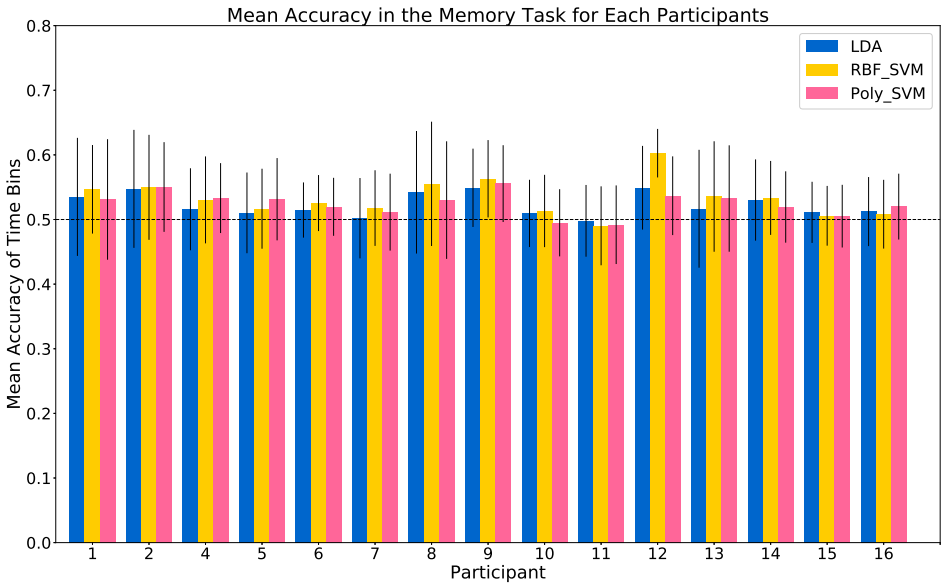



(b) Memory condition.

**Figure 3.3: ROC Curve for the LDA (blue), RBF SVM (yellow), Polynomial SVM (pink) for both the perception and memory task. The x-axis shows the false positive rate and the y-axis shows the true positive rate. The black dashed line depicts classification by chance.**




**(a) Perception task.**




**(b) Memory task.**

**Figure 3.4: Mean accuracy of each participant for both tasks. The mean accuracy over the time series can be seen on the y-axis. The participant can be seen on the x-axis. The blue bars indicate the classification done by LDA, the yellow by RBF-SVM and the pink by Polynomial SVM. There is also a dashed line at 50% accuracy to show classification done by chance.**

ent classifiers for this participant. This shows that, RBF SVM classifies best for participant 12 and this result is also significant ($p < 0.05$).

**Table 3.5: Summary of the mean accuracy and p-values for significance for participant 12 (memory task).**

| Participant 12 | | |
|---|---|---|
| **Classifier** | **Mean Accuracy** | **p-value** |
| LDA | 54% | 0.20 |
| RBF | 60% | <0.001 |
| Poly | 55% | 0.09 |

# 4 Discussion

The purpose of this experiment was to determine whether decisions can be predicted from EEG data. Based on the results, it can be seen that decisions can only be predicted above chance for some participants, not all. In particular, decisions can be predicted quite above chance for participant 10 and 14 in the perception task and participant 12 in the memory task. One possible reason why the results are not consistent across all the participants is because the electrodes were not placed in the same regions for all the participants. As mentioned earlier, the electrode placement differs since the participants suffer from epilepsy which does not occur in the same parts of the brain for all the participants.

The decisions in the perceptual task is best predicted by Polynomial SVM and the decisions in the memory task is best predicted by RBF SVM as discussed in Section 3.1.2. Simply looking at the mean accuracies of the classification of the decisions in the perception task and the memory task, it seems classifier performs better in the perceptual condition than the memory condition. But, these results are not significant as discussed in Section 3.2.

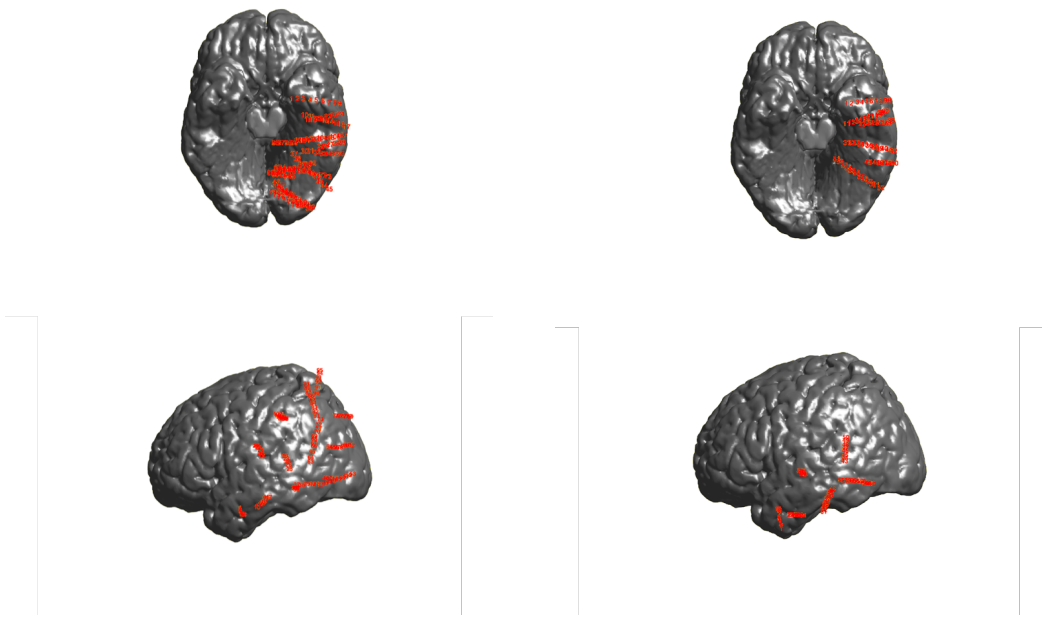
All the classifiers seem to follow the same trend across participants so it can be said that there is some agreement across the different classifiers.

Since the electrodes were not placed in the same positions for all the participants, a qualitative investigation was done to see where the electrodes were placed in the participants for whom the classification was done best. Figure 4.1a shows that for the best participants in the perception task the electrodes are placed in the right temporal lobe, left superior temporal area and also some in the parietal regions (which is known to show activity as expected in evidence accumulation). For the best participants in the memory task, the electrodes are placed in the right temporal lobe and the left inferior frontal area (Figure 4.1b). This could suggest that classification on participants with electrodes in areas known to show activity with evidence accumulation, proposed by the DDM, such as the parietal regions of the brain, tend to classify better.

Based on all these results, collecting more data will provide more insight on how these different classifiers perform as well as how decisions can be decoded from the brain. Having more data will also allow for using deep learning classifiers such as Convolutional Neural Network (CNN) which tend to have an accuracy around 81-88% for EEG data (Craik et al., 2019).

This study does not give much insight into how evidence accumulation can take place in the brain as proposed by the DDM. As mentioned earlier in Section 1, EEG oscillations occurring in the parietal 4-9Hz theta band show activity as expected with evidence accumulation. Hence, collecting sufficient data with recordings from the parietal 4-9Hz theta band and running machine learning classifiers on it will help uncover neural signals that would show activity of evidence accumulation as proposed by DDM thereby providing more information about decision making and also better classification rate. This was also suggested with the qualitative analysis done on the electrode placements earlier.

Due to the limiting amounts of data and other reasons discussed above, this study was not able to predict how decisions are made. Nevertheless, this study gave insight as to what to expect when predicting decisions and what to look for in order to be able to predict them.



**(a) Perception task.** **(b) Memory task.**

**Figure 4.1: Electrode placing for the best participant in both the task.**

# References


Craik, A., He, Y., and Contreras-Vidal, J. L. (2019). Deep learning for electroencephalogram (EEG) classification tasks: a review. *Journal of Neural Engineering*, 16(3):031001.

Crowley, J. L. (2016). Kernel functions and support vector machines. `http://www-prima.imag.fr/Prima/Homepages/jlc/Courses/2015/ENSI2.SIRR/ENSI2.SIRR.S5.pdf`.

Elhabian, S. and Farag, A. A. (2009). A tutorial on data reduction: Linear discriminant analysis (lda). `http://www.sci.utah.edu/~shireen/pdfs/tutorials/Elhabian_LDA09.pdf`.

Hakim, A., Klorfeld, S., Sela, T., Friedman, D., Shabat-Simon, M., and Levy, D. J. (2018). Pathways to consumers minds: Using machine learning and multiple EEG metrics to increase preference prediction above and beyond traditional measurements.

Ho, T. C., Brown, S., and Serences, J. T. (2009). Domain general mechanisms of perceptual decision making in human cortex. *Journal of Neuroscience*, 29(27):8675–8687.

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2017). *An introduction to statistical learning: with applications in R*. Springer.

Kołodziej, M., Majkowski, A., and Rak, R. (2012). Linear discriminant analysis as eeg features reduction technique for brain-computer interfaces. *Przeglad Elektrotechniczny*.

Lemm, S., Blankertz, B., Dickhaus, T., and Müller, K.-R. (2011). Introduction to machine learning for brain imaging. *NeuroImage*, 56(2):387–399.

Leskovec, J., Rajaraman, A., and Ullman, J. D. (2015). *Mining of massive datasets*. Cambridge University Press.

Lotte, F., Congedo, M., Lécuyer, A., Lamarche, F., and Arnaldi, B. (2007). A review of classification algorithms for EEG-based brain–computer interfaces. *Journal of Neural Engineering*, 4(2):R1–R13.

Maglogiannis, I. G. (2007). *Emerging artificial intelligence applications in computer engineering: real word AI systems with applications in eHealth, HCI, information retrieval and pervasive technologies*. IOS Press.

Mulder, M., van Maanen, L., and Forstmann, B. (2014). Perceptual decision neurosciences – a model-based review. *Neuroscience*, 277:872–884.

Paysan, P., Knothe, R., Amberg, B., Romdhani, S., and Vetter, T. (2009). A 3d face model for pose and illumination invariant face recognition. In *2009 Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance*. IEEE.

Quarizmi (2015). A full introduction to the linear fisher discriminant analysis. `https://medium.com/@QuarizmiAdTech/a-full-introduction-to-the-linear-fisher-discriminant-analysis-848530dce336`.

Ratcliff, R. and McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, 20(4):873–922.

Shadlen, M. and Newsome, W. (1996). Motion perception: Seeing and deciding. *Proceedings of the National Academy of Sciences of the United States of America*, 93(2):628–633.

van Vugt, M. K., M., Brandt, A., and Schulze-Bonhage, A. (2018). Tracking perceptual and memory decisions by decoding brain activity. *Communications in Computer and Information Science Artificial Intelligence*, page 76–85.

van Vugt, M. K., Simen, P., Nystrom, L. E., Holmes, P., and Cohen, J. D. (2012). Eeg oscillations reveal neural correlates of evidence accumulation. *Frontiers in Neuroscience*, 6.

Vasily, K. (2018). Introduction to drift diffusion model. *Coursera*. `https://www.coursera.org/lecture/neuroeconomics/introduction-to-diffusion-model-5EQD7`.

Verplancke, T., Looy, S. V., Benoit, D., Vansteelandt, S., Depuydt, P., Turck, F. D., and Decruyenaere, J. (2008). Support vector machine versus logistic regression modeling for prediction of hospital mortality in critically ill patients with haematological malignancies. *BMC Medical Informatics and Decision Making*, 8(1).

Wang, Y., Wang, Y., Patel, S., and Patel, D. (2006). A layered reference model of the brain (LRMB). *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)*, 36(2):124–133.

# A Appendix

The data set used and the code for this study can be found on Github: `https://github.com/sukhleen-kaur/bachelor_project`