Nigerian Journal of Technology (NIJOTECH)
Vol. 40, No. 6 November, 2021, pp1079 –1085
www.nijotech.com

Print ISSN: 0331-8443
Electronic ISSN: 2467-8821
http://dx.doi.org/10.4314/njt.v40i6.10

In Detecting Deception using Decision Tree and SVM across Different Cues

O. Owolafe[1, *], O. A. Ayeni[2]

[1,2]Cyber Security Department, Federal University of Technology, Akure, Ondo State, NIGERIA

Abstract

Past and present researches in deception detection make use of questioning to ascertain the truth and deceit in the response of the interviewee. In such questioning method, either the verbal or nonverbal cues are closely monitored and analysed to arrive at a decision. Since no single verbal or nonverbal cue is able to reliably detect deception the research proposes to use both the verbal and nonverbal cues to detect deception. Therefore, this research aims to develop a Support Vector Machine and Decision Tree Model to classify the extracted verbal, nonverbal and VerbNon features as deceptive or truthful. The verbal cues capture the speech of the suspect while the nonverbal cues capture the facial expressions of the suspect. The Praat (a tool for speech analysis) was used in extracting all the verbal cues while the nonverbal features were extracted using the Active Shape Model (ASM). The work was implemented in 2015a MatLab. The analysis of the result shows that Decision Tree performs better than SVM in the classification with a percentage score of 93.5% for Nonverbal cues as against that of SVM having percentage score of 91.9%. For verbal and VerbNon cues, Decision Tree recorded 89.9% and 97.6% accuracy while SVM recorded 89.2% and 97.1% accuracy.

Keywords: Deception, SVM, Decision Tree, Verbal Cues, Facial Expression

1.0 INTRODUCTION

Jaume *et al.* (2004) defined deception elaborately as the deliberate attempt, whether successful or not, to conceal, fabricate, and/or manipulate in any other way, factual and/or emotional information, by verbal and/or nonverbal means, in order to create or maintain in another or others a belief that the communicator himself or herself considers false. While Baron-Cohen *et al.* (2001) simply stated that deception is the act of deceiving and Vrij (2008) a notable scholar in the field of deception defined the concept as "a successful or unsuccessful deliberate attempt, without forewarning, to create in another a belief which the communicator considers to be untrue".

A study found that lying takes longer than telling the truth, and thus the time to answer a question may be used as a method of lie detection. However, it has also been shown that instant-answers can be proof of a prepared lie. The only compromise is to try to surprise the victim and find a midway answer, not too quick, nor too long (Newman *et al.*, 2003).

Repeated studies have shown that traditional methods of detecting deception during interviews succeed only 50% of the time, even for experienced law enforcement officers. In spite of this, investigators still need the ability to test the veracity of those they interview. To do so, investigators require a model that incorporates research with empirical experience to differentiate honesty from deception. They can use an alternative paradigm for detecting deception based on four critical domains: comfort/discomfort, emphasis, synchrony, and perception management rather than merely trying to detect traditional signs of deception, which, in some cases, may be misleading.

In real life problems are solved by thinking about them, therefore, dealing with the emulation of human thought by a computer program becomes paramount. Since humans do not think about problems as conventional computers do, dealing constantly with uncertainties, ambiguities, and contradictions arises. Sometimes deductive logic is used, but more often we think intuitively, assembling information relevant to a problem, scanning it and coming to a conclusion. Besides this, humans often learn from experience but in many ways computers could be better at detecting deceptions than people because of their tremendous logical analysis capability and the fact that the logical processes used by

***Corresponding author (**Tel:** +234 (0)7033513174**)**
Email addresses: oiyare@futa.edu.ng (Owolafe Otasowie) and oaayeni@futa.edu.ng (Ayeni Olaniyi A.)

computers are quite different from the processes used by people.

The question that remains a subject of controversy and which this research will tend to address is whether deception can reliably be detected through verbal or nonverbal means or a combination of both.

2.0 LITERATURE REVIEW

So many researches in various fields have been carried concerning deception detection. In psychology, Hicks and Ulvestad (2011) carried out a study that detects deception accuracy using verbal or nonverbal cues. Their result recorded no significant effect of cue type (verbal or nonverbal) or statement type (truthful or deceptive) and participants' accuracy was no better that chance.

Mulay *et al*. (2010) applied support vector machine and decision tree for Intrusion Detection System. This research proposes the decision tree based algorithm to construct multiclass intrusion detection system.

Perez-Rosas *et al.* (2015) developed a new deception dataset consisting of videos from real life scenarios, and build deception tools relying on verbal and nonverbal features. The result of their classification ranges from 77-82% when using a model that extracts and fuses features from the linguistic and visual modalities.

A decision tree is a machine-learning algorithm that uses the hierarchical structure consisting of nodes and directed edges (Gehrke *et al*., 2000). The nodes can be root node (which is the starting nodes), internal node and leaf or terminal nodes (decision node). The leaf node is assigned a class label while the root and internal nodes contain attribute test conditions to separate records that have different characteristic. An example of a decision tree used for mammal classification is shown in Figure 1.

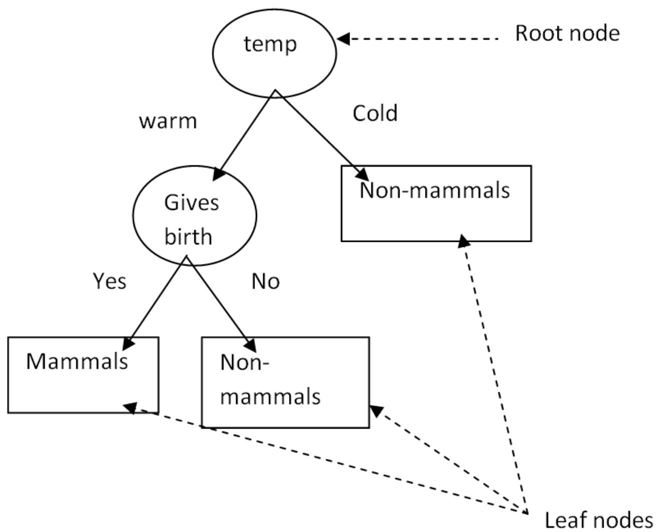

Figure 1: A decision tree for mammal classification problem

SVM is a classification technique that gives the best split of the datasets under consideration. It optimally classifies linearly separable binary sets into two classes using hyperplane. Different hyperlapnes can be drawn but the one that leaves the maximum margin from both classes will be the one selected. The hyperplane is a line dividing a plane in two parts where each class lies in either side of the hyperplane. Given a plot of two label classes as shown in Figure 2, the hyperplane separates the plots into their respective classes.

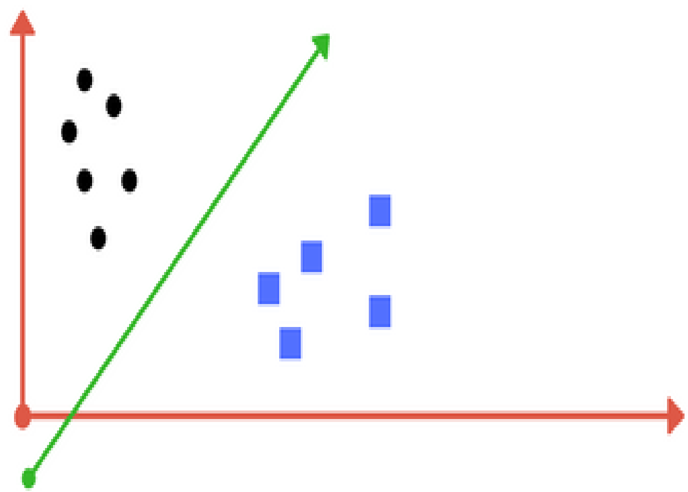

Figure 2: Hyperlpane showing two linearly separable classes

3.0 MODEL DESIGN

The system extracted desired features from the dataset of Perez-Rosas *et. al.* (2016). The dataset consists of real-life trial videos of statements made by exonerees after exoneration and a few statements from defendants during crime-related TV episodes. The speakers in the videos are either defendants or witnesses. The video clips are labelled as deceptive or truthful based on a guilty verdict, not-guilty verdict, and exoneration. The dataset consists of 121 videos including 61 deceptive and 60 truthful trial clips. The average length of the videos in the dataset is 28.0 seconds. The average video length is 27.7 seconds and 28.3 seconds for the deceptive and truthful clips, respectively. The system was designed using decision tree and support vector machine technique. In the research, three deceptive cues were used for detecting deception. They are Verbal, Non-verbal and VerbNon cues. The verbal cues capture the speech of the suspect while the nonverbal cues capture the facial expressions of the suspect. The VerbNon cue is a combination of Verbal and Nonverbal cues. The verbal cues include the voice pitch (in terms of variations), frequency perturbation also known as jitters, pauses (voice or silent), and speechrate (is defined as the rate at which the suspect is speaking).

The PRAAT (a tool for speech analysis) was used in extracting all the verbal cues. After the feature extraction stage, the classification model was designed as shown in section (a) and (b).

a) *Decision Tree*

In Decision trees questions are asked and classification is done based on the answer given, the answer is usually yes or no, true or false. The classification can be categorical or numeric.

In Decision trees questions are asked and classification is done based on the answer given, the answer is usually yes or no, true or false. The classification can be categorical or numeric.

Model formulation:

Given set of training samples $x_i \in R^S$

And label sample $y \in R^i$

A decision tree partitions the space such that the samples with the same labels are grouped together.

Let the data at node m be rep. by Q. For each candidate split $\theta = (j, t_m)$ consisting of a feature j and threshold t_m, the data is partitioned $Q_1(\theta)$ and $Q_2(\theta)$ subset. Where $Q_1(\theta) = (x, y)|\ x_j \leq t_m$ and $Q_2(\theta) = Q/Q_1(\theta)$

The impurity at m is computed using an impurity function in equation (1).

$$G = 1 - \sum_i P^2(i) \qquad (1)$$

If the value of G is 0.5 that means there is no difference between the classes. The best classifier is when the value of G is close to 0.

The process continues for subsets $Q_1(\theta)$ and $Q_2(\theta)$ until the maximum allowable depth is reached. Figure 3 shows some sample decision rule that was generated.

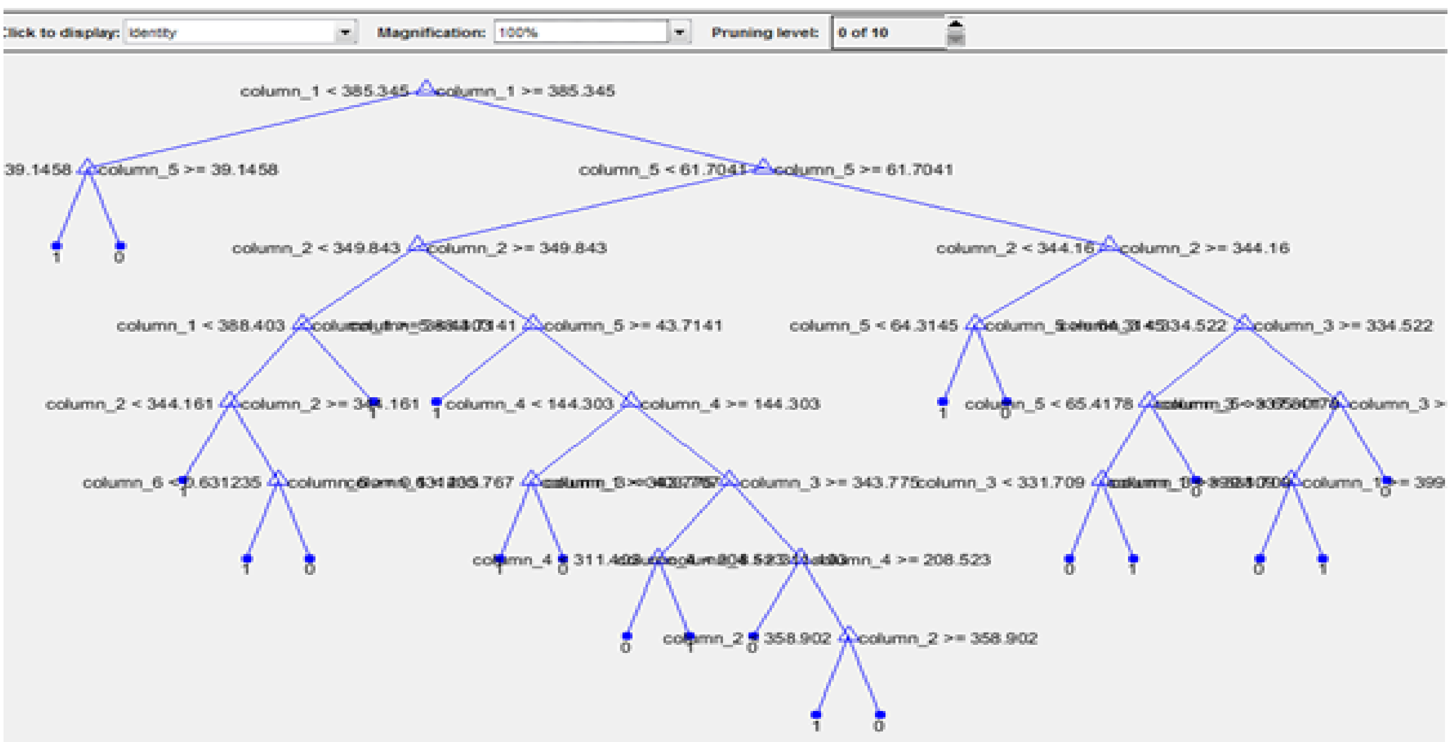

Figure 3: Sample Decision Rule

b) *Support Vector Machines*

It is a classification technique that gives the best split of the datasets under consideration. It optimally classifies linearly separable binary sets into two classes using hyperplane. Two hyperplanes are drawn but the one that leaves the maximum margin from both classes will be the one selected. The margin is the distance between the closest classes from the hyperplane. The separating hyperplane is calculated using equation 2.

$$g(x) = \omega x + \omega_0 \quad \ni \quad g(x)\begin{cases} \geq 1 \ \forall\ x \in class1 \\ \leq -1 \ \forall\ x \in class2 \end{cases} \qquad (2)$$

ω is the vector of weights and ω_0 is the bias.

The distance between the closest element of class 1 and 2 to the hyperplane is calculated using equation 3.

$$d = \frac{1}{||\ \omega||} + \frac{1}{||\ \omega||} = \frac{2}{||\ \omega||} \quad (3)$$

The value of ω is minimized using the langrage multiplier:

$$\omega = \sum_{i=0}^{N} \lambda_i y_i x_i \quad (4)$$

4.0 RESULT TESTING AND EVALUATION

The implementation of the system was carried out in MatLab environment, 933 of the verbal dataset, 2998 of the nonverbal and 1000 of the verbnon dataset was used for training the model. After training, the model was tested using dataset with known classification. Details of the analysis are shown in Table 1 and the graph is shown in Figure 1. Nonverbal dataset have reduced training error as well as reduce testing error.

Table 1: Training versus Testing Error across datasets

	Total dataset	Training dataset	Testing dataset	Training Error	Testing Error
Verbal	1693	933	760	0.3354	0.8745
Nonverbal	5133	2998	2135	0.26156	0.5432
VerbNon	1353	1000	353	0.15214	1.8580

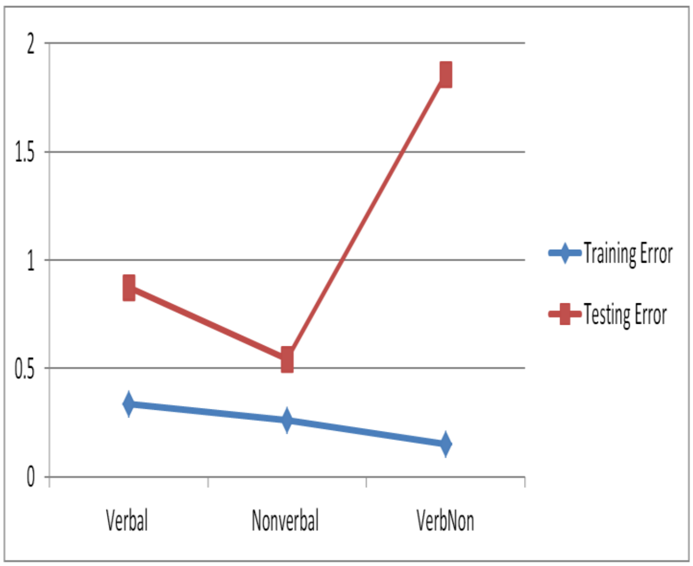

Figure 4: Training versus Testing Error across datasets.

4.1 Performance Metrices

The metrics used for carrying out the performance evaluation are listed as:

1. False Positive Rate (FPR):

$$F_{N,R} = \frac{F_P}{T_P + F_P} \quad (5)$$

2. True Positive Rate (TPR):

$$T_{P,R} = \frac{T_P}{T_P + F_N} \quad (6)$$

3. Accuracy:

The overall accuracy is given by the sum of true and false utterances correctly classified, out of all the classifications carried out. It is the number of correct predictions over the total number of predictions.

$$Accuracy = \sum \left(\frac{T_P + F_P}{T_P + T_N + F_P + F_N} \right) \quad (7)$$

Where $T_P, T_N, F_P\ and\ F_N$ are the True positive, True negative, False positive and False negative values respectively.

4. Confusion Matrix:

It is a table used to describe the performance of the classification model on the dataset.

Table 2 shows the extracted confusion matrix for each of the datasets.

Table 2: Confusion Matrix for Verbal, Nonverbal and VerbNon dataset

	Training	Validation	Test	All
Nonverbal (N)	97.1%	97.2%	97.2%	97.1%
Verbal (V)	84.4%	86.6%	81.9%	84.3%
VerbNon	92.7%	92.8%	91.6%	92.5%

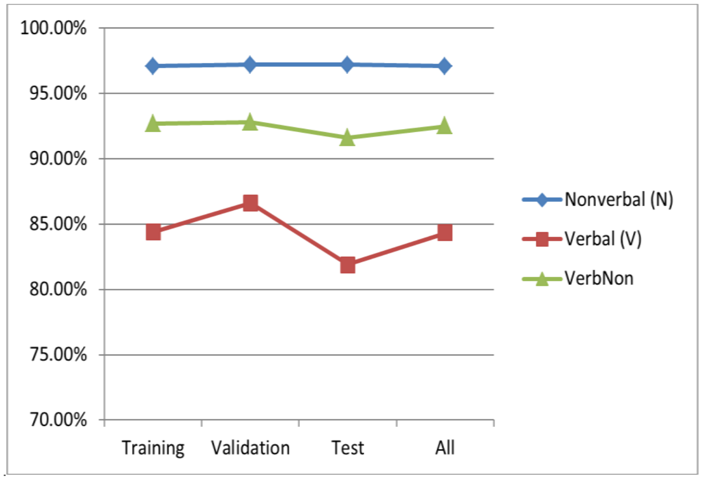

Figure 5: Confusion matrix for verbal, nonverbal and VerbNon dataset

4.2 *Performance of Different Classifiers on the Datasets*

The different datasets were passed through different classifiers to ascertain the performance of the classifiers on the dataset.

Table 3: Comparative Analysis of different classifiers on each dataset

Verbal Cues

	SVM	Decision Tree
Overall Accuracy	89.2%	89.9%
Overall Error	10.8%	10.1%
Total Dataset used	1693	1693

Nonverbal Cues

	SVM	Decision Tree
Overall Accuracy	91.9%	93.5%
Overall Error	8.1%	6.5%
Total Dataset used	5133	5133

VerbNon Cues

	SVM	Decision Tree
Overall Accuracy	97.1%	97.6%
Overall Error	2.9%	2.4%
Total Dataset used	1353	1353

Table 3 shows the performance of various classifiers on each of the datasets while Figures 6, 7 and 8 gives graphical representations of the performance. From the table, it is observed that Decision Tree performs better than SVM using the different datasets.

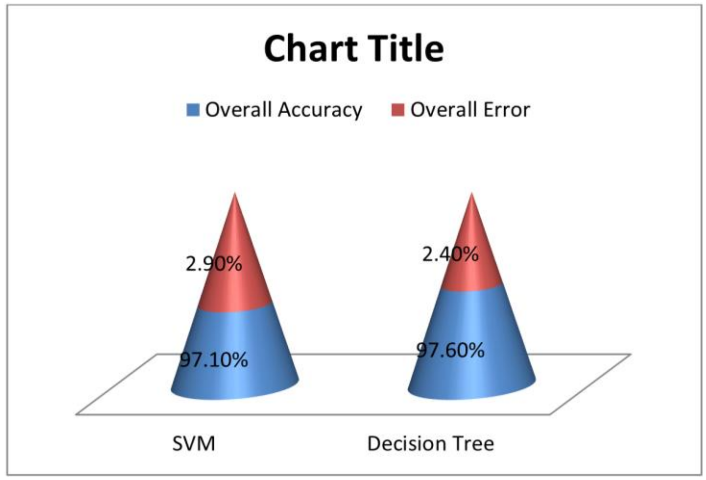

Figure 6: Performance of classifiers on VerbNon dataset

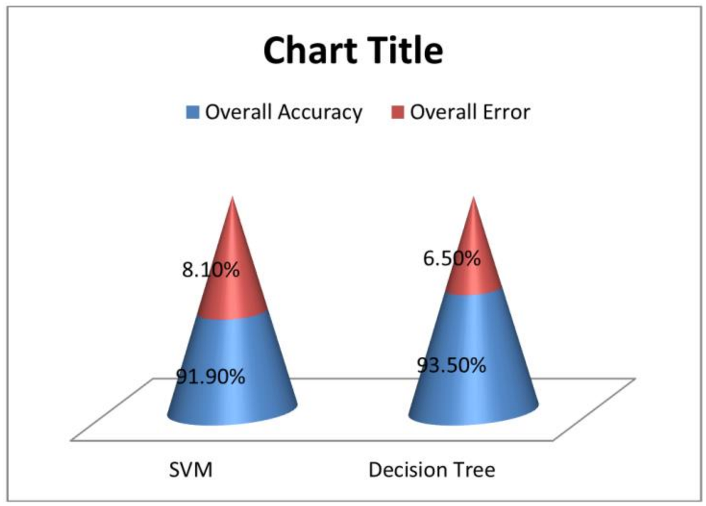

Figure 7: Performance of classifiers on Nonverbal dataset

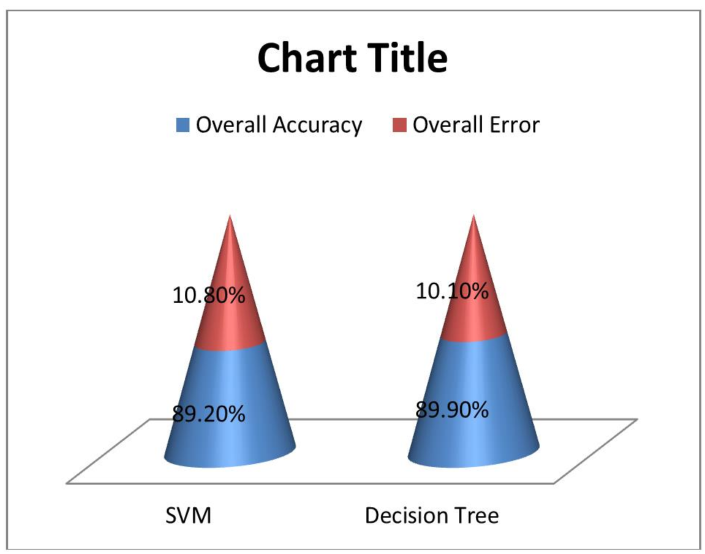

Figure 8: Performance of classifiers on Verbal dataset

5.0 CONCLUSION

Deception detection is an involved social issue because to successfully deceive the deceiver has to formulate a story that is internally consistent while hiding emotions and true intentions. Facial expressions and voice play a critical role in the identification of deception as shown in this research. Previous research made use of only one cue but this research made use of both verbal and nonverbal cues. The developed system was able to perform better than chance and trained professionals with a result difference of 47.6% when a combination of verbal and nonverbal (verbnon) dataset was used.

This work uses verbal, nonverbal cues and a combination of both cues to detect deception. The verbal cues was extracted using Praat while the nonverbal was extracted using Active Shape Model. The classification was done using SVM and Decision Tree and the performance was compared. Decision Tree recorded the best performance with the different datasets.

The system was implemented using Matlab 2015a on window 7 with 2GB RAM. The extracted data was divided into training data and test data. The SVM and Decision Tree model was trained using the training data while the functionality of the model was ascertained using the test data. At the end of the comparative analysis it was discovered that Decision Tree performed better on all the dataset to detect deception. The result obtained using only verbal cue was 89.9% while that of nonverbal cue was 93.5% but on VerbNon yielded 97.6% which is far better than the chance level of 50%.

REFERENCES

[1] Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y. and Plumb, I. "The Reading the Mind in the Eyes" Test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *42*(2), 2001, pp.241-251.

[2] Gehrke J., Ramakrishnan R. And Ganti V. "Rainforest-A framework for Fast Decision Tree Construction of Large Datasets", *Data Mining and Knowledge Discovery*, 4(2/3), 2000, pp.127-162.

[3] Gerschlager C. "Deception in Markets: An Economic Analysis" *Basingstoke: Palgrave Macmillan*, 2005.

[4] Hicks, C. and Ulvestad, N., Deception Detection Accuracy Using Verbal or Nonverbal Cues. *The Journal of Undergraduate Research*, *9*(1), 2011. pp.9.

[5] Jaume M, Eugenio G, and Carmen H. "Defining Deception". *Journal of anales de psicologia*, 20(1), 2004, pp.147-171.

[6] Mulay, S.A., Devale, P.R. and Garje, G.V. "Intrusion detection system using support vector machine and decision tree". *International journal of computer applications*, *3*(3), 2010, pp.40-43.

[7] Newman, M. L., Pennebaker, J. W., Berry, D. S. and J. M. Richards "Lying words: predicting deception from linguistic styles" *Personality and Social Psychology Bulletin*, 29, 2003, pp.665-675.

[8] Pérez-Rosas, V., Abouelenien, M., Mihalcea, R., Xiao, Y., Linton, C.J. and Burzo, M., "Verbal and nonverbal clues for real-life deception detection". *Proceedings of the Conference on Empirical Methods in Natural Language Processing,* September, 2015, pp.2336–2346, Lisbon, Portugal, 17-21

[9] Vrij, A. "Detecting lies and deceit: Pitfalls and opportunities". Chichester, United Kingdom: Wiley, 2008.

APPENDIX

Table 4.3: Nonverbal dataset

Mouth	Nose	Eyelid	Eyebrown
372.4577	350.1620	333.8423	320.9387
372.4825	350.3103	334.1362	321.3512
372.4825	350.3103	334.1362	321.3512
372.4825	350.3103	334.1362	321.3512
372.4825	350.3103	334.1362	321.3512
372.4825	350.3103	334.1362	321.3512
349.4173	332.3308	318.3695	306.8928
349.6762	332.4309	318.2936	306.7356
361.1741	338.1432	330.0042	318.1857
361.1065	338.1345	329.9477	318.1282

361.1065	338.1345	329.9477	318.1282
378.4289	356.5483	339.8712	327.2850
378.368	356.3525	339.7399	327.1270
378.368	356.3525	339.7399	327.1270
391.6167	370.0022	347.4528	334.7716
392.0534	370.6655	347.9038	335.5397
392.0534	370.6655	347.9038	335.5397
404.0472	380.7212	353.2601	340.1458
404.0704	380.7212	353.2601	339.8325
404.0704	380.7212	353.2601	339.8325
399.4222	376.3006	349.6747	335.8231
399.4018	376.4027	349.6715	335.5445
399.4018	376.4027	349.6715	335.5445
387.2334	363.7798	341.9400	328.7511
387.0478	363.7800	342.3307	328.8101
387.0478	363.7800	342.3307	328.8101
369.3888	347.9349	333.8453	322.1860