www.revistageintec.net
ISSN: 2237-0722
revista GEINTEC Gestão, Inovação e Tecnologias

An Enhanced Novel GA-based Malware Detection in End Systems Using Structured and Unstructured Data by Comparing Support Vector Machine and Neural Network

T. Sai Tejeshwar Reddy[1]; A. Sivanesh Kumar[2*]

[1]Research Scholar, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamil Nadu, India.

[1]tsai6476@gmail.com

[2*]Assistant Professor, Project Guide, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamil Nadu, India.

[2*]kas.sivanesh@gmail.com

Abstract

Aim: *The aim of the work is to perform android malware detection using Structured and Unstructured data by comparing Neural Network algorithms and SVM.*

Materials and Methods: *consider two groups such as Support Vector Machine and Neural Network. For each algorithm take N=10 samples from the dataset collected and perform two iterations on each algorithm to identify the Malware Detection.* ***Result:*** *The accuracy results of the Neural Network model has potential up to (82.91%) and the Support Vector Machine algorithm has an accuracy of (79.67%) for Android malware detection with the significance value of (p=0.007).* ***Conclusion:*** *classification of android malware detection using Neural Network algorithm shows better accuracy than SVM.*

Key-words: Malware Detection, Neural Network, Support Vector Machine, Machine Learning, SPSS statistical tool, Machine Learning.

1. Introduction

Main research of this work is to detect android malware phishing attacks using machine learning techniques (Wen and Yu 2017) applications of this work is End -to -End secure data transmission. Importance of the research in this field is to increase the use of smartphones (Yerima, Sezer, and Muttik 2015) applications of this work is prevention of data loss and it helps in

improvisation of security feature. The main applications of this paper is preservation of privacy (Liu et al. 2020), Data security, mode of transmission (Liu et al. 2020), reliable and secure (Jung et al. 2018).

There are nearly 21 articles published in google scholar and 14 articles published in IEEE Xplore related to Android malware detection. Android applications are created quickly over the portable biological system, but Android malware is additionally rising in a perpetual stream (Yuan, Lu, and Xue 2016). With thousands of Android apps, Here completely test Droid Detector and perform an in-depth examination on the highlights that profound learning basically abuses to characterize malware Pernicious applications pose a risk to the security of the Android stage (Westyarian et al. 2015). The creating entirety and contrasts of these applications render conventional watches for the most part incapable and in this way Android smartphones habitually remain unprotected from novel malware(Tahtaci and Canbay 2020). These highlights are embedded in a joint vector space, such that ordinary plans characteristic for malware can be subsequently recognized and utilized for clarifying the choices of our procedure ("Android Malware Detection Using Machine Learning" 2020). Android malware seriously debilitates framework and client security in terms of benefit acceleration, inaccessible control, duty burglary, and security spillage. Subsequently, it is of awesome significance and needs to distinguish Android malware (Ma et al. 2019). The work in paper is display a combination strategy for Android malware discovery based on the machine learning algorithm (Ma et al. 2019).Among the all the research papers most cited article is android malware detection using deep learning. In study of all research papers the best paper is An Android malware detection system based on machine learning (Wen and Yu 2017)

Previously our team has a rich experience in working on various research projects across multiple disciplines (Sathish and Karthick 2020; Varghese, Ramesh, and Veeraiyan 2019; S. R. Samuel, Acharya, and Rao 2020; Venu, Raju, and Subramani 2019; M. S. Samuel et al. 2019; Venu, Subramani, and Raju 2019; Mehta et al. 2019; Sharma et al. 2019; Malli Sureshbabu et al. 2019; Krishnaswamy et al. 2020; Muthukrishnan et al. 2020; Gheena and Ezhilarasan 2019; Vignesh et al. 2019; Ke et al. 2019; Vijayakumar Jain et al. 2019; Jose, Ajitha, and Subbaiyan 2020). Now the growing trend in this area motivated us to pursue this project.

The methods and techniques involved in this study have invoked Support vector machine algorithm. The performance of the SVM method lags while it is implemented on the unstructured data set. It also takes enormous time to train SVM models with the size of datasets to identify and detect the android malware detection provided by Neural Network (Lekssays, Falah, and Abufardeh

2020). The proposed Neural Network has improved the accuracy of Android malware detection better than SVM.

2. Materials and Methods

The paper study was done at Saveetha School of Engineering, SIMATS. The number of groups used for the study is 2. Group 1 is Neural Network and group 2 is Support Vector Machine. Sample size (N=10) for each group and 10 iterations are made for two groups to calculate the accuracy of the model.

Dataset Description

The dataset which is used is "Android Malware dataset". The dataset was collected from Github (https://github.com/cskamil/Android-Application-Dataset-for-Malware-Detection). This is a dataset used to predict the accuracy of android malware detection. The main attributes that have been used to predict the Malware detection accuracy (%)."Flow ID" refers to Transaction ID , "Source IP" refers to Sender Address, "Source Port" refers to Sender's End Socket port number, "Destination IP" refers to Receiver Address, "Destination Port" refers to Receiver's End Socket port number, "protocol" refers to the standard that it follows, "Flow Duration" refers to time taken to transfer from source to destination, "Total Fwd Packet " refers to total number of packets transferred from Source to destination," Total Backward Packets" refers to total number of packets transferred from destination to Source, " Label" refers to virus name, This is the overall description of the Android Malware dataset.

Neural Network

NN is rarely used for prediction modelling. The reason is that Neural Networks usually try to over-fit the relationship provided. NN is by and large utilized in cases where what has happened within the past is repeated nearly precisely within the same way. Neural Systems have numerous diverse coefficients, which it can optimize. It can handle more variability than existing models.

Pseudocode for Neural Network

Input: Trained dataset

Output: Classifier trained accuracy

1. classifier = Sequential()
2. classifier.add(Dense(output_dim = 6, activation = 'relu', input_dim = 10))
3. classifier.add(Dense(output_dim = 6,activation = 'relu'))
4. classifier.add(Dense(output_dim = 10 ,activation = 'softmax'))
5. classifier.compileoptimizer='adam',loss='sparse_categorical_crossentropy', metrics
6. classifier.fit(X_train, y_train, batch_size = 10, nb_epoch = 500)
7. y_pred = classifier.predict(X_test)

Support Vector Machine

SVM is one of the directed machine learning calculations which is utilized for both classification and relapse. The svm classifier works based on isolating the two classes. The reason for the SVM calculation is to form the decision boundary that can separate n-dimensional space into classes, and it makes it simple to put modern data focuses in that class in future.

Pseudocode for Support Vector Machine

Input: Trained dataset

Output: Classifier trained accuracy

1. svc.fit(X_train,y_train)
2. y_pred = svc.predict(X_test)
3. cm = confusion_matrix(y_test, y_pred)
4. accuracy = accuracy_score(y_test, y_pred)
5. print("Accuracy : " , accuracy)

Experiment Setup

The platform used to evaluate the machine learning algorithms was jupyter lab. The hardware configurations were intel core i5 processor with a RAM size of 4GB was used. The system type used

was a 64-bit, OS, X64 based processor with an HDD of 917 GB. The operating system used was Windows and the tool used was jupyter lab with python programming language.

The independent variables of the dataset are “Timestamp”,”Fwd packet length max”,”Fwd packet length mean”,”Fwd packet length std”. And Dependent variables are ”Flow ID”, ”Source IP”,”Source Port”, “Destination IP” “Destination Port”,”protocol”,“Flow Duration”,”Total Fwd Packet ”,” Total Backward Packets”,” Label”.

The dataset is Android malware. Data preprocessing has to be done. Data cleaning like removing the unnecessary attributes from the dataset and concatenating and shuffling also need to be done. The dataset collected was cleaned by removing null values and duplicate values and data pre-processing was done. After data pre-processing the dataset was splited into two parts as testing set and training set. 30% of the dataset was taken as a testing set and 70% of the dataset was taken as a training set. After splitting the dataset the algorithm was fitted. By evaluating the algorithm with train and test sets the required parameter accuracy percentage was predicted.

3. Results

The analysis was done using IBM SPSS tool. SPSS tool is a statistical software used for data analysis. For both proposed and existing algorithms 10 iteration was done and for each iteration the predicted accuracy was noted for analysing in the SPSS tool. With the data obtained from the iterations an independent sample T-test was performed.

The Table-1 is Accuracy Table (NN, SVM), the accuracy of the Neural Network is approximately (82.91%) and Support vector machine is approximately (79.67%). The accuracy varies for different test sizes in decimals. The accuracy varies due to random change in the test size of the algorithm.

Table 1- Accuracy of the NN Algorithm is Approximately (82.91%), and for SVM are Approximately (79.67%). The Accuracy varies for different Sample Sizes

Sample	Neural Network Accuracy (%)	SVM Algorithm Accuracy (%)
1	68.25	59.68
2	69.98	60.56
3	70.92	61.77
4	73.98	61.35
5	77.76	62.58
6	78.13	68.29
7	79.03	71.48
8	81.69	74.56
9	82.45	78.82
10	82.91	79.67

Table-2 is Independent Samples Test, the comparison of accuracy for Android malware detection classification using Neural Network algorithm and Support vector machine with significance less than $p<0.05$ and standard error difference 0.9111. When compared with the other algorithms, the performance of the proposed Neural Network model achieved better performance than the Support vector machine classifier.

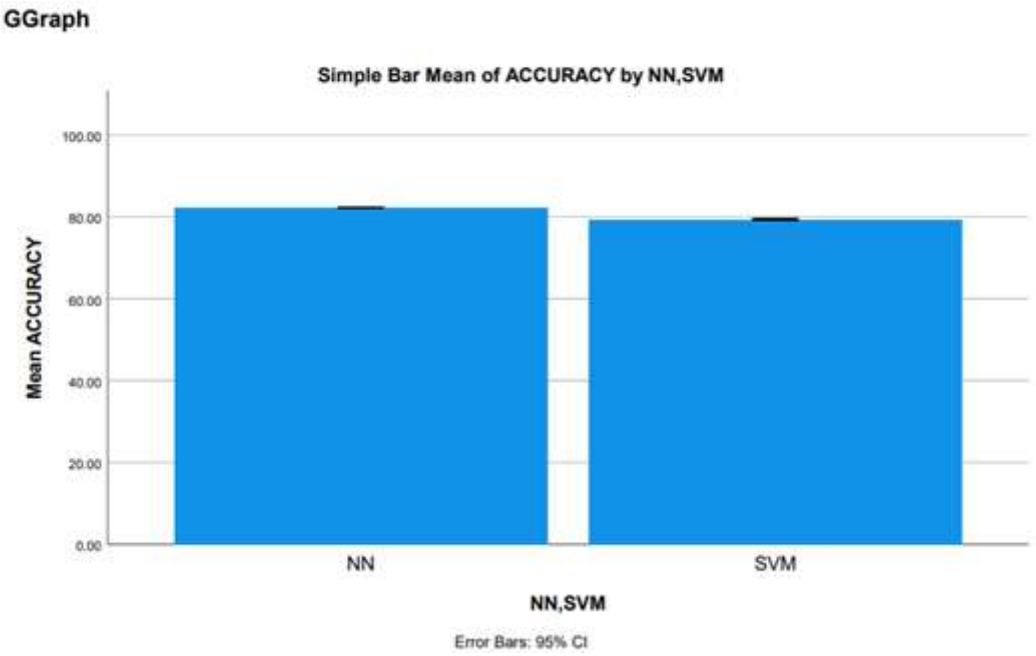
Table-3 is Group Statistics, the mean accuracy and standard deviation for Neural Network is 82.22701 and 0.06783. The Support vector machine algorithm is 79.37801 and 0.28003. In performing statistical analysis of 10 samples, NN obtained 0.06783 standard deviation with 0.02145 standard error while SVM obtained 79.37801 standard deviation with 0.08855 standard error (Table 2). The significance value smaller than 0.05 showed that our hypothesis holds good.

Table 2- Independent Samples T-Test Depicts that NN Proved with Mean difference=2.84900, Std Error Difference=0.09111 than SVM and NN got Better Significance than SVm with Value of (p=0.007)

		Levene's Test for Equality of Variances		**t-test for Equality of Means**						
		F	**Sig.**	**t**	**df**	**Sig. (2-tailed)**	**Mean Difference**	**Std. Error Difference**	**95% Confidence Interval of the Difference**	
									Lower	**Upper**
Accuracy	Equal variances assumed	9.318	.007	31.268	18	.000	2.84900	0.9111	2.65758	3.04042
	Equal variances not assumed			31.268	10.053	.000	2.84900	0.9111	2.64163	3.05187

The fig 1 is Bar Graph, Simple Bar Mean of Accuracy by NN, SVM the bar chart representing the comparison of mean accuracy of Neural Network algorithm is 82.91%. Independent t-test was used to compare the accuracy of two algorithms and a statistically significant difference was noticed $P < 0.05$. The Neural Network model obtained 82.91% accuracy. When compared with the other algorithms, the performance of the proposed Neural network model achieved better performance than the Support Vector machine classifier.

Fig. 1- Bar Graph Represents Comparison of Mean Accuracy of Models Neural Network (82.91%) and SVM (79.67%), the Indications of Graph are (x-axis NN, SVM), (y-axis, Mean Accuracy) i.e.; with Error bars:95% CI.



Neural Network and SVM algorithms were run different times in Jupyter with a sample size of 10 and accuracy was calculated (Table 1). It is observed that the mean accuracy of NN was 82.91% and SVM was 79.67 %. The statistical analysis was done using IBM SPSS version 21. It is a statistical software tool used for data analysis. For both SVM and Neural Network algorithms different iterations were done with a sample size of 10 and for each iteration the predicted accuracy was noted for analysing the accuracy. With value obtained from the iterations Independent Sample T-test (Table 2) and Group Statistics was performed (Table 3). It can be concluded that the Neural Network appears to produce a significant difference than SVM with the value of p=0.007 and with 95% confidence interval (Table 2). The Mean of Neural Network appears to be better when Compared with SVM with a standard deviation of 0.06783 and 0.28003 respectively (Table 3). The Graphs (fig 1) shows the comparison graph of Accuracy percentage of NN (82.91) and SVM (79.67) and Error Difference of NN and SVM was 0.9111 and 0.9111 respectively.

Table 3- Group Statistics Results Depicts NN has an Mean Accuracy (82.22701%), Std. Deviation (0.06783), Whereas for SVM has Mean Accuracy (79.37801%), Std. Deviation-(0.28003) for Sample Size (N=10) for the Accuracy of the Groups

	NN, SVM	N	Mean	Std. Deviation	Std. Error Mean
Accuracy	Neural Network	10	82.22701	0.06783	0.02145
	Support Vector Machine	10	79.37801	0.28003	0.08855

Received: 28.02.2021 – Accepted: 06.04.2021

4. Discussion

In this study of Android Malware Detection, the Neural Network model has higher accuracy approximately (83%) in comparison to Support vector machine algorithm approximately (80%). The Neural Network algorithm has better significance ($p<0.05$) which is achieved while using the SPSS tool for statistical calculations.

The fig 1 can observe that the bar chart represents the comparison of mean accuracy of Neural Network algorithm and Support vector machine algorithm. Neural Network appears to produce more consistent results with minimal standard deviation. The significant difference was less than 0.05 in independent samples tested by the SPSS statistical tool.

The factors that are affecting the proposed work to increase the results are SourceIP, Destination Port with a significant value of (0.75), Source Port, Destination IP with a significant value of (0.48), Flow duration, flow IAT Max with a significant value of (0.95), Fwd IAT total, Flow Duration with a significant value of (0.67). These specific factors are the factors for influencing the accuracy in the proposed model.

In the finding of (Han et al. 2020), Neural Network has an Accuracy (75%) and SVM Algorithm has an accuracy (72%) which are made for detection of android malware. In proposed work of (Mahindru and Sangal 2020), Neural Network has an Accuracy (73%) and SVM Algorithm has an accuracy (70%) which are made for profiling the malware detection in handy computing devices. The work proposed by (Ma et al. 2019) has a different finding as compared to proposed work where it has findings like, Decision Tree Algorithm has an accuracy (80%) and Logistic Regression Algorithm has an Accuracy (78%) which are made used for classification of android malware detection. It limits the consideration of features for training the model.

Our institution is passionate about high quality evidence based research and has excelled in various fields (Vijayashree Priyadharsini 2019; Ezhilarasan, Apoorva, and Ashok Vardhan 2019; Ramesh et al. 2018; Mathew et al. 2020; Sridharan et al. 2019; Pc, Marimuthu, and Devadoss 2018; Ramadurai et al. 2019). We hope this study adds to this rich legacy.

The Suggested model cannot give appropriate results for smaller train data. Here the model is not able to consider all the given feature variable parameters for training of the model. This limits the suggested model for classification of Android Malware Detection.

The future scope of proposed work will be the prediction of malware to be classified under the class for which has to be given and all the given parameters of the future variable with higher rate of accuracy and together with advanced technique and lesser time complexity.

5. Conclusion

The proposed model for classification of Android Malware Detection has accuracy(%) 82.91% for Neural Network model compared with Support Vector Machine accuracy of(%)79.67%.The proposed model proves that Neural Network Model has greater significant accuracy than SVM to this classification model.

Declarations

Conflict of Interests

No conflict of interest in this manuscript.

Author Contributions

Author T. Sai Tejeshwar Reddy was involved in data collection, data analysis, and manuscript writing. Author Dr.A. Sivanesh Kumar was involved in conceptualization, data validation, and critical review of manuscript.

Acknowledgments

The authors would like to express their gratitude towards Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences (Formerly known as Saveetha University) for providing the necessary infrastructure to carry out this work successfully.

Funding: We thank the following organizations for providing financial support that enabled us to complete the study.

1. Surya Informatics Solutions Pvt. Ltd. Chennai.
2. Saveetha University.
3. Saveetha Institute of Medical and Technical Sciences.
4. Saveetha School of Engineering.

References

Agrawal, R., Shah, V., Chavan, S., Gourshete, G., & Shaikh, N. (2020). Android Malware Detection Using Machine Learning. In *2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE),* 1-4. https://doi.org/10.35940/ijrte.b1011.0982s1219

Ezhilarasan, D., Apoorva, V.S., & Ashok Vardhan, N. (2019). Syzygium cumini extract induced reactive oxygen species‐mediated apoptosis in human oral squamous carcinoma cells. *Journal of Oral Pathology & Medicine: Official Publication of the International Association of Oral Pathologists and the American Academy of Oral Pathology, 48*(2), 115-121.

Gheena, S., & Ezhilarasan, D. (2019). Syringic acid triggers reactive oxygen species–mediated cytotoxicity in HepG2 cells. *Human & experimental toxicology, 38*(6), 694-702.

Han, H., Lim, S., Suh, K., Park, S., Cho, S.J., & Park, M. (2020). Enhanced Android Malware Detection: An SVM-Based Machine Learning Approach. *In 2020 IEEE International Conference on Big Data and Smart Computing (BigComp),* 75-81.

https://doi.org/10.1109/bigcomp48618.2020.00-96

Jose, J., & Subbaiyan, H. (2020). Different treatment modalities followed by dental practitioners for ellis class 2 fracture–A questionnaire-based survey. *The Open Dentistry Journal, 14*(1), 59–65.

Jung, J., Kim, H., Shin, D., Lee, M., Lee, H., Cho, S.J., & Suh, K. (2018). Android malware detection based on useful API calls and machine learning. *In 2018 IEEE First International Conference on Artificial Intelligence and Knowledge Engineering (AIKE),* 175-178.

https://doi.org/10.1109/aike.2018.00041

Ke, Y., Al Aboody, M.S., Alturaiki, W., Alsagaby, S.A., Alfaiz, F.A., Veeraraghavan, V.P., & Mickymaray, S. (2019). Photosynthesized gold nanoparticles from Catharanthus roseus induces caspase-mediated apoptosis in cervical cancer cells (HeLa). *Artificial cells, nanomedicine, and biotechnology*, *47*(1), 1938-1946.

Krishnaswamy, H., Muthukrishnan, S., Thanikodi, S., Antony, G.A., & Venkatraman, V. (2020). Investigation of air conditioning temperature variation by modifying the structure of passenger car using computational fluid dynamics. *Thermal Science, 24*(1 Part B), 495-498.

Lekssays, A., Falah, B., & Abufardeh, S. (2020). A Novel Approach for Android Malware Detection and Classification using Convolutional Neural Networks. *Proceedings of the 15th International Conference on Software Technologies.* https://doi.org/10.5220/0009822906060614

Liu, K., Xu, S., Xu, G., Zhang, M., Sun, D., & Liu, H. (2020). A Review of Android Malware Detection Approaches Based on Machine Learning. *IEEE Access, 8,* 124579-124607. https://doi.org/10.1109/access.2020.3006143

Mahindru, A., & Sangal, A.L. (2020). SOMDROID: android malware detection by artificial neural network trained using unsupervised learning. *Evolutionary Intelligence,* 1-31.

https://doi.org/10.1007/s12065-020-00518-1.

Malli Sureshbabu, N., Selvarasu, K., Nandakumar, M., & Selvam, D. (2019). Concentrated growth factors as an ingenious biomaterial in regeneration of bony defects after periapical surgery: A report of two cases. *Case reports in dentistry.*

Mathew, M.G., Samuel, S.R., Soni, A.J., & Roopa, K.B. (2020). Evaluation of adhesion of Streptococcus mutans, plaque accumulation on zirconia and stainless steel crowns, and surrounding gingival inflammation in primary molars: Randomized controlled trial. *Clinical oral investigations, 24*(9), 3275-3280. https://link.springer.com/article/10.1007/s00784-020-03204-9.

Ma, Z., Ge, H., Liu, Y., Zhao, M., & Ma, J. (2019). A combination method for android malware detection based on control flow graphs and machine learning algorithms. *IEEE access, 7,* 21235-21245. https://doi.org/10.1109/access.2019.2896003.

Mehta, M., Tewari, D., Gupta, G., Awasthi, R., Singh, H., Pandey, P., & Satija, S. (2019). Oligonucleotide therapy: an emerging focus area for drug delivery in chronic inflammatory respiratory diseases. *Chemico-biological interactions, 308,* 206-215.

Muthukrishnan, S., Krishnaswamy, H., Thanikodi, S., Sundaresan, D., & Venkatraman, V. (2020). Support vector machine for modelling and simulation of Heat exchangers. *Thermal Science, 24* (1 Part B), 499-503.

PC, J., Marimuthu, T., Devadoss, P., & Kumar, S.M. (2018). Prevalence and measurement of anterior loop of the mandibular canal using CBCT: A cross sectional study. *Clinical implant dentistry and related research, 20*(4), 531-534. https://europepmc.org/article/med/29624863

Ramadurai, N., Gurunathan, D., Samuel, A.V., Subramanian, E., & Rodrigues, S.J. (2019). Effectiveness of 2% Articaine as an anesthetic agent in children: randomized controlled trial. *Clinical oral investigations, 23*(9), 3543-3550.

Ramesh, A., Varghese, S., Jayakumar, N.D., & Malaiappan, S. (2018). Comparative estimation of sulfiredoxin levels between chronic periodontitis and healthy patients–A case‐control study. *Journal of periodontology, 89*(10), 1241-1248.

Samuel, M.S., Bhattacharya, J., Raj, S., Santhanam, N., Singh, H., & Singh, N.P. (2019). Efficient removal of Chromium (VI) from aqueous solution using chitosan grafted graphene oxide (CS-GO) nanocomposite. *International journal of biological macromolecules, 121,* 285-292.

Samuel, S.R., Acharya, S., & Rao, J.C. (2020). School Interventions–based Prevention of Early‐Childhood Caries among 3–5‐year‐old children from very low socioeconomic status: Two‐year randomized trial. *Journal of public health dentistry, 80*(1), 51-60.

Sathish, T., & Karthick, S. (2020). Wear behaviour analysis on aluminium alloy 7050 with reinforced SiC through taguchi approach. *Journal of Materials Research and Technology, 9*(3), 3481-3487.

Sharma, P., Mehta, M., Dhanjal, D.S., Kaur, S., Gupta, G., Singh, H., & Satija, S. (2019). Emerging trends in the novel drug delivery approaches for the treatment of lung cancer. *Chemico-biological interactions, 309,* 108720.

Sridharan, G., Ramani, P., Patankar, S., & Vijayaraghavan, R. (2019). Evaluation of salivary metabolomics in oral leukoplakia and oral squamous cell carcinoma. *Journal of Oral Pathology & Medicine: Official Publication of the International Association of Oral Pathologists and the American Academy of Oral Pathology, 48*(4), 299-306.

Tahtaci, B., & Canbay, B. (2020). Android Malware Detection Using Machine Learning. *In 2020 Innovations in Intelligent Systems and Applications Conference (ASYU),* 1-6.

https://doi.org/10.1109/asyu50717.2020.9259834.

Varghese, S.S., Ramesh, A., & Veeraiyan, D.N. (2019). Blended Module‐Based Teaching in Biostatistics and Research Methodology: A Retrospective Study with Postgraduate Dental Students. *Journal of dental education*, *83*(4), 445-450.

Venu, H., Raju, V.D., & Subramani, L. (2019). Combined effect of influence of nano additives, combustion chamber geometry and injection timing in a DI diesel engine fuelled with ternary (diesel-biodiesel-ethanol) blends. *Energy*, *174*, 386-406.

Venu, H., Subramani, L., & Raju, V.D. (2019). Emission reduction in a DI diesel engine using exhaust gas recirculation (EGR) of palm biodiesel blended with TiO2 nano additives. *Renewable Energy*, *140*, 245-263.

Vignesh, R., Ditto Sharmin, C., Annamalai, S., & Baghkomeh, P.N. (2019). Management of complicated crown-root fracture by extra-oral fragment reattachment and intentional reimplantation with 2 years review. *Contemporary clinical dentistry*, *10*(2), 397-401.

Jain, S.V., Muthusekhar, M.R., Baig, M.F., Senthilnathan, P., Loganathan, S., Wahab, P.A., & Vohra, Y. (2019). Evaluation of three-dimensional changes in pharyngeal airway following isolated lefort one osteotomy for the correction of vertical maxillary excess: a prospective study. *Journal of maxillofacial and oral surgery*, *18*(1), 139-146.

Vijayashree Priyadharsini, J. (2019). In silico validation of the non‐antibiotic drugs acetaminophen and ibuprofen as antibacterial agents against red complex pathogens. *Journal of periodontology*, *90*(12), 1441-1448.

Wen, L., & Yu, H. (2017). An Android malware detection system based on machine learning. *In AIP Conference Proceedings, AIP Publishing LLC*, *1864*(1), 020136. https://doi.org/10.1063/1.4992953

Rosmansyah, Y., & Dabarsyah, B. (2015, August). Malware detection on android smartphones using API class and machine learning. *In 2015 International Conference on Electrical Engineering and Informatics (ICEEI)*, 294-297. https://doi.org/10.1109/iceei.2015.7352513

Yerima, S.Y., Sezer, S., & Muttik, I. (2015). High accuracy android malware detection using ensemble learning. *IET Information Security*, *9*(6), 313-320.

https://doi.org/10.1049/iet-ifs.2014.0099.

Yuan, Z., Lu, Y., & Xue, Y. (2016). Droiddetector: android malware characterization and detection using deep learning. *Tsinghua Science and Technology*, *21*(1), 114-123.

https://doi.org/10.1109/tst.2016.7399288