2023 International Conference on Computer Communication and Informatics (ICCCI), Jan 23-25, 2023, Coimbatore, India

E-Commerce Customer Churn Prediction Scheme Based on Customer Behaviour Using Machine Learning

Nagaraj P
Department of Computer Science and Engineering
Kalasalingam Academy of Research andEducation
Krishnankoil, Virudhunagar, India
nagaraj.p@klu.ac.in

Muneeswaran V
Department of Electronics and Communication Engineering
Kalasalingam Academy of research and Education
Krishnankoil, Virudhunagar, India
munees.klu@gmail.com

Dharanidharan A
Department of Computer Science and Engineering
Kalasalingam Academy of Research andEducation
Krishnankoil, Virudhunagar, India
dharanidharanashok@gmail.com

Aakash M
Department of Computer Science and Engineering
Kalasalingam Academy of Research andEducation
Krishnankoil, Virudhunagar, India
aakashraina2002@gmail.com

Balananthanan K
Department of Computer Science and Engineering
Kalasalingam Academy of Research andEducation
Krishnankoil, Virudhunagar, India
nanthananbala@gmail.com

Rajkumar C
Department of Computer Science and Engineering
Kalasalingam Academy of Research andEducation
Krishnankoil, Virudhunagar, India
stephenrajkumar0705@gmail.com

2023 International Conference on Computer Communication and Informatics (ICCCI) | 979-8-3503-4821-7/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICCCI56745.2023.10128498

***Abstract*—**

Customer retention is crucial for any company to succeed given the rapid growth in the number of companies in established sectors. To analyze and research client retention, numerous approaches (such as data mining and statistics) have been offered. Building digital (CRM) which is known as Customer Relationship Management systems is a new trend that is growing in the global economy as a result of the explosive rise of digital systems and related information technologies. In the telecoms sector, where businesses are progressively going digital, this trend is especially pronounced. Predicting customer attrition is a key function of contemporary telecom CRM systems. Keeping consumers has grown increasingly crucial for insurance businesses, and causes for churning are challenging, according to competition in Iran's insurance industry in recent years and the entry of the private sector. Churners have continually been a serious drawback for each business that provides services. Churning drives up a company's expenses whereas additionally lowering its gross margin. Customers WHO request service termination are usually exhibiting client attrition. those that offer the information keep in government databases, furthermore because the organizations WHO fund the gathering of such information, are at the same time turning into a lot of conscious of however tools that enhance analytical capabilities additionally offer threats to the privacy of information records. However, it is the potential to forecast in case a customer wishes to avoid the service using predictive analysis according to previous utilization, the performance of the service, and other kinds of patterns. Customer churn is a problem for many sectors, but it is especially severe in the fiercely competitive and currently widely liberalized mobile telecommunications sector. The average monthly churn count for mobile telecommunications is reportedly 2.2%. In addition to potential costs from lower revenues, losing customers also increases the demand for acquiring new ones.

Keywords— Customer retention, Churn prediction, E-Commerce, retention rate, predictive analysis.

I. INTRODUCTION

Churn prediction is the process of identifying consumers who are most likely to leave your business or end a service subscription based on their usage of your product. To properly anticipate customer churn, you'll need to combine and use the key signs that your team has identified. This will let you know when a client is likely to leave so that your business can take appropriate action. Long-distance phone calls and periodicals are two subscription services that typically fall under the category of businesses with a high churn rate. Understanding consumer behavior is made easier with the use of churn analysis. Forecast the possibility that users may unsubscribe and transfer their business to a rival. In addition to estimating staff attrition, there are other uses. Finding patterns in data is known as data mining. As the collection of data is gathered the number of data increases every year because of this data mining is becoming a highly used way of converting this data into information. This is mostly used with various profiling methods such as marketing, surveillance, fraud detection, and scientific research. Data mining has a wide range of applications in both the public as well as private sectors. The churning analysis method is one unique feature of data mining. The rate of client attrition for any business is calculated in this way. Finding out which clients are often likely to cease using a product or service is required for this. Churn analysis has a productive impact on the creation of a firm's effective and long-term client retention strategy. It is straightforward to use churn analysis to create a thorough study of the reasons for the churn rate when a company is conscious of the percentage of consumers that stop doing business with them within a certain time frame this aids the organization in creating efficient customer retention campaigns. In addition to identifying at-risk consumers, churn prediction also identifies the pain areas that precede churn, assisting in boosting overall customer happiness and retention. Churn can be avoided by foreseeing it. As a result, preventing churn offers businesses a terrific way to generate income. The world's markets are growing more and more saturated as more and more customers switch their registered services between rival businesses. Companies have thus learned that client retention should be the primary goal of their marketing initiatives rather than customer acquisition. Studies have revealed that a company spends a lot more money trying to attract new consumers than it would if it were trying to keep its current ones. High-risk clients who intend to stop using your services or go to competitors can be the focus of customer retention initiatives. Customer churn is a term used to describe the consequence of customer loss. Framed can identify high-risk consumers before they churn by using machine learning. This churn evaluation is carried out monthly for a particular business so that it can later implement a tailored marketing campaign to keep these consumers. Various businesses interested in developing more sophisticated client retention tactics are offered the service of this high-risk customer identification process.

II. RELATED WORKS

Huang et al. [1] In the realm of computer science, machine learning is constantly revolutionizing research. Machine learning is used in the modern world to address real-time problems by storing, manipulating, extracting, and retrieving data from vast sources. This paper focuses on the efficient and successful creation of machine-learning models using electronics, mathematics, and statistics. We use numerous machine learning tools in daily life, such as Google Translate and Google Maps. Many of us use these apps without realizing how machine learning is combined with fields like math, statistics, and electronics. To comprehend the capabilities of the aforementioned disciplines, we presented an example of weather forecasting, which is a difficult task.

Xai et al. [2] Researching client churn prediction cases each in home and foreign carriers, the strategy was compared with an artificial neural network, call tree, provision regression, and naive theorem classifier. it's found that the strategy enjoys the simplest accuracy rate, hit rate, covering rate, and raise constant, and so, provides an efficient mensuration for client churn prediction.

Hassouna et al. [3] Mobile providers covertly employ data mining tools like regression and decision trees to address churn issues. The experimental findings of this study show that the decision tree model outperforms all logistic regression models examined, including the model developed by a data analytics team working for the cell operator. The decision tree model's lift value at the 30% percentile is 1.598 as opposed to the data analytics team's 1.4 at the same percentile. It is found that decision trees are a superior tool for analyzing customer turnover for this issue and related data sets. More importantly, although data mining provides incredibly valuable insight into customer attrition, limitations in terms of significance, causation, data evolution, and model complexity are clear.

LU et al. [4] This study conducts an experimental investigation into forecasting customer attrition using real data sets. Unlike existing churn prediction techniques, our system allows for the creation of an "Implementation Zone" where customers with the highest likelihood of leaving can be targeted for retention strategies. This study's application of the boosting technique is another unique feature. Instead of attempting to boost a base learner directly using a boosting algorithm, as most researchers did, this study aims to divide the training data according to the difficulty of fitting a base learner and build a specific prediction model for each defined cluster. The outcomes are tested on a live database to compare the findings with a logistic regression model fitted by the entire training set.

Soeini et al. [5] In today's competitive markets, churn customers are crucial for insurance companies. A form was used to collect data because there wasn't enough relevant information. The majority of the variables were demographic and conceptual, and fewer environmental and behavioral. First, determinative customers were identified by the group, then goal indexes were assigned by a predetermined one, and patterns were extracted by call tree, revealing that officers or engineers account for nearly all churn customers. The majority of their business is with Persia Insurance, but they are dissatisfied with the services they receive and eager to leave.

Lemmens et al. [6] apply the two procedures to a client data set of a mysterious U.S. remote broadcast communications organization, and both essentially further develop precision in foreseeing agitate. This higher prescient presentation could eventually prompt steady benefits for organizations that utilize these techniques. Moreover, the outcomes suggest the utilization of a reasonable testing plan while foreseeing an uncommon occasion from huge informational indexes, however, this requires a fitting predisposition rectification.

Fontaine et al. [7] The examination results showed that the utilization of the Gullible Bayes classifier beats the Brain Network-based framework. Taking note that the NB is a significant classifier turns out to be more exact when the data connected with family wages aren't thought of as a key input highlight. For both NN and NB frameworks, the data on the item or administration to be sold is significant for the exactness of expectation.

Perianez et al. [8] Endurance examination centers around foreseeing the hour of the event of a specific occasion, beat for our situation. Old-style strategies, like relapses, could be applied when the game is get emptied by the absence of players. The test emerges for datasets with inadequate beating data for every player, as the greater part of their interface with the game. This is known as a blue-penciled information issue and is in the idea of a stir. Blue penciling is regularly managed endurance examination procedures, yet because of the rigidity of the endurance measurable calculations, the precision accomplished is much of the time poor. Conversely, novel group learning procedures, progressively famous in different logical fields, give fashionable forecast results. In this work, we create, without precedent for the social games space, an endurance troupe model which furnishes a thorough examination along with an exact expectation of stir.

Liu et al.[9] We can improve our results by using vast datasets and cutting-edge profound learning-based techniques. However, because of their complex designs, these techniques demand greater processing resources. Later, this study can be expanded using advanced learning-based techniques to handle upgrading the client stir expectation display for the telecom industry. Future efforts will focus on improving the efficiency and suitability of this system.

Marin Diaz et al.[10] We used the implemented model to analyze data from a company that manufactures products and has a sizable network of partners (customers) spread across the globe. Numerous predictive models have been used with incomplete data; the beat rate covers 1.06% of the entire example. To solve the problem of unequal grouping, information-adjusting methods must be used. The dataset must also be subjected to calculations for subsampling and oversampling at the same time.

III. Materials and Methods

In this project, we are faced with different types of steps to construct a customer churn prediction model. As in the first step, the dataset starts analysis and it will clean the data through data preprocessing where the data will be cleaned to get the best performance during modeling. Finally, machine learning algorithms are used for creating a customer churn prediction model [21-41].

3.1 Data Collection

An e-commerce website called Kaggle is where the information for this research was gathered. For analysis and forecasting, the consumption information of customers who made purchases on the website is chosen. For analysis and forecasting, the consumption information of customers who made purchases on the website is chosen.

3.2. Data Description

Data is composed of a foremost E-commerce platform, it was historical data containing customers in this data consisting of attributes, descriptions, and type, and it's used for churn prediction.

3.3. Data Cleaning

After data exploration, some elements of data cleaning can be carried out to assist built highly error-free machine-learning models and remove bias. Data cleaning has the added benefit of almost eliminating unused and unneeded data that has been there for a while. It also provides accurate input data. All missing data have now been erased to prevent errors from occurring while creating the models.

Table 1: Data description

Attributes	**Description**	**Type**
ID used by the Customer	ID is given in a unique for everyone	Numeric
Churn	Churn Flag	Numeric
Tenure	Based on organization	Numeric
Comfortable Login Device	Preferred login device based on customer preference	Character
City Tier	City Tire	Numeric
Warehouse To Home	Distance between one end to other	Numeric
Payment Mode	As per the preference of the Customer	Character
Gender	Gender of the participants	character
Hour Spend On App	Calculates the time spent in that app	Numeric
Count Of Device Registered	The total number of deceives registered on a particular customer	Numeric
Selected Order Cat	Favored category of order based on the customer in last month	Character
Satisfaction Score	The adequate count of the customer on service	Numeric
Status based on marital life	Marital status of the customer	Character
Number Of Address	Complete count added on a particular customer	Numeric
Complain	Collection of a complaint which is been raised in last month	Numeric
Order Amount Hike based on the previous last Year	Percentage increases in order from last year	Numeric
Coupon Used	A complete count of coupons has been used in the last month	Numeric
Order Count	orders which have been placed in the previous month	Numeric
Last Order date	Day Since the last order by the customer	Numeric
Cashback Amount	Average cashback in the last month	Numeric

3.4 Model Building

The Following Steps consist of so many stages that must be taken to begin with splitting the data into training steps and testing data. Then, we have going to appeal to machine learning techniques to build the model to compare the accuracy of this model.

3.4.1. Decision Tree

A non-supervisory machine literacy method that can be used for bracketing or retrogression is decision trees. A decision tree's goal is to create a model that forecasts the value of the outgrowth by learning the straightforward decision rules presumptively derived from data attributes to demonstrate the model's delicate nature.

3.4.2. Random Forest

Random forest is a machine learning technique that merges the output of numerous decision training for a single output of the model, it controls some over lifting and bias issues associated with a decision tree and it will give accurate predictions especially when individual trees are uncorrelated with some other data.

3.4.3. Support Vector Machine (SVM).

Support Vector Machine is a supervised machine learning technique used for both regression and classification. The SVM steps are used to import the library package and to import the dataset it will extract the X and Y separately and the dataset will divide into train and test the data it will initialize the SVM classifier model and the coming up with a prediction of the model.

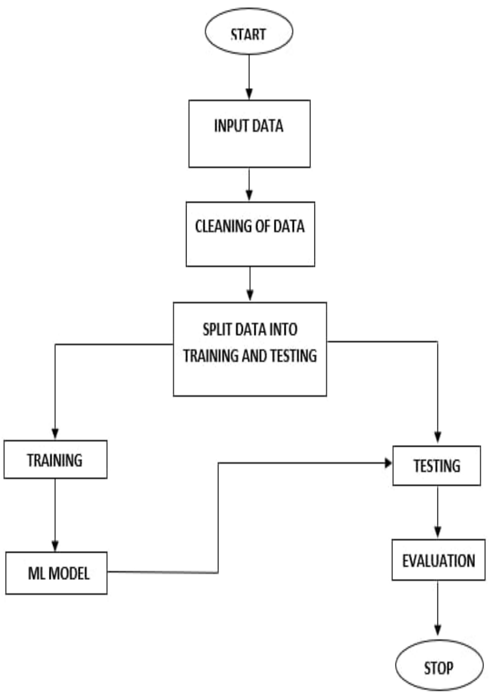

Fig 1: A generic flow of proposed work

3.5. Data Preprocessing

After gathering the data, we next took efforts to weed out any null data before building the models. These processes were necessary to ensure that the models we developed would work as effectively as possible. Therefore, the data should be split into the training and testing data the training data will consist of 75% of the total dataset and the test dataset will have consists of 25% of the total data set. The outcome of the project is will be a churn, so the Customer column will be removed from the predictor column all columns will be converted into numerical values.

Table 2: Comparison table of Existing works

Ref. No.	Description	Drawbacks
[21]	In this paper, the k-fold cross-validation technique is used to select the best model by estimating its performance metrics. Accuracy – 81.72	The accuracy level is too low, so some steps should be taken for increasing the level of accuracy
[22]	In this paper support vector machine (SVM) on structural risk minimization was applied to customer churn prediction. Accuracy – 89.83	insoluble for traditional methods in the customer churn prediction of telecommunication.
[23]	This paper tries to fill this gap by empirically comparing two techniques: Customer churn - decision tree and logistic regression models Accuracy - 78	Churn relates to complex interactions within the population. Examining all the factors affecting customer churn simultaneously and jointly by building a model which is not applicable.
[24]	In this paper, they used Gradient Boosting algorithms and executed with those methods Accuracy - 84	Due to the non-stationary data model phenomenon the results may differ.
[25]	In this paper, the model they proposed is using k-means with the number of clusters they produced optimal results. Accuracy- 91	The method followed in this paper will produce optimal results but this can't be a robust or feasible solution.
[26]	In this paper they used color bagging and stochastic gradient boosting, with classification trees as base classifiers, to the balanced calibration. . Accuracy-83	This method cancels the advantage relative to the balanced sampling. Further experiments showed that reweighting observations of a balanced training.
[27]	In this paper Neural Networks trained with a back-propagation algorithm and Naïve Bayesian (NB) classifiers are used to predict the profile of prospective clients	No drawbacks as the paper has explained about the work done in all papers

[28]	In this paper, they proposed a survival ensemble model which provides a comprehensive analysis together with an accurate prediction of churn Accuracy-94	This is not a flexible technique that does not require a previous manipulation of the data and that can deal efficiently with the temporal dimension of the churn prediction problem
[29]	To enhance the performance of the CCP model, the proposed churn prediction model employs a collection of clustering and classification algorithms. Precision: 93.6	This model performs well only for the classification of customer churn. Accuracy may not be the same if the work is extended to other classes.
[30]	This study examines the binary target variable abandonment in a B2B setting while taking the partner's interactions with the contact center into account. Accuracy – 89.32	There are only a few studies for knowing about the churn rate within a B2B environment
[31]	In this study, SVM-RBF and GBM take advantage of the possibility of more nuanced features reflecting user preferences. Accuracy-86	Research into the nature of the underlying decision-making process may shift in the future from predictive to causal modeling.
[32]	In this paper for the case study, they choose analytical techniques that belong to different categories of learning. Accuracy-93	By including hybrid models and deep learning models, the study can be expanded. It is possible to evaluate performance using additional performance measures.
[33]	This paper provides helps to know about the stage of the feature of unsupervised characteristics engineering ability of DNNs performs across every company when presented with abstract using the input of vectors Accuracy-92	The data representation underperforms when it is compared to the currently employed feature engineering methods at Framed.
[34]	In this paper, they use different ml algorithms.	The feature values produce a low accuracy rate.
[35]	This paper uses AdaBoost an ensemble approach to predict the imbalance in the eCommerce data. Accuracy-93	This paper only describes the approach to predict the imbalance in the data.
[36]	This paper the Stacking Classifier aims at analyzing and anticipating the churn in eCommerce data using various features such as Gender, Tenure, etc. Accuracy-96.34	The model described in the paper doesn't have any drawback where the stack classifier is tested with all the ml algorithms and yet doesn't fail to produce an optimal result.
[37]	To promote electronic commerce, web data mining technology is applied to predict the customer's chum online which can help to reduce the operating costs and enhance the competitive power of the enterprises.	This is only a development plan that provides brief insights into web data mining techniques.
[38]	This paper introduced a new tool that is more powerful for the data extraction knowledge from vast data for this decision tree method is used Accuracy-90.96	They were limited to looking at the variables that were listed in the bank's database. Additionally, it took a long time to extract all the data due to the database's vast number of data and the privacy concerns that went along with it.
[39]	This paper examines the topic of churn analysis while imagining a situation in which a business with private databases wants to use a churn analysis technique on the combination of such databases without disclosing any unwanted information.	With the incomplete dataset, the result that is produced is incorrect which made this a failed model.
[40]	Based on an investigation of time interval distributions in consumer viewpoints, this research gives an empirical analysis of the sale transactions from the 360buy website. Accuracy-96.5	This paper failed to compare differences.

IV. Conclusion

Customer churn is a very important topic in E-commerce. The Churn Prediction is a main indicator for E-commerce like subscription-based companies. The three different algorithms that were applied to predict the E-commerce customer churn are Support Vector Machine (SVM), Decision Tree, and Random Forest. From this algorithm, Random Forest has a good accuracy rate. Using Machine Learning in Customer Churn Prediction will help to increase the customer retention rate. Machine Learning will provide an accurate and profitable way to predict customer churn. This E-commerce customer churn prediction will help to stop many customers leaving from the company. Using these Customer churn predictions, we will easily find the problem and reduce customer churn easily. This churn prediction will help the e-commerce businesses to build or give a better and good product to the customers and also helps to provide good customer support to their customers.

REFERENCES

[1] Huang, B., Kechadi, M. T., & Buckley, B. (2012). Customer churn prediction in telecommunications. *Expert Systems with Applications*, *39*(1), 1414-1425.

[2] Xia, G. E., & Jin, W. D. (2008). Model of customer churn prediction on support vector machine. *Systems Engineering-Theory & Practice*, *28*(1), 71-77.

[3] Hassouna, M., Tarhini, A., Elyas, T., & AbouTrab, M. S. (2016). Customer churn in mobile markets a comparison of techniques. *arXiv preprint arXiv:1607.07792*.

[4] Lu, N., Lin, H., Lu, J., & Zhang, G. (2012). A customer churns prediction model in the telecom industry using boosting. *IEEE Transactions on Industrial Informatics*, *10*(2), 1659-1665.

[5] Soeini, R. A., & Rodpysh, K. V. (2012). Applying data mining to insurance customer churn management. *International Proceedings of Computer Science and Information Technology*, *30*, 82-92.

[6] Lemmens, A., & Croux, C. (2006). Bagging and boosting classification trees to predict churn. *Journal of Marketing Research*, *43*(2), 276-286.

[7] Fontaine, L., & Selouani, S. A. Intelligent Agents for Customer Behavior Prediction to Improve Relationship Marketing.

[8] Periáñez, Á., Saas, A., Guitart, A., & Magne, C. (2016, October). Churn prediction in mobile social games: Towards a complete assessment using survival ensembles. In *2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA)* (pp. 564-573). IEEE.

[9] Liu, R., Ali, S., Bilal, S. F., Sakhawat, Z., Imran, A., Almuhaimeed, A., ... & Sun, G. (2022). An Intelligent Hybrid Scheme for Customer Churn Prediction Integrating Clustering and Classification Algorithms. *Applied Sciences*, *12*(18), 9355.

[10] Marín Díaz, G., Galán, J. J., & Carrasco, R. A. (2022). XAI for Churn Prediction in B2B Models: A Use Case in an Enterprise Software Company. *Mathematics*, *10*(20), 3896.

[11] Fridrich, M., & Dostál, P. (2022). User Churn Model in E-Commerce Retail.

[12] Sabbeh, S. F. (2018). Machine-learning techniques for customer retention: A comparative study. *International Journal of Advanced Computer Science and Applications*, *9*(2).

[13] Spanoudes, P., & Nguyen, T. (2017). Deep learning in customer churn prediction: unsupervised feature learning on abstract company independent feature vectors. *arXiv preprint arXiv:1703.03869*.

[14] Granov, A. (2021). Customer loyalty, return and churn prediction through machine learning methods: for a Swedish fashion and e-commerce company.

[15] Dhote, S., Vichoray, C., Pais, R., Baskar, S., & Mohamed Shakeel, P. (2020). Hybrid geometric sampling and AdaBoost based deep learning approach for data imbalance in E-commerce. *Electronic Commerce Research*, *20*(2), 259-274.

[16] Awasthi, S. (2022). Customer Churn Prediction on E-Commerce Data using Stacking Classifier.

[17] Zhang, W., & Zhu, L. (2017). Electronic Commerce Customer Churn Prediction Model Based on Web Data Mining.

[18] Keramati, A., Ghaneei, H., & Mirmohammadi, S. M. (2016). Developing a prediction model for customer churn from electronic banking services using data mining. *Financial Innovation*, *2*(1), 1-13.

[19] Forhad, N., Hussain, M. S., & Rahman, R. M. (2014, September). Churn analysis: Predicting churners. In *Ninth International Conference on Digital Information Management (ICDIM 2014)* (pp. 237-241). IEEE.

[20] Wang, J., Gao, K., & Li, G. (2010). Empirical analysis of customer behaviors in Chinese e-commerce. *Journal of networks*, *5*(10), 1177-1184.

[21] Nagaraj, P., & Deepalakshmi, P. (2020). A framework for e-healthcare management service using recommender system. *Electronic Government, an International Journal*, *16*(1-2), 84-100.

[22] Nagaraj, P., Deepalakshmi, P., Mansour, R. F., & Almazroa, A. (2021). Artificial flora algorithm-based feature selection with gradient boosted tree model for diabetes classification. *Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy*, *14*, 2789.

[23] Muneeswaran, V., Nagaraj, P., Dhannushree, U., Ishwarya Lakshmi, S., Aishwarya, R., & Sunethra, B. (2021). A Framework for Data Analytics-Based Healthcare Systems. In *Innovative Data Communication Technologies and Application* (pp. 83-96). Springer, Singapore.

[24] Muneeswaran, V., Nagaraj, M. P., Rajasekaran, M. P., Chaithanya, N. S., Babajan, S., & Reddy, S. U. (2021, July). Indigenous Health Tracking Analyzer Using IoT. In *2021 6th International Conference on Communication and Electronics Systems (ICCES)* (pp. 530-533). IEEE.

[25] Pa, N., Mb, A., Kb, B., & Ab, D. (2020). Analysis of data mining techniques in diagnalising heart disease. *Intelligent Systems and Computer Technology*, *37*, 257.

[26] Muneeswaran, V., BenSujitha, B., Sujin, B., & Nagaraj, P. (2020). A compendious study on security challenges in big data and approaches of feature selection. *International Journal of Control and Automation*, *13*(3), 23-31.

[27] Vb, S. K. (2020). Perceptual image super resolution using deep learning and super resolution convolution neural networks (SRCNN). *Intelligent Systems and Computer Technology*, *37*(3).

[28] Nagaraj, P., Muneeswaran, V., Muthamil Sudar, K., Hammed, S., Lokesh, D. L., & Samara Simha Reddy, V. (2022). An Exemplary Template Matching Techniques for Counterfeit Currency Detection. In *International Conference on Image Processing and Capsule Networks* (pp. 370-378). Springer, Cham.

[29] Nagaraj, P., & Deepalakshmi, P. (2021). Diabetes Prediction Using Enhanced SVM and Deep Neural Network Learning Techniques: An Algorithmic Approach for Early Screening of Diabetes. *International Journal of Healthcare Information Systems and Informatics (IJHISI)*, *16*(4), 1-20.

[30] Nagaraj, P., & Deepalakshmi, P. (2022). An intelligent fuzzy inference rule-based expert recommendation system for predictive diabetes diagnosis. *International Journal of Imaging Systems and Technology*.

[31] Birunda, S. S., Nagaraj, P., Narayanan, S. K., Sudar, K. M., Muneeswaran, V., & Ramana, R. (2022, January). Fake Image Detection in Twitter using Flood Fill Algorithm and Deep Neural Networks. In *2022 12th International Conference on Cloud Computing, Data Science & Engineering (Confluence)* (pp. 285-290). IEEE.

[32] Sudar, K. M., Nagaraj, P., Nithisaa, S., Aishwarya, R., Aakash, M., & Lakshmi, S. I. (2022, April). Alzheimer's Disease Analysis using Explainable Artificial Intelligence (XAI). In *2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)* (pp. 419-423). IEEE.

[33] Muneeswaran, V., Nagaraj, P., & Ijaz, M. F. (2022). An Articulated Learning Method Based on Optimization Approach for Gallbladder Segmentation from MRCP Images and an Effective IoT Based Recommendation Framework. In *Connected e-Health* (pp. 165-179). Springer, Cham.

[34] Sunethra, B., Sreeya, C., Dhannushree, U., Nagaraj, P., & Muneeswaran, V. (2022, April). A Systematic Parking System Using bi-class Machine Learning Techniques. In *2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)* (pp. 221-226). IEEE.

[34] Nagaraj, P., Muneeswaran, V., Dharanidharan, A., Balananthanan, K., Arunkumar, M., & Rajkumar, C. (2022, April). A Prediction and Recommendation System for Diabetes Mellitus using XAI-based Lime Explainer. In *2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)* (pp. 1472-1478). IEEE.

[35] Kumar, B. M., Rao, K. R. K., Nagaraj, P., Sudar, K. M., & Muneeswaran, V. (2022, April). Tobacco Plant Disease Detection and Classification using Deep Convolutional Neural Networks. In *2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)* (pp. 490-495). IEEE.

[36] Nagaraj, P., Muneeswaran, V., & Deshik, G. (2022, August). Ensemble Machine Learning (Grid Search & Random Forest) based Enhanced Medical Expert Recommendation System for Diabetes Mellitus Prediction. In *2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC)* (pp. 757-765). IEEE.

[37] Vignesh, K., & Nagaraj, P. (2022). Analysing the Nutritional Facts in Mc. Donald's Menu Items Using Exploratory Data Analysis in R. In *International Conference on Emerging Technologies in Computer Engineering* (pp. 573-583). Springer, Cham.

[38] Nagaraj, P., Deepalakshmi, P., Muneeswaran, V., & Muthamil Sudar, K. (2022). Sentiment Analysis on Diabetes Diagnosis Health Care Using Machine Learning Technique. In *Congress on Intelligent Systems* (pp. 491-502). Springer, Singapore.

[39] Brintha, N. C., Nagaraj, P., Tejasri, A., Durga, B. V., Teja, M. T., & Kumar, M. N. V. P. (2022, June). A Food Recommendation System for Predictive Diabetic Patients using ANN and CNN. In *2022 7th International Conference on Communication and Electronics Systems (ICCES)* (pp. 1364-1371). IEEE.

[40] Deny, J., Rajalakshmi, P., Muneeswaran, V., Sudharsan, R. R., & Nagaraj, P. (2022, August). Automation of Glucose Control for Type-2 Diabetes Mellitus. In *2022 3rd International Conference on Electronics and Sustatainble Communication Systems(ICESC) (pp 79-83)iEEE*