Customer churn analysis using XGBoosted decision trees

Muthupriya Vaudevan[1], Revathi Sathya Narayanan[1], Sabiyath Fatima Nakeeb[1], Abhishek[2]
[1]Department of Computer Science and Engineering, B. S. Abdur Rahman Crescent Institute of Science and Technology, Chennai, India
[2]Department of Computer Applications, B. S. Abdur Rahman Crescent Institute of Science and Technology, Chennai, India

Article Info

ABSTRACT

Customer relationship management (CRM) is an important element in all forms of industry. This process involves ensuring that the customers of a business are satisfied with the product or services that they are paying for. Since most businesses collect and store large volumes of data about their customers; it is easy for the data analysts to use that data and perform predictive analysis. One aspect of this includes customer retention and customer churn. Customer churn is defined as the concept of understanding whether or not a customer of the company will stop using the product or service in future. In this paper a supervised machine learning algorithm has been implemented using Python to perform customer churn analysis on a given data-set of Telco, a mobile telecommunication company. This is achieved by building a decision tree model based on historical data provided by the company on the platform of Kaggle. This report also investigates the utility of extreme gradient boosting (XGBoost) library in the gradient boosting framework (XGB) of Python for its portable and flexible functionality which can be used to solve many data science related problems highly efficiently. The implementation result shows the accuracy is comparatively improved in XGBoost than other learning models.

Corresponding Author:

Muthupriya Vaudevan
Department of Computer Science and Engineering
B. S. Abdur Rahman Crescent Institute of Science and Technology
Seethakathi Extate, GST Road, Vandalur, Chennai-48, India
Email: muthupriya@crescent.education

1. INTRODUCTION

In traditional information technology (IT) projects, the process of development is usually well defined and pretty straightforward. It follows the same procedure of: identifying a business case, developing a system that meets the needs of the business case, drawing timelines for deliverables, and everyone enlisted in the project is tasked with work that must comply with documented requirements. There are few ambiguities in well-constructed IT projects, and everyone understands the order of work. This isn't usually the case in data science projects. Here, business cases can be drawn up but arriving at the desired results isn't always straightforward and predictable. The only hard metric that is applicable for most data science projects is that the results derived from algorithms operating on data must be at least certain percentage "right" when compared with an accepted standard for determining correctness. Several research analyses [1]-[6] were carried out to predict the customer churn in various industries. With that being said it is important to mention that this research proposal is a data science project which involves taking a data set that is available for use and implementing a certain machine learning algorithm on it to successfully achieve a result with desired accuracy. In this paper, the machine learning algorithm used is called XGBoosted decision trees that is used to classify objects into one category or another and the final model built should be able to help in accurately predicting

the customer churn. The paper is organized is such a way that in section 2, the literature survey on the existing work is disseminated. Then in section 3, the proposed model and its design methodologies are discussed. Following in section 4, the implementation details are covered and in section 5, result of the proposed model is analyzed is detail.

Customer churn analysis: churn determinants and mediation effects of partial defection in the Korean mobile telecommunications service industry by Ahna *et al.* [7]. Retaining customers is a crucial challenge in the any industry including mobile telecommunications. Using the customer transaction and billing data captured by companies, studies have investigated the determinants of customer churn in the Korean mobile telecommunications service market. Results indicated that call quality-related are major factors in customer churn; however, factors like customers participating in membership card programs also play a vital role, which further pushes the concept down the process of understanding program effectiveness. Furthermore, it was observed that heavy users also tend to churn.

Customer churn analysis in Telecom industry by Dahiya and Bhatia [8]. There is a lot of scope for researchers in analyzing telecommunication industry data [9]-[13]. Poel and Lariviere [14] surveyed the importance of the economic value of customer retention. Since the major source of profit in any industry are its customers, customer churn plays a significant role in the survival and development of any type industry especially the telecommunications industry. Customer acquisition and retention can be improved by applying customer relationship management (CRM) tools for increasing profit and for supporting analytical tasks [15]. The association of CRM [16]-[18] further helps in capturing data and satisfying needs of soon to be non-customers in future. Understanding churn using data mining also helps these companies to employ effective marketing strategies [19]–[24]. Data mining techniques are applied in telecommunications for CRM because of the rapid growth of the huge amount of data; high pace in the market competition and increase in the churn rate [25]. These industries have suffered from high churn rates and immense churning loss. Although the business loss is unavoidable, but still churn can be managed and kept in an acceptable level. Good methods need to be developed and existing methods have to be enhanced to prevent the telecommunication industry to face challenges.

Many existing methods take plenty of time and yield accuracy below desired levels. To overcome all these challenges, we need a solution that is accurate, fast and reliable in predicting customer churn. The problem is to utilize each of the available alternatives to come up with accuracy levels that are desired while measuring the complexity levels of the taken algorithm. With the complexities involved it is necessary to explore different options available in pursuit of better optimized methods. Some its drawbacks are various levels of complexities, time consuming, varying accuracy.

The paper is organized in such a way that in section 2, the proposed model and its design methodologies are described. Following in section 4, the method and implementation details are covered and in section 5, result of the proposed model is analyzed and discussed.

2.   PROPOSED METHOD

For all businesses, customer retention is important to sustain a profitable growth through an established consumer base. To retain a customer and prevent customer churn, it is first important to identify the set of customers that are likely to leave. This would help the business to focus on these customers and take necessary steps to provide incentive to make the customers stay. Hence identification of possible "soon to be non-customers" is important.

The proposed method involves using XGBoosted decision trees to find out customer churn. Boosting is an ensemble technique for the creation of a collection of predictors. In this technique, trees are built sequentially with early trees fitting simple models to the data and then analyzing data for errors. In other words, consecutive trees are fitted (random sample) and at every step, the goal is to solve for net error from the prior tree. When an input is wrongly classified by a hypothesis, its weight is increased so that next hypothesis is more likely to classify it correctly. By combining the whole set at the end converts weak trees into a better performing model. This paper tries to experiment on the claim of XGBoost classifier to see if an accurate model can be built that outperforms existing model successfully. The proposed method aims to provide efficient and accurate result compared with existing method.

2.1. Design

The Figure 1 shows the general design and Figure 2 explains the detailed design associated with the proposed method. According to the documentation of XGBoost, it is an optimized distributed gradient boosting library designed to be highly efficient, flexible and portable. It implements machine learning algorithms under the gradient boosting framework. XGBoost provides a parallel tree boosting (also known as gradient boosting decision tree (GBDT), gradient boosting machines (GBM)) that solve many data science problems in a fast and

accurate way. The same code runs on major distributed environment (Hadoop, SGE, message passing interface (MPI)) and can solve problems beyond billions of examples.
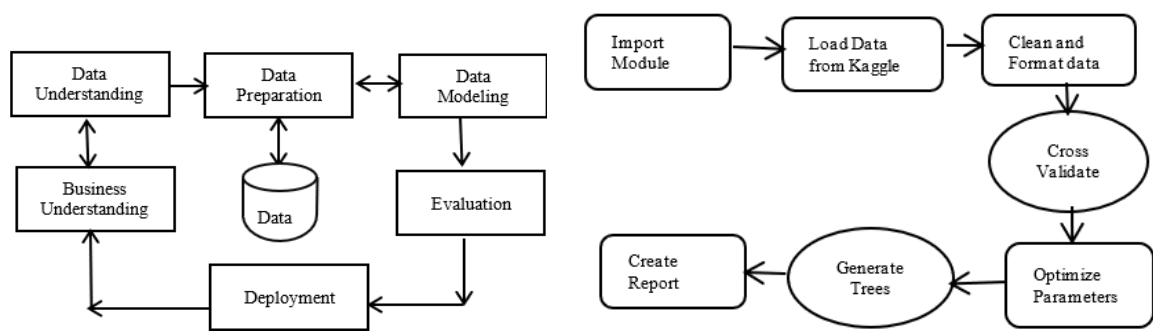

Figure 1. General design of proposed method   Figure 2. Detailed design of proposed method

2.2. Data-set design

The data set has 7043 records and 21 attribute columns. The data set includes details of customers who have left within the last month called churn, services that each customer has signed up for phone, multiple lines, internet, online security, online backup, device protection, tech support, streaming TV, movies, and account information of the customer like how long they've been a customer, contract, payment method, paperless billing, monthly charges, total charges, and demographic information about the customers like gender, age range, and if they have partners and dependents.

3. METHOD

Implementation is the stage in which theoretical design is turned out into a working system. In this section, the details of imported modules and data are given. Also, it provides information on data processing and formatting and further building of preliminary model. Finally, the confusion matrix is used to analyze the behavior of the model.

3.1. Importing modules

The selection of the correct modules/libraries is an important task as pre-written libraries make the analysis easier. Identifying the correct libraries is also crucial as importing unnecessary libraries is a waste of memory. After analysis and help from references, the following modules were installed for use: i) table libraries used library purpose pandas, ii) data manipulation and one hot encoding NumPy quantitative analysis, iii) XGBoost classifier, iv) sklearn model-selection cross validation and algorithm implement, and v) sklearn metrics for confusion matrix.

3.2. Importing data (telco from Kaggle)

After the successful installation of libraries into the notebook, the first step to do is load the data. The loaded data is downloaded from Kaggle.com and stored into a data frame called df. The data frame now contains 7043 records with 21 attribute columns each. For visualization the first five rows and 6 columns of the data set are displayed using the head() function in the Table 1.

Table 1. First five rows of data-set

S.No	Customer Id	Gender	Senior Citizen	Partner	Dependents	Tenure
1	7515	Male	0	Yes	No	1
2	5523	Female	0	No	No	34
3	3924	Male	0	No	No	2
4	9237	Male	1	No	No	45
5	4657	Female	0	No	No	2

3.3. Identifying and dealing with missing data

In Table 2, each row of the data set represents a customer record; each column given in the data set contains the customer's attributes described on the column Metadata. The next step in the analysis is to clean

and format data. For that purpose usage of the info() function takes place to get the meta data of the data set as shown in initial data set column of Table 1. After looking at this column the following conclusions are made.
i) Remove customerID column as it has unique values and will have no contribution to the analysis,
ii) Converting values in churn column from No/Yes to 0/1, and
iii) Then converting the data type of churn column from object to int64.

After filling up the missing values in the total charges column, its Type() was converted to float64 data type. The new meta data for the updated data set after stage 3 is given in updated column of Table 2.

Table 2. Initial and updated data set design

S. No	Column	Not null count	Initial type ()	Updated type ()
1	customerID	7043	Object	Object
2	Gender	7043	Object	Object
3	SeniorCitizen	7043	Int64	Int64
4	Partner	7043	Object	Object
5	Dependents	7043	Object	Object
6	Tenure	7043	Object	Int64
7	PhoneService	7043	Object	Object
8	MultipleLines	7043	Object	Object
9	InternetService	7043	Object	Object
10	OnlineSecurity	7043	Object	Object
11	OnlineBackup	7043	Object	Object
12	DeviceProtection	7043	Object	Object
13	TechSupport	7043	Object	Object
14	StreamingTV	7043	Object	Object
15	StreamingMovies	7043	Object	Object
16	Contract	7043	Object	Object
17	PaperlessBilling	7043	Object	Object
18	PaymentMethod	7043	Object	Object
19	MonthlyCharges	7043	Float64	Float64
20	TotalCharges	7043	Object	Float64
21	Churn	7043	Object	Int64

3.4. Formatting and one hot encoding

After the data has been cleaned, the data needed to be brought into a format that was acceptable by the XGB classifier. For this purpose, the data went through the following transformations: removal of white spaces in the data: white spaces are removed as classification in XGB classifier requires continuous labels. Then the data is splitted into dependant and independent variable Y and X respectively. The churn column is taken as the dependant variable Y and the entire data set other than the churn column is taken as independent variable X.

One hot encoding is a process where for making decision trees it is essential to classify categorical variables into 0 and 1 combinations. This means if for a column gender, there are two values male or female, after one hot encoding male and female values will become a column each themselves and if in a new record the value of gender column is male then male column will have value 1 and female column will have value 0. After the splitting of gender column into male and female columns, the gender column gets removed from the data set. Creation of these new columns does not take extra space as XGBoost uses sparse matrices so it doesn't allocate memory to zeros. The data set before and after one hot encoding is shown in Tables 3 and 4.

Table 3. Before one hot encoding

S.No	Customer Id	Male
1	7515	1
2	5523	0
3	3924	0
4	9237	1
5	4657	0

Table 4. After one hot encoding

S.No	Customer Id	Male	Female
1	7515	1	0
2	5523	0	1
3	3924	0	1
4	9237	1	0
5	4657	0	1

3.5. Building preliminary model

Now that the data is formatted, the model can be built by feeding the data into the classifier. This involves splitting the data into training and testing data. Training data is a part of the data set on which the model is built and testing data is a part of the data set on which the model built is tested for accuracy. Before splitting the data, it is essential to maintain the balance in ratio of churn in the entire data set with both ratio of churn in both training and testing data set. After calculating it was found that 27 random state=42. After splitting the data, the model is built in the iterations as,

− Iteration 0: validation_0-aucpr: 0.579067,
− Iteration 1: validation_0 − aucpr: 0.63937,
− Iteration 2- validation_0 − aucpr: 0.63839,
− Till iteration 50: validation_0-aucpr: 0.652923.

The best value is got at iteration 40: validation0−aucpr: 0.654216, XGBClassifier (seed=42). The model was built after gradient boosting of 50 trees and the early stopping rounds was set to 10. This implied that after building 10 more trees without any better aucpr metric (used for evaluation) the process would stop and the (n-10)th iteration is best iteration and in this case: 40th iteration.

3.6. Confusion matrix

Confusion matrix is an essential for understanding the performance of a machine learning model. It is defined as a performance measurement model to understand how well a machine learning model that was built is working. For our model we are aiming at a target: accuracy of 80% in identifying churn (customer who left the company) and the Table 5 shows the confusion matrix for the reading mentioned in Table 6.

Table 5. Confusion matrix for preliminary model

Label	Predicted Did not Leave	Predicted Left
True	1186	108
Did not Leave		
TrueLeft	242	225

Table 6. Statistics of confusion matrix

Label	Total	Predicted	Accuracy
Did not Leave	1294	1186	91.65
Left	467	225	48.1

3.7. Optimizing parameters with cross validation (grid search)

The accuracy for customers not leaving the company was found to be 91.65%. The accuracy of the prediction of people who actually leave must be improved and find the cause only for the same. Then only the company can stop them from leaving. So, in order to achieve this, the optimization and cross validation are done. XGBoost has a lot of hyper parameters which needs to be tweaked in order to set the direction of the processing which yields better accuracy for people who have left the company. Some of them are gamma, max depth, reg lambda, scale post weight, and GridSearchCV has been used in which data is sub sampled by 90% of the data and only 50% of the columns are used for each tree built. This is helps in better cross validation. This is achieved in two rounds of hit and trial which is shown in Table 7.

After building the model with these values it was noticed that the accuracy was going even lower. So the values were increased in opposite direction and the updated values were arrived as given in Table 8. For the updated values of the hyper parameters given in Table 8, an updated final confusion matrix is shown in Table 9. Therefore, it can be observed from Table 10 that the desired accuracy of > 80% has been achieved by tweaking the hyper parameters for the values of hyper parameters in the Table 8.

Table 7. Hyper parameters after two rounds

Round	Gamma	Learning Rate	Max depth	Reg Lambda	Scale pos weight
1	1	0.05	3	0	1
2	0.1	0.1	3	0	0.5

Table 8. Hyper parameters after final round

Round	Gamma	Learning Rate	Max depth	Reg Lambda	Scale pos weight
N	0,25	0.1	4	10	3

Table 9. Final confusion matrix

Label	Predicted Did not Leave	Predicted Left
True Did not Leave	934	360
True Left	84	383

Table 10. Final statistics from final confusion matrix

Label	Total	Predicted	Accuracy
Did not Leave	1294	934	72.17
Left	467	383	82.1

4. RESULTS AND DISCUSSION

The customer churn analysis is one of the important challenging areas in research. It has its many applications in banking sectors, super marks, telecommunications and other customer related applications. In this paper this is implemented using supervised machine learning algorithm using Python on a given data-set of Telco, a mobile telecommunication company. The implementation shows that using XGBoost, it gives comparatively more accurate prediction than other learning models. The Figure 3 gives comparison of accuracy prediction in different learning models. It can be analyzed from the graph that the prediction of accuracy on customer churn analysis is more in XGBoost learning model and so by using this model, reasons for customer leaving the company can be analyzed and based on that proper solution can be achieved.
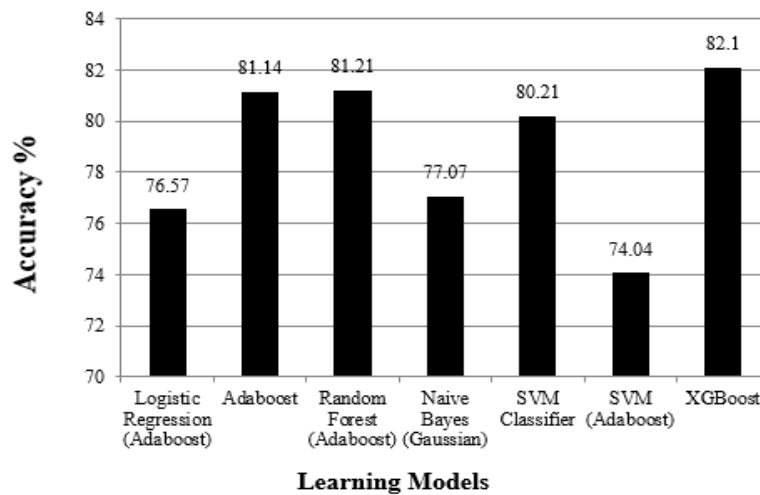
Figure 3. Comparative analysis of accuracy % in different learning models

5. CONCLUSION

Telecommunication industry usually suffers from high rates of customer churn. Although the business loss is unavoidable, but still churn can be managed and kept in an acceptable level. Good methods need to be developed and existing methods have to be enhanced to prevent the telecommunication industry to face challenges. Customer churn prediction becomes a very difficult task for many startups and upcoming companies and so it is very tough to predict the genuine customers of these companies. Therefore, more latest learning models in machine learning and deep learning techniques using assembling models can be used for such predictions with accurate results.

The future enhancements that can be performed in this model involves improving accuracy. Through more rounds of cross validation and working with real time data software like Apache Spark to enhance the model to perform real time customer churn prediction. The user interface (UI) aspect of the application can also be improved from the aspect of making it clearer for business stakeholders.

REFERENCES

[1]   X. Zhao, Y. Shi, J. Lee, H. K. Kim, and H. Lee, "Customer churn prediction based on feature clustering and nonparallel support vector machine," *International Journal of Information Technology & Decision Making*, vol. 13, no. 05, pp. 1013-1027, 2014, doi: 10.1142/S0219622014500680.

[2]   Y. Xu, "Predicting customer churn with extended one-class support vector machine," in *Natural Computation (ICNC), Eighth International Conference on IEEE*, 2012, pp. 97-100, doi: 10.1109/ICNC.2012.6234646.

[3]   T. Vafeiadis, K. I. Diamantaras, G. Sarigiannidis, and K. Ch. Chatzisavvas, "A comparison of machine learning techniques for customer churn prediction," *Simulation Modelling Practice and Theory*, vol. 55, pp. 1-9, June 2015, doi: 10.1016/j.simpat.2015.03.003.

[4]   J. Burez and D. V. D. Poel, "Handling class imbalance in customer churn prediction," *Expert Systems with Applications*, vol. 36, no. 3, pp. 4626-4636, 2009, doi: 10.1.1.477.1151.

[5]   K. W. D. Bock and D. V. D. Poel, "Reconciling performance and interpretability in customer churn prediction using ensemble learning based on generalized additive models," *Expert Systems with Applications*, vol. 39, no. 8, pp. 6816-6826, June 2012, doi: 10.1016/j.eswa.2012.01.014.

[6]   R. Obiedat, M. Alkasassbeh, H. Faris, and O. Harfoushi, "Customer churn prediction using a hybrid genetic programming approach," *Scientific Research and Essays*, vol. 8, no. 27, pp. 1289-1295, Jan 2013, doi:10.5897/SRE2013.5559.

[7]   J. H. Ahn, S. P Han, and Y. S. Lee, "Customer churn analysis: Churn determinants and mediation effects of partial defection in the Korean mobile telecommunications service industry," *Telecommunications Policy 30*, pp. 552–568, 2006, doi: 10.1016/j.telpol.2006.09.006.

[8]   K. Dahiya and S. Bhatia, "Customer churn analysis in telecom industry," *2015 4th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO) (Trends and Future Directions)*, 2015, pp. 1-6, doi: 10.1109/ICRITO.2015.7359318.

[9]   B. Huang, M. T. Kechadi, and B. Buckley, "Customer churn prediction in telecommunications," *Expert Systems with Applications*, vol. 39, no. 1, pp. 1414-1425, 2012, doi: 10.1016/j.eswa.2011.08.024.

[10] A. Keramati, R. Jafari-Marandi, M. Aliannejadi, I. Ahmadian, M. Mozaffari, and U. Abbasi, "Improved churn prediction in telecommunication industry using data mining techniques," *Applied Soft Computing*, vol. 24, pp. 994-1012, 2014, doi: 10.1016/j.asoc.2014.08.041.

[11] G. Li and X. Deng, "Customer churn prediction of china telecom based on cluster analysis and decision tree algorithm," in *Emerging research in artificial intelligence and computational intelligence,Springer Berlin Heidelberg*, vol. 315, pp. 319-327, 2012, doi: 10.1007/978-3-642-34240-0_42.

[12] N. Lu, H. Lin, J. Lu, and G. Zhang, "A customer churn prediction model in telecom industry using boosting," *IEEE Transactions onIndustrial Informatics*, vol. 10, no. 2, pp. 1659-1665, 2014, doi: 10.1109/TII.2012.2224355.

[13] O. Adwan, H. Faris, K. Jaradat, O. Harfoushi, and N. Ghatasheh, "Predicting customer churn in telecom industry using multilayer preceptron neural networks: Modeling and analysis," *Life Science Journal,* vol. 11. no. 3, pp. 75-81, 2014.

[14] D.V.D. Poel and B. Lariviere, "Customer attrition analysis for financial services using proportional hazard models," *European Journal of Operational Research*, vol. 157, no. 1, Aug. 2004, doi: org/10.1016/S0377-2217(03)00069-9.

[15] A. Amin *et al*., "Customer churn prediction in the telecommunication sector using a rough set approach," *Neurocomputing*, vol. 237, pp. 242–254, May 2017, doi: org/10.1016/j.neucom.2016.12.009.

[16] F. Buttle, *Customer Relationship Management Book*, 2nd edition, New York, USA: Taylor & Francis, 2008.

[17] M. A. H. Farquad, V. Ravi ,and S. B. Raju, "Churn prediction using comprehensible support vector machine: An analytical CRM application," *Applied Soft Computing*, vol. 19, pp. 31- 40, June 2014, doi: 10.1016/j.asoc.2014.01.031.

[18] M. R. Ismail, M. K. Awang, M. N. A. Rahman, and M. Makhtar, "A Multi-Layer Perceptron Approach for Customer Churn Prediction," *International Journal of Multimedia and Ubiquitous Engineering*, vol. 10, no. 7, pp. 213-222, 2015, doi: org/10.14257/ijmue.2015.10.7.22.

[19] D. Bhukya and S. Ramachandram, "Decision Tree Induction: An Approach for Data Classification Using AVL-Tree," *International Journal of Computer and Electrical Engineering*, vol. 2, no. 4, pp. 1793-8163, 2010, doi: 10.7763/IJCEE.2010.V2.208.

[20] U. D. Prasad and S. Madhavi, "Prediction of churn behavior of bank customers using data mining tools," *Business Intelligence Journal*, vol. 5, no. 1 pp. 96-101, 2012.

[21] C. C. Günther, I. F. Tvete, K. Aas, G. I. Sandnes, and O. Borgan, "Modeling and predicting customer churn from an insurance company," *Scandinavion Actuarial Journal*, vol. 1, pp. 58-71, 2014, doi: 10.1080/03461238.2011.636502.

[22] S. KhakAbi, M. R. Gholamian, and M. Namvar, "Data Mining Applications in Customer Churn Management," *2010 International Conference on Intelligent Systems, Modelling and Simulation*, 2010, pp. 220-225, doi: 10.1109/ISMS.2010.49.

[23] R. A. Soeini and K. V. Rodpysh, "Applying Data Mining to Insurance Customer Churn Management," *International Proceedings of Computer Science and Information Technology*, vol. 30, pp. 82-92, 2012.

[24] C. F. Tsai and Y. H. Lu, "Data Mining Techniques in Customer Churn Prediction," *Recent Patents on Computer Science*, vol. 3, no. 1, 2009, doi: 10.2174/2213275911003010028.

[25] P. Zerfos, J. Cho, and A. Ntoulas, "Downloading textual hidden web content through keyword queries," *Proceedings of the 5th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL '05)*, 2005, pp. 100-109, doi: 10.1145/1065385.1065407.

BIOGRAPHIES OF AUTHORS

Dr. Muthupriya Vaudevan 🆔 🔍 SC Ⓟ received her B.E. degree in Computer Science Engineering (CSE) from Madras University, India in 1999 and her M.E (CSE) from Madras University, India in 2003. She completed her Ph.D., in Crescent University Chennai. She is currently working as an Assistant Professor in the department of CSE, Crescent University Chennai. She has 21 years of teaching experience and her areas of interest are Wireless Mobile Ad hoc networks, Cryptography and Network security, Machine learning and IoT She is a life member of Indian Society for Technical Education (ISTE), the System Society of India. She can be contacted at email: muthupriya@crescent.education.

Dr. Revathi Sathya Narayanan 🆔 🔗 SC Ⓟ received her B.E. degree in Computer Science Engineering (CSE) from Bharathidasan University, India in 1994 and her M.E (CSE) from Madurai Kamarajar University, India in 2000. She completed her Ph.D., in Anna University Chennai in 2014. She is currently working as a professor in the department of CSE, B.S. Abdur Rahman Crescent Institute of Science and Technology, Chennai. She has 26 years of teaching experience and her areas of interest are Wireless Mobile Ad hoc networks, Cryptography and Network security and IoT. She published more than 50 papers in National and International conferences and journals. She is a life member of Indian Society for Technical Education (ISTE), CSI and IAENG. She can be contacted at email: srevathi@crescent.education.

Dr. Sabiyath Fatima Nakeeb 🆔 🔗 SC Ⓟ Associate Professor, Department of Computer Science and Engineering, B.S. Abdur Rahman Crescent Institute of Science & Technology, Chennai. She has professional experience of more than 18 years working in research and teaching. She has published book chapters and more than 30 papers in various National and International peer reviewed journals (IEEE and Springer) and conferences. Acted as resource person, panel member, chief guest, guest of honor and given plenary talk in various industries and institutions as a part of training, seminars, workshops, international and national conferences. She has been active reviewer in various International Journals and Conferences. Her teaching and research expertise covers a wide range of subject area including Mobile Ad Hoc Networks, Data mining, High Performance Computing, IoT, Big data, and Machine learning. She can be contacted at email: sabiyathfathima@crescent.education.

Abhishek 🆔 🔗 SC Ⓟ was born on 18th October 1997 in New Delhi, India. He has received his Bachelor of Computer Application degree in the year 2019 from Maharaja Surajmal Institute affiliated to Guru Gobind Singh Indraprastha University, New Delhi. He has completed his Master of Computer Application degree in B.S. Abdur Rahman University, Chennai, India. His areas of research interest are Machine learning and Data Mining. He can be contacted at email: abhi.official97@gmail.com.