ISSN: 0974-8571 Vol.13 No. 2 December, 2021

International Journal of Computational Intelligence in Control

Sentiment Analysis for Evaluating the Patient Medicine Satisfaction

Sabarmathi . G[1]
Research Scholar, VTU-RC, Department of MCA, CMR Institute of Technology, Bengaluru-560037
Assistant Professor
Computer Science and Applications, Christ Academy Institute for Advanced Studies, Bangalore
sabarganesh@gmail.com

Dr.R.Chinnaiyan[2]
Professor & Head
Department of MCA, AMC Engineering College, Bengaluru-560083
vijayachinns@gmail.com, chinnaiyan.r@amceducation.in

Date of Submission: 21st October 2021 Revised: 11th November 2021 Accepted: 30th November 2021

How to Cite: **Sabarmathi . G and R.Chinnaiyan (2021). Sentiment Analysis for Evaluating the Patient Medicine Satisfaction. International Journal of Computational Intelligence in Control 13(2), pp.113-118.**

Abstract: **Recently, it has been shown that there is a lot of online reviews being considered resulting in many people taking their medication without consulting health professionals, as telemedicine is also involved in taking their own medication based on the online source. Today, we can see that machine learning played an important role in the development of the recommendation system. The aim of this article is to present a medication review recommendation that will help the patient identify medication based on the indicated condition. This patient rating is used to predict medication recommendations based on the stated condition. Ratings and reviews are used to predict the patient's feelings/opinions based on various processes such as corpus, stemmer, text blob used for predicting the models and recommending the drugs is done based on the sentiment polarity of the reviews based on the condition specified.**

Keywords: drug, recommendation, corpus, cosine similarity, stemming, sentiment analysis

1. Introduction:

Recently, there has been a shortage of health workers in countries during this time of the pandemic. Unskilled workers. This has resulted in tele consultations and telemedicine services, online consultations have increased rapidly today, it is found that about a hundred thousand people are affected by drug prescription every year in the US due to the restrictions on drug prescribing since then by health workers agrees with the result on the basis of the knowledge acquired [1,2]. A new study arrives every day with more drugs, tests, and access to clinical staff.. With online business sector, product evaluation has become a mandatory and integrated factor for buyers all over the world. Analyze online reviews of certain articles and decide based on this rating.

Online opinion plays an important role in purchasing the product. This is true of the healthcare sector, with patients reviewing online comments on prescription drugs, their side effects, etc., as many are now more concerned about their health and wellbeing. This has led to the online diagnosis of their health for themselves. A recent survey found that more than 70% of adults search for health-related topics online, and around 40% of consumers look for self-diagnostic methods online. Recommendations based on your self-diagnosis. The drug recommendation framework is very important, as it can help experts and patients expand their knowledge of drugs for specific health conditions. The proposed research is a common system that recommends drugs to users based on their benefits. These frameworks used customer feedback to break down their feelings and provide recommendations for the exact needs of the patients. In the drug recommendation system, sentiment analysis and function engineering[14-41] are used to provide drugs for specific diseases based on patient ratings. The proposed work is divided into segments as an introduction giving a brief idea of the need for this research work, literature review gives an impression of previous views in this area, section 3 talks about the methodology proposed for this research work, section 4 talks about the results acquired by different models used for this work, the discussion section discusses the limitations of this study and finally gives the conclusion of this research work.

2. Proposed Methodology

The proposed methodology is depicted diagrammatically in figure 1. The steps involved in the proposed system is as follows, the data set for drug review is taken from the UCI Machine Learning Repository. This dataset provides patient ratings for specific medications along with associated medical conditions and a 10-star rating that reflects overall patient satisfaction.

Features such as the patient rating of the drug, patient condition, name of the drug, date of use, and a number of useful users, which indicate the number of users who found the review useful. This dataset contains several types of properties that are categorical, numeric, text, and date. Sentiment analysis is the main goal where we have to classify a review as positive or negative. The train and the test suite are first merged because the main goal is sentiment analysis. The opinion of the review is given according to the rating, as the opinion of the reviews is not initially included in the data set. Feature extraction is done to get a good perspective of the reviews to make good predictions. The pre-processing serves to prepare the data for further processing. The data set is then divided into training and test data sets. 70% of the data is used for training and the remaining 30% of the data is used for testing.

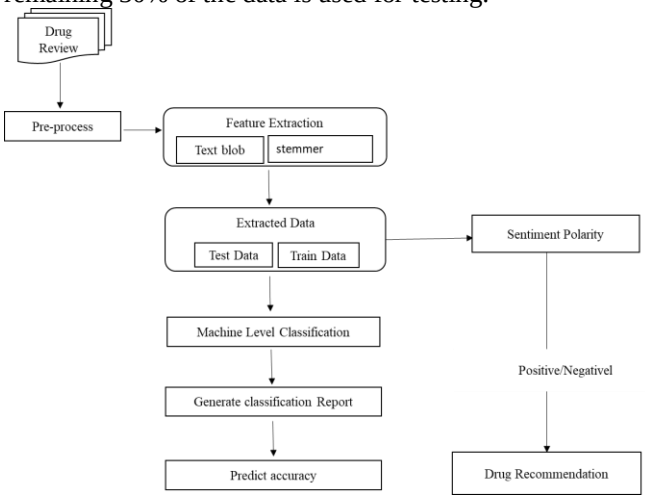

Figure 1. Proposed Methodology

Different machine learning classification algorithms[14-25] were used to create a mood predictor classifier. Therefore, when building the recommendation system, we normalize the useful number through conditions.

Table 1 and figure 2 show the total number of 10-star ratings for the drug. Totally there were 215063 reviews after data pre-processing it has become 213869.

1	2	3	4	5	6	7	8	9	10
28769	9203	8662	6622	10650	8403	12470	24909	36499	67682

Table 1: User Ratings(1-10scale)

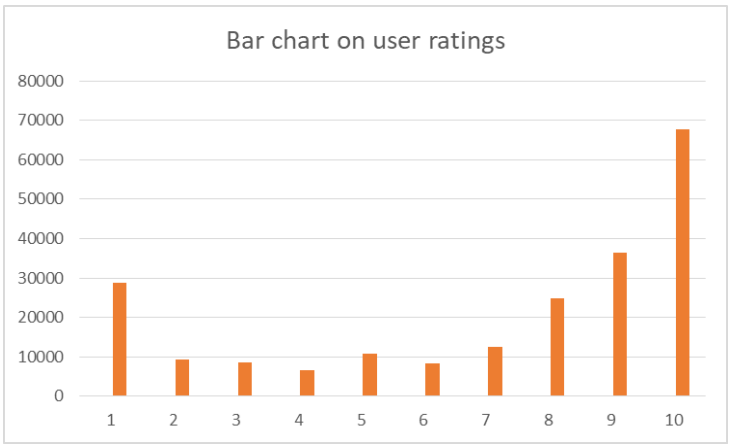

Figure 2: Bar chart on User ratings (1-10)

The ratings are converted into positive and negative sentiment by using the review polarity and it is converted using the condition >=5 as positive and <5 rating as negative. Table 2 and Figure 3 shows the total positive and negative ratings.

User Ratings	
Positive	161491
Negative	53572

Table 2: User ratings after repalacement

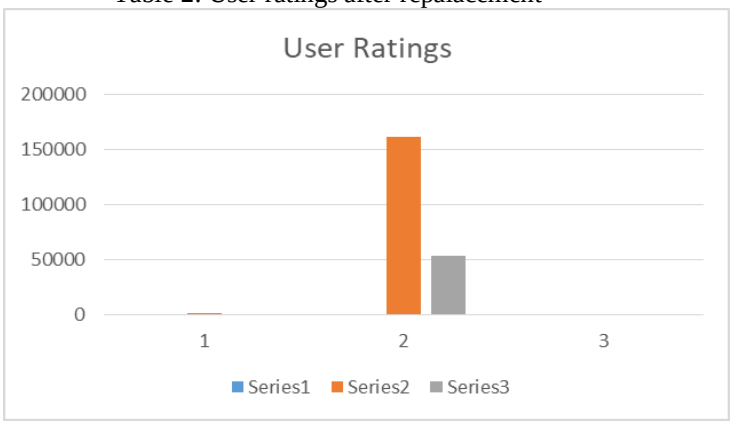

Figure 3: chart on User ratings after replacement (1/0)

Figure 4 shows the top 20 conditions that have a maximum number of drugs review in the data set.

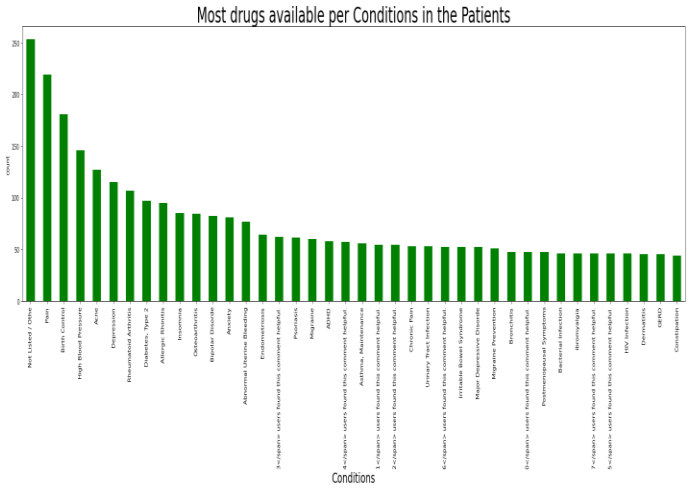

Figure 4: Popular drugs available based on the condition

4. Results:

The classification report for the LGBM model is given below (Figure 10.2). It can be seen that the accuracy for predicting negative mood is 0.85 and for positive mood, it is 0.91. The biggest difference is between their recall value, i.e. 0.73 for negative sentiment and 0.96 for positive sentiment. The accuracy of the LGBM is 0.91249.

```
LGBoost Accuracy
The Accuracy of the model is :
0.9124977946104331
The confusion Matrix is
              precision    recall  f1-
score   support

         0.0       0.85      0.73
0.79     16074
         1.0       0.91      0.96
0.94     48087

    accuracy
0.90     64161
   macro avg       0.88      0.84
0.86     64161
weighted avg       0.90      0.90
0.90     64161
```

Table 4: Classification report on LGBoost

The classification report for the XGBoost model is given below (Figure 10.2). It can be seen that the accuracy for predicting negative opinion is 0.68 and for positive opinion, it is 0.81. The biggest difference is between their recall value, i.e. 0.34 for negative sentiment and 0.95 for positive sentiment. The accuracy of the LGBM is 0.79521

XGB Accuracy

The Accuracy of the model is : 0.7952182790168483

The confusion Matrix is

	precision	recall	f1-score	support
0.0	0.68	0.34	0.46	16074
1.0	0.81	0.95	0.87	48087
accuracy			0.80	64161
macro avg	0.75	0.64	0.67	64161
weighted avg	0.78	0.80	0.77	64161

Table 5: Classification report on XGBoost

Catboost is an algorithm for increasing the gradient in decision trees developed by Yandex researchers and engineers and used for search, recommendation systems, personal assistants, self-driving cars, weather forecast, and many other tasks in Yandex and other businesses. Including CERN., Cloud flare, careem taxi[10]. It's open-source. The Catboost parameters are

CatBoostClassifier (iterations = 10000, learning rate = 0.5)

The classification report for the CatBoost model is given below (Figure 10.2). It can be seen that the accuracy for predicting negative opinion is 0.74 and for positive opinion, it is 0.82. The biggest difference is between their recall value, i.e. 0.74 for negative sentiment and 0.95 for positive sentiment. The accuracy of the CatBoost is 0.79521

CatBoost Accuracy

The Accuracy of the model is : 0.8942971587101198

The confusion Matrix is

	precision	recall	f1-score	support
0.0	0.82	0.74	0.78	16074
1.0	0.92	0.95	0.93	48087
accuracy			0.89	64161
macro avg	0.87	0.84	0.85	64161
weighted avg	0.89	0.89	0.89	64161

Table 6: Classification report on CatBoost

The performance of the models are evaluated using 10-flod cross validation and the score are validated using ROC-AUC curve as shown the figure 5 and figure 6

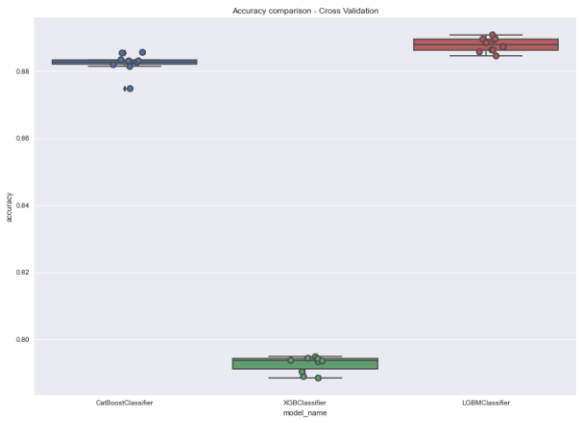

Figure 5: Accuracy Comparison using cross validation

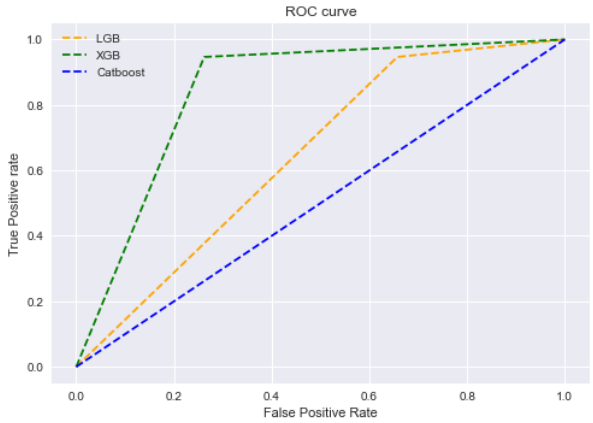

Figure 6: comparison using ROC-AUC curve

Figure 7 show the top 10 drugs based on the specified condition. The recommendation of the drug is calculated using the review sentiment polarity score for each condition divided by total number useful counts that is number of user gave positive review.

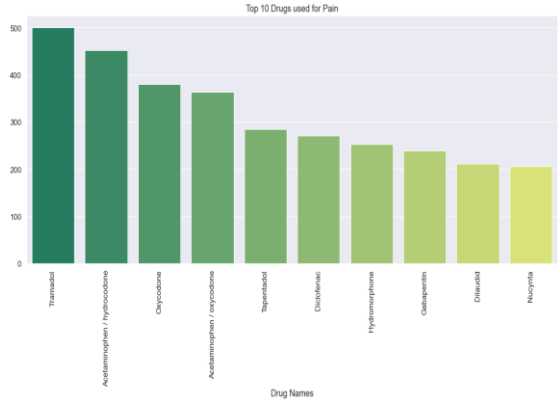

Figure 7: Top 10 drugs recommendation based on condition

3. Discussions:

The results obtained from each classification report shows the performance of the models based on the review sentiments, yet it does not show any recommendation framework for real time applications. It still needs improvements for further insight on data. In the recommendation system we simply taken the polarity score of the review and number of users counts and we tried to propose on recommendation based on the condition. This paper tries to propose one methodology for recommending a drug based on the review sentiment and classification metrics of the model for this data type.

4 Conclusions

The main goal was to predict the opinion of the drug reviews given by the patients. Therefore, feature extraction was performed to obtain more information about the data set, and pre-processing was performed to prepare the data for modelling and data analysis. Therefore, feature extraction was performed on the basis of data pre-processing and patient ratings. The revisions were cleaned up and the characteristics generated. They were trained to be LightGBM, XGBoost, and CatBoost. The best performing model is the LGBM classifier, but its accuracy and classification report are comparable to the other classifier models. The accuracies were 0.91,0.79 and 0.89 respectively. The classification report as well as the cross-validation of the models also serve for a more in-depth analysis of the model, since the accuracies alone do not say much about a classification model. The results are validated using the ROC-AUC curve as shown in the previous figures.

5. References:

1. Ali, Zafar & Khusro, Shah & Ullah, Irfan. (2016). A Hybrid Book Recommender System Based on Table of Contents (ToC) and Association Rule Mining.
2. Doulaverakis, C., Nikolaidis, G., Kleontas, A. et al. GalenOWL: Ontology-based drug recommendations discovery.J Biomed Semant 3, 14 (2012).
3. Garg, Satvik. (2021). Drug Recommendation System based on Sentiment Analysis of Drug Reviews using Machine Learning. 175-181..
4. J. Li, H. Xu, X. He, J. Deng and X. Sun," Tweet modeling with LSTM recurrent neural networks for hashtag recommendation," 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, 2016, pp. 1570-1577.
5. Kurmashov, Nursultan & Latuta, Konstantin & Nussipbekov, Abay. (2015). Online book recommendation system. 1-4.
6. Leilei Sun, Chuanren Liu, Chonghui Guo, Hui Xiong, and Yanming Xie. 2016. Data-driven Automatic Treatment Regimen Development and Recommendation. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16). Association for Computing Machinery, New York, NY, USA, 1865–1874.
7. Powers, David & Ailab,. (2011). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness & correlation. J. Mach. Learn. Technol. 2. 2229-3981.

8. Shimada K, Takada H, Mitsuyama S, et al. Drug-recommendation system for patients with infectious diseases. AMIA Annu Symp Proc.2005.
9. Sumit Mishra"Drug Review Sentiment Analysis usingBoosting Algorithms"Published inInternational Journalof Trend in ScientificResearch andDevelopment (ijtsrd),ISSN: 2456-6470,Volume-5 Issue-4,June 2021, pp.937-941.
10. T. N. Tekade and M. Emmanuel," Probabilistic aspect mining approach for interpretation and evaluation of drug reviews," 2016 International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES), Paralakhemundi, 2016, pp. 1471-1476.
11. Wittich CM, Burkle CM, Lanier WL. Medication errors: an overview for clinicians. Mayo Clin Proc. 2014 Aug;89(8):1116-25.
12. Y. Bao and X. Jiang, "An intelligent medicine recommender systemframework," 2016 IEEE 11th Conference on Industrial Electronics and Applications (ICIEA), Hefei, 2016, pp. 1383-1388.
13. Zhang, Yin & Zhang, Dafang & Hassan, Mohammad & Alamri, Atif & Peng, Limei. (2014). CADRE: Cloud-Assisted Drug Recommendation Service for Online Pharmacies. Mobile Networks and Applications. 20. 348-355.
14. Balachandar S., Chinnaiyan R. (2019) Centralized Reliability and Security Management of Data in Internet of Things (IoT) with Rule Builder. In: Smys S., Bestak R., Chen JZ., Kotuliak I. (eds) International Conference on Computer Networks and Communication Technologies. Lecture Notes on Data Engineering and Communications Technologies, vol 15. Springer, Singapore
15. Balachandar S., Chinnaiyan R. (2019) Reliable Digital Twin for Connected Footballer. In: Smys S., Bestak R., Chen JZ., Kotuliak I. (eds) International Conference on Computer Networks and Communication Technologies. Lecture Notes on Data Engineering and Communications Technologies, vol 15. Springer, Singapore
16. Balachandar S., Chinnaiyan R. (2020) ," Reliable Administration Framework of Drones and IoT Sensors in Agriculture Farmstead using Blockchain and Smart Contracts"BDET 2020: Proceedings of the 2020 2nd International Conference on Big Data Engineering and Technology" January 2020 Pages 106–111
17. Balachandar S., Chinnaiyan R. (2020) Internet of Things Based Reliable Real-Time Disease Monitoring of Poultry Farming Imagery Analytics. In: Pandian A., Senjyu T., Islam S., Wang H. (eds) Proceeding of the International Conference on Computer Networks, Big Data and IoT (ICCBI - 2018). ICCBI 2018. Lecture Notes on Data Engineering and Communications Technologies, vol 31. Springer, Cham
18. Divya R., Chinnaiyan R. (2019) Reliable AI-Based Smart Sensors for Managing Irrigation Resources in Agriculture—A Review. In: Smys S., Bestak R., Chen JZ., Kotuliak I. (eds) International Conference on Computer Networks and Communication Technologies. Lecture Notes on Data Engineering and Communications Technologies, vol 15. Springer, Singapore
19. Divya. R and R. Chinnaiyan(2020), "Reliable Constrained Application Protocol to Sense and Avoid attacks in WSN for IoT Devices," 2019 International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2019, pp. 1898-1901.
20. G. Sabarmathi and R. Chinnaiyan, "Big Data Analytics Framework for Opinion Mining of Patient Health Care Experience," *2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC)*, Erode, India, 2020, pp. 352-357
21. G. Sabarmathi and R. Chinnaiyan, "Investigations on big data features research challenges and applications," *2017 International Conference on Intelligent Computing and Control Systems (ICICCS)*, Madurai, 2017, pp. 782-786
22. G. Sabarmathi and R. Chinnaiyan, "Reliable Machine Learning Approach to Predict Patient Satisfaction for Optimal Decision Making and Quality Health Care," *2019 International Conference on Communication and Electronics Systems (ICCES)*, Coimbatore, India, 2019, pp. 1489-1493
23. Hari Pranav A; M. Latha; Ashwin. M. S; R. Chinnaiyan**,** "BlockchainAs a Service (BaaS) Framework for Government Funded Projects e-Tendering Process Administration and Quality Assurance using Smart Contracts," *2021 International Conference on Computer Communication and Informatics (ICCCI)*, 2021, pp. 1-4, doi: 10.1109/ICCCI50826.2021.9402348.
24. Hari Pranav A;M. Senthilmurugan;Pradyumna Rahul K;R. Chinnaiyan , "IoT and Machine Learning based Peer to Peer Platform for Crop Growth and Disease Monitoring System using Blockchain," *2021 International Conference on Computer Communication and Informatics (ICCCI)*, 2021, pp. 1-5..
25. M. Swarnamugi and R. Chinnaiyan(2020), "Smart and Reliable Transportation System based on Message Queuing Telemetry Transport Protocol," 2019 International Conference on Intelligent Computing and Control Systems (ICCS), Madurai, India, 2019, pp. 918-922.

26. M. Swarnamugi and R. Chinnaiyan, "IoT Hybrid Computing Model for Intelligent Transportation System (ITS)," 2018 Second International Conference on Computing Methodologies and Communication (ICCMC), Erode, 2018, pp. 802-806.
27. Preetika B, ; M. Latha; M. Senthilmurugan; R. Chinnaiyan, , "MRI Image based Brain Tumour Segmentation using Machine Learning Classifiers," *2021 International Conference on Computer Communication and Informatics (ICCCI)*, 2021, pp. 1-9, doi: 10.1109/ICCCI50826.2021.9402508
28. Preetika B; R. Chinnaiyan**,** "An Open Source Distributed Virtualization for Enterprise Infrastructure using OVIRT," *2021 International Conference on Computer Communication and Informatics (ICCCI)*, 2021, pp. 1-5, doi: 10.1109/ICCCI50826.2021.9402547.
29. R Chinnaiyan, S Somasundaram, "Monte Carlo Simulation For Reliability Assessment of Component Based Software Systems", i-Manager's Journal on Software Engineering, 2010.
30. R Chinnaiyan, S Somasundaram, "Reliability Estimation Of Component Based Software Systems Through Markov Process", International Journal of Mathematics, Computer Sciences and Information Technology.
31. R. Chinnaiyan and S. Alex, "Machine Learning Approaches for Early Diagnosis and Prediction of Fetal Abnormalities," *2021 International Conference on Computer Communication and Informatics (ICCCI)*, 2021, pp. 1-3, doi: 10.1109/ICCCI50826.2021.9402317.
32. R.Chinnaiyan, Dr.S.Somasundaram, "Reliability of Component Based Software With Similar Software Components - A Review", i-Manager's Journal on Software Engineering; Nagercoil5.2 (Oct-Dec 2010).
33. R.Chinnaiyan, S.Somasundaram(2009) "Reliability of Object Oriented Software Systems using communication variables – a review " ,International Journal of Software Engineering, Vol.2, No.2,PP.87-96.
34. R.Chinnaiyan, S.Somasundaram(2010) "Evaluating the Reliability of Component Based Software Systems " ,International Journal of Quality and Reliability Management , Vol. 27, No. 1., pp. 78-88
35. R.Chinnaiyan, S.Somasundaram(2011), "An Experimental Study on Reliability Estimation of GNU Compiler Components - A Review", International Journal of Computer Applications,Vol.25, No.3, July 2011, pp.13-16.
36. S. Balachandar , Dr.R.Chinnaiyan , A Reliable Troubleshooting Model for IoT Devices with Sensors and Voice Based Chatbot Application, International Journal for Research in Applied Science & Engineering Technology (IJRASET), Volume 6 Issue II, February 2018, PP.1406-1409 ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887.
37. S.Balachandar, Dr. R.Chinnaiyan, A Reliable Agnostic Data Compression Model Between Edge and Cloud IOT Platform, International Journal of Computer Engineering and Applications, Volume XII, Issue I, Jan. 18,PP.139-143. ISSN 2321-3469
38. Sabarmathi G., Chinnaiyan R. (2020) Envisagation and Analysis of Mosquito Borne Fevers: A Health Monitoring System by Envisagative Computing Using Big Data Analytics. In: Pandian A., Senjyu T., Islam S., Wang H. (eds) Proceeding of the International Conference on Computer Networks, Big Data and IoT (ICCBI - 2018). ICCBI 2018. Lecture Notes on Data Engineering and Communications Technologies, vol 31. Springer, Cham.
39. Tilak Shantaram Nayak, Vykuntam Harini, Ganesh Hegde, Harshith Y N, R.Chinnaiyan (2019)," IOT based Leaf Disease Detection and Fertilizer Recommendation", IJITEE , Vol.9., Issue-2, pp.132-136, December 2019
40. Vardhini B;Shreyas N Dass;Sahana R;R. Chinnaiyan**,** "A Blockchain based Electronic Medical Health Records Framework using Smart Contracts," 2021 International Conference on Computer Communication and Informatics (ICCCI), 2021, pp. 1-4, doi: 10.1109/ICCCI50826.2021.9402689.
41. Yamini U., Taha Sufiyan, Thankam Paul, Suyash Gupta, R. Chinnaiyan (2020),"Smart Crop Monitor using Internet of Things, Cloud, Machine Learning and Android App" , IJEAT, Vol.9, Issue-3 , pp.3213-3217,February 2020

.