EasyChair Preprint

№ 8173

Biometrics Based Secured Online Voting System Using Machine Learning Method

Tazaeen Shaikh, Hritika Ranadhir, Suyash Gugale, Vrushali Patil and Omkaresh Kulkarni

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

June 1, 2022

Biometrics Based Secured Online Voting System Using Machine Learning Method

Tazaeen Ilyas Shaikh
School of Computer Science and Technology
MITWPU
Pune, India
tazaeen.shaikh@gmail.com

Hritika Kamalakar Ranadhir
School of Computer Science and Technology
MITWPU
Pune, India
hritikaranadhir@gmail.com

Suyash Santosh Gugale
School of Computer Science and Technology
MITWPU
Pune, India
suyashsg.21@gmail.com

Vrushali Prakash Patil
School of Computer Science and Technology
MITWPU
Pune, India
vrushalipatil612@gmail.com

Omkaresh Kulkarni
School of Computer Science and Technology
MITWPU
Pune, India
omkaresh.kulkarni@gmail.com

Abstract— **Electronic voting or e-voting has been used in varying forms since 1970s with fundamental benefits over paper-based systems such as increased efficiency and reduced errors. However, there remain challenges to achieve wide spread adoption of such systems especially with respect to improving their resilience against potential faults. Over the course of time, electronic voting has evolved as a substitute for paper ballot voting to decrease redundancies and inconsistencies. Due to the many security and privacy vulnerabilities experienced over time, the past results of e-voting in the last three decades indicate that it has not been very successful. A novel hybrid design based electronic voting system is proposed, implemented and analyzed. The proposed system uses two voter verification techniques to give better results in comparison to single identification-based systems. Finger print and facial recognition-based methods are used for voter identification. For fingerprint verification MANTRA MFS 100 device is used and for facial recognition the automatic voting system uses Convolutional Neural Network (CNN). Also, all votes are encrypted using homomorphic based Paillier cryptosystem. This work is targeted to replace the manual verification system with biometric verification system. The developed system also examines carefully whether the voter has voted once or more.**

Keywords—e-voting, homomorphic encryption, deep learning algorithm, face recognition, fingerprint verification.

I. Introduction

Voting is the democratic right of every citizen that allows them to choose tomorrow's leaders. Voting not only enables people to vote for political parties but also encourages them to understand the value of citizenship [1]. A lot of people do not vote, assuming that one vote does not make a difference, but actually does. The democratic structures of the nation are established by means of elections. Voting is a crucial mechanism that keeps a nation's political structure functioning. This also gives the individual the right to question the government. Voting is a way to express the opinion of a citizen in a democratic country. Voting is crucial to the activation of a democratic process. Literature shows that the voting method has undergone significant changes in order to improve its flexibility, security, speed, cost and availability, especially during the registration and authentication of the elector, the casting of votes and the tabulation phases.

Electoral Systems empower the citizens of a country to elect parliament members of their choice. Paper based electoral system is a classical method to accomplish the said task. In this method, printed votes are submitted to various election booths of country at least one day before the election. After the election timings, sealed boxes containing votes are opened in front of all the legitimate members of booth and counted. The information of counted votes is submitted to a centralized station along with bags of paper votes. The central station compiles and publishes the names of winners and losers through television and radio stations. This method is useful only if the whole process is completed in a transparent way. However, there are some drawbacks to this system. These include higher expenses, longer time to complete the voting process, fraudulent practices by the authorities administering elections as well as malpractices by the voters [1]. These challenges result in manipulated election results.

This paper discusses novel online voting system. This system is based on biometric authentication and Homomorphic encryption. It is completely online voting process where user can vote from anywhere with higher level of security. New system will ensure user privacy, as voters can cast their votes using electronic devices like laptop, and computers. It is an open system. So it gives higher level of trust. It will lead to large number of participant in voting system. This work also concentrates on machine learning and deep learning for face recognition for login to voting system.

This paper is organised as follows. Section 2 is literature survey and related works done by researchers. Section 3 presents proposed system architecture. Section 4 is Result and Discussion details of our system and Section 5 presents Conclusion and future development in the field of E-Voting.

II. Literature Review

Electronic Voting Systems provide efficient and reliable technique to empower citizens of a country or members of an organization to select a person of their choice. These systems can be classified into supervised, hybrid and remote voting styles. Supervised voting also known as offline voting is typically administered by electoral organizations. In this scheme, voting machines are located at polling machines. However, these machines are not connected with a centralized system for cross-verification or any other purpose. Hybrid voting schemes are supervised by election organizing members, however, the machines are connected with internet, Remote voting refers to the schemes which are not administered by any supervising staff and the machines are connected with internet [2]. Benefits of using Biometrics in a voting system is to accurately recognize the voter which enables the election administrators to reduce the error rates by reducing fraudulent and bogus votes. Besides, it also results in cost efficiency, improving physical safety and increasing convenience to the users [3].

Various authors have developed the electronic voting systems. A smart card based voting system is developed by [4]. This smart card system has temporary and permanent storage facilities. To address fraudulent practices, this card also contains biometric information of the end user which can be authenticated by the system.

In [5], biometric voting systems that used an Aadhaar card, voter's iris and fingerprint for enrolment and authentication were developed. Diseases affect iris recognition and the performance of a multimodal system depends on data type and the fusion technique used. The authentication mechanism of [5] requires high computation, large memory capacity, and high cost for implementation. This paper presented an enhanced approach to enrolment and authentication using a metadata (fusion of voter identification number (VIN) and fingerprint) technique that did not expensive and highly secured against attacks. The proposed e-voting model was developed and evaluated to address the identified problems.

Unique finger impression recognition or fingerprint authentication indicates the mechanized strategy for checking a match between two human fingerprints [6]. The examination of fingerprints for coordinating purposes requires the correlation of components of the print design. The extracted parameters of a finger pattern include edges and minutia focuses [7]. These distinct features of a biological pattern give uniqueness to a human being. The mechanized method for the verification of a fingerprint is done by using an electronic device called Fingerprint Verification Module, which captures the unique pattern of a fingerprint in the form of a computerized digital image. The digitally captured images are then processed to prepare a biometric template. This biometric layout is an accumulation of extricated elements which is stored and utilized for coordinating and matching [8]. The proposed system uses a finger print verification module developed by Future Electronics Egypt.

The method employed by the finger print module is the optical method. Optical finger print verification technique maybe defined as the formation of a biometric template from the digitally computerized image for verification using visible light [9]. The surface for scanning the finger print is called as touch surface and underneath there is a light-transmitting phosphor layer which enlightens the surface of the finger. The light reflected from the finger goes through the phosphor layer to a variety of strong state pixels which captures a visual picture of the finger print [10].

Facial recognition system or facial acknowledgement framework is defined as an application capable of detecting and recognizing a person from a digitally processed image [11]. This unit comprises of facial recognition algorithms which includes facial detection, facial feature extraction, formation of biometric template by compression and formation of Eigen vectors and their comparison. Many popular facial recognition algorithms are available in literature that include PCA (Principal Component Analysis) using Eigen faces, LDA (Linear Discriminate Analysis), Fisher-face algorithm and Dynamic link matching [12].

The algorithm used by authors in [13] are based on the principle of feature extraction. Feature extraction in image processing may be defined as being a set of initial value derived from an object in the form of a pattern which is informative and useful for machine learning. The algorithm can be implemented using three steps i.e. Haar feature selection, creation of an integral image and Adaboost training.

Facial recognition is implemented through a cascaded classifier of GPCA and KNN algorithms. KNN is a non-parametric formula used in classification of data. It is also used in pattern recognition. It is one of the simplest algorithms of machine learning for pattern recognition [14]. PCA is an algorithm that converts the correlated elements to linearly uncorrelated elements through orthogonal transformation. In Generalized PCA, the condition of orthogonally is removed to consider an arbitrary number of spaces of unknown and different dimensions [15].

To improve the confidentiality and privacy of the electronic voting systems, most of the systems use Mixnet or homomorphic encryption techniques [16]. Additionally, authors also claim that the homomorphic encryption is more appropriate for the situation with several election candidates as well as elections with neutral votes. The electronic voting system is implemented extensively in developed countries such as USA.

Y. Subba Reddy and P. Govindarajulu in [17] discussed related to machine learning techniques that will provide the evolution of recent voting recommendation techniques. Authors provided new user centric preferences in order to overcome the limitations of current system. It will also lead to improve the reliability and dependability of recommendation systems. This new mechanism can be embedded by calculating weight for views of users and rating.

Dinesh Kumar et al., in [18] proposed new enhanced ensemble classification algorithm which combines various classifiers based on new voting strategy. Final ensemble is constructed after applying various classifiers and selecting best classifiers. In order increase the performance of the ensemble classifiers, association and ensemble concepts are integrated.

III. Proposed System

The Proposed system was developed using Python and Jupyter Notebook. The system uses a set of images of voters as input and uses OpenCV and CNN to train the model and to predict output. The main purpose of this system is to ensure that the election is conducted ethically, this is done by taking voters' image as input and the face recognition model is trained using these input datasets. Then the trained model is utilized for detecting voters faces and ensuring that they vote only once. The Figure 1 represents the approach followed for implementing the Automatic Voting System.

This section discusses the design methodology of the Automatic Voting System using Convolutional Neural Network.

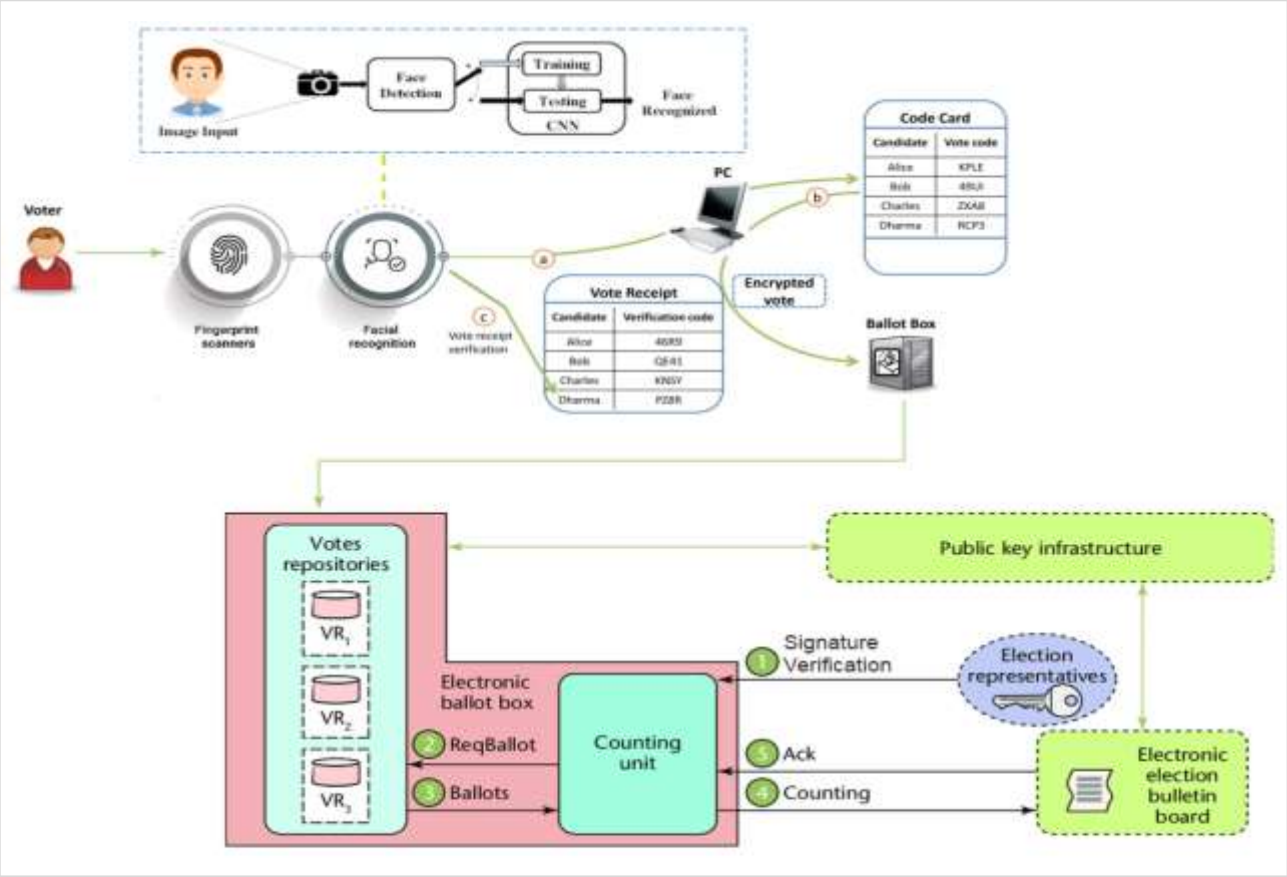

Fig.1 Proposed E-voting system architecture

Input Collection and Creating Dataset: The images of voters are collected by capturing 50 pictures of them and this is considered as the input dataset. The dataset is used for the purpose of training the model. Create the dataset with respect person name with person image are created using opencv module file using python programming.

Image Pre-processing: In an image, the recognition of objects should be done. This approach will likely begin with image processing methods such as removal of noise, followed by filtering of (low-level) features to find lines, regions and potentially areas with certain textures. The key is to view these shapes' sets as single objects. The reason is, an Artificial Intelligence concern is that, when seen from different angles or lighting, an object will look different. These functions are often unintentionally done by the human visual system, but a computer requires highly skilled programming and a range of computational power to approach human performance. Manipulating data through many potential strategies in the form of an image. A picture is generally perceived as a two-dimensional sequence of pixel intensities, and such structures as those of a visual print are most frequently depicted. A picture can be interpreted by a computer, either optically or digitally. The below are the three categories of photographs used in digital image processing:

1. Binary Image - Two-level or bi-level images are also known as binary images. This means that a single bit represents each pixel (0 or 1). The two colours of a binary picture are normally black and white, but any two colours can be used.

2. Gray Scale Image - A input layer is a black-and-white digital image in which each pixel's value is a single sample and the intensity is the only data transported. The strength will vary from 0 to 255. The value 0 indicates the lowest value, while 255 indicates the highest value.

3. Color Image - It will always be a digital image in which the value for each pixel is given based on three primary colours Red, Green and Blue. The intensities of each color quantified by a number between 0 and 255.

Architecture of Convolutional Neural Networks:

CNN are used for facial recognition. It recommended reducing the number of parameters and modifying the architecture of the network exclusively for vision activities. Typically, convolutional neural networks consist of a series of layers that can be grouped by their features. and rms do not have to be defined. Do not use abbreviations in the title or heads unless they are unavoidable.

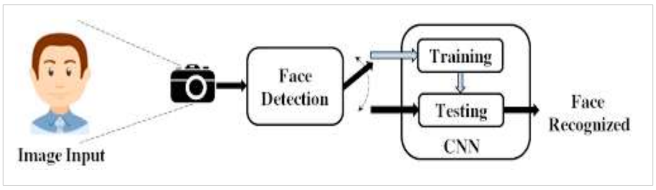

Fig .2 Face Recognition using CNN.

1. Convolution Layer - Processing a 2D convolution on the inputs. "Integrated" across "channels" are the "dot products" between weights and inputs. Across receptive fields, filter weights are exchanged. The filter number would have the same "width" as the output volume.

2. Activation Layer - In order to increase the network non-linearity without upsetting the network's responsive fields of

convolutional layers, it favors ReLU, resulting in faster training. The vanishing gradient question is discussed by LeakyReLU.

3. Softmax - A special type of activation plate, normally at the end of the FC layer. Outputs can be used as a Normalized exponential function. Produce a discreet and convenient probability distribution vector when combined with cross-entropy loss.

4. Pooling Layer - Convolutional layers provide activation maps, but the pooling layer uses nonlinear down sampling on activation maps. Pooling is competitive, Filter sizes are becoming smaller, and pooling is being phased out.

5. FC Layer -It's a typical neural network that can be viewed as the end result of a learning process, where maps extract visual features to the desired outputs. It is typically adaptive to the activities of classification/encoding. The typical output is a vector, which is then passed through softmax to reflect classification trust. Outputs can also be used as "bottleneck".

Homomorphic Encryption Technique (Paillier Cryptosystem)

Key Generation:

- Select two large prime numbers a and b arbitrary and independent of each other such that gcd(n, Φ (n)) = 1, where Φ (n) is Euler Function and n=pq.
- Calculate RSA modulus n = pq and Carmichael's function is given by
  λ = lcm (p-1, q-1).
- Select g called generator where g$\in\mathbb{Z}^*_{n2.}$ Select α and β randomly from a set $\mathbb{Z}_n^*$ then calculate g = (αn + 1) $\beta^n mod\ n^2$.
- Compute the following modular multiplicative inverse μ = (L (g$^\lambda$mod n^2)$^{-1}$ mod n. Where the function L is defined as L(u) = (u−1)/n.
- The public (encryption) key is (n and g).
- The private (decryption) key is (λ and μ).

Encryption:

- Let mess be a message to be encrypted where mess $\in \mathbb{Z}_{n.}$
- Select random r where r $\in \mathbb{Z}^*_{n2.}$
- The cipher text can be calculated as:
  - cipher = g$^{mess}\cdot$r^n.mod n^2.

Decryption:

- Cipher text c$\in\mathbb{Z}^*_{n}{}^2$
- Original message: mess = L (cipher$^\lambda$ mod n^2). μ mod n.

IV. Result and Discussion

A. Experimental Setup

All the experiments are performed on environment of Intel i5, 3.2 GHz, 8 GBs of RAM, 1 TB of hard disk, 512 GB of SSD and Windows 10 operating system. Eclipse IDE is used and Java technology (JSP / Servlet) & Python Technology is used.

B. Performance Parameters

The formula for calculating these measures is as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1Score = \frac{2 * Precision * Recall}{Precision + Recall}$$

True positive, true negative, false positive, and false negative values are represented as TP, TN, FP, and FN, respectively.

C. Proposed System Result

The results of the performance evaluation of CNN techniques is shown in Figure 3.

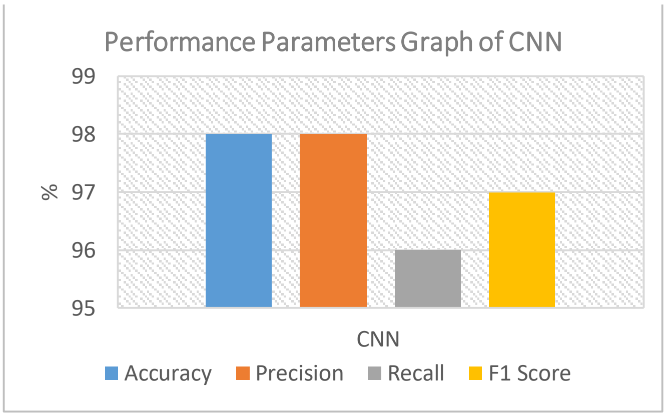

Fig 3. Performance Measures graph of CNN algorithms

Figure 4 shows the training and validation accuracy comparison of CNN model for 10 epochs. The validation accuracy after 10 epochs is 98.20 %. With increase in number of epoch the accuracy remains nearly same as shown in figure 4.

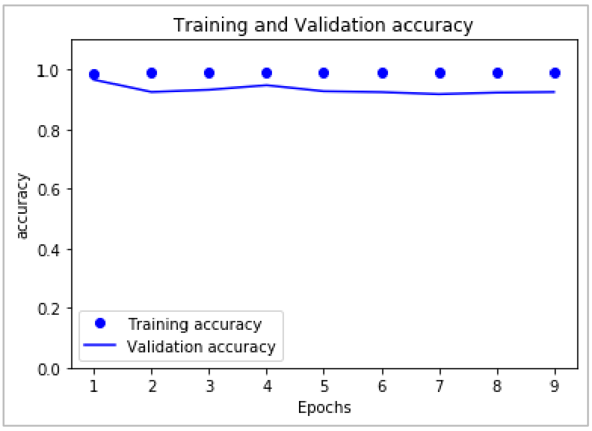

Fig 4. Training and validation accuracy comparison of CNN.

Figure 5 shows the training and validation loss comparison of CNN model for 10 epochs. The validation loss after 10 epochs is 0.08. With increase in number of epoch the loss is reduced which can be seen in figure 5.

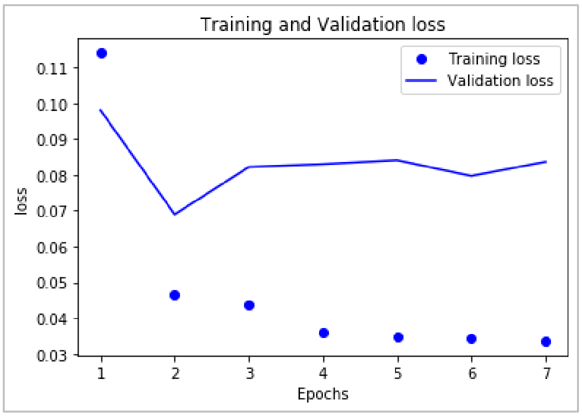

Fig. 5 Training and validation loss comparison of CNN.

Our findings reveal that CNN classifier (for face recognition) beats other approaches with 10 fold and 80 % – 20 % data splitting range and gives maximum accuracy of 98.20 %.

V. Conclusion

The automatic voting system using convolutional neural networks makes the voting process easy as it requires minimum to no human intervention. This Automatic Voting system would take the place of the time-consuming and difficult-to-maintain manual process. Each and every minute, this machine produces modified data. It, therefore, necessitates less capital and manpower. This research was on a secured electronic voting system based on fingerprint and face recognition techniques. A highly secured e-voting model was designed, implemented, tested and evaluated. The results indicated an accuracy of 98.50% making the proposed system reliable in tackling the identified problems, restoring confidence and improve citizens' participation in the electoral process.

References

[1] M. P. Wattenberg, is voting for young people? Routledge, 2020.

[2] Traoré, J., "An Introduction to Electronic Voting Application to Single Transferable Vote", Orange Labs, July, 2014.

[3] Whither Biometrics Committee, "Biometric Recognition: Challenges and Opportunities", National Academies Press, 2010.

[4] Drexler, J., and Dyball, C.J., "Anti-Fraud Voter Registration and Voting System using a Data Card", Google Patents, 1995.

[5] Aman, J., Yojna, A., Jitendra, P., Sachin, Y., & Konark, S. (2020). Design and development of biometric enabled advanced voting system. International Journal of Innovative Research in Computer Science & Technology (IJIRCST), Vol. 8, No. 3, pp. 50-53.

[6] Kumar, R., Gowri, B., and Kumar, K.V., "Biometric Security Based Application Development and Emulation Framework-IEEE 802.15.4 for Intensive Care Units", International Conference on Recent Advances in Computing and Software Systems, pp. 228-232, 2012.

[7] Thornton, J., "Latent Fingerprints, Setting Standards in the Comparison and Identification", Proceedings of 84th Annual Training Conference, California, State Division of IAI, 2000.

[8] Khan, M.K., "Fingerprint Biometric-Based Self Authentication and Deniable Authentication Schemes for the Electronic World", IETE Technical Review, Volume 26, pp. 191-195, 2009.

[9] Wasserman P.D., "Solid-State Fingerprint Scanners", Presentation, NIST, 2005.

[10] Setlak, D.R., "Advances in Biometric Fingerprint Technology Are Driving Rapid Adoption", Consumer Marketplace, Retrieved, Volume 13, December, 2005.

[11] Bansal, A., Mehta, K., and Arora, S., "Face Recognition Using PCA and LDA Algorithm", 2nd International Conference on Advanced Computing & Communication Technologies, pp. 251-254, 2012.

[12] Brunelli, R., and Poggio, T., "Face Recognition: Features Versus Templates", IEEE Transactions on Pattern Analysis & Machine Intelligence, pp. 1042-1052, 1993.

[13] Viola, P., and Jones, M., "Rapid Object Detection Using a Boosted Cascade of Simple Features", Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Volume 1, pp. I-511-I-518, 2001.

[14] Peterson, L.E., "K-Nearest Neighbor", Scholarpedia, Volume 4, pp. 1883, 2009.

[15] Vidal, R., Ma, Y., and Sastry, S., "Generalized Principal Component Analysis (GPCA)", IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 27, pp. 1945-1959, 2005.

[16] Azougaghe, A., Hedabou, M., and Belkasmi M.," An Electronic Voting System Based on Homomorphic Encryption and Prime Numbers", 11thInternational Conference on Information Assurance and Security, pp. 140-145, 2015

[17] Y. Subba Reddy and Prof. P. Govindarajulu, "A survey on data mining and machine learning techniques for internet voting and product/service selection", IJCSNS International Journal of Computer Science and Network Security, Vol. 7, No.9, 2017.

[18] T. Dinesh Kumar, E. Prabhakar, K. Nandhagopal, "An Enhanced Ensemble Classification Algorithm (EECA) for Airline Services Big Data Sentiment Analysis", Journal of Advanced Research in Dynamical & Control Systems, Volume 11, Issue 7, ISSN 1943-023X, Page No.1461-1467, August 2019.