ISSN No:-2456-2165

RV College of Engineering®, Bengaluru-59

(Autonomous Institution Affiliated to VTU)

Department of Electronics and Communication Engineering

16EC81 MAJOR PROJECT
Synopsis

Disease Prediction by Machine Learning

By

Name of the students	USN
PRATEEK	**1RV14EC112**
BANNAPPA A	**1RV16EC401**
NITIN KUMAR	**1RV16EC416**
SARAVANA V	**1RV16EC431**

In house Project	
Guide Name	PROF. SUBRAHMANYA KN
Guide Designation	ASSISTANT PROFESSOR

Abstract:- Disease Prediction utilizing Machine Learning is a framework which predicts the infection dependent on the data or the indications he/she go into the framework and gives the exact outcomes dependent on that data. In the event that the patient isn't a lot of genuine and the client simply needs to know the kind of infection, he/she has experienced. It is a framework which gives the client the tips and deceives to keep up the wellbeing arrangement of the client and it gives an approach to discover the sickness utilizing this forecast. Presently a day's wellbeing industry assumes significant part in restoring the infections of the patients so this is additionally some sort of help for the wellbeing business to tell the client and furthermore it is valuable for the client in the event that he/she would not like to go to the clinic or some other centers, so by entering the manifestations and any remaining helpful data the client can become acquainted with the illness he/she is experiencing and the wellbeing business can likewise get advantage from this framework simply by asking the indications from the client and entering in the framework and in only couple of moments they can advise the specific and dependent upon some degree the precise sicknesses. This Disease Prediction Using Machine Learning is totally finished with the assistance of Machine Learning and Python Programming language with Tkinter Interface for it and furthermore utilizing the dataset that is accessible beforehand by the emergency clinics utilizing that we will anticipate the illness.

I. INTRODUCTION

In this paper, we smooth out AI calculations for compelling expectation of infection episode in sickness regular networks. We Proposed a medical services framework utilizing brilliant dress for practical wellbeing checking. Calm al. had altogether examined the heterogeneous frameworks and accomplished the best outcomes for cost minimization on tree and straightforward way cases for heterogeneous frameworks. Patients' factual data, test results and illness history are recorded in the EHR, empowering us to recognize potential information driven answers for diminish the expenses of clinical contextual analyses Handle organized information, we talk with clinic specialists to extricate helpful highlights. For unstructured content information, we select the highlights naturally utilizing choice tree calculation. At last, we propose a novel choice tree calculation for organized and unstructured information. The sickness hazard model is gotten by the mix of organized and unstructured highlights. Through the

ISSN No:-2456-2165

examination, we make a determination that the exhibition of choice tree is superior to other existing strategies.

II. LITERATURE SURVEY

Smart clothing - Connecting human with clouds and big data for sustainable health monitoring: Wellbeing checking through conventional wearable gadgets is difficult to be economical. To give unavoidable knowledge to savvy attire framework, portable medical care cloud stage is built by the utilization of versatile web, distributed computing and large information examination[1].

Disease prediction using Machine Learning over Big Data: Convolutional neural network (CNN)- based multimodal infection hazard forecast calculation utilizing organized and unstructured information from medical clinic[2].

Big Data – A Survey: Information age, information securing, information stockpiling, and information analysis.Enterprise the executives, Internet of Things, online interpersonal organizations, average applications, aggregate knowledge, and shrewd lattice[3].

Wearable 2.0 - Enable human-cloud integration in next generation healthcare system: We propose a Wearable 2.0 medical services framework to improve QoE and QoS of the cutting edge medical services framework[4].

The 'Big Data 'Revolution in Healthcare: Accelerating Value and Innovation: Information driven rule to accomplish the "5R" objective for levelheaded medication use and clinical pathways. the exploration system of huge information investigation in medical care, examination of clinical cycle, and the writing synopsis of therapy design mining[5].

OBECTIVES

- To collect and create a database of symptoms of patients and predicted diseases.
- To compare the performances of applicable AI/ML algorithms (Decision Tree, Rain Forest, Naive Bayes) on the created database.
- To apply and validate the best algorithm.
- To develop an user interface (UI) to assist the doctor in identifying the exact disease.

III. METHODOLOGY

The machine learning algorithm will be used to predict the risk of symptoms in terms of percentage and by using several tasks.

The tasks that we are going to carry out are as follows:-

A. Collect related dataset (CSV file)
B. Preparing the dataset
C. Data Preprocessing
D. Applying Algorithm
E. Classification
F. Prediction

Decision Tree Algorithm :-

- Decision Tree Algorithm has a place with the group of directed learning calculations.
- The objective of utilizing a Decision Tree is to make a preparation model that can be utilized to foresee the class or worth of the objective variable by taking in basic choice guidelines alluded from set of qualities.
- In Decision Trees, for foreseeing a class mark for a record we start from the base of the tree. The upsides of the root trait are contrasted and the record's property.

Naive Bayes Algorithm :-

- Naive Bayes calculation is an administered learning calculation, which depends on Bayes Theorem utilized for tackling arrangement issue.
- Naive Bayes calculation is one of the straightforward and best Classification calculations which helps in building the quick AI models that can make fast forecasts.
- It performs well in Multi-class forecasts when contrasted with different Algorithms.
- It is utilized in text arrangement, for example, Spam sifting and Sentiment investigation.

Random Forest Algorithm :-

- Random woodland is classifier that contains various choice trees on different subsets of given dataset and takes the normal to improve the prescient exactness of the dataset.
- It can be valuable for tackling choice related issues.
- It assists with contemplating every one of the potential results for an issue.
- It has less prerequisite of information cleaning contrasted with different calculations.

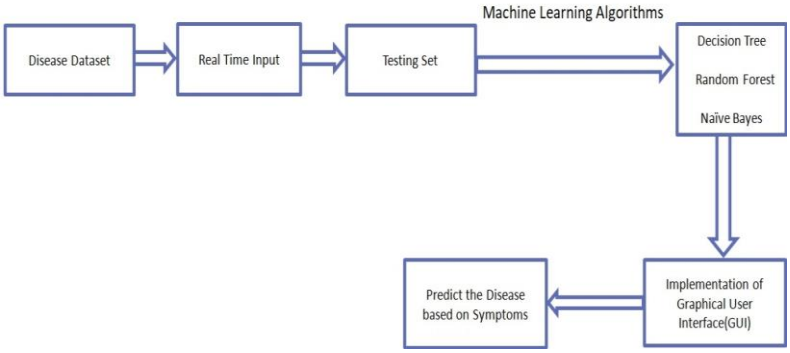
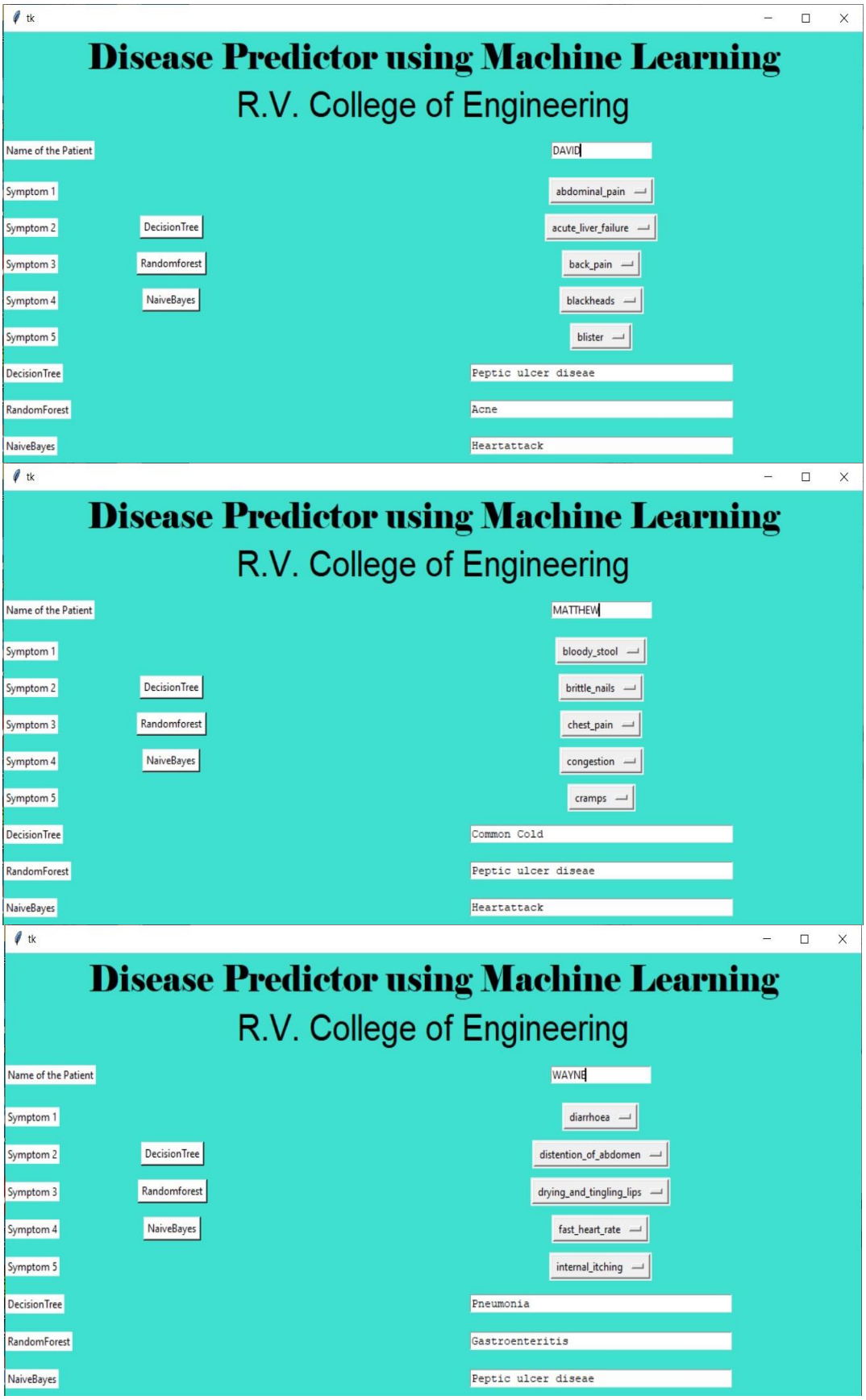
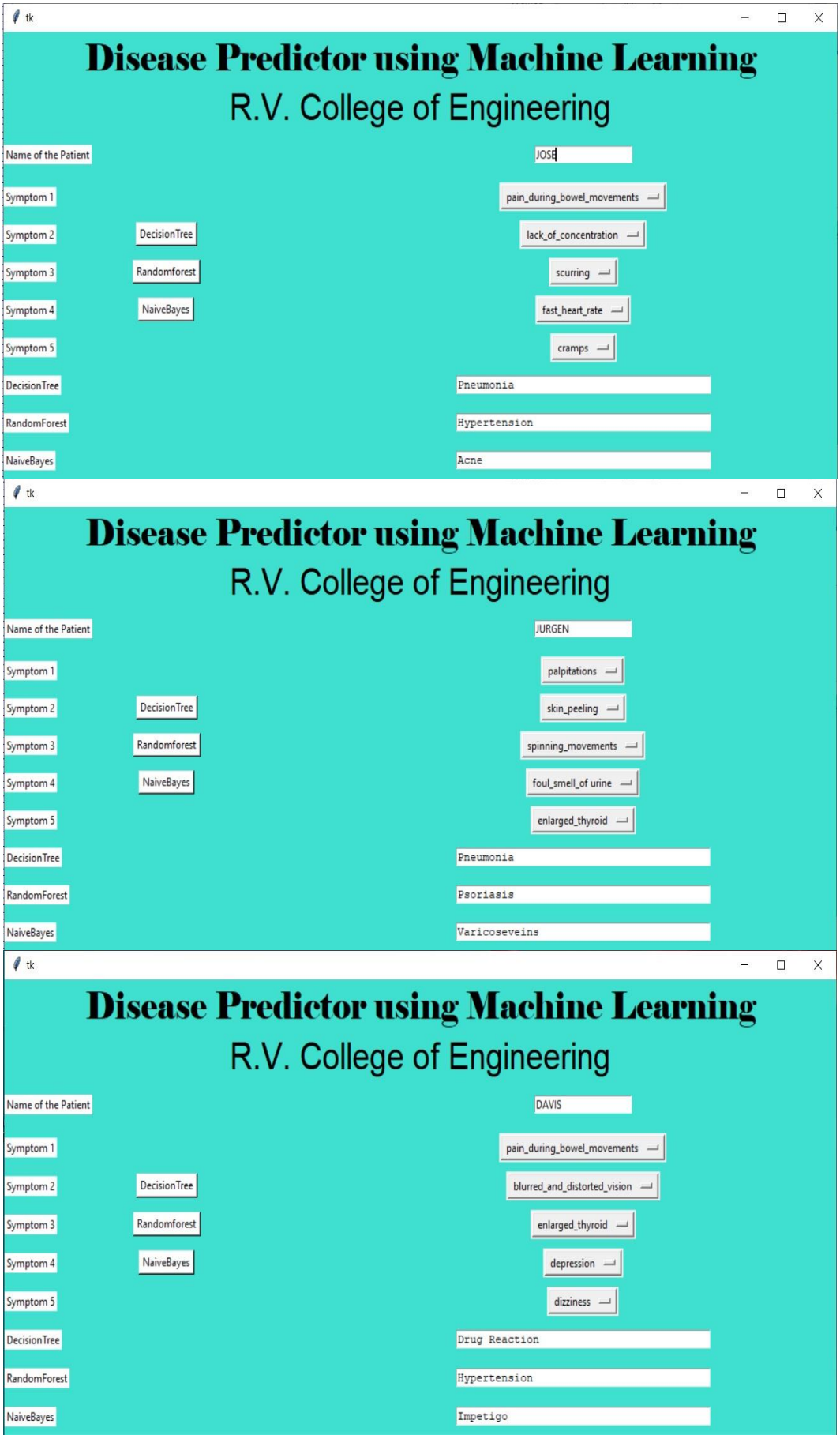


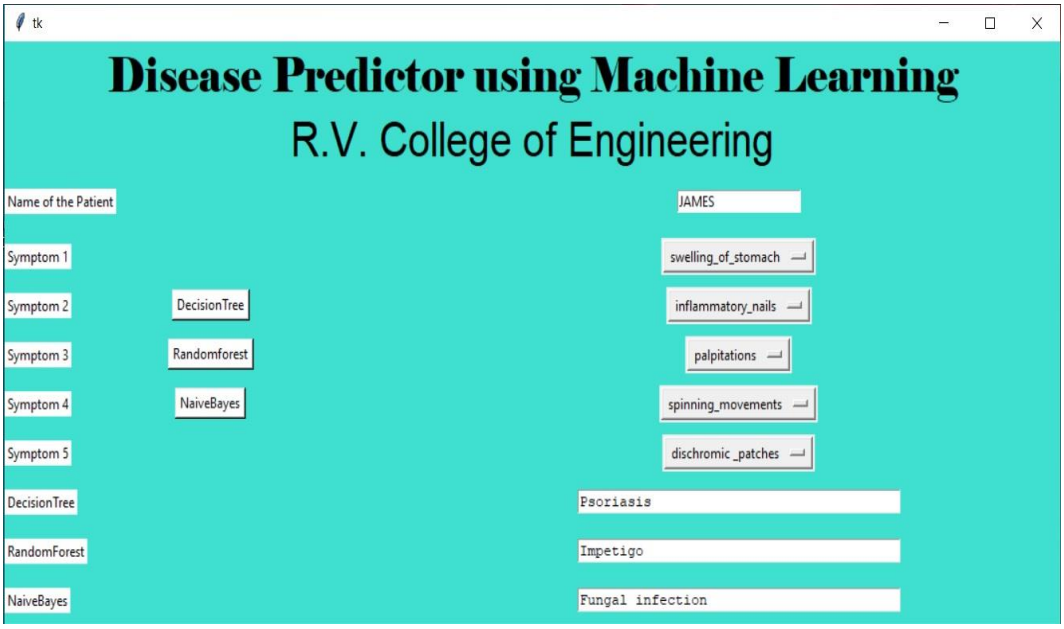
Fig. 1. Flow Diagram of disease prediction

IV. EXPERIMENTAL RESULTS



ISSN No:-2456-2165





All the above figures are the results where different patients with their symptoms are taken and according to the used algorithms, predicted disease have been displayed.

SOFTWARE REQUIREMENTS

OS : Windows 8 or higher
Technology : Machine Learning
Front End : Tkinter (GUI)
IDE : Python 3.9 or higher

V. CONCLUSION

Accordingly, Finally I wrap up by saying that, this project Disease prediction using AI is a ton of important in everyone's regular day to day existence and it is in a general sense more critical for the clinical consideration region, since they are the one that step by step uses these systems to expect the sicknesses of the patients reliant upon their general information and there signs that they are capable. As of now daily's prosperity industry accepts critical part in reestablishing the diseases of the patients so this is in like manner some kind of help for the prosperity business to tell the customer and moreover it is significant for the customer if he/she might not want to go to the crisis facility or some different focuses, so by entering the signs and any leftover supportive information the customer can turn out to be more familiar with the disorder he/she is encountering and the prosperity business can similarly get advantage from this system basically by asking the results from the customer and entering in the structure and in several minutes they can exhort the particular and ward upon some degree the specific sicknesses. If prosperity industry accepts this endeavor, created by the experts can be lessened and they can without a doubt expect the contamination of the patient. The disease gauge is to offer assumption to the extraordinary and overall happening ailments that when unchecked and sometimes neglected can changes into dangerous sickness and cause package of issue to the patient and similarly as their family members.

REFERENCES

[1]. M. Chen, Y. Ma, J. Song, C. Lai, and B. Hu, ``Smart clothing: Connecting human with clouds and big data for sustainable health monitor-ing," ACM/Springer Mobile Netw. Appl., vol. 21, no. 5, pp. 825_845,2016.
[2]. M. Chen, Y. Hao, K. Hwang, L. Wang, and L. Wang, "Disease prediction using Machine Learning over Big Data." IEEE Access (Volume: 5), pp. 8869 – 8879.
[3]. M. Chen, S. Mao, and Y. Liu, ``Big data: A survey," Mobile Netw. Appl.,vol. 19, no. 2, pp. 171_209, Apr. 2014.
[4]. M. Chen, Y. Ma, Y. Li, D. Wu, Y. Zhang, and C. Youn, ``Wearable 2.0:Enable human-cloud integration in next generation healthcare system,"IEEE Commun., vol. 55, no. 1, pp. 54_61, Jan. 2017.
[5]. P. Groves, B. Kayyali, D. Knott, and S. van Kuiken, The`Big Data'Revolution in Healthcare: Accelerating Value and Innovation. USA:Center for US Health System Reform Business Technology Office, 2016.
[6]. M. Qiu and E. H.-M.Sha, ``Cost minimization while satisfying hard/softtiming constraints for heterogeneous embedded systems," ACM Trans.Design Autom. Electron.Syst., vol. 14, no. 2, p. 25, 2009.VOLUME 5, 2017 8877M. Chen et al.: Disease Prediction by Machine Learning Over Big Data From Healthcare Communities
[7]. J. Wang, M. Qiu, and B. Guo, ``Enabling real-time information serviceon telehealth system over cloud-based big data platform," J. Syst. Archit.,vol. 72, pp. 69_79, Jan. 2017.
[8]. D. W. Bates, S. Saria, L. Ohno-Machado, A. Shah, and G. Escobar, ``Bigdata in health care: Using analytics to identify and manage high-riskand high-cost patients," Health Affairs, vol. 33, no. 7, pp. 1123_1131,2014.
[9]. L. Qiu, K. Gai, and M. Qiu, ``Optimal big data sharing approach for tele-health in cloud computing," in Proc. IEEE Int. Conf. Smart Cloud (Smart-Cloud), Nov. 2016, pp. 184_189.

[10]. Y. Zhang, M. Qiu, C.-W. Tsai, M. M. Hassan, and A. Alamri, ``Health-CPS: Healthcare cyber-physical system assisted by cloud and big data,"IEEE Syst. J., vol. 11, no. 1, pp. 88_95, Mar. 2017.
[11]. K. Lin, J. Luo, L. Hu, M. S. Hossain, and A. Ghoneim, ``Localization basedon social big data analysis in the vehicular networks," IEEE Trans. Ind.Informat., to be published, doi: 10.1109/TII.2016.2641467.
[12]. K. Lin, M. Chen, J. Deng, M. M. Hassan, and G. Fortino, ``Enhanced fingerprinting and trajectory prediction for iot localization in smart buildings," IEEE Trans. Autom. Sci. Eng., vol. 13, no. 3, pp. 1294_1307,Jul. 2016.
[13]. D. Oliver, F. Daly, F. C. Martin, and M. E. McMurdo, ``Risk factors andrisk assessment tools for falls in hospital in-patients: A systematic review,"Age Ageing, vol. 33, no. 2, pp. 122_130, 2004.
[14]. M. Chen, P. Zhou, and G. Fortino, ``Emotion communication system," IEEE Access, vol. 5, pp. 326_337, 2017,doi: 10.1109/ACCESS.2016.2641480.
[15]. D. Tian, J. Zhou, Y. Wang, Y. Lu, H. Xia, and Z. Yi, ``A dynamic andself-adaptive network selection method for multimode communicationsin heterogeneous vehicular telematics," IEEE Trans. Intell. Transp. Syst.,vol. 16, no. 6, pp. 3033_3049, Dec. 2015.