International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-8 Issue-6, August, 2019

Human Behavioral Analyzer using Machine Learning Technique

Vikash Yadav, Rati Shukla

Abstract*: This research paper aims to the variety of people suffering from medium or low level of mental agitation i.e. being stress, depression etc. As countries like India in which more than 65% of the population is under the age of 35 [1] are continuously falling down the rank in the World Happiness Report, In 2018, India ranked on 133rd [2] position, and it can be concluded that the majority of population is facing mental health issues and does not have proper methods to analyze their mental health and take appropriate precautions and also to provide automated solutions to the Industry for hiring a productive group of people those are cool minded and sensible, the purpose of this research is to analyze the mental health of a person using behavioral traits of the person that are entered by the person or chosen from a list of given options throughout the analyses procedure of the application in which surveyed data is tested through Machine Learning to determine the status of mental health of a person and associated stress levels and suggesting the user with appropriate recommendations.

***Keywords*: Stress, Machine Learning, Support Vector Machine, Classifier.**

I. INTRODUCTION

To analyze the mental health of a person using his/her behavioral traits and determine the level of mental agitation and stress in the person and suggesting him with appropriate recommendations. In the modern times and the phase of growing technology, People are busy with their hectic and packed daily routines, resulting of which they are corroding their mental health and the symptoms of which are high stress levels, frustrated and irritated living and the consequences of which are catastrophic.

The paper provide an automated solutions to the Industry for hiring a productive group of people those are cool minded and sensible and can also be used for regular stress check of their employees. This will lead to very less number of persons being fired from their jobs and resulting to a higher Level of happiness in the country.

The Paper comprises of 5 Sections. Section 1 contains introduction of the problem domain of the Human Behavioral Analyzer, also the motivation for this is explained along with its varied scope. Section 2 summarized the literature survey carried out in development of the project, multiple perspectives of Stress are briefly explained, Optimum ML Algorithms for different stress Stimuli are mentioned and the development of Assessment of stress through multiple questionnaires is explained, Dataset for Survey corresponding to the Assessment is displayed along some general examples of Stress are included along with Stress Busters. Section 3 describes the Project's System Design and Architecture of the System along with its diagrammatical view, Flowcharts. Methodology is explained and accuracies are mentioned for various algorithms used in training the ML model and classification of Data.

Section 4 contains the Software and Hardware Requirements for the Successful Implementation of the system and Images of the Front end Implementation of this system and associated results are also there in this chapter along with the assumptions made during development of project and its dependencies. Section 5 contains the Conclusion and future work of this paper.

II. LITERATURE SURVEY

The Research Study for Behavior Analysis, and to find the optimum Algorithm for Implementation [3]. There are 3 perspectives to assess stress i.e. the biological perspective, the behavioral perspective and the phenomenological one. Data can be annotated through self assessment and behavioral perspective of a person by recruiting several assessments. Collecting data of good quality is mandatory to have a robust analysis, due to the presence of an external observer the results may vary that is called the Hawthorne effect and this should need to be minimized and there are also ethical limits to a person when dealing with negative state of emotions which are needed to be analyzed in order to work upon them [4]. Unlike pattern recognition it is almost impossible to assess a phenomenon that is affective and also it cannot be annotated. Ultimately the quality of data is important for robust analysis [5,6]. There can be several assessment strategies like self-assessment, external perception, also physiological markers but in this work these markers are not our major concern. Emotional expressions differ majorly from person to person and are highly multimodal. Same emotion can be expressed in a number of ways by a single person. However, the interesting point is that these different emotions can easily be recognized by human beings despite of any inter individual differences among them [7]. Out of the 3 perspectives we are focusing on phenomenological and the behavioral one, self perception is the key aspect in the phenomenological perspective [8].

Revised Manuscript Received on October 30, 2019.
* Correspondence Author
Vikash Yadav, Computer Science & Engineering Department, ABES Engineering College, Ghaziabad, India.
Rati Shukla, GIS Cell, MNNIT Allahabad, Prayagraj, India.

According to Cannon –Bard Theory that states that "Stress can occur even when the body changes are not present because the physiological response of the body and also Lazarus states that Stress occurs when an individual perceives that the demands of an external".

"Situations are beyond his or her ability to cope with them"

Individual perception is the major basis of our definition, stress can be measured with the usage of questionnaires and visual analogue scales. Behavioral perspective states that behaviors like scratching face touches and lip biting and many other displacement activities are associated with stress and its experiences and they may provide very important information about the subject's state of emotion than its verbal statements and expressions [9].

The State Trait Anxiety Inventory: The State-Trait Anxiety Inventory is a type of questionnaire designed by Charles Spielberger and it contains 40 self report type items [14] which helps in measuring stress and anxiety of a person. Two types of anxiety can be evaluated that are state and trait anxiety.

State anxiety can be defined as the anxiety which is felt at a particular instant only, while trait anxiety can be defined as person's relatively enduring disposition making him/her to feel anxious or not, also the questionnaires of other general tests has been added in this to enhance the functionality and diversify the assessment.

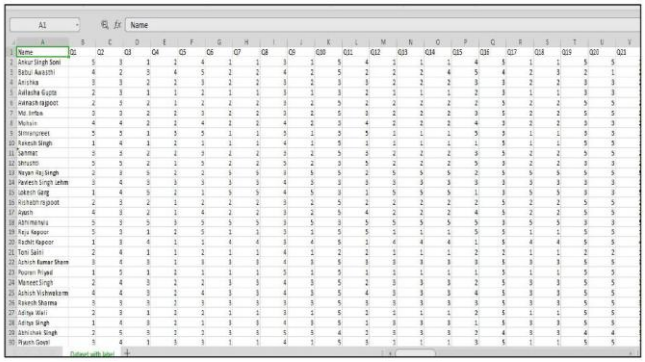


Fig: **1. Data set containing records of 500 people**

Some Examples of stress are as follows:

• Routine stress majorly concerned with the work pressure, school, family, friends and other responsibilities of life.

• Stress generated due to some negative changes, example being fired from a job, beaten up by strangers, or any sudden illness.

Some Examples of stress Busters are as follows:

• Do Exercise on Regular Basis: Around 25-30 minutes of exercise in a day or just walking can help enhance your mood and result in reduction of stress levels.

• Some Relaxing Activities: Some Relaxing activities like yoga, meditation can help reduce stress.

• Make a schedule and follow it religiously for relaxing and Physical activities like jogging etc.

• Have a good social network and try to be connected and socialized with people who are always there to support you, this may help you to reduce stress, ask for help whenever you need.

The data used in this work is heterogeneous kind of data which can be categorized in the form of Big Data [15,16,17].

III. SYSTEM DESIGN & METHODOLOGY

The System for the Human Behavioral Analyzer is designed using Robotics Process Automation Technology, So that it can be implemented on multiple systems using RPA Robots that are based on Artificial Intelligence.

The System Design is based on RPA, through which the whole system can be controlled through AI Robots and a model is trained through Machine Learning is then used to analyze the class of Stress the user might be facing and as per the level of stress the assessment report is sent to the user which also contains appropriate recommendations. The detailed Diagrammatical View of the System is as follows through Fig. 2.

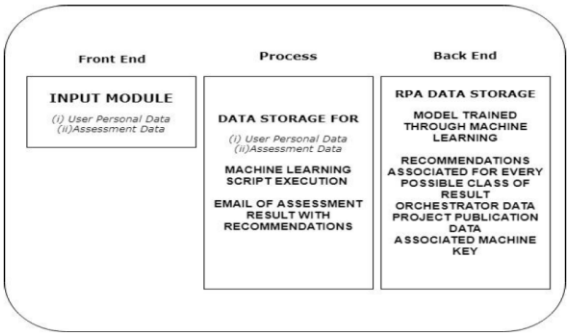


Fig: 2. System Architecture/Diagrammatical View

Algorithms used in Training the Machine Learning Model

3.1.1. Random Forest Classifier

Random decision forests or Random forests is an ensemble learning method for classification, and other for every possible works that runs by the construction of multitude of decision trees and especially at training time and resulting into output of mode of the classes or mean prediction of each and every distinct trees. Random decision forests correct for decision trees' habit of outfitting to their training set.

Flow Chart

Random subspace method [2], the first algorithm for random decision forests was created by Tin Kam Ho using the which, in Ho's formulation, is a way to implement the "stochastic discrimination" approach to classification proposed by Eugene Kleinberg.

An extension of the algorithm was developed by Leo Breiman and Adele Cutler, who registered "Random Forests" as a trademark (as of 2019, owned by Minitab, Inc.). The extension combines Breiman's "bagging" idea and random selection of features, introduced first by Ho and later independently by Amit and Geman in order to construct a collection of decision trees with controlled variance [10].

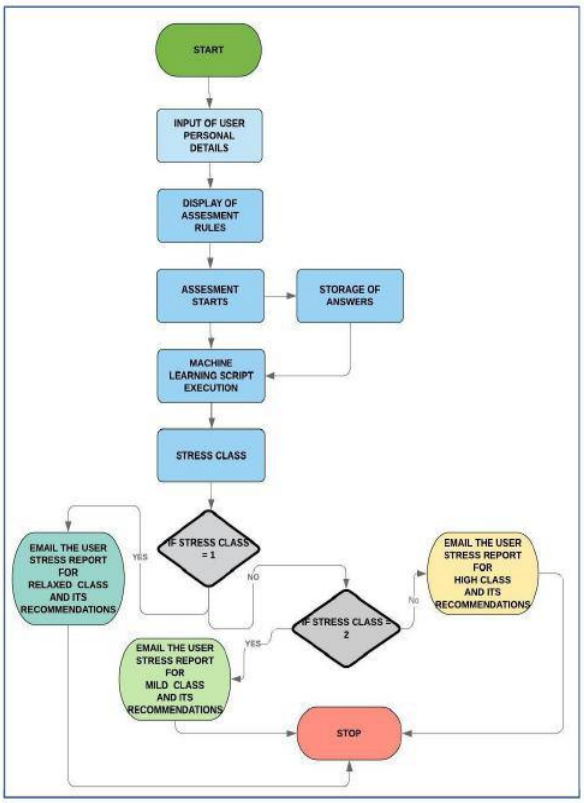


Fig: 3. Flow Diagram

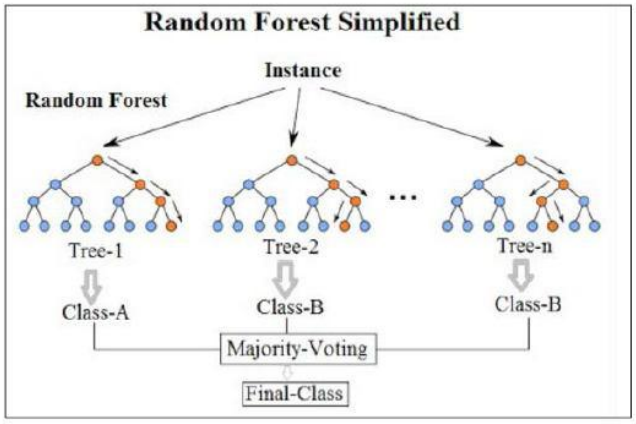


Fig: 4. Random Forest Classifier [11]

The Model's Accuracy with Random Forest Classifier is as follows:

```
RandomForestClassifier(bootstrap=True, class_weight=None, criterion='gini',
                       max_depth=None, max_features='auto', max_leaf_nodes=None,
                       min_impurity_decrease=0.0, min_impurity_split=None,
                       min_samples_leaf=1, min_samples_split=2,
                       min_weight_fraction_leaf=0.0, n_estimators=10,
                       n_jobs=None, oob_score=False, random_state=None,
                       verbose=0, warm_start=False)

print ("Accuracy with random forest  is{}".format(accuracy_score(y_test,clf.predict(X_test))))
Accuracy with random forest  is0.6933333333333334
```

Fig: 5. Accuracy with Random Forest Classifier

3.1.2. Support Vector Machine

Machine learning has some widely used supervised learning Models, one of them are Support vector machines, SVMs use associated learning algorithms that analyze data used for regression analysis and classification. For example In a set of examples used in training and each sample being categorized to either of the two categories, SVM algorithm here will build a model that assigns every new examples to either of the one category or the other category during training, which eventually make it a non-probabilistic binary linear classifier.

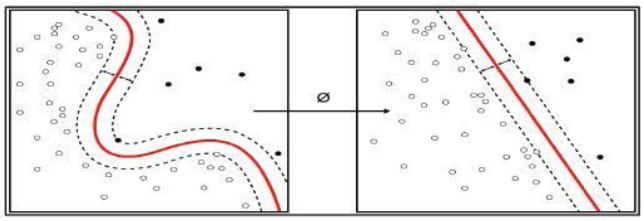


Fig: 6. Support **Vector Machine [12]**

SVM models basically represents examples as points in space, mapped properly so that every example which are assigned to their respective categories are divided by a clear wide gap as shown in Fig. 6 Mapping of new examples are done into that same space and they are predicted to belong to a category based on which side of the gap they fall.
SVMs can efficiently perform a non-linear classification apart from linear classification, using the kernel trick, implicitly mapping their inputs into high-dimensional feature spaces [13].

The Model's Accuracy with Support Vector Machine is as follows:

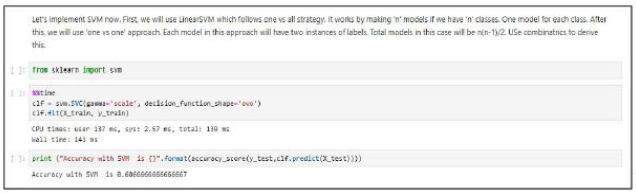


Fig: 7. Accuracy with Support Vector Machine

3.2. Methodology

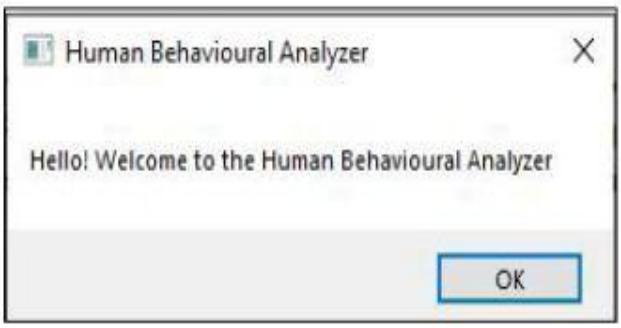
The Proposed method includes collection of data through a survey from not less than 500 undergraduate people that will undergo the Stress Assessment questionnaire and the same data will be used to train the model through Machine Learning algorithms [Percentage Split=70%], So as to use the Trained Model to determine the level of Stress in a Person according to the answers given by him/her.
The results will be sent to the user over Email along with appropriate recommendations as per the level of stress corresponding to that particular user.

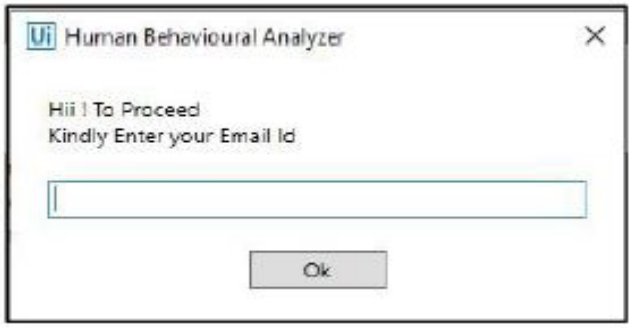
IV. IMPLEMENTATION AND RESULTS

Assumptions and dependencies

1. The Credibility of surveyed data might be low.
2. The project only predicts well only for persons of teen age group.

Front End Implementation and Tool Testing





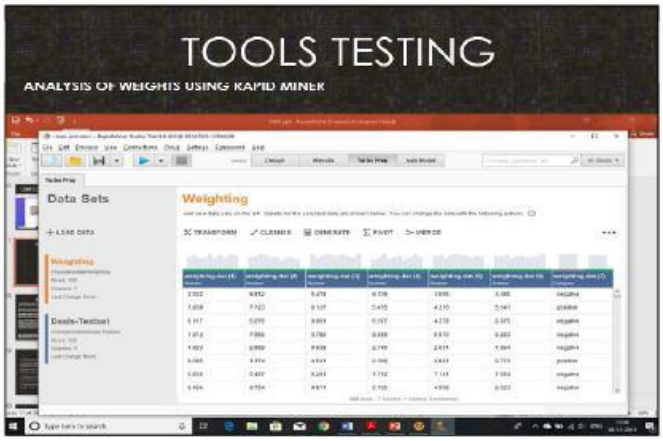
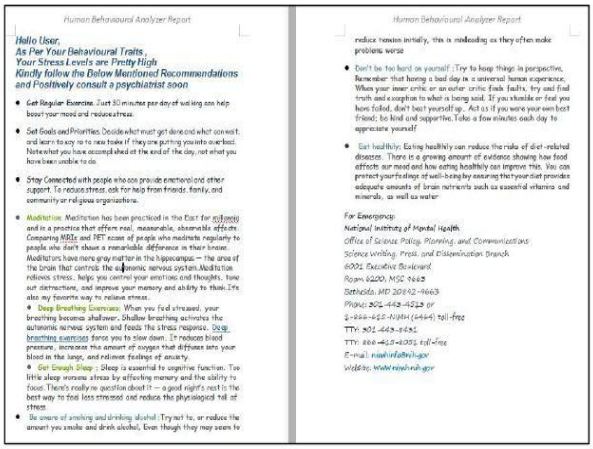


Fig: 8. Screenshot of various performed operations

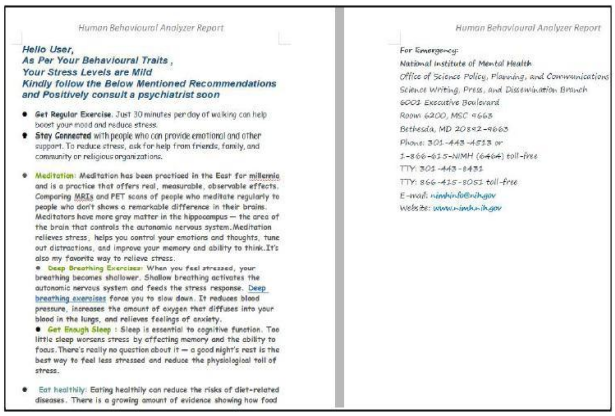
V. RESULTS

The Result of the system has categorized in three states: For highly stressed person, for mild stressed person & for relaxed person. Various screen shots of the system has shown below:

For Highly Stressed Person



For Mild Stressed Person



For Relaxed Person

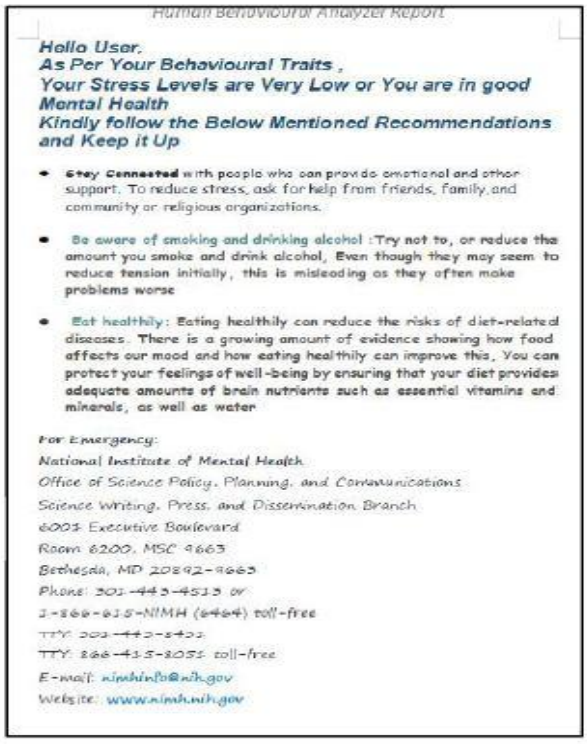


Fig: 9. Results of various stress levels

Robot Generated Email

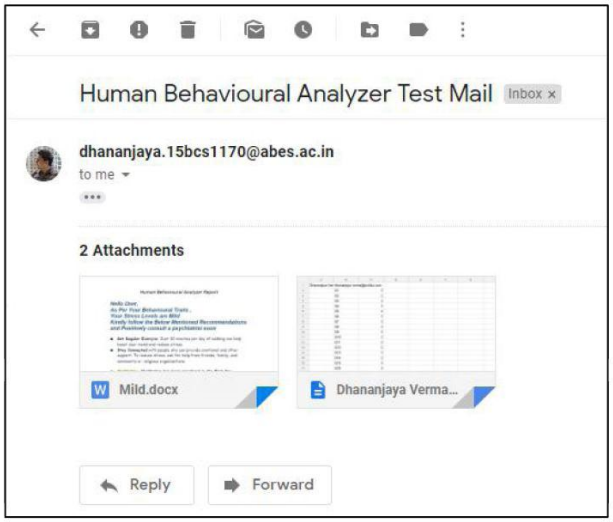


Fig: 10. Robot Generated Email

VI. CONCLUSION

The Algorithms used in the classification, that are Random Forest Classifier and Support Vector Machine has classified the data into 3 Stress Classes that are Relaxed (Numeric Value 1), Mild (Numeric Value 2), High (Numeric Value 3) with accuracy of 69.33 % and 60.66% Respectively and with the usage of Robotics Process automation in the implementation of research enhances the scalability of this work and makes it feasible to run at multiple systems simultaneously accessing the common trained machine learning model.

The Implementation of this work opens of ventures for much more accurate and economically feasible Human Mental Health Analyzers which can be used by anyone who wants to analyze his/her mental health and also this work has the scope of being freely distributed by the Government of the Country to its citizens, to promote a healthy Nation.

In future this work have the scope of being its Machine Learning Engine to be rigorously trained with much higher number of records and assessment data of persons of all age group and the model can predict the most accurate output irrespective of the age, colour, sex of a person.

REFERENCES

1. https://en.wikipedia.org/wiki/Demographics_of_India Demographics of India.
2. https://en.wikipedia.org/wiki/World_Happiness_Report Worlds Happiness Report.
3. https://www.nimh.nih.gov/health/publications/stress/index.shtml National Institute of mental Health, 5 Things You Should Know About Stress.
4. Multimodal detection of stress evaluation of the impact of several assessment strategies by Jonathan Aigrain, These de doctorat Date de soutenance: 05/12/2016 Universite Pierre et Marie Curie Ecole Doctorale numero 130: Informatique, Telecommunications et Electronique de Paris.
5. I. Lefter, G. Burghouts, and L. Rothkrantz. Recognizing stress using semantics and modulation of speech and gestures. IEEE Transactions on Affective Computing, 3045(c):1-1, 2015.
6. P. Kontschieder, M. Fiterau, A.Criminisi, and S. Rota Bulo. Deep neural decision forests. In International Conference on Computer Vision, pages 1467-1475, 2015.
7. J. M. Koolhaas, a. Bartolomucci, B. Buwalda, S. F. de Boer, G. Flugge, S. M. Korte, P. Meerlo, R. Murison, B. Olivier, P. Palanza, G. Richter-Levin, a. Sgoifo, T. Steimer, O. Stiedl, G. van Dijk, M. Wohr, and E. Fuchs. Stress revisited: a critical evaluation of the stress concept. Neuroscience and biobehavioral reviews, 35(5):1291-301, Apr. 2011.
8. J. R. Landis and G. G. Koch, The measurement of observer agreement for categorical data. Biometrics, 33(1):159-174, 1977.
9. R. S. Lazarus. Psychological stress and the coping process. 1966.
10. https://en.wikipedia.org/wiki/Random_forest#Algorithm.
11. https://medium.com/@williamkoehrsen/random-forest-simple-explanation-377895a60d2d.
12. https://en.wikipedia.org/wiki/Support_vector_machine#/media/File:Kernel_Machine.svg.
13. https://en.wikipedia.org/wiki/Support-vector_machine.
14. State-Trait Anxiety Inventory for Adults Sampler Set Manual, Instrument and Scoring Guide Developed by Charles D. Spielberger in collaboration with R.L. Gorsuch, R. Lushene, P.R. Vagg, and G.A. Jacobs Published by Mind Garden, Inc.
15. Vikash Yadav et al, "Role of IoT and Big Data Support in Healthcare" 5th InternationalConference on Computational Intelligence and Communication Technology (CICT-2019), Organized by Department of Computer Science & Engineering, ABES Engineering College, Ghaziabad, India on February 22-23, 2019.
16. Vikash Yadav et al, "A Biometric approach to Secure Big Data", International Conference on Innovation and Challenges in Cyber Security (ICICCS-2016), ISBN: 978-93-84935-69-6, pp. 75-79, February 03-05, 2016.
17. Vikash Yadav et al, "Big Data Analytics for Health Systems", IEEE International Conference on Green Computing and Internet of Things (ICGCIOT-15), ISBN: 978-1-4673-7909-0, pp. 253-258, October 08-10, 2015.

AUTHORS PROFILE

Dr. Vikash Yadav received his B.Tech (Computer Science & Engineering) degree in 2009 from Dr. Ambedkar Institute of Technology for Handicapped, Kanpur (U.P. India), M.Tech (Software Engineering) degree in 2013 from Motilal Nehru National Institute of Technology, Allahabad (U.P. India) and Ph.D (Computer Science & Engineering) degree from Dr. A.P.J Abdul Kalam University (Formerly U. P. Technical University) Lucknow, (U.P. India) in 2017. He is currently working as an Assistant Professor in the Department of Computer Science & Engineering, ABES Engineering College, Ghaziabad, India.

Rati Shukla received her B.Sc in 2006, MCA degree in 2009 from U.P. Technical University Lucknow and M.Tech (Computer Science and Engineering) degree in 2014 from Motilal Nehru National Institute of Technology, Allahabad (U.P. India). She is pursuing her Ph.D at GIS Cell Motilal Nehru National Institute of Technology, Allahabad (U.P. India).