Happiness Index Prediction of Students Using Machine Learning

N.Venkata Sailaja[1*], Kallam Lalithendar Reddy[2], Gottipati Aditya[3], Boppana Shashank[4], Vattikuti Hareen Sai[5,]

[1] Assistant Professor, Department of Computer Science and Engineering, VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad, India.
[2,5] Department of Computer Science and Engineering, VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad, India.
[3] Department of Computer Science, George Mason University, Fairfax County, Virginia.
[4] Department of CS & EMSE, George Washington University, Washington D.C, United States.

sailaja_nv@vnrvjiet.in

Abstract. The National Education Policy emphasizes the importance of providing joyful, holistic, and engaging education in schools. According to research, it is possible to teach children how to develop positive emotions, wellbeing, and engagement in their school experience. The goal is to promote happiness among students, teachers, and parents alike. There are various factors that can contribute to the happiness quotient, including the adoption of more activity-based learning methods rather than the traditional chalk and board approach, ensuring a safe and secure school environment, providing clean and hygienic toilets, creating cheerful and inviting classrooms and surroundings, allowing teachers to treat each other and students with respect and dignity, and promoting excellence in diverse areas, schools where parents are involved in some manner in the delivery of qualitative education, etc. By cultivating greater happiness, children can learn to discover and develop their strengths while boosting their sense of optimism. The development of a comprehensive framework for identifying and measuring the various factors that contribute to student happiness is a critical component in advancing the quality of education that schools provide. The identification of these critical indicators will allow for a predictive analysis that can identify outstanding schools and evaluate the different factors that make them better than others. In an effort to facilitate this process, we have undertaken the development of a website that provides a self-assessment tool for schools to use in optimizing their systems and improving overall student satisfaction.

Keywords: Happiness Index, Activity-based learning, safe and secure school environment, Student Satisfaction.

C. Kiran Mai et al. (eds.), *Proceedings of the Fourth International Conference on Advances in Computer Engineering and Communication Systems (ICACECS 2023)*, Atlantis Highlights in Computer Sciences 18,
https://doi.org/10.2991/978-94-6463-314-6_9

1 Introduction

Predicting the happiness index of students using machine learning is to develop a model that can accurately predict the level of happiness of students based on various factors such as their academic performance, social interactions, personal characteristics, and environmental factors. The model will use data of students, to learn and identify patterns that are associated with happiness levels.

The ultimate goal of this work is to create a predictive model that can help educational institutions identify students who are at risk of experiencing low levels of happiness and take necessary interventions to improve their overall well-being. This model can also be used to monitor the effectiveness of various interventions over time and improve the overall learning experience of students and improve the happiness index of the students.

The success of this work will be measured by the accuracy of the model in predicting the happiness index of students, and its ability to identify students who are at risk of experiencing low levels of happiness. The work will require data collection, data preparation, feature engineering, and selecting an appropriate machine learning algorithm to train the model.

The scope for happiness index prediction of students in schools using machine learning is vast and has the potential to bring about significant positive changes in the educational landscape. Here are some potential scopes of this application:

1. Early Identification of At-Risk Students
2. Customized Learning Experience
3. Monitoring the Effectiveness of Interventions
4. Creating a Positive School Culture

2 Related Work

Lexin You, et al [1] has proposed a study using machine learning algorithms to predict happiness based on demographic and socio-economic factors. The study analyzes survey data from over 2000 participants in the United States using regression analysis models such as linear regression, ridge regression, and lasso regression. The study also employs classification models such as random forest (RF) and support vector machine (SVM) to classify respondents as "happy" or "not happy" based on their survey responses . The study finds that machine learning algorithms can accurately predict happiness based on factors such as income, age, education level, and relationship status. Income is identified as the most significant predictor of happiness, followed by relationship status and age. Prashanthi et al. [7] the support vector machine model is found to be the most accurate in predicting happiness, with an accuracy rate of 77.4%. The study suggests that the use of machine learning algorithms can provide valuable insights into the factors that contribute to individual well-being and highlights the potential of

machine learning models in developing personalized interventions aimed at improving happiness and well-being.

Maria Fernanda Durón-Ramos et al [2] explores the relation between happiness orientation and university students engagement. The study involved a sample of 266 university students in Mexico, who completed surveys measuring their happiness orientation and engagement in academic activities. The study used a quantitative research methodology and several machine learning models to analyze the data. Hierarchical linear regression was used to identify the contribution of happiness orientation to engagement, while decision tree and artificial neural network models were used to predict engagement based on happiness orientation.

The results of the study revealed that happiness orientation significantly predicted university students' engagement in academic activities. With a classification accuracy of 76.7%, the decision tree model was shown to be the most reliable in predicting engagement. The artificial neural network model also demonstrated high accuracy, with a classification accuracy of 73.2%. The study suggests that happiness orientation is an important factor in predicting university students' engagement in academic activities. The findings also highlight the potential of machine learning models in predicting engagement based on happiness orientation, providing insight into the factors that influence student engagement in higher education.

Akane Sano et al [3] predicted the mood, based on daily behaviors and previous sleep history of people through the exploration of machine learning's capabilities. The study used a dataset collected from 29 participants who wore a wristband sensor to collect data on their daily activities, sleep, and mood. In order to forecast mood based on the gathered data, the study employed three machine learning models, including linear regression, decision trees, and random forests. According to the findings, the random forest algorithm was the most accurate in predicting mood from the chosen characteristics. The study found that sleep duration, social interaction, and physical activity were the most important features in predicting mood. The study suggests that machine learning algorithms can be used to predict mood based on daily behaviors and sleep history, which could lead to new interventions and recommendations for individuals who may be at risk of low mood.

Rabia et al [9] proposed a technique for calculating National Happiness Index (NHI) of Pakistan using Roman Urdu messages posted on social media platforms. They classified a Roman Urdu sentiment that has three sub-opinions using a variety of classifiers, including rule-based classifiers, 3-RBC, and supervised machine learning techniques using unigram, bigram, and parts of speech tags in order to obtain the NHI score

Nam, et al [5] research paper investigates the factors that influence the happiness index of hospital nurses in South Korea. The study involved 677 hospital nurses who completed a survey that included measures of job satisfaction, social support, self-effi-

cacy, and happiness. The authors used regression analysis to identify significant predictors of nurse happiness, and found that job satisfaction, social support, and self-efficacy were significant factors influencing the happiness index of hospital nurses. Specifically, nurses who reported higher levels of job satisfaction, social support, and self-efficacy had higher levels of happiness. In contrast, nurses who reported lower levels of job satisfaction, social support, and self-efficacy had lower levels of happiness.

Zongwen Fan et al [6] have suggested a hierarchical machine learning model for predicting happiness levels based on demographic and psychological factors. The model uses a dataset of 18,000 survey responses from the Chinese population and is based on a tree-based approach that utilizes different algorithms at each level of the hierarchy. The first level of the hierarchy predicts overall happiness level based on demographic factors such as age, gender, and income. The second level predicts happiness within each demographic group based on psychological factors such as social support, self-esteem, and life satisfaction. The model was trained and evaluated using various performance metrics such as accuracy, precision, and recall. The findings demonstrated that in terms of forecasting happiness levels, the hierarchical model performed better than more established machine learning models like linear regression and decision trees. Laura et al. [4] the accuracy of the model was assessed using a 10-fold cross-validation technique, and the overall accuracy achieved was 85.4%. The precision and recall values were also reported for each level of the hierarchy, which indicated the model's ability to predict happiness within each demographic group accurately.

3 Methodology

3.1 Data Set

Our study aimed to predict the student happiness index of government schools in the city, taking into account the perspectives of students, parents, and teachers. The questionnaires were designed to collect data on student, teacher, and parent satisfaction with the school environment, education, sports facilities, food and drinking facilities, teaching quality, and overall happiness. To collect the necessary data, we traveled to every government school in the city. We distributed survey forms to students, teachers, and parents, explained the purpose of the survey, and collected additional data on school facilities through interviews and observations. The team's commitment to accuracy and diligence in data entry was essential for the success of the work, as it required a high level of attention to detail and expertise in data management and analysis. The final dataset contains 2145 student entries, 263 parents entries, and 49 teachers entries.

The proposed solution is visually depicted in the figure 1, showcasing the fundamental framework of the system. It comprises three crucial modules, each playing a significant role in the functionality of the system. The first module, the website module, acts as the user interface, providing a seamless experience for the end-user. The second module, the storage, manages and stores all user data, ensuring efficient access and retrieval. The third and final module, the machine learning module, employs sophisticated algorithms to process vast amounts of data, providing valuable insights into user

behavior and preferences. Together, these modules create a robust solution capable of meeting complex user demands.

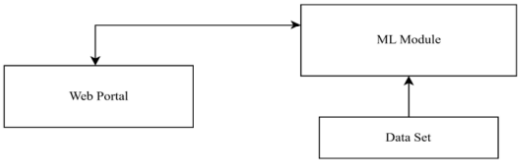

Fig. 1. High Level View of Framework

3.2 Web Portal

This module is the first point of contact in the system for the users. The module is designed as a web application that provides an interface between the services and the users. The development of the Web application involved the utilization of Streamlit, an open-source framework well-suited for creating data-rich web applications. Drawing on this powerful tool, the application was further divided into multiple sections, each intended to offer a comprehensive description of the Happiness Index. These sections are designed to provide a detailed analysis of this vital measure of student well-being, giving users the information, they need to make more informed decisions and better understand the world around them. It offers the unique opportunity for users to engage with the data by taking part in a comprehensive form consisting of 10-15 questions, which in turn yields a personalized happiness index score. The questions are unique to the type of user and there are three sections of questions for the students, teachers. In addition to the sections designed for parents, the web application also includes sections tailored to the needs of educators and administrators. This information can be used to inform lesson planning and classroom management strategies, helping to create a more positive and supportive learning environment for students. By aggregating data from multiple schools, the application can provide a comprehensive view of the happiness index of students across an entire district or region. This can be used to identify areas where additional resources or support may be needed, as well as to track progress over time as improvements are made. The application analyzes data trends to identify common issues affecting student wellbeing, aiding in policy decisions and developing targeted interventions for improved student outcomes.

3.2.1 Student Module

The student module is responsible for gathering information from the user by presenting a series of questions which is a form that delves into the intricacies of their educational journey. To gain a comprehensive understanding of their situation, the module focuses on various aspects such as the level of education they are undertaking, the challenges they are facing, and their current standing.

3.2.2 Parent Module

The parent module is also responsible for collecting relevant data from parents pertaining to their child's academic growth, emotions, and overall happiness index. This information is collected through a series of questions presented in a user-friendly form format. The data collected provides insights into the child's overall progress, as seen through the eyes of the parent.

3.2.3 Teacher Module

. The teacher module also plays a vital role in predicting the happiness index within a school. Through the use of a set of questions tailored to assess the well-being of educators, users gain insight into how the institution values its teachers and their overall satisfaction. As those responsible for imparting knowledge to students,

3.2.4 Feedback Module

The feedback module in the portal provides a comprehensive analysis of government schools, generating rankings and student happiness indexes. It highlights areas for improvement to improve education quality and learning environment. The module visually represents large datasets, enabling stakeholders to understand student mental health and education status. This collaboration helps improve education quality and student welfare.

3.3 ML Module

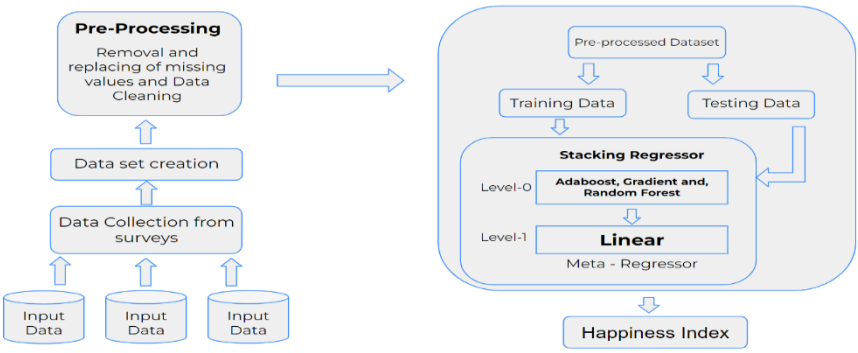

Fig. 2. System Architecture

The Machine Learning module is a major module Tan et al. [8] that is designed to carry out various ML tasks and efficiently run different models for giving the most accurate prediction possible. To achieve this, the module initially receives a categorical dataset

that is then transformed into numeric values through the application of Label Encoding methodology, a powerful Ml technique, making it easier to compute and visualize the happiness index. With precise and reliable results, this module is a crucial component of the data analysis process and an essential tool for any individual or organization seeking to analyze and predict their data 33 accurately. The ML module is a crucial part of the education portal, seamlessly connected with other modules such as the student, teacher, and parent modules. Users engage with the portal by answering questions displayed in their respective modules. Once users complete their responses, they can click the predict button to see their happiness index. However, behind the scenes, the ML module utilizes Stacking regression models to provide accurate predictions. This technique in figure 2 combines different models to improve its accuracy, ensuring users receive the best possible results. With its impressive technology, the education portal provides a comprehensive and reliable user experience.

4 Results

Our application leverages the power of machine learning to analyze a range of features related to the happiness index. We use a comprehensive dataset collected through a survey of government schools to provide our Stacking Regressor algorithm (Sailaja et al.) [10] [11] with the necessary data to accurately analyze the relationship between a student's happiness and the various factors that impact it. By utilizing a range of algorithms, we are able to make insightful predictions regarding happiness levels that are consistent with actual values. Our work aims to identify the core components of happiness that can be used to improve student well-being and overall educational outcomes. To achieve this, we are utilizing sophisticated analytical tools such as the Student's Happiness Index and Python programming, and Tableau allowing us to explore and visualize key data from our extensive dataset. The results are given in table 1, table 2, table 3.

Table 1. Student Happiness Prediction Results

Metrics	Score
Coefficient of Determination (R2)	0.978
Mean Square Error (MSE)	0.037
Mean Absolute Error (MAE)	0.124
Root Mean Square Error (RMSE)	0.193

Table 2. Parent Happiness Prediction Results

Metrics	Score
Coefficient of Determination (R2)	0.982
Mean Square Error (MSE)	0.014

Mean Absolute Error (MAE)	0.062
Root Mean Square Error (RMSE)	0.121

Table 3. Teacher Happiness Prediction Results

Metrics	Score
Coefficient of Determination (R2)	0.869
Mean Square Error (MSE)	0.166
Mean Absolute Error (MAE)	0.297
Root Mean Square Error (RMSE)	0.408

Our system's Web-Portal component is a Stream-lit-based website. The website is thoughtfully created to satisfy all the criteria.

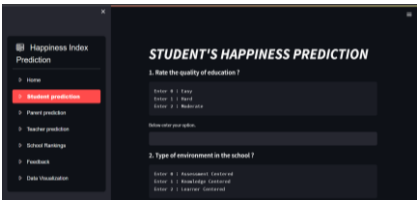

Fig. 3. Student's Happiness Prediction

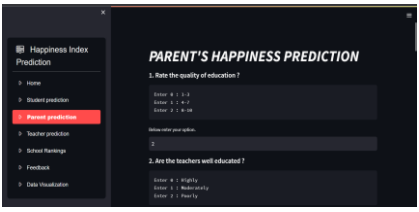

Fig. 4. Parent's Happiness Prediction

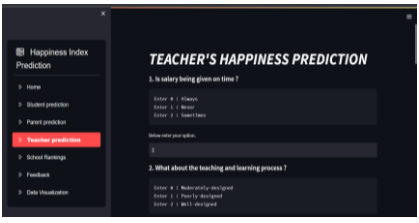

Fig. 5. Teacher's Happiness Prediction

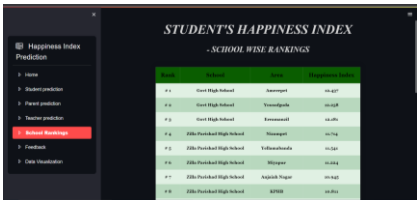

Fig. 6. School Rankings

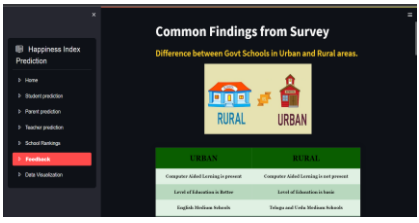

Fig. 7. FeedBack

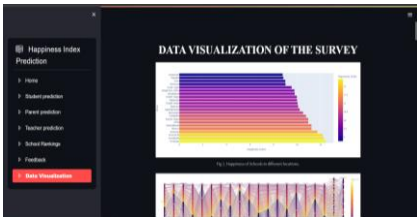

Fig. 8. Data Visualization

The Student happiness Prediction is the page where students are allowed to check out their happiness index by answering the questionnaire. Fig. 3, Fig. 4, Fig. 5, Fig. 6, Fig. 7, and Fig. 8 show the user interface of the website.

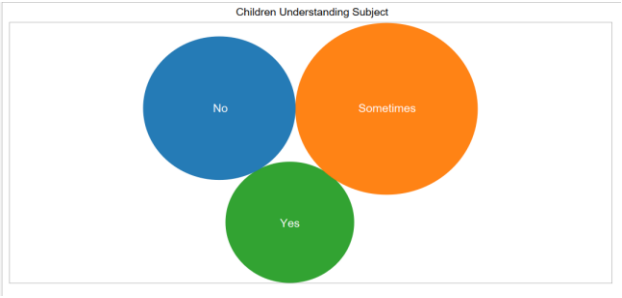

Fig. 9. Bubble chart on how well children are understanding the subject

Fig. 9 depicts the comfortableness of the student in understanding the subjects. Figure describe that only few students are comfortable in understanding all the subjects; this might be due to improper explanation of the faculty or the language that is being teached.

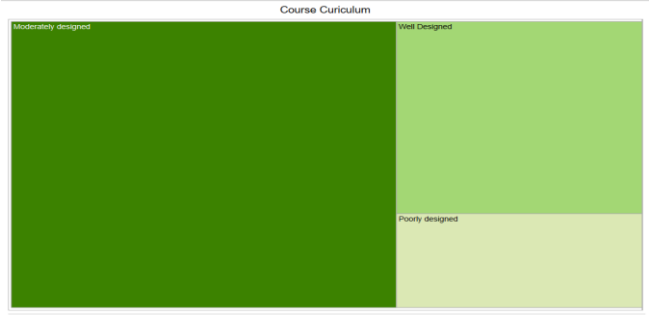

Fig. 10. Tree Map on Course and Curriculum

The above Tree map fig. 10 describes the course curriculum that is designed to teach for the students from the perspective of teachers. But the results also show that teachers also feel that the course is well designed and covers all the important topics. Also with making it easy for students to learn, very few teachers feel it is a poorly designed course.

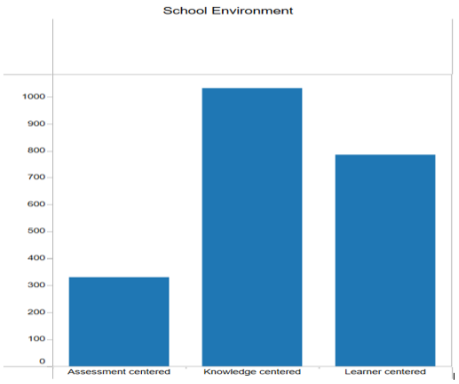

Fig. 11. Bar chart on Type of environment in school

The fig. 11 describes the environment that is observed by the students of the school. The school environment is divided into 3 categories based on the way teachers teach the students.

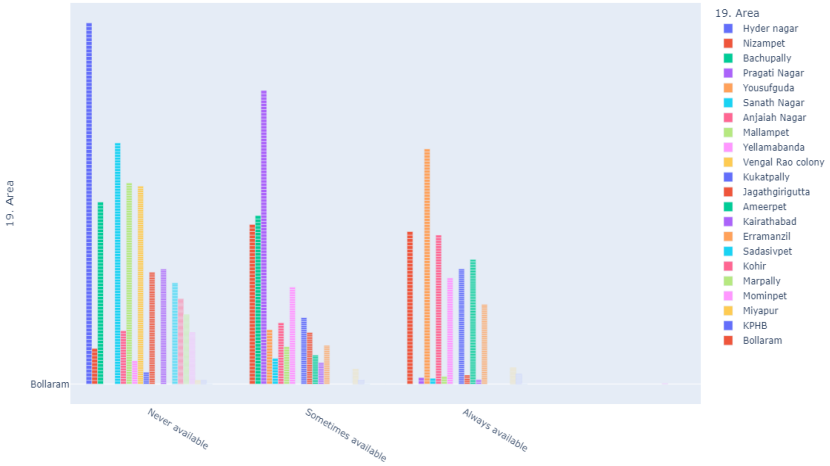

Fig 12. Grouped bar chart on availability of water facilities in school

The fig. 12 describes about the availability of water facilities in the different area of school, the results say that the school in yousufguda always make the water facility available every time and there other other schools where this was not considered. Being an important basic need of every student school are needed to provide this facility and hydernagar school failed to provide that facility and which might be the reason for some decrement in the happiness index of student.

5 Conclusion

This study was conducted to predict the Happiness Index of students, parents and teachers. And to identify the predictive factors affecting the happiness index of students by identifying the level of courses, teacher's competitive attitude, basic facilities provided by the school, infrastructure, etc. and identifying the relationship between them. As a result of the study, the happiness index of students was a moderate score, and it was analyzed that there was a difference in the happiness index in different areas. The teachers and parents happiness index also indicates that they are having various issues with the school indicating that happiness scores can be improved. Through the above results, it is necessary to make some initiatives to develop government schools to improve the happiness index of students, like:

1. The buildings of all schools must be checked and repaired to give students and teachers a safe environment.
2. Hire more teachers.
3. Please train teachers as per latest methods.
4. All schools must have computer labs integrated with-in curriculum.
5. Mid-day meals must be top quality and must provide necessary nutrients.
6. Syllabus must be improved and program delivery must be according to the latest technologies and the teacher must own the process.
7. Regular quality checks must be done by independent agencies anonymously.
8. Very strict laws must be ensured for all teaching and mid-day meal staff for actual program delivery and eradicate pilferage.

Future Scope

As we are providing a framework to predict the Happiness Index of Students using ML, our future scope would be:

1. Make school management aware of issues facing by students: Government schools lack basic facilities like drinking water, mid-day meal schemes, and clean toilets etc. Addressing these issues could lead to improved results.
2. Weekend Classes: Special weekend classes for children can enhance learning through innovative approaches, incorporating academic or non-academic content, aiming to share knowledge and improve their learning experience.
3. Improve English Standards: Teachers should teach students quality English to enhance communication skills, including listening, understanding, pronunciation, fluency, and clear expression, and may receive training to improve spoken English.
4. Improve Teacher's Competitiveness: The concept of learning, unlearning, and relearning is crucial in teaching, requiring a shift from rote learning to conceptual methods, enhancing teaching abilities for better results.
5. Make schools attractive to inspire children to come to school: Enhancing school quality fosters lifelong learning, enabling students to remain enrolled and contribute to nation-building through their learning and productivity.

References

1. You, L. (2021, March). Utilizing machine learning to predict happiness index. In 2021 2nd International Conference on E-Commerce and Internet Technology (ECIT) (pp. 233-238). IEEE
2. Durón-Ramos, M. F., & García-Vázquez, F. (2018). Orientation to Happiness as Predictor of University Students' Engagement. International Journal of Evaluation and Research in Education, 7(4), 294-298
3. Sano, A., Amy, Z. Y., McHill, A. W., Phillips, A. J., Taylor, S., Jaques, N., ... & Picard, R. W. (2015, August). Prediction of happy-sad mood from daily behaviors and previous sleep history. In 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 6796-6799). IEEE
4. Musikanski, Laura, et al. "Happiness index methodology." Journal of Sustainable Social Change 9.1 (2017): 2.
5. Nam, Moon Hee, and Young Chae Kwon. "Factors influencing happiness index of hospital nurses." Journal of Korean academy of nursing administration 19.3 (2013): 329-339.
6. Fan, Zongwen, Fenlin Wu, and Yaxuan Tang. "A hierarchy-based machine learning model for happiness prediction." Applied Intelligence (2022): 1-10.
7. Prashanthi, B., & Ponnusamy, R. (2019). Future Prediction of World Countries Emotions Status to Understand Economic Status using Happiness Index and SVM Kernel. Future Ozakar, R., Gazanfer, R. E., & Hanay, Y. S. (2020). Measuring Happiness Around the World Through Artificial Intelligence. Journal of Business Research and Management, 2(2), 1-11.
8. Tan, Y., Singhapreecha, C., & Yamaka, W. (2020). Applying Machine Learning to Predict Happiness: A case study of 20 Countries. Mukht Shabd Journal, 9(6), 3433-3437.
9. Habiba, Rabia, Dr Muhammad Awais, and Dr Muhammad Shoaib. "A technique to calculate national happiness index by analyzing Roman Urdu messages posted on social media." ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP) 19.6 (2020): 1-16.
10. N. V. Sailaja, M. Karakavalasa, M. Katkam, D. M, S. M and D. N. Vasundhara, "Hybrid Regression Model for Medical Insurance Cost Prediction and Recommendation," 2021 IEEE International Conference on Intelligent Systems, Smart and Green Technologies (ICISSGT), 2021, pp. 93-98, doi: 10.1109/ICISSGT52025.2021.00029.
11. Sailaja N.V., Yelamarthi M., Chandana Y.H., Karadi P., Yedla S. (2021) Early Detection of Sepsis on Clinical Data Using Multi-layer Perceptron. In: Mai C.K., Reddy A.B., Raju K.S. (eds) Machine Learning Technologies and Applications. Algorithms for Intelligent Systems. Springer, Singapore.