Efficiency Predictor: Predicting the Consumption Efficiency of Humans by Machine Learning Technique

[1]Dr. S. Vidya, [2]Dr. T. Veeramakali, [3]Dr. N. C. Brintha, [4]S. Inbasekar
[1]CSE, Sri Sairam Institute of Technology
Chennai, India
e-mail: vidya.cse@sairamit.edu.in
[2]Data Science and Business Systems, SRM Institute of Science & Technology
Chennai, India
e-mail: veeramat@srmist.edu.in
[3]CSE, Kalasalingam Academy of Research and Education
Srivilliputhur, India
e-mail: brinthachris2k@gmail.com
[4]CSE, Sri Sairam Institute of Technology
Chennai, India
e-mail: sit20cs033@sairamtap.edu.in

Abstract—As computer science advances and integrates with statistics in the field of machine learning, the predictability of future events is increasing. Our project focuses on leveraging this domain to forecast human performance using a minimal set of attributes, thereby reducing the need for extensive labels. As present solutions in machine learning helped humanity to predict natural events there is no accurate existing solution to predict the same for human beings. Human efficiency may include the development of an individual or the development of a team or collaboration. Making progress in a work without knowing the success rate might be a challenge as the final output may or may not give the expected results. The amount of hard work engaged in work that may fail in the future causes a great loss of time and energy. The involvement of computers integrated with the statistical models motivates and helps to predict the final output. So, we have taken the initiative to predict the future performance of a person in a more accurate and precise manner. This project aims to predict the consumption efficiency performance of a person using a machine learning algorithm by Ensemble-based Progressive Prediction.

Keywords - Ensemble Based Progressive Predictor, Naive Bayes, C4.5 Classification, Performance Prediction, Future Performance.

I. INTRODUCTION

Nowadays the performance of human beings is crucial in any sector as the world is moving towards robots for more performance. But before a decade human performance is not a concern in industries since the technology has not evolved in such a way that it can overcome humans in any aspect. It completely changed after the evolution of AI, IoT, etc. These technologies have the capability to overcome the performance of human beings in every aspect. So now it is a concern for humans in the aspect of their performance in their respective works. Hence, Machine Learning (ML) is proposed for this problem. With the help of ML, humans can able to know what will be their future performance in their work or respective fields. So, we develop an ML model that will predict the human's future performance in an optimized and more accurate manner. It may give the wrong prediction for a few individuals, to overcome that we develop an EPP model that will be able to overcome the above errors. the applicable criteria that follow.

II. APPROACHES OF PREDICTION OF HUMAN PERFORMANCES

In today's world, optimizing human consumption performance efficiency is a pressing concern for state sanitary organizations and similar institutions dedicated to public health and safety. To address this critical need, we are currently in the process of developing a Machine Learning (ML) model tailored to predict and enhance the performance of various aspects related to human consumption.

This initiative goes beyond the boundaries of a specific industry or sector, with the potential to benefit a wide range of organizations, including state sanitary organizations, etc.

As our initial step, we have undertaken the development of a sophisticated ML model, known as the Ensembled base Progressive Predictor (EPP) model, which has been diligently trained using data collected during the prior phase of this project. The primary objective of this model is to predict and optimize human consumption performance, with a particular focus on public health and safety standards upheld by sanitary organizations.

The data set used for training comprises various performance metrics and parameters related to human consumption practices. This data set empowers our trained model to provide accurate predictions regarding human consumption efficiency for individuals who may not have been part of the initial data set.

The EPP model, previously designed in the project's earlier phase is now fully equipped to predict and optimize the performance of various aspects related to human consumption. State sanitary organizations and other similar entities will find this tool invaluable for assessing and improving public health and safety standards, enhancing resource allocation, and making informed decisions based on data-driven insights.

This application of ML holds immense potential for transforming the operations of sanitary organizations and elevating the standards of health and safety in our communities. By harnessing data related to various consumption parameters, organizations can proactively address issues, allocate resources effectively, and elevate overall human consumption performance efficiency.

II.I PREDICTION BASED ON C4.5 ALGORITHM

C4.5 is sometimes termed a statistical classifier due to its ability to generate decision trees for categorization, similar to ID3. Like ID3, C4.5 leverages the concept of information entropy to build decision trees from a training dataset. At each node of the tree, C4.5 chooses the data feature that optimally separates its sample collection into subsets that are more concentrated in one or more classes. The splitting criterion used is the normalized information gain, which quantifies the difference in entropy. The feature with the highest normalized information gain is then chosen to make the decision.

II.II PREDICTION BASED ON NAIVE BAYES ALGORITHM

In order to address classification issues based on the Bayes theorem, the Naive Bayes method employs supported learning. For text categorization problems requiring a large training dataset, it is primarily utilized. The Naive Bayes Classifier is a simple and effective classification technique that can be used to quickly develop machine learning models that can make predictions. Given that it is a probabilistic classifier, its predictions are predicated on the probability that an item will occur. For spam filtration, sentiment analysis, and article categorization, the Naive Bayes Algorithm is frequently employed.

II.III REGRESSION TREE

Binary recursive partitioning is an iterative procedure that initially divides data into branches or partitions. As the algorithm proceeds up each branch, it further splits each partition into smaller groups, ultimately forming a regression tree. To establish the tree's structure, the pre-classified data in the Training Set are initially separated into the same divisions. Subsequently, the algorithm divides the data into the first two branches, considering potential binary splits on each field. The method selects the partition that minimizes the sum of squared deviations between the two distinct partitions. This same principle is then applied to divide each of the new branches.

II.IV ENSEMBLED-BASED PROGRESSIVE PREDICTOR

The EPP is a function-dependent algorithm that runs on top of the Gaussian Naive-Bayes, C4.5, and Regression Tree. The student data is divided into categories and given a random variable according to their performance. Because the output of one function may also be the output of another, the data frame for each process flow will be updated continuously. The dataset's pre-classification and subsequent processing enhance this algorithm. Each of the functions listed below consists of two main parts: input and output. The input will be a dynamically updated data frame according to predetermined criteria, and the output will be the anticipated result. In addition to the main module, the supporting modules will help with more precise predictions.

TABLE 1: COMPARISON OF HUMAN PERFORMANCE PREDICTION APPROACHES

Metrics	EPP	REGRESSION
MAE	0.09	6.12
MSE	0.19	1.33
RMSE	0.3	2.47
RAE	0.028	0.22
RRSE	0.12	0.99
MEAN	10.66	10.66
MAE/MEAN	1.87	12.531
RMSE/MEAN	2.90	23.205

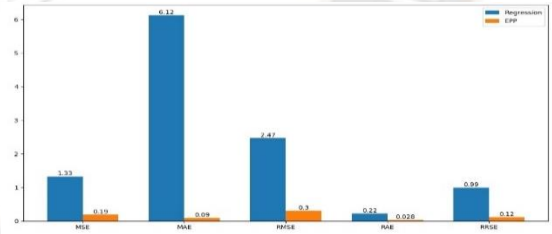

FIGURE 1: COMPARISON GRAPH OF VARIOUS METRICS BETWEEN REGRESSION AND EPP

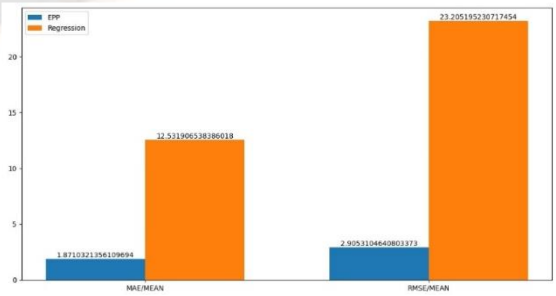

FIGURE 2: COMPARISON GRAPH OF MAE/MEAN AND RMSE/MEAN METRICS BETWEEN REGRESSION AND EPP

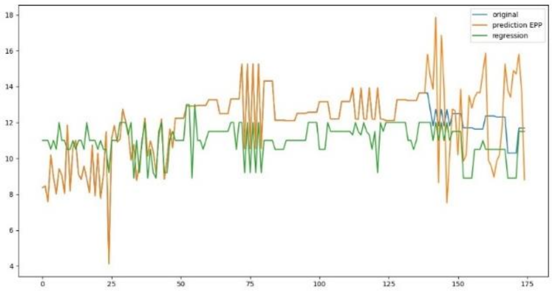

FIGURE 3: COMPARISON GRAPH OF PREDICTION VALUES OF VARIOUS PREDICTION APPROACHES

III FUTURE ENHANCEMENT

Since this research is case-specific, we will eventually create a generic method that can be used to forecast future performance in any field without developing a unique algorithm for every issue or industry.

IV CONCLUSION

This project presents an efficient method to predict the future performance of humans in an optimized and accurate manner. This enables industries, educational institutions, etc., to analyze their employees, student, etc., performance. Hence, the performance can be improved in terms of our prediction model.

REFERENCES

[1] Dr.S. Vidya, S. Inbasekar, U. Vasanthakumaran AND A. Venkatesh. “Performance Predictor: Predicting the future performance of students by Machine Learning approach” 2022.SCOPUS. eISSN 1303-5150 doi:10.14704/nq.2022.20.10. NQ551078.

[2] C.-Y. Ko and F.-Y. Leu, "Examining Successful Attributes for Undergraduate Students by Applying Machine Learning Techniques," IEEE Transactions on Education, vol. 64, no. 1, February 2021, pp. 50 - 57, 2021.

[3] H. H. a. A. V. R. L. Ahadi, "Exploring machine learning methods to automatically identify students in need of assistance," in ICER '15: Proceedings of the eleventh annual International Conference on International Computing Education Research, Omaha Nebraska USA, 2015.

[4] C. Y. K. a. F. Y. Leu, "Analyzing attributes of successful learners by using machine learning in an undergraduate computer course," in IEEE, Krakow, Poland, 2018.

[5] K. Q. a. S. Bergin, "Programming: Further factors that influence success," St Catharine's College, University of Cambridge, UK, 2016.

[6] S. K. a. D. Kanellopoulos, "Association rules mining: A recent overview," International Transactions on Computer Science and Engineering, vol. 32, no. 1, pp.71-82, 2006.

[7] J.-L. H. a. K. Zhang, "Revealing online learning behaviors and activity patterns and making predictions with data mining techniques in online teaching," MERLOT Journal of Online Learning and Teaching, vol. 4, no. 4, pp. 426-436, 2008.

[8] K. E. B. S. K. a. S. B. A. Ezen-Can, "Unsupervised modeling for understanding MOOC discussion forums: A learning analytics approach," in Association for Computing Machinery, New York, NY, United States, 2015.

[9] M.-I. L. J.-M. L. a. S. V. C. Romero, "Predicting students’ final performance from participation in online discussion forums," COMPUTERS & EDUCATION, vol. 68, no. -, pp. 458-472, 2013.

[10] M. C. W. S. W. M. A. Y. L. a. W. H. J.-L. Hung, "Identifying at-risk students for early interventions—A time-series clustering approach," IEEE Transactions on Emerging Topics in Computing, vol. 5, no. 1, pp. 45-55, 2015.

[11] A. M. S. A. A. a. N. G. H. R. Asif, "Analyzing undergraduate students’ performance using educational data mining," Computer Education, vol. 113, no. -, pp.177-194, 2017.

[12] S. A. a. C. Conati, "Unsupervised and supervised machine learning in user modeling for intelligent learning environments," in Association for Computing Machinery, New York NY United States, 2007.

[13] B. J. Zimmerman, "Investigating self-regulation and motivation: Historical background methodological developments and future prospects," American Educational Research Journal, vol. 45, no. 1, pp. 166-183, 2008.

[14] R. L. A. M. Dembo, "The relationship between self-regulation and online learning in a blended learning context," International Review of Research in Open and Distance Learning, vol. 5, no. 2, pp. 1-16, 2004.

[15] D. H. Schunk, "Self-regulated learning: The educational legacy of Paul R. Pintrich," Educational Psychologist, vol. 40, no. 2, pp. 85-94, 2005.

[16] A. W. A. F. H. A. Kitsantas, "Self-regulation and ability predictors of academic success during college: A predictive validity study," Journal of Advanced Academics, vol. 20, no. 1, pp. 42-68, 2008.

[17] Z. A. P. P. S. B. Kotsiantis, "Machine learning: A review of classification and combining techniques," Artificial Intelligence Review, vol. 26, no. 3, pp. 159-190, 2006.