International Journal of Research Publication and Reviews

Life Expectancy Prediction Using Machine Learning

Bongi Pooja[1], Ms .U .Archana[2]

[1](21341a1223), [2]Assistant Professor

Department of Information Technology, GMR Institute of Technology, G.M.R. Nagar, Rajam-532127, A.P

ABSTRACT

Life expectancy refers to the average number of years a person or a population can expect to live based on various demographic and statistical factors. It provides an estimate of the

typical lifespan, considering factors like age, gender, location, and current mortality rates. Life expectancy prediction involves complex algorithms and statistical models that take into account various factors and data sources to estimate how long a person or a population is likely to live. The commonly used algorithms in life expectancy prediction is Decision trees, KNN, and Random forest. In the existing system, the life expectancy can be predicted using decision tree .But there are some limitations observed in the life prediction. They are some drawbacks of random forest algorithm it takes much time for training as it combines a lot of decision trees to determine the class. To overcome the limitations, here proposed random forest algorithm which works on a life expectancy prediction. The accuracy of random forest in life expectancy prediction is more.

Keywords: Decision tress, KNN, life expectancy, random forest.

I, INTRODUCTION

Human an incredible creation of god. Every creature in the world has a limited life span, to achieve something in the world. We have a limited life span to survive in the current world. To preserve our self from the consequences, even though lot of inventions has been made by human, to prevent from diseases is a major question mark. Life span prediction has a greater impact in our modern society because of our food habits, different types of diseases and environmental conditions. Investigations about the life span of vertebrates have been made, except the human (HOMO SAPIENS). It is an emerging research area that is gaining interest but involved lot of challenges due to the limited amount of resources (i.e., datasets) available. Life expectancy prediction using Machine Learning (ML) is a fascinating and important application of data science and artificial intelligence. It involves the development of predictive models to estimate an individual's or a population's expected lifespan based on various factors and data sources. life expectancy prediction using ML leverages advanced data analysis techniques to provide estimates of how long individuals are likely to live based on a combination of demographic, health, and environmental factors. This has significant implications for healthcare planning, policy-making, and personalized health recommendations

2. LITERATURE SURVEY

Following research papers are studied in details to understands the proposed recommendation technique and experimental result for predicting the output .

[1] Kasichainula Vydehi, Keerthi Manchikanti, and T. Vijay Kumar. "Life Expectancy Prediction through Analysis of Immunization and HDI Factors using Machine Learning Regression Algorithms." In 2021 4th International Conference on Inventive Computation Technologies (ICICT), 2021.

Kasichainula vyedhi, Keerthi and vijaykumar, proposed the idea that human beings are remarkable creations   and that all living creatures have a limited lifespan.

2.1 OBJECTIVE:

✓ It highlights the significance of trying to achieve something meaningful during our lifetime and the challenges humans face in preserving their health and well-being, especially in the face of various diseases and environmental conditions.

✓ Additionally, it suggests that despite numerous inventions and advancements in medicine, preventing diseases remains a significant challenge.

✓ The mention of "life span prediction" implies the importance of estimating how long an individual is likely to live, considering factors such as lifestyle, health habits, and environmental influences.

[2] Mahumud R.A., Hossain, G., Hossain, R., Islam, N. and Rawal, "impact of life Expectancy on economic Growth and health care Expenditure in Banglesh.," Universal journal of public health,2019.

✓ Mahumud and Hossain proposed that the correlation between attributes like diseases, gender, ages.

3. OBJECTIVE:

✓ They emphasize the importance of examining the correlation between various attributes, such as diseases, gender, age, and environmental factors, in relation to predicting human lifespan. They specifically mention the use of the Decision algorithm for this purpose.

✓ The accuracy rate of decision tree algorithm is 91%

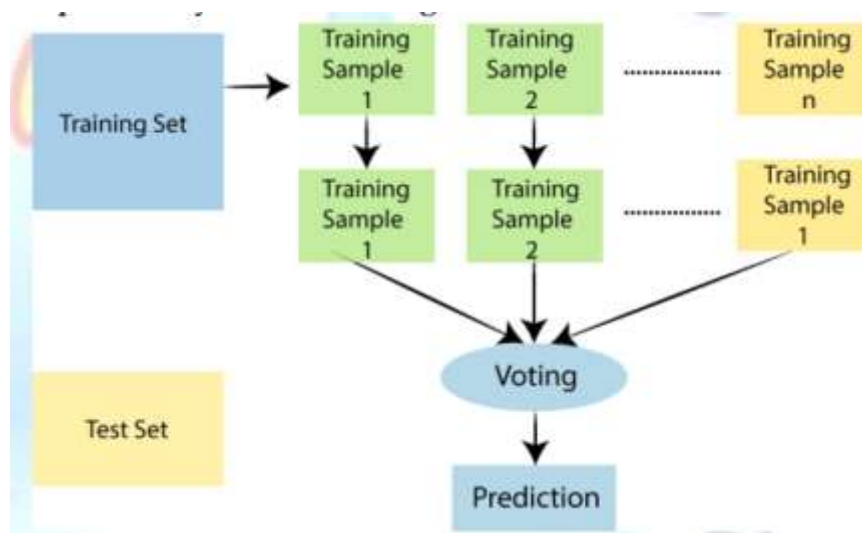✓ the primary goal of the research is to find or predict human lifespan with greater accuracy.

6. METHODOLOGY

Reference Paper-1: Life Expectancy Prediction using Machine Learning

6.1 Novelties :

✓ Random forest

✓ Evaluation Metrics for Decision tree

✓ Challenges in Research

✓ Human Lifespan Prediction

✓ ANN, and RNN

6.2 ARCHITECTURE:

6.3 PROCEDURE

Data Preparation

Collect a labeled dataset for your task.

Split the data into training and testing sets.

Decision Tree Training

Randomly select a subset of the training data (Bootstrapped Dataset).

Randomly select a subset of features (m features out of n, where m << n).

Decision Tree Predictions

For each tree in the forest:

Use the trained decision tree to predict outcomes for the test dataset.

Aggregate Predictions (Classification, Regression)

The most popular class label among the trees is the final prediction.

Formula: Final Prediction = mode(predictions from all trees)

Tuning Hyperparameters

Experiment with hyperparameters like the number of trees, maximum depth of trees, and minimum samples per leaf to optimize the model's performance.

Reference Paper-2: Life Expectancy Prediction through Analysis of Immunization and HDI Factors using Machine Learning Regression Algorithms

Novelties :

- ✓ Naïve bayes
- ✓ Linear regression
- ✓ Support vector machine
- ✓ Random forest
- ✓ Decision trees

Architecture:
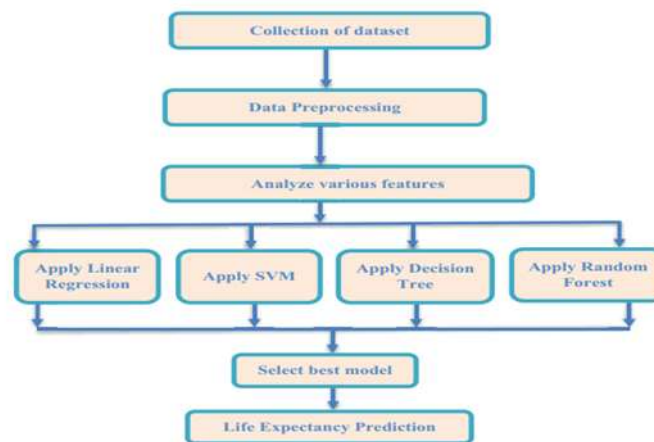
- ✓ **Apply Linear Regression**: This step involves applying a linear regression model to the data. This model assumes a linear relationship between the independent variables (features) and the dependent variable (life expectancy).

- ✓ **Apply SVM:** This step involves applying a support vector machine (SVM) model to the data. SVMs are a type of machine learning algorithm that can be used for classification and regression tasks. They are known for their ability to handle nonlinear relationships between variables.

- ✓ **Apply Decision Tree**: This step involves applying a decision tree model to the data. Decision trees are a type of machine learning algorithm that can be used for classification and regression tasks. They are easy to interpret and can handle both categorical and numerical data.

7. RESULT:

[1] Kasichainula Vydehi, Keerthi Manchikanti, and T. Vijay Kumar. "Life Expectancy Prediction through Analysis of Immunization and HDI Factors using Machine Learning Regression Algorithms." In 2021 4th International Conference on Inventive Computation Technologies (ICICT), 2021.
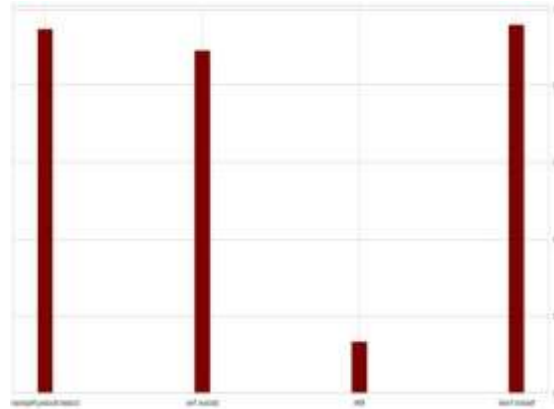
- The baseline model and doctors were correct in 20% of the cases, with a margin of 33% around the actual life expectancy as the target. The keyword model achieved an accuracy of 29% in its prognoses.

The best-performing model used a skip-gram architecture, a cut-off frequency boundary of 10, a window size of 5, and 300 dimensions. It produced the best results on the analogy task.

The authors discussed their results in terms of precision and recall, rather than sensitivity and specificity, due to the skewed distribution of the data.

[2] Mahumud R.A., Hossain, G., Hossain, R., Islam, N. and Rawal, "impact of life Expectancy on economic Growth and health care Expenditure in Banglesh.," Universal journal of public health,2019.
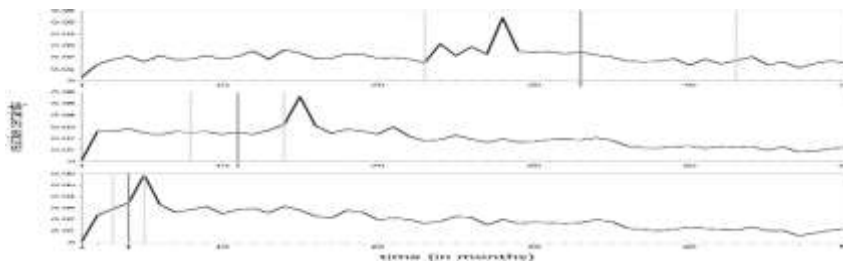
- Frailty Indices (FIs) in mice were measured longitudinally until death, and machine learning was used to develop two clocks: the FRIGHT (Frailty Inferred Geriatric Health Timeline) clock, which predicts chronological age, and the AFRAID (Analysis of Frailty and Death) clock, which accurately predicts life expectancy and the efficacy of lifespan-extending interventions.

- The study used random forest algorithms for modeling, which are robust to outliers and noise and allow for complex non-parametric modeling. These models showed improved prediction over simple linear regression and were able to detect variation in predicted chronological age for mice of the same actual age, representing biological age

8.CONCLUSION

In this paper a system of the human lifespan can be predicted earlier. By employing data through datasets, the correlation between attributes like diseases, gender, ages and environmental factor are monitored. The Random Forest algorithm is achieved in order to forecast the human lifespan with more precise. The advantage of Random Forest algorithm, gives more flexibility without obtaining the processed data and accurate. Thus, We have analyzed the lifespan among human beings based on some of the health and environmental factors. By prognosticate the human lifespan with dissimilar models Random Forest algorithm gives more precise .These results clearly show and prove the importance of health, education, and economic features on Life expectancy. But there is still some room for improvement by including the other features such as environmental and geographical features. The inclusion and dependency of these suggested features on life expectancy is still a matter of debate and a future part of research in this particular domain Furthermore, the future enhancement can be made by using deep learning algorithm which may give better solution.

9. REFERENCES

[1] Kasichainula Vydehi, Keerthi Manchikanti, and T. Vijay Kumar. "Life Expectancy Prediction through Analysis of Immunization and HDI Factors using Machine Learning Regression Algorithms." In 2021 4th International Conference on Inventive Computation Technologies (ICICT), 2021.

[2] Mahumud R.A., Hossain, G., Hossain, R., Islam, N. and Rawal, "impact of life Expectancy on economic Growth and health care Expenditure in Banglesh.," Universal journal of public health,2019.

[3] Bhosale , A.A. and Sundaram, k.k, "life prediction equation for human being",international conference on Bioinformatics and Biomedical technology, 2019.