Ann. For. Res. 66(1): 4035-4043, 2023
ISSN: 18448135, 20652445

ANNALS OF FOREST RESEARCH
www.e-afr.org

HUMAN LIFE EXPECTANCY PREDICTION USING MACHINE LEARNING

Dr. S. Selvakumar Raja[*1]
Principal & Professor, Department of ECE

M. Nagarani [2]
Associate Prof & HOD, Department of CSE

B. Kavitha [3]
Assistant Professor, Department of CSE

P. Nagarjun [4]
Assistant Professor, Department of CSE, Kakatiya Institute of technology and science for women, Nizamabad 1,2,3,4, Telangana, India.

ABSTRACT:

Many countries' societal and economic structures are profoundly impacted by life expectancy (LE) models. Life expectancy forecasts have been shown to have far-reaching effects on healthcare system administration and social policy in several research. The healthcare system and the mechanism for advanced care planning may both benefit greatly from the insights provided by these models. Many current determinants were once thought to be sufficient for predicting the longevity of the generic set of population, but it has become clear that this is not the case. Previous models relied on data gleaned from the population's death rate. Despite improvements in forecasting methods and years of hard work, some researchers have pointed out that many factors beyond mortality rate need to be taken into account before Predicted Life Expectancy (PrLE) models can be derived. As a result, researchers are examining life expectancy alongside a wider range of topics, including healthcare, economics, and social support.

The authors of the Analysis use a variety of machine learning methods to improve accuracy depending on specific characteristics of the dataset.

Many countries' societal and economic structures are profoundly impacted by life expectancy (LE) models. Life expectancy forecasts have been shown to have far-reaching effects on healthcare system administration and social policy in several research. The healthcare system and the mechanism for advanced care planning may both benefit greatly from the insights provided by these models. Many current determinants were once thought to be sufficient for predicting the longevity of the generic set of population, but it has become clear that this is not the case. Previous models relied on data gleaned from the population's death rate. Despite improvements in forecasting methods and years of hard work, some researchers have pointed out that many factors beyond mortality rate need to be taken into account before Predicted Life Expectancy (PrLE) models can be derived. As a result, researchers are examining life expectancy alongside a wider range of topics, including healthcare, economics, and social support. The authors of the Analysis use a variety of machine learning methods to improve accuracy depending on specific

characteristics of the dataset.Life expectancy (LE) models have vast effects on the social and financial structures of many countries around the world. Many studies have suggested the essential implications of Life expectancy predictions on social aspects and healthcare system management around the globe. These models provide many ways to improve healthcare and advanced care planning mechanism related to society.
However, with time, it was observed that many present determinants were not enough to predict the longevity of the generic set of population. Previous models were based upon mortality-based knowledge of the targeted sampling population.
With the advancement in forecasting technologies and rigorous work of the past, individuals have proposed this fact that other than mortality rate, there are still many factors needed to be addressed in order to deduce the standard Predicted Life Expectancy Models (PrLE). Due to this, now Life expectancy is being studied with some additional set of interests into educational, health, economic, and social welfare services. In the Analysis, the authors have implemented different machine learning algorithms and have achieved better accuracy based on pertinent features of the dataset
Life expectancy (LE) models have vast effects on the social and financial structures of many countries around the world. Many studies have suggested the essential implications of Life expectancy predictions on social aspects and healthcare system management around the globe. These models provide many ways to improve healthcare and advanced care planning mechanism related to society.
However, with time, it was observed that many present determinants were not enough to predict the longevity of the generic set of population. Previous models were based upon mortality-based knowledge of the targeted sampling population.
With the advancement in forecasting technologies and rigorous work of the past, individuals have proposed this fact that other than mortality rate, there are still many factors needed to be addressed in order to deduce the standard Predicted Life Expectancy Models (PrLE). Due to this, now Life expectancy is being studied with some additional set of interests into educational, health, economic, and social welfare services. In the Analysis, the authors have implemented different machine learning algorithms and have achieved better accuracy based on pertinent features of the dataset

KEYWORDS: Machine Learning, Life Expectancy, Regression.

1. INTRODUCTION

Nobody has ever known how long the average human life span really is. People have no idea how long their life will last or what they will do with that time. It seems that life expectancy is the most important indicator of population health. The term "life expectancy" is used to describe how long a person is likely to live. According to, life expectancy has more than doubled since 1900, when it was only approximately 30 years worldwide (Roser et al., 2019). There was a worldwide gain of almost six years in life expectancy between the years 2000 and 2019 (World Health Organisation, n.d.-b). Several variables affect one's life expectancy, including one's socioeconomic status, access

to and quality of healthcare, lifestyle, diet, social and genetic background, and the environment. Higher levels of socioeconomic advantage and more excellent healthcare resources of the people were more likely to increase life expectancy in a study conducted on the Asian population, including Singapore, Malaysia, and Thailand (Chan & Kamala Devi, 2015). Thanks to the development of big data and analytics, researchers now have access to a wealth of information that can be used to develop accurate projections of human longevity. Data mining refers to the process of studying big databases in search of previously unseen patterns, correlations, or anomalies. Despite their similarities, data mining and machine learning are not the same thing. Data mining is the process of looking for patterns within data. Machine learning, on the other hand, doesn't rely only on data from the past when making predictions. The result is based on the machine's inferences from the available data. Disease prediction (Nalluri et al., 2020; Shuja et al., 2020;[1] Verma et al., 2020), sociological studies (Song & Song, 2021; Vanlalawmpuia & Lalhmingliana, 2020), agricultural practises (Kaur et al., 2021), pedagogical practises (Ahmad Tarmizi et al., 2019; Basheer et al. Decision Tree, Naive Bayes, k-Nearest Neighbour, and Support Vector Machine are just few of the data mining classifiers that have been employed in prior studies of longevity (Mohammad Suhaimi et al., 2019; Sharma et al., 2016). This research uses a database of national life expectancy rates in Asian nations to evaluate the accuracy of mortality forecasts. This research aims to classify life expectancy in the Asian population using machine learning algorithms (MLAs) based on the identified factors, and to compare the performance of these MLAs based on these factors. There are three Machine Learning Algorithms (MLAs) utilised to create the forecasting models: J48, Random Tree, and Random Forest. Various measures of precision are used to assess the MLAs' effectiveness. The literature review, methods, results, and conclusions, as well as the conclusion and recommendations for future research are presented below [2].

2. LITERATURE REVIEW

In their study on forecasting life expectancy in global health metrics, Kyle J. Forma et al. (2018) made this suggestion. Their evaluation results, which included 195 nations from 1990 to 2016, were mostly focused on GBD, risk factors, and injury analyses. It is interesting how many health-related aspects have been included, and it is demonstrated how each trait may be extrapolated into a health scenario. In the estimation, 79 health drivers were taken into account. Based on factors related to health, they developed an ARMA (Auto Regressive Moving Average) model to forecast life expectancy. The demographic and educational characteristics of each nation are not discussed or taken into account in this specific research. Moving average and autoregression analysis are used to time series datasets (with good results) in the ARMA model.[3].

Aishwarya N. and others. He emphasized how amazing of a creation man is. Every living creature has a certain amount of time in which to do anything in the world. Even though we have developed numerous technologies to fight sickness, it is still unclear how we will survive the repercussions. In our contemporary culture, dietary practices, numerous illnesses, and environmental factors all have a significant impact on life expectancy. [4].

VM Shkolnikov and others. Estimating the human lifetime is thought to be an important first step. Although there is increased interest in this field of research, there are numerous obstacles to

overcome because there are little resources accessible. (i.e. datasets). Numerous research are carried out to evaluate human stability by noting date of birth, environmental variables, eating habits, illnesses, and medical history.[5].

D.F. Andrews. Data approaches are said to be resistant to minute fractional changes. Otherwise, the strategies will be more effective if the static efficiency is maintained at a high level. The forecast outcome is accurate if the accuracy grade is great. [6]

3. SYSTEM ANALYSIS

3.1 EXISTING SYSTEM:

Our lifespans have increased. Life expectancy and the oldest age of death have both risen dramatically in the last several decades. Changing human DNA or restricting food consumption alone to increase longevity is not yet feasible, helpful, or ethical. Extending human life is one of the greatest problems of the 21st century, and the maximum potential human life span is still a matter of much controversy. Many researchers think that human lifespan is capped at 85, 100, or 150 years old, but they can't agree on which number is correct. While the record of older age is that of death, it is generally postulated that the maximum human life span is around 125 years.

3.2 PROPOSED SYSTEM:

This is a classic example of using past data to make predictions about the future using machine learning, or Regression.We will use machine learning to make inferences from the provided data set and forecast the life expectancy rate for a certain nation.

ADVANTAGES:

- 99% accuracy.
- Less time taking

4. SYSTEM ARCHITECTURE:

To illustrate the interdependence of the system's parts, an architectural diagram would be drawn. They are designed to depict the cooperation between hardware and software components in a system, and hence are often drawn for such systems. However, it can also be designed for use in web-based programmes.

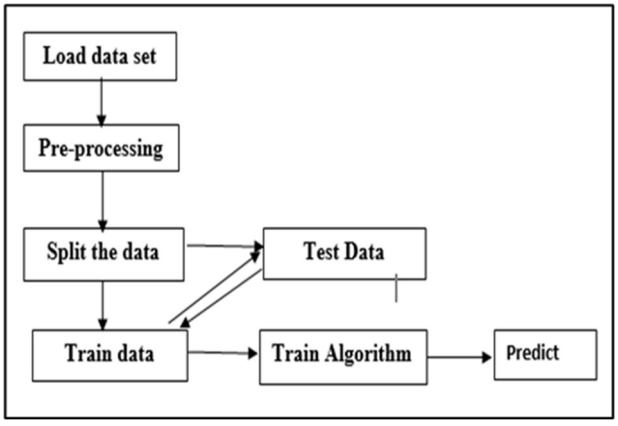

Fig4.1: Block Diagram

LOAD DATASET

The dataset contains a lot of separate data, but can be used to train an algorithm to find predictable patterns across the dataset.

PREPROCESSING

Preparing raw data so that it may be used in a machine learning model is known as "data pre-processing." This is the very first and most important stage in developing an ML model. It's not always the case that we come across clean and prepared data while developing a machine learning project. It is also crucial to clean and prepare data before running any actions on it. The data pre-processing activity is used for this purpose. Unfortunately, machine learning models cannot always be applied directly to real-world data due to issues such as noise, missing values, and unusable formats. In order to improve the accuracy and efficiency of a machine learning model, data pre-processing is an essential step in the data preparation process.

SPLITTING THE DATASET

In this procedure, a dataset is divided in half. The first group is used to train the model, hence it's termed the training data set. The second subset's input is given to the model, and the results are compared to those predicted by the model. A test data set is the term for the second batch of information.

TRAIN ALGORITHM

Model training, in ML parlance, is the practise of providing an ML algorithm with data in order to assist it discover and learn appropriate values for all of the variables in question. Supervised and unsupervised models are the most prevalent forms of machine learning algorithms. When both input and output values are included in the training data, supervised learning becomes feasible. A "heartbeat" is a record of the system's predicted inputs and outputs. When inputs are fed into a model, the model is trained based on how much they deviate from the documented result during processing.

PREDICT

You are now ready to use your Machine learning model to derive results in real world scenarios.

5. IMPLEMENTATION

Everything that is done to make the transition from the old system to the new system is part of the implementation. The suggested approach makes it exceedingly challenging to do the manual procedures that make up the previous system. To provide a dependable system that caters to the requirements of the organisation, proper implementation is required.

- Just like any other ML model, this model also needs to be trained using some data-set.
- Import all the necessary modules.
- Read the data from “csv” file into Pandas data-frame.
- Pre-process the data.
- Classify the data.
- Split the data-set into Train and Test data
- Train the Model using Regression Algorithm with Train data.
- Test the model with test data to check the Accuracy of the model

- Create the User Interface using Django Framework
- Deploy the ML model into UI

6. RESULTS

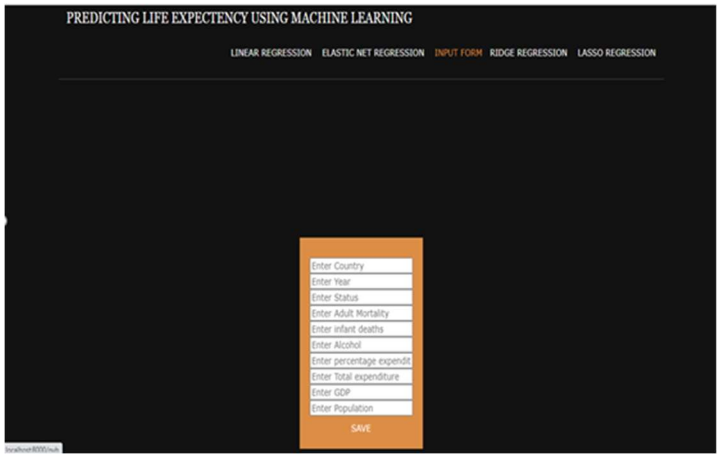

Fig 6.1: Input parameters

The initial stage is to provide the algorithm with the parameters it needs to preprocess and partition the data, and if everything checks out, the programme will go on to testing the data without further processing.

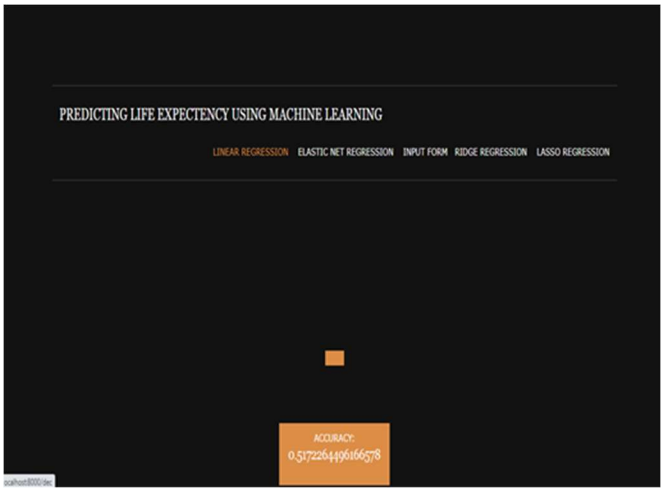

Fig6.2: Accuracy value for the life expectancy 0.517

Form the given input parameter we can see above figure with a predicted life expectancy value of 0.517.

Fig 6.3: Accuracy value for the life expectancy 0.507

Form the given input parameter we can see above figure with a predicted life expectancy value of 0.507.

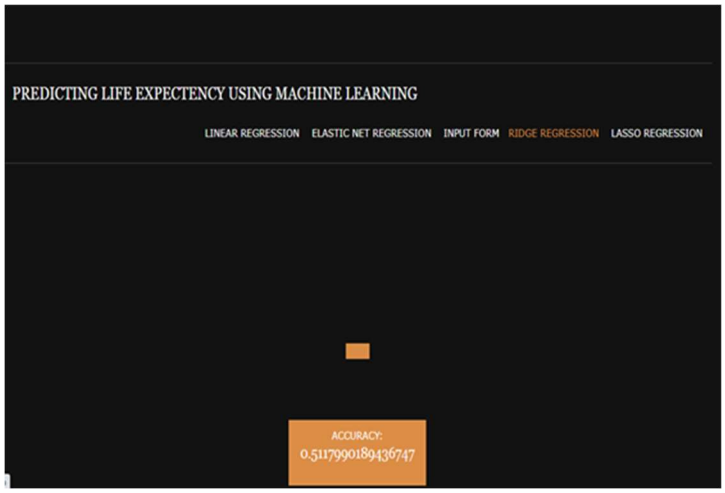

Fig 6.4: Accuracy value for the life expectancy 0.5117

Form the given input parameter we can see above figure with a predicted life expectancy value of 0.5117.

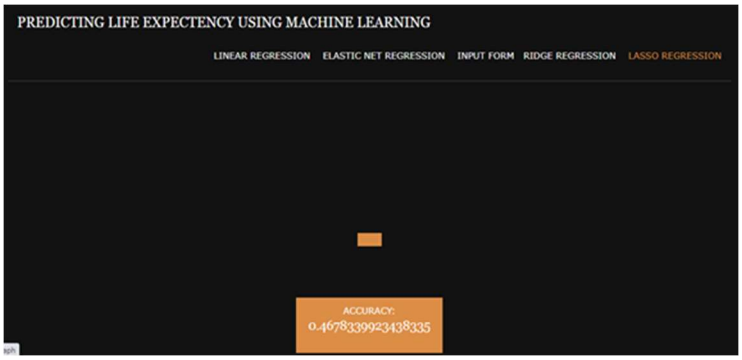

Fig 6.5: Accuracy value for the life expectancy 0.467

Form the given input parameter we can see above figure with a predicted life expectancy value of 0.467.

CONCLUSION

A system of the human lifespan can be anticipated earlier in this work. The relationship between characteristics like illnesses, gender, ages, and environmental factors is tracked by using data from databases. the was accomplished to provide more accurate predictions about human longevity. The Random Forest algorithm's benefit is that it provides greater flexibility without requiring processed data and is accurate. We have thus examined human lifetime in light of various environmental and health aspects. These findings unequivocally demonstrate and support the impact of economic, educational, and health factors on life expectancy. The inclusion of other elements, such as environmental and geographical factors, should yet be improved. It is still up for dispute whether or not these recommended characteristics should be included in studies of life expectancy in this particular field. Additionally, applying deep learning algorithms for future improvements may result in superior solutions.

REFERENCES:

1. Nalluri, S., Vijaya Saraswathi, R., Ramasubbareddy, S., Govinda, K., & Swetha, E. (2020). Chronic Heart Disease Prediction Using Data Mining Techniques. Advances in Intelligent Systems and Computing, 1079(June), 903–912.
2. https://bigcloud.io/machine-learning-can-accurately-predict-your-death-date/
3. Kyle J Foreman Forecasting life expectancy, years of life lost, and all-cause and cause-specific mortality for 250 causes of death, 2018 Nov 10;392(10159):2052-2090. doi: 10.1016/S0140-6736(18)31694-5. Epub 2018 Oct 16.
4. Aishwarya.n, “Determinants of life expectancy in most polluted countries: Exploring the effect of environmental degradation, Published online 2022 Jan 21. doi: 10.1371/journal.pone.0262802”.
5. Vladimir M Shkolnikov “Losses of expected lifetime in the United States and other developed countries: methods and empirical analyses, Demography 2011 Feb;48(1):211-39. doi: 10.1007/s13524-011-0015-6.”
6. D.f. Andrews “The Minimum Inhibitory Concentration of Antibiotics: Methods, Interpretation, Clinical Relevance, Published online 2021 Feb 4. doi: 10.3390/pathogens10020165”.
7. Noorhannah Boodhun, Manoj Jayabalan, ‚Risk prediction in life insurance industry using supervised learning algorithms,‘ Complex & Intelligent Systems, vol. 4, no. 2, pp. 145-154, 2018.
8. Mahumud, R.A., Hossain, G., Hossain,R., Islam, N. and Rawal, L.,;, ‚Impact of Life Expectancy on Economic Growth and Health Care Expenditures in Bangladesh.,‘ Universal Journal of Public Health, vol. 1, no. 4, pp. 180-186, 2013
9. Bhosale, A.A. and Sundaram, K.K., ‚Life prediction equation for human beings,‘ International Conference on Bioinformatics and BiomedicalTechnology, vol. IEEE, pp. 266-268, 2019

10. Aggarwal, D., Mittal, S., Bali V., ‚Prediction Model for Classifying Students Based on Performance using Machine Learning Techniques,' International Journal of Recent Technology and Engineering, vol.8 p.2S7, 496-503, 2019
11. Chen, Tianqi; Guestrin, Carlos, ‚XGBoost: A Scalable Tree Boosting System,' Proceedings of the 22nd ACM SIGKDD
12. https://bigcloud.global/
13. https://health.economictimes.indiatimes.com/
14. https://www.directhit.com/web?qo=semQuery&ad=semA&q=life%20expectancy&o=777014&ag=fw4&an=msn_s&rch=intl222&rtb=29655.