Statistical Analysis on Prediction and Analysis of Life Expectancy

Dr. Nikam Gitanjali Ganpatrao[1], Monishendra Singh[2], Sunny Kumar jaiswal[3], Santosh[4]

[1,2,3,4]National Institute of Technology, Kurukshetra, Haryana, India.
Email: gitanjalichalak@gmail.com[1], s2jaiswal.sj@gmail.com[2] , monishendra.in@gmail.com[3],
sahu18456@gmail.com[4]

Abstract—**Life Expectancy (LE) models have vast effects on many countries' social and financial structures 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 generic set of populations' longevity. Previous models were based upon mortality-based knowledge of the targeted sampling population. With the advancement in forecasting technologies and rigorous work of past individuals has proposed that other than mortality rate, there are still many factors needed to be addressed to deduce the standard Predicated 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. Our objective here will be to briefly discuss the previously used machine learning, Ensemble methods and hybrid methods used in building life expectancy prediction models. We will aim to test the potential and accuracy of existing ML techniques and their probabilistic projections to predict the generic Life expectancy of a particular region, state.**

Keywords— *Ensemble methods, Life expectancy (LE), Machine learning (ML), Predicated Life expectancy (PrLE).*

I.   INTRODUCTION

Life Expectancy is a statistical measure of longevity of a population depending upon distinct factors. Over the years, Life expectancy observations are being vastly used in the medical, healthcare planning and pensions related services by concerned government and private bodies. Advancement in forecasting, predictive analysis techniques, and data-science technologies have now made it possible to develop accurate predictive models. In many countries, it is a matter of political debate about how to decide the retirement age and how to manage the financial issues related to the public matter. Life expectancy predictions provide solutions related to these issues in many developed countries. With the advancement in new systematic, accurate, efficient and result-oriented techniques in the field of Data Science, now predictions of the life expectancy of the selected region are becoming more prominent in demand for the government and the private bodies and their policymaking. Studies have suggested that in early life or the pre-modern era, human beings' average lifespan was around 30 years in approximately all parts of the world. Since then, industrial enterprise and modernization have valued the rapid increase within the lifespan all round the world. The advancement of technology, better healthcare facilities, and education all have led to positive changes in the lifestyle of people. Which, in turn, increases the expected average age of a human being. However, there were stillmany countries with less life expectancy than the rest of the world in the early 1900s. The whole reason for such inequality was the disoriented healthcare facilitiesin these countries. Developed countries have speedily improved their health care and also the public distribution mechanism. This inequality between developed and developing countries has led to such an improper distribution of life expectancy around the globe.
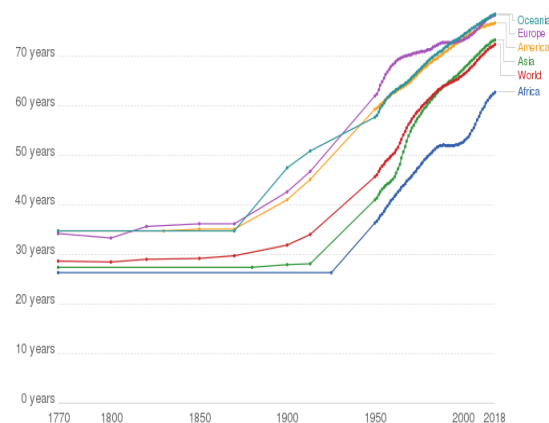

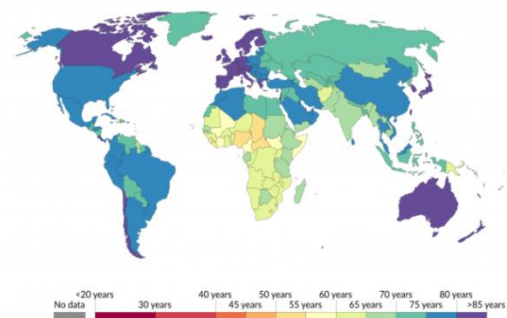Fig 1.Shown above is the average life span in a few countries (Source: UN population division, 2019).

Fig 2.Average no. of years a newborn expected to live if mortality rate remains the same throughout the year.(Source: UN Population Division,2019).

Due to certain developments in public healthcare, now emerging countries also catching up with the other developed countries in terms of life expectancy. In 2019, most of the Central African countries have a low life expectancy of around 52-55 years, whereas, in Japan, recent statistics have shown

thatlife expectancy is around 87 years for women. The lifespan of South Korea was twenty-three years, a century past. Nevertheless, as of today, the life expectancy of India has almost tripled, and in South Korea, it has almost quadrupled since then.

There have been many vast improvements in the field of Data-science and analytical techniques, which explains the rise in life expectancy around the world. These major improvements in the predictive analysis techniques have also led us to more ways so that we can improve the life expectancy of the distinct population. These improvements were solely dependent upon certain indicators. Our extensive research into the prior life expectancy models has suggested us the inclusion of many-more indicators than we expected, such as; GDP(gross domestic product), healthcare expenditure, family income, educational expenditure, infant mortality rate, adult mortality rate, healthcare plans and population of the selected region. Recent studies have also shown the impact of geographical factors, climate conditions on life expectancy. Implicitly, the educational background of people, health plans, economic stability, the burden of diseases, BMI, and environment variables also affect the lifestyle of the people.

There have been many vast improvements in the field of Data-science and analytical techniques, which explains the rise in life expectancy around the world. These significant improvements in the predictive analysis techniques have also led us to more ways so that we can improve the life expectancy of the distinct population. These improvements were solely dependent upon specific indicators. Our extensive research into the prior life expectancy models has suggested us the inclusion of many-more indicators than we expected, such as; GDP(gross domestic product), healthcare expenditure, family income, educational expenditure, infant mortality rate, adult mortality rate, healthcare plans and population of the selected region. Recent studies have also revealed the impact of geographical factors, climate conditions on life expectancy. Implicitly, the educational background of people, health plans, economic stability, the burden of diseases, BMI, and environment variables also affect the people's lifestyle.By Summarizing all the factors mentioned earlier, we may be able to create a distinct set of datasets that will help us reach our final goal to predict the desired population's life expectancy.In the final stages, selecting the correct and accurate ML algorithmsis probably the most tedious job of this prediction model. Accuracy and reliability factors of final results depend upon the methodology used in the demonstration and as well as the correctness of the dataset(which in turn depends upon the selection of independent variables and their sources of origin).

In this section, we have discussed the various interrelated aspects of the life expectancy prediction model.In upcoming sections, we will discuss some of the importance of predictive models, their use cases, and the exhaustive research part of prior predictive models. We will summarize some of the essential aspects of those papers. Later, we will conclude the implications and assertions made from those research papers.

II. APPLICATIONS AND USE CASES

Life expectancy results are not just the merely figures derived from the datasets; instead, these are quite influential in solving real-world problems.Most demographers use forecasting models to predict males and females' average life span for specific countries.

Table 1. Summarized review of previous research papers with their limitations and their references:

Country	Male	Female	Difference
Russia	64.7	76.3	11.6
Belarus	66.5	78.0	11.5
Lithuania	68.1	79.1	11.0
Rwanda	60.9	71.1	10.2
Syria	59.9	69.9	10.0
Ukraine	66.3	76.1	9.8
Latvia	69.6	79.2	9.6
Estonia	72.7	82.0	9.3
El Salvador	68.8	77.9	9.1

These results are quite useful in studying the life structure and living patterns of the population.It helps in drawing appropriate insights from those results. In general, life expectancy predictions used in the field of research and policymaking decisions. Numerous government policies all around the world are made using these results, such as in healthcare services, human resource management decisions, public health wellness, and maintenance.It can also affect the decisions of the policymakers, for example, Health expenditure by the concerned authority. Generic life expectancy predictions can also help the population improve their lifestyles and make efficient, healthier decisions individually.For example, smoking's adverse effects can endanger human health and may cause some lung-related severe diseases to the individual. Hence, smoking should be avoided to live a healthy life long.

Life insurance companies are one of the major stakeholders in using these life expectancy predictions. Risk assessment is quite a critical part of the life insurance profession. They classify their applicants based upon certain criteria by using these results[1]. They aim to enhance the process of risk assessment using predictive analytics. Some of the stronghold life insurance organizations still rely on the conventional mortality rate-based prediction models. TheseFirms decide the premium of their plans as per the risk-based on the history of their customers. They categorize their customers on their estimation of survival inferred from their medical history and several other affecting factors.

By taking an example of Bangladesh, we will explore a case study implication of life expectancy. In Bangladesh, these results have some policy implications for the country, and it suggested that life expectancy is dependent on economic growth, per person expenditure on health and individual having a higher education is expected to have a healthier life than the others[2]. Therefore, the study suggested that to have

political stability, adequate and satisfactory public health policies,the government needs to be involved more often in decision making to increase life expectancy and provide economic stability. Generally, this scenario is real in a sense for all countries around the world(not just in this particular case). Countries having higher GDP pertain to higher LE than compare to others. Nations, which are spending more wealth on Medical and health care expenditures, are expected to have a higher life expectancy than others. Life expectancy results could be used for statistical comparisons based on illustrative features of the results.

One of the most proficiently used real-time systems using life expectancy analytics is into the cancer diagnosis based upon socio-economic factors. In this, the patient's EHR is collected and the available data is used for the estimation of lost lifespan for each patient due to different types of cancer. These results are quite helpful in providing the real-world measure of survival of the cancer patient. The results can be estimated at even diagnosis stages of the patients. It also helps in the calculation of future trends of life loss of a patient. They might also help the policymakers identify the disease burden and motivate the decision-making efforts to balance socio-demographic inequalities. The consistency criteria have prominently relied upon the variables that affect life expectancy, prediction models. Prior works of literature have shown the introduction of more realistic and significant variables can improve the performance of models.

III. INVOLVED CHALLENGES

Creating life expectancy forecasting models may be a tedious job to try and do. There are several challenging factors involved in this process. From identifying indicators, collecting the datasets, preprocessing the available data, and at last, the final phase of implementation is applying appropriate machine learning algorithms to the datasets. Testing the accuracy is another matter of discussion, but it is still one of the involved challenges in the process. We will discuss each of the challenges briefly below one by one :

1) The first significant challenge in front of us will be to prepare the correct and error-free dataset—the sources of data needed to be legit and trustworthy.

2) After identification of legitimacy of sources, dataset need to be categorized the two categories;

   a) *Structured data:* this kind of data is stored in EMR or UN official records as well as on the government's official websites for public reference.

   b) *Unstructured data consists of numerous non-validated information that needs to be converted into measurable and statistically discrete attributes to accommodate the other dataset's other attributes.*

3) Usage of numerous non-standardized abbreviations and improper attributes(with no proper definition in the context of life expectancy) can create ambiguity in the final results.

Elimination of redundant information is also an essential process in the initial stages and also a challenging one.

4) The selection of appropriate techniques, machine learning algorithms, and technologies is a big hurdle to create the LE prediction model. This phase will determine the consistency and efficiency of the model into the future.

These challenges mentioned above present the tedious task of building such forecasting models.

IV. INVOLVED PARAMETERS

All these factors are some of the major parameters involved in the estimation of life expectancy.

Table 2. List of Attributes used in Life expectancy Predictive models.

Attribute	Brief Description
Infant mortality rate	It depicts the death of kids beneath a certain age one per a thousand live births.
Alcohol consumption	It represents the average consumption of alcoholic beverages by the net population.
Percentage of expenditure	The proportion of total average family expenditure described by associate item (budget share).
BMI	It is an estimation of body fat in accordance with the weight and height applicable to adult males and females.
GDP	It is also referred to because of the estimated price of all the products and services created in a year by a country.
Under-five deaths	It represents the estimation of deaths of children before reaching age 5 per 1000 live births.
Total income composition	The relative share of every income supply or group of sources is expressed as a proportion of that cluster or region's aggregate total income.
Educational expenditures	It refers to the total sum of expenses done on the educational services and subsidies acquired by each group of people of that region.
Health care expenditures	Whole consumption of health-related services, expenses on the health care plans (including personal care and family healthcare plans).
Population	It simply describes the number of people (active entities) in the region.
Environmental criteria	Major environmental factors, for example, climate change, modernization, and altitude.
Schooling	Total expenditure on educational services by people.
Thinness	Prevalence of thinness among children and adolescents.

V. DESCRIPTION OF ML TECHNIQUES

The prime objective of any predictive model is to predict a dependent variable's outcome due to the number of independent variables. In the domain of machine learning, numerous techniques can be applied to solve such a task. Here, we shall briefly discuss some of the widely used ML

techniques for such continuous prediction problems in the Life expectancy prediction domain.

- *Linear regression:* It is one of the simple techniques, which is used to predict the outcome of the dependent variable depending upon single or more independent features. Multiple regression analysis allows us to develop a mathematical model dependent on multiple features. The stepwise regression method is composed of iteratively adding or removal of independent features from the set, in the end, giving us the best performing model. Bhosale et al.(2010) predicted humans' life expectancy based upon their heart rate, respiration rate, and blood pressure using the linear mathematical model [19].

- *Regression trees:* It requires a lot of tree-based algorithms to produce the resultant dependent variable, such as Random forests, Decision trees, XG Boost [3]. These techniques can help in predicting the complex and non-linear relationships among the variables and they're resultant. Kerdprasop, N. et al.(2017) adopted the CHAID algorithm, which is a kind of classification and regression tree method used for building predictive models[15].

- *Neural Networks:* It helps in finding solutions for pattern and speech recognition problems. It can also help in some way with the predictive analysis. It has an associated input layer and an output layer. In between, it has several hidden layers. These hidden layers and combinations of neurons help us to find the required solution to the problem. The basic formulation of any neural network is:

$$y = f(x)$$

Here, 'y' is the outcome that can be generated from any relationship 'f' follows by feeding input 'x.' A neural network consists of many layers and each layer made up of nodes. These nodes combine the input with pre-defined weights and then summed up. These input-weight product values are passed on to activator-function. It helps the network to determine the ultimate progress of that particular input in the network.
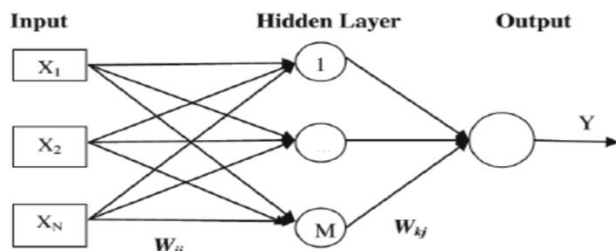

Fig 3.The basic architecture of a neural network model.

In [4], D. Hainaut presented an approach to estimate mortality forecasting using a neural network model.

Some of the most useful time-series forecasting techniques are :

- *Artificial neural networks:*This algorithm works similarly to neurons in our brain and works in parallel for data

processing and recollection. These neurons have some associated weights with them. These weights store some information, which helps in the final estimation of outcomes. The prime purpose of this analysis is to identify these 'weights' to fit the model mathematically. Adya and Collopy gave an early review of ANN [5]. This phenomenon was first acknowledged by the McCulloch and Pitts around 1943 [6]. Rosenblatt presented the concept of calculation related to single mathematical neuron [7]. Mhd. Khafiroh ZamzamySormin et al. predicted the life expectancy of the world population using Cyclical Order Weight/Bias (On of ANN methods). This methodrained the model with burdensome-learning rules and bias with some additional updates after the data is presented in the input [8].

In the real-world, for predicting the life expectancy problem's continuous outcomes, above mentioned regression models are not enough due to the time-series datasets. Most of the time-series datasets are non-stationary, which means properties like mean and variance are not constant with time. Successive study of many review papers has suggested to us that the Life expectancy problem's outcomes can be optimized using the Time-series forecasting techniques. Time-series analysis is employed in several fields like census analysis, value prognostication, economic forecasting, and energy consumption forecasting in the future [10]. Life expectancy falls under the category of multivariate analysis of the time-series domain. Some of the most useful time-series forecasting techniques are:

- *Moving Average and Exponential smoothing(MA and EA):* Both techniques are similar in a manner, such that both assume the time series as stationary locally with a slowly moving average [11]. EA assigns higher weights to the new values while MA treats each as equal. It also has many applications in creating a basis for other time-forecasting techniques such as ARMA, ARIMA.

- *ARIMA(Autoregressive Integrated Moving Average):* This model approaches the task as the autoregressive and moving average model [9]. If parameters involved in the forecasting task have some correlations, the ARIMA model is a perfect analysis technique for forecasting. It can represent the stationary and non-stationary time series. The basic idea behind the ARIMA model is transforming the time-series into stationary by the differencing process. These models follow a standard and exemplary approach, details of which can be found in the literature of Box and Jenkins [10]. The mathematical formulation of the ARIMA(p,d,q) model can be given as(using lag polynomials):

$$\varphi(L)(1-L)^d y_t = \theta(L)\varepsilon_t$$

$$\left(1 - \sum_{i=1}^{p} \varphi L^i\right)(1-L)^d y^t = \left(1 + \sum_{j=1}^{q} \theta_j L^j\right)\varepsilon_t$$

Here p,d and q are the order of AR, integrated and MA part of the ARIMA model. 'd' variable controls the model's level of difference. Kyle J. Foreman et al.(2018) presented a paper in Global health metrics. He proposed a solution to model life expectancy across 195 countries using the ARMA model[13].

VI. COMPARATIVE LITERATURE-BASED STUDY

Advancement in data visualization techniques has led the researchers to attain the optimum solution. We will now discuss some of the significant contributions to this section of research by numerous researchers and briefly describe their mechanism and inferences deduced from their work.

In their paper, Aggarwal et al.(2017)resented a 5-year predictive tool of patients from a recent one or two hospital visits[12]. This is one of the most accurate predictive models in recent years, providing results up to 5 years. Their work was based upon the ensemble of different machine learning models. Ensemble technique is a machine learning technique in which various base models are combined in order to achieve the one final predictive model. This was the reason behind the final model's deployment with precise implementation because it was ensembled with 75 individual models. To attain such a level of precision, they need to be specific about patients and need to have access to each individual's EMR data. It performed far better than the previous models.

Kyle J. Foreman et al.(2018) proposed a paper related to life expectancy estimation in 'Global health metrics'[13]. In which their prediction results were mostly centered on the GBD, risk factors and injuries analysis of 195 countries from 1990 to 2016. Consideration of that number of health-related factors was quite astonishing, and the projection of each attribute against the health scenario was shown. They took 79 health drivers into account for the estimation. They used the ARMA(autoregressive moving average model) for the analysis of life expectancy based on the health drivers. Each country's demographic and educational factors were not mentioned or taken into account in this particular paper. In the ARMA model, both autoregression analysis and moving average methods are used and applied to a well-behaved time-series dataset.

Beeksma et al. presented this paper in BMC Medical Informatics[14]. It is one of the recent researches in the concerned subject area. They proposed the idea of using supervised machine learning using recurrent neural networks on deceased patients by using their medical records. They approached the task with supervised-machine learning. Then, they trained and tested the data on the LSTM(Long short-term memory) recurrent neural network. LSTM is a type of RNN(recurrent neural network). In RNN,the output from the last step is fed as input within the current step. LSTM was designed to beat the matter of long-term dependencies of RNN. It is mostly used in the time-series data due to the lags between the important unknown events in time-series. Their model was based on non-text and text-features of medical records. The first one was the base model(containing non-text features), and another one was the keyword model(containing text-features in EMR). The Keyword model proposed a better accuracy of 29% than 20% of the base model. But, some limitations of this model were the data availability and not generalized in predictions.

N. Kerdprasop et al. (2017), in their paper, suggested that there is an association of economic and environmental factors to life expectancy[15]. They used the Chi-square Automatic Interaction Detector (CHAID) algorithm for categorical values and regression on continuous and numerical values. The underlying principles under the CHAID algorithm isa Decision tree. It is a tool used to discover the relationship between the variables. Using this technique, they attempted to show the relationship between resources and economic growth of a country to its people's life expectancy. They predicted the life expectancy of people living in the countries along the Mekong River. Eventually, their results revealed a strong relationship between environment and GDP growth to people's life expectancy.

In their paper, Pieter van et al.(2016)presented a model to make a base forecast for sex-specific and SES-specific LE in the Netherlands using a set of critical assumptions about the future group-specific and typical time trends[16]. They were using the traditional Li-Lee model for the estimation. This Li-Lee model is an extension of the original Lee-Carter model. The Lee-Carter model was introduced in 1992, and as of today, it is used by many authors to determine future mortality trends. This model is fed with a matrix of age-specific mortality rates and the output is another matrix of forecasted mortality rate.

These works of literature mentioned above and their extensive contributions are some of the significant developments in this context.

VII. SCOPE OF WORK AREA

Availability of all health-related data, education, and economic expenditure stats have made possible the proper and error-free estimation of life expectancy models. Earlier life expectancy models were dependent on very few variables and due to the unavailability of advanced data exploratory and validation techniques, the trained model was not so much accurate. Now, in recent studies, various milestones have been achieved in this field. For example, the ensemble of various base models, which includes the results of individual base models, the inclusion of ANN, and RNN techniques in solving such tasks. So, the accuracy and inclusion of decisive variables into the final life expectancy model will be far two most significant areas of concern in our further research part.

This paper focuses on presenting the current scenario going into the field and try to propose the generic solution to test the life expectancy models on selected datasets.Validation and accuracy of the trained model are to be tested by the exploitation of numerous model validation techniques and

finally, the effect of various indicators will come later into the scenario of discussion.The main aim of our future work in this fieldwill bethe optimization of results with the inclusion of a few more variables by not affecting the overall performance, complexity, and accuracy of the trained-model.
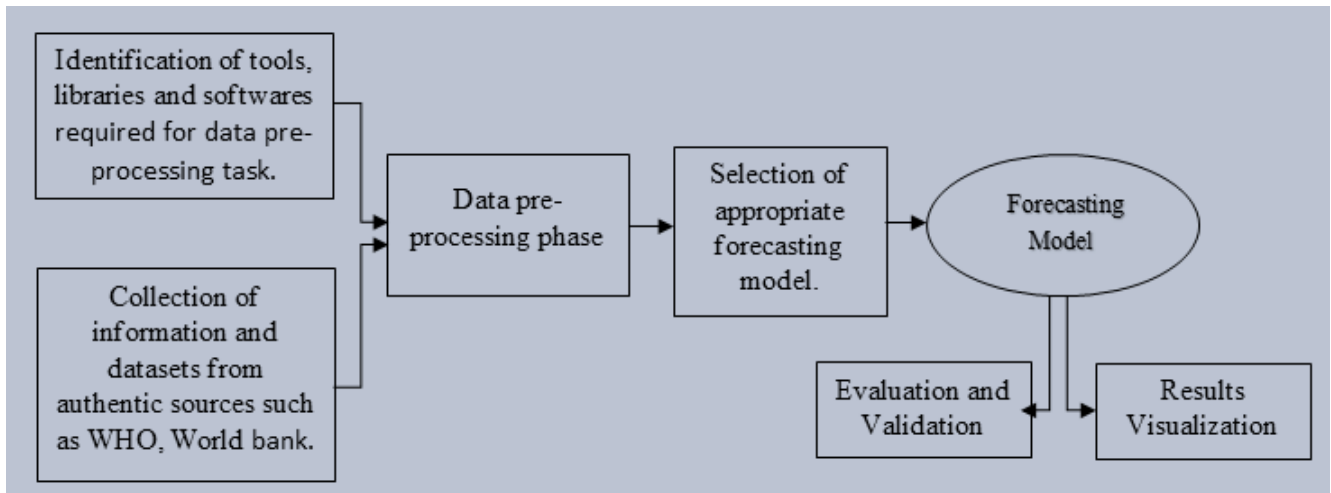

Fig4. A flow chart presenting design & description of each phase of the Basic Forecasting model.

Table 3. Summarized review of prior research papers with their limitations and their references:

Index	Authors	Objective	Technologies used	Limitations
[1]	Beeksma et al.(2019)	To test the potential of ML and NLP techniques for predicting life expectancy using EMR.	RNN on the text and non-textual dataset.	The model's performance is far from perfect (baseline model 20% accuracy & Keyword model with 29%).
[2]	KJ Foreman et al. (2018)	To provide a novel approach to modeling life expectancy, all-cause mortality, and cause of death forecasts (Based on GBD data) [13].	ARIMA model.	Many factors were not included; forecasts were limited by the validity of these simple extrapolations of past trends.
[3]	Pascariu et al. (2018)	To construct a model for a forecasting life expectancy of female and male life expectancy at any age.	Time-series forecasting and Linear regression [17].	Indication of improper utilization of Linear regularity to determine LE at age 65 or above.
[4]	Kerdprasop et al. (2017)	To analyze the effect of economic and environmental factors on the LE of people.	Chi-squared automatic interaction detection (CHAID) algorithm and regression [15].	Absence of health and education-related variables in the final prediction.
[5]	Vasilis Kontis et al. (2017)	To predict age-specific future LE in 35 industrialized countries [18].	Ensemble of 21 forecasting models.	Inability to account for completely unexpected events and changes in the social, technological, and health systems determinants of health.
[6]	Aggarwal et al. (2016)	To design an online health informatics tool to predict the 5-year life expectancy of older adults.	An Ensemble of 75 predictive model.	The proposed work was based on patient-specific data rather than disease-specific.
[7]	Pieter van et al. (2016)	To present a base prediction for gender-specific and SES-specific LE in the Netherlands using a set of key assumptions about the	Using the traditional Li-Lee model.	Unavailability of required information in a few attributes.

		future group-specific and common time trends [16].		
[8]	Bhosale et al. (2010)	Constructing a LE model based on sampling populations [19].	Using regression models.	LE model was based on only three limited variables.

VIII. CONCLUSION

In this review paper, explorations of essential implications of Life expectancy pre-dictions in real-world situations have been discussed. Many previous studies have suggested the adverse effects of poor health and economic conditions on life expectancy. We discussed many similar kinds of life affecting indicators gathered from the prior research in this field. We also briefly described these variables and their importance in our proposed model. The global statistics of life expectancy and their inequalities among different parts of the world have been briefly discussed. Numerous applications and use cases of life expectancy results are thoroughly stated in our study, which is utilized in the various organizations and how these inferences can help develop social healthcare services and policies. Based on our literature study, scholars' contributions and their limitations are referred to in the current scenario with their references.

After our extensive literature survey, we might suggest that by applying necessary machine learning prediction algorithms on the selected dataset, It is possible to successfully design the life expectancy estimation model, which will provide results with a better degree of accuracy (or comparable to previous models).

IX. REFERENCES

[1] 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.

[2] Mahumud, R.A., Hossain, G., Hossain, R., Islam, N. and Rawal, L.,;, "Impact of Life Expectancy on Economics Growth and Health Care Expenditures in Bangladesh.," *Universal Journal of Public Health,* vol. 1, no. 4, pp. 180-186, 2013.

[3] Chen, Tianqi; Guestrin, Carlos, "XGBoost: A Scalable Tree Boosting System," *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,* p. 785–794, 2016. doi:10.1145/2939672.2939785.

[4] Hainaut, D., "A neural-network analyzer for mortality forecast," *ASTIN Bulletin: The Journal of the IAA,* vol. 48, no. 2, pp. 481-508, 2018.

[5] Adya, M. and Collopy, F., "How effective are neural networks at forecasting and prediction? A review and evaluation," *Journal of forecasting,* vol. 17, no. 5-6, pp. 481-495, 1998.

[6] M. R. Hebb, "The Organization of Behaviour," New York, Wiley, 1949, p. 437.

[7] Rosenblatt, F., "The perceptron: a probabilistic model for information storage and organization in the brain," *Psychological Review,* vol. 65, no. 6, p. 386, 1958.

[8] Sormin, M.K.Z., Sihombing, P., Amalia, A., Wanto, A., Hartama, D. and Chan, D.M., "Predictions of World Population Life Expectancy Using Cyclical Order Weight/Bias," *Physics: Conference Series (IOP Publishing),* vol. 1255, no. No. 1, p. 012017, 2019.

[9] Nath, B., Dhakre, D.S. and Bhattacharya, D., "Forecasting wheat production in India: An ARIMA modelling approach," *Journal of Pharmacognosy and Phytochemistry,* vol. 8, no. 1, pp. 2158-2165, 2019.

[10] Box, G.E., Jenkins, G.M., Reinsel, G.C. and Ljung, G.M., Time series analysis: forecasting and control, John Wiley & Sons., 2015.

[11] Deb, C., Zhang, F., Yang, J., Lee, S.E. and Shah, K.W., "A review on time series forecasting techniques for building energy consumption," *Renewable and Sustainable Energy Reviews,* vol. 74, pp. 902-924, 2017.

[12] Mathias, J.S., Agrawal, A., Feinglass, J., Cooper, A.J., Baker, DW and Choudhary, A., "Development of a 5-year life expectancy index in older adults using predictive mining of electronic health record data," *Journal of the American Medical Informatics Association,* vol. 20, no. e1, pp. 118-124, 2013.

[13] Foreman, K.J., Marquez, N., Dolgert, A., Fukutaki, K., Fullman, N., McGaughey, M., Pletcher, M.A., Smith, A.E., Tang, K., Yuan, C.W. and Brown, J.C., "Forecasting life expectancy, years of life lost, and all-cause and cause-specific mortality for 250 causes of death: reference and alternative scenarios for 2016–40 for 195 countries and territories," *The Lancet,* vol. 392, no. 10159, pp. 2052-2090, 2018.

[14] 1. Beeksma, M., Verberne, S., van den Bosch, A., Das, E., Hendrickx, I. and Groenewoud, S., " Predicting life expectancy with a long short-term memory recurrent neural network using electronic medical records," *BMC medical informatics and decision making,* vol. 19, no. 1, p. 36, 2019.

[15] Kerdprasop, N. and Kerdprasop, K., "Association of economic and environmental factors to life expectancy of people in the Mekong basin," *IEEE Conference on Industrial Electronics and Applications (ICIEA),* pp. 1984-1989, 2017.

[16] van Baal, P., Peters, F., Mackenbach, J. and Nusselder, W., "Forecasting differences in life expectancy by education," *Population studies,* vol. 70, no. 2, pp. 201-216, 2016.

[17] Pascariu, M.D., Canudas-Romo, V. and Vaupel, J.W., "The double-gap life expectancy forecasting model," *Insurance: Mathematics and Economics,* vol. 78, pp. 339-350, 2018.

[18] Kontis, V., Bennett, J.E., Mathers, C.D., Li, G., Foreman, K. and Ezzati, M., "Future life expectancy in 35 industrialised countries: projections with a Bayesian model ensemble," *The Lancet,* vol. 389, no. 10076, pp. 1323-1335, 2017.

[19] Bhosale, A.A. and Sundaram, K.K., "Life prediction equation for human beings," *International Conference on Bioinformatics and Biomedical Technology,* vol. IEEE, pp. 266-268, 2010.

[20] Gustin, M., McLeod, R.S. and Lomas, K.J., "Prediction of Internal Temperatures During Hot Summer Conditions with Time Series Forecasting Models," 2018.