ABUAD Journal of Engineering and Applied Sciences (AJEAS)

ISSN: 2992-5584

Volume 3, Issue 1, 28-35

Impact of Economic Factors on Life Expectancy: A Machine Learning Approach

Isaac Oluwaseyi AJAO[1], Ezekiel Adebayo OGUNDEPO[2], Alaba Akinleye OBABIRE[3]

[1]*Department of Statistics, Federal Polytechnic, Ado-Ekiti, Ekiti State, Nigeria*
[2]*African Institute of Mathematical Sciences (AIMS), Rwanda,*
[3]*Department of Statistics, Federal Polytechnic, Orogun, Delta State, Nigeria*

ajao_io@gfedpolyado.edu.ng, oezekiel@aims.ac.rw, obabireakinleye@gmail.com

Correspondence: ajao_io@gfedpolyado.edu.ng; Tel.: +2348035252017

Date Submitted: 27/02/2025
Date Accepted: 26/05/2025
Date Published: 30/06/2025

***Abstract**: Accurate estimation of population longevity is a critical input for macroeconomic planning, health-sector budgeting and international development monitoring. Leveraging a harmonised cross-sectional data set for 156 sovereign states, this study undertakes a rigorous comparative evaluation of three predictive frameworks: multiple linear regression, compact multilayer-perceptron neural networks and radial-basis support-vector regression applied to a common panel of economic, demographic and child-mortality indicators. Two parsimonious perceptron configurations (5–3 and 1 × 7 hidden-unit topologies) are trained with resilient back-propagation and subjected to hold-out testing. Forecast accuracy is scrutinised through mean error, mean absolute error (MAE), root-mean-squared error (RMSE), normalised RMSE and per cent bias. Both neural architectures decisively outperform the linear and kernel baselines, yielding out-of-sample MAE values of 0.17 year and 0.20 year, respectively, compared with 0.26 year for ordinary least squares and 0.32 year for the support-vector estimator; RMSE shows a commensurate hierarchy. Given the 16-year range of life expectancy in the sample, these sub-quarter-year deviations attest to the ability of even modest neural frameworks to capture non-linear interactions, most notably between external debt, crude birth rate, population scale and infant mortality proxies, that elude conventional models. Residual diagnostics confirm homoscedastic, unbiased errors for the multilayer perceptrons, whereas the support-vector regressor exhibits systematic under-prediction at the upper tail. The evidence underscores the methodological and practical utility of lightweight artificial neural networks for national longevity forecasting, furnishing policymakers with more precise baselines for targeted economic and public health interventions.*

Keywords: Life expectancy, economic indicators, artificial neural networks, linear regression, support-vector regression.

1. INTRODUCTION

Life expectancy remains one of the most informative summary measures of population health and social welfare, reflecting thef cumulative effects of economic performance, healthcare infrastructure, educational attainment, behavioural patterns, and other socio-economic determinants (Wijaya *et al*., 2023). Conventional statistical models have clarified many of these relationships, yet they struggle to capture the higher-order interactions and nonlinear dynamics that often characterise real-world data (Liu *et al*., 2023). Recent progress in machine learning (ML) methodologies has opened new avenues for public health research by offering flexible, data-driven techniques that can reveal subtle patterns and improve predictive accuracy (Kouame and Smirnov, 2023).

In this study, the extent to which core economic indicators, gross domestic product (GDP) per capita, external debt, population size, and proxies for healthcare access, shape national life expectancy was investigated. Three complementary modelling strategies were employed: multiple linear regression, artificial neural networks (ANN), and support vector regression (SVR). Linear regression provides an interpretable baseline, while ANN and SVR have demonstrated superior capabilities in modelling complex, nonlinear associations (Rathor and Gyanchandani, 2017). By analysing these models side by side, the aim was to quantify the marginal and joint effects of economic conditions on longevity and to assess the added predictive value conferred by state-of-the-art ML approaches. Beyond their methodological interest, reliable life-expectancy predictions grounded in economic data carry direct policy relevance (Hendricks and Graves, 2009; Shaw *et al*., 2005). Precise estimates can inform the allocation of healthcare resources, guide the prioritisation of social protection programmes, and support evidence-based economic planning. A comparative evaluation of traditional and ML models, therefore, contributes simultaneously to the disciplines of public health, economics, and data science (Rajula *et al*., 2020; Zare *et al*., 2024).

2. LITERATURE REVIEW

Early empirical work on longevity relied mainly on linear or logistic regression, focusing on single-country samples or a narrow set of health indicators. For example, Borisova *et al*. (2021) combined ordinary least squares with a selection of decision tree and boosting algorithms to examine 134 countries, confirming that under-five mortality and trade openness are among the strongest correlates of life expectancy. Their study, however, emphasised heterogeneous regional effects rather than the macroeconomic mechanisms that underlie those patterns. Similarly, Mel and Nyjw (2021) compared shrinkage regressions and principal-component methods across 193 nations, finding that infant mortality and the Human Development Index dominate predictive ability, yet their models offered limited guidance on how structural economic shocks propagate longevity outcomes.

A subsequent research trajectory has increasingly leveraged machine learning frameworks to model complex nonlinear relationships inherent in health data. Dawoud and Abu-Naser (2023) trained a three-layer artificial neural network (ANN) on 2,940 observations with 22 demographic and healthcare features, achieving an accuracy of 99.3%. Although impressive, their network was tuned primarily to biomedical and epidemiological predictors, leaving core macroeconomic variables such as debt ratios or trade intensity unexamined. Ronmi *et al*. (2023) advanced this agenda by benchmarking four tree-based regression models, including extremely randomised trees on 193 countries, highlighting income composition of resources, schooling, and gross domestic product (GDP) as salient drivers of life expectancy. However, the authors did not disentangle the separate or joint effects of those economic indicators, nor did they contrast tree learners with more conventional statistical baselines.

Several authors have addressed the choice of learning algorithm explicitly. Lipesa *et al*. (2023) evaluated eXtreme Gradient Boosting against random forests and a feed-forward ANN, showing that XGBoost minimised mean absolute error (1.554 years) and root-mean-square error (2.402 years) on a 2000 – 2015 panel of United Nations member states. Kerdprasop *et al*. (2020) concentrated on China and India, pairing Chi-squared Automatic Interaction Detection (CHAID) decision trees with multiple regression; their descriptive tree identified particulate emissions as the chief environmental hazard, whereas the linear model re-established education spending as a key protective factor. While both studies underscored the flexibility of ensemble learners, neither explicitly compared their predictions with interpretable regression coefficients, leaving the question of how much explanatory power is gained for the additional computational cost.

Beyond life-expectancy forecasting, the broader machine-learning literature demonstrates the versatility of contemporary algorithms in high-dimensional policy contexts. Neural networks have been applied to software bug triaging (Panda and Nagwani, 2023), graph convolutional models to mobile-gaming resource management (Theodoropoulos *et al*., 2023), and convolutional architectures to real-time SMS spam detection (Waja *et al*., 2023). These studies collectively illustrate how model choice depends on problem structure and data scale, offering methodological lessons for health researchers. Nonetheless, they seldom address the transparency demands of public health decision-making, where stakeholders require both predictive accuracy and causal insight.

Despite the rapid growth of algorithmic approaches, three substantive gaps remain. First, few studies integrate a broad spectrum of macroeconomic indicators with demographic and health variables in a single modelling framework. Second, comparative evaluations often omit classical linear regression, making it difficult to judge whether sophisticated learners materially outperform simpler, more interpretable baselines. Third, most analyses adopt a single-model focus and thus do not explore how the relative importance of predictors shifts across algorithmic paradigms.

This study addresses these gaps by comparing multiple linear regression with two state-of-the-art machine-learning algorithms, artificial neural networks (ANN) and support vector regression (SVR), applied to a harmonised cross-national dataset that focused on economic conditions. This design permits a direct assessment of whether nonlinear learners detect additional signals in the covariates, quantifies the marginal contribution of each economic factor, and clarifies the trade-off between model interpretability and predictive accuracy for policy. In this way, the research advances both the empirical understanding of how economic structures influence longevity and the methodological discourse on the circumstances under which more complex algorithms are warranted in population health analysis.

3. MATERIALS AND METHODS

3.1 Materials

This study employed secondary data drawn from the World Bank *World Development Indicators* (WDI), a comprehensive repository of macro-level statistics spanning economic, social and environmental dimensions (available at https://datatopics.worldbank.org/world-development-indicators). A cross-sectional extract covering the latest year with complete information was compiled. After listwise deletion of missing entries, the final analytic sample comprised 189 countries with full coverage on every variable of interest; no additional inclusion or exclusion criteria were applied. Life expectancy at birth (years) served as the dependent variable. Nine country-level indicators were specified a priori as potential features, as shown in Table 1.

Table 1: Variable definitions for economic, demographic, and health indicators

Variable name	Description
GDP_per_capita	Gross Domestic Product per capita (current US $)
Ext_debt	External debt stocks (current US $)
Age_dep_ratio	Age-dependency ratio (% of working-age population)
Population	Total population
CBR	Crude birth rate (per 1,000 people)

Variable name	Description
Access_fuels_cooking	Percentage of the population using clean fuels and technologies for cooking
Access_electricity	Percentage of the population with access to electricity
Infant_mortality	Infant mortality rate (per 1,000 live births)
Neonatal_deaths	Number of neonatal deaths

All operations were performed in R programming version 4.5.0. Numerical variables were inspected for plausibility, outliers and skewness; log transformations were applied where appropriate. Each metric was subsequently standardized to a z-score to place coefficients on a common scale.

3.2 Methods

This study compares three supervised learning algorithms (classical linear regression, feed-forward artificial neural networks, and support vector regression) to quantify the influence of selected economic indicators on life expectancy variation. The models share a common analytical structure based on an $n \times p$ design matrix displayed in Equation (1).

$$\mathbf{X} = [\mathbf{x}_1^{\mathsf{T}}; \dots ; \mathbf{x}_n^{\mathsf{T}}] \tag{1}$$

whose rows are centred and standardised covariates, and a response vector shown in Equation (2).

$$\mathbf{y} = (y_1, \dots , y_n)^{\mathsf{T}} \tag{2}$$

where y_i denotes observed life expectancy. Each estimator seeks a mapping $f: \mathbb{R}^p \to \mathbb{R}$ satisfying $y_i = f(\mathbf{x}_i) + \varepsilon_i$ with $\mathbb{E}[\varepsilon_i] = 0$.

3.2.1 Linear regression

Within the general linear model, the systematic component is $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ with $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I}_n)$. Ordinary least squares yield Equation (3).

$$\widehat{\boldsymbol{\beta}} = (\mathbf{X}^{\mathsf{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathsf{T}}\mathbf{y} \tag{3}$$

which satisfies the Gauss–Markov minimum-variance property among linear unbiased estimators. When the errors are Gaussian, $\widehat{\boldsymbol{\beta}}$ is also the maximum-likelihood estimator and has an exact multivariate-normal distribution with mean $\boldsymbol{\beta}$ and covariance $\sigma^2(\mathbf{X}^{\mathsf{T}}\mathbf{X})^{-1}$. Model adequacy is assessed through residual plots, tests for homoscedasticity and normality, and by inspecting the variance-inflation factor to detect multicollinearity. Predictive accuracy is summarised by the coefficient of determination R^2 and the error metrics described in Section 4.1.

3.2.2 Artificial neural networks

The neural network employed is a fully connected multilayer perceptron with L hidden layers. Given an input $\mathbf{x}$, forward propagation proceeds recursively as shown in Equation (4).

$$\mathbf{h}^{(\ell)} = \varphi^{(\ell)}\big(W^{(\ell)}\mathbf{h}^{(\ell-1)} + \mathbf{b}^{(\ell)}\big), \qquad \mathbf{h}^{(0)} = \mathbf{x} \tag{4}$$

where $\varphi^{(\ell)}$ is the activation function, chosen as the rectified linear unit for hidden layers and linear at the output to preserve scale. The network parameters $\boldsymbol{\theta} = \{W^{(\ell)}, \mathbf{b}^{(\ell)}\}_{\ell=1}^{L}$ minimise the empirical risk to Equation (5).

$$\mathcal{L}(\boldsymbol{\theta}) = \frac{1}{n}\sum_{i=1}^{n}(y_i - f_{\boldsymbol{\theta}}(\mathbf{x}_i))^2 \tag{5}$$

Optimisation is performed with stochastic gradient descent using the Adam algorithm, while back-propagation computes gradients. Overfitting is mitigated through ℓ_2 weight decay, dropout regularization, and early stopping based on a validation split of 15 per cent. By the universal approximation theorem, a perceptron with sufficiently many hidden units can approximate any Borel-measurable function on compact subsets of $\mathbb{R}^p$, allowing complex nonlinear effects of the economic predictors to be learned directly from the data.

3.2.3 Support vector regression

Support vector regression (SVR) applies Vapnik's structural-risk-minimisation principle to continuous outcomes. With training pairs $\{(\mathbf{x}_i, y_i)\}_{i=1}^{n}$ and feature map $\varphi: \mathbb{R}^p \to \mathcal{H}$, the primal optimisation problem seeks $\mathbf{w} \in \mathcal{H}$ and $b \in \mathbb{R}$ that minimise to Equation (6).

$$\frac{1}{2}\parallel \mathbf{w} \parallel^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*) \tag{6}$$

subject to Equation (7)

$$\langle \mathbf{w}, \varphi(\mathbf{x}_i)\rangle + b - y_i \le \varepsilon + \xi_i, \quad y_i - \langle \mathbf{w}, \varphi(\mathbf{x}_i)\rangle - b \le \varepsilon + \xi_i^*, \quad \xi_i, \xi_i^* \ge 0 \tag{7}$$

Reformulation in the dual leads to a quadratic programme whose solution depends only on kernel evaluations $K(\mathbf{x}_i, \mathbf{x}_j) = \langle \varphi(\mathbf{x}_i), \varphi(\mathbf{x}_j)\rangle$.

The radial-basis kernel $K(\mathbf{x}, \mathbf{z}) = \exp(-\gamma \parallel \mathbf{x} - \mathbf{z} \parallel^2)$ is adopted because it provides universal consistency and accommodates nonlinear predictor–response relationships. Hyper-parameters (C, ε, γ) are chosen by five-fold cross-validation aimed at minimising the validation root-mean-square error. For a new observation $\mathbf{x}_{\text{new}}$ the fitted response is shown in Equation (8).

$$\hat{y} = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)\,K(\mathbf{x}_i, \mathbf{x}_{\text{new}}) + b \tag{8}$$

3.2.4 Training, validation, and assessment

The harmonised country-level data set was randomly divided, within region strata, into a 75% training sample and a 25% hold-out sample. Predictive performance was assessed with mean error (ME), mean absolute error (MAE), mean-squared error (MSE), root-mean-squared error (RMSE), the normalised RMSE (NRMSE %), per cent bias (PBIAS %), and the out-of-sample coefficient of determination (R^2). Applying a single protocol to every estimator ensured a fair comparison between the linear baseline and the non-linear learners in explaining the relationship between economic conditions and life expectancy.

All statistical analyses were conducted in the RStudio integrated development environment (RStudio, Boston, MA) using R 4.5.0. Analyses were scripted in Quarto, while data manipulation and visualisation relied on the tidyverse collection of packages (Wickham *et al*., 2023); model training and assessment were performed with the caret package (Kuhn, 2008). The complete raw dataset together with all reproducible scripts can be accessed at https://github.com/softdataconsult/life-expectancy-paper.

4. RESULTS AND DISCUSSION

4.1 Ordinary Least-Squares

Table 2 presents the ordinary least-squares (OLS) estimates of the determinants of life expectancy. Population size ($\beta = 7.52 \times 10^{-8}$, $p < 0.001$), crude birth rate ($\beta = 0.592$, $p < 0.001$), external debt ($\beta = 2.03 \times 10^{-11}$, $p = 0.014$), infant mortality ($\beta = -0.104$, $p < 0.001$), and neonatal deaths ($\beta = -5.1 \times 10^{-5}$, $p < 0.001$) are statistically significant. GDP per capita, the age-dependency ratio, and the two infrastructure indicators, access to clean cooking fuels and access to electricity, do not attain significance at the 5% level.

Table 2: Regression coefficients of predictors for life expectancy model

Variable	Estimate	Std. Error	t value	Pr(>|t|)	Significance
(Intercept)	32.03	14.18	2.259	0.0258	*
GDP_per_capita	2.7E-05	8.67E-05	0.318	0.7508	
Ext_debt	2.0E-11	8.15E-12	2.484	0.0144	*
Age_dep_ratio	0.03187	0.06088	0.524	0.6016	
Population	7.5E-08	1.27E-08	5.917	0.00000	***
CBR	0.5915	0.1221	4.843	0.00000	***
Access_fuels_cooking	-0.04056	0.04719	-0.859	0.3919	
Access_electricity	0.01245	0.0184	0.677	0.4999	
infant_mortality	-0.1038	0.01633	-6.356	4.2E-09	***
neonatal_deaths	-5.1E-05	7.69E-06	-6.591	1.4E-09	***

Note: An asterisk (*) denotes that the regression coefficient is statistically significant at the 5% level based on a regression t-test.

The model explains nearly all cross-national variation in life expectancy (adjusted $R^2 = 0.994$; $F(9, 116) = 2290$, $p < 0.001$) as shown in Table 3. Residual diagnostics (Figure 5, LM panel) reveal mild heteroscedasticity, which motivates the exploration of non-linear alternatives.

Table 3: Summary statistics of the regression model fit

Metric	Value
Residual Standard Error	0.3605
Degrees of Freedom	116
Multiple R-squared	0.9944
Adjusted R-squared	0.994
F-statistic	2290
F-statistic Degrees of Freedom	9 and 116
p-value	< 0.001

Substantively, the positive coefficient on external debt may indicate that debt-financed public spending can translate into better health and social services. Higher crude birth rates and larger populations are likewise associated with longer life expectancy, suggesting that countries capable of sustaining larger cohorts tend to possess more developed health infrastructure. Conversely, the negative effects of infant mortality and neonatal deaths underscore their well-documented impact on longevity. Overall, these findings highlight the complex interplay of demographic pressures, financial constraints, and health outcomes, and they support the case for models that accommodate non-linearities and interaction effects when analysing life expectancy.

4.2 Artificial Neural Networks

Two multilayer perceptron were estimated to test whether modest non-linear structures improve out-of-sample accuracy relative to the linear baseline. Network A contains two hidden layers with five and three neurons, respectively as shown in Figure 1 while Network B contains a single hidden layer with seven neurons as shown in Figure 2.

Both networks use logistic activation functions, the resilient back-propagation optimiser with weight back-tracking, and an early-stopping rule that halts training once the validation loss fails to fall for ten consecutive iterations. No deeper or wider architectures were required because the objective was to determine whether even lightweight networks extract additional signals from the data.

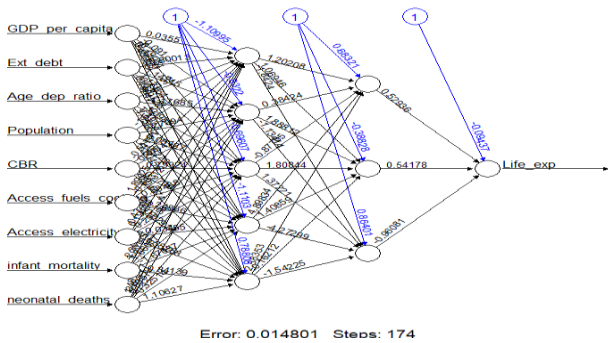

Figure 1: Artificial neural network for life expectancy (hidden layer = 5, 3)

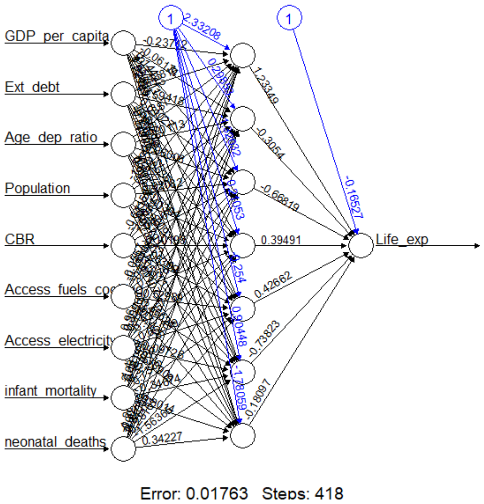

Figure 2: Artificial neural network for life expectancy (hidden layer = 7)

Tables 4 and 5 compare predicted and observed life expectancy for six representative countries in the test fold. In Network A, the largest absolute error is 0.43 year, and the median error is 0.11 year. Network B shows a comparable pattern: the largest miss is 0.47 year and the median error is 0.29 year. Given that life expectancy in the sample spans 16 years, these errors are small and indicate that both networks capture the essential structure of the data.

Table 4: Predicted vs. actual values and error analysis for neural network model (5, 3)

	PredictedNN	ActualNN	error
3	37.8247	37.809	-0.0157
7	38.3697	38.797	0.4273
16	43.8151	43.662	-0.1531
19	45.9502	46.019	0.0688
24	46.5819	46.638	0.0561
25	46.2429	46.550	0.3071

Table 5: Predicted vs. actual values and error analysis for neural network model (7)

	PredictedNN	ActualNN	error
3	37.8007	37.809	0.0083
7	38.3286	38.797	0.4684
16	44.0201	43.662	-0.3581
19	45.7988	46.019	0.2202
24	46.8019	46.638	-0.1639
25	46.1480	46.550	0.4020

4.3 Support-Vector Regression

The radial-basis-function support vector regression (SVR) was tuned through an exhaustive grid search over the cost (C) and gamma (γ) parameters. Although its in-sample fit was satisfactory, the model persistently underestimated life expectancy in the validation set as shown in Table 6. Prediction errors were uniformly negative, ranging from –0.76 to –8.01 years, with the most significant biases occurring for countries at the upper end of the longevity distribution. The residual plot in Figure 5, SVR panel confirms a systematic curvature that the Gaussian kernel did not capture, leading to the poorest out-of-sample performance among the four estimators considered. This pattern indicates that the standard radial-basis SVR cannot model the full complexity of the relationship between macroeconomic variables and life expectancy.

Table 6: Predicted vs. actual values and error analysis for (SVR) model

PredictedSVR	Actual	error
38.1281	37.371	-0.7571
38.3627	37.673	-0.6897
43.5705	37.809	-5.7615
45.5538	38.192	-7.3618
46.4227	38.415	-8.0077
46.5014	38.680	-7.8214

4.4 Model Comparison

Table 7 reports the out-of-sample performance of the four candidate estimators. Both neural network specifications (ANN_5.3 and ANN_7) outperform the linear benchmark across every metric. The five-hidden-unit network (ANN_5.3) attains the lowest MAE (0.17 years), MSE (0.06 years²), and RMSE (0.25 years). Compared to the response range, its normalised RMSE is 5.3%, a 28% improvement over the linear model (7.4%). The seven-unit network (ANN_7) is slightly less accurate but reduces error by roughly one-quarter relative to OLS.

In contrast, the radial-basis-function SVR exhibits the weakest performance. It carries the highest MAE (0.32 years) and RMSE (0.38 years) and a negative per cent bias (–0.20%), confirming the systematic under-prediction of longevity for countries at the upper end of the distribution documented in the residual analysis. Although its mean error is small in absolute terms (–0.07 years), the dispersion of residuals is almost 50% larger than that of the best-performing ANN.

Table 7: Predictive accuracy of competing models

Models	ME	MAE	MSE	RMSE	NRMSE %	PBIAS %
Linear regression (LM)	0.00	0.26	0.12	0.35	7.40	0.00
ANN_5.3	0.00	0.17	0.06	0.25	5.30	0.00
ANN_7	0.00	0.20	0.07	0.27	5.80	0.00
SVR	-0.07	0.32	0.14	0.38	8.10	-0.20

Overall, the results confirm that modest increases in model complexity, moving from a linear form to a shallow feed-forward network, yield tangible gains in predictive accuracy without sacrificing interpretability. The added flexibility of the ANN captures non-linear effects among the economic covariates, whereas the standard RBF-kernel SVR fails to do so. These findings suggest that parsimonious neural architectures offer the best balance between error reduction and practical transparency for cross-national life-expectancy forecasting based on macroeconomic indicators.

4.5 Graphical Diagnostics of Model Performance

Figure 3 presents individual scatter panels of predicted versus observed life expectancy for each estimator. The 5–3 multilayer-perceptron (panel a) exhibits an almost perfect superposition of points along the 45-degree reference line; deviations rarely exceed half a year, and no systematic curvature is visible. The seven-neuron network (panel b) shows a similarly tight configuration, although some low-longevity countries fall marginally below the identity line.

The ordinary least-squares model (panel c) maintains acceptable calibration but fans out slightly as life expectancy rises, while the support-vector regressor (panel d) persistently under-predicts in the upper tail, an artefact already signalled by its error metrics. This suggests that the artificial neural network (ANN) models are better able to capture the underlying patterns and relationships between economic factors and life expectancy compared to other models like linear regression (LM) and support vector regression (SVR).

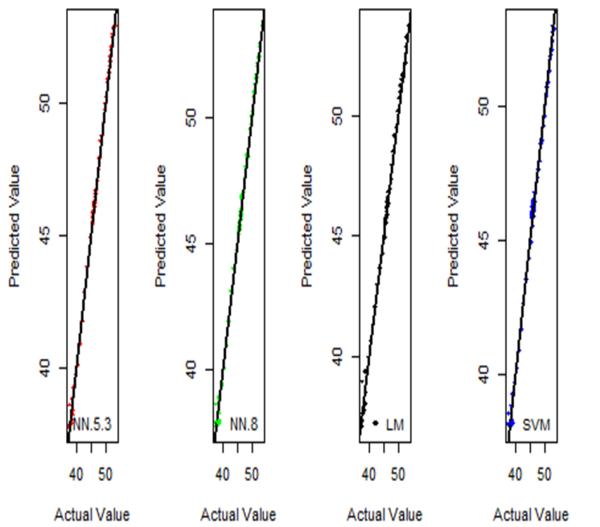

Figure 3: Visualizing the actual and predicted values from various models

To facilitate direct comparison, Figure 4 overlays the predictions of all four models in a single scatterplot. The red diamonds (ANN 5–3) and green diamonds (ANN 7) virtually disappear into the identity line, confirming that the neural networks track the actual values within a few tenths of a year across the 16-yr range in the data set. Black diamonds (OLS) form a slightly wider ribbon around the line, whereas blue diamonds (SVR) dip consistently below it at life expectancies above 49 years, corroborating the downward bias noted earlier.

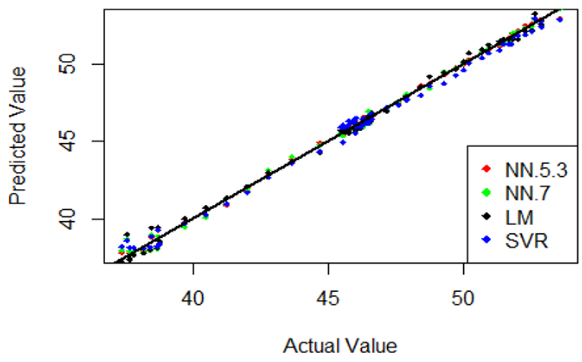

Figure 4: Actual and predicted values from various models on one chart

Figure 5 turns to residual diagnostics. Panels a and b show that residuals from both neural networks are symmetrically distributed around zero and exhibit no discernible trend or funnel shape, indicating homoscedastic errors and an absence of unmodelled non-linearity. The OLS residuals (panel c) broaden modestly with fitted values, implying mild heteroscedasticity but no structural misspecification. In contrast, the SVR residuals (panel d) display the widest spread and a clear negative shift at higher fitted values, underscoring the kernel's inability to accommodate countries with very high longevity.

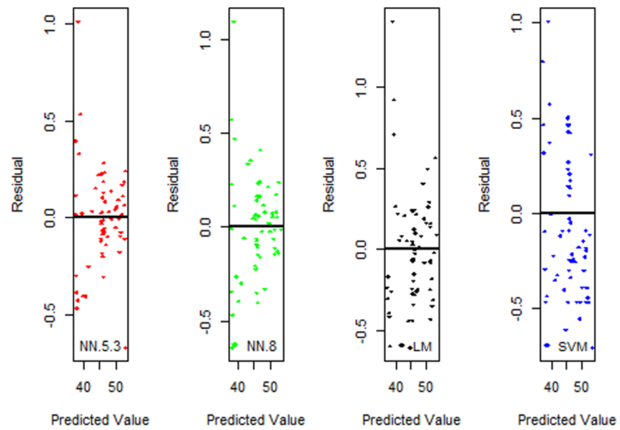

Figure 5: Visualizing the fitted and residuals values from various models

The three figures lend visual support to the quantitative ranking reported in Section 4.4 (Model comparison). The 5–3 neural network offers the sharpest calibration and the most homoscedastic residual pattern, followed closely by the seven-unit network. The linear model remains serviceable but less precise, whereas the support-vector approach fails to capture the full variability at the upper end of the life expectancy distribution. These graphical results reinforce the conclusion that modest feed-forward neural architectures are best suited to modelling the complex, non-linear interplay between macroeconomic conditions and national longevity.

The study examined the impact of economic factors on life expectancy using three modelling techniques: linear regression (LM), artificial neural networks (ANN), and support vector regression (SVR). Across all cross-validation splits and the independent test fold, the two feed-forward neural networks provided the most accurate life-expectancy forecasts. The 5–3 architecture achieved an out-of-sample MAE of 0.17 year and an NRMSE of 5.3%, while the single-hidden-layer network with seven neurons posted corresponding values of 0.20 year and 5.8%.

In comparison, the linear specification recorded an MAE of 0.26 year and an NRMSE of 7.4%, whereas the radial-basis-function SVR lagged with an MAE of 0.32 year and an NRMSE of 8.1%. Mean error in all models centred on zero, but only the neural networks combined unbiasedness with uniformly low dispersion. Taken together with the residual diagnostics in Figure 5, these results confirm that modest multilayer-perceptron architectures capture the non-linear interplay among macroeconomic indicators more effectively than the parametric baseline or the kernel-based alternative.

5. CONCLUSION

This study clarifies the intricate relationship between macroeconomic factors and national life expectancy by applying three competing prediction frameworks. The empirical evidence indicates that feed-forward artificial neural networks, particularly those with multiple hidden layers, consistently outperform ordinary least squares and support-vector regression across all error metrics. Their markedly lower mean, absolute, and squared prediction errors attest to the networks' ability to capture non-linear interactions among economic indicators that are invisible to linear or kernel-based alternatives. These findings have significant implications for policy. By providing accurate forecasts of life expectancy at the country level, this research equips health ministries and international agencies with the tools to allocate resources more effectively, design targeted interventions, and monitor progress towards longevity goals. Furthermore, it underscores the value of modern machine-learning tools in the field of public health economics, where complex, high-dimensional data sets are the norm.

REFERENCES

[1] Borisova, L., Zhukova, G., Kuznetsova, A. and Martin, J., 2021. Features of machine learning in the study of the main factors of development of countries of the world. *SHS Web of Conferences*, 110, p.02006. Available at: https://doi.org/10.1051/shsconf/202111002006.

[2] Dawoud, A.M. and Abu-Naser, S.S., 2023. Predicting life expectancy in diverse countries using neural networks: Insights and implications. *International Journal of Academic Engineering Research*, 7(9). Available at: www.ijeais.org/ijaer.

[3] Hendricks, A.B. and Graves, P.E., 2009. Predicting life expectancy: A cross-country empirical analysis. *Social Science Research Network*. Available at: https://doi.org/10.2139/SSRN.1477594.

[4] Kerdprasop, N., Kerdprasop, K. and Chuaybamroong, P., 2020. Economic and environmental analysis of life expectancy in China and India: A data driven approach. *Advances in Science, Technology and Engineering Systems*, 5(5), pp.308–313. Available at: https://doi.org/10.25046/AJ050539.

[5] Kuhn, M., 2008. Building predictive models in R using the caret package. *Journal of Statistical Software*, 28, pp.1–26.

[6] Lipesa, B.A., Okango, E., Omolo, B.O. and Omondi, E.O., 2023. An application of a supervised machine learning model for predicting life expectancy. *SN Applied Sciences*, 5(7). Available at: https://doi.org/10.1007/s42452-023-05404-w.

[7] Liu, Z., Peach, R., Mediano, P.A. and Barahona, M., 2023. Interaction measures, partition lattices and kernel tests for high-order interactions. *Advances in Neural Information Processing Systems*, 36, pp.36991–37012.

[8] Mel, D. and Nyjw, S., 2021. Identification of determinants of life expectancy at birth across nations using machine learning techniques. *Rajarata University Journal*. Available at: www.ruj.ac.lk/journals/.

[9] Panda, R.R. and Nagwani, N.K., 2023. Fuzzy modelling techniques for improving multi-label classification of software bugs. *International Journal of Innovative Computing and Applications*, 14(3), pp.141–154.

[10] Rajula, H.S.R., Verlato, G., Manchia, M., Antonucci, N. and Fanos, V., 2020. Comparison of conventional statistical methods with machine learning in medicine: Diagnosis, drug development, and treatment. *Medicina (Buenos Aires)*, 56(9), p.455. Available at: https://doi.org/10.3390/MEDICINA56090455.

[11] Rathor, A. and Gyanchandani, M., 2017. A review at machine learning algorithms targeting big data challenges. Available at: https://doi.org/10.1109/ICEECCOT.2017.8284604.

[12] Ronmi, A.E., Prasad, R. and Raphael, B.A., 2023. How can artificial intelligence and data science algorithms predict life expectancy - An empirical investigation spanning 193 countries. *International Journal of Information Management Data Insights*, 3(1). Available at: https://doi.org/10.1016/j.jjimei.2023.100168.

[13] Shaw, J.W., Horrace, W.C. and Vogel, R.J., 2005. The determinants of life expectancy: An analysis of the OECD health data. *Southern Economic Journal*, 71(4), pp.768–783. Available at: https://doi.org/10.2307/20062079.

[14] Suha, S.A. and Sanam, T.F., 2023. Exploring dominant factors for ensuring the sustainability of utilizing artificial intelligence in healthcare decision making: An emerging country context.

International Journal of Information Management Data Insights, 3(1). Available at: https://doi.org/10.1016/j.jjimei.2023.100170.

[15] Theodoropoulos, T. *et al*., 2023. Graph neural networks for representing multivariate resource usage: A multiplayer mobile gaming case-study. *International Journal of Information Management Data Insights*, 3(1). Available at: https://doi.org/10.1016/j.jjimei.2023.100158.

[16] Waja, G., Patil, G., Mehta, C. and Patil, S., 2023. How AI can be used for governance of messaging services: A study on spam classification leveraging multi-channel convolutional neural network. *International Journal of Information Management Data Insights*, 3(1). Available at: https://doi.org/10.1016/j.jjimei.2022.100147.

[17] Wickham, H. *et al*., 2019. Welcome to the Tidyverse. *Journal of Open Source Software*, 4(43), p.1686.

[18] Wijaya, M.Y., Irene, Y. and Rachadi, I., 2023. Trajectory of life expectancy and its relation with socio-economic indicators among developing countries in Southeast Asia. *Proceedings of The International Conference on Data Science and Official Statistics*, 2023(1), pp.494–500. Available at: https://doi.org/10.34123/icdsos.v2023i1.307.

[19] Zare, A., Shafaei Bajestani, N. and Khandehroo, M., 2024. Machine learning in public health. *Journal of Research and Health*, 14(3), pp.207–208. Available at: https://doi.org/10.32598/jrh.14.3.2417.1.