Predictive Modelling of Claim Frequency Social Protection (SOCSO'S Survivors' Pension Benefits)

M. Z. A. Chek[1], I. L. Ismail[2], M. S. Asrulsani[3], H. Hasim[4], Z. H. Zulkifli[5]

[1],[3]Actuarial Science Department, UiTM Perak Branch

[2]Department of Statistics and Decision Science, UiTM Perak Branch

[4]School of Mathematical & Computer Sciences Heriot-Watt University, UK

[5]Actuarial Partners Consulting, Malaysia

DOI: https://dx.doi.org/10.47772/IJRISS.2025.90300067

Received: 23 February 2025; Accepted: 27 February 2025; Published: 31 March 2025

ABSTRACT

This study employs predictive analytics to forecast the frequency of claims for SOCSO's Survivors' Pension Benefits from 2025 to 2030. Using historical data from 1985 to 2024, we develop predictive models employing ARIMA, decision trees, and artificial neural networks (ANN). The study aims to provide a comprehensive understanding of claim frequency trends and their potential impact on SOCSO's financial sustainability. The ARIMA model captures time-series patterns, decision trees identify key determinants influencing claim fluctuations, and ANN detects complex, non-linear relationships within the dataset. The study's findings indicate a continuous upward trajectory in claim numbers, with projections revealing an annual growth rate of approximately 5.5%. These results highlight the necessity for SOCSO to adjust its financial strategies, optimize actuarial assumptions, and implement data-driven policy interventions to ensure the long-term sustainability of the fund. By integrating predictive analytics, SOCSO can proactively manage risk, optimize contribution rates, and maintain financial stability while continuing to provide adequate support for beneficiaries. The insights generated from this study will be instrumental for policymakers in making informed decisions regarding future pension fund allocations and social security planning.

Keywords - Survivors' Pension, Predictive Analytics, Time-Series, ARIMA, Neural Networks

INTRODUCTION

The SOCSO Survivors' Pension Benefit is a crucial component of Malaysia's social security system, designed to provide financial support to the dependents of insured individuals upon their passing. The scheme ensures continued financial stability for beneficiaries, mitigating the economic challenges associated with the loss of a primary income earner. Over the past four decades, claim frequencies have increased significantly due to demographic shifts, economic transformations, and policy amendments that have broadened the scope of social security coverage [1].

From 1985 to 2024, SOCSO's Survivors' Pension Benefit witnessed an exponential rise in claims, reaching an all-time high of 383,101 in 2024. Various factors have contributed to this trend, including increased workforce participation, extended life expectancy, and evolving labor market dynamics. While the scheme continues to serve as a vital financial safety net, the surge in claims poses significant challenges to SOCSO's financial sustainability. Policymakers must anticipate future claim trends to ensure that the fund remains viable and can fulfill its long-term obligations to beneficiaries [2].

This study aims to apply predictive analytics techniques to forecast claim frequencies from 2025 to 2030. By leveraging time-series forecasting models such as ARIMA, decision trees, and artificial neural networks (ANN), this research seeks to provide policymakers and financial planners with a robust framework for anticipating future pension fund liabilities. The findings will be instrumental in optimizing resource allocation,

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume IX Issue III March 2025

adjusting contribution rates, and formulating evidence-based policy recommendations to safeguard the financial stability of SOCSO's Survivors' Pension Benefit [3]–[5].

LITERATURE REVIEW

Survivors' Pension Schemes

Survivors' pension schemes exist globally as a crucial aspect of social security systems, providing financial protection to the dependents of deceased insured individuals. Various countries have implemented different models, ranging from defined-benefit schemes to contribution-based programs, each with unique eligibility criteria and payout structures. Studies indicate that Malaysia's SOCSO scheme aligns with global best practices but faces sustainability challenges due to increasing claims. Comparative research highlights that actuarial forecasting and data-driven adjustments play a vital role in maintaining pension fund solvency [6]–[10].

Predictive Analytics in Social Security

Predictive analytics has been widely used in social security systems to optimize financial planning and forecast future pension fund liabilities. Techniques such as ARIMA models, machine learning algorithms, and regression analyses have been employed to predict claim trends and ensure that social security funds remain financially viable. Research indicates that implementing data-driven actuarial models can significantly improve the efficiency and accuracy of forecasting pension claims. Additionally, studies show that machine learning techniques, including neural networks, can identify complex patterns in social security data that traditional statistical models may overlook [11]–[15].

Historical Trends in SOCSO Claims

SOCSO's historical data from 1985 to 2024 reveals a steep upward trend in the number of Survivors' Pension Benefit claims. This trend has been influenced by demographic shifts, policy changes, and socio-economic factors such as rising mortality rates and increased workforce participation. Past research has examined the impact of these variables on claim rates, concluding that predictive modeling can help anticipate future financial burdens on the social security system. The growing dependency on SOCSO's pension benefits necessitates a proactive approach to managing its long-term sustainability [16]–[19].

METHODOLOGY

Data Collection

This study utilizes a comprehensive dataset obtained from SOCSO's official records and publicly available sources, spanning from 1985 to 2024. The dataset includes historical claim frequencies, demographic attributes, and relevant socio-economic indicators that influence claim patterns. Data cleaning and preprocessing were conducted to ensure accuracy and consistency, including handling missing values, normalizing numerical variables, and detecting anomalies that could affect model predictions [20].

Forecasting Techniques

Three key predictive analytics techniques were applied to model future claim trends:

Autoregressive Integrated Moving Average (ARIMA) Model:

A statistical method for time-series forecasting analyzes historical claim data to capture linear trends, seasonality, and stochastic components affecting claim frequencies. The ARIMA model was optimized through parameter tuning, including determining the order of autoregression (p), differencing (d), and moving average (q) to achieve the lowest error rate [21].

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume IX Issue III March 2025

Decision Trees:

A supervised learning technique identifies significant factors contributing to variations in claim frequencies. The model was trained on historical data to classify and predict claim behaviors based on key determinants such as age group, income level, and policy changes. Analyzing feature importance analysis was conducted to enhance model interpretability and improve decision-making for SOCSO policymakers [22].

Artificial Neural Networks (ANN): A deep learning approach used to capture complex non-linear relationships within the dataset. The ANN model was structured with multiple hidden layers, activation functions, and backpropagation to refine prediction accuracy. Hyperparameter tuning, including learning rate adjustments and dropout regularization, was employed to enhance model performance [5], [23].

Model Evaluation Metrics

To assess the accuracy and reliability of each forecasting model, the following evaluation metrics were employed:

Mean Absolute Percentage Error (MAPE):

Measures the average percentage deviation of predicted values from actual values, providing insight into model accuracy [24].

Root Mean Square Error (RMSE):

Quantifies the standard deviation of residuals, highlighting the dispersion between observed and predicted values [25].

Coefficient of Determination (R^2): Evaluates the proportion of variance in claim frequency explained by the model, indicating predictive strength. Comparative analysis of model performance was conducted to determine the most effective approach for forecasting claim frequencies from 2025 to 2030 [26].

RESULTS AND DISCUSSION

The results of this study provide an in-depth analysis of the trends in claims for SOCSO's Survivors' Pension Benefit over a span of forty years from 1985 to 2024 (Table 1).

Table 1 Frequency Socso's Survivors' Pension Benefit

Year	Fre	Year	Fre	Year	Fre
1985	432	**1999**	75,189	**2012**	191,006
1986	2,145	**2000**	82,113	**2013**	203,454
1987	4,341	**2001**	88,281	**2014**	216,001
1988	7,990	**2002**	94,890	**2015**	231,279
1989	11,760	**2003**	106,531	**2016**	249,018
1990	17,813	**2004**	114,336	**2017**	268,540
1991	22,820	**2005**	120,655	**2018**	284,885
1992	26,750	**2006**	129,515	**2019**	305,227

1993	32,338	**2007**	138,717	**2020**	320,520
1994	36,005	**2008**	145,178	**2021**	340,957
1995	43,377	**2009**	160,179	**2022**	357,359
1996	51,043	**2010**	175,564	**2023**	366,722
1997	59,194	**2011**	182,713	**2024**	383,101
1998	67,034				

The total number of claims increased substantially over time, showing a steady rise in the early years followed by a rapid escalation after 2000. The statistical analysis highlights significant fluctuations in claim numbers, influenced by socio-economic conditions, policy changes, and demographic shifts [10], [27].

Forecasted Claim Frequency (2025-2030)

The forecasting models used in this study projected a continuous increase in SOCSO Survivors' Pension claims from 2025 to 2030. The ARIMA model, which captures time-series trends, estimated that the number of claims would rise steadily, while the Decision Tree model identified critical factors contributing to claim variations. The ANN model, leveraging non-linear relationships, provided additional insights into the claim frequency trajectory [28]. The projected claim frequencies are as follows (Table 2):

Table 2 The Forecasted Claim Frequencies

Year	ARIMA Forecast	Decision Tree Forecast	ANN Forecast
2025	402,500	405,200	408,300
2026	423,100	428,000	432,500
2027	447,800	454,300	460,000
2028	475,600	482,700	489,500
2029	507,300	515,100	523,200
2030	542,900	550,800	560,100

These projections indicate an average annual increase of approximately 5.5%, reflecting demographic and economic changes influencing SOCSO's beneficiaries.

Model Performance Evaluation

To ensure the reliability of the predictions, the study assessed model accuracy using multiple performance metrics. The evaluation of the last five years of historical data revealed the following results:

Table 3 Model Performance Evaluation

Model	MAPE (%)	RMSE	R^2 Score
ARIMA	3.5	9,800	0.94

Decision Tree	4.1	10,500	0.92
ANN	3.2	8,900	0.96

The ANN model exhibited the highest accuracy, achieving the lowest Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE), along with the highest R^2 score, suggesting that it effectively captured complex patterns in claim data. ARIMA, while reliable for time-series trends, performed slightly less accurately than ANN. The Decision Tree model, though useful for identifying critical factors, had a relatively lower accuracy compared to ANN and ARIMA [29].

Discussion of Trends and Policy Implications

The observed growth in claim frequencies underscores the necessity for SOCSO to adopt proactive financial strategies. The projected increase in claims suggests that contribution rates and fund reserves should be reviewed to ensure the sustainability of the pension scheme. The ANN model's predictive strength highlights the importance of leveraging machine learning techniques for pension fund forecasting. Policymakers should consider integrating such models into SOCSO's financial planning framework to enhance risk management and optimize resource allocation [6].

CONCLUSION AND RECOMMENDATIONS

This study underscores the projected increase in the number of SOCSO's Survivors' Pension Benefit claims over the next six years, highlighting the importance of data-driven policymaking to ensure fund sustainability. The findings indicate that claim frequencies will continue to rise at an estimated annual rate of 5.5%, placing increasing financial pressure on SOCSO's resources [3].

To address these challenges, SOCSO should consider implementing actuarial recalibrations and adjusting contribution rates to ensure long-term sustainability. Additionally, the organization must optimize its financial planning strategies by integrating predictive analytics into decision-making processes. The use of artificial intelligence and machine learning models, such as ANN, will enhance forecasting accuracy and provide deeper insights into future claim trends [2].

Furthermore, policymakers should explore the development of risk mitigation strategies, such as diversifying revenue sources, revising eligibility criteria, and refining payout structures to align with projected demographic and economic trends. Establishing reserve funds that can accommodate claim surges will also be essential in maintaining financial stability [9].

Future research should focus on incorporating macroeconomic indicators, such as employment trends, wage growth, and mortality rates, to refine predictive models further. Comparative studies with global pension schemes could also provide insights into best practices that SOCSO can adopt to strengthen its fund management strategies [5].

By proactively leveraging predictive analytics and actuarial methodologies, SOCSO can ensure the continued sustainability of its Survivors' Pension Benefit while safeguarding the financial well-being of its beneficiaries [7].

ACKNOWLEDGEMENT

We would like to express our sincere appreciation to the Social Security Organization (SOCSO) for providing access to valuable datasets and resources that made this research possible. Special thanks are extended to the actuarial and data analytics teams at SOCSO for their insightful contributions and assistance in data validation. Additionally, we acknowledge the support from Universiti Teknologi MARA (UiTM) Perak Branch for facilitating this research and providing a conducive academic environment. We also appreciate the

collaborative efforts of research peers and statisticians who provided constructive feedback that enhanced the rigor and depth of this study. Lastly, we are grateful to our families and colleagues for their unwavering encouragement throughout this research process.

REFERENCES

1. M. Z. A. Chek and I. L. Ismail, "Retirement Planning Issues, Problems, and Opportunities in Malaysia," Int. J. Res. Innov. Soc. Sci., vol. 7, no. 9, pp. 1926–1931, 2023, doi: 10.47772/IJRISS.2023.71055.
2. I. L. Ismail and M. Z. Awang Chek, "Reframing Social Protection within Development Paradigms: A Study of Malaysia's Survivors' Pension Scheme," Int. J. Res. Innov. Soc. Sci., vol. 8, no. 6, pp. 816–825, 2024, doi: 10.47772/IJRISS.2024.806064.
3. R. I. Ibrahim, N. M. Nordin, and M. Z. A. Chek, "Investing Sensitivity Effects of Actuarial Assumption on Pension Liabilities in Malaysia," Baghdad Sci. J., vol. 18, no. 1, pp. 830–835, 2021.
4. M. Z. A. Chek, I. L. Ismail, H. Hasim, and A. F. Mansor, "Profiling Return – to – Work (RTW) Recipients in Malaysia," Int. J. Acad. Res. Bus. Soc. Sci., vol. 12, no. 7, pp. 925–935, 2022.
5. M. Z. A. Chek and I. L. Ismail, "Understanding strategic enhancements for the coverage, efficiency, and sustainability of the social security system in Malaysia," Int. J. Res. Innov. Soc. Sci., vol. 8, no. 4, pp. 1723–1730, 2024, doi: 10.47772/IJRISS.2024.804122.
6. M. Z. A. Chek and I. L. Ismail, Understanding Social Protection of Survivors' Pension Benefit. Eudoxia Research Centre, 2024.
7. M. Syakir, M. Z. A. Chek, and I. L. Ismail, "Understanding A Long-Term Care towards Ageing Population in Malaysia," Int. J. Acad. Res. Bus. Soc. Sci., vol. 13, no. 12, pp. 4744–4754, 2023, doi: 10.6007/ijarbss/v13-i12/20328.
8. I. L. Ismail, M. Zaki, A. Chek, and M. Syakir, "Understanding the Employment Insurance Scheme in Malaysia," vol. 13, no. 11, pp. 2137–2143, 2023, doi: 10.6007/IJARBSS/v13-i11/19622.
9. M. Z. A. Chek and I. L. Ismail, "Retirement Planning Issues, Problems, and Opportunities in Malaysia," Retire. Plan. Issues, Probl. Oppor. Malaysia, vol. VII, no. 2454, pp. 1926–1932, 2023, doi: 10.47772/IJRISS.
10. S. C. Seng, "Social Security : Challenges and issues," Kuala Lumpur, 2014–1, 2014.
11. S. Haberman, "Pension funding modelling and stochastic investment returns," no. March, 1994, [Online]. Available: http://www.sias.org.uk/data/papers/stochastic1.pdf/DownloadPDF
12. A. Sa-ngasoongsong and J. Chongwatpol, "An analysis of diabetes risk factors using data mining approach," pp. 1–11, 2012.
13. D. Y. Lin, "Linear regression analysis of censored medical costs," Biostatistics, vol. 1, no. 1, pp. 35–47, 2000, doi: 10.1093/biostatistics/1.1.35.
14. W. J. Stevenson and C. Ozgur, Introduction to Management Science with spreadsheets, 1st ed. United States: Mc Graw Hill, 2007.
15. I. Kose, M. Gokturk, and K. Kilic, "An interactive machine-learning-based electronic fraud and abuse detection system in healthcare insurance," Appl. Soft Comput. J., 2015, doi: 10.1016/j.asoc.2015.07.018.
16. I.T.Jumaniyazov and A.Xaydarov, "The importance of social insurance in social protection," Sci. Educ. Sci. J., vol. 4, no. 1, pp. 1033–1043, 2023.
17. P. W. Cheong, A. A. Jemain, and N. Ismail, "Practice and pricing in non-life insurance:The malaysian experience," J. Qual. Meas. Anal., vol. 4, no. 1, pp. 11–24, 2008.
18. M. Z. A. Chek and I. L. Ismail, "Issues and Challenges Social Insurance in Malaysia," Int. J. Res. Innov. Soc. Sci., vol. 05, no. 04, pp. 278–281, 2021, doi: 10.47772/ijriss.2021.5410.
19. F. Zulkifli et al., "Time series forecasting of future claims amount of SOCSO's Employment Injury Scheme (EIS)," in AIP Conference Proceedings, 2012, vol. 1482. doi: 10.1063/1.4757502.
20. M. Z. Awang Chek, I. L. Ismail, and N. F. Jamal, "Optimising Contribution Rate for SOCSO ' s Invalidity Pension Scheme : Actuarial Present Value (APV)," Int. J. Eng. Technol., vol. 7, pp. 83–92, 2018, doi: 10.14419/ijet.v7i4.33.23491.
21. S. Haberman and A. Renshaw, "Parametric mortality improvement rate modelling and projecting," Insur. Math. Econ., vol. 50, no. 3, pp. 309–333, 2012.
22. A. Gepp, J. H. Wilson, K. Kumar, and S. Bhattacharya, "A Comparative Analysis of Decision Trees Vis

a-vis Other Computational Data Mining Techniques in Automotive Insurance Fraud Detection," J. Data Sci., 2012.
23. F. M. Liou, Y. C. Tang, and J. Y. Chen, "Detecting hospital fraud and claim abuse through diabetic outpatient services," Health Care Manag. Sci., 2008, doi: 10.1007/s10729-008-9054-y.
24. M. Z. Awang Chek, I. L. Ismail, N. F. Jamal, M. Z. A. Chek, I. L. Ismail, and N. F. Jamal, "Estimating Severity of SOCSO's Invalidity Pension Scheme (IPS)," Int. J. Acad. Res. Bus. Soc. Sci., vol. 11, no. 4, pp. 618–625, 2021, doi: 10.6007/ijarbss/v11-i4/9708.
25. M. Z. A. M. Z. A. Chek, I. L. I. L. Ismail, and N. F. N. F. Jamal, "Descriptive research on SOCSO's Invalidity Pension Scheme (IPS) claims payment," Int. J. Recent Technol. Eng., vol. 8, no. 2 Special Issue 11, pp. 660–663, 2019, doi: 10.35940/ijrte.B1104.0982S1119.
26. Nurulhuda.J, Ho.J.S, and Jamilah.M.M, "A survey of risk of accidents in Malaysia," in 14th Road Engineering Association of Asia and Australasia (REAAA) Conference, 2013, pp. 1–10.
27. M. Z. A. Chek et al., "Univariate time series modeling and an application to future claims amount in SOCSO's invalidity pension scheme," in AIP Conference Proceedings, 2012, vol. 1482. doi: 10.1063/1.4757501.
28. B. Render, J. Ralph M. Stair, M. E. Hanna, and T. S. Hale, Quantitative analysis for management, 12th ed. United States: Pearson Education Limited, 2015.
29. N. Azreen, A. Razak, A. Khamis, M. Asrul, and A. Abdullah, "ARIMA and VAR Modeling to Forecast Malaysian Economic Growth," J. Sci. Technol., vol. 9, no. 3, pp. 16–24, 2017.