Maximizing Public Budget Allocation Efficiency with Machine Learning: Towards Evidence-Based Public Policy

Momoha Hirose
mhirose20@earlham.edu
Earlham College
Richmond, Indiana, USA

ABSTRACT

The implementation of policy technology tools that leverage artificial intelligence to improve decision making is gaining attention in public policy. However, tools that not only evaluate historical data and policies, but also propose and simulate policies based on them, have not been fully developed. This study focuses on the research and development of an algorithm for allocating Japan's national budget based on past data, using four different machine learning models and genetic algorithm optimization methods. It takes as inputs classified categories and Japan's past budgets and economic indicators, such as the country's GDP growth rate and inflation rate, and returns a proposal for specific public budget allocations f for the following year.

1 INTRODUCTION

Budget allocation stands as a pivotal domain where AI holds immense promise as a tool to bolster government decision-making [13]. Despite the paramount importance of efficiently allocating a nation's budget and assessing its impact, this process often lacks transparency and falls short of being evidence-based. In the pursuit of developing a policy tech tool to imbue national budget allocation with greater evidence-backed efficiency, it becomes imperative to explore the algorithms and AI methods applicable to each policy assessment [2].

In my research, I would like to examine what machine learning models/methods can be suitable for effective public budget allocation, specifically for the Japanese government. This paper will first identify the public budgeting process in the context of the Japanese government, including the specific stages of the budget allocation process that these algorithms and AI techniques can influence. It then examines the various machine learning models, including Random Forest, Gradient Boosting, Multi-Layer Perceptron neural networks, and Support Vector, and genetic algorithm optimization methods, that have been considered for this purpose. Several existing studies in this area will be discussed, mainly focusing on a study by Valle-Cruz et al. that investigates data analysis through the application of artificial neural networks and genetic algorithms to analyze the budget of the Mexican federal government [14]. Finally, this paper provides an overview of an analysis of the results obtained and a comparison of machine learning models to determine which one is suitable for this context. My contribution is to develop an algorithm that not only evaluates past public budgets and reflects their impact, but also uses machine learning methods to suggest specific potential budget allocations for the following year. Additionally, my another contribution is that I am applying the artificial intelligence methods to a new domain, the public budget of Japan.

Table 1: Main Categories of Japanese Public Budget

Number	Budget Category
1	Government-run agency fees
2	Local government finance
3	Defence-related expenses
4	Land preservation and development
5	Industry and economy
6	Education and culture
7	Social security
8	National compensation system fees
9	National debt service fees
10	Other

2 BACKGROUND – UNDERSTANDING PUBLIC BUDGETING IN JAPAN

In Japan, the national government budget is classified into the following nine issues; 1) Social security, 2) Public works, 3) Education and sciences, 4) Agriculture and small business, 5) Economic corporation, 6) National defense, 7) Measure for energy, 8) Local finance, 9) National debt [3]. The process of budget preparation for each fiscal year consists of three stages, and it takes approximately eight months, with the process starting in June. In Stage I, each ministry and government agency submits budget requests to the Ministry of Finance (MOF). In Stage II, after cabinet decision on the general principles of budget formulation in December, MOF presents the budget proposal to the cabinet. After revival negotiations, MOF and other ministries prepare official budget documents, and they will be presented to the cabinet in Stage III. After that, the Diet will make the final decision. Algorithmic and artificial intelligence methods can play a role in the planning process for Stages I and II, proposing optimal budget allocations for each ministry based on the evaluation of each policy and historical data [3] [10]. My research intends to focus mainly on Stage I and II, to give the Ministry of Finance an overview and a suggestion of how allocate public budget should effectively allocated for the upcoming fiscal year.

3 METHODS – FOUR MACHINE LEARNING MODELS AND GENETIC ALGORITHM

Four machine learning regression models used in this research are Random Forest, Gradient Boosting, Multi-Layer Perceptron, and Support Vector. These machine learning models are classified under supervised learning, ensemble learning, and neural network, and they are commonly and widely used for various machine learning

Table 2: Machine Learning Models

Model	Accuracy	Interpretability	Computational Complexity	Robustness to Overfitting
Random Forest Regressor	High	Moderate	Moderaet	Moderate
Gradient Boosting Regressor	High	Low	Moderate-High	Low
Multi-Layer Perceptron Regressor	High	Low	High	Low
Support Vector Regressor	Moderate-High	Moderate	Moderate	high

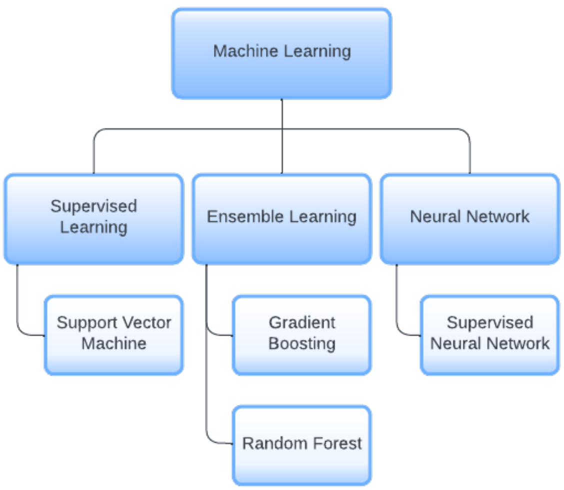

Figure 1: Machine Learning Models

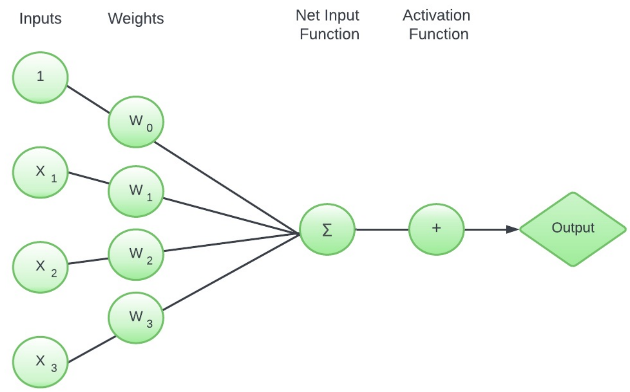

Figure 2: Neural Networks

examples. Table 1 discusses the general characteristics of these models, and this section provides details on each model [9] [11] [6].

3.1 Random Forest Regressor

A random forest is a meta-estimator that fits a set of decision tree regressors to different subsamples of the data set and uses averaging to improve predictive accuracy and control over-fitting.

It is a supervised learning algorithm and bagging technique that uses an ensemble learning method for regression in machine learning. The trees in random forests run in parallel, meaning that there is no interaction between the trees as they are built.

3.2 Gradient Boosting Regressor

Gradient boosting is one of the most popular machine learning algorithms for tabular data, and it has shown considerable success in a wide range of practical applications [7]. It is powerful enough to find any nonlinear relationship between any given model's target and features, and has a great usability that can deal with missing values, outliers, and high cardinality categorical values on features without any special treatment. Furthermore, it is highly customizable to the particular needs of the application, like being learned with respect to different loss functions [7].

3.3 Neural Networks and Multi-Layer Perceptron

Artificial neural networks (ANNs) are powerful tools commonly employed for solving prediction and classification problems. They draw inspiration from the intricate functioning of the human brain's information processing capabilities. Neural Networks are composed of individual processing units known as neurons, and they work collectively to tackle specific tasks (Figure 2). In neural networks, each neuron receives a set of inputs and, through a series of computations, generates an output [14]. These computations are often nonlinear and involve complex transformations of the input data. The strength and nature of the connections between neurons are critical in neural networks. These connections, referred to as synaptic weights or synapses, play a pivotal role in shaping the network's behavior. Specifically, each input to a neuron, denoted as x_i, is associated with a weight parameter w_i. The synaptic weights determine the impact of each input on the neuron's response and are adjusted during the training process to optimize the network's performance.

Artificial Neural Networks are known for their ability to learn from data, adapt to patterns, and generalize to make predictions or classifications on new, unseen data. As neural network architectures have evolved and grown in complexity, they have become indispensable tools in the field of machine learning and artificial intelligence. Researchers and practitioners continually explore innovative ways to improve and utilize ANNs for a wide range of real-world problems.

In the public budget allocation research conducted by Valle-Cruz et al., ANN framework is used to calculate the impact of each

Table 3: Artificial Neural Network Tests Example

Test	Hidden Layers	Activation Function (hidden layer)	Activation Function (output layer)	Error Sum of Squares (Testing)
1	1	hyperbolic tangent	sigmoid	0.147
2	1	hyperbolic tangent	hyperbolic tangent	0.083
3	1	hyperbolic tangent	Identity	5.546
4	2	hyperbolic tangent	Identity	1.635
5	2	hyperbolic tangent	Sigmoid	0.259
6	2	hyperbolic tangent	hyperbolic tangent	0.271

classified budget item (e.g., Energy, Health) on a defined indicator, such as GDP and inflation, as inputs.

Then, an activation function can be designed to determine the maximum or minimum impact according to the architectural model of the ANN as deemed appropriate for the research. Table 3 includes the neural network tests conducted by Valle-Cruz et al. to determine 2 hidden layers with hyperbolic tangent activation functions is the best model for their case with the least error sum of squares.

Multi-Layer Perceptron (MLP) is a common type of artificial neural network that consists of fully connected neurons with a nonlinear type of activation function, organized in at least three layers, characterized by the ability to discriminate data that are not linearly separable. MLP approach is both nonparametric and stochastic and offers high flexibility [1].

3.4 Support Vector Regressor

Support Vector Regression (SVR) is a type of Support Vector Machine (SVM) used for regression tasks. SVR attempts to find a function that best predicts the continuous output value for a given input value.

In machine learning, Support Vector Machines are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis[6].

SVR can use both linear and nonlinear kernels. A linear kernel is a simple dot product between two input vectors, while a nonlinear kernel is a more complex function that can capture more complicated patterns in the data. The choice of kernel depends on the characteristics of the data and the complexity of the task. SVMs reduce most machine learning problems to optimization problems and optimization lies at the heart of SVMs [12].

3.5 Genetic Algorithms and Optimizations

Genetic algorithms (GAs) primarily serve the purpose of tackling optimization problems through adaptive processes inspired by natural systems. They excel at finding solutions within intricate adaptive landscapes by employing stochastic techniques. The representation of GAs can take various forms, including binary strings, networks, ordered lists, or real parameter vectors. In the case of binary strings, chromosomes are constructed from binary units referred to as genes. Each gene has a specific value known as an allele, and its position on the chromosome is termed the locus [5]. The complete genetic makeup is termed the genotype, and its interaction with the environment is known as the phenotype. This interaction ultimately leads to the interpretation of the chromosome, yielding an alternative solution.

Table 4: Genetic Algorithm Tests

Scenario	1	2	3	4
Population Size	100	500	1000	1000
Generations	100	100	100	100000
Mutation Chance	0.01	0.01	0.01	0.01
Elitism	20	20 50	100	

In the case of this public budget allocation research analyzed by Valle-Cruz et al, a genetic algorithm is used to try out several scenarios to optimize the activation formula and determine the best fitness for a given indicator (in this case GDP and inflation) after analysis with ANN [14]. In another example of existing research aimed at the efficient allocation of government R&D budgets, a robust optimization technique was used to hedge against uncertainty in the predicted value of R&D output [4].

4 RESEARCH DESIGN

While much of the existing research focuses on studies that retrospectively evaluate past budget allocations and suggest which areas should receive more money, I would like to go further and create an algorithm that suggests specific and optimal national budget allocations for each category. Following the analysis steps shown in the data architecture diagram (Figure 3), I will apply four different machine learning model – Random Forest Regressor, Gradient Boosting Regressor, and Multi-Layer Perceptron, and Support Vector Regressor – and genetic algorithms to develop a machine learning algorithm to propose public budget allocations for the upcoming year based on an analysis of past public policy evaluation patterns.

4.1 Research Steps

As for the budget data to be used in the analysis, this research will use the General Accounts Expenditure Budgetary Objectives Classification Summary Tables for 24 years since 2000, which are publicly available on the Japanese Ministry of Finance's website. These data include the ten major categories of Japan's public budget, and they are appropriate for this research because each category contains the exact budget allocation for the past 24 years and the tables are formatted in tabular form. Additionally, corresponding 24 years of data for features and economic indicators that are explained in section 4.2 below will be used as inputs to train machine learning models.

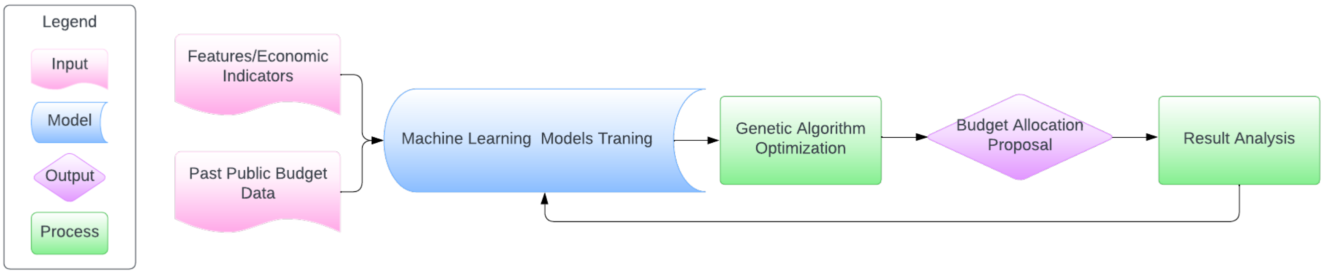

Figure 3: Data Architecture Diagram

In this study, 100 epochs and a batch size of 32 were used to train each model because the data set size is relatively small, with the number of columns being less than 20.

After running four different machine learning models (Table 2), four genetic algorithm tests (Table 4) are applied to see which scenario has the best fitness that maximizes the impact of each economic indicator and features. Based on the best scenario, we will analyze how much percentage of the total budget should be spent on each category of the budget. Then, the process of comparing the results from each machine learning model, analyzing the differences and their causes, and adding more economic indicators/features and changing the weights is repeated until the best model is found. This way, we aim to provide efficient specific budget allocations as final outputs [8].

Through this research, I hope to create a national budget allocation algorithm for the Japanese government and develop policy tools in this area using artificial intelligence.

4.2 Factors and Features for Machine Learning Model Design

As shown in the data architecture (Figure 3), machine learning models for this study take as initial input the amount of budget by objective and considers the following factors to propose a better/more efficient budget as output for the next year:

- GDP growth rate
- Inflation rate
- General flexibility of the national public budget
- Approval rating for the government
- National debt balance
- Tax income
- Population
- Aging rate
- Poverty rate
- Unemployment rate
- The percentage of the population who is on welfare
- Homeless rate
- Average wage
- Policy areas that the current Prime Minister is focusing on

The above list includes some of the most important quantitative indicators for measuring a country's status. In addition, by setting different weights for each economic characteristic/indicator, the algorithm can simulate budget allocations with certain areas of the budget weighted more heavily. the GDP growth rate and the inflation rate are weighted more heavily than other economic factors because they are common and explicit measures for evaluating national fiscal performance.

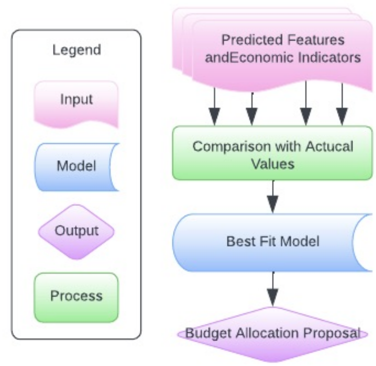

Figure 4: Result Analysis Detail

4.3 Metrics and Evaluation of Models

As a way to evaluate the performance of each machine learning model, this research puts the output results to the same machine learning model and determines which model maximizes the generated (predicted) values for economic features/indicators.

In this way, four different machine learning algorithms can be compared and suitable machine learning models can be found among these four. Since this study applies the models for the following fiscal year, another comparison can be made after the year is over and new economic features/indicators are available by checking how the values maximized by the models compare to the actual data to evaluate the model performance (Figure 4).

5 RESULTS AND ANALYSIS

The results obtained are shown in Table 5, as the ratio (to 1) of the total budget that should be spent on each category of the budget for the following fiscal year, 2025.

With the same input data, the four machine learning models have different output figures and different categories in which they

Table 5: Machine Learning Models Result

Number	Budget Category	Random Forest	Boosting	MLP	SVR
1	Government-run agency fees	0.387065	0.027313	0.032091	0.022329
2	Local government finance	0.000323	0.008152	0.003834	0.012777
3	Defense-related expenses	0.016816	0.223133	0.475689	0.012761
4	Land preservation and development	0.033812	0.019343	0.217680	0.020645
5	Industry and economy	0.000729	0.017500	0.088364	0.112672
6	Education and culture	0.108449	0.255527	0.147684	0.132018
7	Social security	0.171992	0.301440	0.010477	0.484742
8	National compensation system fees	0.023825	0.032486	0.022505	0.023714
9	National debt service fees	0.247716	0.114961	0.000162	0.016890
10	Other	0.009268	0.000145	0.001509	0.161447

suggest spending the most budget – Random Forest: Government-run agency fees, Gradient Boosting: Social security, Multi-Layer Perceptron: Defense-related expenses, and Support Vector: Social Security. As described in section 4.3 above, the final comparison will be made after the economic feature/indicator data for the fiscal year 2025 is available. However, by taking 23 years of data from 2000 to 2023 and doing the same comparison and training processes, this research found out that Multi-Layer Perceptron generated the budget allocation proposal that produced the most maximized results compared to those data with actual allocation. As the Multi-Layer Perceptron has the flexibility to incorporate more features/factors, this seems to be the best fit for this research purpose and for the future work listed in the last section.

6 FUTURE WORKS

Future efforts will focus on refining and optimizing algorithms to improve the accuracy and efficiency of proposed budget allocations by further integrating policy evaluations and running the model over time. Other potential future work includes incorporating budget data and economic indicators/features from other countries to further train the model and increase its applicability. As this research continues, it is very important to emphasize and be transparent about what data are used as inputs for machine learning models and how each of the data are weighted, in order to ethically and responsibly use artificial intelligence for more democratic policymaking.

ACKNOWLEDGMENTS

I would like to thank Dr. Charlie Peck, Dr. David Barbella, and Dr. Javier Orduz for their support and feedback on the project.

REFERENCES

[1] Álvaro Costa and Raphael N Markellos. 1997. Evaluating public transport efficiency with neural network models. *Transportation Research Part C: Emerging Technologies* 5, 5 (1997), 301–312.

[2] Bo-Young Heo, Min Jung Kim, and Won-Ho Heo. 2020. An algorithm for validation of the efficiency of disaster and safety management budget investment in South Korea. *International journal of disaster risk reduction* 47 (2020), 101566.

[3] Hiromitsu Ishi. 1996. Budgets and the budgetary process in Japan. *Hitotsubashi Journal of Economics* (1996), 1–19.

[4] Hoon Jang. 2019. A decision support framework for robust R&D budget allocation using machine learning and optimization. *Decision Support Systems* 121 (2019), 1–12.

[5] Wuttipong Kusonkhum, Korb Srinavin, Narong Leungbootnak, Preenithi Aksorn, and Tanayut Chaitongrat. 2022. Government construction project budget prediction using machine learning. *Journal of Advances in Information Technology Vol* 13, 1 (2022).

[6] Batta Mahesh. 2020. Machine learning algorithms-a review. *International Journal of Science and Research (IJSR).[Internet]* 9, 1 (2020), 381–386.

[7] Alexey Natekin and Alois Knoll. 2013. Gradient boosting machines, a tutorial. *Frontiers in neurorobotics* 7 (2013), 21.

[8] A Nedić and Dimitri P Bertsekas. 2003. Least squares policy evaluation algorithms with linear function approximation. *Discrete Event Dynamic Systems* 13, 1-2 (2003), 79–110.

[9] Victor Rodriguez-Galiano, Manuel Sanchez-Castillo, M Chica-Olmo, and MJOGR Chica-Rivas. 2015. Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. *Ore Geology Reviews* 71 (2015), 804–818.

[10] Mickie Schoch and Corina den Broeder. 2013. Linking information on policy effectiveness and efficiency to budget decisions in the Netherlands. *OECD Journal on Budgeting* 12, 3 (2013), 1–22.

[11] Emily Sullivan. 2022. Understanding from machine learning models. *The British Journal for the Philosophy of Science* (2022).

[12] Yingjie Tian, Yong Shi, and Xiaohui Liu. 2012. Recent advances on support vector machines research. *Technological and economic development of Economy* 18, 1 (2012), 5–33.

[13] David Valle-Cruz, Vanessa Fernandez-Cortez, and J. Ramon Gil-Garcia. 2022. From E-budgeting to smart budgeting: Exploring the potential of artificial intelligence in government decision-making for resource allocation. *Government Information Quarterly* 39, 2 (2022), 101644.

[14] David Valle-Cruz, J. Ramon Gil-Garcia, and Vanessa Fernandez-Cortez. 2020. Towards smarter public budgeting? Understanding the potential of artificial intelligence techniques to support decision making in government. In *The 21st Annual International Conference on Digital Government Research*. 232–242.