A Machine Learning Decision-Support System for Optimizing Healthcare Budgetary Financing in Uzbekistan

Durdona Khaydarova[1], Erkinjon Buriboev[1], Kamola Sherbekova[1], Muyassar Ismoilova[1] and Ibrat Usmonov[2]
[1]*Department of Accounting and Finance, Gulistan State University, Gagarin Str. 4, 120100 Gulistan, Uzbekistan*
[2]*Department of Economics, Gulistan State University, Gagarin Str. 4, 120100 Gulistan, Uzbekistan*
durdonaxaydarova19991995@gmail.com, boribayeverkinjon@gmail.com, sherbekovakamola494@gmail.com, ismoilovamuyassar91@gmail.com, usmonovibrat315@gmail.com

Keywords: Healthcare Financing, State Budget, Machine Learning, Predictive Modeling, Health Insurance, Decision Support Systems, Data Analytics, Capitation.

Abstract: This article analyzes state budget financing in the healthcare sector of the Republic of Uzbekistan and introduces an applied Information Technology (IT) component to support future healthcare financing decisions. As Uzbekistan continues its transition toward state health insurance and modern reimbursement models such as capitation, data-driven decision-support systems may complement traditional budget analysis. To support this transition, we developed a Machine Learning (ML) binary classification model to predict citizens' Willingness-to-Pay (WTP) for the upcoming mandatory medical insurance system. Using a structured digital survey dataset, we compared a Logistic Regression baseline with a Random Forest Classifier. The Random Forest model outperformed the baseline, achieving 75% accuracy and 0.92 recall for the positive class. Feature importance analysis showed that historical out-of-pocket medical expenses and perceptions of insurance benefits were stronger predictors of WTP than standard demographic variables. These pilot findings suggest that predictive ML models may support ongoing and future healthcare financing reforms in Uzbekistan by helping policymakers design more targeted awareness campaigns and subsidy strategies. However, the model should be interpreted as a pilot decision-support tool and not as a direct evaluation of earlier budget decisions.

1 INTRODUCTION

To achieve sustainable economic growth, it is essential not merely to reduce budgetary expenditures, but to strategically optimize their structure in a manner that effectively promotes societal development. In recent years, the Republic of Uzbekistan has made extensive use of budgetary resources to finance the healthcare sector, alongside the expansion of paid medical services. To mitigate the adverse effects of the COVID-19 pandemic and strengthen social infrastructure, significant priority was placed on healthcare financing, including the allocation of 526 billion UZS to the Fund for the Material Incentivization and Development of Medical Institutions.

As part of the ongoing healthcare financing reforms in Uzbekistan, the establishment of a fair, competitive environment is expected to create favorable conditions for both public and private medical institutions [1]. A cornerstone of this reform is the current transition to a state health insurance system, which moves the sector away from traditional input-based funding toward modern, output-based reimbursement mechanisms, such as capitation and case-based financing.

However, the success of such comprehensive economic reforms relies heavily on public adaptation and efficient fund allocation. Traditional economic analyses, which rely purely on descriptive statistics of past state budgets, are insufficient for predicting future public behavior. Therefore, this study aims to bridge the gap between public finance and Information Technology (IT) by developing a Machine Learning (ML) predictive model. By analyzing citizens' Willingness-to-Pay (WTP) for the new medical insurance system, this paper proposes a data-driven Decision Support System (DDSS). The

specific scientific contributions of this study are threefold:

- Applied ML Framework: Introducing a supervised machine learning pipeline to forecast citizen participation in state medical insurance.
- Feature Identification: Empirically demonstrating that historical out-of-pocket medical expenses, rather than static demographic factors, are the primary drivers of WTP in Uzbekistan.
- Digital Policy Integration: Proposing a conceptual e-health dashboard workflow that may help translate algorithmic predictions into targeted healthcare subsidies and awareness campaigns.

2 LITERATURE REVIEW

The transition of healthcare financing mechanisms has been widely discussed within the context of public finance and socioeconomic reform. Previous literature emphasizes that the social sector is a foundational element of societal development, making healthcare expenditures a critical component of state financial planning [2]. Because medical services play a vital preventive role in maintaining the health of the population-the main driving force of the economy-scholars assert that state participation in the management and financing of healthcare remains an objective necessity, regardless of the level of market privatization [3].

Recent structural reforms in Uzbekistan's healthcare sector have focused on the implementation of a state medical insurance system. Recent policy documents in Uzbekistan outline a transition toward capitation-based financing, preferential patient targeting, and the development of a unified electronic healthcare system [1], [4]. The integration of local medical institutions into a centralized electronic network is expected to ensure transparent tracking of medical supplies and transparent fee determination. Furthermore, pilot projects, such as the new healthcare organization model tested in the Syrdarya region, demonstrate that shifting financial responsibility to medical institutions through insurance mechanisms directly increases the quality of medical services [5].

Despite the extensive economic and legal analysis of these reforms [6], [7], there remains a critical gap in the literature regarding the application of advanced predictive analytics to healthcare finance in Uzbekistan. While current methodologies effectively summarize retrospective budget allocations [8], they lack the algorithmic capacity to predict citizen engagement with new financial models. This paper addresses this gap by introducing an applied IT framework, utilizing ensemble Machine Learning techniques to forecast public willingness to participate in the co-financing of state medical insurance.

Furthermore, the integration of Artificial Intelligence (AI) and ML into public health finance is gaining international momentum. Recent global studies highlight the efficacy of causal machine learning frameworks in evaluating the heterogeneous impacts of subsidized health insurance [9], and the use of ML to predict health insurance enrollment in developing economies [10]. These digital decision-support systems provide policymakers with the necessary computational tools to shift from broad, untargeted healthcare spending to precision-based financial protection mechanisms.

3 RESEARCH METHODOLOGY

While traditional economic analyses rely on descriptive statistics to evaluate budget allocations, modern healthcare financing requires data-driven Decision Support Systems (DDSS). To address this, this study introduces an applied Information Technology (IT) component: a Machine Learning (ML) binary classification model designed to predict citizens' Willingness-to-Pay (WTP) for the upcoming state medical insurance system.

3.1 Data Collection and Preprocessing

A structured dataset was compiled from a digital survey of 120 respondents in Uzbekistan, collected through convenience sampling in early 2026. The dataset should therefore be interpreted as a pilot sample rather than a nationally representative cohort. The target distribution was moderately balanced, with approximately 60% of respondents indicating willingness to pay for state health insurance and 40% indicating unwillingness.

The input feature set consisted of four main groups:

1) demographic variables, including age-related and employment-related categories;
2) self-reported medical expenditure indicators, particularly historical out-of-pocket healthcare spending;

3) insurance perception variables, including perceived improvements in service quality and expected reduction in financial burden;
4) Likert-scale survey responses measuring attitudes toward the proposed insurance system.

The target variable, Willingness-to-Pay (WTP), was binarized as follows: WTP = 1 for respondents who answered “Yes” and WTP = 0 for respondents who answered “No”. Missing or noisy textual entries in numerical fields (e.g., self-reported medical expenses recorded as “1mln” or “none”) were cleaned using regular expressions and imputed with median values. Categorical variables were transformed using One-Hot Encoding, while Likert-scale responses were converted into numerical integer values. To reduce scale-related distortion in the baseline statistical model, numerical variables were standardized using StandardScaler.

3.2 Algorithmic Framework

To evaluate the predictive capacity of the data, two distinct mathematical models were deployed:

- Logistic Regression (Econometric Baseline): Utilized as a standard statistical benchmark to represent traditional policy analysis. To ensure reproducibility, the model was configured with solver='lbfgs' and random_state=42.
- Random Forest Classifier (Ensemble ML Model). An advanced decision-tree-based algorithm deployed to capture complex, non-linear relationships between a citizen's socioeconomic profile and their WTP. The model was configured with n_estimators=100, max_depth=None, and random_state=42; these hyperparameters were selected to balance model complexity and generalization performance given the limited sample size, and should be further tuned via systematic grid search on larger datasets.

The input feature set included demographic variables, self-reported medical expenditure indicators, and Likert-scale perceptions of insurance benefits. The target variable, Willingness-to-Pay (WTP), was strictly binarized as 1 for respondents willing to pay ("Yes") and 0 otherwise ("No").

To ensure a robust machine learning pipeline, all data preprocessing (such as imputation, encoding, and scaling) was applied strictly before model fitting to prevent data leakage. The dataset was partitioned into training and test subsets using an 80/20 split. Five-fold cross-validation was performed only on the training subset for hyperparameter validation, while final model performance was evaluated on the unseen hold-out test subset.

3.3 Conceptual IT Architecture and Integration

To bridge the gap between algorithmic prediction and public finance policy, the proposed Random Forest model is designed to function as the backend intelligence for a Ministry of Health e-health dashboard. As illustrated in Figure 1, the conceptual workflow automatically ingests citizen survey and demographic data, processes the features through the ML classifier, and outputs a WTP probability score.

Policymakers can then utilize this dashboard to trigger targeted interventions-such as personalized digital awareness campaigns or dynamic premium subsidies-for specific population clusters identified as "low probability" for voluntary enrollment.

4 ANALYSIS AND RESULTS

4.1 The Economic Context: Healthcare Budget Allocations

Before applying the predictive algorithm, it is critical to contextualize the financial landscape of Uzbekistan’s healthcare system. Initial state healthcare expenditures for 2024 were projected to reach 33,493.3 billion UZS, representing an increase of 17.6% compared to 2023 levels [4]. The 2025 Budget for Citizens also provides an important fiscal basis for assessing the continuation of healthcare financing priorities in Uzbekistan [11]. The predominant share of these expenditures-approximately 78.4% (26,164.0 billion UZS)-is allocated to mandatory fixed costs, primarily wages and equivalent payments. Spending on medicines and medical supplies constitutes the second-largest category at 3,759.0 billion UZS [4], [8].

Because the high proportion of fixed expenditures significantly constrains opportunities for cost rationalization, the burden of optimization falls primarily on the successful implementation of the new medical insurance system and the reimbursement mechanism. For this transition to succeed without placing undue strain on the state budget, the government must accurately forecast and stimulate citizen participation.

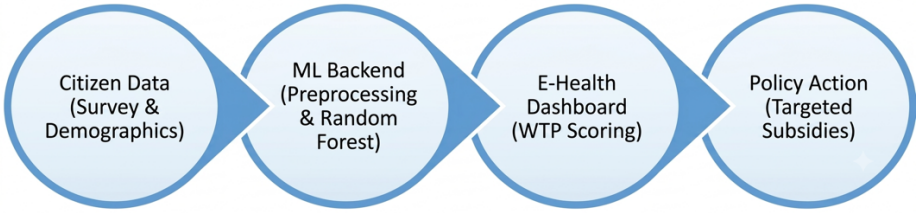

Figure 1: Conceptual workflow of the decision support system.

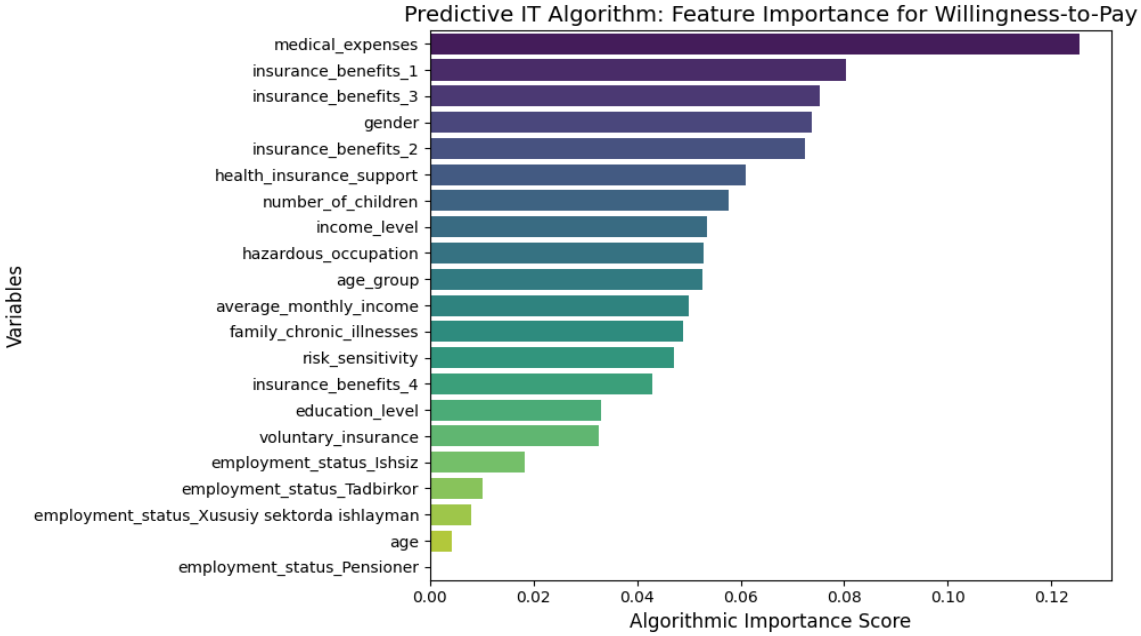

Figure 2: Feature importance for predicting willingness-to-Pay for state health insurance.

4.2 Predictive Modeling Results and IT Evaluation

The computational results demonstrate a significant performance advantage when applying Machine Learning to healthcare financing predictions. As shown in Table 1, the traditional Logistic Regression baseline achieved an accuracy of only 58%. In contrast, the Random Forest ensemble model achieved a predictive accuracy of 75%.

Class 1 represents citizens expressing Willingness-to-Pay. Most notably, the Random Forest model achieved a robust recall score of 0.92 for the positive class. In the context of national health insurance rollouts, Recall was prioritized as the primary performance metric. A False Negative (failing to identify a citizen who is actually willing to pay) represents lost potential state revenue and missed policy momentum. A high recall of 0.92 ensures that the vast majority of the receptive population is correctly identified, allowing the state to secure their enrollment efficiently.

Table 1: Algorithmic performance comparison.

Evaluation Metric	Logistic Regression (Baseline)	Random Forest Classifier (ML)
Overall Accuracy	0.58 (58%)	0.75 (75%)
Precision (Class 1)	0.57	0.69
Recall (Class 1)	0.67	0.92
F1-Score (Class 1)	0.62	0.79

4.3 Algorithmic Feature Importance

To provide actionable insights for the Ministry of Health, algorithmic Feature Importance was extracted from the Random Forest model (see Fig. 2).

The algorithmic evaluation revealed that standard demographic factors-such as age group and employment status (e.g., pensioner, unemployed), which reflect broader national distributions [12]-carried the lowest predictive weight. Conversely, the

most critical algorithmic predictor of WTP was the user's historical out-of-pocket medical_expenses. This was immediately followed by variables representing the citizens' algorithmic perception of insurance benefits, specifically whether they believe the insurance will improve medical quality (insurance_benefits_1) and reduce overall financial burden (insurance_benefits_3).

5 DISCUSSION

The results of this pilot algorithmic evaluation may offer useful insights for ongoing healthcare financing reforms in Uzbekistan. Existing policy initiatives, including the Syrdarya region pilot project and the broader transition toward state medical insurance, highlight the importance of combining financial reform with public acceptance and institutional trust [5]. In this context, the proposed ML framework should be interpreted as a decision-support instrument for subsequent implementation stages rather than as a direct evaluation of earlier budget decisions.

The predictive results suggest that ongoing and future healthcare financing measures may benefit from more targeted digital awareness efforts rather than fully generalized public communication. In particular, the relatively low importance of standard demographic variables and the stronger contribution of out-of-pocket medical expenses indicate that behavioral and financial experience variables may be more informative for identifying groups with different levels of willingness to participate.

Similarly, because perceived improvements in medical quality and expected reductions in financial burden emerged as important predictors, policy communication may be more effective when these specific benefits are emphasized. If integrated carefully into an e-health dashboard environment, this type of ML-based system can help policymakers identify population segments that may require additional awareness measures or financial support. However, these implications should be interpreted cautiously given the pilot sample size and convenience-based data collection strategy [8].

6 CONCLUSIONS

This study examined the potential of Machine Learning as a decision-support tool for healthcare financing reform in Uzbekistan. Using a pilot survey dataset, Logistic Regression and Random Forest models were evaluated for predicting citizens' Willingness-to-Pay (WTP) for the proposed state medical insurance system. The results showed that Random Forest achieved better predictive performance, while historical out-of-pocket medical expenses and perceived insurance benefits were identified as the most important predictors.

The findings indicate that ML-based approaches can support healthcare policy planning by providing evidence-based insights into citizen preferences. Such tools may assist policymakers in developing targeted awareness campaigns, improving communication strategies, and designing more effective subsidy mechanisms during future stages of health insurance implementation.

7 LIMITATIONS AND FUTURE WORK

The study has several limitations. The dataset included only 120 respondents collected through convenience sampling in early 2026, which may not fully represent rural, low-connectivity, or other underrepresented population groups. Therefore, the proposed model should be considered a pilot analytical framework rather than a basis for nationwide policy conclusions.

Future research should expand the dataset with larger and more representative samples, include additional socioeconomic and regional factors, and validate the proposed ML framework using real healthcare administration data.

REFERENCES

[1] Decree of the President of the Republic of Uzbekistan No. PF-158, "On the Strategy 'Uzbekistan - 2030'," Sep. 11, 2023, [Online]. Available: https://lex.uz/.

[2] E. V. Fakhrutdinova, "The role of the social sphere and social policy and the organization of sustainable social and economic development," Economic Science, no. 8 (57), p. 7, 2009.

[3] B. I. Nurmuhamedova, "The need for state participation in the management and financing of the health care system," Finance, no. 2, p. 92, 2012.

[4] Ministry of Economy and Finance of the Republic of Uzbekistan, "Budget for Citizens 2024 Draft," api.mf.uz, 2024, [Online]. Available: https://api.mf.uz/media/document_files/Budjet_P_24_uz.pdf.

[5] Decree of the President of the Republic of Uzbekistan No. 4890, "On measures to introduce a new model of organizing the healthcare system and state medical insurance mechanisms in the Syrdarya region," Nov. 12, 2020.

[6] G. A. Kasimova, Improving the Methodology of Financial Support of the Social Sphere in the Treasury System, Monograph. Tashkent, Uzbekistan: Economics and Finance, 2021.

[7] G. A. Kasimova, State Social Policy: Financing Education and Health Care, Monograph. Tashkent, Uzbekistan: TDIU, 2024.

[8] Ministry of Economy and Finance of the Republic of Uzbekistan, "Fiscal Strategy 2023-2025," admin.openbudget.uz, 2023, [Online]. Available: https://admin.openbudget.uz/media/post_attachments/2023_2025.

[9] V. Shah and N. Kreif, "Estimating heterogeneous impacts of subsidised health insurance: A causal machine learning approach," PLoS One, vol. 20, no. 9, Art. no. e0315057, 2025.

[10] United Nations, "Using machine learning to predict health insurance enrolment and take-up of health services," UN Dept. of Economic and Social Affairs, 2021.

[11] Ministry of Economy and Finance of the Republic of Uzbekistan, "Budget for Citizens 2025 Project," api.mf.uz, 2025, [Online]. Available: https://api.mf.uz/media/document_files/Budjet_P_25_uz.pdf.

[12] Statistics Agency under the President of the Republic of Uzbekistan, "Official Website," stat.uz, 2026, [Online]. Available: https://www.stat.uz.