Khazanah Journal of Religion and Technology
Universitas Islam Negeri Sunan Gunung Djati Bandung, Indonesia
https://journal.uinsgd.ac.id/kjrt
Online ISSN: 2987-6060
Volume 3 Issue 2 (2025)
DOI: https://doi.org/10.15575/kjrt.v3i2.1539

XGBoost and Mixed Effect Model: Can *Zakat* (Alms) Improve the Human Development Index in Indonesia?

Nazwa Hemalia Putri
Komunitas Pendidikan Nasional (KOMPENAS)
Garut, Indonesia
naf50951@gmail.com

Rizkco Fauzan Adhim
Student Unity Network (SUN)
Bandung, Indonesia
rizkcofauzana@gmail.com

Saftana Fitri
Community of Santri Scholars of Ministry of Religious Affairs (CSSMoRa)
Ministry of Religious Affairs
Jakarta, Indonesia
saftanafitri25@gmail.com

Sifa Mutiasya Hendayana Puteri
Google Developer Group (GDG)
Jakarta, Indonesia
sifamutiasya14@gmail.com

Suci Ihtisabi Hida Nursyifa
Polyglot Indonesia
Bandung, Indonesia
sucinursyifa064@gmail.com

Abstract— This study aims to analyze the effect of *zakat* distribution on the Human Development Index (HDI) in Indonesia by comparing two modeling approaches: the Mixed Effect Model (LMM) and Extreme Gradient Boosting (XGBoost). The data used are panel data from 34 provinces during the 2021–2024 period, which includes sectoral *zakat* distribution and HDI variables. Initial exploration results indicate spatial variation and a non-linear correlation between *zakat* distribution and HDI. The Mixed Effect model demonstrated more stable predictive performance with an RMSE of 1.633 and an R^2 of 0.762 on the test data. While XGBoost, despite its high accuracy on the training data, exhibited overfitting with a decrease in accuracy on the test data. These findings indicate that the LMM approach is superior in capturing spatial and temporal variations between regions and provides more applicable interpretations in the context of public policy. This study contributes to the use of statistical approaches and machine learning to evaluate the effectiveness of *zakat* in supporting sustainable human development in Indonesia.

Keywords- Human Development Index (HDI), Mixed Effect Model, panel data, sustainable development, XGBoost, Zakat.

I. INTRODUCTION

Sustainable social and economic development is a priority for developing countries like Indonesia. However, poverty and social inequality remain major challenges that need to be addressed. For Muslims, *zakat* plays a crucial role as an instrument for wealth distribution, a form of worship. *Zakat* is a certain portion of wealth that every Muslim must pay upon meeting certain requirements. As one of the pillars of Islam, *zakat* is paid to be distributed to those entitled to receive it. Besides being an act of worship, *zakat* also serves to reduce social inequality and support community welfare. In Indonesia, *zakat* distribution has proven to have a significant impact. According to the National *Zakat* Agency (BAZNAS) (2024) through its *Zakat* and Poverty Alleviation Report, in 2023, *zakat* empowered 54,081 beneficiaries, with 58.76% of them escaping extreme poverty. *Zakat* distribution

also contributed to reducing the national poverty rate from 26.36 million (September 2022) to 25.90 million (March 2023), a decrease of 460,000 people [1].

Research by Nurul Agniya et al. (2024) shows that *zakat* distribution significantly reduced poverty rates in 34 provinces in Indonesia during the 2019–2021 period. In the panel regression model used, a 1 billion rupiah increase in *zakat* distribution was associated with a 0.043 percentage point decrease in poverty rates at the provincial level, after controlling for macroeconomic variables such as inflation, unemployment, and economic growth [2]. Meanwhile, Majid and Gissay (2024) in their study in East Nusa Tenggara Province from 2013 to 2022 found that *zakat* distribution reduced poverty in urban areas with an elasticity effect of -0.12. This means that every 10% increase in *zakat* distribution contributes to a 1.2% reduction in poverty. However, in remote areas, *zakat* effectiveness is hampered by limited governance and unequal access to *zakat* distribution [3].

The concept of *zakat* as a tool for human development has now been expanded into an Islamic social finance framework that includes a sustainability dimension. Widiastuti et al. (2022) developed a sustainability index for *zakat* and waqf programs based on maqāṣid al-sharīʿah, which emphasizes the strategic role of *zakat* in supporting the multidimensional well-being of society [4].

Thus, the role of *zakat* also contributes directly to the achievement of the Sustainable Development Goals (SDGs) initiated by the United Nations (UN) [5]. The SDGs encompass 17 global goals, including poverty eradication (SDG 1), zero hunger (SDG 2), good health (SDG 3), quality education (SDG 4), and reduced inequality (SDG 10). Therefore, effective *zakat* distribution not only impacts the beneficiaries (mustahik) but also strengthens national commitment to the global sustainable development agenda.

Zakat, as part of Islamic social finance, is now being studied more broadly in the context of the Sustainable Development Goals (SDGs). Tok et al. (2022) emphasize the importance of converging *zakat* policies with the national development agenda through a coherent and measurable approach [6].

The Human Development Index (HDI) is a composite indicator widely used by governments and international institutions to evaluate the achievement of several key goals in the SDGs, particularly in the areas of health, education, and a decent standard of living. Therefore, the influence of *zakat* on the HDI can be used as an important benchmark in assessing *zakat*'s contribution to the success of sustainable development. Research by Mukadar et al. (2023) conducted a case study in Bengkulu Province using a t-test on individual HDI data before and after *zakat* distribution. The results showed an increase in the average HDI score from 0.612 to 0.658 ($p<0.05$) among mustahik after receiving *zakat* [7]. It can be concluded that *zakat* has significant potential to support human development in Indonesia. However, there is still a gap in comprehensive analysis that uses a quantitative and adaptive approach to complex data patterns.

Socio-economic data such as *zakat* and the Human Development Index (HDI) have complex relationships and are full of non-linear interactions, so traditional methods such as linear regression are often inadequate. Advanced machine learning algorithms such as Extreme Gradient Boosting (XGBoost) are a machine learning algorithm based on gradient boosting decision trees designed for efficiency and high performance. XGBoost combines many decision trees in a step-by-step manner, where each new tree attempts to correct the errors of the previous tree. The XGBoost algorithm has proven effective in capturing non-linearities and variable interactions in various social applications, such as life expectancy in Indonesia [8] and prediction of HDI classification at the district/city level in Indonesia [7]. A global study on life expectancy prediction using XGBoost reported an MAE of 1.554 and an RMSE of 2.402, outperforming other models such as Random Forest (MAE 7.938; RMSE 11.304) and ANN (MAE 3.860; RMSE 5.002), thus illustrating the superiority of XGBoost in handling macro-level socio-economic variables [9].

Meanwhile, the Mixed Effects Model or Linear Mixed Model (LMM) is a statistical model that combines fixed effects and random effects within a single framework. Fixed effects are universally applicable independent variables (e.g., the amount of *zakat* distribution), while random effects capture variation between units (e.g., between provinces or years). This model is particularly useful in panel or hierarchical data analysis where correlations exist between groups. The Mixed Effects Model is able to handle dependencies in longitudinal data, reducing bias and improving estimation accuracy [10]. A panel regression study using the Generalized Method of Moments (GMM) approach on data from 34 provinces during 2010–2020 shows that per capita income and average years of schooling are the main determinants of HDI improvement [11]. These findings confirm that economic and social factors play a crucial role in human development, and reinforce the urgency of using hierarchical statistical methods such as Linear Mixed Models (LMM) to capture regional variations in HDI data.

In analyzing panel data across provinces, it is important to consider structural differences such as *zakat* distribution policies, management capacity, and varying socio-economic conditions. To address heterogeneity and correlation across regions and time, approaches such as Mixed Effect Models (LMMs), which combine fixed and random effects, are relevant. However, because the relationships between variables such as *zakat*, education, and health are often complex and non-linear, alternative approaches such as Extreme Gradient Boosting (XGBoost) are also needed. XGBoost is able to capture non-linear interactions and produce accurate predictions without assuming data distribution, making it suitable for use in dynamic and interconnected socio-economic contexts [12].

Therefore, this study aims to compare the effectiveness of the Mixed Effect Model and Extreme Gradient Boosting (XGBoost) approaches in analyzing and predicting the influence of *zakat* distribution on the Human Development Index (HDI) at the provincial level in Indonesia. The Mixed Effect Model will be examined in terms of interpretability and its ability to control for variations across regions and time, while XGBoost will be evaluated based on its predictive

accuracy and its ability to capture non-linear patterns and complex interactions in the data.

II. Related Works

Various previous studies have examined the relationship between *zakat* and social welfare indicators such as poverty and the Human Development Index (HDI). Research by Rusanti et al. (2023) shows that *zakat* distribution in the humanitarian and education sectors has a significant positive effect on the national HDI. Based on an OLS regression model from 2012–2022, each increase in *zakat* in the humanitarian sector is associated with a 1.32% increase in the HDI (coefficient = 0.013197, $p < 0.001$), while *zakat* in the education sector contributes 1.15% (coefficient = 0.011470, $p < 0.001$). This confirms that targeted *zakat* distribution can drive better human development outcomes [13].

Support for the role of *zakat* in development is further strengthened by research by Mutmainah et al. (2024), who analyzed data from 2013–2022 and found that every 1% increase in *zakat* correlated with a 1.03% decrease in poverty. These results indicate that *zakat* not only impacts education and health indicators, which comprise the Human Development Index (HDI), but also directly reduces poverty, a central element of the sustainable development agenda [14].

In a broader national study, Wardani et al. (2021) applied path analysis to 2015–2019 data and concluded that *zakat*, Islamic bank financing, and education spending have both direct and indirect effects on the HDI through economic growth. Statistical test results show that the effect of *zakat* on HDI through economic growth has a t-value of 4.560, which is greater than the t-table (2.014), with a significance level of $\alpha < 0.05$. This finding confirms that *zakat* has a multiplier effect: in addition to directly driving poverty reduction, *zakat* also accelerates human development by strengthening economic growth [15].

Research by Hasan and Ulfatun (2023) used a spatial panel approach to examine the impact of *zakat* distribution on the Human Development Index (HDI) in Indonesia. The results showed that optimal *zakat* distribution significantly improves the quality of life of communities at the provincial level [16]. Majid and Gissay (2024) added empirical evidence from eastern Indonesia that *zakat* is able to reduce poverty significantly in urban areas, but is less effective in remote areas due to limited distribution. [17].

On the other hand, a study by Kurniawan and Indahyanti (2024) applied the Extreme Gradient Boosting (XGBoost) algorithm to predict life expectancy in Indonesia. This algorithm was proven capable of capturing the complex non-linear relationships between socioeconomic variables [18]. Indah and Yanti (2025) also used XGBoost in the classification of HDI at the district/city level and found that this model had advantages in terms of accuracy compared to traditional statistical methods [10].

In addition to machine learning approaches, Linear Mixed Effect Models (LMMs) are increasingly used in socio-economic panel data analysis in Indonesia. Several studies have shown that LMMs are capable of capturing random effects across regions and years, while simultaneously controlling for spatial and temporal heterogeneity [19] [20]. This approach has proven effective in evaluating the impact of *zakat* distribution on human development indicators, such as the Human Development Index (HDI), across various Indonesian provinces. These studies confirm that LMM not only improves the accuracy of predictive models but also provides a deeper understanding of socio-economic variations across regions.

In another comparative study, Rahman and Putra (2022) used various machine learning algorithms to model the impact of *zakat* on poverty reduction and found that non-linear models had better predictive performance than linear models [21]. Zuhdi et al. (2023) emphasize the importance of *zakat* in supporting the achievement of the SDGs by showing the direct contribution of *zakat* to SDG 1 (No Poverty) and SDG 10 [22].

Outside the Indonesian context, a study by Abdurrahman et al. (2023) in Malaysia showed that the effectiveness of *zakat* is highly dependent on a decentralized distribution strategy and transparency of digital-based *zakat* management [23]. This shows the importance of the governance aspect in strengthening the influence of *zakat* on human development.

A study by Satyakti (2023) presents evidence that the maqāṣid al-sharīʿah framework can be used to assess Islamic social financing using a panel data approach. This opens up similar methodological opportunities for evaluating *zakat* distribution against the human development index [24]. Finally, research by Noor et al. (2022) models the effect of *zakat* on the economic welfare of Muslim communities in Bangladesh using dynamic panel regression, and recommends the integration of *zakat* into national fiscal policy [25].

However, most existing research is still limited to the use of classical statistical methods such as linear regression or path analysis, and has not accommodated the complexity of non-linear relationships and the hierarchical panel data structure across regions and time. Therefore, this study attempts to fill this gap by comparing the Mixed Effect Model (LMM) and XGBoost approaches in explaining the influence of *zakat* on the HDI in a more comprehensive and adaptive manner to the characteristics of provincial panel data in Indonesia.

III. Research Methods

A. Data Collection

The Human Development Index (HDI), population data per province, and total *zakat* collection and distribution per province are used. The data comprises data from 34 provinces covering the 2021-2024 period. Several data sources used include West Java Open Data for HDI and population data, and the National *Zakat* Management Report published by the PPID BAZNAS of the Republic of Indonesia for *zakat*-related data. The attributes of the data used are as follows.

Table 1. Dataset Attributes

Data Attribute	Description
id	Unique identifier for each data row

Data Attribute	Description
province_code	West Java provincial code according to BPS regulations
province_name	West Java provincial name according to BPS regulations
human_development_index	Human development index value by province
unit	Human development index measurement unit (POIN)
year	Year of data production
population	Population by province
total_*zakat*	Total *zakat* collected by province
humanity	Total *zakat* distributed to humanitarian sectors by province
health	Total *zakat* distributed to health sectors by province
education	Total *zakat* distributed to education sectors by province
economy	Total *zakat* distributed to economic sectors by province

B. Data Understanding

At this stage, several basic checks were conducted to understand the structure, data quality, and key characteristics of the data. These steps were as follows:

- Checking the number of columns and rows of data, data types, and value distribution.
- Identifying missing values and duplicate data, and validating the conformity of province codes and province names.
- Identifying outliers in numeric variables using the IQR (Interquartile Range) method to detect extreme values in the data distribution.
- Correlation analysis between target features and independent variables was conducted to understand the initial relationship patterns between features.
- Evaluating variability between provinces to determine the extent of variation in the human development index caused by spatial differences using the Intraclass Correlation Coefficient (ICC) calculation.

Initial dataset exploration results showed the dataset consisted of 136 rows of data and 11 columns. The initial check revealed one missing value in the total_*zakat*, humanity, education, health, and economy columns. Validation of the province_code and province_name columns met BPS administrative standards, as this dataset covers a total of 34 Indonesian provinces.

Identification of outliers using the IQR method revealed a number of extreme values in various columns, with a total of 35 rows of data containing at least one outlier. However, upon closer examination, these extreme values were deemed valid because they represented the actual state of inter-regional inequality and were not the result of input errors.

Correlation analysis showed a positive relationship between the Human Development Index (HDI) and all *zakat* distribution variables, although not particularly strong. The highest correlation was found in total_*zakat* (0.28), followed by *zakat* distribution in education, humanitarianism, economics, and health.

The results of the inter-provincial variability evaluation showed an Intraclass Correlation Coefficient (ICC) of 0.889, or 88.9%, indicating variation in the HDI between provinces and reinforcing the assumption that spatial factors, including differences between provinces, significantly contribute to differences in HDI values in Indonesia.

C. Data Preparation

Based on the results of the previous stage, a series of data improvements and transformations will be carried out in this stage to address previously identified data quality issues. The following steps will be taken:

- Handling missing values and duplicate data. Missing values are handled using group mean imputation based on province code groups.
- Creating derived variables or feature engineering to provide additional information to the model. The derived variables added include: *zakat*_per_capita (the ratio of total *zakat* to population) to capture the intensity of *zakat* distribution per individual, and tahun_num, which converts the year to a numeric value relative to the initial year of the data (2021) to capture temporal trends.
- Exploratory Data Analysis (EDA) includes variable distribution analysis, correlation analysis and relationships between variables, and other data analysis.

EDA findings indicate quite striking differences in the geographic sector. The provinces with the highest total *zakat*, namely West Java, Central Java, East Java, Aceh, and Jakarta, are not comparable to the provinces with the highest Human Development Index (HDI), namely Jakarta, Yogyakarta, East Kalimantan, Riau Islands, and Bali. Furthermore, the provinces with the highest *zakat* per capita, namely Aceh, North Kalimantan, West Sumatra, West Nusa Tenggara, and Riau, do not overlap at all with the list of provinces with the highest HDI.

Meanwhile, based on the average calculation across all provinces, the largest proportion of *zakat* distribution nationally is found in the humanitarian sector, followed by education, economics, and health.

The geographic disparities revealed by this EDA indicate that the relationship between *zakat* (both total and per capita) and the HDI exhibits non-linear characteristics, the effectiveness of *zakat* on the HDI may differ significantly between Java and outside Java, and the distribution of *zakat* across sectors has the potential to be a stronger predictor than total *zakat*, given that allocations to education and the economy likely have a direct impact on HDI components.

D. Modelling

1) Pre-Modeling Data Preparation

Initial modeling preparation was carried out by dividing the data into training and test data. The data was divided based on time, with the last year used as test data, while the remaining year served as training data. This approach was chosen to maintain temporal validity in the predictions and

also to allow for a more honest evaluation of the model's generalizability to previously unseen time periods.

The predictor data variables used in the modeling consisted of eight variables: total_*zakat*, humanity, health, education, economy, *zakat*_per_capita, year_num, and province_code. These were selected based on theoretical assumptions and previous exploratory findings that *zakat* distribution across various sectors can influence human development achievements at the provincial level. Meanwhile, the target variable being predicted was the human development index.

In addition to data division, a distribution transformation process for numeric features exhibiting high skewness was also performed: the total_*zakat*, humanity, health, education, economy, and *zakat*_per_capita columns, using the Yeo-Johnson method. This method handles data with zero or negative values to make the feature distribution more symmetrical and closer to normal. Next, numeric features were standardized using StandardScaler to ensure each feature has a proportional scale with a mean of zero and a standard deviation of one.

The transformation and standardization process was carried out using an approach that differentiates between training and test data. The training data was fitted with the fit_transform() process, while the test data was only used with transform(), with parameters calculated from the training data to prevent data leakage and ensure valid model evaluation.

2) Mixed Effects Modeling

To explain the influence of *zakat* distribution on the Human Development Index (HDI), the mixed effects model allows for the separation of fixed effects, which apply to all observations, and random effects, which reflect variations between provincial groups.

The initial modeling strategy began with a random intercept model as a baseline, allowing each province to have a different starting point for the HDI while maintaining the assumption that the slope or influence of each predictor is uniform. Next, several random slope configurations were explored by incorporating random components into specific *zakat* distribution variables: humanitarian, health, education, and economic. This stepwise approach aimed to identify whether the effectiveness of *zakat* distribution in improving the HDI was consistent across regions or influenced by local characteristics.

3) XGBoost Modeling

As a comparison to mixed effects, modeling was also conducted using XGBoost, which has the ability to automatically capture non-linear relationships and feature interactions. However, because XGBoost does not directly support categorical features, one-hot encoding of the provincial region variable was performed to capture differences in regional contributions to the HDI without the random structure of the Mixed Effects Model.

Initially, the model was built with default parameters (baseline) to obtain a baseline performance. Next, to achieve a balance between predictive accuracy and generalizability to test data, optimal hyperparameters were searched for through grid search or random search.

E. Model Evaluation

The performance and effectiveness of the two models were evaluated by comparing the following two main evaluation matrices:

- Root Mean Squared Error (RMSE), a matrix that measures the average error in the model's predicted values to identify whether the model's predictions are significantly off the actual values.
- R-squared (R^2), a matrix that indicates how well the model captures existing patterns by calculating the proportion of variation in the target variable that the model successfully explains.

In addition to prediction accuracy, the evaluation also considers the extent to which the model experiences overfitting, where the model performs very well on the test data but not on the training data. Therefore, a side-by-side evaluation of the performance differences on the training and test data will be conducted to assess model stability and generalization capabilities.

IV. Result and Discussion

This study aims to evaluate and compare two modeling approaches—the Mixed Effect Model (MEM) and Extreme Gradient Boosting (XGBoost)—in explaining and predicting the influence of *zakat* distribution on the Human Development Index (HDI) in Indonesia at the provincial level, using panel data from 2021–2024. The main focus of this discussion is to assess model performance in terms of predictive accuracy, generalizability, and ability to capture the complexity of relationships between variables. It also considers interpretability and model stability in the context of public policymaking.

A. Mixed Effect Model Results

Initial modeling was conducted using a Mixed Effect Model (MEM) approach, combining fixed effects (such as total *zakat* distribution across various sectors: education, health, humanitarianism, and the economy) and random effects (differences between provinces). The modeling strategy began by building a basic model using a random intercept, followed by a more complex model with random slopes for several *zakat* distribution variables.

Initial evaluation results on the training data showed that the MEM model, with its complex structure (e.g., random slopes across the four *zakat* sectors), had very high accuracy, with an RMSE of only 0.021 and an R^2 of 0.99997. However, when the model was tested on 2024 data (never before seen), the prediction performance declined drastically, with the RMSE rising to 1.76 and the R^2 dropping to 0.72. This phenomenon indicates that the model experienced severe

overfitting, meaning it overfitted the training data and lost its ability to generalize to new data.

To address overfitting, adjustments were made to the shrinkage parameters. This involved first testing several shrinkage parameters to find the optimal shrinkage value, which in this case was found to be 0.4. This adjustment significantly improved the model's stability. One of the best configurations was found with a model with a random slope for the *zakat* distribution across three sectors: health, economics, and education. This model demonstrates a good balance between accuracy and generalization ability, with an RMSE value of 1.53 on the training data and 1.64 on the test data (Figure 1). The R^2 value of 0.84 on the training data and 0.76 on the test data indicates stable performance and no overfitting (Figure 2).

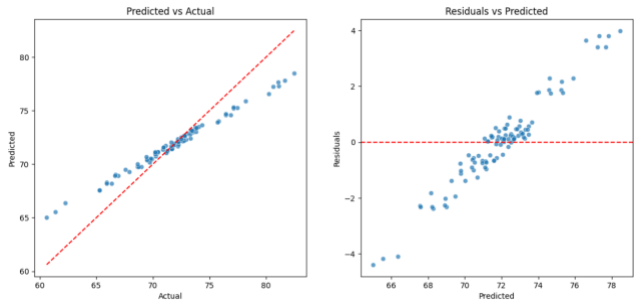

Fig 1. Plot of MEM predicting train data

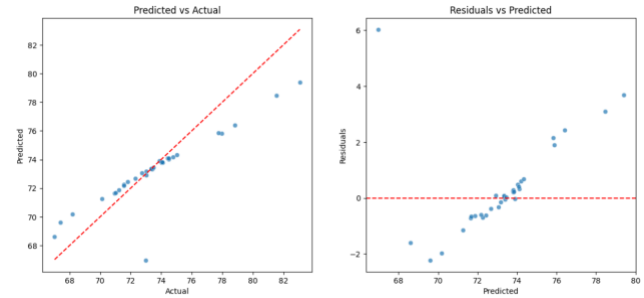

Fig 2. Plot MEM predict test data

In general, the MEM approach offers advantages in interpretability, as the model structure allows for direct observation of the specific influence of each *zakat* distribution sector on the HDI, while also accounting for variations between provinces that may be caused by differences in regional policies, *zakat* management capacity, or local socio-economic conditions.

B. XGBoost Results

For comparison, modeling was also performed using the Extreme Gradient Boosting (XGBoost) algorithm, a machine learning approach designed to handle non-linear relationships and complex interactions between variables without assuming data distribution. Initially, XGBoost was built with default parameters. The results showed very high accuracy on the training data (RMSE: 0.00096, R^2: 1.00), but this performance was not reflected in the testing data, which showed a significant decrease in accuracy (RMSE: 1.76, R^2: 0.72). This indicates that the model suffered from severe overfitting, with the RMSE ratio between the training and testing data reaching over 1800x.

To reduce overfitting, parameter tuning was performed using RandomizedSearchCV and GridSearchCV. Several configurations were tested, including adding L1(α) and L2(λ) regularization, as well as reducing tree complexity. The best parameter tuning results were found in the configuration with higher regularization provides in Table 2

Table 2. Evaluation Result o XGBoost Tuning

Train RMSE	0,875
Test RMSE	1,988
Train R^2	0,948
Test R^2	0,647

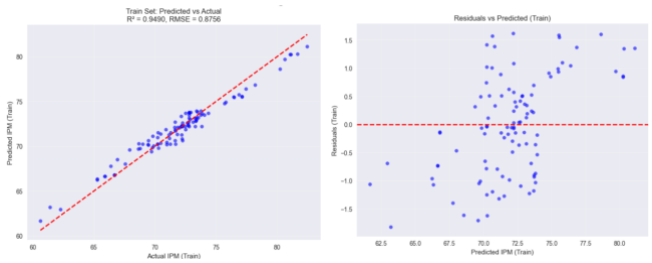
Fig 3. Plot XGBoost predict train data

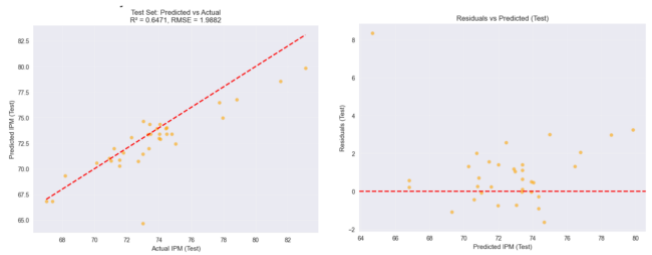
Fig 4. Plot XGBoost predict test data

Despite the reduction in overfitting, the model's predictive performance remains below MEM, particularly in terms of performance stability from training data to testing data. This indicates that XGBoost, while capable of capturing complex non-linear patterns, requires careful tuning to avoid overfitting, especially when the data is not as large as in this study.

Furthermore, another weakness of XGBoost is its lack of direct interpretability, particularly in the context of public policy. It is not easy to directly interpret how each *zakat* sector contributes to the increase in the HDI, as the model focuses more on predictive accuracy than interpretation.

C. Perbandingan dan Implikasi

Based on the comparison table of the two models, the Mixed Effects Model (MEM) demonstrated more stable performance and generalization than XGBoost for the study of *zakat* distribution and the effect of *zakat* on the HDI.

Table 3. Model evaluation comparison results

	MEM	XGBoost
Train RMSE	1,521	0,875
Test RMSE	1,633	1,988
Train R^2	0,846	0,948
Test R^2	0,762	0,647

Although XGBoost had excellent RMSE and R^2 values on the training data (RMSE = 0.875 and R^2 = 0.948), its performance declined drastically on the testing data (RMSE = 1.988 and R^2 = 0.647), indicating overfitting, where the model generalizes less well to new data. Conversely, MEM performed better on new data and was more stable in capturing existing patterns in the data, with relatively small differences between the training and testing data (RMSE = 1.521 vs. 1.633 and R^2 = 0.846 vs. 0.762).

These modeling results align with previous EDA findings that identified geographic disparities and non-linear characteristics in the *zakat*-HDI relationship. Although EDA exhibits nonlinear characteristics that should benefit XGBoost, MEM actually outperforms it in terms of generalization. This occurs because XGBoost, with its highly flexible nonlinear capabilities, is "stuck" in capturing noise and specific patterns from the limited training data (136 observations), thus losing its ability to generalize to true patterns. Furthermore, given the hierarchical and geographic structure of the data, even one-hot encoding of the province variable does not allow the model to assume that each region may have a different baseline or response to *zakat* distribution.

In contrast, MEM, with its simpler structure but still able to capture variation between provinces through random effects, successfully identified more robust and generalizable patterns. Modeling the structure through random intercept and random slope components makes this model superior in capturing complex spatial and socioeconomic dynamics, even though its model is mathematically simpler than XGBoost.

D. Deployment

The *zakat* fund allocation-based HDI prediction system, developed using the Mixed Effects Model (MEM) that has gone through evaluation and validation stages, is then deployed through the Streamlit platform as an interactive web application. The deployment process begins by saving the evaluated MEM model in pickle (.pkl) format for easy reloading. The web application was developed using Streamlit by creating an app.py file that serves as the main user interface. The interactive dashboard allows users to input parameters such as population size, total *zakat* funds, and the percentage of allocation to the education, health, economic, and humanitarian sectors. The system also displays the predicted HDI value from the collected *zakat* and its distribution as well as an analysis of the contribution of each sector to changes in the HDI and provides strategic recommendations for optimizing *zakat* distribution.

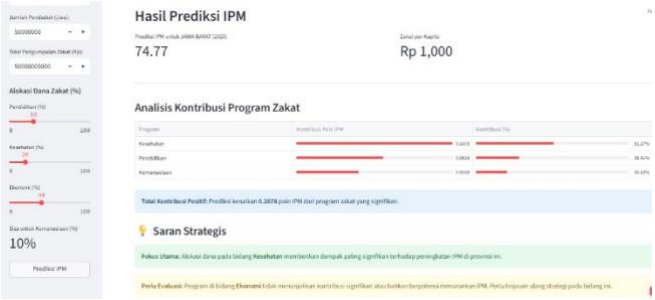

Fig 5. Deployment project

V. CONCLUSION

This study comprehensively examines the influence of *zakat* distribution on the Human Development Index (HDI) in Indonesia using a quantitative approach based on panel data from 34 provinces during the 2021–2024 period. By comparing two modeling methods—the Mixed Effect Model (MEM) and Extreme Gradient Boosting (XGBoost), this study assesses the effectiveness and accuracy of each model in explaining the complex relationship between sectoral *zakat* distribution and the HDI.

The results indicate that *zakat* distribution has a positive effect on increasing the HDI. However, the relationship between the two is complex and not always linear, as demonstrated by data exploration results that found a weak to moderate correlation and significant spatial disparities between provinces. In this context, the Mixed Effect Model proved superior to XGBoost. MEM is able to capture spatial and temporal variation through a combination of fixed and random effects and demonstrates good predictive stability with relatively balanced RMSE and R^2 between training and test data. Conversely, although XGBoost is highly accurate in learning from training data, it tends to experience significant overfitting on test data, primarily due to the limited size of the dataset.

Furthermore, the MEM's advantage lies in its interpretability. With an explicit model structure, policymakers can trace the contribution of *zakat* distribution sectors such as education, health, economics, and humanitarianism to improving the Human Development Index (HDI) in each province. This is crucial for designing a more targeted and evidence-based *zakat* distribution strategy. In the context of sustainable development, strategic *zakat* distribution can accelerate the achievement of several Sustainable Development Goals (SDGs), particularly in the areas of poverty alleviation, quality education, and inequality reduction.

Methodologically, this study also makes an important contribution by integrating classical statistical approaches (MEM) and machine learning (XGBoost), demonstrating the limitations and advantages of each in handling socio-economic data. These findings encourage the adoption of models that are not only predictively accurate but also capable of providing applicable policy insights. Therefore, the Mixed Effects Model approach is recommended as a more

relevant and effective analytical framework for evaluating the impact of *zakat* distribution on human development in Indonesia.

REFERENCES

[1] "BAZNAS RI 2023 Zakat and Poverty Alleviation Report," pp. 0–24, 2024.

[2] N. Agniya, T. Badina, and M. A. Najib, "The Effect Of Zakat Performance And Macroeconomic Variables On Poverty Levels (Empirical Study In 34 Provinces In Indonesia)," *International Collaboration Conference on Islamic Economics*, pp. 1–14, 2023.

[3] T. Widiastuti and L. N. Rani, "Evaluating the Impact of Zakat on Asnaf's Welfare," *Global Journal Al-Thaqafah*, vol. 5, no. SpecialIssue, pp. 91–99, 2020.

[4] T. Widiastuti, A. Prasetyo, A. Robani, I. Mawardi, R. Rosida, and M. U. Al Mustofa, "Toward developing a sustainability index for the Islamic Social Finance program: An empirical investigation," *PLoS One*, vol. 17, no. 11 November, pp. 1–26, 2022, doi: 10.1371/journal.pone.0276876.

[5] M. R. Gunawan and A. S. Rusydiana, "Zakat and Sustainable Development Goals (SDGs)," *Indonesian Conference of Zakat ...*, no. 2017, pp. 39–49, 2023.

[6] E. Tok, A. J. Yesuf, and A. Mohamed, "Sustainable Development Goals and Islamic Social Finance: From Policy Divide to Policy Coherence and Convergence," *Sustainability (Switzerland)*, vol. 14, no. 11, 2022, doi: 10.3390/su14116875.

[7] A. Mukadar, M. Marini, and K. Pramadeka, "Pengaruh Zakat Terhadap Indeks Pembangunan Manusia (Studi Kasus Pada BAZNAS Provinsi Bengkulu)," *EKOMBIS REVIEW: Jurnal Ilmiah Ekonomi dan Bisnis*, vol. 11, no. 2, pp. 1271–1282, 2023, doi: 10.37676/ekombis.v11i2.4072.

[8] W. Kurniawan and U. Indahyanti, "Prediksi Angka Harapan Hidup Penduduk Menggunakan Metode XGBoost," *Indonesian Journal of Applied Technology*, vol. 1, no. 2, p. 18, 2024, doi: 10.47134/ijat.v1i2.3045.

[9] B. A. Lipesa, E. Okango, B. O. Omolo, and E. O. Omondi, "An application of a supervised machine learning model for predicting life expectancy," *SN Appl. Sci.*, vol. 5, no. 7, 2023, doi: 10.1007/s42452-023-05404-w.

[10] Y. M. Indah, R. Aristawidya, and A. Fitrianto, "Comparison of Random Forest , XGBoost and LightGBM Methods on the Human Development Index Classification," *Jambura Journal of Mathematics (JJM)*, 2025.

[11] S. Setyadi, S. Didu, L. Indriyani, A. K. Fitri, and A. Widiastuti, "Modeling Life Expectancy in Indonesia Using System GMM Model," *Review of Applied Socio-Economic Research*, vol. 25, no. 1, pp. 83–98, 2023, doi: 10.54609/reaser.v25i1.338.

[12] P. Jarupunphol, "Assessing decision tree , random forest , and XGBoost models for human capital readiness predictions in low-income areas," *Multidisciplinary Science Journal*, vol. 18, 2025.

[13] E. Rusanti, A. S. Anindya, N. Atiya, and T. Widiastuti, "Does Zakat Impact on Human Development Index? Empirical Evidence from Indonesia," *International Journal of Zakat*, vol. 8, no. 2, pp. 36–47, 2023.

[14] S. Mutmainah, N. W. I. Rahayu, F. Fauzan, and ..., "The influence of zakat, Human Development Index, open unemployment rate, and income on poverty in Indonesia," *... of Islamic Economics ...*, vol. 10, no. 1, pp. 363–382, 2024.

[15] I. I. Wardani and M. N. R. Al Arif, "The Effect of Sharia Bank Financing, Zakat, and Education Expense, on Economic Growth and Human Development Index in Indonesia 2015-2019," *EKONOMIKA SYARIAH : Journal of Economic Studies*, vol. 5, no. 1, p. 1, 2021, doi: 10.30983/es.v5i1.4096.

[16] Z. Hasan, "The Effect of Economic Growth and Human Development Index on Poverty in Indonesia," *Journal of Economics and Sustainability*, vol. 3, no. No.1, pp. 42–53, 2021, doi: 10.32890/jes2021.3.1.5.

[17] E. A. N. Choiriyah, A. Kafi, I. F. Hikmah, and I. W. Indrawan, "Zakat and Poverty Alleviation in Indonesia: a Panel Analysis At Provincial Level," *Journal of Islamic Monetary Economics and Finance*, vol. 6, no. 4, pp. 811–832, 2020, doi: 10.21098/jimf.v6i4.1122.

[18] A. Das, M. Uddin, and R. Karim, "Predicting Life Expectancy using Machine Learning Techniques * Correspondence should be addressed to Md . Rezaul Karim," *International Journal of Statistical Sciences*, vol. 25, no. 1, pp. 55–70, 2025.

[19] I. K. B. Akmal, M. A. S. Majid, and E. Gunawan, "Does Zakat matter for human development? An empirical evidence from Indonesia," *Regional Science Inquiry*, vol. 12, no. 2, pp. 195–208, 2021.

[20] I. Isman and Y. Yanuri, "Maq āṣ id Al-Syar ī ' ah Analysis of the Implementation of Food Social Assistance on Multidimensional Welfare in Indonesia (2020-2023) Maqā ṣ id Al-Syarī ' ah Analysis of the Implementation of Food Social Assistance on Multidimensional Welfare in Indone," *Al-Iktisab: Journal of Islamic Economic Law*, no. May, 2025, doi: 10.21111/aliktisab.v8i2.12972.

[21] N. H. Abd Samad, R. Abdul Rahman, S. Masrom, N. Omar, and H. Hasan, "Predicting the intention to adopt e-zakat payment services: a machine learning approach," *Bulletin of Electrical Engineering and Informatics*, vol. 14, pp. 2330–2337, 2025, doi: 10.11591/eei.v14i3.8512.

[22] Z. Muttaqin and M. D. A. Nasir, "Can Zakat Contribute to Achieving Sustainable Development Goals? A Case Study on Java Island, Indonesia.," *Jurnal Ekonomi Syariah Teori dan Terapan*, vol. 11, no. 1, pp. 35–53, 2024, doi: 10.20473/vol11iss20241pp35-53.

[23] M. I. Rosele, A. Muneem, N. N. B. A. Rahman, and A. K. Ali, "The Digitalized Zakat Management System in Malaysia and the Way Forward," *Al-Ihkam: Jurnal Hukum dan Pranata Sosial*, vol. 17, no. 1, pp. 242–272, 2022, doi: 10.19105/AL-LHKAM.V17I1.5365.

[24] Y. Satyakti, "The Effect of Applying Sustainability (Maqasid Shariah) and Competition on Islamic Bank Financing," *Sustainability (Switzerland)*, vol. 15, no. 17, 2023, doi: 10.3390/su151712994.

[25] Paulina, M. Erwan, and S. D. Sadikin, "Zakat and Its Impact on Economic Growth in Indonesia and Malaysia," *The Seybold Report*, vol. 18, no. 10, pp. 1187–120, 2023, doi: 10.5281/zenodo.10171370.