Application of the Multinomial Logistic Regression Model to Analyze the Impact of E-Government on Youth in Albania

ERALDA XHAFKA[1], JONIDA TETA[1], ERALDA GJIKA[2]
[1]Department Production and Management, Faculty of Mechanical Engineering,
Polytechnic University of Tirana,
Mother Tereza Square 1, Tirana
ALBANIA

[2]Department of Informatics,
Faculty of Natural Sciences, University of Tirana
Bulevardi Zogu I, Tirana,
ALBANIA

Abstract: - E-Government is using digital technology to improve services and increase civic participation. By digitalizing public services and transforming e-governance, governments are changing the interaction between people and institutions, to make operations more transparent, efficient, and accessible. More important is the fact that digitalized transactions reduce the burden on citizens and businesses by making public services faster, more convenient, and less costly. Understanding the e-government level of use and the experience of young people is important for successfully implementing public service digitization. This study aims to analyze the various factors that affect young users' experience of e-government and how the digitalization of public services impacts Albanian youth, focusing on their perceptions, experiences, and how these services influence their integration into the digital world. Employing a multinomial logistic regression analysis through a survey of young people from different parts of the country, the objective is to understand how they feel about the e-governance experience. The study explores the relationship between user experience (binary dependent variable) and key independent variables, including ease of use, data security, cost reduction, time reduction, bureaucracy, human error reduction, system malfunctions, lack of skills, and access to the internet. To identify the most relevant predictors for the multinomial outcome variable, we conducted a stepwise selection process using multinomial logistic regression which considers all possible variables in the reduced dataset from the survey. To ensure the robustness of our selected model, we performed 10-fold cross-validation using the `cv.glm` function in R. Variance Inflation Factor (VIF) scores were calculated to evaluate multicollinearity among predictors. Results: The logistic regression model reveals that Corruption and system malfunctions significantly influence users' experience, of young people and are more likely to report satisfaction. In the final step of the stepwise model selection, only two variables were found to be important and with a low VIF: Corruption and System Malfunction. Corruption and system malfunctions emerge as strong predictors of experience satisfaction with e-government. The model demonstrates good fit and predictive ability, as indicated by statistical tests and graphical analyses, including a Predicted Probability Plot. The findings show that many young people view e-governance as a positive step towards modernizing public services and promoting transparency. However, there are still challenges, like ensuring access to technology and improving digital literacy. Addressing these issues will help maximize the benefits of digital transformation for all young citizens.

Key-Words: - E-government, Albanian youth, access to public services, corruption, system malfunction, logistic regression model.

Received: August 19, 2024. Revised: May 17, 2025. Accepted: June 15, 2025. Published: August 4, 2025.

1 Introduction

By most accounts, the economies of the Western Balkans are lagging behind the EU's digital transformation. If properly supported and managed, the digital transformation in the Western Balkans can have clear benefits. From an economic perspective, this includes higher growth and greater levels of employment. Furthermore, it can have political repercussions by contributing to lower

discretion, more rule-based administration, greater regional connectivity and harmonization, and accelerated EU integration, [1]. Despite challenges such as low digital literacy, limited internet connectivity, poor ICT infrastructure, and budget constraints, e-government projects have shown potential to reduce bureaucratic inefficiencies, overcome geographical barriers, and improve responsiveness. E-Government (electronic government) refers to the use of digital technology to improve the delivery of public services and increase civic participation. This has a positive socio-economic impact, promoting inclusive development, empowerment of marginalized communities, and grassroots good governance practices, [2]. The governance of digitalization, therefore, requires governments to govern the use of digital tools and data in the public sector and beyond on the one hand and manage their digital transformation of the public sector with new and adapted governance approaches on the other hand, [3]. E-government is a broad concept that refers to electronic interactions between governments, citizens, businesses, and other government agencies. It's about how technology can reshape the way government functions, making processes more efficient and accessible. In Albania's journey towards European Union membership, the country has undertaken significant reforms, embracing technology to meet international and European standards.

In recent years, Albania has seen strong public support for the use of information technology, which has improved the quality and efficiency of public services. Systems for business registration, public procurement, tax payments, the civil registry, and obtaining criminal records are now digital, making these processes more transparent, faster, and less costly for both citizens and businesses.

The digitalization strategy aimed to introduce electronic transactions within the public administration's internal processes, ensuring that essential services for individuals and businesses could be accessed online. The goal was to provide simple, fast, and free access to information about public services, improving transparency and efficiency in how government functions.

Through these initiatives, Albania seeks to modernize governance and create a more open and sustainable society. This aligns with Albania's regional cooperation efforts and its coordination with the European Union as part of its integration process.

Today, e-government in Albania includes platforms like e-Albania (the official platform for accessing public services online), the National Agency for the Information Society (AKSHI) which is in charge of the implementation of digital public services, and the Agency for the Provision of Integrated Services (ADISA), which provide digital services for individuals and businesses both within and outside the country. The government's investment in digitizing public services has greatly reduced the financial and time burdens on citizens by eliminating the need for physical certificates and long lines at government offices.

Since May 1, 2022, 95% of all public services provided by the Albanian state administration have gone online through the e-Albania portal. According to INSTAT (Institute of Statistics of Albania) [4], there are now 1,227 online services available from 83 state institutions, supported by 58 digital registers. Over 2.8 million users are registered on e-Albania, including 220,000 businesses.

With 99.9% of the country's territory covered by online services, the government's priority now is to continue improving the quality of these services, expand them to cover local services across all 61 municipalities, and reduce the reliance on cash-based transactions. The data we analyzed is from a government-mandated platform that provides essential services: accessible information, taxes, education, payrolls, pensions, utility bill payments, and judicial information. This is not a platform whose use is optional; rather, citizens must use it to access critical services. With that in mind, our focus was on the identification of which variables drive a "positive" response is to say, "yes, we use the platform"-against responses classified as "neutral" and "negative" aggregated. Given this, our choice of multinomial logistic regression is more apt to handle binary-like behavior than the levels of satisfaction ordered strictly.

2 Problem Formulation

According to OECD Digital government explores and supports developing and implementing digital government strategies and initiatives that bring governments closer to citizens and businesses. It recognizes that technology is a strategic drive to improve public sector efficiency and make policies more effective and governments more open, transparent, innovative, participatory, and trustworthy, [5].

The 2021 OECD Survey on Trust in Public Institutions offers a clear picture of what makes people trust institutions and how we can make democracies more resilient. One key takeaway from the data is that young people tend to have less trust

in public institutions. Many also feel like their voices aren't being heard. For example, only 22% of members of parliament in OECD countries are under 40, and in some countries, less than 5% of the central government workforce is under 35.

The COVID-19 pandemic hit young people especially hard, which likely impacted their trust in public institutions even further. Despite this, young people have been at the forefront of social movements, especially in pushing for stronger action on climate change. They also played a vital role in helping communities stay resilient during the pandemic. The criteria for the adoption of e-government from both citizens and government perspectives is highly dependent on two major factors, trust and security, [6]. Trust is probably one of the most important aspects in the implementation of e-government plans. Without trust, citizens will not participate in the e-government process, [7].

Other factors impact the building of trust. For instance, studying the individual (citizen) behavioral attributes would improve institutional attributes, developing internal technology, quality of services provided, and fulfillment of services. Such expectations put more load on governments to deliver systems that are efficient and effective.

E-participation, as a concept, has been in the public domain for almost two decades. Growing evidence points to the rapid expansion of e-participation as a tool for engagement and strengthened collaboration between governments and citizens.

The UN defines e-participation as "the process of engaging citizens through ICTs in policy, decision-making, and service design and delivery to make it participatory, inclusive, and deliberative". The objective is to improve access to information and public services as well as to promote participation in

policymaking, both for the empowerment of individual citizens and the benefit of society as a whole.

E-participation is not only a subset of participation and e-governance but is also closely related to other elements of governance such as citizen inclusion, transparency, and accountability. The national information technology plan for each country focuses on Information Communication Technology (ICT) as a tool to reform public organizations. Therefore, the objectives are similar in concept but differ in approach based on the structure of the country, [8]. Strengthening democracy in the digital age is complicated. It requires exploring how digital tools can improve participation, transparency, and accountability in government, while also being mindful of the risks that digitalization might pose.

Many governments are already working on innovative solutions to these challenges, creating new frameworks to protect digital rights and democracy, [9]. Albania has made the modernization of information technology one of its top seven long-term goals as part of its National Development Plan to help transition to a market economy.

The level of digitalization development in Albania, unlike economic, social, and legal parameters, has progressed at a relatively satisfactory pace compared to developed countries. In some sectors, it is almost at the same level and stage of development. According to data from INSTAT (Institute of Statistics of Albania) in 2016, the vast majority of enterprises (95.6%) and a significant portion of employees (28%) in Albania used computers for work purposes. Additionally, 96.8% of enterprises had Internet access, and 7.1% of them sold products through the Internet, [10].

A wide range of public services are in the process of digitalization, offering not only convenience but also cost reductions for both the state and citizens (e.g., e-Albania). Audiovisual broadcasts are in the process of full digitalization, providing advantages in terms of expanding capacities and services. In the financial sector, cyber systems, mainly represented by banking networks, have seen improvements over the past 12 years, especially after 2005. The progress of electronic payments throughout 2022 has contributed to the fulfillment of the quantitative objectives of the National Strategy of Retail Payments (2018-2022), reaching a level of use of electronic payments percapita at 17.19 out of 10 payments which is the objective of the Strategy, [11]. Certain commercial companies in various sectors are making individual investments to integrate digital technology to take advantage of the benefits it offers in terms of increasing capacities and reducing costs.

Albania has made significant strides in digitalizing its public services, standing out compared to other countries in the region. The e-Albania platform now offers a wide range of services for both businesses and individuals, allowing people to access things like certificates, health cards, and tax status verifications in real-time and at minimal cost. Additionally, services related to vehicles, driving licenses, and real estate permits are available online, making these once-complicated tasks much easier to handle. In the healthcare sector, there's an ongoing project to digitize medical recommendations, prescriptions, and other

administrative tasks, further simplifying processes for citizens.

By 2023, almost all interactions between Albania's public and authorities- such as filing taxes, accessing social benefits, obtaining official documents like ID cards or birth certificates, and even public educational services moved online. The biggest growth was seen in people downloading and printing official forms, with a 2.8% increase compared to 2022. Additionally, there was a 1.9% increase in people getting information from websites and apps, showing that more Albanians are embracing digital services as part of their daily lives, [12].

This shift towards digital services has made life easier for citizens, saving time and reducing the need to wait in line for various services.

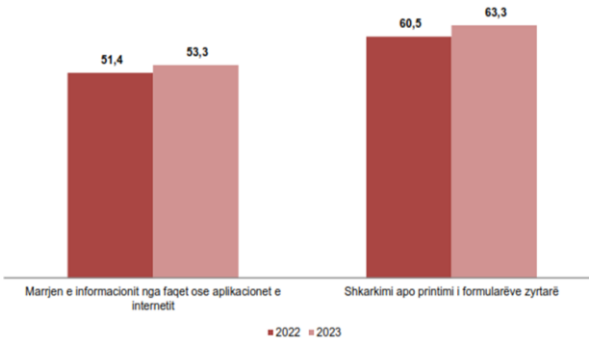

Fig. 1: Use of E-Government 2022-2023 in Albania, [12]

The Figure 1 shows a positive trend in the use of e-government services between 2022 and 2023:

Accessing Information increased slightly from 51.4% to 53.3% (+1.9 percentage points), indicating modest growth.

Downloading or Printing Official Forms rose from 60.5% to 63.3% (+2.8 percentage points), showing stronger engagement.

Factors like improved ICT infrastructure, awareness, and system usability likely contributed to these trends.

Prioritizing raising awareness, improving usability, addressing digital literacy, and broadening access to technology will help to further boost the involvement of citizens. The Government of Albania has identified ICT as an enabler for the social and economic development of the country. They are introducing ICTs and new technology tools in the day-to-day work of Central Ministries and Governmental Agencies. Reports periodic to UNPAs (Network of Administration Public to UN) make a rating to e-government Readiness index and e-participation. These data are based on a survey of the Site's official website, [13].

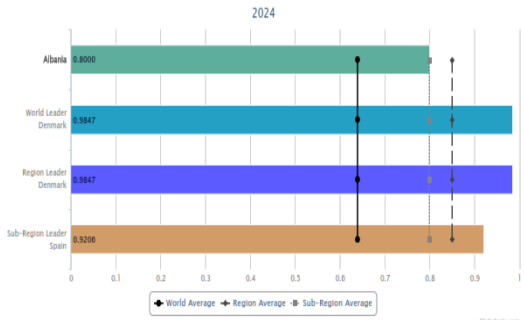

Fig. 2: E-Government Development Index, [14]

Reports periodic to UNPAs (Network of Administration Public to UN) make a rating to e-government Readiness index and e-participation. Results for web access and services online issued. As shown in Figure 2 the Development Index of Albania is 0,8 and it is ranked in the 62nd place of 193, [14]. Table 1 illustrates the rank and the value of Albania measured by the Reports periodic to UNPAs (Network of Administration Public to UN).

Table 1. Albania E-Government Development Index, [14]

E-Government Development Index	2024	2022	2022	2018	2016	2014	2012	2010	2008
Albania (Rank)	62	63	59	74	82	84	86	85	86
Albania (Value)	0.8	0.7413	0.7399	0.6519	0.533	0.5045	0.516	0.4519	0.467

3 Theoretical Framework: Multinomial Logistic Regression Model

In the context of understanding user experiences with various services, a multinomial logistic regression model is employed to assess the impact of several predictor variables on the categorical response variable, which represents different levels of user sentiment: negative (Class 0), neutral (Class 1), and positive (Class 2). This approach is particularly suitable when the dependent variable is nominal with more than two categories, allowing for a comprehensive analysis of factors influencing user perceptions. The multinomial logistic regression model estimates the probability of each outcome category relative to a reference category, often chosen as the most frequent or a baseline class. In this study, the user sentiment is modeled as a function of multiple independent variables, including "Corruption", “System malfunctions”, and other relevant factors such as ease of use, data security, and bureaucratic hurdles. The survey categories were selected based on the relevance of their content to our main research objective into factors that influence platform usage behavior.

Although these aims were best approached given the nature of the dependent variable using a multinomial logistic regression, we recognize that there may be further insights from the examination of other survey questions through the use of traditional statistical techniques, such as t-tests for comparing means across groups and Chi-Square tests for examining relationships among categorical variables-for example, user sentiment and system malfunctions. The theoretical framework assumes that the log-odds of the probability of each sentiment class can be expressed as a linear combination of the predictor variables.

Mathematically, the model can be represented as follows:

$$log\left(\frac{P(Y=k|X)}{P((Y=i|X)}\right) = \beta_0^k + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n \; ; for \; k \in E\{i\} \quad (1)$$

where:

- i is the reference category from the set E of possible categories. In our case $E = \{0,1,2\}$
- k is any category other than i
- Y represents the categorical outcome (user sentiment)
- $P(Y=i|X)$ is the probability of the reference category,
- $P(Y=k|X)$ is the probability of each non-reference category k,
- $X_1, X_2, \ldots, X_n$ are the predictor variables
- β_0^k is the intercept for category k,
- $\beta_1, \beta_2, \ldots, \beta_n$ are the coefficients for each predictor $X_1, X_2, \ldots, X_n$

From the above definition where $Y=i$ is the reference category, the probability for each category $Y=k$ (where $k \neq i$) can be written as follows:

$$(Y = k|X) = \frac{e^{\beta_0^k + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n}}{1 + \sum_{j \in E\{i\}} e^{\beta_0^j + \sum_{m=1}^{n} \beta_m X_m}}$$

And for the reference category $Y=i$, the probability is:

$$(Y = i|X) = \frac{1}{1 + \sum_{j \in E\{i\}} e^{\beta_0^j + \sum_{m=1}^{n} \beta_m X_m}}$$

where:

- E is the set of all possible categories,
- k is any non-reference category ($k \neq i$),
- β_0^j is the intercept term for category j,
- β_m are the coefficients for the predictor variables $X_1, X_2, \ldots, X_n$.

This modeling approach systematically estimates the relationship between user experiences and various factors, providing insights into how these factors contribute to differing sentiments. By employing this model in our data, we aim to identify the factors that significantly impact user sentiment (positive, neutral, or negative) and how strongly each factor contributes to the shape of the sentiment. Thus, gaining deeper insight into how users perceive and interact with the service.

4 Methodology

4.1 Data Collection and Preparation

Logistic regression has emerged as a cornerstone analytical tool in customer satisfaction research, primarily due to its adeptness at handling binary outcome analysis. This study employs logistic regression analysis to investigate the factors influencing the user experience of e-government and it suits dichotomous outcome variables like user experience (satisfied vs. unsatisfied). Its robustness in handling continuous and categorical independent variables makes it an ideal choice for this study, [15]. Logistic regression's capability to provide odds ratios for the likelihood of an outcome given specific predictor variables is particularly useful in interpreting the factors affecting user experience, [16]. Its application extends across a spectrum of factors influencing customer satisfaction, from service quality to customer demographics, offering valuable insights into what drives customer contentment in various contexts, [17]. The study aims to provide information regarding the experiences and perceptions of digital platforms, including benefits, challenges, and suggestions for future improvements. The information gathered will serve to better understand the needs and expectations of users, as well as to develop and further enhance electronic government services. To collect the necessary data from the study, a questionnaire has been designed to evaluate citizens' experiences and opinions regarding the use of e-government platforms. The questionnaire is a common tool for data collection and consists of a form or instrument that includes a set of questions and secure responses that respondents (from a specific population) complete to provide the researcher with the necessary information for the study. The data obtained from a questionnaire cannot be acquired from secondary sources, [18]. Therefore, the design of the questionnaire is of great importance to ensure the collection of accurate data

so that the results are interpretable and generalizable, [19].

The structure of this questionnaire covers a wide range of issues related to the use of e-government platforms. The questionnaire consists of 23 questions, and the first section contains demographic information about age, gender, education level, and employment status. The following section contains questions regarding the use of electronic platforms, users' personal experiences, and suggestions for further improvements.

The questionnaire is based on the Likert scale, where respondents are allowed to choose between a series of alternatives along a predetermined continuum using a scale from 1-5, such as: 1. extremely little, 2. little, 3. sufficient, 4. a lot, 5. extremely much. In this way, key variables such as the quality of services, difficulties in using these platforms, security, and privacy, and citizens' personal experiences with these platforms are assessed. The total sample size is estimated to be 226 individuals who were randomly selected. Respondents were interviewed online using the questionnaire in the Google Forms format. The survey is distributed online via email and social media platforms, as well as in person at selected retail locations. This mixed-mode data collection approach is designed to maximize reach

and response rate, as suggested by [20]. The questionnaire remained online for 45 days, as this was deemed the necessary time to gather the required information and reach a certain number of respondents.

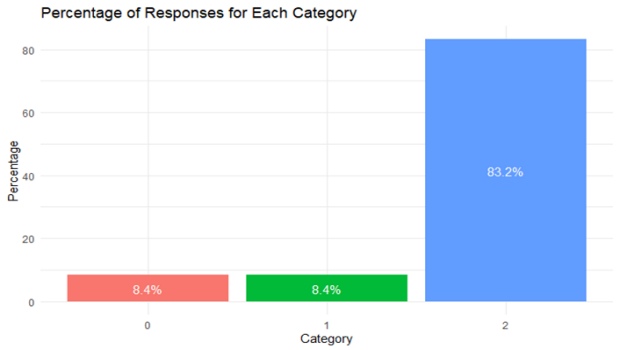

Fig. 3: Experience with e-government services

This study utilized a dataset comprising responses to the survey question, “What is your experience with the digitalization of public services”; which categorized experiences into three classes: '0' (negative), '1' (neutral), and '2' (positive). The dataset included various predictors such as `Corruption`, `malfunction system`, and several additional variables pertinent to the study's context. The distribution of the responses is shown in Figure 3. Most of the reactions 83.2% lies in the Positive experience.

4.2 Stepwise Model Section

For this step and all modeling and tests we used R libraries. The stepwise model selection algorithm goal was to minimize the Akaike Information Criterion, which represents a balance between model complexity and goodness of fit. We also used both forward and backward model selection as part of this stepwise process to decide which variables should be included or excluded from the final model.

The stepwise selection process started with an initial model that could be either null (no predictors) or full (all predictors), depending on the approach. Variables were then added or removed one at a time, based on their contribution to the model fit, measured by changes in the AIC value. Akaike Information Criterion (AIC) is used to compare the relative fit of statistical models, with lower values indicating a better fit. A variable was retained in the model when its retention significantly enhanced the explanatory capability of the model; otherwise, it was excluded if that variable did not contribute significantly.

This iterative process was repeated until no further reductions in AIC could be achieved, ensuring that the final model included only the most informative predictors. By minimizing the AIC, we effectively balanced the trade-off between model parsimony and predictive accuracy, reducing the risk of overfitting. This approach allowed us to derive a robust and interpretable model tailored to our research objectives.

In this study, stepwise selection was chosen for its ability to balance simplicity, interpretability, and efficiency. It provided a systematic approach to identify the most significant predictors, aligning well with the user sentiment and the practical constraints of the dataset and computational resources. Other techniques could have been applied such as Bayesian methods, but they frequently require more advanced computational resources and larger datasets. Stepwise selection, guided by the Akaike Information Criterion (AIC), provides an effective approach and computationally efficient solution that aligns well with the study's objectives and limitations. To identify the most relevant predictors for the multinomial outcome variable, we conducted a stepwise selection process using multinomial logistic regression which considers all possible variables in the reduced dataset from the

survey. The Akaike Information Criterion (AIC) was used as the guiding metric for variable selection. The final stepwise model takes into consideration the variables `Corruption` and ` System Malfunctions ` and based on the lower value of the Akaike Information Criterion (AIC) (AIC=249.47) was selected as the best model. The Residual Deviance of 222.4126 indicates the model's fit, with lower values suggesting a better fit.

The AIC of 266.4126 is used for model comparison, where lower AIC values indicate a better-fitting model. This model retained only predictors that significantly contributed to explaining the variability in user sentiment. Table 2 gives the model coefficients for all variables used in the stepwise model.

Table 2. Model Coefficients for all variables used in the Stepwise model

Variable	**Neutral (1) Coefficient**	**Positive (2) Coefficient**
Intercept	2.316	5.906
Ease of Use	0.864	0.421
Data Security	0.021	0.348
Low Cost	-0.843	-0.46
Bureaucracy	0.009	0.294
Corruption	-0.561	-1.218
Time Reduction	0.692	0.832
Human Error Reduction	-0.225	-0.264
No Internet Access)	-0.083	-0.39
Lack of Skills	-0.274	0.247
System Malfunctions	-0.366	-0.754

4.2.1 Interpretation of Key Findings:

- *Ease of Use*: A one-unit increase in ease of use increases the log odds of having a Neutral or Positive experience relative to a Negative one. The effect is stronger in the Neutral category (0.864) than in the Positive category (0.421) (Table 2).
- *Data Security*: Has a smaller positive effect on the likelihood of experiencing both Neutral and Positive outcomes compared to Negative ones, particularly in the Positive category (0.348) (Table 2).
- *Corruption*: Has a strong negative association with both Neutral and Positive experiences, significantly reducing the likelihood of either outcome, with a larger impact on Positive experiences (-1.218) (Table 2).
- *Time Reduction:* This is positively associated with both Neutral (0.692) and Positive (0.832) experiences, indicating that users who perceive time-saving benefits are more likely to report better experiences (Table 2).

4.2.2 Variable Importance and Multicollinearity Assessment

Variance Inflation Factor (VIF) is a measure used to detect multicollinearity among predictors in a regression model."

First, we computed the VIF for all variables involved in the model and noticed that several predictors had very high VIF values, indicating the presence of multicollinearity (Tab.3). In light of this finding, stepwise regression was adopted to select only the most significant predictors to be taken forward into the final model. Once the final variables were determined, we calculated their respective VIF values again; both were well below the commonly accepted threshold value of 5, as suggested in the literature. This signifies that the final model does not suffer from multicollinearity. Variance Inflation Factor (VIF) (Table 3) scores were calculated to evaluate multicollinearity among predictors, with scores exceeding 5 indicating potential multicollinearity issues. In the final step of the stepwise model selection, only two variables were found to be important and with a low VIF: Corruption and System Malfunction.

Table 3. Variance Inflation Factors (VIF) for Predictor Variables used

	Variance Inflation Factor (VIF)
Ease of Use	92.79
Data Security	82.46
Cost reduction	85.45
Bureaucracy	38.08
Corruption	104.19
Time Reduction	137.66
Human Error Reduction	117.40
No Internet Access	14.87
Lack of Skills	9.21
System Malfunctions	26.91

Table 4. Variance Inflation Factors (VIF) of the final stepwise model

Predictor	Importance score	VIF
Corruption	0.65	4.75
Malfunction systems	0.35	4.20

The variable importance analysis confirmed that the final selected predictors played a crucial role in classifying the outcome. VIF scores indicated that

Corruption ` and ` Malfunction systems `, were within acceptable limits for inclusion in the final model (Table 4).

Table 5. Variance Inflation Factors (VIF) of the final stepwise model

Coefficient	Class 1	Class 2	S.E Class 1	S.E Class 2
Intercept	1.6964	6.5202	1.6975	1.3564
Corruption	-0.0454	-0.3190	0.2622	0.1990
Malfunction systems	-0.3824	| -0.8155	0.3141	0.2527

Many studies empirically prove that there is a positive impact of e-government on corruption when the impact is mediated by ICT development and institutional quality, [21].

As shown in the Table 5 the coefficient for `Corruption` has a negative sign for both classes, indicating that as corruption increases, the log odds of reporting a negative experience (Class 0) versus the reference category decrease. The effect is more pronounced in Class 2, suggesting that greater corruption is associated with lower odds of reporting a positive experience. The same is observed for the `System Malfunction` which indicates that this variable negatively impacts the likelihood of reporting both negative (Class 0) and neutral (Class 1) experiences, again with a stronger effect for Class 2. Overall, the stepwise model suggests that both `Corruption` and `System Malfunction` have a significant negative influence on the likelihood of experiencing a positive or neutral outcome with the services. This sign indicates that higher levels of corruption and system malfunctions are associated with lower odds of positive experiences.

4.2.3 Class-specific Performance

i) Cross-Validation

In our research, we conducted a 10-fold cross-validation on a logistic regression model to test its performance using a predictive approach. However, stratified cross-validation is not considered in this process. Although the standard cross-validation we applied is designed to randomly divide the data into 10 folds, with each fold used once as a test set, it does not explicitly ensure that the sentiment classes are equally represented in every fold.

We would like to point out that the application of stratified cross-validation might strengthen this evaluation by preserving the class distribution proportion in each fold, especially when the dataset suffers from imbalance. Although our results recorded an accuracy of 83.63% using cross-validation as we expand the dataset in future studies, we plan to incorporate stratified cross-validation to enhance the reliability of our results.

Summarizing, to ensure the robustness of our selected model, we performed 10-fold cross-validation using the `cv.glm` function in R. The statistical software and programming language R (version 4.4.0) and its libraries were used for the analysis. The model exhibited an accuracy of 83.63% with a Kappa statistic of 0.0451, indicating slight agreement beyond chance between predicted and actual outcomes.

ii) Prediction and Confusion Matrix

The confusion matrix derived from the predictions of our multinomial model showed the distribution of predicted categories against the actual categories (Table 6). The overall accuracy was determined to be 83.63% with a 95% confidence interval of (0.7815, 0.882). However, the sensitivity analysis revealed that the model performed well in identifying category '2' (sensitivity = 1.000), while it struggled with categories '0' and '1' (sensitivity = 0.0526 and 0.0000, respectively).

Table 6. Prediction and Confusion Matrix

Predicted / Actual	**0**	**1**	**2**
0	0	0	0
1	0	0	0
2	19	19	188

The overall statistics by class showed that:

For Classes 0 and 1, the model demonstrates significant shortcomings, as evidenced by a sensitivity close to 0 for both classes. This indicates a failure of the stepwise model in detection for these two classes. The negative predictive value of 0.916 for both classes indicates that there is a 91.6% chance that these predictions are correct when they do not belong at Class 0 or Class 1. We should emphasize that the representation of these two classes in our data is 8.3% respectively so these results were somehow expectable.

In contrast, Class 2 shows a strong performance with a sensitivity of 1, indicating that all actual instances of Class 2 were accurately identified. This class represents 83.2% of the total data, highlighting its dominance in the dataset. The positive predictive value suggests that 83.2% of the predictions made for Class 2 are correct, reflecting the model's effectiveness in recognizing this class. Despite the high detection rate for Class 2, the model's inability to correctly identify Classes 0 and 1 point to a need for further refinement to achieve balanced

performance across all classes. Increasing the number of surveys may support and increase the prediction rate for Class 0 and Class 1.4.2.4 Predicted Probabilities of User Sentiment by Ease of Use

A bar plot was generated to visualize the predicted categories' proportions, illustrating each category's response percentage (Figure 4). The plot highlighted that category '2' accounted for 83.2% of the predictions, which reflects a skew in the data distribution or model performance. This data is explained by the fact that the use of e-Albania platform is mandatory, so people use it since it is a platform that the government has incorporated into every government service.

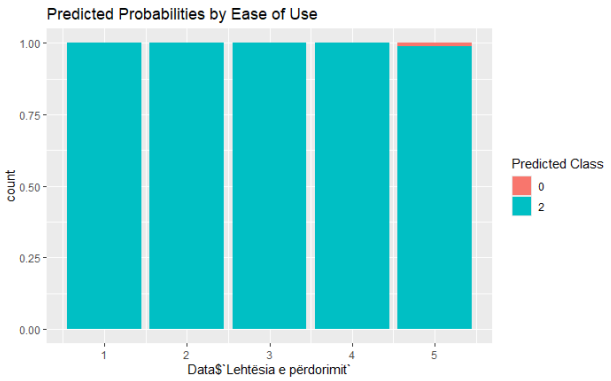

Fig. 4: Predicted Probabilities User Sentiment by Ease of Use

Figure 4 shows the predicted probabilities in the highest ease of use category (Category 5) not reaching 100%. This indicates that even users who rate ease of use highly do not always report positive sentiment, suggesting the influence of other factors. Ease of use remains a significant predictor of sentiment, as shown in the regression results, but the variability in sentiment within higher categories highlights its complex relationship.

5 Limitations of the Study

Class Imbalance in Data:
The dataset is heavily skewed, with 83.2% of responses falling into the "positive" category, making it challenging to predict "negative" and "neutral" experiences accurately.

The sensitivity for these underrepresented classes is very low, which is an indicator that the model has a hard time covering all sentiment classes evenly.

Lack of Stratified Cross-Validation:
Although 10-fold cross-validation was applied, it was not stratified. This may have affected the reliability of the model's performance evaluation.

Simplification of Predictors:
While stepwise selection narrowed the predictors to corruption and system malfunctions, other relevant factors such as ease of use and time reduction may still play a role and require further investigation.

Context-Specific Findings:
The results are specific to Albanian youth and the e-Albania platform. Generalizing these findings to other countries or populations should be done with caution.

Self-Reported Data:
The reliance on survey responses introduces the possibility of bias, as participants may overestimate or underestimate their experiences or the factors influencing them.

6 Conclusion

Many governments are already well on the way to improving digital service delivery, but there is often a gap between rhetoric and reality. There is no shortage of government policies and strategies relating to "digital first," "e-government," and "gov2.0," in addition to digital by default. But governments need more than a strategy. "Going digital" requires leadership at the highest levels, investments in skills and human capital, and cultural and behavioral change, [22]. Therefore, in our research, we considered the data gathered by a survey of young people where a sample of 250 was obtained after the initial phase of data cleaning and pre-processing. In order to gain more knowledge of user perception we constructed a logistic regression model. In our model we considered some of the factors that based on statistics show a higher effect on user perception. After a step-by-step regression model, we ended up with only 2 among 10 predictors that showed a high effect in user experience and their perceptions. Implementing a multinomial logistic regression model in this study has provided valuable insights into E-Albania platform user experiences with various services, particularly concerning the influence of several key factors such as corruption, system malfunctions, and ease of use. By analyzing the relationships between these predictors and user sentiment categorized as negative, neutral, or positive, we have identified critical determinants that significantly impact how users perceive and engage with services.

The stepwise model suggests that both `Corruption` and `Malfunction systems` have a significant negative influence on the likelihood of experiencing a positive or neutral outcome with the services. The negative coefficients indicate that higher levels of corruption and system malfunctions are associated with lower odds of positive experiences. These results could inform policymakers and service providers about areas needing improvement to enhance user satisfaction. The model's relatively low AIC suggests it is a reasonable fit for the data.

The findings from this study underscore the utility of stepwise selection in model optimization for multinomial outcomes. While the final model demonstrates strong overall accuracy, it also reveals challenges in adequately predicting certain categories. These variables show a strong association with dissatisfaction, highlighting the need for targeted reforms to improve user trust and the effectiveness of e-government platforms like e-Albania. Similar to studies in other contexts (e.g., African countries), corruption emerged as a significant factor undermining trust in e-government services. Previous research has established that reducing corruption through digitization improves transparency and public trust in institutions, aligning with this study's findings.

For instance, previous studies highlighted the mediating role of ICT development in reducing corruption, which parallels the observed benefits of Albania's digitization efforts.

The impact of technical issues is consistent with global studies that emphasize the importance of robust ICT infrastructure for successful e-government adoption. In Malawi, for example, technical challenges similarly hindered the effectiveness of digital platforms.

The OECD (2021) findings on youth trust in public institutions resonate with this study's results, where younger users view e-government positively but express concerns over transparency and system reliability. This shared perspective highlights the importance of addressing institutional and technical barriers to sustain youth engagement.

Future research should consider incorporating additional variables or exploring alternative modeling approaches to enhance prediction accuracy for all response categories. Also, explore longitudinal data to assess how these relationships evolve, as well as expand the scope of predictors to include additional variables that may influence user sentiment. This comprehensive understanding can inform strategic decisions and drive improvements in service design and implementation.

7 Practical Implications

Addressing corruption and system malfunctions identified as significant predictors of dissatisfaction can lead to enhanced user experiences and greater trust in e-government platforms.

Reducing system malfunctions could involve investing in ICT infrastructure and ensuring robust system maintenance.

Policy Recommendations: Policymakers should prioritize addressing corruption and system malfunctions to enhance user satisfaction and trust. They should focus on transparency initiatives allocate resources to build a reliable and user-friendly digital ecosystem.

Enhancing digital literacy and equitable access to technology are critical for ensuring all citizens benefit from digital transformation.

Service Optimization:

Investments in ICT infrastructure and technical support systems are critical to improving service reliability. The study highlights the importance of integrating user feedback into the design and implementation of e-government services to address specific pain points effectively.

Future Research Directions: Expanding the dataset to include a broader demographic and increasing the representation of "neutral" and "negative" experiences will provide more balanced insights.

Incorporating additional variables, such as user demographics and psychological factors, can offer a more comprehensive understanding of user sentiment.

Acknowledgement:
My sincere gratitude to the generous support of National Agency for Scientific Research and Innovation (NASRI) in Albania, which enabled us to carry out this study. The financial support received for our project "Ndikimi i qeverisjes elektronike në përmirësimin e shërbimeve në Shqipëri", based on Decision No. 10, date 21.07.2023, on the approval of the financing of winning projects of the National Research and Development Program for the Period 2023-2024", is responsible for the significant success of the study.

Declaration of Generative AI and AI-assisted Technologies in the Writing Process

The authors wrote, reviewed and edited the content as needed and they have not utilised artificial intelligence (AI) tools. The authors take full responsibility for the content of the publication.

References:

[1] Bieber, F., Vizek, M., & Gligorov, V. (2018). *The impact of digital transformation on the Western Balkan – Tackling the challenges towards political stability and economic prosperity*, [Online]. https://www.eizg.hr/userdocsimages/vijesti/vijesti_dogadaji/dt_studija/wb6-policy-paper.pdf (Accessed Date: September 28, 2024).

[2] Chungu, M. F. (2024). E-Government Implementation and Service Delivery in Rural Malawi. i-manager's *Journal on Management*, 18(4),31-35. https://doi.org/10.26634/jmgt.18.4.20678.

[3] OECD (2021), The E-Leaders Handbook on the Governance of Digital Government, OECD Digital Government Studies, *OECD Publishing*, Paris, https://doi.org/10.1787/ac7f2531-en.

[4] Albania-2024 Digital Public Administration Factsheet, EU Commission, July 2024. Interoperability and Data, [Online]. https://interoperable-europe.ec.europa.eu/sites/default/files/inline-files/NIFO_2024%20Supporting%20Document_Albania_vFinal_0.pdf (Accessed Date: September 28, 2024).

[5] OECD (2024), "2023 OECD Digital Government Index: Results and key findings", OECD Public Governance Policy Papers, No. 44, OECD Publishing, Paris, https://doi.org/10.1787/1a89ed5e-en.

[6] Al-Moalla, A., and Li, D. "Organisational Issues with Electronic Government Procurement: A case study of the UAE.", *The Electronic Journal of Information Systems in Developing Countries,* vol. 41, Issue 1, pp. 1-18. https://doi.org/10.1002/j.1681-4835.2010.tb00292.x.

[7] Al-Khouri, A. M. "E-Government Strategies: The case of the United Arab Emirates (UAE)", *European Journal of e Practice,* pp. 126-154, 2012, [Online]. https://interoperable-europe.ec.europa.eu/sites/default/files/document/2014-06/ePractice%20Journal-Vol.%2017-September%202012.pdf (Accessed Date: November 10, 2024).

[8] WSIS Action Lines Supporting The Implementation of the SDGs. United Nations, 2016, [Online]. https://www.itu.int/net4/wsis/forum/2016/Content/documents/outcomes/WSISForum2016%E2%80%94WSISActionLinesSupportingImplementationSDGs.pdf (Accessed Date: April 28, 2024).

[9] Rachinger, M., Rauter, R., Müller, C., Vorraber, W. and Schirgi, E. (2019), "Digitalization and its influence on business model innovation", *Journal of Manufacturing Technology Management,* Vol. 30 No. 8, pp. 1143-1160. https://doi.org/10.1108/JMTM-01-2018-0020.

[10] INSTAT, 2016, Information and Communication Technology Report in Enterprises, 2016s, [Online]. https://www.instat.gov.al/media/2975/information-and-communication-technologies-in-enterprises-2016.pdf (Accessed Date: September 24, 2024).

[11] Annual Report 2022. Bank of Albania, [Online]. https://www.bankofalbania.org/rc/doc/RV_2022_ENG_FINAL_24634.pdf (Accessed Date: April 28, 2024).

[12] Survey on Information and Communication Technologies (ICT) Usage in Households and by Individuals in 2023, [Online]. https://www.instat.gov.al/media/12934/ict-2023__.pdf (Accessed Date: April 28, 2024).

[13] Anila Gjonaj, Ermelinda Shehaj E-Government in Albania. *Academic Journal of Interdisciplinary Studies,* Vol. 5 No 3S1, 456-459, December 2016. doi: 10.5901/ajis.2016.v5n3s1p456.

[14] UN E-Government Knowledgebase, View Country data, Albania, [Online]. https://publicadministration.un.org/egovkb/en-us/Data/Country-Information/id/2-Albania (Accessed Date: October 28, 2024).

[15] Taulant Kullolli, Brunela Trebicka, Shkelqim Fortuzi, Understanding Customer Satisfaction Factors. A Logistic Regression Analysis, *Journal of Educational and Social Research,* Vol 14 No 2, March 2024, pp. 218-231. doi: 10.36941/jesr-2024-0038.

[16] Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). Applied logistic regression. John Wiley & Sons. doi: 10.1002/9781118548387.

[17] Agresti, A. *Categorical Data Analysis (3rd ed.*). John Wiley & Sons. 2013. ISBN: 978-0-470-46363-5.

[18] Pandey, P. &. (2015). Research Methodology: Tools and Techniques (Vol. 1). *Romania: Bridge Center.Technol. Manag.* 30 (8), 1143–1160, [Online]. https://euacademic.org/bookupload/9.pdf (Accessed Date: October 28, 2024).

[19] Ng Chirk Jenn. Design a Questionnaire. *Malays Fam. Physician.* 2006 Apr 30; 1(1):32-5. eCollection 2006.

[20] Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method. John Wiley & Sons. doi: 10.1002/9781394260645.

[21] I.O. Adam, Examining E-Government development effects on corruption in Africa: the mediating effects of ICT development and institutional quality, *Technol. Soc.,* 61 (2020) 101245. https://doi.org/10.1016/j.techsoc.2020.101245

[22] Digital government: turning rhetoric into reality / M. Carrasco, P. Goss, BCG perspectives, 2014, [Online]. https://web-assets.bcg.com/img-src/Digital_Government_Jun_2014_tcm9-63670.pdf (Accessed Date: October 28, 2024).

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

- Eralda Xhafka carried out the conceptualization and the methodology of the paper.
- Jonida Teta has organized and executed the formal analysis.
- Eralda Gjika was responsible for the Statistics and data curation.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

Funded by the National Agency for Scientific Research and Innovation (NASRI) in Albania.

Conflict of Interest

The authors have no conflicts of interest to declare.