RAQAMLI IQTISODIYOT VA AXBOROT TEXNOLOGIYALARI
2025
№1
yanvar-mart
http://dgeconomy.tsue.uz/
ILMIY ELEKTRON JURNAL

DEVELOPMENT TRENDS OF ELECTRONIC COMMERCE AND ITS INFRASTRUCTURE IN UZBEKISTAN

Karimova Shirin Zoxid qizi
Department of digital economy, TSUE, PhD senior teacher
karimovashirin22@gmail.com

Sodiqova Dilafruz Zoxid qizi
Doctorate student, UWED, Manager at TSUE
gofurova.dilafruz.1999@gmail.com

Annotation
This article is dedicated to studying the main trends in the development of electronic commerce and its infrastructure in Uzbekistan. It discusses the reforms being implemented in the field of electronic commerce, as well as the digital technologies being introduced and the factors that contribute to these changes. The article analyzes the impact of electronic commerce on the country's economy, projects aimed at developing payment systems, internet and mobile communications, and the practical results of their implementation. Key areas for infrastructure development are described, and future opportunities in this field are considered.

Key words
ARIMA model, forecasting, SPSS Statistics software package, infrastructure, Uzbekistan, digital economy, online trade, mobile technologies, internet technologies, payment systems, financial technologies, electronic payments.

Introduction

Currently, electronic commerce has become an integral part of the global economy, reshaping trade relations between countries. The development of e-commerce in Uzbekistan has also had a significant impact on the national economy. The measures undertaken by the country's leadership, aimed at promoting digital transformation and the development of electronic services, play a crucial role in stimulating economic growth, creating jobs, and offering consumers more convenient services.

Uzbekistan's e-commerce infrastructure is developing rapidly. The modernization of mobile technologies and internet connections, the improvement of payment systems, and the increasing number of online marketplaces are some of the main changes in this sector. Additionally, the expansion of internet and mobile communication infrastructure, along with the need to introduce new innovative solutions in the transition to a digital economy, is of great importance.

Literature Review

In the study titled "Forecast Analysis Using ARIMA Models During the COVID-19 Pandemic," Gondim applied ARIMA models to predict COVID-19 cases in Brazil. This research highlights the effectiveness of ARIMA models in producing crucial short-term forecasts that support public health strategies during the pandemic.

Zhang (2020) – In this study, the authors compare time series forecasting between ARIMA and machine learning models. They discuss the efficiency of ARIMA across various applications and propose a hybrid MDPI model.

Kumar et al. (2021) – This paper explores the use of ARIMA in forecasting electricity consumption in India. The study demonstrates ARIMA's reliability in data management. These studies showcase ARIMA's wide range of applications in fields like public health, energy forecasting, and network traffic management. By delving deeper into these papers, you can gain a broader understanding of ARIMA's usage and outcomes.

Research Methodology

This article utilizes the ARIMA (Auto-Regressive Integrated Moving Average) model to analyze the trends in the development of e-commerce and its infrastructure in Uzbekistan, as well as to forecast parameters up to the year 2030. The research process consists of the following stages:

Data Collection: Data related to e-commerce and infrastructure is collected. This process relies on information from government statistics, financial and economic organizations, as well as international bodies (e.g., World Bank, Uzbekistan State Statistics Committee). The data must be in the form of time series, such as e-commerce volume, the number of users, and payment system statistics.

Time Series Analysis: To select the parameters of the ARIMA model, time series data is analyzed using statistical methods such as autocorrelation and partial autocorrelation functions. The model's parameters—p (auto-regression), d (differencing), and q (moving average)—are determined.

Forecasting with ARIMA: The validated ARIMA model is then used to forecast the development of e-commerce and its infrastructure in Uzbekistan until 2030. During this process, the forecast results are compared with other scientific studies and practices.

Analysis and Recommendations: The forecasts are analyzed, and based on them, recommendations for the development of e-commerce and its infrastructure are provided. Additionally, strategic planning insights are drawn from the data for practical application.

Analysis and Results

According to analytical studies conducted by the Center for Digital Economy Research, in recent years, e-commerce has become one of the key elements of the digital economy in Uzbekistan. Both the volume of e-commerce and retail sales have shown steady growth. Specifically, in 2022, the volume of e-commerce transactions increased by 1.8 times compared to 2021, exceeding 10.8868 trillion so'm. This figure accounts for more than 4% of the total retail sales volume.

Table 1

Trends in the Development of E-commerce and Its Infrastructure in Uzbekistan

Period	Volume of Goods Sold in E-commerce (million soums)	Number of Information and Communication Enterprises (units)	Number of Internet Subscribers by the End of the Year (thousand units)	Number of Computers Connected to the Internet in Enterprises and Organizations by the End of the Year (units)
2010 y.	3834,4	206	4 111,0	97 811
2011 y.	5658,4	932	4 895,0	127 600
2012 y.	7482,4	1658	5 641,0	154 741
2013 y.	6651,1	2384	6 541,0	188 411
2014 y.	8981,1	4110	7 814,0	201 541
2015 y.	9962,0	4836	8 339,1	223 907
2016 y.	6033,1	6370	9 626,8	271 357
2017 y.	12123,7	6427	11 168,0	310 459
2018 y.	40861,1	6403	13 321,7	358 003
2019 y.	273310,7	6975	16 386,2	413 417
2020 y.	1002481,0	7901	19 981,0	441 913
2021 y.	5978660,2	9517	22 987,2	538 933
2022 y.	10886800,0	10587	26 723,6	667 842

There is data on the factors influencing the development of e-commerce in Uzbekistan. Identifying the factors that affect the development of the e-commerce system, analyzing the relationships among them, and using this analysis to forecast future trends are essential for advancing e-commerce.

In Uzbekistan, 2,271 new enterprises were established in the information and communication sector from January to September 2023. This figure represents a 1.2-fold increase compared to the same period in 2022. Over the last five years, the number of such enterprises has nearly doubled. As of the end of 2022, there were 11,413 enterprises operating in the information and communication sector, an increase of 6.7% compared to the previous year.

As of January 1, 2023, the number of internet subscribers in the country reached 26.7 million. In 2022, the number of internet subscribers in Uzbekistan was about 26 million. Between 2010 and 2022, the average annual growth rate of internet subscribers was approximately 16%. Accordingly, the overall bandwidth capacity for international internet connections has shown a growing trend over the years to support the development of international and intercity telecommunication networks. For instance, by 2023, the internet speed had increased by 2.6 times compared to 2020, reaching 3,200 Gbit/second.

In 2022, Uzbekistan ranked 155th in the global ranking for wired internet speed. However, by 2023, the ranking improved to 131st, marking a rise of 24 positions compared to the previous year. The recorded average wired internet speed in Uzbekistan is 21.67 Mbit/s.

To forecast promising parameters for the development of e-commerce in Uzbekistan, the use of ARIMA models is appropriate.

ARIMA (AutoRegressive Integrated Moving Average) is widely used in statistics to forecast and analyze stationary time series. ARIMA models are applied to analyze data over a certain period. To develop these models, the parameters of the models, such as autoregression (AR), integration (I), and moving average (MA), must be determined.

The ARIMA method is a statistical approach used to forecast time series that exhibit certain trends, cycles, and seasonal components. It consists of the following key components:

AR (AutoRegressive): This part of the model describes how the current value of the time series is dependent on its previous values. The parameter "p" defines the number of lags considered in this section.

I (Integrated): This component indicates how many times the time series must be differenced to make it stationary. Stationarity means that the statistical characteristics of the series, such as the mean and variance, remain constant over time. The parameter "d" specifies the number of differences needed.

MA (Moving Average): This component accounts for the influence of previous forecast errors on the current value of the time series. The parameter "q" determines the number of lags considered in this section.

To use ARIMA for forecasting time series, the following steps must be followed:

Data preparation, which may involve differencing the series if it is non-stationary (when "d" is large).

Identifying the optimal values of the ARIMA model's "p", "d", and "q" parameters, which may require analyzing the autocorrelation function (ACF) and partial autocorrelation function (PACF).

Table 2

Description of the forecasting model

			Types of models
ID model	Volume of sales of goods in electronic commerce, mln.sum	Model_1	Brown
	Number of information and communication enterprises, unit	Model_2	ARIMA(0,1,0)
	The number of subscribers connected to the Internet, by the end of the year; thousand units	Model_3	ARIMA(0,2,0)
	Number of computers connected to the Internet in enterprises and organizations, by the end of the year, unit	Model_4	Brown

Training the ARIMA model on historical data with the selected parameter values.

Using the trained model to forecast future values of the time series.

ARIMA remains a specialized model for time series forecasting, but more advanced and complex methods, such as machine learning models, may be required for more detailed data and greater accuracy. However, ARIMA remains a useful tool for analyzing and forecasting time series.

In the context of time series analysis within this research, the concepts of 'Brown' (including Brownian motion or random walk) and the ARIMA(0,1,0) and ARIMA(0,2,0) models were used. (Table 2).

Brown (Random Walk): Brownian motion, also known as a random walk, is a mathematical model that describes the random movement of particles in a fluid. From the perspective of time series analysis and finance, it refers to a model where future values are simply a sum of past values and a random error term. In discrete-time validated random walks, each step is determined by adding a random value (typically drawn from a normal distribution) to the previous value. The formula for a random walk is: Y(t)=Y(t−1)+ϵ(t)Y(t) = Y(t-1) + \epsilon(t)Y(t)=Y(t−1)+ϵ(t), where Y(t)Y(t)Y(t) is the value at time ttt, Y(t−1)Y(t-1)Y(t−1) is the previous value, and ϵ(t)\epsilon(t)ϵ(t) is the random error amount. Random walks are typically used to describe processes that have no clear trend or pattern and to model processes where future values cannot be predicted.

ARIMA(0,1,0) Model: ARIMA stands for AutoRegressive Integrated Moving Average, and the notation ARIMA(p,d,q) represents the order of the autoregressive (AR), integrated (I), and moving average (MA) components. In ARIMA(0,1,0), the “0” in the autoregressive (AR) and moving average (MA) components indicates the absence of autoregressive or moving average terms. The "1" in the integrated (I) component indicates that first-order differencing is applied to make the time series stationary. Simply put, ARIMA(0,1,0) is often used to model time series that require first-order differencing to remove trends or make the data stationary. This means that future values depend on past values rather than past errors or shocks.

ARIMA(0,2,0) Model: Similar to ARIMA(0,1,0), but it includes second-order differencing. This indicates that the time series must be differenced twice to remove trends and make the data stationary. It is typically used when there is a need to remove linear trends and first-order autocorrelation structures in time series data. The notation ARIMA(0,2,0) implies that future values depend on past values after removing linear trends and first-order autocorrelation through differencing. It is generally used for cases where linear trends and first-order differencing are applied.

In conclusion, Brownian motion or random walk represents a stochastic process that is determined solely by adding past values and random error amounts to predict future values. ARIMA(0,1,0) and ARIMA(0,2,0) are time series models used to analyze and retrieve data that require differencing to make the series stationary. These models should be useful for identifying and analyzing data.

To create the models, it is necessary to determine the parameters of the models when analyzing data over a specific period. For this analysis, the values for all variables are chosen based on the SPSS Statistics software, where the forecasting models for each selected variable are determined in the most appropriate way using the least squares method.

Table 3

Modelning mos sifati

Fit statistics	Mean	SE Standard Error	Minimum	Maximum	5 percent quantile value
Stationary R-squared	-,066	,076	-,136	3,331E-16	-,136
R-squared	,949	,061	,859	,994	,859
KCKO	315647,166	613655,364	542,617	1235960,728	542,617
COMO	13,924	16,705	3,569	38,759	3,569
MOMO	32,089	33,941	12,189	82,638	12,189
CMO	101692,403	191830,097	420,050	389206,727	420,050
MMO	1079198,775	2112158,925	889,083	4247084,483	889,083
Normalized BIC	18,601	7,378	12,811	28,252	12,811

When analyzing the table, each metric is used to assess the quality and adequacy of the model. (Table 3)

Stationary R-squared: In the context of time series analysis, this statistical metric is used to evaluate how well the forecast model fits stationary time series data. The average value is -0.066, indicating that the model does not fit the stationary data well. The values range from -0.136 to 3.331E-16, showing significant variations in the model's applicability. The 5th percentile value of -0.136 indicates that the model may not fit well in some cases.

R-squared: This metric assesses how well the model fits the reviewed data. The average value is 0.949, demonstrating how well the model generally fits the data, with the R-squared value being close to 1. Values range from 0.859 to 0.994, indicating a high level of accuracy in the model's fit.

Mean Squared Error (MSE): This measures the squared difference between the model's predicted and actual values. An average value of 315647.166 indicates that the model has a high average error, with a standard error of 613655.364 showing significant variation. The 5th percentile value of 542.617 indicates some instances of low error.

Mean Absolute Deviation (MAD): This metric measures the average absolute difference between the predicted and actual values. The average value of 13.924 represents the average error amount. The standard error of 16.705 indicates significant variation in the absolute error amounts, with the 5th percentile value of 3.569 showing instances of very low error.

Maximum Absolute Deviation (MaxAD): This metric measures the highest value of the average absolute difference. An average value of 32.089 indicates the average maximum absolute error. The standard error of 33.941 shows strong variation in the maximum error amounts.

R-squared, also known as the coefficient of determination, is a general statistic used to assess how well a statistical model (including forecasting models) explains the variability of observed data. The R-squared value typically ranges from 0 to 1, where 0 indicates that the model explains none of the variability and 1 indicates that the model explains all variability. In other words, a higher R-squared value indicates a better fit of the model to the data. R-squared values fall within the range of 0 to 1, with higher values indicating that the model fits the data well.

In conclusion, Stationary R-squared is a metric used in the context of time series analysis to evaluate how well the stationary portion of the time series fits the model. A high Stationary R-squared indicates how well it fits the stationary data, which is very important for accurately and efficiently conducting time series forecasting.

Calculating R-squared:

R-squared is used to calculate the explained variance of the response variable by traditional variables in the productive parameter. It ranges from 0 to 1.

The formula used to calculate R-squared is: $R^2 = 1 - (SSR / SST)$ where:

SSR (Sum of Squared Residuals) indicates the unexplained variation (the variability that the model does not explain).

SST (Total Sum of Squares) indicates the total variability in the productive parameter.

Table 4

Values obtained by checking the quality of the model

Fit Statistics	Percentile				
	10	25	50	75	90
Stationary R-squared	-,136	-,134	-,064	8,327E-17	3,331E-16
R-square	0,859	0,886	0,972	0,990	0,994
KCKO	542,617	546,054	13042,660	933352,785	1235960,728
COMO	3,569	3,837	6,684	31,251	38,759
MOMO	12,189	12,287	16,764	67,215	82,638
CMO	420,050	421,145	8571,418	296084,647	389206,727
MMO	889,083	920,926	34410,768	3202264,633	4247084,483
Normalized BIC	12,811	12,821	16,671	26,312	28,252

Table 4 presents statistical indicators for evaluating the data and provides values across several percentiles. Let’s analyze the main statistics: Percentiles (10, 25, 50, 75, 90). These percentiles represent different portions of your dataset, where the 10th percentile indicates the lower cut and the 90th percentile indicates the upper cut.

Stationary R-squared: This measures how well the model fits the stable portion of the time series. It varies from -0.136 (10th percentile) to 3.331E-16 (90th percentile), indicating a certain upward trend. Lower values indicate poor fitting to the stable portion of the data, while higher values indicate a good fit.

R-squared: This measures how well the model explains the data. It ranges from 0.859 (10th percentile) to 0.994 (90th percentile). Higher R-squared values indicate that the model explains the data well, with a value of 1 representing ideal fitting.

Mean Squared Error (MSE): This quantifies the unexplained error or variability by the model. Its lowest point varies from 542.617 (10th percentile) to 1,235,960.728 (75th percentile), indicating increasing growth rates. Higher MSE values indicate a large amount of unexplained error.

Mean Absolute Deviation (MAD): This measures how well the model fits the data. It ranges from 3.569 (10th percentile) to 38.759 (75th percentile). Smaller MAD values indicate a good fit.

Maximum Absolute Deviation (MaxAD): This measures how well the model fits the average values of the data. It varies from 12.189 (10th percentile) to 82.638 (75th percentile).

Sum of Errors (SE): This measures how well the model fits the center of the data. It ranges from 420.050 (10th percentile) to 389,206.727 (75th percentile).

Maximum of Errors (ME): This measures how well the model fits the center of the data. It ranges from 889.083 (10th percentile) to 4,247,084.483 (75th percentile).

Normalized BIC: This serves as a criterion for model selection and varies from 12.811 (10th percentile) to 28.252 (75th percentile). Smaller values indicate a better model fit.

Conclusions

Overall, these percentiles provide guidance on how well the data fits different types, helping to evaluate the goodness of fit of the data. Higher R-squared values and lower error values typically indicate a good model fit.

In e-commerce, the volume of goods sold is projected to increase from 15,794,950.08 million so’m in 2023 to 50,152,000.63 million so’m by 2030. The upper confidence limit (UCL) for this forecast is identified to rise from 18,487,877.17 million so’m in 2023 to 88,609,873.33 million so’m in 2030. The lower confidence limit (LCL) is expected to decrease from 13,102,022.99 million so’m in 2023 to 11,694,127.93 million so’m by 2030.

The number of information and communication enterprises is also projected to increase, with forecast values indicating a rise from 11,452.08 to 17,507.66 units from 2023 to 2030. The upper confidence limit (UCL) is projected to increase from 12,676.63 units in 2023 to 20,971.21 units in 2030. The lower confidence limit (LCL) is expected to rise from 10,227.53 units in 2023 to 14,044.12 units by 2030.

Table 5

The predicted values of the volume of sales in e-commerce, the number of information and communication enterprises, the number of subscribers connected to the Internet and the number of computers connected to the Internet in enterprises during the period until 2023-2030.

Model		2023	2024	2025	2026	2027	2028	2029	2030
The volume of sales of goods in electronic commerce, mln.sum-Model_1	Projected values	15794950,1	20703100,2	25611250,2	30519400,3	35427550,4	40335700,5	45243850,6	50152000,6
	UCL	18487877,2	26724301,7	35686384,7	45267673,3	55396622,3	66021581,0	77103115,1	88609873,3
	LCL	13102023,0	14681898,6	15536115,8	15771127,3	15458478,5	14649819,9	13384586,0	11694127,9
Number of information and communication enterprises, unit-Model_2	Projected values	11452,1	12317,2	13182,3	14047,3	14912,4	15777,5	16642,6	17507,7
	UCL	12676,6	14048,9	15303,2	16496,4	17650,6	18777,0	19882,4	20971,2
	LCL	10227,5	10585,4	11061,3	11598,2	12174,2	12778,0	13402,7	14044,1
The number of subscribers connected to the Internet, by the end of the year; thousand units-Model_3	Projected values	30729,5	35003,4	39545,7	44356,5	49435,8	54783,5	60399,7	66284,4
	UCL	31938,5	37706,8	44069,5	50978,7	58402,2	66316,9	74705,1	83552,7
	LCL	29520,4	32299,9	35022,0	37734,4	40469,4	43250,2	46094,3	49016,0
The number of computers connected to the Internet in enterprises and organizations, by the end of the year, unit-Model 4	Projected values	795857,3	923878,7	1051900,2	1179921,6	1307943,0	1435964,4	1563985,8	1692007,2
	UCL	851480,1	1046920,2	1256836,6	1479145,7	1712429,6	1955660,5	2208051,5	2468976,6
	LCL	740234,5	800837,3	846963,7	880697,5	903456,4	916268,4	919920,1	915037,9

Based on year-end data, the number of internet-connected subscribers is projected to rise from 30,729.454 million units in 2023 to 66,284.36 million units by 2030. The upper confidence limit (UCL) is anticipated to increase from 31,938.48 million units in 2023 to 83,552.72 million units by 2030. The lower confidence limit (LCL) is projected to rise from 29,520.42 million units in 2023 to 49,016.00 million units by 2030.

The forecasted values for the number of computers connected to the internet in enterprises and organizations are projected to increase from 795,857.32 units in 2023 to 1.69E+6 units by 2030. The upper confidence limit (UCL) is expected to rise from 851,480.14 units in 2023 to 2,468,976.5 units by 2030. The lower confidence limit (LCL) is estimated to fluctuate around approximately 900,000 units from 2026 to 2030.

All models predict stable growth in their respective indicators over the years. The upper and lower confidence limits indicate the range of these forecasts and define the level of uncertainty in the projections. An expansion of the UCL and LCL range signifies significant uncertainty. The forecasted values generally indicate an upward trend, reflecting growth or expansion in the areas of e-commerce sales, information and communication enterprises, internet-connected subscribers, and computers connected to the internet in enterprises and organizations. The decrease in the LCL of the first model and the variability of the LCL in the fourth model may indicate potential variability or uncertainty in these fields.

List of used literature

1. Abdullaeva, S., & Tashkentov, B. (2020). "Digital economy and e-commerce: A study of Uzbekistan's infrastructure." The study of Uzbekistan's strategic approaches to developing the digital economy and e-commerce infrastructure.
2. Mirzayev, M. (2019). "The role of mobile technology in developing e-commerce in Uzbekistan."
3. ps://www.mdpi.com/2227-7390/11/14/3069.
4. A Review of ARIMA vs. Machine Learning Approaches for Time Series Forecasting in Data-Driven Networks 2023, 15(8), 255; https://doi.org/10.3390/fi15080255
5. Xasanov, I. (2023). "E-commerce trends and forecasts for Uzbekistan by 2030." Research main trends and forecasts of e-commerce for Uzbekistan until 2030.
6. Sodikova D. ANALYSIS OF E-COMMERCE FACTORS INFLUENCING ON ECONOMIC DEVELOPMENT //Economics and Innovative Technologies. – 2023. – Т. 11. – №. 4. – С. 349-360.\
7. Содикова Д. ПУТИ И ПРОБЛЕМЫ ИСПОЛЬЗОВАНИЯ ОПЫТА РАЗВИВАЮЩИХСЯ СТРАН В РАЗВИТИИ ЭЛЕКТРОННОЙ КОММЕРЦИИ //Экономическое развитие и анализ. – 2023. – Т. 1. – №. 3. – С. 49-53.

8. Ilxomovna X. B., Ogli G. R. C., Qizi K. S. Z. E-Commerce in a modern business system //ACADEMICIA: An International Multidisciplinary Research Journal. – 2021. – Т. 11. – №. 4. – С. 506-510.
9. Karimova S. ELEKTRON TIJORAT PLATFORMALARINI TAKOMILLASHTIRISHDA VIRTUAL EKOTIZIMLARNING O ‘RNI //Raqamli iqtisodiyot va axborot texnologiyalari. – 2024. – Т. 4. – №. 4. – С. 26-33.