# Time Collection Prediction Techniques of Sciences in Education Using ARIMA Model Instance from Technical Establishments Case of RP-KIGALI-College

# Mukarugwiro Placidie<sup>1</sup>; Nizeyimana Viateur<sup>2</sup>; Haguminshuti Sylivestre<sup>3</sup>; Gasana Jean Claude<sup>4</sup>; Adrien Nzumvirumukiza<sup>5</sup>; Munezero Joseph<sup>6</sup>

1;2;3;4;5;6 Assistant Lecturer at Rwanda Polytechnic Kigali College

Publication Date: 2025/06/09

Abstract: This study examines the sizeable ability of time series forecasting science in education zone. Through applying ARIMA, Exponential Smoothing, and Seasonal Decomposition fashions to real-world statistics on pupil enrollment, instructional overall performance, attendance, and teacher retention, the look at provides actionable insights for educators, policymakers, and directors. Many researcher research in make the studies at excessive gaining knowledge of organization (RP-Kigali college, TUMBA, MUSANZE, HUYE), they did now not display the case why college students want to fail Sciences in Technical college? what is the reality motive them to fail science? In my research I got here into technical school students want to lose relies upon on the reality that the wide variety of college students is inserted inside the technical college is half of's within the Sciences, it makes the number of the winners are rather relatively averaged whilst enters the technical faculty have a lowest know-how of the Sciences. My contribution lies in providing personalized, engaging and data driven solutions that bridge the gap between technical training and scientific learning, helping students succeed in both fields. This study come up by the result showing that: The successive three years:2024: 82%, 2025: 85%, 2026: 87%. There is sluggish improvement in common test ratings in science at PR KIGALI College High learning Institution, which shows that recent educational reforms or curriculum modifications are having a superb effect.

Keywords: Students Performance, Exponential Smoothing, ARIMA Model.

How to Cite: Mukarugwiro Placidie; Nizeyimana Viateur; Haguminshuti Sylivestre; Gasana Jean Claude; Adrien Nzumvirumukiza; Munezero Joseph (2025). Time Collection Prediction Techniques of Sciences in Education Using ARIMA Model Instance from Technical Establishments Case of RP-KIGALI-College. *International Journal of Innovative Science and Research Technology*, 10(3), 3341-3350. https://doi.org/10.38124/ijisrt/25mar1891

#### I. INTRODUCTION

A time series is a sequential set of records points, measured normally over successive instances. Time collection information may have main forms approach continuous and discrete; a non-stop time series is defined constantly in time. Way defined at each factor in time which includes Air temperature, and the speed of a car (2021). Discrete time collection is one in which the variable being examined is not defined at each factor in time, but most effective at specific time, as an example month-to-month rain fall (El, 2021).

In 2020 ARIMA (Autoregressive integrated transferring average model) changed into used by Sakira, Ma Iquebal and Cattopadhvav in Indian even as modeling and forecasting time collection statistics of college students perfomence in mathematics from 1969 to 2007, in their look at they used Root mean square, Akaike information Criterion and Bayesian Criterion to become aware of the exceptional version . As end result they determined that ARIMA (2,1,0) version executed higher amongst other models of ARIMA circle of relatives, for modeling as well as forecasting reason et al. (2019).

Fatema and Md, Israh in Bangladesh have Assessed students' performance in arithmetic in Bangladesh via ARIMA, Holt's, and ARCH-GARCH models of their evaluation the goodness of in shape changed into achieved thru the Akaike information criteria, Schwartz Bayesian criteria. As result It become seen that the Generalized Autoregressive Conditional Heteroscedastic (ARCH) model changed into advanced to the autoregressive included moving average model and Holt's linear exponential smoothing because the records was characterized by using changing imply and variance et al. (2019).

# ISSN No:-2456-2165

Irfan, M, (2023) and performed studies in which their principal objective turned into to develop a appropriate version after which forecast the education pupil's performance inside the one-of-a-kind universities of Rwanda all through the period of 2000—2008 to 2010—2019 the usage of ARCH circle of relatives fashions given that their preceding mentioned the only ARMA technique in place of the ARCH family fashions et .al (2018). As a result, on the premise of two standards AIC (Akaike facts standards) and SIC (Schwartz statistics standards), the GARCH version became selected for all universities also offered for the forecasting goal, after that they advocated that the university of Rwanda better forecast in technology in comparison to others high learning institution in Rwanda et .al (2018).

Parap at (2013) conducted studies in which their functions had been to discover a model to forecast the scholar's attendance in Hong Kong and to locate the maximum suitable forecasting length whether or not it should be in daily, weekly, month-to-month, or quarterly. they have analyzed facts the use of the ARIMA (Autoregressive included moving common) and ARMA (Autoregressive transferring common) fashions Shwetha at (2023). To pick the perfect forecasting techniques and the maximum appropriate forecasting length they taken into consideration the smallest value of AIC (Akaike information Criterion) and RMSE (Root suggest square mistakes), respectively. because the end results their observe confirmed that the ARIMA version turned into the fine version for finding the maximum healthy able forecasting period in monthly and quarterly. And ARMA model turned into the fine version for finding the most suitable forecasting length in every day and weekly Shwetha at (2023).

Hassle declaration: How can time collection forecasting strategies be carried out to are expecting the enrollment tendencies and educational performance of college students in technological know-how at RP-Kigali high learning Institution (Technical schools) so that you can improve useful resource allocation, school making plans, and normal institutional control?

#### II. RESEARCH OBJECTIVE

#### > The General Objectives:

The general objectives of this research is to evaluate student's performance in sciences at technical establishments.

# Specific Objectives:

Evaluate Key Analytical strategies:

Observe the fundamental time collection forecasting strategies, which includes ARIMA (Autoregressive incorporated transferring common), Exponential Smoothing, and Seasonal Decomposition.

Compare the suitability of each technique for modeling numerous varieties of academic statistics, including pupil enrollment, performance metrics, and attendance patterns.

### Literature Review

Time series forecasting has end up an invaluable tool in the subject of training, supplying statistics-pushed insights that may manual choice-making, resource allocation, and policy system Ekylia at (2014). This paper presents an outline of key analytical techniques utilized in time series forecasting, together with ARIMA (Autoregressive included transferring common), Exponential Smoothing, and Seasonal Decomposition. those methods are examined within the context of their utility to educational facts, together with student enrollment trends, instructional performance, attendance styles, and trainer retention Ekylia at (2014).

https://doi.org/10.38124/ijisrt/25mar1891

Via actual-international case research, the paper demonstrates how time collection forecasting can deal with crucial challenges science in education. those encompass forecasting pupil enrollment to optimize lecture room area and staffing, predicting scholar overall performance to become aware of at-chance novices, and modeling instructor retention to enhance team of workers management and decrease turnover. through looking ahead to destiny trends, academic establishments can higher put together for aid desires and put into effect well timed interventions El.at (2022).

The paper also explores the blessings of time collection forecasting, which include advanced accuracy in lengthytime period making plans, extra efficient use of sources, and the ability to proactively address problems earlier than they improve. furthermore, it discusses future guidelines for the sphere, highlighting the capacity integration of device studying strategies, using actual-time records, and the mixture of pass-area information to enhance forecasting models. those advancements promise to boom the precision of predictions and offer more dynamic, adaptive strategies for addressing evolving academic challenges El.at (2022).

A time series version approach was used in (haye, 2022) to expect student enrollment in Rwanda. Fifty-four records factors were generated from the 1961 to 2014 enrollment dataset provided by Rwanda TVET Board (RTB). Findings showed that ARIMA (0,2,2) model identified the use of AIC forecasted boom in enrollment within the next 5 years, with a sluggish lower after every year thereof. additionally, (haye, 2022) defined an aggregated enrollment prediction method via aid vector system and rule-based totally predictive models. The preliminary predictive consequences are generated by means of the SVM, that are then fed to a tool known as Cubist (have, 2022), which produces the guidelineprimarily based predictive model. consequences of the observe display that SVM and Cubist made fairly correct predictions with MAPE ranged from zero% - 11% for the SVM version and 2. thirteen% to fifteen. Fifty nine% with a median blunders of five. Five% for the Cubist version.

In (Clar, 2023), a Seasonal Autoregressive integrated shifting common (SARIMA) version was used to project enrollment of universities college students in Rwanda. The model is represented as SARIMA(p,d,q) x (P,D,Q) wherein p is the autoregressive phrases, d is the variations, q is the moving common phrases, P is the seasonal autocorrelation, D

### https://doi.org/10.38124/ijisrt/25mar1891

ISSN No:-2456-2165

is the seasonal fashion, Q is the seasonal transferring common and s is seasonal period. thru this version, seasonality was captured through the seasonal enrollment sample of undergraduate college students with the aid of semester, the use of a restrained fifteen to 20 years of records. The observe was capable of establish forecasting version using several crucial signs. The paper (Mohammed, 2024) (Mohammed, 2024) captured enrollment prediction of student candidates at a cohort stage for the universities.

The dataset used for the prediction became accrued from 2003 to 2016 via the seasonal enrollment periods of spring, summer season, and fall semesters. AIC became used to identify the best ARIMA model for the series. The SARIMA (zero,zero,zero) x (1,0,3) version turned into decided on as it has the lowest AIC. It turned into used to predict the enrollment at UNM for the year 2017, ensuing in a dependable accuracy at 80% self belief c programming language. The dataset applied for the prediction spanned from the years 2004 to 2018. Forecasting validation tools like AIC, BIC, and HQ have been employed to perceive the satisfactory prediction model expertise to be used for decision making (Ghosh, 2011).

Prediction plays a crucial function in education management because it attracts inferences applicable for choice and policy-making undertakings. The administration and the researchers have the leeway to apply prediction on various troubles with a ramification of complexity. Prediction, while used for scholar enrollment, mainly aids knowledge for enrollment trend analysis, which constitutes knowledge on the massive effect of enrollment for sales outcomes. The generated records ought to influence destiny approach and aid choices (Keyla, 2020). Scholar enrollment prediction is beneficial for training because the know-how generated should leverage on the use of top-quality selectionmaking strategies wanted for future making plans. But, the choice of the precise prediction strategies for a selected trouble is still a quest for the reason that facts mining algorithms to be used are depending on the availability of the **RTB** statistics.

On this paper, the famous ARIMA(p,d,q) set of rules that's a type of time collection evaluation version is applied since the dataset used is a univariate ancient facts of enrollees from the distinct RP institution, from years 2011-2012 to 2018-2019, received from the organization office of the Registrar. one of a kind ARIMA(p,d,q) models were tested, and the gold standard model for use for forecasting was decided on from it. The motive of this have a look at is to offer a 6-yr forecast on the wide variety of future enrollments on the abovementioned RP Kigali college, specially for the college years 2019-2020 to 2024-2025. The stop result of this have a look at can provide the following benefits to wit: improve the prevailing admission and retention rules of the university, make principal choices on the college's long term enrollment control approach, develop an annual advertising and marketing and recruitment plan, and to determine inner and outside factors affecting the drop and increase of enrollment, among others. The projected enrollment can also serve as a basis for offering extra school room homes to meet

elevated demand. generally, this look at is hoped to make a contribution to the two predominant literatures; (Keyla, 2020) on the usage of the ARIMA algorithm Upon assessment, it's been diagnosed that the excellent model to be used is ARIMA (1,0,0), as it has the minimum values. The outcomes of the observe indicated an insufficient boom in students' enrollment through the years (Lee, 2020) The look at (Lee, 2020) applied the ARIMA version to are expecting scholar admissions at the colleges of Rwanda. records from undergraduate admissions from 1962 to 2016 had been gathered and analyzed. the choice of the fine ARIMA version has been made the usage of facts criteria along with AIC, SBC, AME, RMSE, and MAPE thru splitting records into estimation length and validation length. Findings showed a chronic growth in student population yearly ( (ELSEVIER, 2017).

#### III. METHODOLOGY

#### A. Dataset Technique

The method for this observe is designed to comprehensively verify the use of time series forecasting inside the subject of training, focusing on analytical strategies, actual-international packages. The studies technique includes a combination of quantitative evaluation, case research, and literature evaluation to reap the study's objectives. beneath is an outline of the technique.

#### ➤ Facts Collection

The first step within the method is to collect applicable statistics that can be analyzed the use of time series forecasting strategies. The statistics will be sourced from RP KIGALI College high Learning Institution (e.g., curriculum planning, student Enrollment Forecasting, educational performance Prediction, useful resource making plans and Allocation, Curriculum Optimization) and publicly to be had instructional datasets. Varieties of facts: pupil Enrollment data: historical facts of student enrollments over several years, broken down by way of grade, school, or location. student overall performance statistics: examination ratings, grades, or standardized test results over multiple educational years. teacher Retention data: facts on trainer turnover, retention prices, and staffing stages. Attendance statistics: every day or monthly pupil attendance records to become aware of seasonal and cyclical patterns.

#### • Facts Resources:

Faculty administrative databases (student performance, enrollment, attendance), countrywide or nearby training information (e.g., authorities reports, country wide surveys), academic publications and former studies on training and time series forecasting.

#### ➢ Facts Preprocessing

Before making use of any forecasting techniques, the statistics will undergo preprocessing to make sure it's far clean, entire, and structured well for analysis. This includes: managing lacking facts: Imputation strategies or interpolation techniques may be used to fill missing values inside the datasets, ensuring that the time series stays non-stop and usable for forecasting. Outlier Detection: Statistical methods

# ISSN No:-2456-2165

which includes Z-ratings or the Tukey method might be used to stumble on and handle outliers which can skew the evaluation. Seasonal changes: Adjusting for seasonality (e.g., instructional yr cycles, vacations) wherein relevant, specially in attendance information or performance metrics stricken by seasonal factors. Normalization/Transformation: reworking facts to make certain stationarity, including taking log differences of non-stationary statistics (e.g., scholar performance tendencies), will be considered.

#### > Time Series Evaluation and Forecasting Fashions

As soon as the facts is preprocessed, the take a look at will observe diverse time collection forecasting fashions. these models might be decided on based totally on their applicability to the facts and the studies objectives. ARIMA (Autoregressive integrated moving common). reason: ARIMA fashions will be used for predicting traits in datasets that display non-stationarity over time, along with student enrollment or performance developments.

• Steps:

Stationarity checking out: The Augmented Dickey-Fuller (ADF) check may be carried out to test if the time collection facts is desk bound. Non-desk bound information will be differenced until it becomes desk bound. version selection: The ARIMA version's parameters (p, d, q) could be optimized the use of strategies just like the AIC (Akaike facts Criterion) and BIC (Bayesian facts Criterion). model assessment: The ARIMA model's accuracy may be assessed using forecast blunders metrics along with RMSE (Root suggest Squared errors), MAE (suggest Absolute blunders), and MAPE (mean Absolute percent mistakes).

#### IV. STATISTICS ASSESSMENT AND VERSION VALIDATION

To ensure the robustness and reliability of the time collection forecasting fashions, numerous assessment techniques may be used:

- Errors Metrics: The accuracy of the fashions might be assessed the usage of standard metrics along with RMSE, MAE, and MAPE to ensure reliable predictions.
- Cross-Validation: For more sturdy validation, ok-fold cross-validation will be used to assess version performance, specifically when handling smaller datasets.
- Again trying out: Forecasts might be examined on holdout datasets to evaluate how well the models predict unseen information.

The datasets used in this paper are the historic facts of the overall enrollment be counted from Rwanda polytechnics which include IPRC KIGALI high learning Institution, for S. Y's 2011-2012, 2012-2013,2013-2014, 2014-2015, 2015-2016, 2016-2017, 2017-2018, and 2018-2019. The datasets have been obtained from the Rwanda Polytechnic's workplace of the Registrar

#### A. ARIMA Set of Rules

The ARIMA(p,d,q) version is used in time series forecasting. The p variable denotes the autoregressive order,

https://doi.org/10.38124/ijisrt/25mar1891

# $\phi(B)(w_t - \mu) = \theta(B)a_t$

In which t, is represented as time index and backshift operator for symbol B, autoregressive parameter assigned as, for MA, for d value inside the ARIMA (p, d,q) version and for white noise (Lee, 2020).

➤ Autocorrelation Function (ACF) and Partial

# • Autocorrelation Function (PACF)

Let r (X, Y/Z) = Corr (X, Y/Z) denote the partial correlation coefficient between x and y adjusted for Z (or with z held constant) (Lee, 2020). Denote:

$$\varphi_2 = \operatorname{corr}(X_t, X_{t+2}|X_{t+1})$$

$$\varphi_3 = \operatorname{corr}(X_t, X_{t+3} | X_{t+2})$$

 $\varphi_k = \operatorname{corr}(X_t, \dots, X_{t+k} | X_{t+k-1})$ 

 $\varphi_k$  = Partial autocorrelation coefficient at lag k.

The partial auto correlation function (PACF) is

$$\{\rho_k, \rho_k, \rho_k, \dots\} = \{\rho_k, k \ge 1\}$$

The  $\varphi_k$ 's are related to the  $\rho_k$ 's by

$$\varphi_1 = corr(X_t, X_{t+1}) = \rho_1$$

Recall that 
$$r(X, Y|Z) = \frac{r(X, Y) - r(X, Z)r(Y, Z)}{\sqrt{1 - \gamma^2(X, Z)}\sqrt{1 - \gamma^2(Y, Z)}}$$

Applying this here, using  $X=X_tY = X_{t+2}, Z = X_{t+1}$ 

$$\varphi_2 = \operatorname{corr}(X_t, X_{t+2}|X_{t+1}) = r(X, Y|Z)$$
, along with  $\rho_1 = r(X, Z)$  and  $\rho_2 = r(X, Y)$  yields:

$$\varphi^{2} = \frac{\rho_{1} - \rho_{1}^{2}}{1 - \rho_{1}^{2}}$$

$$P_{k} = \frac{\sum_{t=k+1}^{r} (Y_{t} - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^{r} (Y_{t} - \bar{Y})^{2}}$$

#### ➤ Akaike Information Criterion (AIC)

In various papers where ARIMA modeling is observed, the most appropriate model to be used to forecast is the

International Journal of Innovative Science and Research Technology

#### https://doi.org/10.38124/ijisrt/25mar1891

ISSN No:-2456-2165

candidate model with the lowest AIC value. The AIC is expressed as equation (3) below:

$$AIC = 1n\frac{\sum_{i=1}^{t}\hat{\varepsilon}_i^2}{T-n} + \frac{2n}{T}$$

Where  $\hat{\varepsilon}_i^2$  denotes squared residual estimates, T for observation size within samples, and *n* for the estimated parameters (Lee, 2020).

Table 1 ARIMA (0,1,\*) Selection using AIC

ARIMA Model	AIC
**0,1,1**	102.1462
0,1,2	102.0872
0,1,3	104.0393
0,1,4	104.4132
0,1,5	106.1677

Table 2 ARIMA (0,2,\*) selection using AIC

ARIMA Model	AIC
**0,2,1**	94.68104
0,2,2	95.22491
0,2,3	97.49976
0,2,4	100.3610

The ARIMA (0,1,1) has the bottom AIC price thinking about d=1 at the same time as ARIMA (0,2,1) has the lowest AIC price thinking about d=2 as obtrusive in Tables 1 and a couple of table 3 indicates that ARIMA (zero,2,1) version regarded to be the statistically appropriate version to forecast the enrollment charge of the RP for S.Y. 2019-2020 to2024-2025. The version mounted the lowest AIC price and is premiere for prediction.

# V. OUTCOMES AND DISCUSSION

#### ➤ Version Match

Prediction mistakes for each the schooling and validation information pass down together with more iterations, which suggests that overfitting does not arise. A greater distinctive version assessment, like go-validation, can also be done for both strategies. but, due to the fact our pattern length could be very huge and the range of enter variables is pretty small, it's miles clean that the training, validation, and test data in our have a look at replicate the complete dataset well.



Graph 1 Suggests class consistency at every overall performance level reduce respectively:

https://doi.org/10.38124/ijisrt/25mar1891

Table 3 Comparison of Classification Consistency Index for Two Overall Performance Stage Cuts						
	Performance Level Cut 1		Performance Level Cut 2		Cut 2	
Target Field	Linear	Bayesian	Regression	Linear	Bayesian	Regression
	regression	networks	tree	regression	networks	tree
G8 Math	0.698	0.742	0.788	0.882	0.896	0.901
G8 Reading	0.822	0.815	0.845	0.858	0.863	0.874
G8 Science	0.802	0.804	0.818	0.856	0.878	0.885
G7 Math	0.832	0.839	0.853	0.899	0.903	0.909
G7 Reading	0.820	0.819	0.842	0.856	0.866	0.876
G6 Math	0.780	0.831	0.845	0.882	0.910	0.915
G6 Reading	0.786	0.832	0.847	0.846	0.884	0.889
G5 Math	0.784	0.818	0.822	0.863	0.882	0.885
G5 Reading	0.787	0.828	0.833	0.853	0.877	0.880
G5 Science	0.759	0.808	0.810	0.898	0.910	0.911
G4 Math	0.797	0.823	0.826	0.857	0.884	0.885
G4 Reading	0.803	0.820	0.834	0.830	0.871	0.871

G7 Math scores are higher in all conditions while searching at the first overall performance stage reduce. usually, the consistency of classifications for rating Forecasting using device studying fashions that predicted scale ratings with Bayesian networks are much like those anticipated by regression timber, and lots higher than the ones from linear regression. however, there's one exception: for the Grade eight studying test, the consistency index for anticipated rankings from Bayesian networks at the first overall performance stage cut is lower than that from linear regression.

• Prediction Errors



Graph 2 Prediction Errors

#### https://doi.org/10.38124/ijisrt/25mar1891

# ISSN No:-2456-2165

This histogram suggests that the XGBoost regression tree has the smallest RMSE a number of the three methods. Bayesian networks are slightly worse than XGBoost and better than linear regression for maximum topics and grades, besides for grade eight arithmetic. in addition, we additionally compute the imply mistakes and locate that XGBoost has the most solid and lowest mean absolute mistakes throughout all exams.

• Suggest Absolute Mistakes for all Tests with the Aid of Three Strategies



Graph 3 Suggest Absolute Mistakes for all Tests with the Aid of Three Strategies

The cause why Bayesian networks do not perform nicely is probably that pretty some college students have missing values for their previous-yr rankings, and Bayesian networks could offer horrific predictions for these students. at the opposite, XGBoost gives mean mistakes Linear Regression Bayesian internet XGB Regression Tree rating Forecasting with system gaining knowledge of model higher predictions for college kids with missing information.

• Graphical and Statistical Strategies



Fig 1 Graphical and Statistical Strategies

Shows the time collection plot for Rwanda polytechnic high learning enrollment rate from S.Y. 2011-2012 to 2018-2019. An upward fashion is obvious within the graph, which denotes more and more enrollees over the years. It is also show us the higher performance students in sciences at different years. The year 2018-2019 increases compare to the other remaining years. Which means that students in the year 2018 up 2019 performed well in sciences compare to the remain years in education at Rwanda polytechnic Kigali college as actual enrollment.

ISSN No:-2456-2165

https://doi.org/10.38124/ijisrt/25mar1891









Figure 3, indicates the ACF and PACF correlogram plot for (MA)(q) and (AR)(p) value for lags 1 to eight. The ACF and PACF showed a decaying pattern throughout the lags. The ACF is decaying, whilst the PACF is likewise decaying but now not in an abrupt manner. This denotes a zero value for autoregressive (p) and a technique for the moving average (q). meanwhile, a fashion is obvious in determine 6, making the information non-stationary, therefore including value in d within the ARIMA(p, d,q) model. The foremost p,d,q version is selected based totally at the model with the bottom Akaike information Criterion (AIC), as shown in Tables 1 and 2.



Fig 4 The Forecasted Number of Enrollees for Students Performance at Different YearsError! Reference source not found.

#### International Journal of Innovative Science and Research Technology

### https://doi.org/10.38124/ijisrt/25mar1891

ISSN No:-2456-2165

Figur4 suggests the graphical illustration of the forecasted enrollment facts of the RP Kigali College Institution, with its 95% self-assurance c programming language the use of the ARIMA (0,2,1) version. An growing fashion is projected from S.Y. 2019-2020 onwards with the specific forecasted values for S.Y2019-2020 to 2024-2025, as proven in table 5, and additionally in the figure 4, this suggest that there's a amazing overall performance in technological know-how from the years 2022 up 2024. this means that students perfomed nicely in sciences considering to the ultimate years.

As the it shows in 2022 to 2024 students performed well in sciences considering to the remaining years. In 2024 students performed at 95 percentage compare to the other remaining years which implies that students entered at RP Kigali College in 2024, the largest number are those who had background knowledge in sciences. This implication have made them to perform very well in this 2024. Which mean that when students being selected at RP Kigali College are those without any knowledge in sciences this will make to fail science in successively years.

Table 4 Forecasted Enrollment while variety with a 55% Confidence e Frogramming Language						
School Year	Forecast	Lo 95	Hi 95			
2019-2020	2,350	1,697	3,003			
2020-2021	2,616	1,692	3,540			
2021-2022	2,918	1,786	4,049			
2022-2023	3,256	1,949	4,562			
2023-2024	3,630	2,169	5,090			
2024-2025	4,040	2,440	5,640			

# Table 4 Forecasted Enrollment Wide Variety with a 95% Confidence c Programming Language

# VI. CONCLUSION AND RECOMMENDATION

On this paper, the most advantageous ARIMA (p, d, q) model turned into identified to forecast the destiny enrollment fashion in Rwanda Polytechnic Kigali college high learning Institution, for the faculty years 2019-2020 to 2024-2025. The ACF, time difference and the PACF procedure were diagnosed as a way to come up with the proper version for prediction. The choice of the version with the bottom AIC fee was discovered so that you can come up with the fine model. The simulation outcomes confirmed that ARIMA (0,2,1)model changed into recognized as the first-rate ARIMA (p,d,q) model to forecast enrollment in Rwanda Polytechnic Kigali College High Learning Institutiton. The forecast confirmed an growing fashion in enrollment for the succeeding sciences RP Kigali college years. Destiny researchers may additionally take into account predicting precise enrollment counts per faculties for a higher understanding of enrollment fashion analysis and expertise extraction. it is recommended that the usage of statistics analysis techniques and algorithms be discovered for similarly information extraction.

#### ABBREVIATIONS

- ACF: autocorrelation feature
- AR: autoregressive
- ARIMA version: automatic regressive incorporated shifting average model
- EQUATOR: improving the satisfactory and Transparency Of fitness studies
- ITS: interrupted time series
- MA: transferring common
- PC: partial autocorrelation feature
- RCT: randomized managed trial

#### REFERENCES

- [1]. Anitha. (2018). group of workers forecasting models: a systematic overview.
- [2]. Arteta, G. D. (August 22, 2). profits Inequality in university Enrollment and diploma Attainment at some point of and After the splendid Recession Years.
- [3]. Clar, W. (2023). Seasonal Autoregressive integrated transferring average .
- [4]. ELSEVIER. (2017). analyzing undergraduate students' overall performance the usage of instructional data mining.
- [5]. Ghosh, J. ok. (December 2011). AIC, BIC and current advances in version choice.
- [6]. Hayes, a. (2022). what is a Time collection and how Is It Used to analyze facts?
- [7]. Kyla, e. (2020). instructional massive records: Predictions, packages and challenges.
- [8]. Lee, J. C. (August 22, 20). profits Inequality in university Enrollment and degree Attainment during and After the notable Recession Years.
- [9]. Mohammed. (2024). university scholar Enrollment Prediction: A machine mastering Framework.
- [10]. Pearson, T. &. (22 March 2021). Interrupted time series analysis using autoregressive incorporated.
- [11]. Shwetha, B. N. (2023). analysis of character family strength intake Forecasting the use of ARIMA model, CNN and LSTM model.
- [12]. tds. (Jul 29, 2021). a thorough guide to Time collection evaluation | by way of Fangyi . Togo: Kelly,El.
- [13]. Boken, V.K. (2000). Forecasting spring wheat yield using time series analysis: a case study for the Canadian Prairies. Agron. J, 92:1047-1053.

ISSN No:-2456-2165

- [14]. Chrisbooks (2008), Introductory Econometrics for finance, the ICMA Centre, University of Reading, second edition Fatema Tuz Jhohura and Md. Israt rayhan (2009), An assessment of renewable energy in Bangladesh through ARIMA, Holt's, ARCH-GARCH Models, Institute of Statistical Research and Training (ISRT)
- [15]. Iqbal, N., K. Bakhsh, K. Maqbool, and A.S. Ahmad (2000). Use of the ARIMA model for forecasting wheat area and production in Pakistan. Int. J. Agri. Biol. 2: 352-354.
- [16]. Kathiresan.A (2011). Strategies for Sustainable Crop Intensification in Rwanda, Kigali
- [17]. P.J. Brockwell and R.A. Davis (1986). Time Series: Theory and Methods, Springer Series in Statistics.
- [18]. P.J. Diggle (1990). Time Series: A Biostatistical Introduction, Oxford University Press
- [19]. Pasapitch Chujai, Nittaya Kerdprasop, and Kittisak Kerdprasop (2013). Time series Analysis of House hold Electric consumption with ARIMA and ARM A models, IMECS 2013, Vol. 1
- [20]. Indian Journal of Agricultural Sciences 2011 No.81 Vol. 6, pp 520 -523.
- [21]. Samer Saab, Elie Badr and George Nasr (2001). Univariate modeling and forecasting of energy consumption the case of electricity in Lebanon, 2001 Energy, Vol. 26, pp. 1-14.