# An Exploration of Machine Learning Models to Forecast the Unemployment Rate of South Africa: A Univariate Approach

Rudzani Mulaudzi
School of Computer Science
and Applied Mathematics
The University of the Witwatersrand
Johannesburg, South Africa
0601737r@students.wits.ac.za

Ritesh Ajoodha
School of Computer Science
and Applied Mathematics
The University of the Witwatersrand
Johannesburg, South Africa
ritesh.ajoodha@wits.ac.za

***Abstract*—The South African unemployment rate is 29.1%, this is the highest unemployment rate that the country has recorded since the 1970s. The country is in the top ten countries with the highest unemployment rates in the world. COVID-19 threatens to increase the unemployment rate above the 50% mark. A public policy intervention is the most suitable instrument for the country in order to address this problem, however, policy is reliant on accurate and reliable forecasting. This paper explores univariate machine learning techniques to forecast the South African unemployment rate. Six traditional statistical models are compared with seven machine learning models. The multi-layer perceptron achieves the lowest error rate, whilst the ridge regression model achieved the highest $R-squared$. These are closely followed by ARIMA, LASSO, and the elastic net, showing that machine learning models can forecast the South African unemployment rate with higher accuracy than traditional statistical methods.**

***Index Terms*—Forecasting, Machine Learning, ARIMA, Unemployment, Univariate**

## I. Introduction

It is no longer surprising when Statistics South Africa announces an increase in the South African unemployment rate. This has been a constant theme for the past twelve years. Additionally, the advent of the so-called 'Fourth Industrial Revolution' is estimated to lead to higher unemployment rates in the immediate future. Therefore, South Africa is a "social unrest ticking time-bomb" because there is currently no obvious solution to the current unemployment crisis and there is already evidence of frustration amongst the unemployed [1].

Public policy is one of the only known methods that could alleviate unemployment in South Africa. This is because policy implementations can result in systemic changes across the whole country [2]. One example is the broad-based black economic empowerment policy in South Africa, which gives procurement preferences to companies that are owned, governed, and run by designated people groups i.e. black people, women, youths, and persons with disability [3]. In order for policy to be developed, accurate and reliable forecasting is relied upon [2].

Machine learning techniques are capable of performing the requisite forecasts as well as capturing non-linear patterns in the data. Neural networks and regression models have thus far been applied for unemployment forecasting in North America, Europe, and Asia [4]–[9].

The performance of machine learning forecasts is typically benchmarked against traditional statistical models. Six traditional statistical models are common benchmarks and are used as benchmarks in this paper as well: the naïve model, the moving average, simple exponential smoothing (SES), Holt-Winters, autoregressive integrated moving average (ARIMA), and Seasonal ARIMA models. These are univariate models that use the lag values of the unemployment rate as predictors of the unemployment rate.

This research explores i) the utility of machine learning in forecasting the South African unemployment rate and ii) the ability of machine leaning models to capture the underlying patterns in the South African unemployment rate. Seven machine learning models are utilised for this research: linear regression, elastic net, ridge regression, least absolute shrinkage and selection operator (LASSO), support vector regression (SVR), multi-layer perceptron (MLP), and recurrent neural networks (LSTM). The lags of the South African unemployment rate are used as the feature set whilst the unemployment rate itself is the target.

The rest of this paper is structured as follows. Section II discusses univariate statistical models and machine learning models that have been used to forecast unemployment rates across the world. Section III discusses the results that were achieved by employing machine learning models to forecast unemployment in South Africa. Section IV is the final section which discusses the contribution of this research as well as opportunities for future research.

## II. Literature Review

In South Africa, there is a limited body of literature on forecasting the unemployment rate, more so on forecasting using machine learning. The machine learning techniques are mostly applied in North America, Europe, and Asia. Both univariate and multivariate approaches are used in literature. However, this paper focuses on univariate approaches as these are commonly used in literature [4], [5], [10], [11].

In cases where multivariate models are used, the data is typically accessed from a single source such as the Federal Reserve Bank of St Louis (FRED) [6], [7], [12]–[15]. These databases are often 'model ready' as the data is already in a clean tabular structure and in some cases the data can even be accessed through application programming interfaces. However, in South Africa, reliable and accurate data is in multiple data sources such as Statistics South Africa, South African Reserve Bank, Bureau of Economic Research, and Quantec.

The South African data sources are in different number formats, time frequencies, and organisational structures. Therefore, significant data wrangling is required in order to use a multivariate approach in South Africa. Hence a univariate approach is selected for this paper.

Univariate models are linear models that forecast a variable based on its lag values and associated error terms [16]. The ones that are most commonly used for forecasting unemployment rates and utilised in this research will be discussed in this section. Starting with the traditional statistical ones ending with machine learning ones.

**Traditional Statistical Methods**

### A. Naïve

Naïve forecasts are those that assume what has happened in the past will continue happening in the future. Therefore, future unemployment is forecasted based on recently observed unemployment data only. This technique is sometimes used in research as a baseline to determine if alternative unemployment forecasting models are better [17].

### B. ARIMA

The most commonly used univariate statistical model is the ARIMA model. The model states that the dependant variable, $u_t$, can be determined using a linear combination of past and present values of the same variable i.e. $u_{t-i}$, where $t-i$ is a time period $i$-times before $t$ [17]. The ARIMA model is made of a combination of three common time-series models:

1) *Autoregressive (AR) model*, which is the linear regression part of the ARIMA model, where the current variable, $u_t$, depends only on the past values of the variable and an error term [17]. This can be presented as $\phi(L)u_t = \epsilon$, where $\phi(L)$ is a linear combination as shown in equation (1), with $L$ being the lag operator, which returns the previous elements of a series, for example, $L^2$ and $L^i$ would return two previous values and $i$ previous values respectively. The $\epsilon$ is the error term.
2) *Integrated (I)*, is the mechanism to ensure the data being considered is stationary [17]. Stationary data being time-series data with a constant mean and a constant variance. This stationarity is achieved by differencing the series $d$ times.
3) *Moving Average (MA) model*, are linear combinations of the error terms of a regression process. These terms are generated independently of each other, are not correlated with one another, and have a constant mean and variance [17]. The $MA$ term is given by $y_t = \theta(L)u_t + \mu$, where $\theta(L)$ a linear combination as shown in equation (2) [17].

Several researchers and institutions use the ARIMA model to forecast unemployment rates across the world [16]–[18]. In equation (3), a version of the ARIMA equation is shown [17].

$$\phi(L) = 1 - \phi_1 L - ... - \phi_2 L^p \tag{1}$$

$$\theta(L) = 1 - \theta_1 L - ... - \theta_q L^q \tag{2}$$

$$\phi(L)(1-L)^d u_t = c + \theta(L)\epsilon + t \tag{3}$$

Where, $u_t$ is the unemployment rate being estimated. Variables $p$, $d$, and $q$ are natural numbers used to set the autoregression order, the differencing order, and moving average order respectively.

### C. Holt-Winters

Holt is a model that takes trends into account in how it forecasts. Holt-Winters is an extension of Holt, allowing for seasonality to be captured in models [18]. The Holt-Winters model enables trends and seasons to be modelled by using triple exponential smoothing to represent and capture past data [18]. Holt-Winters and ARIMA are standard benchmarks when modelling univariate time-series data [10]. The models are used as orthodox benchmarks across univariate unemployment forecasting research [16]–[18].

**Machine Learning Methods**

### D. Elastic Net (ENet)

Core to machine learning algorithms is minimising errors relating to bias and variance, which coincidentally have a trade-off relationship. The elastic net (ENet) model was designed to specifically minimize both these errors through regularization. The model has penalties for over-fitting by penalising the model if it is overly complex (too many variables) or has over-reliance on particular variables. Through this process, the ENet learns which features are most important in the data without a need for the researcher to make assumptions as is required in the traditional forecasting approaches [12].

It was demostrated in the United States of America (USA) that ENet can forecast unemployment over 3, 6, 9, 12, and 24-month horizons more accurately than traditional statistical forecasting models and professional forecasters which were used as baseline models [12]. This model was able to identify unemployment shifts over recessions and booms more accurately than the traditional statistical models. Through the regularisation process of ENet, the model is able to identify key variables that predict unemployment by setting all other coefficients to zero [12].

### E. Least Absolute Shrinkage and Selection Operator (LASSO)

LASSO regression has been used to forecast unemployment rates in the USA. This regression model minimizes prediction error through a regularization process. The model penalises the reliance on particular variables for prediction, therefore, reducing over-fitting. LASSO improved the forecasting accuracy of unemployment in the USA as well as identifying variables which are most important in forecasting unemployment [7].

### F. Support Vector Regression (SVR)

Support vector machines (SVMs) are an approach to supervised learning that enables non-linear data to be classified using a hyperplane [6]. The approach outputs a class identity, categorising the inputs into varying classes. These models are useful for modelling non-linear data. There are two types of SVMs: support vector classifiers (SVC) and support vector regression (SVR). To use SVMs for regression problems an implementation of the $\varepsilon$-sensitive loss function is required which enables non-linear regression modelling [9]. The loss function is shown in equation (4):

$$L_\varepsilon(x_i) = \begin{cases} 0, & \text{if } \mid y_i - f(x_i) \mid \leq \varepsilon \ , \\ & \varepsilon \geq 0 \\ \mid y_i - f(x_i) \mid - \varepsilon, & \text{if other} \end{cases} \tag{4}$$

Where, $y_i$ is the correct output, $f(x_i)$ is the predicted output, and $\varepsilon$ is the error. The usage of non-linear decision boundaries in SVRs makes them effective in unemployment forecasting tasks with an accuracy that is above linear regression [4], [8].

### G. Multi-layer Perceptron (MLP)

Multi-layer perceptrons (MLPs) are neural networks where the data flows from input to output in a ‘forward’ direction without moving backwards. Their goal is to find a predictor function, $f$, which maps an input $x$ to some output $y$ i.e. $y = f(x)$ [19]. MLPs are foundations to more advanced neural networks and they are referred to by multiple names. MLPs enable non-linear relationships to be effectively modelled, which is important for unemployment data.

Fully connected MLPs were used to predict unemployment rates in the USA and they found them to perform better than Surveys of Professional Forecasters [4]. MLPs offer a significant improvement in forecasting results when compared to traditional unemployment forecasting techniques [7], [9].

In the Mediterranean countries, MLPs were sometimes outperformed by Holt-Winters models on preprocessed data that showed evidence of non-linearity, heteroskedasticity, and non-normality [10]. The key reason given for the under performance by the neural network was the simplicity of the MLP structure: a single hidden layer model with 1 to 10 nodes, and a sigmoid activation function. It was also observed that large data sets are required for neural networks to outperform traditional statistical methods [20]. Therefore, it is clear that neural networks need sufficient data and an appropriate architecture to perform effectively.

### H. Recurrent Neural Networks (RNN)

Recurrent Neural Networks (RNN) are a type of neural network architectures that enables effective modelling of sequences of data as they possess an internal memory structure [19]. At their core they are essentially MLPs with feedback loops [19]. RNNs were developed for natural language processing requirements such as language translation. These models can forecast unemployment accurately by using a variation of RNNs referred to as Long Short Term Memory (LSTM) networks [4]. The model outperformed the ARIMA model, using only past unemployment rates in the USA as the input to the model [4]. They were noted as the top performing neural network architecture when forecasting the USA unemployment rate.

This section provides an overview of the traditional statistical models and machine learning models that have been employed to forecast unemployment rates across the world. The next section will discuss results from the application of traditional statistical models and machine learning models in South Africa.

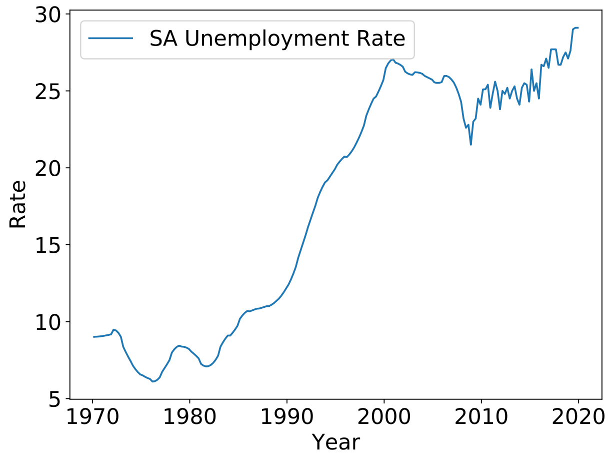
## III. Results and Discussion

A univariate time-series analysis was undertaken to forecast the South African unemployment rate, this is similar to what was used by several researchers [4], [5], [10], [11]. The model’s performance was measured based on their mean absolute percentage error ($MAPE$): a common performance measure in time-series analysis. The following subsection will provide the descriptive statistics of unemployment as well as the results of the analysis.

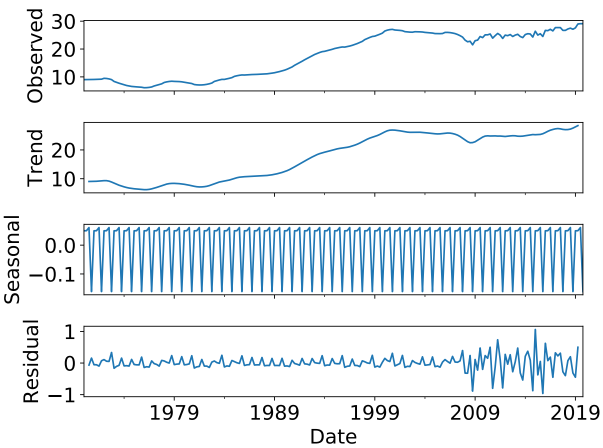
### A. Statistical Description of the South African Unemployment Rate

The unemployment data used for this analysis was accessed from the Bureau of Economic Research (BER). The time-series had a total of 209 quarterly observations of the South African unemployment rate from January 1970 to December 2019. Figure 1 (a) depicts this data. The average unemployment rate over the time period is 17.7%, the lowest was 6.1% in the 1970s, and the highest rate was 29.1% which is the current unemployment rate (reported December 2019).

Time-series data can be decomposed into three key parts: the trend, seasonality, and residuals [18]. It was observed that the South African unemployment rate trends upwards and is seasonal where repeated movements in the data are observed each year: depicted in Figure 1 (b). The change in variation is visible by studying the residuals, where it can be seen that between 2009 and 2019 there was greater variability in the data than all the previous periods.




(a) A depiction of the unemployment rate from January 1970 to December 2019.




(b) The decomposition of the South African unemployment into the trend, seasonality, and residuals.

Fig. 1: The descriptive visualisation of the South African unemployment rate.

From Figure 2, it is clear that the South African unemployment rate does not follow a normal distribution. The autoregressive nature of the univariate traditional statistical methods generally assumes that the data follows a normal distribution with a zero mean and a constant variance. Kolmogorof and Lilliefors tests were also conducted, which are well-known tests for normality, and they both confirmed that the unemployment rate is not normally distributed. Figure 2 shows the unemployment rate kernel distribution function whose shape suggests that the unemployment rate is bi-modal.

It was also observed that the unemployment rate is also non-stationary over the observed period. Stationarity was tested through Augmented Dickey–Fuller ($ADF$) test, whose null-hypothesis is that the series is non-stationary. The $ADF\ test$ returned a $p-value$ of 0.923 which means that we accept the null-hypothesis that the South African unemployment rate is non-stationary. This result was confirmed with two other tests: Dickey-Fuller GLS and Phillips Perron test, both of which confirmed the $ADF$ results.

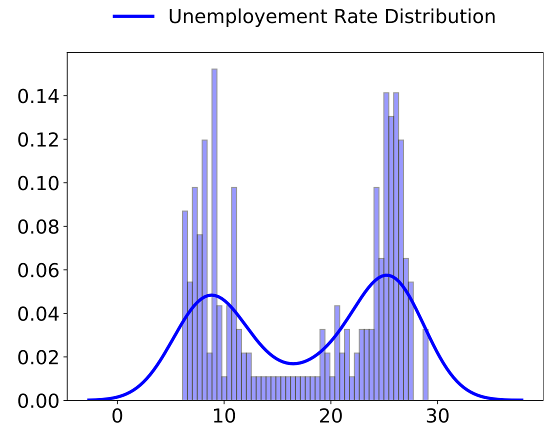



Fig. 2: The South African unemployment rate bi-modal probability distribution.

### *B. Analysis of the South African Unemployment Rate*

Six traditional statistical models were trained to forecast South Africa's unemployment rate: the naïve model, the moving average, simple exponential smoothing (SES), Holt-Winters, autoregressive integrated moving average (ARIMA), and Seasonal ARIMA (SARIMA) model. Seven machine learning models were compared against the traditional statistical models: linear regression, elastic net, ridge regression, least absolute shrinkage and selection operator (LASSO), support vector regression (SVR), multi-layer perceptron (MLP), and a recurrent neural network (LSTM variation).

The models were trained and tested on time-series data comprising of 209 quarterly observations of the South African unemployment rate from January 1970 to December 2019. This data was split into training and testing, where, January 1970 to October 2004 was the testing data and January 2005 to December 2019 the testing data: an 80% training and 20% test data split. The performance measures used were the mean absolute percentage error ($MAPE$) and R-squared ($R^2$) which are standard measures in time-series analysis.

$R^2$ evaluated each model's ability to explain the South African unemployment rate. $MAPE$ is selected because it is orthodox and can be scaled and compared across different time-series [21]. The results are presented in Table I showing that the multi-layer perceptron (MLP) has the lowest $MAPE$ with the second highest $R^2$. This is followed by the ARIMA model. ARIMA and Holt-Winters are standard benchmarks when modelling univariate time-series data [10]. Even though ARIMA had relatively low error rates, Holt-Winters had relatively higher error rates. The machine learning models – ridge regression (ridge), elastic net (ENet), MLP, LASSO, linear regression, and SVR – performed better than Holt-Winters as measured by $MAPE$ and $R^2$.

Table I also shows that ridge has the highest $R^2$, meaning that the model was able to capture the properties of the South African unemployment rate. This is because the South African unemployment rate is highly correlated with its lag values (Figure 3). This correlation typically causes high variance in

TABLE I: Model evaluation table that compares the performance of traditional statistical models and machine learning models for forecasting the South African unemployment rate.

| Model | MAPE | $R^2$ |
|---|---|---|
| MLP | 2.387 | 0.696 |
| ARIMA | 2.391 | 0.696 |
| Ridge | 2.468 | 0.703 |
| ENet | 2.483 | 0.698 |
| LASSO | 2.655 | 0.633 |
| Linear Regression | 2.873 | 0.588 |
| SVR | 2.965 | 0.597 |
| Holt Winters | 3.056 | 0.481 |
| LSTM | 3.391 | 0.404 |
| SARIMA | 3.745 | 0.155 |
| SES | 6.947 | -1.660 |
| ARMA | 6.964 | -2.116 |
| Naive | 7.040 | -1.712 |

Machine Learning are models highlighted in green.

machine learning models, which can be controlled through regularization. Ridge, LASSO, and ENet have regularization terms that reduce variance. However, LASSO penalises variances by setting some of the correlated terms to zero to avoid model complexity: making the model in this case behave more like linear regression as only two lag values were used. Ridge on the other hand drives the contribution of the correlated terms close to zero but not to zero, through $L2$ regularisation. Ridge, therefore, captures the contribution of all the lag variables, whilst LASSO does not. ENet combines the regularisation approaches of both LASSO and ridge in an attempt to capture the benefits offered by both models.

The appropriate parameters of ridge, LASSO, and ENet were discovered through grid search: over 100 000 different regularization rates were explored. Grid search was also used to determine the parameters to be used in the ARIMA model. Every possible combination from 0 to 15 was considered to determine the appropriate parameters: $p$, $d$, and $q$. However, the most suitable parameters were discovered by studying the Auto-correlation Function (ACF) and Partial ACF functions to determine $q$ and $p$ respectively. The $d$ was determined by testing the stationarity of the time-series (using ADF test) after differencing it. The series was stationary after differencing it once, therefore, $d$ was set to one.

Trial and error was used to determine the parameters and hyperparameters for the MLP, LSTM, Holt-Winters, and SES models. A standard (default) model was tried at first and tuned until continual tuning did not result in a reduction in the error rate. In some cases grid search was used to find the general range which were likely to lead to the lowest error rates. However, grid search, was not feasible with the neural network approaches as this was too computationally expensive. The final MLP model had three hidden layers, each with a 100 nodes, rectified linear unit (ReLU) was used as the activation function, with an adaptive learning rate.

Unlike traditional statistical models, the machine learning models require that there be features and targets. The lag values of the unemployment rate were used as features of the model with the observed unemployment rate being the target variable. The number of lags were selected by trial and error. Lags between 2 and 7 were tried. These lags were treated as a parameter that was set and updated based on the performance of the model. Two lags were the best performing across the models and therefore this was used across the models i.e. two features and one target variable. Figure 3 shows that the unemployment rate has a high correlation with its lag values. Therefore, there is evidence of temporal dependencies in the South African unemployment rate.

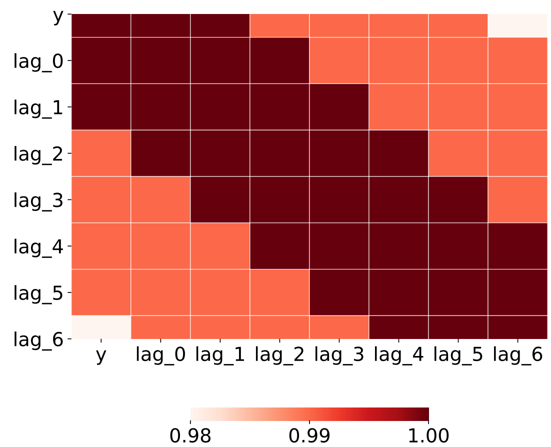



Fig. 3: The Correlation Map for the lag values of unemployment rate showing that the unemployment rate is highly correlated with the lag values.

The worst performing machine learning model was the Long Short Term Memory (LSTM). The hyperparameters of the model were tuned through trial and error. The depth and width of the LSTM were treated as tunable parameters and were updated until there was no improvement in the error rate. Dropout was also applied as there was evidence of overfitting without dropout: the dropout rate was set to 20%. The LSTM is expected to perform well with time-series data as observed when forecasting the USA unemployment rate: were the LSTM outperformed all other machine learning models [12]. The key difference between the USA and the South African forecasts is that the data in this paper was less than 30% what was used in the USA research. The LSTM under performance in this paper is consistent with research that states that the data size matters when it comes to forecasting time-series using deep neural networks [20]. Furthermore, the USA unemployment rate is not trendy, and has a fairly constant mean and variation, it exhibits stationary traits whilst the South African unemployment rate is non-stationary.

Table II shows predictions over the last four periods in the test data. The actual unemployment rate is highlighted in blue whilst models that were closest in that quarter are highlighted in green. It is clear to see that the ARIMA model, although it was the second best performing model, it seems to have centered its predictions between 23% and 24%. The ARIMA does not capture the non-linear movements of the

South African unemployment rate. The Holt-Winters model captures this better, even though its $MAPE$ was lower than most models. Figure 4 depicted this, where the graphs show that the machine learning models are able to capture the underlying patterns in the unemployment rate, whilst ARIMA is unable to.

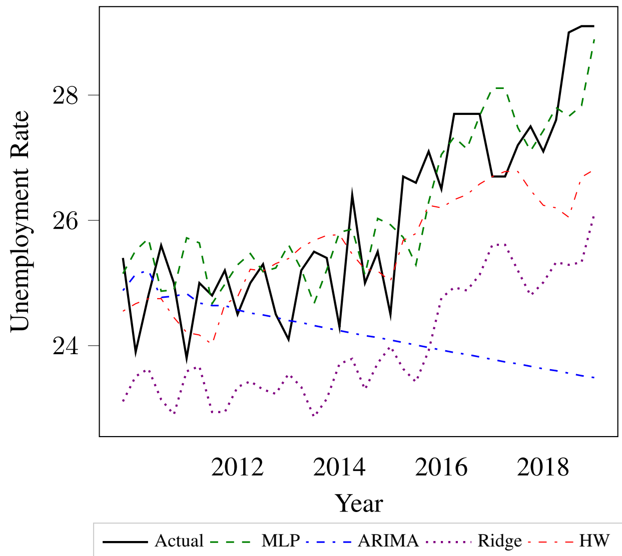



Fig. 4: The forecast of the machine learning models and statistical models over the test period, depicting the ability of machine learning models to capture the underlying properties of the South African unemployment rate.

TABLE II: A comparison of the machine learning and traditional statistical models over the last four periods of the test data.

| | Test Date | | | |
|---|---|---|---|---|
| **Model** | Mar-19 | Jun-19 | Sep-19 | Dec-19 |
| Actual | 27.6 | 29.0 | 29.1 | 29.1 |
| MLP | 27.8 | 27.7 | 27.8 | 28.9 |
| ARIMA | 23.6 | 23.6 | 23.5 | 23.5 |
| Ridge | 25.3 | 25.3 | 25.3 | 26.1 |
| ENet | 27.6 | 27.3 | 27.6 | 28.8 |
| LASSO | 27.7 | 27.3 | 27.8 | 29.2 |
| Linear Regression | 27.9 | 26.7 | 28.3 | 30.7 |
| SVR | 27.3 | 27.2 | 27.3 | 28.2 |
| Holt Winters | 26.2 | 26.0 | 26.7 | 26.8 |
| LSTM | 26.7 | 26.7 | 26.5 | 26.6 |
| SARIMA | 25.8 | 25.8 | 26.0 | 25.7 |

Therefore, machine learning models offer an added advantage in that they are able to capture the underlying movements in the South African unemployment rate. These models are able to do so over long periods of time. ARIMA, SARIMA, and SES tend to center their forecast around an average unemployment rate. Therefore, machine learning models are better suited for modelling and forecasting the South African unemployment rate.

These models should be considered by South African policy makers when formulating unemployment policies. Based on the results obtained, machine learning forecasts should be incorporated as a critical input to policy setting as these forecasts can provide a data-driven view that is able to capture non-linear patterns in the data as well as improving the forecasting accuracy.

## IV. Conclusion

Unemployment is one of the biggest problems in South Africa. It is a problem that has been growing without any obvious solutions over the last two decades. Policy is one of the most effective instruments to alleviate unemployment in the country. However, policy relies on accurate and meaningful forecasting of the unemployment rate. Currently, these forecasts are carried out through traditional statistical methods, amongst other tools. This paper demonstrated that machine learning models are able to not only reduce the error rates in univariate forecasting of the South African unemployment rate, but are also able to capture the underlying patterns in the movement of unemployment rates.

The multi-layer perceptron (MLP) with three hidden layers was able to forecast unemployment rates more accurately than all the traditional statistical methods. The model captures the non-linear structure of the South African unemployment rate. The MLP forecasts that the South African unemployment rate should exceed 32% this year whilst the Bureau for Economic Research (BER) forecasted that South Africa would only reach 32% by 2024 (based on the January 2020 forecast). The ridge regression model achieved the highest $R^2$ and was able to capture and represent the high correlations that the South African unemployment rate has with its lag values. Six of the seven machine learning models were able to achieve lower forecasting error rates than the Holt-Winters model. The Holt-Winters and ARIMA models are considered orthodox benchmarks when forecasting unemployment rates.

Therefore, this research demonstrated that machine learning techniques are a viable alternative to univariate forecasting of the South African unemployment rate. These techniques were deployed on the unemployment time-series data without significant prior analysis and pre-processing of the data structure and data proprieties as was required by the traditional statistical methods. The models were able to capture the non-linear movements of unemployment rates and produce forecasts that had lower error rates than the traditional statistical methods in methods in most cases.

This research is a motivation for a multivariate approach to be explored because it demonstrated that univariate forecasting can be successfully achieved through machine learning. Machine learning techniques offer additional benefits in multivariate contexts because models such as elastic net and LASSO that are able to automatically determine which features predict unemployment movements.


## Acknowledgment

This work is based on the research supported in part by the National Research Foundation of South Africa (Grant number: 121835).

## References


[1] D. F. Meyer, “Job creation, a mission impossible? The South African case,” Mediterranean Journal of Social Sciences, 5(16):65–77, 2014.

[2] J. Levinsohn. “Two Policies to Alleviate Unemployment in South Africa,”. Technical report, 2007.

[3] B-BBEE Commission. “Broad-Based Black Economic Empowerment Act,”. theDti (www.bbbeecommission.co.za), 2013.

[4] T. R. Cook, A. S. Hall. “Macroeconomic Indicator Forecasting with Deep Neural Networks,” Unpublished, 2017.

[5] K. Mahipan, N. Chutiman, B. Kumphon. “A forecasting model for Thailand’s’s unemployment rate,” Modern Applied Science, 7(7):10–16, 2013.

[6] G. Sermpinis, C. Stasinakis, K. Theofilatos, A. Karathanasopoulos. “Inflation and Unemployment Forecasting with Genetic Support Vector Regression,” Journal of Forecasting, 33(6):471–487, 2014.

[7] A. Kreiner, J. V. Duca. “Can machine learning on economic data better forecast the unemployment rate?,” Applied Economics Letters, 2019.

[8] C. Katris. “Prediction of Unemployment Rates with time-series and Machine Learning Techniques,” Computational Economics, 2019.

[9] M. Aiken. “A neural network to predict civilian unemployment rates,” Technical report, 1996.

[10] C. Katris, “Forecasting the Unemployment of Med Counties using Time-Series and Neural Network models Forecasting the Unemployment of Med Counties using time-series,” Journal of Statistical and Econometric Methods, 8(2):37–49, 2019.

[11] E. Olmedo. “Forecasting Spanish Unemployment Using Near Neighbour and Neural Net Techniques,” Computational Economics, 43(2):183–197, 2014.

[12] A. S. Hall. “Machine Learning Approaches to Macroeconomic Forecasting,” Economic Review (Kansas City, MO), 2018.

[13] G. S. Atsalakis, C. Ioana, C. H. Skiadas. “Forecasting Unemployment Rate Using a Neural Network with Fuzzy Inference System,”. ICAP, 2007.

[14] G. N. Kouziokas “Unemployment Prediction in UK by Using a Feed-forward Multilayer Perceptron,”. Unpublished, pages 65:74, 2019.

[15] R. Diana, I. N. Budiantara, S. Darmesto. “Statistical modeling for unemployment rate using smoothing spline in semiparametric multivariable regression model with Bayesian approach,” Model Assisted Statistics and Applications, 9(2):159(166), 2014.

[16] C. Brooks, “Introductory Econometrics for Finance,” Cambridge UniversityPress, 3rd edition, 2014.

[17] A. L. Montgomery, V. Zarnowitz, R. S. Tsay, G. C. Tiao, F‘orecasting the U.S. Unemployment Rate.Journal of the American Statistical Association, 93(442):478, jun 1998.

[18] R. J. Hyndman, G. Athanasopoulos, “Forecasting: Principles and Practice,” OTexts, Melbourne, Australia, 2 edition, 2018.

[19] I. Goodfellow, Y. Bengio, A. Courville. “Deep Learning,” MIT Press, 2016.

[20] V. Cerqueira, L. Torgo, C. Soares. “Machine Learning vs Statistical Methods for time-series Forecasting: Size Matters,”. arXiv, 2019.

[21] R. J. Hyndman, “Another look at forecast-accuracy metrics for intermittent demand,” unpublished, 2006.