# AN ENHANCED CUSTOMERS' BEHAVIOUR USING RECURRENT NEURAL NETWORK (RNN)

**[1]Nwozor, B.U**

Department of Computer Science, Faculty of Natural and Applied Sciences, School of Post Graduate Studies, Ignatius Ajuru University of Education (IAUE), Rivers State, Nigeria

**Email:** bnwozor@gmail.com[1]

**[2]Onuodu, F.E**

Department of Computer Sciences, Faculty of Natural Sciences, University of Port Harcourt, Rivers State, Nigeria

**Email:** gonuodu@gmail.com[2]



**Abstract**

Predicting and Managing Customers' Behaviour poses a problem in business. This is because, the way a customer is treated determines whether he or she would return or not. Smart business organizations have centralized databases where customers' records are stored for future references. However, there is need to address the mentioned issue using machine learning approach. In this paper, an Enhanced Customers' behaviour using recurrent neural network was developed. Furthermore, the waterfall methodology was adopted in this approach, and our expected results showed an efficiency rate of 98% when evaluated with the existing system model in terms of Time Complexity (TC), Speed (S) Latency Reduction (LR) and Risk Assessment (RA). In addition, the study also recommended the importance of recurrent neural networks for machine learning and behavioral pattern recognition.

### 1.1 Introduction

Predicting and Managing Customers' Behaviour poses a difficult problem in business. This is because, the way a customer is treated determines whether he or she would return or not. Smart business organizations have centralized databases where customers' records are stored for future references. However, there is need to address the mentioned issue using machine learning approach.

Expert systems using Long Short-term Memory is an appropriate model for training datasets for clients, and further using the trained datasets for predicting customers' behaviour (Daniel, 2019)

LSTM is an artificial Recurrent Neural Network (RNN, i.e. a part of artificial intelligence that involves directed cycles in memory) architecture which is applied in the field of deep learning. LSTM devises the capability to process structured and unstructured datasets such as handwriting recognition, speech recognition, videos, etc. In addition, the lstm components include a neuron, an input gate, an output gate and a forgot gate.

The cell is the memory unit of the LSTM which processes data sequentially and keeps its hidden state through time; while the input, output and forget gates are known as the regulators in charge of information flow in the LSTM. Networks of the LSTM are also designed for grouping; processing and making predictions based on time series (series of data points indexed in time order) data (Ayako, 2016)

The LSTM input gate takes care of the new flow into the cell, the forgotten gate controls the extent to which a value stays in the cell and the output gate handles the extent to which a cell value is utilized to calculate the unit of LSTM output activation. The function that activates the LSTM gates is known as the logistic function (i.e. a logistic curve is a common "S" shape with equation) which is given as:

$f(x) = L/1+e^{-k}(x - x_0)$ Equation (1)

LSTM concepts have been analyzed by numerous researchers. Abdelhadi et al (2019) discussed a Regular Network Architecture for Internet traffic Matrix Analysis on Long Short-Term Memory. Most of the decisions that network operators make depend on how the traffic flows in their network. In any case, despite the fact that it is essential to precisely gauge traffic parameters, cur-rent switches and system gadgets don't give the likelihood to constant checking, thus web administrators can't respond successfully to the traffic changes. To deal with this problem, prediction techniques were applied to predict network parameters and can therefore react to changes in the network in near real time. The predictability of system traffic parameters is determined primarily by their statistical characteristics and a strong comparison between chronologically ordered values.

Network traffic is defined by: self-similarity, multi-scalarity, long-range dependency and strongly non-linear existence (inadequately represented by Poisson and Gaussian models). A network TM shows the volume of traffic at a certain time t between all the network nodes of origin and destination (OD) pairs. Points of Proximity (PoPs), servers or connections can be the nodes in a traffic matrix.

### 1.2 Statement of the Problem

Predicting users' behaviour in business poses a problem in business due to the following reasons:

i) lack of software model for managing customer relationship
ii) lack of an expert system that uses Recurrent Neural Network

iii) lack of the right clustering algorithm for managing data

### 1.3 Aim and Objectives of the Study

The aim of this work is to develop an Enhanced Customer's Behaviour using Recurrent Neural Network. The specific objectives of the study include to:

i) design a model that predicts customers' behaviour using Recurrent Neural Network.

ii) deploy text-based datasets to the proposed system model

iii) compare our results with other existing models.

### 1.4 Significance of the Study

This paper could be beneficial to Entrepreneurs and Software Developers.

a) How the beneficiaries would benefit from the study:

**i) Entrepreneurs:**

The proposed model will enable entrepreneurs grow their business

**ii) Software Developer:**

Software developers will use the relevant information from the study to update existing software to predict users' behavior

## 2.0 General Concept

### 2.1 Recurrent Neural Networks

Recurrent Neural Networks (RNNs) are part of Artificial Intelligence that is deployed to solve the problem of Natural Language Processing (NLP). Furthermore, RNNs are used for predicting patterns and sequential characteristics.

### 2.2 Overview of Long Short-Term Memory (LSTM)

Long Short-Term Memory units (LSTM) are part of a recurring artificial neural structure. Recurrent neural networks are designed to use certain types of artificial memory processes that can help these artificial intelligence programs to imitate human thinking more efficiently. The recurrent neural network uses long-term memory blocks to provide meaning for how inputs are obtained and outputs generated by the algorithm.

The long-term memory block is a dynamic unit with different components such as weighted inputs, activation functions, previous block inputs and potential outputs. The unit is called a long-term memory block, because the program uses a short-term memory-based structure to create longer-term memory. For example, these systems are often used in the processing of natural languages. The recurrent neural network uses the long-term memory blocks to take a specific word or phoneme and interpret it in a sequence in the sense of others, where memory can be useful to sort and categorize these types of inputs.

Generally speaking, LSTM is an agreed and general principle in the creation of recurrent neural networks. People don't start thinking every second from scratch. This is not possible for conventional neural networks, and it seems like a significant weakness. For instance, imagine an individual wanting to classify what kind of event occurs at every point in a film. It is uncertain how a conventional neural network might utilize its reasoning to inform later events about past events in the film.

Reoccurring neural networks are dealing with the problem. They are networks that have loops in them, which allow information to persist. In addition, a LSTM's components include a cell, an input gate, an output gate, and a forgotten gate. The cell is the memory unit of the LSTM

which processes data sequentially and keeps its hidden state through time; while the input, output and forget gates are known as the regulators in charge of information flow in the LSTM. Networks of the LSTM are also designed for grouping; processing and making predictions based on time series (series of data points indexed in time order) data. The LSTM input gate handles the new flow into the cell, the forget gate handles the degree to which a value stays in the cell, and the output gate handles the extent to which a value is used in the cell to measure the LSTM unit output operation.

### 2.3 Related Works

**Yahaya et al (2017)** researched on mathematical modeling for population projection and management: A case study of Niger State. The work developed a mathematical model to predict population figures of Niger State for the period of 20 years using 2006 census figures and MATLAB software. The result of the work showed that the applied parameter can be used for the population and any given region which helps to check unwanted population increase or decrease. However, they could not compare their parameter with other existing models.

**Abdelhadi et al (2019)** looked at a long-term recurring neural network memory framework for prediction of network traffic matrix. The work suggested a framework for LSTM RNN to predict traffic matrix (TM) in large networks. The results of the work showed that their designed model converged quickly and predicted the performance of state-of - the-art TM for relatively small models. However, they could not illustrate how their model accepts memorized real-time datasets.

**James et al (2016)** looked at new tools for predicting economic growth using machine learning: a guide for theory and policy. They developed a model for machine learning to forecast growth and solve economic problems. The result of the work showed that their developed model will complement economic growth prediction for proper planning. However, they could not further interpret the generated charts from their model.

**Osonde et al (2017)** researched on an artificial intelligence / machine learning angle on social simulation. The work examined the current state of the art in data infrastructure and artificial intelligence approaches that could be useful for social and behavioral modeling. The results of the work showed key identification of challenges on social simulation through the development of three models. However, they could not show the interface of their models to real-life problem solving.

**Julian et al (2018)** researched on reading China: predicting policy change with machine learning. The work developed a quantitative indicator of the Chinese government's policy priorities over a long period of time called the Policy Change Index (PCI) for China. The result of the work showed that the developed PCI is flexible and efficient since it does not require the understanding of Chinese Text, which also suggests a wide range of other software. However, they could not convince the government for adequate fund for future improvement on the work.

**Samir et al (2017)** researched on population projection by age, sex and educational attainment in rural and urban region of 35 provinces of India, 2011 – 2101. The work explained methodological approaches and technical details of how population projections are conducted in India. The result of the work showed that their methodological approach provides essential information for planning and implementing government policies. However, they could not perform comparative analysis of their methodology with other existing methodologies.

**Marinko et al (2017)** looked at population and economic growth: a review essay. The research put the topic in a historical, empirical and structural context to identify population as an important factor in socio-economic prosperity, often measured in terms of economic growth. The result of the work showed that population effect should be measured only in terms of population growth, measured by the increase in number of inhabitants. However, they could not practically implement the reviewed study.

**Ida et al (2018)** looked at identification of population growth and distribution based on urban zone functions. The work analyzed zones with the highest population in Indonesia. The result of the work showed that the zones with the highest population growth and accommodate the highest population increase. However, they could not illustrate any population model-type used in the work.

**Markus et al (2013)** looked at income and population growth. The study estimated the impact on increasing population of crises on national income, credibly exogenous and difficult to be induced by technological change. The result of the work showed that there is significant relationship between oil induced income growth and population growth. However, they could not illustrate any population model type used in the work.

**David (2011)** looked at population dynamics in India and Implication. The work discussed theoretical and empirical literatures on the effect of demographics on labor supply, savings, and economic growth in India. The result of the work showed that failure to take advantage of the opportunities inherent in demographic change can lead to economic stagnation. However, they could not practically implement the reviewed study.

**Samir et al (2010)** researched on projection of population by level of educational attainment, age and sex for 120 countries for 2005 – 2050. The work used demographic multi-state and cohort-component methods to produce projections for 120 countries. The result of the work showed the possible range of future educational social planning at national and international levels, and to the assessment of the feasibility of international education goals. However, they could not implement with a population prediction and forecast model.

**Folorunso et al (2010)** researched on population prediction using artificial neural network. The study employed an artificial neural network for population prediction which manages incomplete and contradictory data usually encountered in the use of mathematical and demographic models when forecasting the population. The result of the work showed that the applied artificial neural network percentage accuracies ranged between 81.02% and 99.15% which implies better performance than the demographic model. However, they could not show how they clustered and stored real-time datasets used in their implementation.

**Brian et al (2010)** looked at a guide to global population projections. The work presented a guide to such projections aimed at researchers and educators. The result of the work showed that new approach to demographics proved to be efficient. However, they could not practically implement the discussed study.

**Bruno et al (2016)** researched on global population growth technology and Malthusian constraints: a qualitative growth theoretic perspective. The work estimated a two-sector Schumperian growth model with endogenous population and finite land reserves to study long run of global population. The result of the work showed that the Schumperian growth

model is significant to population growth. However, they could not compare the Schumperian growth model with any other existing model.

**Shilpa et al (2014)** analyzed Indian population growth and projected predictions. The works estimated India's population growth from 2001 to 2012 using the logistic design strategy and comparing it with India's actual population for the same period. The result of the work showed an insight to the changing trends with respect to India's population growth. However, they could not practically implement the discussed issue.

**Huabin et al (2015)** researched on a study of two models of population growth and an assessment of influencing factors population growth in China. The work reviewed two population growth models and tries to find a proper way to explain and predict population growth models and tries to find a proper way to explain and predict population growth. The results of the work showed that two factors, the degree of urbanization and the sex ratio, have a major impact on population growth in China. However, they could not show how they clustered and stored real-time datasets used in their implementation.

**Yemi et al (2016)** discussed on Actuaries interested in population issues. The work outlined how actuarial work is affected by demographic issues and suggests reasons why other organizations interested in population issues talk to actuaries. The result of the work showed that close details to demographics would improve population prediction growth. However, they could not practically implement the work.

**Nwosu (2014)** looked at the effects of population growth on economic growth in Nigeria. The work examined the time series role of population growth in Nigeria and how economic growth is achieved by population growth. The result of the work showed that there is a sustainable long run equilibrium relationship. However, they could not implement the discussion with a model.

**Quamrul (2008)** Reviewed on the dynamics of Malthusian's population: theory and evidence. Theory and Evidence. The work checks the presence of dynamics of the Malthusian population in the pre-industrial revolution period. The result of the work showed that the application of the Malthusian theory had an increased population size over time. However, they could not practically implement the reviewed study.

**Pamela (2017)** researched on Utah's Long-Term Demographic and Economic Projection Summary. The result of the work showed the importance of the Utah's Long-Term Demographic model to population growth. However, they could not practically implement the discussed work.

**Oramah (2014)** looked at the effects of population growth in Nigeria. The study addressed the use of double-time growth analysis to understand the need for population control in Nigeria and the potential danger that may emerge from the continued neglect of environmental issues posed by environmentalists and demographers in Nigeria and around the world. The result of the work showed that it is important for a country to invest on tools used for population prediction and forecast. However, they could not practically implement the discussed issues.

**David et al (2010)** looked at implications of population ageing for economic growth. The work looked at media to raise awareness about the consequences for economic growth of population ageing. The results of the research showed that; important factors that indicate the effect of population ageing will not hinder the rate of economic growth in developing countries significantly. However, they could not implement with a population prediction.

**Doris (2017)** researched on overpopulation and the impact on the environment. The work focused on the problem of overcrowding and its environmental impact. The result of the work showed that the solution to overpopulation lies on the efforts of national institutions to implement policies that will correspond to guidelines given by the best of the global community. However, they could not practically implement the discussed study.

**Ulrich et al (2010)** looked at the population history of Germany: research strategy and preliminary results. The work presented an aggregate reconstruction of the population of Germany from the seventeenth to 1840, when all German states began to be fully covered by official statistics. The result of the work showed that cumulative rates of national increase are broadly consistent with independent estimates of population growth in Germany. In addition, the result also consistently suggested the presence of both preventive and the positive check during the eighteenth century. However, they could not implement with a population prediction and forecast model.

**Musa (2015)** researched on an Econometric model on population growth and economic development in India: an empirical analysis. The work has developed an econometric model that takes India as a case study. The result of the work showed that there is need for government to take population as virtue by investing more resources in human capital development through quality education. However, he could not compare the econometric model with any other existing model.

**Adindu (2012)** researched on consequences of population growth on agricultural production in Obingwa Local Government Area of Abia State. The study discussed on various setbacks on uncontrolled population growth. The result of the work showed that there is a significant effect of population growth and food production based on land pattern systems in the locality. However, they could not implement with a population prediction model.

**Frank et al (2007)** discussed about population, labor force and long-term economic growth. The work explored related issues and drew out implications for Canada's economic growth prospects. The result of the work showed that proper economic planning is vital for population forecast. However, they could not practically implement the discussed study.

**Atsaka (2017)** researched on population predictions for Japan (2017). The work looked at factors needed for successful population projection and forecast. The result of the work showed that the application of demographics will go a long way in complementing a successful population forecast. However, they could not implement with a population prediction model.

**Eduardo (2007)** looked at failure trends in a large disk drive population. "Over 90 percent of all new information produced in the world is estimated to be stored on magnetic media, according to the work." The result of the work showed that there is need for urgent backing up of related populated files. However, there was no adequate comparative analysis of the work with other population systems.

**Ogochukwu et al (2019)** researched on using data mining techniques for census analysis to give geo-spatial distribution of Nigeria. The work discussed effort towards harnessing the power of data mining technique to develop mining model applicable to the analysis of census data that could uncover some hidden patterns to get their geo-spatial distribution. The result of the work showed that the applied mining model could provide government with the

intelligence for strategic planning and better policy formulation.

Furthermore, they implemented with a Geographic Information System (GIS) which is only applied in spatial distribution. However, they could not use their model to develop a decision support system using Long Short-Term Memory (LSTM). In addition, the absence of the LSTM from their model may negatively affect decision making after the census analysis and focus.

**Dang (2019)** looked at Customer Churn Prediction in Computer Security Software. Realistically, the churn rate is a calculation of the amount of persons or objects move over a precise duration from a collective group. It has become one of two primary factors that influence the level of stable customers that a company will promote.

The author tried to provide solution to problems associated with churn prediction. Furthermore, the result of the work also showed the importance of handling issues of churn prediction.
The author did not implement their's with an expert system such as Recurrent Neural Network

### 3.1 Analysis of the Existing System

The existing system as developed by Dang (2019) is a standard Churn Rate Prediction Model (CRPM) as illustrated in figure 3.1. The CRPM enables the measurement of the number of people or objects move from a collective group over a specific period of time. One of the two main factors deciding the amount of steady-state customers a company serves.

#### 3.1.1 Advantages of the Existing System

i) The Existing System is scalable in the management of customers' Data

ii) The Churn Model is accurate for Prediction and forecast of customers' behavior.

iii) The Churn Model is also efficient in the processing of Customers' data

#### 3.1.2 Disadvantages of the Existing System

i) lack of an enhanced software model for managing customer relationship

ii) lack of an expert system that uses Recurrent Neural Network

iii) lack of the right clustering algorithm for data management

#### 3.1.3 Advantages of the Proposed System

i) An enhanced approach to customer behaviour prediction using Recurrent Neural Network

ii) The proposed system is an expert system that facilitates both Supervised and Unsupervised Learning processes.

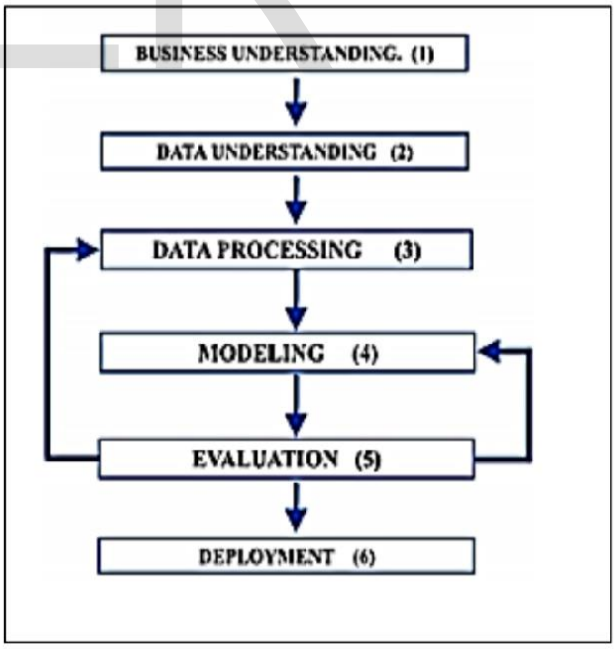


**Figure 3.1:** Churn Prediction Model Architecture for Customer behavior prediction (Existing System Architecture) (**Source:** Dang, 2019)

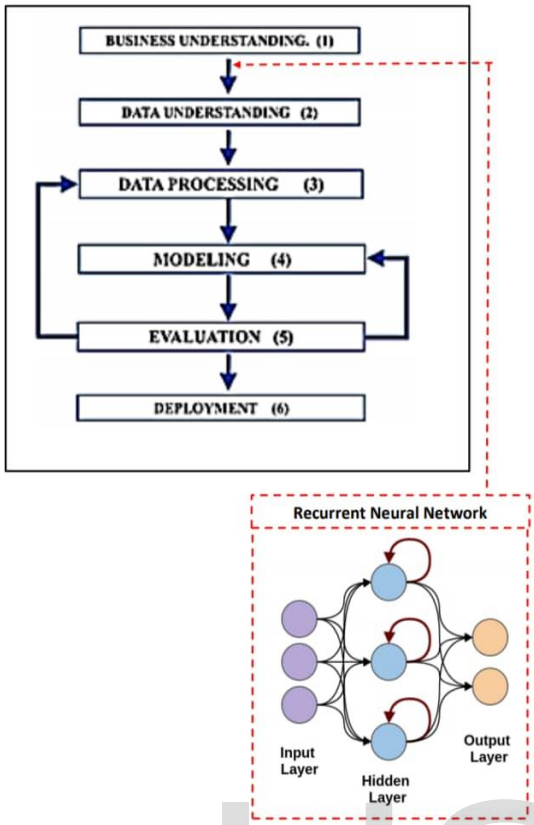


**Figure 3.2:** An Improved Churn Prediction Model Architecture for Customer behavior prediction (Proposed System)

| SN | PERFORMANCE EVALUATION CRITERIA (PEC) | ASSESSED PERFORMANCE RATE (%) |
|---|---|---|
| 1. | Time Complexity (TC) | 21 |
| 2. | Life Cycle Assessment (LCA) | 13 |
| 3. | Benchmarking (B) | 7 |
| 4. | Multi-Criteria Decision Making (MCDM) | 22 |
| 5. | Risk Assessment (RA) | 4 |
| 6. | Cost Benefit Analysis (CBA) | 8 |
| 7. | Speed (S) | 10 |

### 3.3 Methodology

The waterfall model is a breakdown of project activities into linear sequential phases, where each phase depends on the previous one's results and corresponds to a task specialization. The approach is typical of certain engineering design areas.

## 4.0 Results

Table 1.0: Dang (2019): Customer Churn Prediction Model in Computer Security Software:

**Assessed Parameters Summary:**

| | | |
|---|---|---|
| TC | = | 21 |
| LCA | = | 13 |
| B | = | 7 |
| MCDM | = | 22 |
| RA | = | 4 |
| CBA | = | 8 |
| S | = | 10 |
| **TOTAL** | **=** | **85%** |

Table 2.0: Blessing (2019): An Improved Customer Churn Model using Recurrent Neural Network (RNN)

| SN | PERFORMANCE EVALUATION CRITERIA (PEC) | ASSESSED PERFORMANCE RATE (%) |
|---|---|---|
| 1. | Time Complexity (TC) | 25 |
| 2. | Life Cycle Assessment (LCA) | 19 |
| 3. | Benchmarking (B) | 7 |
| 4. | Multi-Criteria Decision Making (MCDM) | 22 |
| 5. | Risk Assessment (RA) | 6 |
| 6. | Cost Benefit Analysis (CBA) | 9 |
| 7. | Speed (S) | 10 |

**Assessed Parameters Summary:**

| | | |
|---|---|---|
| TC | = | 25 |
| LCA | = | 19 |
| B | = | 7 |
| MCDM | = | 22 |
| RA | = | 6 |
| CBA | = | 9 |
| S | = | 10 |
| **TOTAL** | **=** | **98%** |

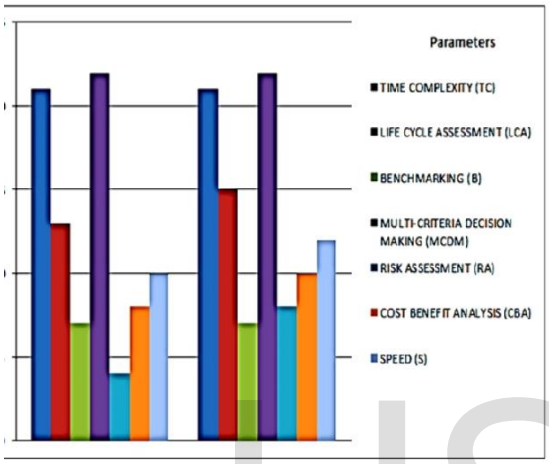


Figure 3.3: Comparative Analysis Chart of the Proposed System against the Existing System

## 5.1 Conclusion

The paper looked at Predicting and Managing Customers' Behaviour using a Recurrent Neural Network. Traditionally, it has been hard to train the recurrent neural networks. The Long Short-Term Memory, or LSTM, network may be the most successful RNN as it overcomes the problems of training a recurring network and in turn has been widely utilized in several applications.

## References

Abdelhadi A. (2019), A Long Short-Term Memory Recurrent Neural Network Framework for Network Traffic Matrix Prediction, *https://arxiv.org/pdf/1705.05690.pdf*

Adindu K. (2012), Consequences of population growth on Agricultural production in Obingwa Local Government Area of Abia State Nigeria, *International Journal of Computer Applications (IJCA) 4(8), 7 – 17*

Akinwale A. (2010), Population prediction using artificial neural network, *African Journal of Mathematics and Computer Science (AJMC), 3(8), 155 – 162*

Atin B. (2016), New tools for predicting economic growth using machine learning: A Guide for theory and Policy; *Research Gate Publications, https://www.researchgate.net/publication/291827961*

Atsaka O. (2017), Population Projections for Japan: *International Journal of Computer Applications (IJCA) 2(7), 4 – 8*

Ayako M. (2016), A Long Short-Term Memory Neural Network Architectures for generating music and Japanese lyrics, *International Journal of Computer Applications (IJCA), 4(8), 14 – 22*

Bart D. (2018), Identification of Population growth and distribution based on urban zone functions, an International Article published at sustainability *2018, 10, 930; doi: 10.3390/su10040930 https://www.mdpi.com/journal/sustainability*

Bilal B. (2010), Population projection by age, sex and educational attainment in rural and urban regions of 35 Provinces of India, 2011 – 2101; *International Institute for Applied Systems, Analysis Schlossplatz 1, A-2361 Laxenburg, Austria, www.iiasa.ac.at*

Boniface E. (2019), Using Data Mining Technique for Census Analysis to give Geo-Spatial Distribution of Nigeria, *IOSR Journal of Computer Engineering (IOSR-JCE), 14(2), 1 – 5*

Brian C. (2010), A guide to global population projections, *demographic research, 4(8), 203 – 228*

Bruno L. (2016), Global population growth, technology and Malthusian constraints: A quantitative growth theoretic perspective, *Grantham Research Institute on Climate change and environment, working paper, No. 161*

Byron S. (2007), Population, Labor force and long-term Economic growth, *International Journal for social Sciences (IJSS) 4(7), 1 – 32*

Dang B. (2019), Customer Churn Prediction in Computer Security Software, *International Publication, https://cdn/.sph.havard.edu/wp-content/uploads/sites/1288/2013/10/PGDA_WP_65 .pdf*

David B. (2010), Implications of population aging for economic growth, *Oxford Review of Economic Policy, 26(4), 583 – 612*

David C. (2010), Implications of population aging for economic growth, *Oxford Review of Economic Policy, 26(4), 583 – 612*

Deborah B. (2010), A guide to global population projections, *demographic research, 4(8), 203 – 228*

Doris B. (2017), Overpopulation and the impact on the environment, dissertations thesis and capstone projects, *https://academicworks.cuny.edu/cgi/viewcontent.cgi?article=2929$context=gc_etds*

Eduardo P. (2007), Failure trends in a large disk drive population, appears in the *proceedings of the 5th USENIX conference on file and storage technologies (FAST'07), February 2007*

Fadi A. (2011), Annals of the University of Craiova, Mathematics and Computer Science Series, *38(2), 18 – 25, ISSN: 1223 - 6934*

Folorunso O. (2010), Population prediction using artificial neural network, *African Journal of Mathematics and Computer Science (AJMC), 3(8), 155 – 162*

Frank D. (2007), Population, Labor force and long-term Economic growth, *International Journal for social Sciences (IJSS) 4(7), 1 – 32*

Georg F. (2010), The population history of Germany: research strategy and preliminary results, MPIDR *WORKING PAPER WP, 2010 – 035 DECEMBER 2010, https://www.demogr.mpg.de/papers/ working/wp_2010_035.pdf*

Guy P. (2019), A Long Short-Term Memory Recurrent Neural Network Framework for Network Traffic Matrix Prediction, *https://arxiv.org/pdf/1705.05690.pdf*