IJDI-ERET

INTERNATIONAL JOURNAL OF DARSHAN INSTITUTE ON ENGINEERING RESEARCH AND EMERGING TECHNOLOGIES
Vol. 12, No. 2, 2023
www.ijdieret.in

(Research Article)

Long Short-Term Memory (LSTM) Network for Efficient Classification of Prostate Cancer and Prostate Enlargement

F. O. Aranuwa[1*], O. J. Popoola[2]

[1*,2]*Department of Computer Science, Adekunle Ajasin University, Akungba – Akoko, Ondo State, NIGERIA*

Abstract

*Long Short-Term Memory (LSTM) network, unlike traditional feed forward neural networks is capable of processing complex and sequential data with minutest error. This classifier possessed feedback connections and its activation patterns in the network change once per time-step, and this characteristics make its algorithm ideal for classification and prediction problems. This study designed a model for efficient classification and prediction of prostate cancer and prostate enlargement using the LSTM network. Prostate cancer and prostate enlargement, also known as b*enign prostatic hyperplasia *(BPH) relatively shared similar risk factors and symptoms that makes their classification and prediction a complex task. Researchers have made efforts to address these issues, however the sequence prediction problems still draw more research attentions. The inability of the classical methodologies to learn and synthesize underlying relationships among the symptoms, always results in low predictive accuracy and high error rates. Data attributes for the study were sourced from selected medical institutions in the South Western, Nigeria. A total of 1,149 datasets were collected, prepared and made suitable for modeling through transformation and normalization processes. Samples were categorized based on specific attribute determinants, whose results could be Cancer, BPH and Normal (1, 2, 0). The dataset was trained and tested on 80/20 ratio on Anaconda platform. The experiment was set to run on 70 epochs, to determine the accuracy and ascertain the model's efficiency and practicality. Results shown that the model correctly classified and predicted the status of the malaise with 96.7% accuracy over the benchmark of 92%. The model, haven suitably addressed the sequence prediction problems, which is an improvement on the previous models considered, is recommended for deployment in various health care sector for efficient classification and prediction of the disease.*

Keywords: Long Short-Term Memory, Accuracy, Prostate Cancer and Enlargement, Deep Learning, Mortality Rate.

1. Introduction

Globally, the rising incidences and mortality rate from prostate cancer and prostate enlargement represents a significant and growing threat to human life. The syndrome are common in men, and relatively shared similar risk factors, that makes their classification and prediction a complex task [1];[2]. Researchers and medical experts have made efforts to proffer solution to these issues, however the inability of these methodologies to learn and synthesize underlying relationships among the symptom, remains a challenges [3].

The economic burden of these maladies is substantially growing, and diagnosis of new cases is increasing at an exponential rate globally [4]. Therefore, the need for an intelligent classifier enthused this study.

The study designed a model for efficient classification and prediction of prostate cancer and prostate enlargement using Long Short-Term Memory (LSTM) network. Common indicators of prostate cancer and enlargement includes trouble in urinating, decreased force in the stream of urine, passing urine often especially at night, blood in the urine and semen, bone pain and erectile dysfunction. Major risk factors includes age, genetic and high-fat diet [5]. The disease's risk increases with age, particularly after the age 50, and for some men, genetic factors may put them at higher risk. According to [6], the syndrome is considered an indolent age-related disease, with incidences increasing from 1 in 350 men under the age of 50, to 10 in 52 men ages 50–59, the incidence rate is nearly 60% in men over the age of 65 years.

Figure 1 and Figure 2 depicts the structure of prostate cancer and prostate enlargement (BPH), respectively.

Generally, the syndrome goes through two main growth periods as a man ages. The first occurs early at puberty, when the prostate doubles in size. The second phase of growth begins around age 25 and continues during most of a man's life. Benign prostatic hyperplasia often occurs with the second growth phase.

[*]*Corresponding Author: e-mail: felix.aranuwa@aaua.edu.ng, Tel-+2347031341911,*
ISSN 2320-7590 (Print) 2583-3863 (Online)

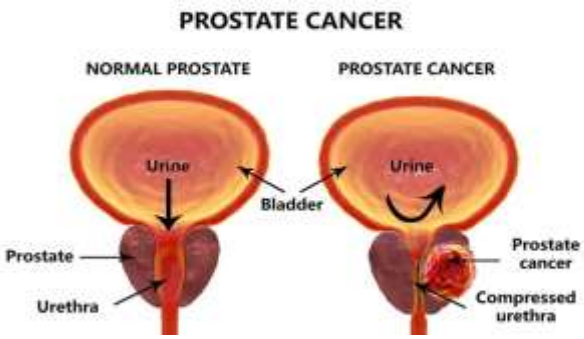


Figure 1. Normal prostate and prostate cancer [7]

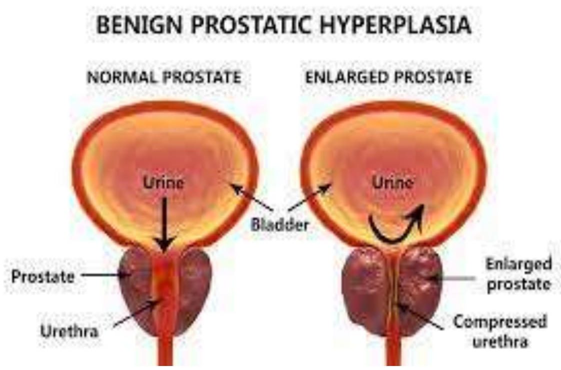


Figure 2: Normal prostate and prostate enlargement [7]

2. Long Short-Term Memory (LSTM) Network

Long Short-Term Memory (LSTM) network is an artificial neural network used in the fields of artificial intelligence and deep learning. LSTM, unlike typical feed forward neural networks, has feedback connections capable of processing complex and sequential data with minutest error [8]. It is a type of recurrent neural network (RNN) that have achieves state-of-the-art results on challenging prediction problems especially in sequence prediction problems. LSTM is very intelligent and was explicitly designed to avoid the long-term dependency problems. It is well suited to learn from experiences that have very long time lags [9]. The units are used as building units for the layers of a RNN, often called an LSTM network. It enable RNNs to remember inputs over a long period of time because it contain information in a memory, much like the memory of a computer. LSTMs effectively can store and access long-term dependencies using a special type of memory cell and gates. Figure 3 and Figure 4 show the LSTM diagram and structure of LSTM Networks, respectively.

According to [10], LSTMs use a series of 'gates' which control how the information in a sequence of data comes into, is stored in and leaves the network. There are three gates in a typical LSTM; forget gate, input gate and output gate. These gates can be thought of as filters and are each their own neural network. The cell state is represented by $C_{(t-1)}$ and $C_{(t)}$ for the previous and current timestamps, respectively. $H_{(t-1)}$ represents the hidden state of the previous timestamp and H_t is the hidden state of the current timestamp. The memory can be seen as a gated cell, meaning that the cell can decides whether or not to store or delete information, based on the importance assigns to the information. Assigning of importance happens through weights, which are also learned by the algorithm. Figure 5 depicts the LSTM cell.

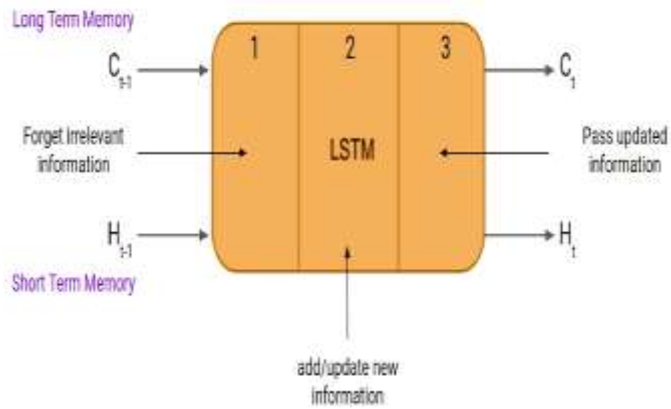


Figure 3. LSTM diagram [8]

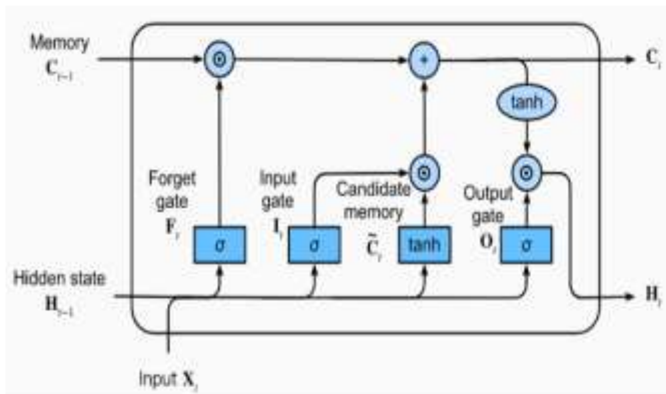


Figure 4. The structure of LSTM networks [10]

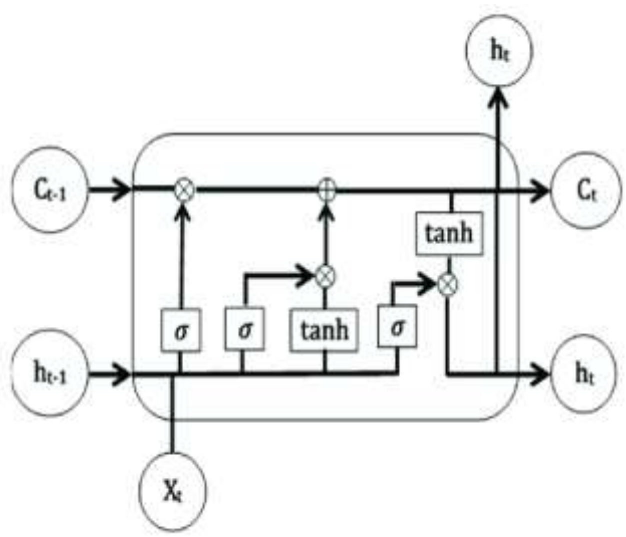


Figure 5. LSTM cell [11]

International Journal of Darshan Institute on Engineering Research and Emerging Technologies
Vol. 12, No. 2, 2023, pp. 01-06

3. Methodology

The conceptual diagram of the model is depicted in Figure 6 consisting of five (5) major components, namely: Data acquisition and pre-processing, model building, classification output, and predictive mechanism. The processes are discussed in section 3.1.

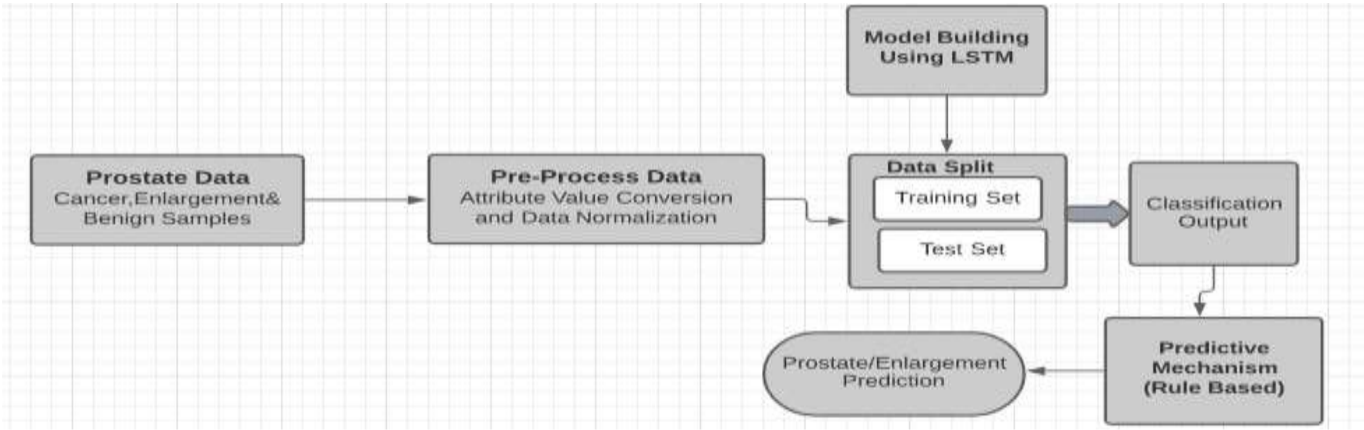


Figure 6. The architecture of the classification model

3.1 Data acquisition and pre-processing: Data attributes for the study were sourced from selected medical institutions in the South Western, Nigeria after due ethical approval. A total of one thousand, one hundred and forty nine (1,149) datasets were acquired. The data samples were categorized based on specific attribute determinants, whose results can be Cancer, BPH and Normal (1, 2, 0). The data attributes considered include: prostate weight, seminal vesicle invasion (svi), capsular penetration (cp), gleason score grade, prostate-specific antigen (psa) level, patient's age, and cancer volume (cv). The datasets were prepared, made suitable for modeling through transformation and normalization processes, and were trained and tested on 80/20 ratio. Anaconda, an open source Python environment for data discovery and analytic platform was used for the experiment. The experiment was set to run on 70 epochs, to determine the accuracy and ascertain the model's efficiency and practicality. The data format is presented in Table 1.

Table 1. Sample data

ID	Weight	Svi	cp	Gleason score grade	psa	Age	Classified status
1	97	0	0	3	42	87	1
2	88	0	0	4	241	70	1
3	125	0	0	0	333	60	2
4	98	0	0	0	25	74	2
5	42	0	0	0	20	77	0
6	67	0	0	0	45	67	2
7	69	0	0	0	3	78	2
8	51	0	0	0	0	60	0
9	103	0	0	0	27	77	2
10	45	0	0	0	33	67	0
11	56	0	0	4	78	70	1
12	119	0	0	3	4	79	2
13	137	0	0	2	4	55	2
14	111	0	0	0	567	89	2
15	51	0	0	0	65	91	1
16	146	0	0	0	19	69	2
17	67	0	0	2	3	40	1

Mathematically, the model is presented as follows: Given a set of samples containing instances of assigned class label as either prostate cancer, benign or normal prostate as captured in equation (1).

$$D = \{(x_i, y_i, z_i)\}_{i=1}^{n} \qquad (1)$$

Where D represents the prostate dataset, x_i is an input for the $i-th$ sample described by set of attributes $x_{i1}, x_{i2}, x_{i3}, \dots, x_{iq}, y_i, z_i \in$ {prostate cancer, benign, and normal prostate}, and n represents the total number of data samples. D was normalized using min-max normalization method in this work as represented in equation (2).

$$x' = \frac{x - min_x}{max_x - min_x} \quad (2)$$

x' represents the new value, x denotes the observed value. The dataset were later partitioned into Training set (T) and Test set (T_s). T is used to train and computes a mapping from input data x_i to output class label y_i at time t by calculating network unit activations using the equation (3) – (9)

$$f_t = \sigma(W_f[x_t, h_{t-1}] + b_f) \quad (3)$$

$$i_t = \sigma(W_i[x_t, h_{t-1}] + b_i) \quad (4)$$

$$z_t = \tanh(W_z\ [x_t, h_{t-1}] + b_z) \quad (5)$$

$$c_t = (z_t * i_t + f_t * c_{t-1}) \quad (6)$$

$$o_t = \sigma(W_o\ [x_t, h_{t-1}] + b_o) \quad (7)$$

$$h_t = \tanh(c_t) * o_t \quad (8)$$

$$\hat{y} = h_t \quad (9)$$

Where σ represents the sigmoid function, I, f, and o are input, forget, and output gate respectively, W represents the weight parameter, b represents bias, h_t represents the output vector of the LSTM unit which is the predicted class label $\hat{y}$. The star ($*$) represents the element-wise product called Hadamard product, c_t represents cell state or updated memory, c_{t-1} represents the previous cell state from the last LSTM unit, and h_{t-1}represents output of last LSTM unit. LSTM uses backward propagation through time (*BPTT*) as training algorithm to minimize classification loss L captured in equation (10) as weights are adjusted when the network propagates backward through gradient descent approach presented in equation (11)

$$L(y, \hat{y}) = \frac{(y - \hat{y})^2}{2} \quad (10)$$

$$\delta = dL/\, dW \quad (11)$$

$\hat{y}$ represents the predicted output, and y is actual output. δ is the partial derivative of model's loss L with respect to its weight W. The performance of the model was evaluated using standard performance metric of accuracy, precision and recall to ascertain the model's efficiency.

4. Experimental Setup

The dataset for the experiment was fed into an Anaconda platform installed on Processor Intel(R) Core(TM) i7-8550U CPU system at 1.80 GHz, 1992 MHz, 4 Core(s), 8 Logical Processor(s) system with 1TB HDD and 8GB RAM for the purpose of the work. The software is an open-source application for management of multiple Python versions on one computer. Anaconda provides a large collection of highly optimized, commonly used data science libraries to carry out experiments. It has a powerful IDE with advanced editing, interactive testing, debugging and introspection features called Spyder.

5. Results Findings

Experimental results showed that the model correctly classified and predicted the status of the malaise with 96.7% accuracy over the benchmark of 92%. Distinct factor that differentiate BPH from prostate cancer is the prostate weight and decreased force in the stream of urine as a result of compressed urethra. The True Positive and False Positive rates stand at 0.967 and 0.3, respectively. This implies that the model can correctly classified and predicted prostate cancer and prostate enlargement in a person with minutest error. Comparison of the experimental results in a bar chart is presented in Figure 7.

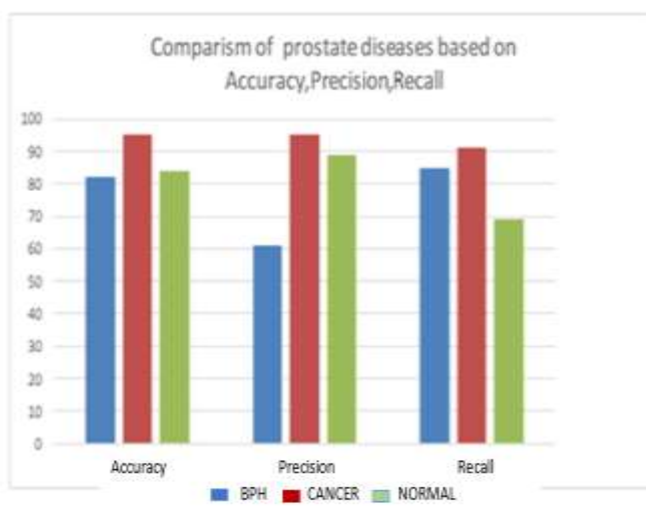


Figure 7. Comparison of the experimental results (accuracy, precision and recall)

The accuracy and error evaluation graphs are presented in Figure 8 and Figure 9.

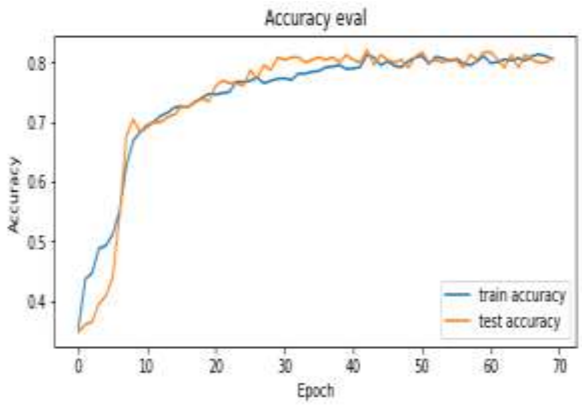


Figure 8. Accuracy evaluation

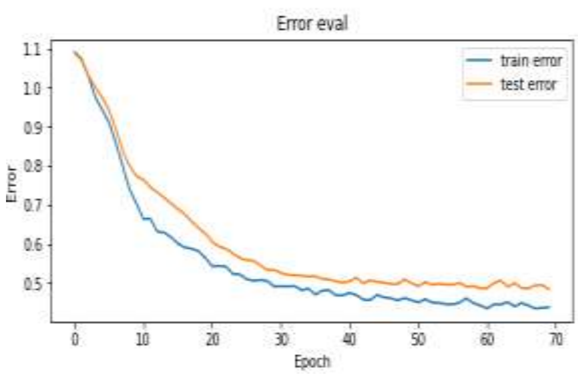


Figure 9. The error evaluation

6. Conclusions

The quest for an efficient technique for accurate classification of prostate cancer and prostate enlargement motivated this study. The syndrome relatively shared similar risk factors, which makes their classification and prediction a complex task. Additionally, the inability of existing methodologies to learn and synthesize underlying relationships among the symptom, remains a challenges. Hence, this study designed a classification model for proper classification and prediction of the syndrome using Long Short-Term Memory (LSTM) network. The datasets acquired for the work were sourced from selected medical centres, prepared and made suitable for modeling through transformation and normalization processes. The dataset were trained and tested on 80/20 ratio. The experiment was set to run at a considerable epochs to determine the accuracy and ascertain the model's efficiency and practicality. Experimental results showed that the model correctly classified and predicted the status of the malaise with 96.7% accuracy. The model haven proved to be an improvement over the previous models, is recommended for deployment for early prediction or and detection of prostate cancer and prostate enlargement in our various health care system. This is believe will help to prevent the growing threat and mortality rate of these syndromes. However, the research is open to future work, especially in the area of sourcing for more indicators and datasets.

References

[1] P. I. Ellsworth, "Prostate Cancer: Symptoms, Survival Rate, Treatment, Causes & Stages," *MedicineNet*, 2021. https://www.medicinenet.com/prostate_cancer/article.htm (accessed Jul. 11, 2023).

[2] P. Rawla, "Epidemiology of Prostate Cancer," *World Journal of Oncology*, vol. 10, no. 2, pp. 63–89, 2019, doi: 10.14740/wjon1191.

[3] N. Sharma, M. Singh, and Y. Gupta, "DISEASE PREDICTOR: ALGORITHMS TO REDUCE THE FAULTY PREDICTIONS," *International Journal of Engineering Applied Sciences and Technology*, vol. 5, no. 1, pp. 174–179, May 2020, doi: 10.33564/IJEAST.2020.v05i01.025.

[4] "Cancer Facts & Figures 2022." https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2022.html

[5] C. Printz, "Men who use assistive reproduction appear to have higher prostate cancer risk," *Cancer*, vol. 126, no. 6, pp. 1148–1149, Mar. 2020, doi: 10.1002/cncr.32780.

[6] "Enlarged prostate: MedlinePlus Medical Encyclopedia." https://medlineplus.gov/ency/article/000381.htm

[7] "Theresa May to launch £75m drive against prostate cancer," *News-Medical.net*, Apr. 10, 2018. https://www.news-medical.net/news/20180410/Theresa-May-to-launch-c2a375m-drive-against-prostate-cancer.aspx (accessed Jul. 26, 2023).

[8] Jason Brownlee, "A Gentle Introduction to Long Short-Term Memory Networks by the Experts," *Machine Learning Mastery*, May 24, 2017. https://machinelearningmastery.com/gentle-introduction-long-short-term-memory-networks-experts/ (accessed Jul. 26, 2023).

[9] S. Shekhar, "What is LSTM for Text Classification?," *Analytics Vidhya*, Jun. 14, 2021. https://www.analyticsvidhya.com/blog/2021/06/lstm-for-text-classification/ (accessed Jul. 26, 2023).

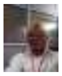
[10] R. Dolphin, "LSTM Networks | A Detailed Explanation," *Medium*, Dec. 12, 2021. https://towardsdatascience.com/lstm-networks-a-detailed-explanation-8fae6aefc7f9

[11] R. N. Toma, M. N. Hasan, A.-A. Nahid, and B. Li, "Electricity Theft Detection to Reduce Non-Technical Loss using Support Vector Machine in Smart Grid," in *2019 1st International Conference on Advances in Science, Engineering and Robotics Technology*

(ICASERT), Dhaka, Bangladesh: IEEE, May 2019, pp. 1–6. doi: 10.1109/ICASERT.2019.8934601.

Biographical notes

Dr. F. O. Aranuwa is a senior faculty member in the Department of Computer Science, Adekunle Ajasin University, Akungba - Akoko, Nigeria· He has held administrative positions as Head of Department of Computer Science. He possessed B.Sc (Hons), M.Sc and Ph.D in Computer. Science in year 1997, 2010 and 2015 respectively. His research interest covers; Biometrics, System Modelling, Data Science, Machine Learning and Human Computer Interaction.



O. J. Popoola is an M.Sc Student at the Department of Computer Science, Adekunle Ajasin University, Akungba, Nigeria. His current research interest covers Data Science and Biometrics.