**Development of a Part of Speech Tagger for Yoruba Language using Deep Neural Network**

Chukwuemeka Christian UGWU

*The Federal University of Technology, Akure, Nigeria*

Abisola Rukayat OYEWOLE

*The Federal University of Technology, Akure, Nigeria*

Olugbemiga Solomon POPOOLA

*Department of Mathematics and Computer Science, University of Medical Sciences, Ondo, Nigeria*

Adebayo Olusola ADETUNMBI

*Bowen University, Iwo, Nigeria*

Ayo ELEBUTE

*Bowen University, Iwo, Nigeria*

ABSTRACT

The pursuit of advancing Yoruba language in the realm of technology has underscored the necessity for an efficient foundational natural language processing (NLP) tool, notable the part-of-speech (POS) tagger. POS tagging serves as the building block to numerous NLP applications, as its capacity to recognize and assign appropriate syntactic tags to words is pivotal to the efficiency of NLP solutions. However, the existing POS taggers for Yoruba language either rely on rule-based approaches, which are limited by the comprehensiveness and accuracy of the defined rules or stochastic approaches, which are extremely redundant in generating sequence of tags. Hence, this paper advocates the utilization of machine learning models to develop robust and highly effective POS taggers tailored to Yoruba text. Specifically, a Feed forward Deep Neural Network (FF-DNN) was employed and trained using curated Yoruba tag set sourced from Yoruba religious and dictionary texts, comprising 20795 words alongside their corresponding POS tags. The evaluation of the model demonstrates an accuracy of 98% and a precision of 99% in predicting appropriate tags, outperforming alternative machine learning models such as random forest, logistic regression, and k-nearest neighbour.

1. **INTRODUCTION**

Natural Language Processing (NLP) serves as a powerful tool utilized across a spectrum of sectors, including educational institutions, healthcare facilities, households, and government agencies, to overcome the challenges posed by language diversity. NLP applications aid in language instruction, text comprehension, enable automated language translation, sentiment analysis, among others. At its core, NLP seeks to bridge the gap between human communication and computational analysis by providing machines the ability to process and comprehend natural language data [3].

Despite the widespread adoption of NLP and its benefits across diverse sectors, low-resource languages have not fully benefited from NLP applications [3]. The Yoruba language is widely recognized as a low-resource language in the context of natural language processing (NLP) and computational linguistics. As one of the many languages spoken in Nigeria and neighboring countries, Yoruba faces several challenges that categorize it as a low-resource language in the digital domain. Yoruba lacks comprehensive linguistic resources such as annotated corpora, lexicons, and limited state-of-the-art NLP models.

It is widely acknowledged that part-of-speech (POS) taggers constitute fundamental elements in the development of any Natural Language Processing (NLP) application, playing a pivotal role in enhancing the accuracy of language analysis systems [4]. These tools serve the purpose of identifying and annotating words or tokens within a sentence with contextually appropriate POS tags, which encompass lexical categories such as nouns, pronouns, verbs, and more [5]. An effective POS tagger is instrumental in generating grammatically correct sentences, thereby facilitating various language processing tasks including machine translation, named entity recognition, automated text summarization, and others. POS tagging significantly mitigates ambiguity, a prominent challenge in NLP, by providing clarity regarding the syntactic roles of words within a given context.

In the Yoruba language, there exist eight basic parts-of-speech categories: Noun (Ọrọ Orukọ), Pronoun (Arọpo Orukọ), Verb (Ọrọ Ise), Adjective (Ọrọ Apọnle), Adverb (Ọrọ Asapejuwe), Preposition (Ọrọ Asọtẹle), Conjunction (Ọrọ Asopọ), and Interjection (Ọrọ Ipinu). Each of these categories plays a crucial role in Yoruba syntax and semantics, contributing to the overall structure and meaning of sentences. By accurately tagging words with their respective parts-of-speech, Yoruba POS taggers facilitate syntactic analysis, language understanding, and automated processing of Yoruba text, enabling the development of advanced NLP applications tailored to the Yoruba language community.

Part-of-speech tagging can be executed either manually or automatically, each method offering distinct advantages depending on the size of the corpus being analyzed. Manual tagging tends to yield higher accuracy, especially when working with smaller datasets. However, as the corpus expands, manual tagging becomes increasingly labor-intensive and impractical. Conversely, automatic tagging, particularly with larger datasets, typically outperforms manual methods. Two primary automatic tagging approaches are rule-based and stochastic-based, both commonly employed in existing Yoruba POS taggers are not without their limitations [3]. The rules-based tagging relies on meticulously crafted linguistic rules, which are challenging to develop and often result in rule sets constrained in size. On the other hand, stochastic-based tagging methods tend to generate overly redundant sequences of tags [6].

The adoption of a machine learning (ML) approach for Yoruba part-of-speech (POS) tagging marks a significant advancement in bolstering the capabilities of Natural Language Processing (NLP) systems tailored for the Yoruba language. By harnessing the capabilities of ML algorithms, this study aims to contribute to the development of POS tagging solutions specifically designed for Yoruba text, characterized by enhanced accuracy, scalability, and linguistic insight. The successful implementation of ML-based POS taggers holds the promise of enriching various NLP applications and broadening access to technology and information for speakers within the Yoruba language community.

ML algorithms are subset of artificial intelligence that extract predictive insights from data and make informed decisions based on patterns and trends [6]. These algorithms encompass various categories, including supervised, unsupervised, semi-supervised, and reinforcement learning, each tailored to specific learning objectives and data characteristics.

In the context of this study, the chosen approach involves the utilization of feed-forward Deep Neural Network (FF-DNN) under supervised learning to fulfill the requirements of an effective POS tagger for Yoruba text. By employing a supervised learning paradigm with FF-DNN architecture, the study aims to train the model using annotated Yoruba text data, allowing the algorithm to learn the intricate relationships between words and their corresponding POS tags. Supervised learning enables the model to generalize patterns from the training data, thereby facilitating accurate tagging of unseen Yoruba text samples. The utilization of FF-DNN architecture offers flexibility and adaptability in capturing complex linguistic features inherent in Yoruba syntax and semantics, contributing to the development of a robust and linguistically informed POS tagging solution.

The structure of the remaining sections of this paper is as follows: Section 2 provides a comprehensive review of related works. Section 3 details the materials and methods utilized in the study, outlining the specific methodologies and approaches employed for experimentation and analysis. In Section 4, the experimental setup, results, and evaluation are presented in detail. Section 5 encapsulates the conclusion of the study.

1. **yoruba language**

Yoruba is renowned for its rich linguistic and cultural heritage, with a long history of written and oral literature, arts, religion, and philosophy. It has a diverse range of dialects, reflecting the geographical and historical diversity of the Yoruba people. The Yorùbá language, with stress on the first syllable, belongs to the kwa branch of the Niger-Congo language family. It is also regarded as one of the twelve (12) languages of the Edekiri sub-branch from the great family of Niger-Congo [3]. It is one of the three main languages in Nigeria. It has about twenty dialects, which show differences in phonology and lexis. Yorùbá dialects are mostly spoken in other places that have the language affinity. Examples are the Yorùbá races in Republic of Benin, Togo and Sierra Leone. It is used as language of immigrants in places like Ghana and Cote-de-Ivoire. Due to the influence of 17th, 18th and 19th centuries’ slave trade, the language has gotten some prominence in Cuba, where it is called *LUKUMI* and Brazil, where it is called *NAGO.* The standard Yorùbá form, which is understood by speakers of the various dialects, is used for educational training on radio and television, and in schools and newspaper publications. The language is evidently showing the signs of extinction that are obviously seen in other languages. In Nigeria, the majority of the speakers of the Yorùbá language resides in the southwestern region [14 ] . It is estimated that the language is natively spoken by over 30 million people in southwest Nigeria, Benin, Togo, the UK, Brazil, and the USA [3]. It is one of Nigeria's most widely used native tongues. The majority of Yorùbá population is today within the country of Nigeria where they make up to 21% of the country’s population [15], making them one of the largest ethnic groups in Africa. The oldest known textual reference to the name Yorùbá could be seen in an essay titled: *Mi’raj al-Su’ud* from a manuscript written by the Berber jurist called Ahmed Baba in 1614 [ 16]. This earliest 1600’s reference is indicative that the name Yorùbá was already in popular demotic use as far back as 1500s.

Many guesses have occurred concerning the derivation of the name Yorùbá and these include *Ya’rub* meaning: “Son of Canaanite”, Joktan by Mohammed Bello [17 ], *Goru Ba* by T.J. Bowen or *Yolla Ba*, a Mende word for the River Niger [18 ]. The English ethnologist Richard Burton noted that the name is derived from *Ori Obba* meaning: the king’s head” [19]. Huge Clapperton moved on to subject the name to early changes in its evolution from the existing Hausa coinage *Yorabá* as was the style of addressing the king of Oyo [ 20]. These guesses have suffered a lack of support by many locals for being alien to the tradition of the Yorùbá people [21 ]. Oral history has recorded that the Yorùbá language developed in-situ out of earlier Mesolithic Volta Niger populations by the 1st millennium BCE[22 ]. Historically, the Yorùbá people were a dominant cultural force as far back as the 11th century [23 ]. The people presently reside in a country called Nigeria and they natively speak a language, which is Niger-Congo based with dwindling number of L1 speakers [24 ].

The Yorùbá language has eight basic part-of-speech, and they are: **Noun** – Ọrọ Orukọ, **Pronoun**–Arọpo Orukọ, **Verb**–Ọrọ Ise, **Adjective**–Ọrọ Apọnle, **Aderb**–Ọrọ Asapejuwe, **Preposition**– Ọrọ Asọtẹle, **Conjunction**–Ọrọ Asopọ, and **Determinant**–Ọrọ Ipinu. Noticeably, classes of words in Yorùbá sentences have been variously used as labels in PoS Taggers and the most important are the verbs, adverbs, adjectives, nouns, pronouns and the three connecting elements such as prepositions, conjunctions, and interjections.

1. **VERB:** is used to express action. Examples are Jump, Run, Sleep and Stand. **Usage**:English-*Ayo* ***jumped*** *on the roof; he did not* ***run*** *before he* ***stopped****, because he was feeling* ***sleepy***. Yorùbá-*Ayo* ***fo*** *si ori orule; ko sa* ***ere*** *ki o to* ***duro****, nitori wipe onse e bi ki* ***osu****n.*

2. **ADVERB:** is used to modify or to describe a verb, adjective or another verb. Examples are gently, extremely, carefully, well and very quickly. **Usage:** English*-Ayo handed the book to me* ***gently*** *and he* ***carefully*** *show respect with* ***well-deserved*** *honour and he later disappeared* ***very quickly*** *afterwards.* Yorùbá-*Ayo se* ***jeje*** *lati fun mi ni iwe, osi* ***rora*** *se aponle pelu iwuri* ***ti o pe iye****, lehin eyi ni o poora pelu* ***isure tete***.

3. **ADJECTIVE:** is used to modify or describe a noun or pronoun. Examples are pretty, old, green and smart. **Usage:** English-*The* ***old*** *woman wears a* ***green*** *blouse; she is* ***pretty*** *and* ***smart***. Yorùbá-*Iya* ***arugbo*** *wo aso* ***olomi ewe****, o* ***rewa*** *osi* ***jafafa****.*

4. **NOUN:** is a name of a person, place, thing or ideas. Examples are ayo (oruko), house (ibi ti eniyan ngbe), thing (ohun) and ideas (imo/ogbon). **Usage:** English-***Ayo*** *went to* ***Lagos*** *for a holiday, his friend offered him* ***rice****, and he ate it with* ***happiness***. Yorùbá-***Ayo*** *lo si ilu* ***Eko*** *fun isinmi, ore re fun ni* ***iresi*** *je, osi je e pelu* ***idunnu***.

5. **PRONOUN:** is a word used in place of a noun. Examples are she, we, they and it. **Usage:** English-***She*** *brought good news from the King,* ***we*** *passed it unto* ***them*** *and* ***they*** *were very happy*. Yorùbá-***Arabirin*** *mu ihin rere wa lati odo Oba,* ***awa*** *so o fun* ***won****, inu* ***won*** *si dun pupo.*

6. **PREPOSITION:** is a word or group of words used before a noun, pronoun, or noun phrase to show direction, time, place, location, and spatial relationships. Examples are by, with, about and until. **Usage:** English-*The lady stood* ***with*** *her friend* ***by*** *the roadside, they stayed silent for a while* ***until*** *they told us* ***about*** *themselves.* Yorùbá-*Arabirin na duro* ***ti*** *ore re* ***ni****-****eba*** *ona, won da ke jeje fun igba die* ***titi*** *won fi so* ***nipa*** *won fun wa.* The typology of prepositions is further tagged and/or labeled in the following narratives for better understanding. (1) Preposition of **Time**: ***During***-Ayo came home **during** the winter (Ayo wa si ile **ni igba** otutu). (2) Preposition of **Place**: ***In***-Ayo met me **in** the house (Ayo pade mi **ninu** ile). (3) Preposition of **Direction**: ***Along***: Ayo met me **along** the road to my house (Ayo pade mi **ni ona** ti olo si ile mi). (4) Preposition of **Movement**: ***From***-Ayo is moving **from** one house to another (Ayo nlo **lati** ile kan si ekeji). (5) Preposition of **Means**: ***by bus***-Ayo travelled **by bus** (Ayo rin ajo **pelu oko ero**). (6) Preposition of **Instrument**: ***with a pen***-Ayo wrote the book **with a pen** (Ayo ko iwe na **pelu gege**). (7) Preposition of **Manner**: ***without hesitation***-Ayo did the job **without hesitation** (Ayo se ise na **lai se iye meji**). (8) Preposition of **State/Condition**: ***on duty***-Ayo met Tolu **on duty** (Ayo ba Tolu **lenu ise**). (9) Preposition of Purpose: ***Search for gold***-Ayo will **search for gold** to better his condition (Ayo **yio wa wura** lati je ki igbe aiye re kodara). There is also idiomatic usage of preposition in which some prepositions are used after certain nouns, an example is ***succeed in***-Ayo will succeed in his work (Ayo **yi o** **se orire** ninu ise re).

7. **CONJUNCTION:** is used to join words, phrase, or clause. Examples are and, but, or, while and because. **Usage:** English-*A lady* ***and*** *her friend stood by the roadside looking at the dramatist* ***while*** *others were moving,* ***but*** *they could not talk* ***because*** *they were shy* ***or*** *angry.* Yorùbá-*Arabinrin* ***ati*** *ore re duro si eba ona lati wo osere na* ***nigbati*** *awon ara ti oku nrin lo,* ***sugbon*** *won ko lee soro* ***nitori*** *ti oju nti won* ***tabi*** *ki won ma binu.*

8. **INTERJECTION:** is used to express emotion. Examples are oh!, wow! and oop! **Usage:** English-Wow! They killed the man. Yorùbá-***Paga!*** *Won pa arakunrin na.*

1. **Related works**

In a couple of reviews conducted on part-of-speech (POS) taggers for the Yoruba language and other low-resource languages such Igbo, Bangla, and Korean, different approaches have been explored to address the challenges inherent in linguistic diversity and resource limitations.

In studies [7, 8], a rule-based approach to POS tagging was employed for translation purposes. The method involved formulating and specifying RE-writes rules using context-free grammar, with the aim of translating English words into their corresponding interpretations in Yoruba. The translation process utilized a bilingual dictionary dataset containing English words paired with their Yoruba equivalents. However, a notable limitation of the system was identified in handling English words with multiple meanings, which can lead to confusion for translators. To mitigate this issue, the authors recommended the implementation of additional rules to disambiguate such cases effectively. Moreover, there was a proposal for a method that considers diacritics in the POS tagging of Yoruba words, particularly for words with multiple meanings.

Meanwhile, in study [9], authors presented a suffix-based POS tagger for the Bangla language, employing suffix analysis as the primary method. The dataset used in the study was partitioned into training and test sets, with the former utilized for model training and the latter for evaluation. The tag sets were derived from rules based on Bangla grammar and sentence patterns, enabling the POS tagger to analyze and annotate Bangla text effectively.

Authors in [10] investigated the performance of six existing POS taggers on Igbo POS-tagged corpora, identifying the most suitable method for the Igbo language with a small dataset. The evaluation, conducted on a corpus containing approximately 300,000 tokens mapped with 67 tags, revealed that taggers performed better on known words but were unsatisfactory on unknown words. In [11], a novel two-step POS tagging method was proposed, employing Long Short-Term Memory (LSTM) encoder-decoder and Bidirectional Long-Short-Term-Memory Conditional Random Fields (BI-LSTM-CRF). This method generates a sequence of lemmatized and recovered morphemes from an input sequence, followed by assigning a POS tag to each generated morpheme using BI-LSTM-CRF.

Additionally, [12] introduced a novel neural architecture for an end-to-end Korean POS tagger using an input feeding and copying mechanism. This method, applied on the Sejong-Korean POS tagging dataset, extended the Seq. to Seq. model by modifying the decoding process and incorporating an attention mechanism to address performance issues with long sentences.

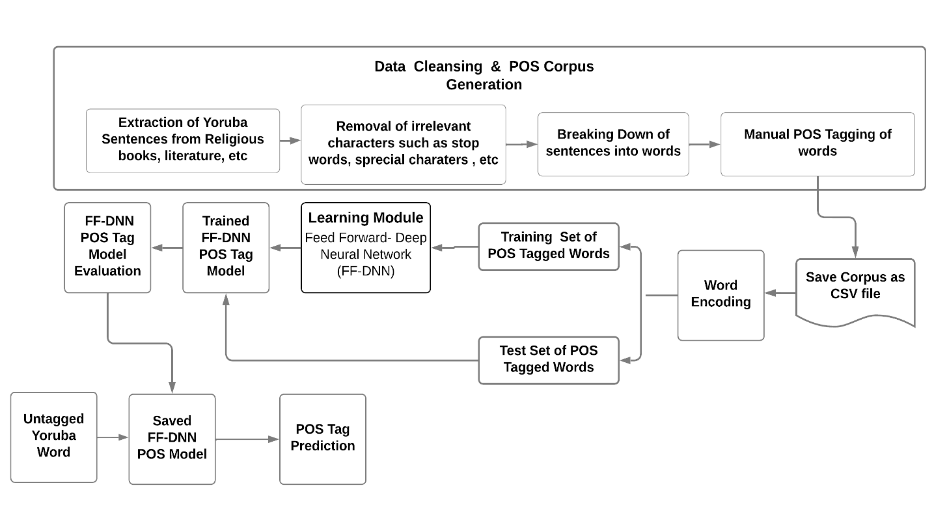
The authors [13] introduced two smoothing techniques to evaluate whether incorporating information from affixes could enhance the performance of a part-of-speech (POS) tagger utilizing the Hidden Markov Model (HMM). The study compared the performance and sustainability of two POS tagging models, namely the HMM-based model and Conditional Random Fields (CRF), particularly in scenarios where data scarcity or other challenges associated with the models arise. To conduct the study, the researchers manually annotated a dataset consisting of 8075 words, following the guidelines provided by the Penn Treebank tagging guideline. Subsequently, both HMM and CRF models were trained and tested using the same dataset, which was split into two sets: 6620 words for training and 1445 words for testing, with 214 of them being unknown words. The results revealed that the HMM-based model achieved an overall precision and recall of 95.12% and 79.41%, respectively, while the CRF model attained 97.19% precision and 83.08% recall. Through performance analysis tests, it was confirmed that the CRF model outperformed the HMM models in terms of accuracy and recall.

The authors in [14] presented a study aimed at designing an annotated corpus for the Yoruba language and developing a tagger using a statistical approach. The Yoruba bible served as the primary source of text, and Perl 5 scripts were employed to extract text and perform segmentation. The system was trained using a Support Vector Machine (SVM) tool called SVMTlearn, and SVMTagger was utilized to annotate the corpus. Manual annotation was conducted for unknown elements, leveraging an English-Yoruba and Yoruba-English dictionary for reference. The study utilized a dataset consisting of 312,562 words and achieved an impressive accuracy rate of 98.04%.

Given the findings from these reviews, there remains a lack of robust human language technologies for Yoruba, with existing tools providing subpar translation outputs. Challenges include the consideration of the eight major parts of speech for Yoruba words and the handling of Yoruba diacritical marks. To address these issues, the study proposed the use of a neural network algorithm to learn words and their corresponding POS tags in a Yoruba dataset, enabling informed decision-making based on learned knowledge. This approach aims to improve the accuracy and effectiveness of Yoruba language processing tools and enhance translation capabilities in low-resource language contexts.

1. **mATERIALS AND METHODS**

The proposed system architecture, illustrated in Figure 1, comprises three modules: data cleansing and POS corpus generation, POS tag model training and prediction, and evaluation. The Yoruba data undergoes preprocessing, primarily involving word encoding. Features are encoded using the Term Frequency Inverse Document Frequency (TFIDF) vectorizer, a text transformation technique that converts words into vectors. For label encoding, a binary-coded method is employed, effectively encoding data into '0' and '1' representations. The training set is then inputted into a Feed Forward Deep Neural Network (FF-DNN), enabling the system to learn each word and its corresponding part-of-speech tag within the training dataset. The utilization of FF-DNN for training facilitates the learning of word-to-tag mappings, enabling the system to predict part-of-speech tags for unseen Yoruba text data. Finally, the proposed system is evaluated to assess its efficiency. This architecture delineates a systematic approach to handle Yoruba language processing tasks, starting from data preparation through to model training and evaluation.



*Figure 1. Proposed architecture*

**Data Cleansing and Corpus Generation**

Given the scarcity of Yoruba Part-of-Speech (POS) corpora for machine learning in publicly available data repositories, the necessity to create one was imperative. To address this, raw data was meticulously gathered from diverse sources, including Yoruba dictionaries, online newspapers, and religious books. A total of 621 sentences were harvested and subsequently subjected to a cleansing process aimed at removing unwanted characters such as stop words, numbers, and special characters (e.g., &, ?, /, \_). The cleansed data, which were originally in sentences were broken down into words using regular expressions to tokenize on whitespace. Thereafter, the tokenized data was manually tagged with their corresponding part of speech. Since the best approaches for developing POS tag corpus are mostly supervised, this study engaged manual annotators to POS-tagged corpus for the language under consideration. The manually annotated text contains 20795 tokens; with 8 distinct POS tags assigned to each token. This approach to corpus creation serves as a foundation for the development of robust machine learning models specifically tailored for the Yoruba language, addressing the data sparsity challenge, and contributing to the advancement of natural language processing capabilities in Yoruba computational linguistic contexts. Table 1 presents the English POS, its Yoruba equivalent, and their POS tags.

Table 1. English word, its Yoruba equivalent and POS Tags

|  |  |  |
| --- | --- | --- |
| English POS | Yoruba POS Equivalent | POS Tag |
| Pronoun | Aropo Oruko | Prn |
| Noun | Oro Oruko | N |
| Adverb | Asapejuwe | Adv |
| Adjective | Oro Aponle | V |
| Preposition | Oro Asotele | Prn |
| Interjection | Oro Ipinu | Interj |
| Verb | Oro Ise | Conj |
| Conjuction | Oro Asopo | Prn |

**

*Figure ???: Sample Data*

*Word Encoding*

In the process of word encoding, Term Frequency Inverse Document Frequency (TFIDF Vectorizer) was employed to transform tokens into vectors. TFIDF is a text transformation technique that assigns weights to words based on their frequency in a document relative to their frequency across multiple documents. This method helps capture the importance of words within a specific context while reducing the impact of common terms.

For label encoding, a binary-coded method was utilized. This method encodes the data into '0' and '1' values, indicating the presence or absence of a particular label, respectively. Specifically, '1' denotes the acceptance of the tag for the presented word, while '0' signifies that the tag is not accepted for the given word. This binary representation enables the neural network to understand and process the association between words and their corresponding labels efficiently.

Once encoded, these vectors serve as input into the neurons within the first layer of the artificial neural network (ANN), which constitutes the input layer. Each neuron in the input layer corresponds to a specific feature or dimension of the input data, in this case, the encoded word vectors. By feeding the encoded vectors into the input layer, the neural network begins the process of learning and extracting meaningful patterns from the data, ultimately facilitating the classification or prediction of part-of-speech tags for the given words.

*Data Split*

The dataset underwent a division into a training dataset and a test dataset, adhering to the common practice in machine learning model development. The training dataset comprised 80% of the entire dataset, while the test dataset constituted the remaining 20%. In essence, the training set contained Yoruba words accompanied by their corresponding tags, serving as the input for training the model. Conversely, the test set comprised words lacking tags, which were utilized to evaluate the performance of the developed Yoruba POS tagger. The training process involved feeding the training set into the Feed Forward Deep Neural Network (FF-DNN), enabling the model to learn the associations between individual words and their respective tags. Subsequently, the test set was utilized to validate the effectiveness of the developed Yoruba POS tagger. By inputting the untagged words from the test set into the trained model, the tagger's performance could be assessed based on its ability to accurately predict the appropriate part-of-speech tags for the given Yoruba words. Table 3 presents the training and test set distribution.

*Table 3: Distribution of dataset*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Entire set** | | **Training set** | | **Test set** | |
| Noun | 11668 | Noun | 9302 | Noun | 2366 |
| Pronoun | 16 | Pronoun | 13 | Pronoun | 3 |
| Verb | 4672 | Verb | 3747 | Verb | 925 |
| Interjection | 23 | Interjection | 19 | Interjection | 4 |
| Preposition | 74 | Preposition | 63 | Adjective | 679 |
| Adjective | 3412 | Adjective | 2733 | Preposition | 11 |
| Adverb | 894 | Adverb | 728 | Adverb | 166 |
| Conjunction | 36 | Conjunction | 31 | Conjunction | 5 |
| Total | 20759 | Total | 16636 | Total | 4159 |

**Learning Model**

The learning model employed to predict the part of speech for possible Yoruba words in this study is the FF-DNN classifier. FF-DNNs are simply stacked neural networks with multiple hidden layers that moves in forward direction. They are widely used for different machine learning based classification and prediction tasks, due to their discriminative and representation learning capabilities [13]. FF-DNN is expected to learn from the tagged dataset (training set) and perform validation using the untagged set (Test set). Suppose the input of FF-DNN is a sequence of untagged token words , and the model is to produce an output sequence , from the set of 8 part of speech categories; denoting the output of the hidden layers as of FF-DNN, the computation of output with hidden layers is given as:

where each pre-activation function is a linear operation with weight matrix and bias ; as captured in Equation (2).

The activation function used in the study is the Rectified Linear Unit (ReLU). In the output layer, the neuron with the highest probability was the one projected, by the neural network as the predicted tag. After computing the output, the network output was compared with the actual output, and the error term was computed as an iterative nonlinear optimization method was employed on weight vectors which minimizes :

(3)

where is the iteration step, is specified with respect to the gradient of the error function which is computed by the error of backpropagation, are weights for the output layer and the hidden layer respectively. This process is repeated for multiple iterations, until it reached the acceptable error rate.

**Performance Metrics**

Accuracy, Precision, Recall and F-measure were considered for the performance evaluation metrics. These metrics are computed as presented in Equations (3), (4), (5) and (6).

*Accuracy*

Accuracy is the ratio of number of correct predictions to total number of predictions.

*Precision*

Precision is calculated as the sum of true positives across all classes divided by the sum of true positives and false positives across all classes.

where TP is true positive, and FP is false positive. True positive is when the predicted class is positive and actual class is positive. If the predicted class is positive, but actual class is negative then it is false positive.

*Recall*

Recall is calculated as the sum of true positives across all classes divided by the sum of true positives and false negatives across all classes.

where TP is true positive, and FN is false negative. True positive is when the predicted class is positive and actual class is positive; and if the actual class is positive, but the predicted class is negative, then it is false negative.

*F-Measure*

F1 Score is the weighted average of Precision and Recall.

**4 Experimental Setup**

The model was experimented using a personal computer system with 2.3GHz Intel (R) Core (TM) i5 4200U CPU @ 1.60 GHz 2.30GHz and 8GB of RAM. The codes were implemented with python 3.7.6 programming language. FF-DNN model was developed using Keras. Number of neurons in the input layer is twelve (12), which is maximum length of a valued word in Yoruba Language. The length of the target vector depends on the number of different grammatical categories that were compiled in the develop annotated corpus, which in this case is eight (8) distinct POS tags.

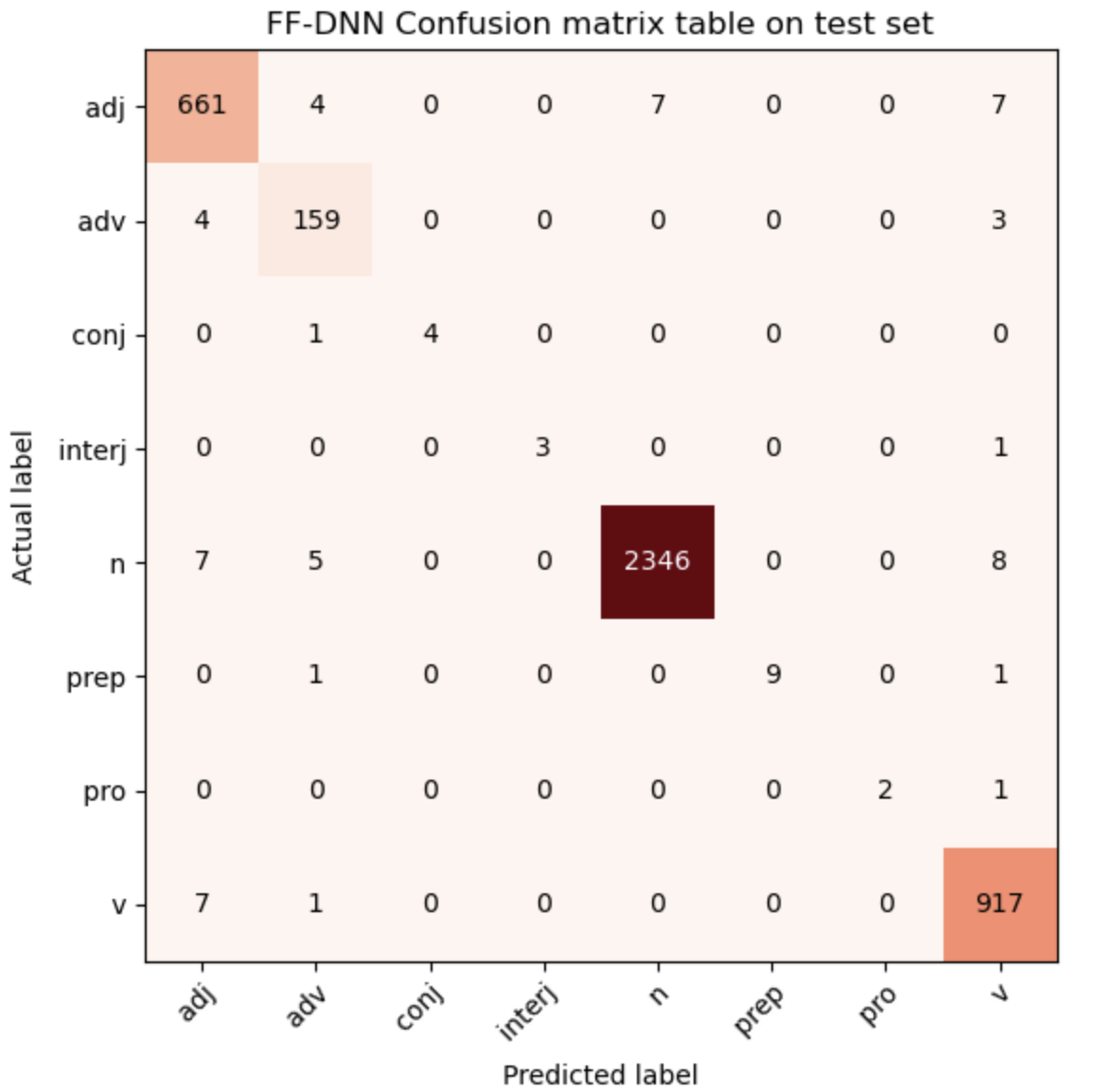
**4.1 Results and Evaluation**

Figure 2 shows the confusion matrix plot of the test data. The confusion matrix displays the rightly predicted tags, and the misclassified tags. The values on the principal diagonal are indicated as the correctly predicted POS tags, while values off the principal diagonal are the wrongly tagged POS. In addition, parameters (true positive, true negative, false positive, and false negatives) for computing performance metrics for this study are obtained from the confusion matrix plot. The overall accuracy of the developed model is 60.3%, which is obtained from the computation of the ratio of correct predictions to the total number of test set. Furthermore, the computation of the macro average of scores yielded a precision of 63%, a recall of 55%, and an F1-score of 0.57, which is fair score; considering the small size of dataset used in the study.

Similarly, the Precision, Recall and F1-measure for each class of POS were computed; and the results are as shown in Table 3.

The Table shows that the model predicted “Noun” POS tagged word much better than the others. This is due to the high quantity of “Noun” tagged words in the training and test sets.

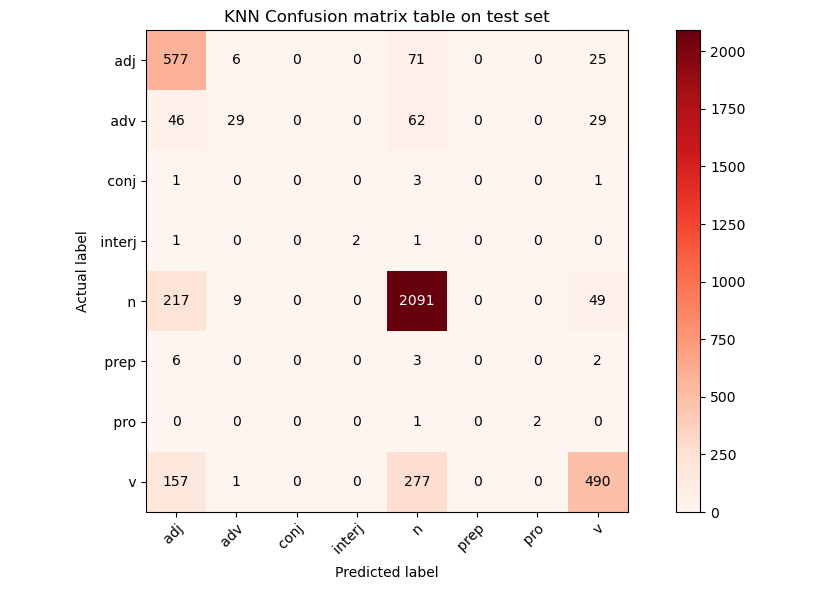
Presented in Table 4 is the prediction of POS by the validated model, when fed with untagged word input directly from external users. The prediction in the Table shows the developed model correctly predicted the POS for the given sentences, which proved the usability of the developed model in real world situations.



*Figure 2: Confusion of Plot on test set*

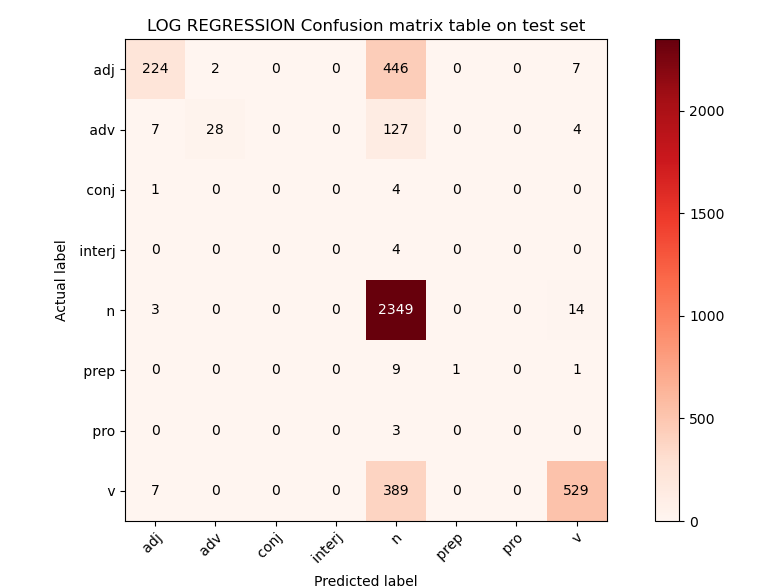
*Table 4. Performance evaluation of FF-DNN*

|  |  |  |  |
| --- | --- | --- | --- |
| **POS Tag** | **Precision** | **Recall** | **F-score** |
| Adj | 0.97 | 0.97 | 0.97 |
| Adv | 0.93 | 0.96 | 0.94 |
| Conj | 1.00 | 0.80 | 0.89 |
| Interj | 1.00 | 0.75 | 0.86 |
| N | 1.00 | 0.99 | 0.99 |
| Prep | 1.00 | 0.82 | 0.90 |
| Prn | 1.00 | 0.67 | 0.80 |
| V | 0.98 | 0.99 | 0.98 |
| **Macro Avg** | **0.98** | **0.87** | **0.92** |

**

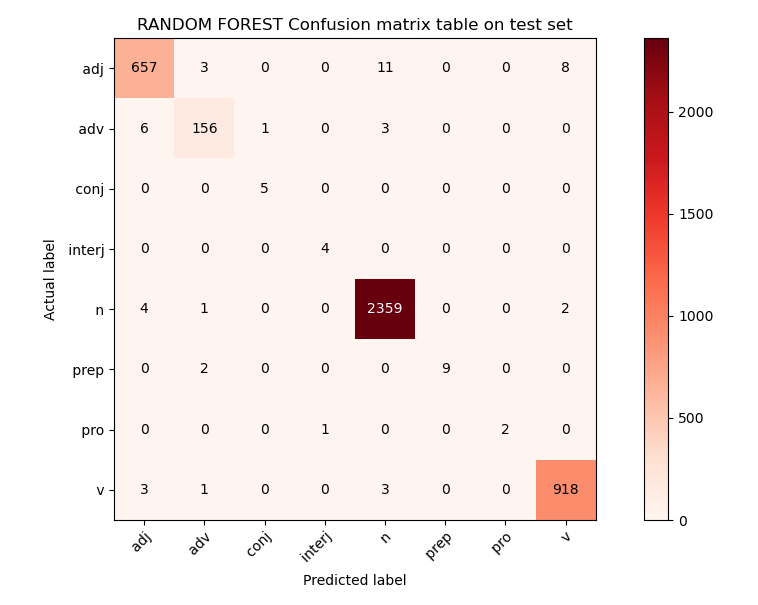
*Table 5. Performance evaluation of KNN*

|  |  |  |  |
| --- | --- | --- | --- |
| **POS Tag** | **Precision** | **Recall** | **F-score** |
| Adj | 0.57 | 0.85 | 0.69 |
| Adv | 0.64 | 0.17 | 0.27 |
| Conj | 0.00 | 0.00 | 0.00 |
| Interj | 1.00 | 0.50 | 0.67 |
| N | 0.83 | 0.88 | 0.86 |
| Prep | 0.00 | 0.00 | 0.00 |
| Prn | 1.00 | 0.67 | 0.80 |
| V | 0.82 | 0.53 | 0.64 |
| **Macro Avg** | **0.61** | **0.45** | **0.49** |

**

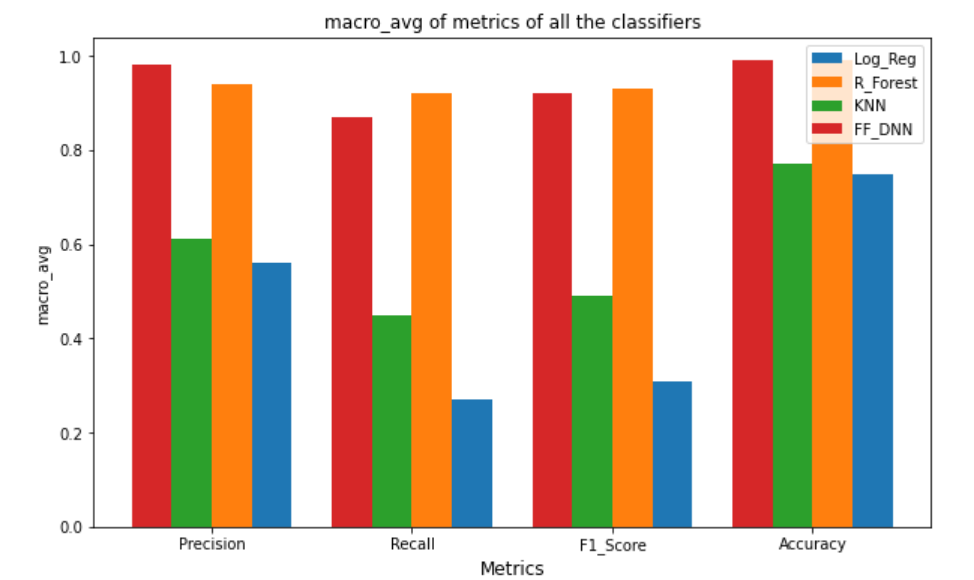
*Table 6. Performance evaluation of LOG REGRESSION*

|  |  |  |  |
| --- | --- | --- | --- |
| **POS Tag** | **Precision** | **Recall** | **F-score** |
| Adj | 0.93 | 0.33 | 0.49 |
| Adv | 0.93 | 0.17 | 0.29 |
| Conj | 0.00 | 0.00 | 0.00 |
| Interj | 0.00 | 0.00 | 0.00 |
| N | 0.71 | 0.99 | 0.82 |
| Prep | 1.00 | 0.09 | 0.17 |
| Prn | 0.00 | 0.00 | 0.00 |
| V | 0.95 | 0.57 | 0.71 |
| **Macro Avg** | **0.56** | **0.27** | **0.31** |

**

*Table 6. Performance evaluation Of RANDOM FGOREST*

|  |  |  |  |
| --- | --- | --- | --- |
| **POS Tag** | **Precision** | **Recall** | **F-score** |
| Adj | 0.98 | 0.97 | 0.97 |
| Adv | 0.96 | 0.94 | 0.95 |
| Conj | 0.83 | 1.00 | 0.91 |
| Interj | 0.80 | 1.00 | 0.89 |
| N | 0.99 | 1.00 | 0.99 |
| Prep | 1.00 | 0.82 | 0.90 |
| Prn | 1.00 | 0.67 | 0.80 |
| V | 0.99 | 0.99 | 0.99 |
| **Macro Avg** | **0.94** | **0.92** | **0.93** |

**

*Table 4. Model prediction of untagged words*

|  |  |
| --- | --- |
| **Untagged Words** | **POS Prediction** |
| Kọ́lá naa Akínlàdé | ['N' 'Det' 'N'] |
| Akínlàdé n jeun | ['N' 'V' 'V'] |
| ọ̀kan nii won | ['Det' 'V' 'Prn'] |
| Kọ́lá n lo | ['N' 'V' 'V'] |
| ilẹ̀ Kọ́lá | ['N' 'N'] |
| èdè | ['N'] |

1. **CONCLUSION**

This study has developed a part of speech tagger for Yoruba language using deep neural network to boost the performances of natural language processing applications or high-level tasks. The study focused on predicting eight distinct Part of speech in Yoruba. Experimentation was based on the data collected from religious, newspapers and other literature books. 621 sentences were extracted, preprocessed, and tokenized into 13757 words. Furthermore, the tokenized sentences were manually tagged with the correct POS and fed into deep neural network for POS prediction of untagged words. The results obtained from the experiment were promising; as an accuracy of 60.3%, a recall of 63%, a precision of 55% and an F1-score of 0.57 were recorded. However, research with interest in developing POS taggers can improve on this study by increasing number of POS set and employing more sophisticated deep learning approach.

**REFERENCES**

[1] A. Sirbu, The significance of language as a tool of communication. *Naval Academy Scientific Bulletin, 2*(2015), 405 – 406.

[2] N. U. Emeka-Nobia, The place of indigenous Nigeria languages in national development. *Research on Humanities and Social Sciences, Department of Linguistics and Literary Studies, Ebonyi State University, Abakaliki, Nigeria, 5*(2015), 112 – 117.

[3] A. A. Semiyou, J. O. Aoga, M. A. Igue, Part of speech tagging of Yoruba standard language of Niger-Congo family. *Research Journal of Computer and Information Technology Sciences, 1*(2013), 2 – 5.

[4] J. Markam, V. Rus, On the implementation of a baseline part-of-speech tagger*. Journal of Computing Sciences in Colleges, 20*(2005), 151 – 159.

[5] T. Adegbola, Building capacities in human language technology. *African Language Technology, 2*(2009), 53 – 54.

[6] N. J. Nilsson, Introduction *to machine learning*. Stanford University, Stanford, CA 94305, 1998.

[7] S.I. Eludiora, A.O. Adeyangi, D.T. Ojediran, Word sense disambiguation in English to Yoruba machine learning translation system. *Journal of Multidisciplinary Engineering Science and Technology (JMEST), 2*(2015), 1814 – 1819.

[8] O.I. Akinwale, A.O. Adetunmbi, O.O. Obe, A.T. Adesuyi, Web-based English to Yoruba machine translation. *International Journal of Language and Linguistics, 3*(2015), 154 – 159.

[9] M.K. Roy, M. K. Paul, S. R. H. Noori, S. M. H. Mahmud, Suffix based automated part of speech tagging for Bangla language. *International Conference on Electrical, Computer and Communication Engineering, (EECE), 1*(2019), 7 – 9.

[10] I.E. Onyenwe, M. Hepple, U. Chinedu, I. Ezeani, Toward an effective igbo part-of-speech tagger. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18*(2019), 1-26.

[11] H. Song, S. Park, Korean part of speech tagging based on morpheme generation. *ACM Trans. Asian Low-Resour. Lang. Inf. Process, 19 (2020)*, 3.

[12] S. Jung, C. Lee, H. Huang, End to end Korean part-of-speech tagging using copying mechanism. *ACM Trans. Asian Low-Resour. Lang. Inf. Process, 17* (2018).

[13] M. Imran, Advantages of neural network-benefits of AI and deep learning (2020). <https://www.folio3.ai/blog/advantages-of-neural-networks/> (accessed July, 3rd, 2021).

[14] Fabunmi, F. A., & Salawu, A. O. (2007). Is Yoruba an Endangered Language?. *Trends in the Study of Languages and Linguistics in Nigeria*, 245-260.

[15] Central Intelligence Agency Nigeria country profile in The World Fact-book 18th January 2022

[16] Ahmed Baba Slavery JSTOR 25653344 of October 2021

[17] Jefferys, M.D.W (2021) Braima Alias Abraham: A study in Diffusion, Folklore 70 (1)

[18] Bowen J.T. (1857) Central Africa: Adventures and Missionary Labour, pg. 264 ISBN: 9780598721280

[19] Burton R.F. Abeokuta and the Cameroons Mountains pg.229 retrieved October 2021\

[20] Huge Clapperton (1829) A second expenditure into the interior of Africa John Murray pp.41

[21] Parrinder et al (1878) Folklore in JSTOR 1259324 of October 2021

[22] Adeyemi A. (2016) Migration and the Yoruba Myth of Origin European Journal of Arts pp.36-45

[23] Robin Walker(2006) When we ruled: The ancient and medieval history of black civilization Indiana University: Every Generation Media pp. 323

[24] Bendor-Samuel, J.T. Benue-Congo languages Encyclopedia Britannica

A.E. Adenike, O.A. Odetunji, & S.S. Akeem, Lexical ambiguity resolution system for standard Yoruba verbs. *American Journal of Engineering Research (AJSER), 7*(2018), 170 – 176.

P. Passban, Q. Liu, A. Way, Boosting neural pos tagger for Farsi using morphological information. *ACM Trans. Asian Low-Resour. lang. Inf. Process, 16 (2016)*, 1.

**ADDITIONAL READINGS**

Abiola, Oyelere (2016). *Rule based parts of speech tagging of Yorùbá simple sentences*. Munich, GRIN Verlag. https://www.grin.com/document/347022.

Akinwonmi, A. E. (2021). Development of a prosodic read speech syllabic corpus of the Yoruba language. *Development, 7*(36), 13 – 32.

Anna, Maria Di Sciullo (2000). Parsing asymmetries. In D. N. Christodoulakis (Ed.), *Natural Language Processing – NLP 2000: Second International Conference Patras, Greece, June 2-4, 2000 Proceedings* (pp. 1 – 15). Springer-Verlag Berlin Heidelberg.

Cardenas, R., Lin, Y., Ji, H., & May, J. (2019). *A grounded unsupervised universal part-of-speech tagger for low-resource languages*. arXiv preprint arXiv:1904.05426.

Hassani, H. (2022). *Part of speech tagging (POST) of a low-resource language using another language: Developing a POS-tagged lexicon for Kurdish (Sorani) using a tagged Persian (Farsi) corpus*. arXiv preprint arXiv:2201.12793.

Kim, Y. B., Snyder, B., & Sarikaya, R. (2015). Part-of-speech taggers for low-resource languages using CCA features. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing* (pp. 1292 – 1302).

Maosong, S., Yang, L., & Jun, Z. (Eds.). (2017). *Chinese computational linguistics and natural language processing based on naturally annotated big data*. Lecture Notes in Artificial Intelligence. (Springer International Publishing).

Onyenwe, I. E., Hepple, M., Chinedu, U., & Ezeani, I. (2019). Toward an effective Igbo part-of-speech tagger. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18*(4), 1 – 26. https://doi.org/10.1145/3314942.

**KEY TERMS AND DEFINITIONS**

**Feed Forward Neural Network (FF-DNN):** Stacked layers of feed forward neural networks with the capabilities of complex feature extractions from data.

**Generative Adversarial Network (GAN):** A type of neural network that synthetically generate image, speech, video or textual data from a given distribution of data. The adversarial technique is composed of two neural components namely: the generative and discriminative networks.

**Morpheme:** An indivisible grammatical component of a language. It could be a word or a meaningful part of a word.

**Natural Language Processing (NLP):** A field of artificial intelligence in computer science that studies the ability of computer program in the understanding various forms of languages (i.e., text or spoken).

**Part of Speech (POS) Tagging:** A process of assigning part of speech categories, such as noun, pronoun, verb, adverb, conjunction, adjectives, preposition, ???, to the given word in a sentence or literature, based on their contextual representation.

**Term Frequency-Inverse Document Frequency (TFIDF):** A measure used in machine learning to quantify how relevant a word is to a text in a corpus. It is a metric for texts vectorization process; with which words within a text document are transformed into importance numbers.

**Transfer Learning:** A machine learning process in which a trained model from a given task is reused for another related task. In this process, the machine applies the knowledge acquired from a prior task to improve the generalization of subsequent task.