International Journal of Applied Information Systems (IJAIS) – ISSN : 2249-0868
Foundation of Computer Science FCS, New York, USA
Volume 12– No.30, June 2020 – www.ijais.org

The Performance Evaluation of an Igbo Text-Based Intelligent System

Ifeanyi-Reuben Nkechi J.
Department of Computer Science,
Rhema University, Nigeria

Odikwa Henry
Department of Computer Science
Abia State University Uturu, Abia State, Nigeria

Benson-Emenike Mercy E.
Department of Computer Science
Abia State Polytechnic, Nigeria

ABSTRACT
The expansion in Information Technology (IT) has inculcated Igbo, one of the three Nigeria major languages in text-based intelligent systems such as text classification, data/information retrieval and natural language processing. The evaluation of an intelligent system and its processes is very important for the progress of the system. This paper presents a performance evaluation of an Igbo text-based intelligent system. A system performance evaluation is the procedure by which a system's resources and results are measured to find out if the system is operating at an optimal level. The performance evaluation of the system was done on classification results of an Igbo text represented with Unigram and Bigram Language Models. Object-Oriented design methodology is used for the work and is implemented with the Python programming language with tools from Natural Language Toolkit (NLTK). The system performance is assessed by calculating the precision, recall and F1-measure of the classification result obtained on Unigram, Bigram and Trigram represented text. Result shows classification on an Igbo Bigram represented text has higher level of precision and accuracy.

Keywords
Performance Evaluation, Igbo Language, Text Classification, Unigram Model, Bigram Model, Trigram Model

1. INTRODUCTION
The development in Information Technology (IT) has encouraged the use of Igbo Language in text creation, online news reporting, online searching and articles publications [1]. It has advanced to the extent one can also operate (Windows 7 and above operating system) using Igbo language and creating room for more generation of data with the language [2].

Igbo is one of the three Nigeria major languages. It is an agglutinative language by which its words are formed by stringing up different words. Igbo vocabulary is made up of a large number of compound words. The individual meaning of words in a phrase or compound word does not entail the context it is being used for. Standard Igbo has thirty-six (36) alphabets (a, b, ch, d, e, f, g, gb, gh, gw, h, i, ị, j, k, kw, kp, l, m, n, nw, ny, ṅ, o, ọ, p, r, s, sh, t, u, ụ, v, w, y, z) [3].

Information Technology has dramatically changed the knowledge dissemination process and many of them lack a standard evaluation process to validate the system's performance [4]. The evaluation of an intelligent system and its processes is very important for the progress of the system. System performance is the efficiency of a system or can be seen as how fit a system executes considering all its processes [5]. It is the procedure by which a system's resources and results are measured to find out if the system is operating at an optimal level.

In assessing a system performance, a number of parameters are used to find out the result. System has standards, point for reference, called benchmark, used against the generated parameters to assess the performance of a system [6]. As the technological characteristics of IT are changing, benchmark for evaluating the performance of any given system has to be constantly updated and this makes a system evaluation a difficult work. This paper presents a performance evaluation of an Igbo text-based intelligent system

1.1 Benchmarks for Evaluating System Performance
Benchmark is how to determine the value of system change. [7] explained some of the factors that can serve as benchmarks for evaluating system performance as follows:

i. **Responsiveness:** This measures how quickly a given task can be accomplished by the system. Some of the feasible parameters of responsiveness are waiting time and queue length.

ii. **Usage Level:** This measures how well the different components of the system are being used. The Possible parameters are throughput and utilization of diverse resources.

iii. **Missionability:** Missionability is a measure that indicates if the system would remain constantly operational for a given time. The likely measures of missionability are the allocation of the work accomplished throughout the mission time, interval accessibility/availability (possibility that the system will keep operating efficiently during the mission time) and the life-time (time when the probability of undesirable behaviour increases ahead of defined threshold).

iv. **Dependability:** This measure points out how reliable the system is over the long run. The likely parameters used in this measure are the total of failures in a day, Mean Time to Failure (MTTF), Mean Time to Repair (MTTR), long-term availability, and failure cost.

v. **Productivity:** Productivity indicates how effectively a user can get his or her work accomplished and its possible measures are user friendliness, maintainability and understandability.

The benchmark for evaluating the performance of an intelligent system can either be sub-system or application based. The performance evaluation of an Igbo text-based intelligent system addressed in this work is application based. In this benchmark procedure, certain parameters are used to compute the level of performance of the system.

2. RELATED WORKS

Some papers related to the work were studied, analyzed and discussed as follows:

The model for evaluation of the performance of text classification system using TF*IDF, LSI (Latent Semantic Indexing) and multiple-words text representation was proposed in [8]. The model was experimented and evaluated with Chinese and English text corpora in text retrieval and text classification. Their result showed that LSI produced optimal performance in retrieving English text and also produced best performance for Chinese text classification.

The performance of rule-based classification algorithm is evaluated in [9]. The work compared the five algorithms of three performance parameters, number of classified instances, accuracy and error rate are considered. The results of experiment are presented in tabular and graphical form. From this study it is found that PART is best algorithm for classification.

The similar thesaurus on Arabic language with full word and stemmed methods is designed and built in [10]. The result was compared and it proved that similar thesaurus using stemmed method is more efficient than traditional full words method with higher levels of recall and precision.

[3] adopted Euclidean similarity measure and determined the similarities between Igbo text documents represented with two word-based n-gram text representation (unigram and bigram) models. The result obtained revealed that Igbo text document similarity measured on bigram represented text gives accurate similarity result and will likely give better, effective and accurate result when used for tasks that requires computation of similarity between documents on Igbo language.

The part-of-speech tag set for an Igbo language was developed in [11]. The work serves as a linguistic resource to support computational Natural Language Processing (NLP) research on the Igbo language and was used in a part-of-speech annotation task for the development of POS tagged Igbo corpus.

[12] developed an Arabic text classification system using K-Nearest Neighbour model with vector space model. They analysed and compared results gotten from three different similarity measures; cosine, dice and jaccard. The performance of the system was tested, measured and it was confirmed that the Cosine performed better than Dice and Jaccard. This system was done in Arabic text; the present research will improve on this and experiment will be carried out in Igbo textual documents.

3. SYSTEM METHODOLOGY

This section discusses the procedures and models involved in developing and implementing the proposed system for evaluating the performance of an Igbo text-based intelligent system. The performance is done on classified Igbo text represented with unigram and bigram models.

3.1 Unigram Language Model

A unigram language model represents words or terms independently. It represents document in single words. This can also be described as using BOW in text representation.

The probability of representing text in unigram is approximated as follows:

$$P(s_1,s_2, ..., s_T) \approx P(s_1)\, P(s_2) \ldots P(s_T) \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 1$$

This implies that

$$P(s_{1,}\ s_2\ \ldots\ s_n) = \prod_i P(s_i) \qquad \ldots\ldots\ldots \quad .2$$

For example: P(w=ụlọ akwụkwọ) = P(ụlọ) * P(akwụkwọ)

The text “ụlọ akwụkwọ” will be represented with unigram model as “ụlọ” and “akwụkwọ” = 2 features.

3.2 Bigram Language Model

Bigram estimates the likelihood a word occurring in the context of a previous word. This implies the possibility of a word occurring depends on the possibility of the previous word.

Definition 1: For a given string sequence S={s1,s2, ...s_n}, the possibility of a string sequence in bigram language model (n=2) is given as

$$P(s_1,s_2,...,s_n) =$$

$$P(s_1 \mid s_2,s_3,...,s_n)P(s_2 \mid s_3,...,s_n)...P(s_{n-1} \mid s_n)P(s_n) \quad \ldots\ldots\ldots\ldots\ldots \quad 3$$

It is simply the product of the conditional possibilities of its bigrams i.e.

$$P(s_{1,} \ldots s_n) = \prod_{i=1}^{n} P(s_i \mid s_{i-1}) \ldots\ldots\ldots\ldots\ldots\ldots \quad 4$$

The maximum likelihood estimation in bigram is given as

$$P(s_2 \mid s_1) = count\ (s_1, s_2) / count\ (s_1) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 5$$

The Bigram model probability in Igbo Text is illustrated below:

Igbo: Ihu akwụkwọ Weebụ dị ugbua na-agba mbọ ịmepe otu saịtị na Intaneeti

English: The current webpage is trying to open a site on the Internet

P(Ihu akwụkwọ Weebụ dị ugbua na-agba mbọ ịmepe otu saịtị na Intaneeti) = P(akwụkwọ │ Ihu) P(Weebụ │akwụkwọ) P(dị │ Weebụ) P(ugbua │ dị) P(na │ ugbua) P(agba │na) P(mbọ │ agba) P(ịmepe │ mbọ) P(otu │ imepe) P(saịtị │otu) P(na │ saịtị) P(Intaneeti │ na)

The above text features will be represented in bigram as: Ihu akwụkwọ, akwụkwọ weebu, Weebụ dị, dị ugbua, ugbua na, na agba, agba mbọ, mbọ ịmepe, imepe otu, otu saịtị, saịtị na, na Intaneeti = 12 bigram-based features

3.3 Trigram Language Model

Trigram model employs the most recent 2 words of the document to condition the possibility of the subsequent word. The probability of word order in trigram is

$$P(s_3 \mid s_{1,} s_2) =$$

$$count\ (s_1, s_{2,} s_3) / count\ (s_{1,} s_2) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 6$$

The count (s_1, $s_{2,}$ s_3) is the total occurrence of the sequence of words {s_1, $s_{2,}$ s_3} and count ($s_{1,}$ s_2) is the total occurrence of the sequence of {s_1, s_2} in the text document.

3.4 System Performance Evaluation Process

The performance of an Igbo text-based intelligent system is assessed using the classification result on Igbo text. The percentage of correctly classified document is obtained and used as a means of determining the performance of the system. This is illustrated in table 1. It gives a two-way contingency table for measuring performance.

Table 1: Contingency table for Performance Measure

	Document actually belongs to a class / category	Document actually does not belong to a class
Classifier says document belongs to a class	True (Correct) Positive – TP	False (Fake) Positive - FP
Classifier says document does not belong to a class	False (Fake) Negative – FN	True (Correct) Negative -TN

In this work, a True Positive (TP) is a document that is properly assigned a class label; False or Fake Positive-FP is one that is improperly given affirmative document class label; a Fake Negative-FN is textual document that really fits in to a class, but it was not predictable by the text classification system; True (Correct) Negative-TN is textual document that does not actually belong to document class but the text classification system predicted it does not belong to it.

Sum of TP and FP gives the real quantity of text documents the classification system put in a particular document class, where as the total TP and FN gives the quantity of textual documents that really fit in to a particular class, despite the classification system prediction.

3.3.1 Precision, F1-Measure and Recall

Precision of a proposed text classification system is defined as the quotient of total TPs and sum of total TPs and FPs.

$$\textbf{\textit{Precision}} = \frac{\textbf{TP}}{\textbf{TP+FP}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 7$$

Precision point is known to as a point of correctness.

Recall of the classification system is described as the quotient of total TPs and sum of total TPs and total FNs. Recall level measures completeness.

$$\textbf{\textit{Recall}} = \frac{\textbf{TP}}{\textbf{TP + FN}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 8$$

F1-Measure is single function that joins recall and precision points. When the F1-measure is high, it means that the overall text classification system is high.

F1-Measure =

(2*Precision*Recall) / (Precision + Recall) …………………9

=2TP / (2 TP + FP + FN) ………………………………… 10

4. EXPERIMENTS

The performance evaluation of an Igbo text-based system is implemented with Python and tools from Natural Language Toolkit (NLTK). The performance is measured by computing the precision, recall and F1-measure.

Figure 1: Performance Measure Result

Table 2: Performance Measure Result

Category	TP	TN	FP	FN	Precision	Recall	F1-Measure
Unigram	8	0	2	0	0.80	1.00	0.89
Bigram	9	0	1	0	0.90	1.00	0.95
Trigram	7	0	3	0	0.70	1.00	0.82

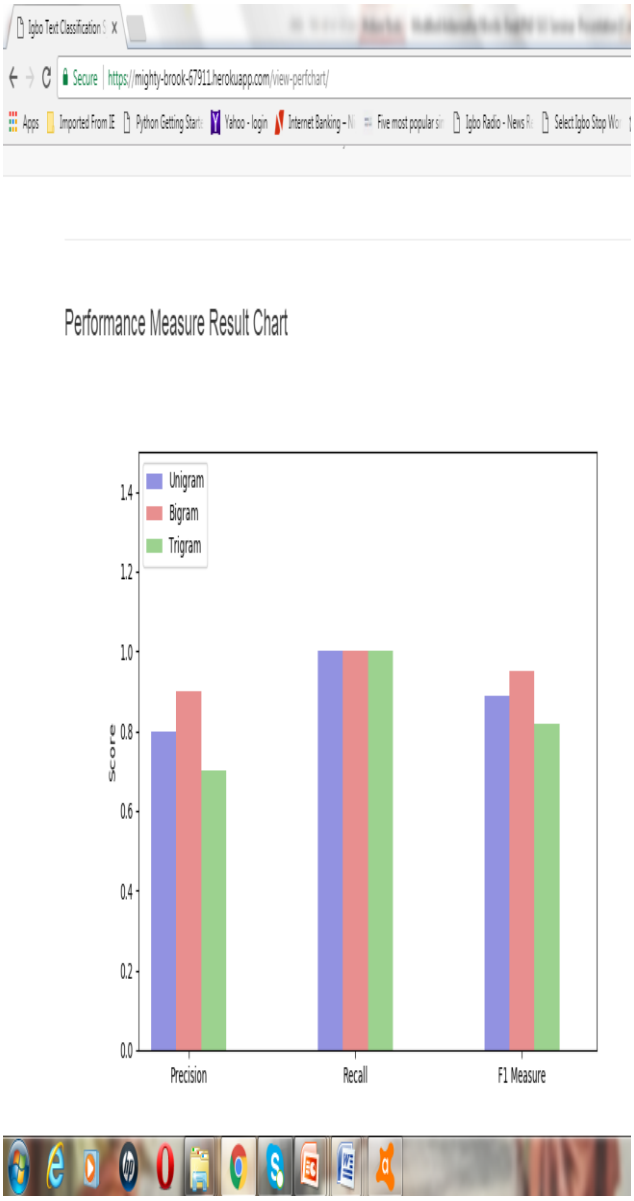

Figure 2: Chart showing Performance Measure Result

5. RESULT ANALYSIS

Figure 1 shows the performance measure result of the developed text-based system. Table 2 shows the detailed analysis of the performance measure result extracted from the system. It displays the TP, TN, FP, FN, recall, precision and F1-measure for the classification result obtained in unigram, bigram and trigram represented Igbo text. In Unigram (TP=8,TN=0,FP=2,FN=0), Bigram (TP =8, TN=0, FP=2, FN=0) and Trigram (TP=7,TN=0,FP=3, FN=0). The recall, precision and F1-measure for unigram are 1.00, 0.80 and 0.89 respectively. The recall, precision and F1-measure for bigram are 1.00, 0.90 and 0.95 respectively. The recall, precision and F1 for Trigram are 1.00, 0.62 and 0.82 respectively.

Figure 2 shows the system performance result chart. Recall measures the degree of completeness while precision measures the degree of exactness. The result shows the performance on the text represented with the three language models (unigram, bigram and trigram) has the same level of recall (completeness). This means all the text documents that were given to the classifier, were given a label name.

The classification with bigram has highest degree (0.90) of exactness (precision) while trigram has the lowest degree (0.62) of exactness.

F1 measures the system accuracy by taking into consideration the precision and recall to calculate its value. F1-measure is at its finest score (value) at 1 and at its worst at 0. Bigram represented text classification has the highest value (0.95) of F1 while Trigram has the lowest value (0.82).

6. CONCLUSION

This paper analyzed, studied and presented a measure for evaluating the performance of any text-based intelligent system on Igbo language. The evaluation of an intelligent system and its processes is very important for the progress of the system. The system performance is assessed by calculating the precision, recall and F1-measure of the classification result obtained on Unigram, Bigram and Trigram represented text. Result shows classification on an Igbo Bigram represented text has higher level of precision and accuracy. The performance result shows that Bigram language model is highly recommended for representing Igbo text for any intelligent system for efficient, robust and effective operation.

7. ACKNOWLEDGEMENTS

The authors wish to express their great appreciation to the reviewers of this paper. We are grateful of your useful comments and contributions that assisted in improving the value of this work.

8. REFERENCES

[1] Ifeanyi-Reuben Nkechi J. and Benson-Emenike Mercy E.(2018). An Efficient Feature Selection Model for Igbo Text. International Journal of Data Mining & Knowledge Management Process (IJDKP). 8(6): 19 -33

[2] Ifeanyi-Reuben N.J., Ugwu C. and Adegbola T. (2017). Analysis and representation of Igbo text document for a text-based system. International Journal of Data Mining Techniques and Applications (IJDMTA). 6(1): 26-32

[3] Ifeanyi-Reuben Nkechi J., Ugwu Chidiebere and Nwachukwu E.O. (2017). Comparative Analysis of N-gram Text Representation on Igbo Text Document Similarity. International Journal of Applied Information Systems (IJAIS). 12(9): 1-7

[4] Hsiao-Ya Chiu, Chieh-Chung Sheng and An-Pin Chen (2008). Modeling agent-based performance evaluation for e-learning systems. The Electronic Library. 26 (3): 345-362

[5] Raj Jain (1991). The Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation, and Modeling. ISBN: 978-0-471-50336-1

[6] Manicka Chezian R. and Kanakalakshmi C. (2015). Performance Evaluation of Machine LearningmTechniques for Text Classification. Proceedings of the UGC Sponsored National Conference on Advanced Networking and Applications. Special Issue Published in Int. Jnl. Of Advanced Networking and Applications (IJANA). 53 – 57

[7] K. Kant (1992). Introduction to Computer System Performance Evaluation, McGraw-Hill Inc.

[8] Zhang W., Yoshida T. and Tang X. (2008). Text Classification based on multi-word with support vector machine. Journal of Knowledge-based systems. 21 (2008): 879-886

[9] Aditi Mahajan and Anita Ganpati (2015) . Performance Evaluation of Rule Based Classification Algorithms. International Journal of Advanced Research in Computer Engineering & Technology (IJARCET). 3(10): 3546 - 3550

[10] Essam, S. H. (2013). Similar Thesaurus based on Arabic document: An overview and comparison. International Journal of Computer Science, Engineering and Applications (IJCSEA). 3(2): 1-10.

[11] Onyenwe, I. E., Uchechukwu, C. and Hepple, M. (2014). Part-of-Speech Tagset and Corpus Development for Igbo, an African language. The 8th linguistic annotation workshop, Dublin, Ireland, pp 93-98.

[12] Jafar A., Omar A., Wael H., Nidhal K. T. E. and Ali A. (2014). Vector space models to classify Arabic text. International Journal of Computer Trends and Technology (IJCTT). 7(4): 219 – 223