1. **Applied Method**

The increased demand for computing power from the business world, pharmacological entities, agricultural industries, and scientific communities’ bases on E-services to classify huge applications and process huge data is on the rise. One of the biggest challenges in achieving this is the time constraint [1-3]. Present-day applications take a lot of time to process the available data. Sometimes it goes for days and weeks to provide the much-needed data to reach a decision.

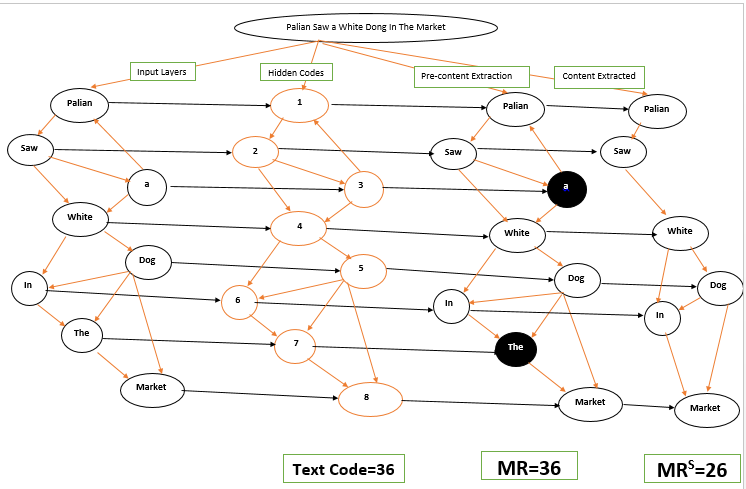


Figure 2 represents a selected text content made up of eight (8) words. Each word is coded with a number. The interconnected words formed a graph depending on the relationship between words. The standard rule to build this graph using numbers allows the study to measure and evaluate content. Each text allows the extraction to evaluate the richest word in a sentence. Figure 2 indicates (pre-content extraction) which according to the study is the total of the input dataset called metrics range while (content extracted) according to the study represents the metrics range substitute. A metrics range substitute is the rich content extracted from the same given text called metrics range. The total number of words according to hidden codes represented by numbers is added to give us twenty-six (26) while the whole content added all together as per hidden codes gives us thirty-six (36). This system of data extraction can evaluate and extract data from all types of the dataset without bias.

1. **Main Contribution**

The study is a proposed model called (scoring model) with the objective to modify the functionality of the bag-of-words model. The study is aimed at simplifying representation used in natural language processing and information retrieval (IR) applied deep learning.

The bag-of-words model is commonly used in methods of document classification where the (frequency of) occurrence of each word is used as a feature for training a classifier.

This study makes it possible to classify, summarize and categorize text or sentences or documents based on the word frequency or occurrence by weighing its relevance, providing grading scale, aligning a pattern of representation and providing a distinct pattern and format that differentiates it from similar or same content type. Brief explanation of the core objectives.

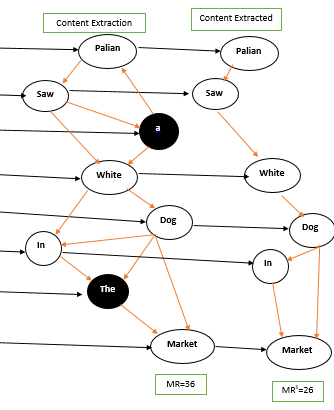
* Weighing document relevance takes care of grammatical errors unlike bag of words model that doesn't consider grammatical errors
* Providing grading scale make sure two or more documents shouldn't be the same except they received the same score.
* Providing a distinct pattern and format that differentiates one content or document from similar or same content type. The study with the help of a scoring model realized that there's a need for decoding information especially for the purpose of security. A lot of cyber security challenges merits the need to code text. The pattern identified in this study can be confidential between producer and end users.

The (scoring model) enables that when applying algorithms in NLP, which works on numbers, we make sure we automatically feed our text into that algorithm. This model does not require the use of a preprocessing of text by converting it into numbers unlike the bag of words, which keeps a count of the total occurrences of most frequently used words. This model automatically generates numbers to all the words or sentences and begins extraction. This automatic generation of numbers to words makes it easier to evaluate content relevance.

**NB:** The scoring model assumes that information comes from humans and deserves a grade to distinct, and recognize intellectual ability and capabilities

**2. Results.**

**In this section, we provide information on how text or speech content can be extracted and classified into different parts of the speech based on data obtained via NLP means that help in decision-making. We applied (behavior-oriented drive and influential functions of the internet of thoughts on*content extraction*). The brief steps taken by the study to come out with the results indicated in the abstract are as follows. Determination of Behavior oriented drive and influential function of text categorization based on NLP and DL**

****

**Figure 1, metrics range and metrics range substitute.**

**Above figure 1, provide data classified into different groups. The data label as (content extracted) are metric range substitute (MRS) while the (content extraction) is metric range (MR). To obtain influence rate, metric range substitute is divided by sum of metric range then multiple by behavior score of value five (5). Behavior score represent human five sense organ. *MR =36 and* =26. . BIF=Behavior oriented drive and influential function on text categorization based on NLP and DL**

**F=push factors of E-services**

***MR =36 and* =26**

**D=Dependent parameters**

**MR=Metrics Range**

**= Metrics Range Substitute**

**BS=Behavior Score or five sense organs**

**KBS=key benefits score**

***=***

**====))))**

**====)))**

**The statistics above detailed how relevant each word in a text content matters. From the statistics we can say that the influence score is grade “Good” as per classification. The score 3.61 is the influential function determining the weight of text content.**

**Abstract**

**Background:** This paper introduces an inductive evaluation of categorizing natural language documents into a graph-like structure to a rich content category. Text categorization has very important in information retrieval thanks to natural language processing systems.

**Objective**: Aims to develop a text categorization system where each word in a text context matters in the extraction process. Also, aims to eliminate focus on frequently used words (FUW) by most algorithms.

**Problem**: The study realizes that tokenization, classification, text tagging, and summarisation have been biased on document context. Many algorithms focus on frequently used words (FUW) which creates a bias to text content categorization present online and on strong databases. Weak databases and less frequently used words (LFUW) with rich content end up being unnoticed.

**Method**: The study used text content, natural language processing, and deep learning techniques to construct a word-to-graph model name (scoring model). Each text word was coded with figures and measured with behavior-oriented drive and influential function.

**Results**: Based on the behavior-oriented drive and influential function, a score grade of 3.61 out of 5 was recorded.

**Conclusion**: The study concluded that content categorization evaluation of words-to-graph-like structure is a perfect model that will achieve unbiased extraction. The study also concluded that this novel model is excellent because it takes into consideration each word in the text content. This approach weighs each word and doesn't focus on popular and frequently used words as other algorithms do.

**Recommendations**: The study recommended professionals in the field of natural language processing and deep learning to review this novel model, train it and implement it as a modern approach to document categorization techniques.

**4. Conclusion**

This includes the synthesis of intelligible verbal and gestural acts per see and each combination allows a continuous flow of human-like multimodal utterance as digital services streamline. The behavior layers of content extraction charge of generating coordinated verbal gestural and facial behaviors for realizing an important of each utterance behavior to the require context. This help the extractor or practitioner determine influence of each utterances or each word in a document.. The main research purpose of this algorithm is to analyze the degree of richness user content has on each word measurement in a document.

**Analysis of a novel model of a text-to-graph categorization based on natural language processing and deep learning techniques**

**Pascal Muam MAH**

AGH University of Science and Technology, Krakow, Poland, [**mah@agh.edu.pl**](mailto:mah@agh.edu.pl)

Supervisors

**Iwona SKALNA**

AGH University of Science and Technology, Krakow, Poland, [skalna@agh.edu.pl](mailto:skalna@agh.edu.pl)

**Tomasz Pełech-Pilichowski**

AGH University of Science and Technology, Krakow, [tomek@agh.edu.pl](mailto:tomek@agh.edu.pl)

