ISSN: 2755-6190

Review Article

Open Access Journal of Artificial Intelligence and Technology

Hybrid Method for Text Graph Classification using 6DoS-CC Framework: A Case Study of Nigerian Hate Speech

Oluwatobi Balogun

Bells University of Technology, Ota, Ogun State, Nigeria

Corresponding author
Oluwatobi Balogun, Bells University of Technology, Ota, Ogun State, Nigeria.

Received: January 08, 2026; **Accepted:** January 14, 2026; **Published:** January 23, 2026

ABSTRACT
In the use of the frequency of words sub-graph approach, the lack of instance feature representation constitutes a major obstacle. This experimental work makes use of the 6DoS-CC framework, a hybridization of the methods of the six degrees of separation for the validation of word graphs and the Collatz Conjecture for identifying vectorized frequency feature of hate-speech graphs based on the number of characters that form each word for the categorization of text. Results show a strong Pearson's Correlation Coefficient value of 0.517462073 indicating a large strength of correlation figure based on the observed high disparity within odd hail stone numbers for Nigerian Hate speech.

Keywords: Classification, Feature extraction, Graphs, Hate Speech, Six Degrees of Separation

Introduction

Hate speech covers many forms of expressions which spread, incite, promote or justify hatred, violence and discrimination against a person or group of persons [1]. The Nigerian social media space is currently rife with such atrocities making it receive the rapt attention of Nigeria's legislative arm of the government [2]. Combating hate speech and fake news are the most pressing societal issues [3]. Various machine learning techniques have been used to automatically classify text which is examined at the document, sentence, and aspect (context) levels. The goal of this process is to automatically classify opinions within the text [4]. Natural languages like the English language's written text have knowledge embedded in them in ambiguous forms A graph is made up of nodes, which stand for items, and edges, which indicate the relationships between those objects [5]. It is an effective model for explaining the intricate structures of things found in the real world [5]. Text-based graphs can be derived for any text corpus as has been widely used in the area of graph theory [6].

The experimental process for this work limits the definition of isomorphism in graphs as looking into whether graphs with comparable structures and/or characteristics can identify a text's class. Generally speaking, there is an isomorphism between graphs and subgraphs if two or more graphs have the same or equal numbers of vertices, edges, in and out degrees, linked components, loops, and parallel edges [7]. One of the problems with easily identifying hate speech lies with the ambiguous nature of natural language which makes classification harder for humans and by extension machines. used tweets to categorize hate speech directed towards black individuals in general as either racist or non-racist [3]. Their poll was created to evaluate the dependability of agreement and the difficulty of identifying hate speech using Fleiss' Kappa. Their findings revealed a 33% overall agreement, which suggests that opinions on what constitutes hate speech are not widely held. This illustrates how challenging it will be for computers to categorize hate speech at the context level.

The lack of instance feature representation is a bane in the use of graphs for classification which informs the use of Machine

Citation: Oluwatobi Balogun. Hybrid Method for Text Graph Classification using 6DoS-CC Framework: A Case Study of Nigerian Hate Speech. Open Access J Artif Intel Tech. 2026. 2(1): 1-6. DOI: doi.org/10.61440/OAJAIT.2026.v2.21

Learning to enhance it [6]. Subgraph feature-based techniques translate a graph into vector space using a set of subgraph features, thereby enabling the application of pre-existing machine-learning techniques. Machine Learning which works well with subgraph feature-based methods and Lexicon-based hybrid approach have been used for sentiment analysis [8]. On the other hand, distance-based methods for graph classification use distance measures to assess similarities between each pair of graphs for classification. Generally, subgraph feature-based methods are superior to distance-based algorithms and feature extraction for scalar-based graphs is often overlooked [6,10]. posited that graph classification is becoming increasingly popular due to its involving dynamically changing graph records [11]. When selecting subgraph features, common methods use evaluation metrics, such as the frequency, to choose several important subgraph features. Hence, a template for identifying hate speech will first be used to properly identify hate speech from a Nigerian perspective and then the sub-graph features will be identified using NLTK library.

The first phase of the framework uses the Six Degrees of Separation (6DoS) to verify that all identified sentences that are classified as hate speech fall within the spectrum of a graph network calculating its average path length described in 3.1. This concept is widely used in the area of human networking and social networks based on the creation of a solid and useful network of contacts implying that anyone is one degree away from acquaintances, two degrees away from everyone they know, and so on [9]. In order to determine the frequency of occurrence of a text graph that has been vectorized as a feature, the second part of the framework uses the Collatz Conjecture (CC), an unproven hypothesis that states that for all positive whole numbers, if odd, is multiplied by 3 and added to 1 but if even is divided by 2, will always fall back to a 4,2,1 loop. The conjecture adds some orderliness to the randomness of words to identify specific vector patterns.

Literature Review

Humans are considered intelligently unique because they can form sophisticated ideas and thoughts, speak and understand complex communication, and articulate and explain themselves [13,14]. However, other species also possess the ability to communicate. Natural language processing's (NLP) only goals are human-machine communication and comprehension [5]. demonstrated how Neural Machine Translation (NMT) builds target words iteratively by using context words to predict words that come next. The outcomes of their experiment on Chinese to English and Workshop on Statistical Machine Translation (WMT'14) English to German translation tasks show that their approach can achieve notable improvements on multiple datasets. examined the process of Emotion Cause Extraction (ECE), which is to identify the probable reasons for specific emotions in text [13]. They demonstrated its two main shortcomings: first, it requires annotation of the emotion before extracting the cause in ECE, which significantly restricts its use in real-world scenarios; second, it ignores the fact that the two steps are mutually indicative. To extract potential pairs of emotions and corresponding causes in a document, they proposed a new task called emotion-cause pair extraction (ECPE). Their approach's efficacy and the task's viability were demonstrated by their results on a benchmark emotion cause corpus. suggested that two real-world, albeit little-studied, issues with dialogue state tracking are an excessive reliance on domain ontology and a deficiency of knowledge sharing between domains [9].

For multi-domain dialogue state tracking, they introduced a transferable dialogue state generator (TRADE), which learns to track states without any predefined domain ontology. used sub-graph features for text classification in a streaming environment [11]. They combined labelled and unlabeled graphs with features instance weighting, subgraph feature selection, and ensemble updating. They also found informative subgraphs with the least amount of redundancy and captured concept drifting in streams. Their findings demonstrate the distinct advantages of their classifier's use of minimum-redundancy subgraph features in the construction of precise classifiers through the application of instance-level weighting. According to, most graph applications are based on unlabelled graphs, so all vertices are treated as equals [8]. This makes graph isomorphism time-consuming, especially when only a portion of the graph is searched to determine the class. This highlights the need for more straightforward techniques based on feature or distance methods that require less computing power. made use of a VEAM (Vertex and Edge Approximate Graph Miner) [10].

With it, the approximate matching between edge label sets in frequent subgraph mining is included in the mining process. Also, a framework for graph-based image classification was introduced. The experimentation results show the method gets better results than graph-based image classification. Their work didn't show how this can be applied to text data [6]. developed a new architecture of subgraph feature extraction named GMADL. Dictionary learning approaches were used to extract the features of graph data and to enhance the discrimination of the model. To improve the efficiency of extraction, the analysis dictionary was designed as a bridge to generate the sparse code directly. Each sparse code represents the feature matrix of a graph. Their work did not indicate how well it would work with unlabelled data and as such the analysis dictionary may not be a proper fit for scalar-based vectors.

classified star graphs by the label category of star-centred node and the feature information of adjacent nodes [9]. Experimental results show the effectiveness of the proposed method in index creation time, space usage and subgraph query and that the use of machine learning techniques is still a key feature Given the randomness of text data, we introduce orderliness in order to identify the frequency feature of the numerical representation of the data gathered by making use of the Collatz Conjecture. Collatz Conjecture explains a sequence that eventually ends in 1. The sequence is called the hailstone sequence which is defined as "Start with any positive integer, if that integer is an odd number, then triple it and add one to get the next term, but if the selected number is even, divide it by 2 [10]. This study also implemented the six degrees of separation theory to identify that the text graph data conforms to established network rules. The six degrees of separation theory was first suggested by Hungarian writer Frigyes Karinthy, in the year 1930 [11]. It has found its application in networks, especially in social networking and suggests that the number of people a person knows grows exponentially with the number of chains in a network and so a few lines are sufficient to get to any node in a network. Network experiments carried out by Primetime and Microsoft show this theory to be true indicating that the number of connections from a node within a network to a target node does not require more than 6 degrees of connection [9].

Note that the average path length between two nodes in a random network is equal to

$$\frac{\ln N}{\ln K} \tag{1}$$

where N = total nodes and K = acquaintances per node.

Methodology

This work first makes use of the United Nations standard as a template in fusion with Boolean Logic to identify Hate speech in the Nigerian context sourced from the work of [12]. The UN standard shows that hate speech can be direct and public incitement to commit genocide, as defined by international criminal law or any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence", as defined in Article 20 (2) of the International Covenant on Civil and Political Rights.

The following steps were taken to achieve the objective of this work

- Text collation was achieved in the Microsoft Excel environment using the get external data command from the www.nairaland.com discussion forum.
- The identified Hate Speech downloaded text was imported into the Jupyter Notebook environment using the pandas library of Python programming language.
- The tokenization process was carried out using the Natural Language Toolkit (NLTK) sentence and word tokenizer methods
- Pre-processing involves the removal of stop-words and getting the new count of words present in the sentence is carried out using the NLTK stop-word methods
- Calculating the degree of separation using the result from (4) by implementing the equation is carried out.
- The text graph is implemented using the Networkx, NumPy, and matplotlib libraries with Python programming language by making use of the length of words within a sentence. They are also used to calculate the shortest path and unique path (Using Kruskal's algorithm) for each graph.
- Application of the Collatz Conjecture rule to the number of characters to derive the frequency feature is then carried out. Results are then analyzed using Equation 8

This work's framework model is summarily described in Figure 1.

Hate Speech Text Identification Template

A template for words that usually border around hate speech (similar to feature extraction) within the Nigerian-English context was developed to aid in this process. Hate Speech Variables identification with respect to this work include:

A: online discriminatory textual form of expression
B: intent to spread such form of expression
C: incite violence against a group of people
D: Justification for such violence or hatred
E: Influential origintor (huge audience)

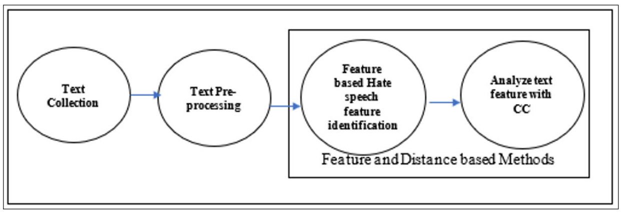

Figure 1: The 6doscc Framework for Hate-Speech Frequency Feature Identification

The five (5) factors as shown above by covers all the bases for which hate speech can be identified [12].

Hence, from our variables, we see that concerning hate speech

(A.B)+E→Hate speech (H) (2)
(A.C)+E→Hate speech (H) (3)
(A.D)+E→Hate speech (H) (4)
A(B+C+D)+E→Hate speech (H) (5)
(A.B.C.D)+E→Hate speech (H) (6)
A.B.C.D.E→Hate speech (H) (7)

Collatz Conjecture

This work applies the Collatz Conjecture principle to our graphs by taking the total cost based on the character level for each word. To identify peculiarities in hail stone numbers, we generate hail stone numbers for all numbers between 1 and 2500 due to the specification of the computer in use for the experiments. We omit the numbers 1,2, and 4 due to the loop they form and their presence for all positive natural numbers and then compare the frequency of occurrence between the Nigerian hate speech data and natural hailstone numbers. The following are the results gotten

Parameter of Performance

The Pearson's Correlation Coefficient, defined by, as if two subgraph patterns g_p and g_q, and a graph set D, the Pearson's Correlation Coefficient between g_p and g_q over the graph, set D is defined as [11].

$$\phi(g_p, g_q) = \frac{N^D N^D_{g_p,g_q} - N^D_{g_p} N^D_{g_q}}{\sqrt{N^D_{g_p}(N^D N^D_{g_p}) - N^D_{g_q}(N^D N^D_{g_q})}} \tag{8}$$

N^D denotes the total number of graphs in graph set D.
$N^D\ g_p,g_q$; $N^D\ g_p$ and $N^D\ g_q$ denotes the number of graphs in D containing g_p,g_q; g_p, and g_q respectively.

Result

We describe the results obtained after implementation in this section

Hate Speech Text Collection and Pre-Processing Result

From the 500 posts downloaded from www.nairaland.com gave a total of 104 hate speeches. Figure 5 combines the concept of hate speech alongside multiple words to identify hate speech statements. Figure 6 from the analysis shows that the recurring factors that Nigerian hate speech contains borders on criminality, intelligence, and physical attributes. Figure 6 shows its distribution from the data analyzed. Due to the limitation of the NLTK word tokenizer, preprocessing involved the concatenation of multiple words of hate speech to accommodate them in the hate speech corpus.

Analysis shows that the average amount of words used in the construction of a Nigerian hate speech sentence is 28 words. We make use of the Natural Language Tool Kit library to minimize the word vertices that will form the graph and then test the data with regard to its conformity with the 6 degrees of separation.

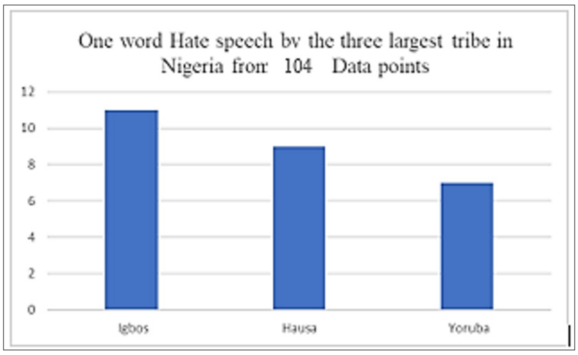

Figure 2: One-Word Hate Speech Targeted at The Three Largest Nigerian Tribe

Figure 3: Multiple Words Hate Speech Targeted at The Three Largest Nigerian Tribes

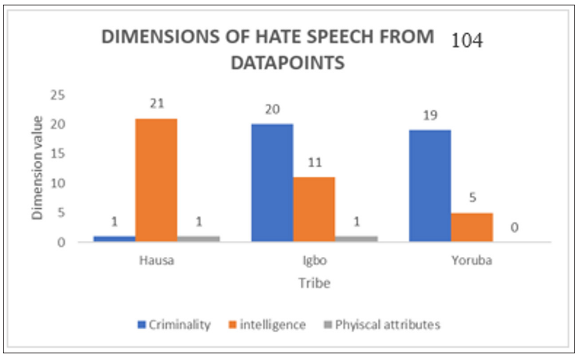

Figure 4: Dimensionality of Hate Speech

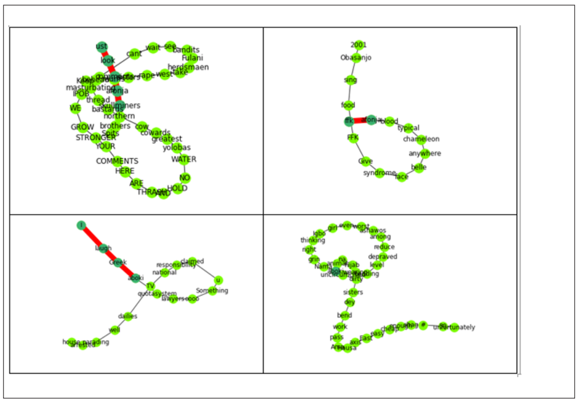

Figure 5: Cross-section of Text Graph Showing the Shortest Path

Going by the graph with a maximum number of edges, results show that the 6 degrees of separation holds for all datasets with a maximum value of 6.044394119358453 which is ≈ 6. Also, 99.03846154 of Nigerian hate speech graphs formed showed to be simple, directed and complete graphs indicating that the majority of Hate speech text does not usually come with repeated words as shown in figure 6.

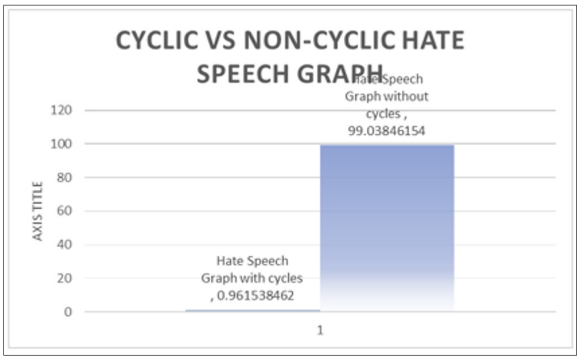

Figure 6: Cyclic vs non-cyclic Nigerian hate speech graph

This study checked to verify if the graph is a connected graph by applying Kruskal's algorithm to find a minimum spanning tree (unique path) between every set of nodes in each graph and conclude that the cost to traverse all nodes within a Nigerian text-based hate speech graph is constant. We notice that the more the number of words, the more several leave-like cyclic figures become represented in the tree graph construction as seen in Figure 7. The graph with the largest number of words at first observation shows an iris plant shape as identified by Microsoft Bing. This presents an interesting approach to text graph feature analysis for dense text graphs.

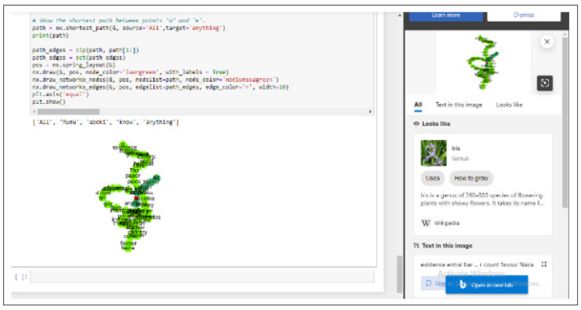

Figure 7: Graph Shape Representation and Identification Using Computer Vision

Result of Collatz Conjecture Application

After the application of the Collatz Conjecture, results show that the highest hailstone number reached was 9232. It was observed that not all natural numbers from the Nigerian Hate speech hail storm numbers are accounted for as the query returns a zero for some as can be seen in the case of 6, 9, 12, 14, 15, 18, 19 from figure 10. It was observed that there is a peculiar frequency difference between odd numbers of the Nigerian hate speech and that of the naturally occurring hail stone numbers for numbers accounted for as seen in figure 8. This is despite the observation that numbers from 92 and above are not accounted for. Pearson's Correlation Coefficient gives a value of 0.517462073 indicating a large strength of correlation figure based on the observed high disparity within odd hail stone numbers for Nigerian hate speech using figure 9 as a guide.

Table 1: Cross Section of Frequency Count of Nigerian Hate Speech Generated Hail Stone Number and Naturally Occurring Hailstone Numbers

fCount		
3	2	15
5	71	1861
6	0	13
7	2	315
8	79	1991
9	0	11
10	62	1858
11	9	815
12	0	11
13	18	911
14	0	313
15	0	11
16	45	1988
17	6	871
18	0	9
19	0	341

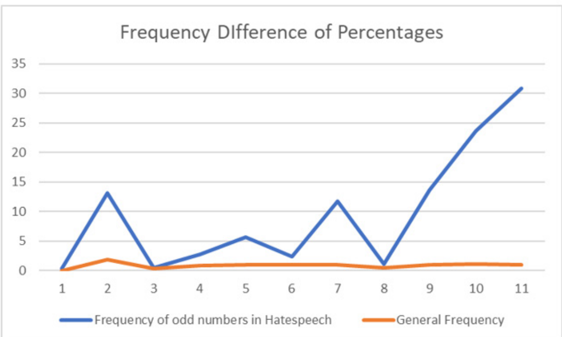

Figure 8: The Frequency Difference Between Odd Hail Stone Numbers

Strength of Association	Coefficient, *r*	
	Positive	Negative
Small	.1 to .3	-0.1 to -0.3
Medium	.3 to .5	-0.3 to -0.5
Large	.5 to 1.0	-0.5 to -1.0

Figure 9: Pearson's correlation guide [13]

Conclusion And Recommendations

Many of the Machine learning and graph approaches require mapping and annotation for their effectiveness within and across contexts often requiring more computational resources. This work introduces the 6DoSCC framework to identify the features of hate speech from a Nigerian perspective by hybridizing the strengths of both the feature-based and distance-based methods. It was observed that Nigerian hate speech does not, up to an over 99% value, contain Cycles and have no unique paths that connect all nodes and the shortest path in Nigerian hate speech graphs is irrelevant as the cost to traverse it remains constant but the same cost was instrumental to getting the odd number frequency difference between Nigerian hate speech hailstone numbers and naturally occurring hail stone numbers. This work was able to identify unique features that would help in the classification process. It is shown that the disparity in odd number frequency is strong.

Giving the uniqueness of ambiguity in natural language, it is recommended that more work should be put into expanding the capabilities of the Nigerian hate speech with more specific annotations while a larger dataset which will generate a larger cost for traversing the graph should be looked into. Also hate speech from the Nigerian perspective needs a more robust and evolving Nigerian hate speech database that can be referenced for future work and clarification.

References

1. United Nations, United Nations Strategy and Plan of Action on Hate Speech: Detailed Guidance. 2020.
2. Bakare S. Nigeria: Bills on hate speech and social media are dangerous attacks on freedom of expression. 2019.
3. Kwok I, Wang Y. Locate the Hate: Detecting Tweets against Blacks, in Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence, Wellesley. 2013.
4. Fortuna P, Nunes S. A Survey on Automatic Detection of Hate Speech in Text, ACM Comput. Surv. 2019. 51: 1-30.
5. Zhang W, Feng Y, Meng F, You D, Liu Q. Bridging the Gap between Training and Inference, Association for Computational Linguistics. 2019.
6. Zheng X, Liang S, Liu B, Xiong X, Hua X, et al. Subgraph feature extraction based on multi-view dictionary learning for sub-graph classification, Knowledge Bases systems. 2020. 1-12.
7. Blum A, Chawla S. Learning from Labeled and Unlabeled Data using Graph Mincuts. 2018.
8. Somkunwar R, Vaze VM. A Comparative Study of Graph Isomorphism, International Journal of Computer Applications. 2017. 162: 34-37.
9. Shan X, Ma J, Gao J, Xu Z, Song B. A Subgraph Query Method Based on Adjacent Node Features on Large-Scale

Label Graphs, Web Information Systems and Applications. 2019. 226-237.
10. Acosta-Mendoza N, Gago-Alonso A, Medina-Pagola JE. Frequent approximate subgraphs as features for graph-based image classification, Knowledge-Based Systems. 2011. 381-392.
11. Pan S, Zhu X, Zhang C, Yu PS. Graph Stream Classification using Labeled and Unlabeled Graphs. in ICDE. 2013.
12. PennState. The Six Degrees of Separation. 2014.
13. Xia R, Ding Z. Emotion-Cause Pair Extraction: A New Task to Emotion Analysis in Texts, Association for Computational Linguistics. 2019.
14. Nishad MT. Mathematical Proof of Collatz Conjecture, International Journal of Mathematics Trends and Technology. 2021. 67: 178-182.
15. Exploring your mind. The Six Degrees of Separation Theory. 2019.
16. Balogun O, Awodele O, Onuiri E, Nzenwata U. "ENHANCED HATE SPEECH CLASSIFICATION MODEL USING LONG SHORT-TERM MEMORY AND CONSENSUS APPROACH," American Journal of Computer Sciences and Applications. 2024. 4.
17. Laerd Statistics. Pearson Product-Moment Correlation. 2018.