

Spring 5-1-2021

## Semantic Classification of Multidialectal Arabic Social Media

Tom Rishel  
*University of Southern Mississippi*

Follow this and additional works at: <https://aquila.usm.edu/dissertations>



Part of the [Data Science Commons](#)

---

### Recommended Citation

Rishel, Tom, "Semantic Classification of Multidialectal Arabic Social Media" (2021). *Dissertations*. 1892.  
<https://aquila.usm.edu/dissertations/1892>

This Dissertation is brought to you for free and open access by The Aquila Digital Community. It has been accepted for inclusion in Dissertations by an authorized administrator of The Aquila Digital Community. For more information, please contact [Joshua.Cromwell@usm.edu](mailto:Joshua.Cromwell@usm.edu).

# SEMANTIC CLASSIFICATION OF MULTIDIALECTAL ARABIC SOCIAL MEDIA

by

Thomas H. Rishel

A Dissertation  
Submitted to the Graduate School,  
the College of Arts and Sciences  
and the School of Computing Sciences and Computer Engineering  
at The University of Southern Mississippi  
in Partial Fulfillment of the Requirements  
for the Degree of Doctor of Philosophy

Approved by:

Dr. Andrew Sung, Committee Chair

Dr. Iliyan Iliev

Dr. Bo Li

Dr. Ramakalavathi Marapareddy

Dr. Chaoyang Zhang

May 2021

COPYRIGHT BY

Thomas H. Rishel

2021

*Published by the Graduate School*



THE UNIVERSITY OF  
**SOUTHERN**  
**MISSISSIPPI®**

## ABSTRACT

Arabic is one of the most widely used languages in the world, but due in part to its morphological and syntactic richness, resources for automated processing of Arabic are relatively rare. Arabic takes three primary forms: Classical Arabic as seen in the Qur'an and other classical texts; Modern Standard Arabic (MSA) as seen in newspapers, formal documents, and other written text intended for widespread distribution; and dialectal Arabic as used in common speech and informal communication. Social media posts are often written in informal language and may include non-standard spellings, abbreviations, emoticons, hashtags, and emojis. Dialectal Arabic is commonly used in social media.

Semantic classification is the task of assigning a label to a text based on its primary semantic content. Given the increased use of dialectal Arabic on social media platforms in recent years, there is an urgent need for semantic classification of dialectal Arabic. Even compared to MSA there are few resources for automated processing of dialectal Arabic. The prior work dealing with automated processing of dialectal Arabic are limited to only one or two dialects. One of the major obstacles to doing semantic classification of multi-dialectal Arabic is the lack of a large, multi-dialectal, tagged corpus. To the best of our knowledge there are no automated processes for semantic classification of multi-dialectal Arabic social media texts.

We gather a data set of more than one million tweets collected from 449 accounts located in 12 Arabic-speaking countries. We group those tweets into 21,791 documents by country, account, and month. We first construct a query to represent a particular semantic concept. Then, using Latent Semantic Analysis (LSA) we rank the documents by semantic similarity to the query. Next, we use that ranking to train a deep neural

network classifier to identify documents whose text is semantically similar to the query. Experiments demonstrate an overall accuracy of 98.075% and a positive accuracy of 88.178% have been achieved by this approach to semantic classification of multi-dialectal Arabic. The source code and the data set are provided on GitHub at <https://github.com/therishel/ArabLeader>.

## ACKNOWLEDGMENTS

I would like to thank Dr. Bikram Banerjee for helping me organize my PhD completion plan and advising me regarding requirements along the way. I would also like to thank my dissertation committee: Dr. Andrew Sung, Committee Chair, Dr. Iliyan Iliev, Dr. Bo Li, Dr. Ramakalavathi Marapareddy, and Dr. Chaoyang Zhang. I would particularly like to thank Dr. Li for his generous help, guidance and encouragement throughout the preparation of this dissertation and my oral defense presentation. I would also like to thank Dr. Iliev for his continued support as my research colleague and mentor for the past three years.

## DEDICATION

This dissertation is gratefully dedicated to Dr. A. Louise Perkins. I would never have considered pursuing a PhD in Computer Science without the encouragement of Dr. Perkins. She first saw in me the potential for this accomplishment. She actively nurtured that potential, guiding and mentoring me in completing my master's degree and hiring me for my first academic position. She continued to support me as a research colleague and mentor and inspired this work, as well as future work that I still hope to complete.

## TABLE OF CONTENTS

ABSTRACT .....	ii
ACKNOWLEDGMENTS .....	iv
DEDICATION .....	v
LIST OF TABLES .....	viii
LIST OF ILLUSTRATIONS .....	ix
LIST OF ABBREVIATIONS .....	x
CHAPTER I - INTRODUCTION .....	1
CHAPTER II - LITERATURE REVIEW .....	8
2.1 Sentiment analysis .....	10
2.2 Neural Network Semantic Classification .....	11
2.3 Dialectal Arabic Processing .....	13
CHAPTER III - BUILDING THE DATA SET .....	14
3.1 Candidate Twitter account selection .....	14
3.2 Tweet selection and downloading .....	15
3.3 Tweet organization and processing .....	15
CHAPTER IV - LATENT SEMANTIC ANALYSIS .....	17
4.1 Removing stop words and applying TF-IDF .....	17
4.2 Latent Semantic Analysis as ground truth .....	22
4.2.1 Term Document Matrix .....	25



4.3 Singular Value Decomposition .....	29
4.4 Building a query.....	30
CHAPTER V - NEURAL NETWORK TRAINING AND TESTING .....	32
5.1 Finding a threshold .....	32
5.2 Neural network architecture and training trials .....	36
CHAPTER VI – RESULTS.....	43
6.1 Multi-dialectal Arabic social media data set.....	43
6.2 LSA scoring of documents.....	44
6.3 Neural network classification of documents.....	48
CHAPTER VII – CONCLUSION, DISCUSSION, AND FUTURE WORK .....	51
7.1 Discussion .....	51
7.2 Future work.....	52
7.3 Conclusion .....	53
APPENDIX A - Arabic Stop Words.....	55
APPENDIX B – Arabic query text .....	59
APPENDIX C - Example of document tagged as semantically related to terrorism.....	62
REFERENCES .....	124

## LIST OF TABLES

Table 4.1 The term document matrix generated from three example sentences. ....	26
Table 4.2 Some examples of common “stop words” in English.....	27
Table 4.3 Tabulated word frequencies.....	28
Table 5.1 Number and size of hidden layers.....	39
Table 6.1 Top results of terrorism query .....	44
Table 6.2 Top results of religion query .....	45
Table 6.3 Top results of soccer query .....	46
Table 6.4 Descriptive statistics for cosine similarities of documents to terror query .....	48
Table A.1 List of Arabic stop words.....	55
Table B.1 Arabic terrorism query terms .....	59
Table B.2 Arabic religion query terms .....	59
Table B.3 Arabic soccer query terms.....	60

## LIST OF ILLUSTRATIONS

Figure 5.1 Histogram of cosine similarities to terrorism query document .....	35
Figure 5.2 Diagram of our neural network architecture .....	37
Figure 5.3 Loss curves without learning rate decay .....	41
Figure 5.4 Loss curves with learning rate decay.....	41
Figure 6.1 Scatter plot of cosine similarities of twitter documents to query document ...	48
Figure 6.2 Loss curve for our most accurate neural network model .....	49

## LIST OF ABBREVIATIONS

<i>ALT</i>	Arab leaders' tweets
<i>BOW</i>	Bag of words
<i>GRU</i>	Gated Recurrent Unit
<i>KNN</i>	K-Nearest neighbor
<i>LSA</i>	Latent Semantic Analysis
<i>MSA</i>	Modern Standard Arabic
<i>MSE</i>	Mean Square Error
<i>NB</i>	Naïve Bayes
<i>NLP</i>	Natural language processing
<i>SVD</i>	Singular value decomposition
<i>SVM</i>	Support vector machine
<i>TF-IDF</i>	Term frequency-Inverse document frequency

## CHAPTER I - INTRODUCTION

Arabic<sup>1</sup> is one of the world's most widely spoken languages. Arabic ranks fifth among both the most commonly spoken languages and most used on the internet. Over 422 million people speak Arabic around the world – over 6.6% of the global population [1]. In spite of this popularity, machine processing resources for Arabic remain relatively rare and are still in the early stages of development. For example, a recent article noted that Apple's voice assistant, Siri, barely gets by in Arabic while Amazon's Alexa and Microsoft's Cortana don't speak Arabic at all [2]. More formal research shows a similar picture [1] [3]. As recently as 2018 Boukil et al. [1] noted, "It [Arabic] is one of the most challenging languages in the world with its rich morphology, its complex syntax, and its difficult semantics. This makes its analysis and automatic processing very hard and complex."

This dissertation only deals with written language. In general, written Arabic may be divided into three large categories, classical Arabic as seen in the Qur'an and other classical texts, Modern Standard Arabic (MSA) which is typically used for all formal written communication such as newspapers, magazines, contracts, government documents, etc., and dialectal Arabic [4]. Until relatively recently dialectal Arabic was

---

<sup>1</sup> Disclaimer: The author of this dissertation does not speak, read, or write Arabic. Therefore, any Arabic presented in this dissertation is from an external source and has not been verified, screened, or otherwise checked for accuracy or appropriateness. For this reason, most examples used to illustrate concepts will be presented in English. Examples of Arabic that are included are taken directly from the research.

almost exclusively spoken. If something was written or published it was usually written in MSA. The rise of social media and the now ubiquitous act of extremely casual writing and publication on social media sites (e.g. Twitter, Facebook, online forums, etc.) have generated large volumes of written dialectal Arabic. This has led to an urgent need for automated processing of text in dialectal Arabic.

Automated processing of dialectal Arabic poses many challenges. The first is that it is not a single language. Estimates vary, but some authors identify 22 different dialects [5]. Most research focused on automated processing of Arabic deals exclusively with MSA. In recent years more attention has been given to research on dialectal Arabic, however, it remains an emerging area of research. The challenges of automated processing of dialectal Arabic compound those of MSA. Unlike Modern Standard Arabic (MSA), dialectal Arabic has no orthographic standard [6]. In various dialects the same word may be written in different ways.

Most written dialectal Arabic is in the context of social media. The lack of writing standards and the specific inclusion of non-standard communication methods (e.g. intentional misspellings, repeated characters, hashtags, abbreviations, emoticons, and other symbols ('<3', '/s')) add nuance and overtones to social media posts that prove challenging for automated systems in any language. When these challenges are layered over the pre-existing challenges of multi-dialectal Arabic, the difficulty becomes substantial. We will detail some of these challenges below and present our approach to processing multi-dialectal Arabic social media posts.

Aggarwal and Zhai [7] define text classification as:

“The problem of classification is defined as follows. We have a set of training records  $D = \{X_1, \dots, X_N\}$ , such that each record is labeled with a class value drawn from a set of  $k$  different discrete values indexed by  $\{1 \dots k\}$ . The training data is used in order to construct a classification model, which relates the features in the underlying record to one of the class labels. For a given test instance for which the class is unknown, the training model is used to predict a class label for this instance.”

As this definition makes clear, one of the requirements of text classification is a set of training records labeled with a class value. For our task of semantic classification of multi-dialectal Arabic social media, *the lack of a well-developed training data set is a significant challenge*. In this work we develop a multi-dialectal Arabic social media data set, tag it using a semi-automated process, and use it as a training set for a neural network semantic classifier.

Sentiment analysis is closely related to semantic classification. Sentiment analysis is a classification problem in which the classes are defined by the emotional tone of the content. Typically, sentiment classification divides texts into “positive” and “negative” classes. A common example of sentiment classification is the classification of film, television, or product reviews into groups of those which praise the product or those which are critical of the product. Like semantic classification, sentiment classification relies on some ground truth that characterizes the fundamental view of the review. In sentiment analysis this ground truth is often represented by a numeric rating system of some kind that accompanies each review (e.g., a five star review).

One problem with semantic classification is the identification and validation of this ground truth. There are two broad categories of approaches for automated semantic classification. The first is the manual tagging of documents with semantic tags. The manually tagged documents are then gathered into corpora and published for use by the research community. The second is the use of thesauruses or lexical databases such as the Arabic WordNet [8] [9]. *As compared to other languages with a similar number of speakers, there are not many language processing resources for Arabic [10]. The majority of the resources that are available are for MSA. There is a substantial lack of freely available dialectal Arabic language resources.* In [11], Zaghoulani highlights the lack of easily accessible dialectal Arabic corpora. At that time (2017) he identified only two easily accessible NLP resources for dialectal Arabic [12] [13]. This lack of resources is a contributing factor to the relative lack of natural language processing (NLP) work done in dialectal Arabic.

The lack of NLP resources for dialectal Arabic is exacerbated by multiple factors. One challenge is that until relatively recently dialectal Arabic was nearly exclusively a spoken language. In general, if something was written down in Arabic, it was written in MSA. This includes newspapers, contracts, government documents, textbooks, etc. The lack of readily available large collections of digital dialectal Arabic text made developing resources for automated processing of dialectal Arabic nearly impossible. The rise of social media over the past several years has provided both the digital text required to develop resources as well as the motivation for effective semantic classification tools for dialectal Arabic. A second challenge is the number, variety, inconsistency, and geographic distribution of Arabic dialects. Estimates of the number of distinct dialects



range from seven to 22 [5] or more depending on how various authors group or divide dialects. Most authors recognize that there are not clear boundaries between dialects and that they merge, combine, and separate in various ways and for various reasons over time. Therefore, any clear division of dialectal Arabic into distinct groups is likely to be somewhat arbitrary. While these distinctions are widely acknowledged, it is also true that all forms of Arabic are clearly identifiable as Arabic and are more closely related to each other than to any other language.

Common conventions in social media posts add to the difficulty of automated processing of dialectal Arabic. Social media posts are often very informal and thus, when written in dialectal Arabic may not include diacritic marks or other more formal encodings. Repetition of characters and use of other non-standard spelling is common in social media. The use of slang titles and terms is also common. Social media posts often use abbreviations, hashtags, emojis, and various colloquialisms to communicate sometimes subtle nuances that may strongly affect the specific message of a post. Examples from English include the use of the abbreviation “LOL” to indicate that a statement should not be taken seriously; one or more of a possibly infinite number of hashtags (e.g. #dogsofinsta), numbers to indicate opinions (e.g. “1/5 would not recommend” or “11/10 good boi” in reference to a photo of a dog); the use of the “/s” tag to indicate “end sarcasm”, thus suggesting that the previous text is sarcastic in nature and should not be understood at face value; and the use of the “cup of tea” emoji to indicate skepticism about the truth of a statement. These conventions vary widely by language, culture, age, and sub-group and change quickly over time as trends rise and fall.

A final challenge is the actual encoding used for some characters depending on the keyboard mapping on a particular device. Different keyboard mappings encode different shapes of the same Arabic letter or diacritics as illustrated by the following excerpt:

“Besides, Arabic NLP applications face the challenge of encoding, which is the representation of the language symbols in computers, especially when representing the different shapes of the same letter or the diacritics. Unicode is the actual current standard for encoding a large number of language symbols including Arabic, such as the Arabic letter ك) U+0643) and the Persian ك) U+06A9) using the same shape ك, which adds confusion when the Arabic letter is written using a Persian keyboard.” [14]

We suggest that these challenges make the development and use of formal ontologies, dictionaries, thesauruses, and grammars impractical for semantic classification of social media posts in multi-dialectal Arabic. A variety of approaches have been implemented in other attempts to classify this data. They include several different varieties of stemming, character replacement, removing repeated characters, named entity recognition and resolution, stop word removal, removal of URLs, removal of numbers and non-text characters including emoji, removal of punctuation including emoticons, the use of dialect-specific corpora, and partial translation or word replacement. In addition to the current impracticality of developing and maintaining such resources for all of the various dialects of Arabic, research suggest that these techniques may not provide the most accurate results. In the literature review below we will present previous work that found limited success from such approaches. One of the goals of this work is to create a data set of multi-dialectal Arabic social media posts organized to facilitate various levels of analysis. We also present an approach for automatically

tagging elements of the data set with a semantic label generated through Latent Semantic Analysis (LSA).

The remainder of this dissertation is organized as follows. Chapter II provides a review of the relevant literature. *Our review found no previous work attempting semantic classification of multi-dialectal Arabic social media.* Chapter III presents the process that we used to build and organize our data set. We gathered over 1.2 million tweets from 449 accounts in 12 countries and organized them into 21,791 documents grouped by country, account, and month. Chapter IV describes LSA and our use of it to order our data set by semantic similarity to a query document. Chapter V explains our neural network architecture and the model we trained and tested to classify documents as semantically related to our query. Chapter VI reviews our results. In Chapter VII we present discussion of our work, future work, and conclusions.

## CHAPTER II - LITERATURE REVIEW

In 2015 Hmeidi et al. [15] conducted a comprehensive comparative study of automatic Arabic text categorization in which they reviewed approximately 39 publications related to text classification. Of these 39 publications, 21 consider text classification of Arabic (the other references are provided to demonstrate the advanced state of the art in English). They only consider MSA and do not deal at all with dialectal Arabic. The publication dates of the 21 papers dealing with MSA range from 1999 to 2013. In addition to reviewing work related to text classification of MSA documents, the authors also conducted experiments to compare Arabic text classification using Naïve Bayes (NB), Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Decision Tree (J48), and Decision Table classifiers. They concluded that using a SVM with the Light10 stemmer produced the most accurate classification although it should be noted that a NB classifier performed nearly as well as the SVM classifier. It is also worth noting that stemming in general gave very modest improvements ( $< \sim 1\%$ ) compared to using the original documents and, in some cases, decreased the accuracy of the classification. A different comparative survey of MSA text classification from 2013 [16] also found that a SVM produced the most accurate classification of 97% on the Islamic Topics Dataset but dropped to 61% for the Arabic Poems data set. This variance in performance demonstrates the challenges of performing text classification on non-standard text, even within the formally recognized structures of MSA.

Both [17] and [18] found that the use of semantic relations was superior to more grammar-based approaches using, for example, n-grams, stemmers and part-of-speech taggers. Yousif, Elkabani, Samawi, and Zantout found a more than 12.6% improvement

in the F1 measure when using semantic relations over other methods. Their approach used semantic relations defined by Arabic WordNet [19]. Alowaidi, Saleh, and Abulnaja also used Arabic WordNet [17]. Their position is that a fundamental weakness of word-based approaches is that they are “semantically weak” in that they do not consider semantic relationships between words. In other words, if two synonymous but different words are used in two different documents those documents would not be identified as semantically related. They tested three different approaches. In one they added the terms from the identified WordNet synset to the original term from the data set. In another they replaced the original term with all of the terms from the identified synset. In the third they used the extracted concept only without any of its associated words. In a separate study the same authors found that classification using a List of Pertinent Words improved accuracy of classification more than using a List of Pertinent Synsets, a Bag of Words, or a Bag of Concepts. [20]. The biggest barrier to implementing such approaches when classifying dialectal Arabic is that there are few if any similar resources for dialectal Arabic. There is a WordNet for the Iraqi dialect [21] and work continues on developing language resources for dialectal Arabic.

In related work, Khalil, Halaby, Hammad, and El Beltagy [22] found that a Bag of Words approach was most accurate when doing sentiment analysis with MSA. They evaluated weighting schemes, n-gram combinations, feature selection using information gain, and simple data pre-processing steps on three different Arabic data sets using two different Bayesian classifiers as well as an SVM. "Our experiments showed that simple text cleaning and filtration steps like replacing links and mentions with placeholders,

removing English characters, and Arabic characters normalization do not seem to affect the accuracy significantly." [22]

LSA is a semi-automated procedure that identifies semantic concepts in a corpus of documents and that may be used to score either the constituent documents or a query document that is external to the corpus by identifying the extent to which each of the identified semantic concepts contributes to the document being scored. In 2017, Al-Anzi and AbuZeina [23] used cosine similarity and LSA to enhance classification of a collection of MSA documents. They evaluated nine different classifiers and found that an SVM gave the highest accuracy of 84.75%.

## **2.1 Sentiment analysis**

Sentiment analysis is closely related to semantic classification. Sentiment analysis is a classification problem in which the classes are defined by the emotional tone of the content. Typically, sentiment classification divides texts into “positive” and “negative” classes. A common example of sentiment classification is the classification of film, television, or product reviews into groups of those which praise the product or those which are critical of the product. Like semantic classification, sentiment classification relies on some ground truth that characterizes the fundamental view of the review. In sentiment analysis this ground truth is often represented by a numeric rating system of some kind that accompanies each review (such as a five-star review).

Several authors have conducted studies of sentiment analysis both in MSA and in dialectal Arabic. In 2019 AlFarasani, AlHarthi, and AlHumoud used semantic classification of Twitter data to identify road conditions in Saudi Arabia [14]. Alowaidi, Saleh, and Abulnaja noted that Bag of Words (BOW) approaches to sentiment analysis

are semantically weak [17]. They used Arabic Word Net to supplement individual words with their associated semantic concepts and then evaluated their data using a SVM as well as a NB classifier. They found that the inclusion of semantic concepts improved the performance of their classifiers by 4.48% and 5.78% respectively.

Tartir and Abdul-Nabi [24] studied sentiment analysis in the Jordanian Arabic dialect. They created an ontology of 24 words divided into positive and negative sentiment classes. Using this ontology, they gathered and classified 1,000 tweets into either positive, negative, or neutral classes across three topic areas. They achieved an average precision across their topic areas of 75% and an average recall of 72%. They found these results to be consistent with other studies that they reviewed. Their work is representative of studies of dialectal Arabic in that the sample size is relatively small, they created their own hand-scored data set for ground truth comparison, they used a classification tool which is small and specific to their work, and their study is limited to a single dialect.

## **2.2 Neural Network Semantic Classification**

In 2010 Harrag and Al-Qawasmah [25] used a neural network to classify a collection of 453 Arabic documents arranged in 14 categories. Their data set of 5743 words was taken from a corpus of prophetic traditions or Hadiths' collected from the Prophetic encyclopedia (Alkotob Altissâa, "The Nine Book"). They chose that data in part because in its original form it was organized into the 14 categories, thus providing them with a ground truth to test against. After preprocessing including stop word removal and light stemming, they had 1065 unique tokens. They used Term Frequency-Inverse Document Frequency (TF-IDF) term weighting and Singular Value Decomposition

(SVD) to reduce the size and dimensionality of the data before building a forward-feed neural network model and training it on one half of their data set. The maximum number of training documents that they had for any single category was 34. Their best results were F1 scores of 49% for a neural network trained on un-reduced vectors (1065 terms) and 53% for a network trained on vectors produced by SVD (530 dimensions). In related work, Harrag, El-Qawasmah, and Al-Salman evaluated various types of stemming as a feature reduction technique to preprocess Arabic text for classification by a neural network [26]. They found that the use of a light stemmer gave improved results with a neural network classifier.

In more recent work, Boukil et al. used a convolutional neural network (CNN) to classify a collection of MSA documents [1]. They found that, depending on the size of the data set, the accuracy of their CNN varied from 86.3% on a data set of 27,932 documents, to 92.94% on a data set of 111,728 documents. Elnagar, Al-Debsi, and Einea compared several deep learning models for both single-label and multi-label text classification of MSA text gathered from three popular Arabic news websites [27]. Most of their data was tagged by the sources. They supplemented the tags where necessary, performed some mild data cleaning to remove Latin characters and punctuation marks, and created two data sets; one in which each article is tagged with only a single label to be used for single topic categorization and a second in which the articles are tagged with multiple labels to be used for multi-label text categorization. They achieved their best accuracy of 96.94% for single-label categorization with attention-GRU and 88.68% for multi-label categorization.



### **2.3 Dialectal Arabic Processing**

In a 2015 survey of NLP work with dialectal Arabic, Shoufan and Al-Ameri reviewed 91 publications published between 2000 and 2015. They organize their findings into four categories of research. Those categories are basic language analyses, building language resources, semantic-level analysis and synthesis, and identifying Arabic dialects. Of these four categories, the third, semantic-level analysis and synthesis, most closely relates to the current work. The authors identified a total of 25 articles within the category of semantic-level analysis and synthesis. Of that total, they categorized 15 as related to machine translation and ten as “others”. Since our work does not involve machine translation, we did not review the machine translation papers. Of the ten “others”, four deal with sentiment analysis ( [28], [10], [29], [3]) rather than semantic analysis. An additional three articles focus on named entity recognition in dialectal Arabic ( [30], [31], [32]). The remaining three articles deal with mining slang comments [33], summarization of twitter data [34], and information retrieval in dialectal Arabic [35]. In reviewing 91 articles published over a fifteen year span the authors did not identify a single article that is directly related to semantic classification of dialectal Arabic. The work that most closely parallels semantic classification is summarization, however, in summarization the topic is derived from the content of the text(s) rather than being provided externally. Our literature review also found no research using a neural network to do semantic classification of social media posts in multi-dialectal Arabic.

## CHAPTER III - BUILDING THE DATA SET

Our first step was to build a data set named Arab Leaders' Tweets (ALT) data set. As noted in the literature review there is a lack of automated language processing resources for Arabic in general and dialectal Arabic more specifically. Furthermore, the majority of the data sets we reviewed for dialectal Arabic cover only a single dialect. We did identify one resource that includes two dialects . We construct a data set of more than one million social media posts from 14 Arabic-speaking countries. The data set is organized by country, individual, and month to allow for various levels of analysis. ALT is created via the following process.

### **3.1 Candidate Twitter account selection**

We conducted an extensive search of Twitter accounts from various Arab speaking countries and their social media presence. We created a dedicated Twitter account specifically for use in this research. We identified 449 individual Twitter accounts from 12 countries (Bahrain, Egypt, Iraq, Kuwait, Lebanon, Libya, Oman, Qatar, Saudi Arabia, Syria, UAE, and Yemen). These individuals were identified as political, social, or religious leaders and have the most significant Twitter following from these countries with the number of followers ranging from the tens of thousands to millions. They have a significant reach to large numbers of people through their social media presence. The selected countries have some of the largest Arab speaking populations, over a wide range of dialects for the analysis, as well as divergent socio-political conditions.

### **3.2 Tweet selection and downloading**

We used the Tweepy Python library [36] to download the most recent 3,240 tweets for each identified Twitter account. The algorithm we developed harvests the most recent tweet first and works backward chronologically until 3,240 tweets have been downloaded or there are no more available tweets, whichever comes first. Due to variations in the frequency of tweets the time frame covered by this approach varies from account to account. Some accounts did not have 3,240 tweets. In those cases, we downloaded all of the tweets for that account.

### **3.3 Tweet organization and processing**

The total number of tweets in the database is more than 1.1 million. We produced Unicode files organized by country, then by author, then by month. This allows further analysis with NLP tools at the individual or country level, and over time. We define a document as a UTF-8 encoded text file containing the text of all of the tweets of a given user during a specific month. Using UTF-8 encoding allowed us to preserve the tweets in the original language of the user including URLs, emojis that have UTF-8 encodings, and all other symbols encoded with UTF-8 Unicode. Saving the tweets into files, as opposed to harvesting and analyzing tweets in real time, allows us to build an archive of tweets over time that may prove useful to analyze longer term patterns of social media activity within or among the twitter users in our data set. We then iterate over the parent folder containing all of the tweets, read each file into a single string, and construct a Python list where each element of the list is a document. We then construct a Python dictionary from the remaining text and build our corpus using this dictionary.

Grouping the tweets by month offers a solution to the general sparseness of social media data (most Twitter users do not post every day), as well as the short statement problem when working with automated textual analysis [37]. The data allows us to track subtle changes in expression over an extended period of time.

Considering multiple tweets within a given month as a single document allows us to both consider the activity as a time series and to minimize the impact of one of the challenges of NLP with Twitter data, the short length of individual tweets. This approach provides a counterbalance to smooth the distribution of tweet frequency between those users who tweet very often and those who tweet less often. If we consider individual tweets as documents, then those users who post very frequently take on a proportionally greater weight in the document library. Since we aggregate tweets by month the frequency of tweets converts into greater document length. Our data set contains an average of 96,144.5 tweets per month from all accounts. The maximum number of tweets that we recorded for any one month was 134,006 while the minimum recorded for any one month was 70,615. Our current data set consists of 21,791 documents.

## CHAPTER IV - LATENT SEMANTIC ANALYSIS

After collecting tweets and organizing them into files as described above, our next task was to perform LSA to organize the data set by semantic vectors. First, we removed the stop words. Then we created a term by document matrix populated with the TF-IDF weighted values of each remaining term. We decomposed that matrix using singular value decomposition to create a hyperdimensional vector space containing the semantic vectors representing a selected number of concepts. Next, we created a query of terms related to terrorism, converted the query into a semantic vector, and compared it to the documents in the newly created semantic space, ranking the original document vectors by decreasing cosine similarity to the query vector. The result was a sorted list of document IDs ranked by semantic similarity to the query that we had constructed. We have provided the list of terms that we used as the query document in Appendix B.

This ranking provides a weak semantic ordering that does not rely on manual tagging of the data set or on specific terms. It does consider semantic content and the semantic relationships among various terms within the documents without requiring the use of an ontology, thesaurus, dictionary, or wordnet. Furthermore, it is updateable with relative ease, so may be maintained to remain current with social media terms, slang, abbreviations, emoticons, etc. We used the Gensim library [38] to implement LSA in Python. The remainder of this chapter will detail each step in this process.

### **4.1 Removing stop words and applying TF-IDF**

To increase the accuracy of our semantic analysis, we use a two-stage process to decrease the influence of frequently occurring words. The first stage is the standard

approach of removing stop words. The second is adjusting the weight of the remaining terms using TF-IDF.

Stop words are words that occur with great frequency in a language. Such words do not contribute to the information encoded in a text and are generally not useful for classification or discrimination among texts. Our focus is on identifying activity within a narrow range of topics, thus removing stop words is necessary to improve the accuracy of the algorithm through greater topic discrimination. We downloaded a list of Arabic stop words for this purpose. After reading the files into our data structure the first step in processing our data is the removal of any words on the list of stop words. Our stop word list is in MSA and therefore may miss many dialectal words that would be included as stop words if we had access to such a list in all dialects. We have provided the list of Arabic stop words that we used in Appendix A.

Due to variations in vocabulary, spelling, character encoding, and formality across dialectal Arabic social media posts, removing stop words alone is not an effective approach when working with social media posts in dialectal Arabic. The removal of stop words using a list does not address possible alternate spellings or encodings of words across dialects of Arabic. Arabic includes the use of diacritic marks that alter the meaning of the words to which they are applied. Native speakers often omit the diacritic marks when typing text in social media posts. Thus, the same word may be encoded differently, with and without the diacritic marks, in social media. The inclusion or exclusion of diacritics would change the UTF-8 encoding even though the meaning of the post would not change. The informal encoding may prevent very common words from being identified for removal by our stop word list. The same word may be written in different

ways in different Arabic dialects. This further complicates the use of a list of stop words. Another practice that decreases the effectiveness of stop word removal is the use of various keyboards for writing Arabic. Furthermore, social media posts often use abbreviations, repeated letters, and other non-standard spellings that make a stop word list less effective.

The second stage in our term weighting management process is using Term Frequency-Inverse Document Frequency (TF-IDF) weighting to reduce the influence of terms that occur frequently within the corpus. TF-IDF was first proposed by Salton and Buckley in 1989 [39]. TF-IDF is based on the idea that documents are represented by term vectors of the form

$$D = (t_i, t_j, \dots, t_p)$$

where each  $t_k$  represents a term from some document  $D$ .

Common measures of the effectiveness of an information retrieval approach are recall and precision. Recall is the ratio of the number of relevant items retrieved to the total number of relevant items in the corpus. Precision is the ratio of the number of relevant retrieved items to the total number of retrieved items. Recall is generally calculated by

$$recall = \frac{tp}{tp + fn}$$

where  $tp$  represents the true positive items and  $fn$  represents the false negative items.

Precision is generally calculated by

$$precision = \frac{tp}{tp + fp}$$

where  $tp$  represents the number of true positive items and  $fp$  represents the number of false positives [40].

The conflicting requirements of recall and precision necessitate the use of a composite term weighting value that includes both recall and precision. This composite should take into account three factors:

- 1) Frequently occurring terms increase recall, so a weighted measure of term frequency (TF) should have a place. TF is generally calculated as:

$$TF_t = \frac{\text{number of times a word occurs in a text}}{\text{total number of words in the text}}$$

- 2) Term frequency alone is not sufficient. Especially if the terms are distributed evenly across the entire corpus rather than concentrated in a few documents, they cause many documents to be retrieved, thus decreasing precision. Inverse document frequency (IDF) compensates for this tendency. IDF is generally calculated as:

$$IDF_t = \log \left( \frac{\text{total number of documents in } C}{\text{number of documents containing } t} \right)$$



where  $C$  is a collection of documents (a corpus) and  $t$  is a term that occurs in  $C$ .

- 3) Our goal is to identify terms that discriminate among documents within a corpus. The terms that discriminate the best occur frequently in selected documents, but infrequently across the corpus.



Thus we use TF-IDF, which is simply the product of  $TF_t * IDF_t$ , which balances precision and recall and more heavily weights those terms with the most power to discriminate among documents within a corpus.

One effect of the use of TF-IDF with social media posts that have not been heavily pre-processed is the reduction in weight of terms that are unique or nearly unique while preserving the weight of those that are used by convention, even if that convention is undocumented or has recently emerged. For example, if a document contained an intentional misspelling (e.g. haaaappppy) that is relatively unique, that term would have a very low term frequency and would thus generate a low TF-IDF value. The low value of that specific string would not affect the value of either the same word spelled correctly, or a term that is non-standard, but not misspelled (e.g. covfefe).

This also applies to the use of emojis. Emojis have become somewhat standardized to those symbols that have corresponding Unicode values. Thus, if a given emoji is consistently associated with documents in particular semantic context (e.g.  associated with soccer or football), but does not occur equally frequently in other documents across the corpus, that character would generate a high TF-IDF value and could be used as a discriminating term for semantic classification. Conversely, an emoji that occurs with approximately equal frequency across semantic groupings (e.g. ) would generate a low TF-IDF value and would be less heavily weighted as a discriminator.

A potential weakness of this approach when dealing with multi-dialectal Arabic is the possibility that a term that is unique to a particular dialect may not occur frequently enough to contribute its true value to the semantic classification. This is more likely to be

the case if the number of words in each dialect is not balanced across the corpus (i.e., if the corpus contains 25,000 words in the Egyptian dialect and only 1,000 words in the Tunisian dialect, terms that are distinct to the Tunisian dialect but which do not occur in the majority of the Tunisian documents may not be weighted heavily enough to be considered as a discriminator). One possible remedy for this is the inclusion of a weighting factor based on the portion of the corpus represented by each dialect. This approach assumes that the various dialects can be accurately identified during processing.

#### **4.2 Latent Semantic Analysis as ground truth**

One of the fundamental challenges of semantic classification of multi-dialectal Arabic is the lack of a data set representing the ground truth against which to test the classification. This has been noted by many authors. The problem is compounded by the fact that the majority of written dialectal Arabic is on social media and is thus very informal language including abbreviations, emoticons, emoji, slang terms, hashtags, and various other atypical language constructs. Furthermore, given the rapid pace of change in social media trends, terms, symbols, and topics, it is likely impossible to create a comprehensive and current dictionary or thesaurus of terms used on social media. Two examples from English that illustrate this are the rapid spread and subsequent decline of the term “covfefe” in June 2017 and the similar spread of the hashtag #metoo in October 2017 and following months. Most common approaches to the noisy data of multi-dialectal social media text are stemming to reduce words to their root meanings, removal of hashtags, removal of URLs, removal of emojis, removal or replacement of emoticons, removal of repeated characters, and removal of all punctuation which would include semantically significant symbols such as “/s” to indicate sarcasm. In less complex

languages such as English these approaches have been shown to increase the accuracy of text classification. They have not been as effective in Arabic. In 2015 Khalil, Halaby, Hammad, and El Beltagy [22] found that techniques including replacing URLs and mentions, removing English, and normalizing Arabic (i.e., stemming, repeated character replacement, standardizing spellings) did not significantly improve the accuracy of their classification. Furthermore, more widely studied languages have large, well-developed corpora representing a gold standard against which to compare results. There is no such corpus for multi-dialectal Arabic.

Our approach to this problem is to use Latent Semantic Analysis to score a set of documents by similarity to a query document. If the query document is constructed to represent a semantic class, then the similarity scores may be used to rank the documents by similarity to a semantic topic. That ranking can then be used as a ground truth to train a neural network classifier to identify social media documents related to the chosen semantic topic. This approach has been tried by others with success in MSA [25]. Harrag and Al Qawasmah used SVD as a feature reduction technique before classification of a 5,743-word corpus into 14 categories. Their average F1 score when using SVD was between 22% and 53% (averaged across the various dimensionalities of their experiments).

In this dissertation we view NLP as a mapping between the written textual representation of a language and its associated semantic meaning. Latent Semantic Analysis takes advantage of the mapping between terms and semantic content across a library of documents to identify ideas or concepts expressed within said library. Since the technique (described below) does not rely on a 1:1 identification of semantic content with

terms, it allows a more flexible and accurate semantic matching than term-matching queries. Furthermore, it does not rely on dictionaries, ontologies, thesauruses, WordNets, or other such pre-formulated language resources except for a single query document used to establish the semantic context and rank the semantic document vectors generated by SVD.

Latent Semantic Analysis constructs a term-by-document matrix for the library to be analyzed. A weighting factor equalizes term frequencies across documents. The result is a large, sparse matrix  $X$  which is populated with the weighted frequency of terms within documents. This matrix is decomposed into the product of three other matrices:  $X = T_o S_o D_o'$  such that  $T_o$  and  $D_o$  have orthonormal columns and  $S_o$  is diagonal. This is the singular value decomposition of  $X$ . Each of the component matrices is of full rank. The singular values in  $S_o$  are ordered by size from largest to smallest. The largest  $k$  values are kept, and the remainders are set to zero. The new matrix,  $X_{hihat}$ , is approximately equal to  $X$  and is of rank  $k$ . We remove the zeroed-out rows and columns of  $S_o$  as well as the corresponding columns of  $T_o$  and  $D_o$ . The results of this operation are two new matrices,  $T$  and  $S$  such that  $X \approx X_{hihat} = TSD'$ . This model gives us the  $k$ -most frequently occurring topics from the original library. These topics are semantic ideas represented by one or more terms in the original library.

The terms occur with approximately equal frequency in the documents from which they were drawn. The singular value groups them because of their frequency co-occurrence. Their similar frequency also provides the rationale for their semantic similarity. The underlying assumption is that if a set of terms occurs together with a similar relative frequency across a set of documents, then those terms are somehow

related not only to the containing documents, but also to each other. Further, they are assumed to represent a related idea or set of ideas.

#### 4.2.1 Term Document Matrix

A term/document matrix is a matrix composed of columns of documents and rows of the terms that occur in each of the documents. Table 1 is the term/document matrix generated as a result of the following three–document set (These short examples are for illustrative purposes only). The numerical cell values are scaled frequencies of words within documents. The last row is the sum of the scaled frequencies. Scaling normalizes term frequencies across document lengths and is found by the equation  $\frac{count_t}{length_d}$  where  $count_t$  is the count of the occurrences of a term within a document and  $length_d$  is the total length of the document. Preprocessing of these three documents included removing the punctuation. (We did not remove the “stop words” because of the length of the examples.)

Document 1: After the first day I felt a spring in my step.

Document 2: The first day of Spring was a beautiful day.

Document 3: When she first comes in, spring up and shout, “Surprise!”

Table 4.1 *The term document matrix generated from three example sentences.*

		Document 1	Document 2	Document 3
T1	After	0.09	0	0
T2	The	0.09	0.11	0
T3	First	0.09	0.11	0.1
T4	Day	0.09	0.22	0
T5	I	0.09	0	0
T6	Felt	0.09	0	0
T7	A	0.09	0.11	0
T8	Spring	0.09	0.11	0.1
T9	In	0.09	0	0.1
T10	My	0.09	0	0
T11	Step	0.09	0	0
T12	Of	0	0.11	0
T13	Was	0	0.11	0
T14	beautiful	0	0.11	0
T15	When	0	0	0.1
T16	She	0	0	0.1
T17	Comes	0	0	0.1
T18	Up	0	0	0.1
T19	And	0	0	0.1
T20	Shout	0	0	0.1
T21	surprise	0	0	0.1
		0.99	0.99	1

Words that are very common, known as “stop words,” are usually excluded from the matrix. The value in each cell of the matrix is the scaled frequency of the term in the document. To reduce the cost of query comparisons, the singular value decomposition is truncated after an *a priori* limited number of matrix entries (300 in the current work). From this truncated matrix, we extract the most significant 300 orthogonal factors. The original matrix can be approximated by linear combination of these factors. [41].

After this reduction, a vector of factor weights represents each document. The number of items in each vector equals the number of factors into which the original matrix was decomposed (300 in the above–mentioned discussion). Queries are represented in a manner similar to documents. Query vectors are built from the scaled combination of the terms within the query.

Table 4.2 *Some examples of common “stop words” in English.*

the	Or	with	At	that
be	As	by	For	and
from	under	such	There	of
other	whether	also	Than	which
now	where	these	When	we
an	To	but	Upon	then
If	Is	it	Can	this

Another example from Ecclesiastes 3:1–8 (Holy Bible New International Version, 1988, p. 592):

There is a time for everything, and a season for every activity under heaven:

a time to be born and a time to die,  
 a time to plant and a time to uproot,  
 a time to kill and a time to heal,  
 a time to tear down and a time to build,  
 a time to weep and a time to laugh,  
 a time to mourn and a time to dance,  
 a time to scatter stones and a time to gather them,  
 a time to embrace and a time to refrain,  
 a time to search and a time to give up,  
 a time to keep and a time to throw away,  
 a time to tear and a time to mend,  
 a time to be silent and a time to speak,  
 a time to love and a time to hate,  
 a time for war and a time for peace.

Table 4.3 *Tabulated word frequencies.*

30	A
29	Time
26	To
15	And
4	For
2	be, tear
1	season, every, activity, under, heaven, there, everything, born, die, plant, uproot, kill, heal, is, down, build, weep, laugh, mourn, dance, scatter, stones, gather, them, embrace, refrain, search, give, up, keep, throw, away, mend, silent, speak, love, hate, war, peace



After the word frequency count (Table 4.3) is performed and placed into a vector, all documents are aggregated into a matrix. The matrix is decomposed into its singular values. The result is a grouping of frequently used words from each document. For example, in the text above, the three words, “a,” “time,” and “for,” might form one group {a, time, for) and would match other documents that repeated that phrase frequently. Suppose you had three documents that have been analyzed as above and their matrix values reduced to three groups of words with the following weights:

D1: [.2 .6 .2]

D2: [.1 0 .9]

D3: [.3 .3 .4]

and your query produced the following match on those same three groups of words,

Q: [.1 .1 .8].

then the relative rankings of the documents against the query would produce

D1: .24

D2: .73

D3: .38

so that the system would return the rankings: D2, D3, D1. We would then consider D2 as the most likely match or context for our query.

### **4.3 Singular Value Decomposition**

LSA begins with a large, sparse, term-by-document matrix containing scaled word frequencies computed from a document library. (See Table 4.1 for a brief example.) SVD is performed on this matrix to generate a set of orthogonal factors from which the original matrix can be approximated by linear combination.

Each of these orthogonal factors represents a dimension of similarity among the terms and documents of the original library. Thus, it is possible to construct a vector of factor weights to approximately represent a document from the library. Furthermore, a vector may be constructed to represent any subset of the terms that were included in the original library. In constructing the vector, the factors are scaled appropriately to accurately represent the terms of the document or query.

#### **4.4 Building a query**

The ideal query is a naturally occurring text that is similar to the data set but contains a known semantic content. Since we do not read, write, or speak Arabic we resorted to a list of terms translated from English into Arabic by Google Translate to formulate our queries. Using this technique we constructed queries related to terrorism, religion, and soccer. We chose terrorism as an obvious topic of interest, religion as a politically, culturally, and socially important topic, and soccer as a more neutral control. The query documents for each of these topics are in Appendix B.

To find documents similar to a query, a vector is constructed from the terms of the query and compared to the vectors of all documents in the library. The documents whose vectors most closely match the query vector are returned as possible matches. Note that the match is not based on the query and documents containing the same terms but rather on the query and documents containing similar weighting factors. We consider each principal component vector as an abstraction of a concept or idea that is represented in the original document library. Hence, a document's vector represents a combination of one or more of these concepts (i.e., a semantic entity). The similarity of the vectors is evaluated by cosine similarity. LSA provides a method to order the documents in a

corpus without direct text matching by comparing the similarity of the semantic ideas represented by the documents to the semantic content of a query document.

## CHAPTER V - NEURAL NETWORK TRAINING AND TESTING

After using LSA to score our data set on similarity to a query document, we trained a deep neural network to classify the documents into one of two classes of “positive” (meaning semantically related to terrorism) or “negative” (meaning unrelated to terrorism). Our inputs consisted of documents from the previously constructed data set. We selected documents to classify as “positive” or “negative” based on a calculated threshold separating the complete data set into two subsets. We used Otsu’s method to calculate the threshold. We performed a series of experiments varying the number of features from 30 to 500, varying the number of training epochs from 200 to 4,000, and conducted trials with and without a step-based learning rate decay algorithm.

### 5.1 Finding a threshold

The first step in our classification process was to divide the data set into “positive” and “negative” subsets. In this context “positive” means more closely related to the chosen semantic category, in our case, terrorism. “Negative” means less closely related. For our early experiments we simply chose the top-ranked 1,000 documents to be the positive set and chose 1,000 documents with ranks from 3,000 to 4,000 to be the negative set. That simple approach has two advantages. It creates two sets that are more likely to be of similar sizes and it guarantees that there will be some distance between the two sets.

Our next approach was to divide the data set using a threshold,  $T$ . All documents with rankings above the threshold were labeled as positive and all documents with rankings below the threshold were labeled as negative.

*If  $f(x, y) > T$  then  $label(x, y) = "negative"$  else  $label(x, y) = "positive"$*

In an effort to make our threshold selection more rigorous and supported by research we implemented Otsu's thresholding [42]. Otsu's thresholding method was published in 1979 and was developed as a technique to binarize a gray scale image into foreground and background components. To implement Otsu's method, we first created a histogram of our values. Since Otsu's method was developed for gray scale image processing, the most common number of bins for the histogram is 256, which is what we used as well. Otsu's method divides the data set in two such that the distance between the means of the two subsets is maximized.

We arbitrarily choose a number of bins,  $L$ , and create a normalized histogram of our cosine similarities (i.e., for each bin  $i$ ,  $P(i)$  is the normalized frequency of the set of cosine similarities in  $i$ ). Assuming that we have set the threshold at  $T$ , the normalized fraction of documents that will be labeled as negative will be:

$$q_{neg}(T) = \sum_{i=1}^{bin(T)} P(i)$$

The normalized fraction of documents that will be labeled as positive will be:

$$q_{pos}(T) = \sum_{i=bin(T)+1}^L P(i)$$

Since we have constructed a probability distribution

$$q_{neg}(T) + q_{pos}(T) = 1$$

The mean cosine similarity of the negative subset will be:

$$\mu_{neg} = \frac{1}{q_{neg}(T)} \sum_{i=1}^{bin(T)} c(i) * P(i)$$

The mean cosine similarity of the positive subset will be:

$$\mu_{pos} = \frac{1}{q_{pos}(T)} \sum_{i=bin(T)+1}^L c(i) * P(i)$$

The variance of the negative subset will be:

$$\sigma_{neg}^2(T) = \frac{1}{q_{neg}(T)} \sum_{i=1}^{bin(T)} (c(i) - \mu_{neg})^2 P(i)$$

The variance of the positive subset will be:

$$\sigma_{pos}^2(T) = \frac{1}{q_{pos}(T)} \sum_{i=bin(T)+1}^L (c(i) - \mu_{pos})^2 P(i)$$

The variance of the entire data set will be:

$$\sigma^2 = \sum_{i=1}^L (c(i) - \mu)^2 P(i)$$

The variance can be written:

$$\begin{aligned} \sigma^2 &= q_{neg}(T)\sigma_{neg}^2(T) + q_{pos}(T)\sigma_{pos}^2(T) + q_{neg}(T)(\mu_{neg}(T) - \mu)^2 + q_{pos}(T)(\mu_{pos}(T) - \mu)^2 \\ &= \sigma_W^2(T) + \sigma_B^2(T) \end{aligned}$$

Where  $\sigma_W^2(T)$  is defined to be the within-class variance and  $\sigma_B^2(T)$  is defined to be the between-class variance. Since the total variance does not depend on  $T$ , the  $T$  minimizing  $\sigma_W^2(T)$  will be the  $T$  maximizing  $\sigma_B^2(T)$ .

We can re-write  $\sigma_B^2(T)$  as follows:

$$\sigma_B^2(T) = \frac{[\mu(T) - \mu q_B(T)]^2}{q_B(T)q_{pos}(T)}$$

Where

$$\mu(T) = \sum_{i=1}^{bin(T)} c(i) * P(i)$$

To accomplish this, we iterate through the histogram calculating the variance of each subset assuming the set is divided at the average of the values in the current bin of the histogram. We continue to iterate over the bins of the histogram selecting as our threshold the average of the values in the bin that maximizes the between-class variance and minimizes the within-class variance. For our experiments we used the MatLab implementation of Otsu's method to select a threshold. Figure 5.1 shows the histogram of cosine similarities generated by the LSA process described above and the selected threshold.

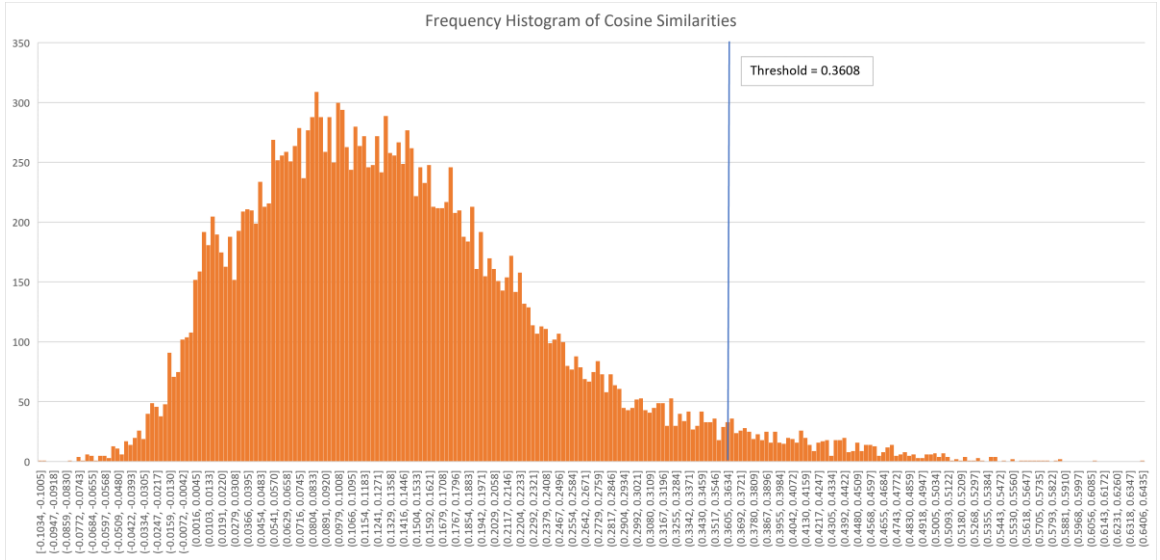


Figure 5.1 Histogram of cosine similarities to terrorism query document

Figure 5.1 shows the 256-bin histogram of the cosine similarities of the Twitter documents to the terrorism query document. We used the MatLab implementation of Otsu's method to calculate the displayed threshold of 0.3608.

Otsu's method is well-accepted and widely used in image processing to establish a threshold for binarizing grayscale images. However, analysis has shown that it works best when the histogram of the image is a bimodal distribution [43]. Our distribution is monomodal with a long right tail. Given this distribution a concern is whether Otsu's method finds the optimal threshold for this purpose. We noted that this threshold divides the data set into segments of 96.22% negative and 3.78% positive. These values seem intuitively correct as it is likely that only a small minority of tweets express ideas related to terrorism. While the threshold may be intuitively correct, the division of the data set in this skewed manner is not optimal for processing. Multiple trials with the full data set did not yield satisfactory results. To balance the sizes of the two document classes more closely we limited the size of the negative class to the same number of documents contained in the positive class. We chose the set of documents with the lowest similarity scores for the negative class. This provided multiple benefits. First, it more closely balanced the sizes of the positive and negative classes. It should be noted that even after this balancing, the positive class (108,061) was still substantially larger than the negative class (22,621). We speculate that this was due to greater frequency and length of tweets in the positive class. The second benefit was decreasing the size of the training and test sets. The smaller size made processing much faster. The third benefit of splitting the classes this was increasing the semantic distance between the two classes. The accuracy of the classifier improved with the use of the more diverse data set.

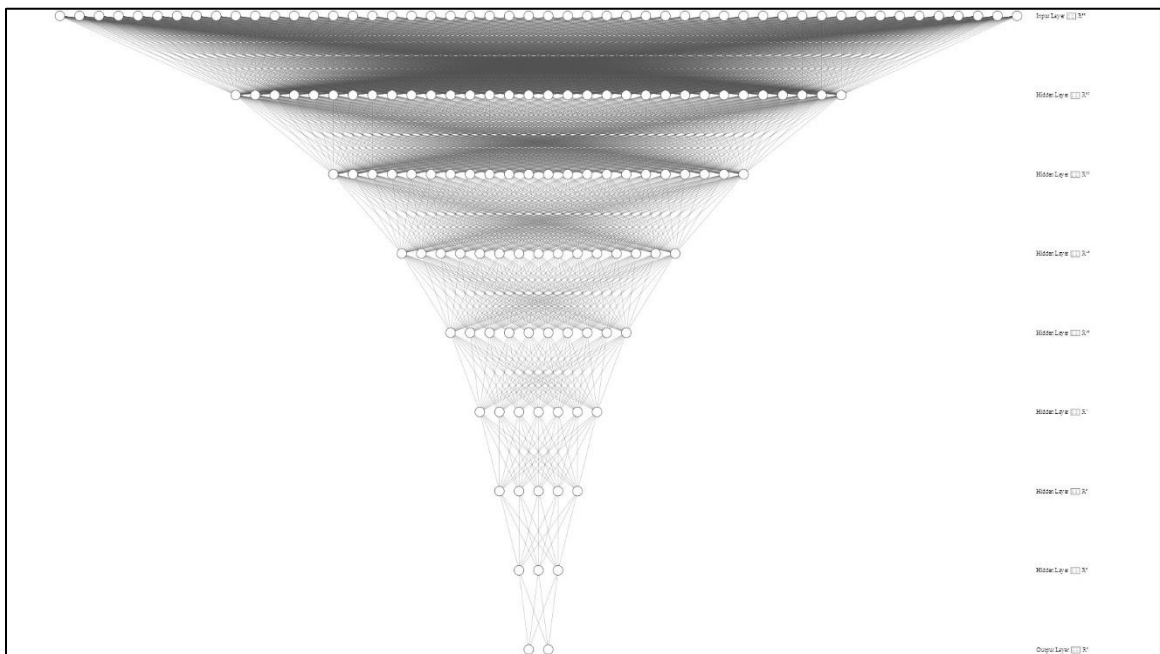
## **5.2 Neural network architecture and training trials**

Artificial neural networks have been used for a wide variety of classification problems across many domains of knowledge. The general structure of a deep learning



neural network is as a series of layers where each layer is made up of a set of nodes. The first layer is the input layer and receives the input. The last layer is the output layer which provides the final result. The layers in between are called hidden layers. Each node in a hidden layer holds a value or set of values determined by the value of the nodes in the previous layer of the network modified by a weight and bias that relates the two nodes. In the case of the input nodes, which have no previous node, the values are simply the input values. All nodes in a hidden or output layer receive input from all nodes in the previous layer, calculate their value by summing the weighted inputs and applying the bias, apply the activation function, and pass that value on to the nodes in the next layer.

Figure 5.2 *Diagram of our neural network architecture*



Deep learning neural network architecture with 50 input nodes, seven hidden layers, and two output nodes.

Training a neural network consists of evaluating the error after each forward pass through the network and then propagating the error back through the network and

adjusting the parameters of each node to decrease the error on subsequent passes. Each cycle of forward and backward passes is called an epoch. After a feed forward pass through the network the predicted value of each element in the training data set is compared to the known “true” value. The difference between the predicted values and the known “true” values are the error or loss. Various formulae have been used to calculate the loss. We used Mean Squared Error (MSE) to calculate the loss. We calculate the MSE as follows:

$$L(y, \hat{y}) = \frac{1}{N} \sum_{i=0}^N (y - \hat{y})^2$$

Where  $y$  is the true value and  $\hat{y}$  is the predicted value.

The loss values are then used to adjust the weights at each layer back through the network. This process is known as back-propagation and is one part of what makes the network a learning network. Back-propagation works by calculating the gradient of the loss function with respect to the parameters of the network [44]. A separate algorithm is used to adjust the values of the parameters to decrease the loss. We used the stochastic gradient descent algorithm.

Many different functions have been used to configure artificial neural networks. We used the linear function between the nodes of each hidden layer and the ReLu function as the activation function for each node. The ReLu function is simply  $\max(z, 0)$  or:

$$R(z) = \begin{cases} z & z > 0 \\ 0 & z \leq 0 \end{cases}$$

We varied the number of hidden layers in different trials depending on the number of features being used. Our general rule was to approximate a 2/3 reduction in the number of

nodes at each successive layer until we reached two output nodes. Table 5.1 shows the number and size of hidden layers that we used for various numbers of input features.

We constructed our network as a binary classification network, so our final output layer consisted of two nodes, one for positive and one for negative. We note that a different approach to accomplishing the same goal is to use a structure that delivers a continuous value to a single output node rather than a discrete value to two nodes. Such a value can then be evaluated to determine the class of the item. We reviewed such an approach but chose the two-node output model.

Table 5.1 *Number and size of hidden layers*

Number of features	Size of each hidden layer
30	20, 12, 8, 5, 3
50	33, 22, 15, 10, 7, 5, 3
100	66, 40, 27, 18, 12, 8, 6, 4
250	150, 90, 54, 32, 19, 11, 6, 4
500	350, 250, 167, 115, 90, 60, 40, 27, 18, 12, 8, 6, 3
1,000	700, 500, 350, 125, 90, 60, 40, 13, 9, 6, 4
2,500	1000, 500, 243, 81, 27, 9
5,000	4000, 3000, 2000, 1000, 500, 243, 81, 27, 9

We recognize that not all of the entries in this table conform to the decrease to 0.66 rule mentioned in the text. As we progressed in training, we adjusted the ratio. Our final trials used the architectures with the 0.66 rule.

In our network the input values are constructed as vectors such that the number of elements in each input vector is equal to the number of terms selected from the data set with the following limitations. We first limited the terms to those with a document frequency of 85% or less. In other words, we excluded any term that occurs in more than 85% of documents. The rationale for this exclusion is the assumption that a term that occurs in more than 85% of documents does not contribute to discriminating among the

classes of documents. This is a mathematical approach that simulates the removal of stop words as described in Chapter IV. The 85% value is a configurable parameter. We also experimented with a 70% document frequency but did not see any change in accuracy. We then calculate the TF-IDF score for each term. The input vectors are the TF-IDF scores of each term.

One of the parameters that we adjusted on various trials of our network was the number of features selected. Feature selection consisted of selecting the terms with the highest TF-IDF scores. On various trials we used a minimum of 30 features up to a maximum of 5,000 features. The greatest accuracy was obtained with 50 features. Using a greater number of features did not improve the accuracy of the classification.

Another parameter that we experimented with was the number of training epochs. We tried epoch numbers from as low as 40 to as high as 10,000. The execution time required for high numbers of epochs combined with high numbers of features limited trials of those combinations. One trial of 5,000 features and 4,000 epochs took approximately six hours to complete execution. We did not systematically try all combinations of parameters.

In addition to experimenting with various combinations of numbers of features and numbers of epochs, after trying several configurations we implemented a learning rate decay factor to control the speed at which the network adjusted the weights between the nodes of the hidden layers as training progressed. Pytorch offers several different approaches for learning rate decay. We implemented a step-based decay algorithm that takes a `step_size` and a decay rate, `gamma`, as its parameters. The Pytorch implementation

decays the learning rate of each parameter group by gamma every step\_size epochs. We set our initial learning rate at 0.0001. After experimenting with various learning rate decay and step\_size values, we found that with a step\_size determined by  $\frac{\text{number of epochs}}{40}$  a gamma of 0.85 consistently produced a desirable learning curve. See Figure 5.3 for examples of learning curves before implementing learning rate decay and Figure 5.4 for examples of learning curves after implementing learning rate decay.

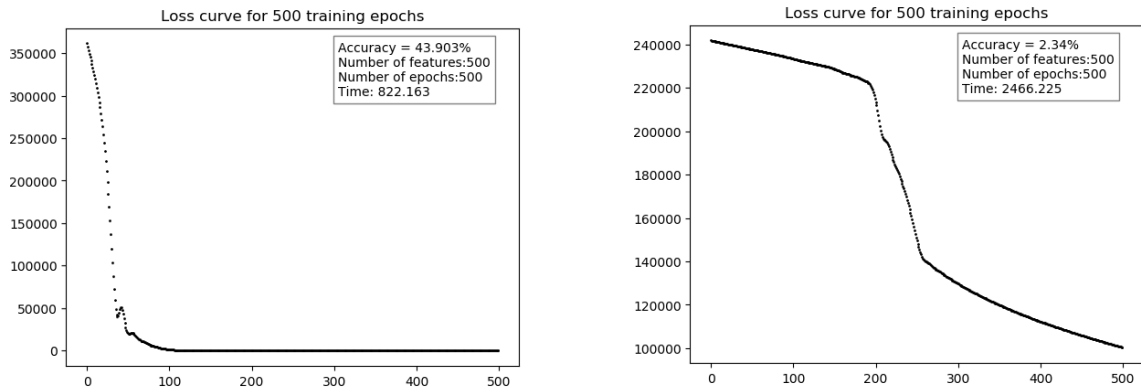


Figure 5.3 Loss curves without learning rate decay

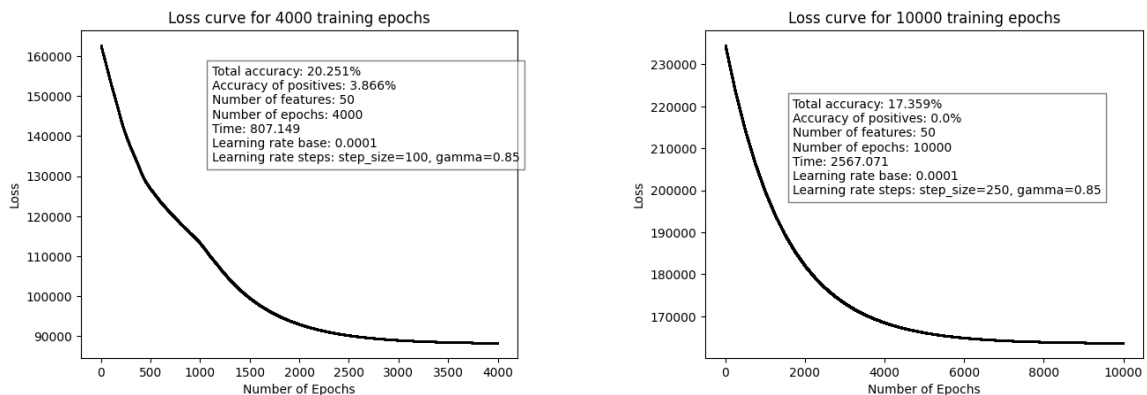


Figure 5.4 Loss curves with learning rate decay

We calculated two different measures of accuracy: total accuracy and positive accuracy. Total accuracy or overall accuracy is the number of correctly classified items as a percentage of the total number of items classified.

$$accuracy_{total} = \left( \frac{\hat{y}}{y} \right) * 100$$

where  $y$  is the total number of items and  $\hat{y}$  is the number of items classified correctly.

Positive accuracy is the number of items predicted to be positive divided by the number of actual positive items.

$$accuracy_{total} = \left( \frac{\hat{p}}{p} \right) * 100$$

where  $p$  is the number of positive items and  $\hat{p}$  is the number of items classified as positive. To implement our neural network, we used the Pytorch library [45] running on top of Numpy [46] and SciPy [47] in a Python implementation on Visual Studio Community 2019.

## CHAPTER VI – RESULTS

The results of the work covered in this dissertation take three forms. First is the data set itself. Second are the results of the LSA classification. Third are the results of the neural network classification. We will cover each result in the sections below.

### **6.1 Multi-dialectal Arabic social media data set**

We developed a script using Python and the Tweepy library to collect a data set of over 1.2 million tweets ranging in date from April of 2009 to February of 2021. The tweets are organized into documents by country, then by user, then by month. We define a document as the collection of tweets issued by one user during one month. At the time of this writing we had 21,791 documents in our database with a total file size of approximately 214 megabytes.

We selected Twitter accounts from 12 Arabic-speaking countries (Bahrain, Egypt, Iraq, Kuwait, Lebanon, Libya, Oman, Qatar, Saudi Arabia, Syria, UAE, and Yemen). Our goal in selecting accounts was to identify those accounts that are most likely to influence and reflect socio-political opinions within their country of origin and throughout the region. As such, our primary criterion for selecting accounts was number of followers. All of the accounts in our database have a minimum of 10,000 followers. Some accounts have followers numbering in the millions.

We did not attempt to limit our collection to any specific language. Indeed, a significant portion of the data set contains tweets either completely or partially in English. We also did not attempt dialect identification on the data set, so do not know what percentage of the tweets are in dialectal Arabic vs. MSA or which dialects, if any, are present, missing, or more prominent. The partial review of our documents by a native

Arabic speaker revealed that many documents combine MSA and dialectal Arabic. Given the informal nature of social media posts in general, this is not surprising and supports our approach of not screening or excessively cleaning our data set to ensure that it contains only one dialect or language. Our review of the literature revealed that there are very few resources available for NLP research on multi-dialectal Arabic. The resources that are available are primarily limited to one or two dialects and are generally small by modern corpora standards. Our data set is a significant contribution to the NLP resources for research in multi-dialectal Arabic and analysis of language, opinions, and social media trends in the region.

## 6.2 LSA scoring of documents

We used the Gensim library [38] and Python to execute LSA on our library of 21,791 documents. Cosine similarities to our query document ranged from a maximum of 0.6435 to a minimum of -0.1034. Applying Otsu’s method [42] to a 256-bin histogram of the cosine similarities yielded a threshold of 0.3608. This threshold divides the data set into a “positive” set of 788 documents (3.62%) and a “negative” set of 21,003 documents (96.38%). In tables 6.1, 6.2, and 6.3 we list the top ten documents ranked by cosine similarity to the terrorism, religion, and soccer queries, respectively.

Table 6.1 *Top results of terrorism query*

Document ID	Cosine Similarity	Twitter Handle	Date
9337	0.643487	@Dr_alwasmi	August 2011
7984	0.60886	@ khameskhanjar	October 2017



Table 6.1 continued

9411	0.587528	@Dr_alwasmi	December 2017
9335	0.5852213	@Dr_alwasmi	June 2011
9336	0.58505744	@Dr_alwasmi	July 2011
9272	0.5764522	@DrHAKEM	January 2017
9333	0.57361585	@Dr_alwasmi	March 2011
9269	0.57274026	@DrHAKEM	September 2016
7933	0.57023716	@khameskhanjar	April 2013
7930	0.5649872	@khameskhanjar	January 2013

Table 6.2 *Top results of religion query*

Document ID	Cosine Similarity	Twitter Handle	Date
1098	0.692725	@amrdiab	August 2011
9135	0.65976	@KingSalman	June 2017
2356	0.659475	@Hamaki	June 2017
1734	0.652885	@DrAliGomaa	October 2012
1732	0.648555	@DrAliGomaa	August 2012
1156	0.648043	@amrdiab	July 2016
1148	0.643648	@amrdiab	November 2015

Table 6.2 continued

1167	0.638005	@amrdiab	June 2017
8713	0.635798	@Dr_amosleh	September 2015
8734	0.631563	@Dr_amosleh	June 2017

Table 6.3 *Top results of soccer query*

<b>Document ID</b>	<b>Cosine Similarity</b>	<b>Twitter Handle</b>	<b>Date</b>
7490	0.73908	@aymanjada	May 2017
7464	0.669448	@aymanjada	March 2015
9582	0.665062	@SamiAlJaber	January 2017
7483	0.662481	@aymanjada	October 2016
7455	0.65716	@aymanjada	May 2014
8854	0.654736	@faisalbinturki1	May 2012
7493	0.652406	@aymanjada	August 2017
3488	0.633253	@trikaofficial	November 2015
7481	0.630216	@aymanjada	August 2016
7465	0.629053	@aymanjada	April 2015

One of the challenges of automated processing of social media posts in dialectal Arabic is the lack of authoritative corpora against which to compare results. Because we do not have access to such a corpus, we cannot compare our LSA scoring results to an established “ground truth”. As a partial validation of our approach, we had a native Arabic speaker read through the top ten documents ranked in descending order by cosine similarity to our terrorism query document. He scored nine of the ten documents as relevant to terrorism and noted that three of the top ten may be interpreted as a “call to action” while a fourth “openly advocates a revolution in Iraq”. This verification serves as a validation of our technique. Appendix C contains the original Arabic text as well as an English translation from Google Translate of part of a document that our algorithm ranked as semantically similar to terrorism. This document is the collection of tweets by the account DrHAKEM from October of 2016.

Figure 6.1 shows a scatter plot of the cosine similarities of the document vectors in our data set to the query document we constructed with terms related to terrorism. Table 6.4 shows descriptive statistics for the cosine similarities of our data set to our terrorism query document.

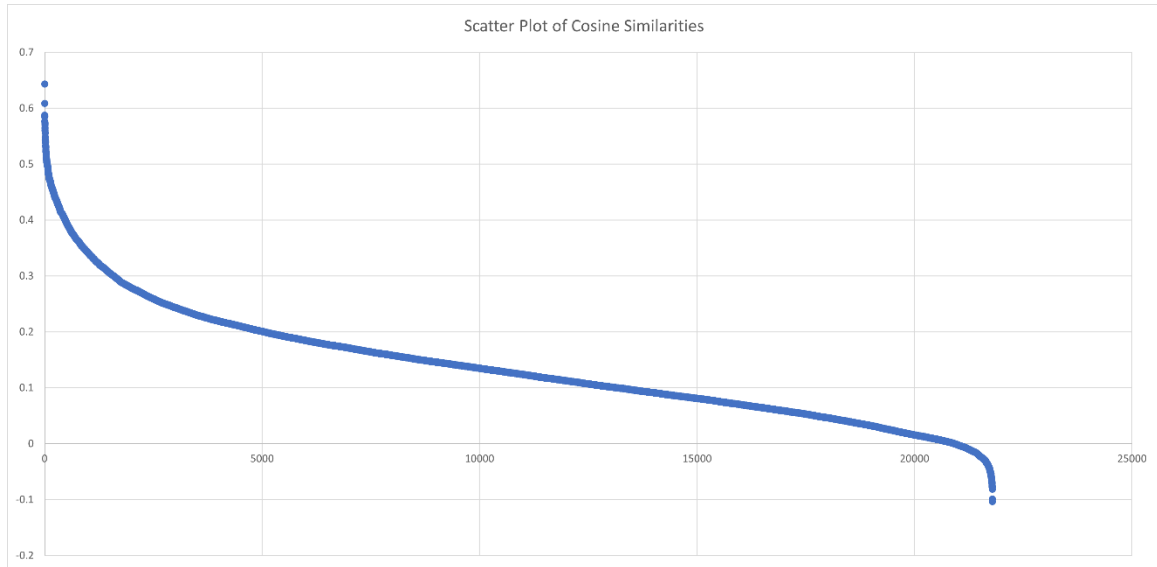


Figure 6.1 Scatter plot of cosine similarities of twitter documents to query document

Table 6.4 *Descriptive statistics for cosine similarities of documents to terror query*

Count	21,791
Maximum	0.6434865
Minimum	-0.10344128
Average	0.138874628
Range	0.74692778

### 6.3 Neural network classification of documents

The most accurate classification that we achieved with our neural network was 88.178% total accuracy and 98.075% accuracy in classifying documents that were semantically related to terrorism (positive accuracy). In Figure 6.2 we present the training curve, some of the parameters, and the results of that model. The parameters for that neural network model were 50 features run for 10,000 epochs with a base learning rate of 0.0001, a learning decay gamma of 0.85, and a step\_size of 250 epochs.

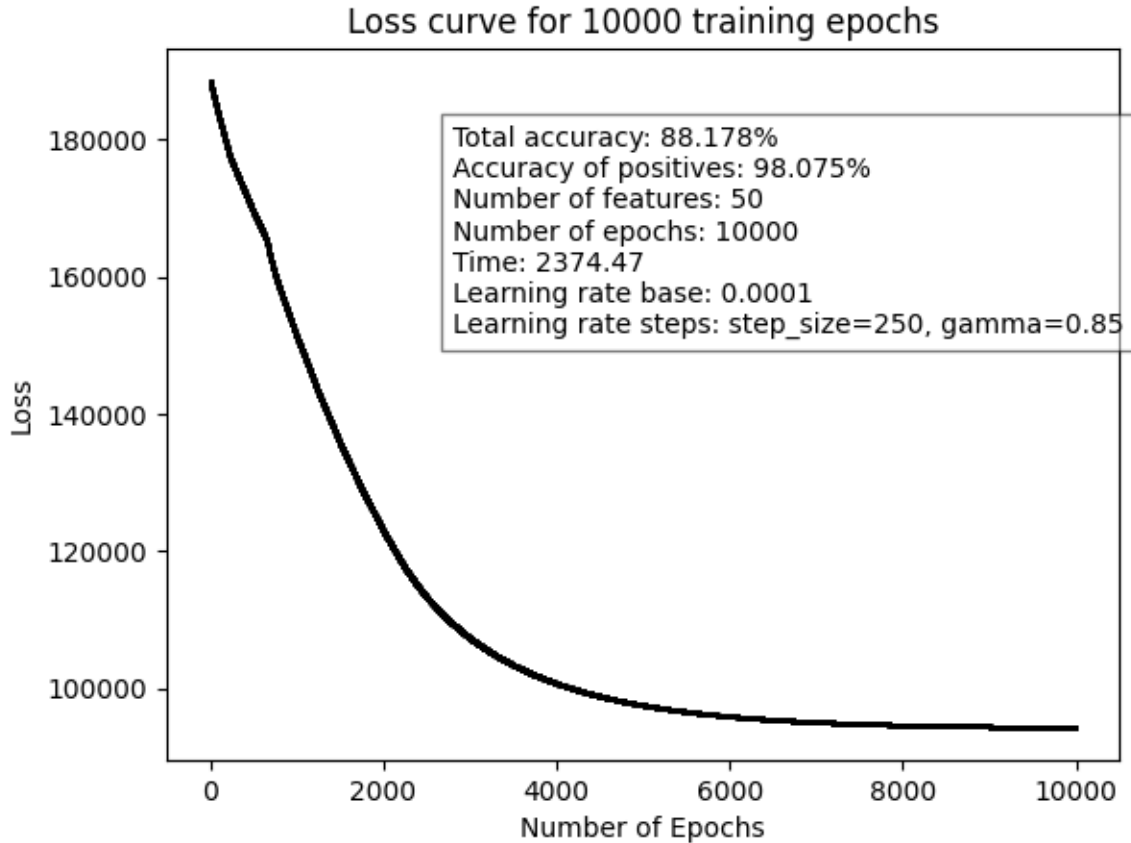


Figure 6.2 *Loss curve for our most accurate neural network model*

In 2020, Elnagar, Al-Debsi, and Einea [27] achieved 96.94% accuracy for single-label categorization with attention-GRU and 88.68% for multi-label categorization also using attention-GRU on a data set of MSA harvested from news websites. In a similar study, Boukil et al. [1] got maximum accuracies of 86.3%, 88.2%, and 92.9% using a logistic regression classifier, support vector machine classifier, and convolutional neural network classifier respectively. Boukil et al. were also working only with MSA. In 2013 Khorsheed and Al-Thubaity [16] found that an SVM gave them the best accuracy of 97% for the Islamic Topics data set, although they also reported a 61% accuracy for the

Arabic Poems data set. Our best accuracy of 88% overall and 98% positive accuracy are very competitive with these results.

## CHAPTER VII – CONCLUSION, DISCUSSION, AND FUTURE WORK

### 7.1 Discussion

To the best of our knowledge, this is the first study of semantic classification of multi-dialectal Arabic social media posts using a neural network enhanced by LSA. This approach has several advantages. It provides a mechanism for semi-automated labeling of large corpora. It incorporates semantic content without requiring the use of either string matching, stemming, or the use of word nets, dictionaries, ontologies, or similar resources. It is language independent. Any language that is computer-encoded can be processed with these techniques including emoji, emoticons, and hashtags. The only requirement of the language is that it be used with great enough consistency across the corpus to encode semantic meaning. Indeed, the current study includes documents in at least two dialects of Arabic as well as English. Furthermore, languages can be mixed within the corpus, as long as there is a large enough sample of each language to capture the semantic concepts. The corpus and the semantic vector space can be easily updated with new documents and the query re-run to generate new findings as new data becomes available. When working with social media text it is important to be able to update the corpus easily as terminology and trends change quickly on social media platforms. Our approach does not require extensive pre-processing of text to clean the data. In fact, we argue that substantial cleaning of social media posts reduces the semantic richness of the text. Furthermore, some research shows that pre-processing such as replacing URLs and mentions, removing English, and normalizing Arabic (i.e. stemming, repeated character replacement, standardizing spellings) did not significantly improve the accuracy of

classification [22]. Other studies found that stemming did increase the accuracy of the classification.

Our study also had several weaknesses. LSA provides “weak” semantic labels as compared to hand-tagged data or data that uses word nets, ontologies, dictionaries, or other similar resources. In the absence of such resources for multi-dialectal Arabic we use LSA to provide labels. Social media data is very noisy. Non-standard language and misspellings (intentional and unintentional) are common. Punctuation is often used to convey meaning apart from its normal usage (e.g. emoticons, /s to indicate sarcasm). Hashtags are prevalent. Numbers are often used to indicate sentiment (1/5, 11/10). Emojis are widely used and often convey subtle messages that vary by age, language, and culture of the user. The simplest means of dealing with all of this noise is by simply removing it, however, the semantic content of the text is decreased and there is some evidence that the removal does not increase accuracy of classification. Training a neural network is computationally expensive and if a classifier is going to remain current with social media trends it will have to be updated repeatedly. Other approaches may offer similar performance for less computational cost.

## **7.2 Future work**

Future work on this project may include seeking out collaborators with expertise in Arabic, specifically dialectal Arabic. Greater collaboration with dialectal Arabic speakers would provide benefits including creating or curating a more effective query document and more complete validation of results. Evaluation of selected samples from the LSA labeling as well as from the neural network classification would improve the confidence of the results.



Using other classifiers (e.g. SVM, NB) to categorize the documents in our corpus and comparing the results to the current work would provide further data as to the value of neural network classification. Extending the neural network research into other forms of networks would also be of interest. Another refinement that may be undertaken in future work is removal of URLs as part of preprocessing. When URLs are processed as words in the corpus they contribute little semantic value but add noise. Since URLs are typically unique they should receive a very low TF-IDF score and thus be excluded from feature vectors, but removing them entirely is a simple pre-processing step that may improve processing time as well as accuracy.

Since our approach is independent of language it may be effective for other languages and dialects as well. For example, it may be useful for semantic classification of dialects of Spanish, French, and/or Telugu. Specific properties of these languages and currently available resources for processing them would need to be explored. Another final direction for future work is using semantic vectors from LSA directly as features in a neural network classifier. Direct classification of semantic vectors may improve speed and accuracy.

### **7.3 Conclusion**

We constructed a large data set of multi-dialectal Arabic social media posts harvested from 449 Twitter accounts located in 12 Arabic-speaking countries. We organized the tweets into documents and performed LSA on the corpus to reduce the dimensionality as well as to identify the principal semantic concepts addressed by the corpus. We compared the semantic document vectors to the vector of a constructed query document containing terms related to terrorism and ranked the corpus document via

cosine similarity to the query vector. We then divided that ranked list by automatic thresholding to separate a relatively small “positive” subset from the remainder of the corpus. To create a “negative” subset we selected an equal number of documents from the bottom of the similarity rankings. We divided these two labeled sets of documents into training and test sets and used the training set to train a backpropagation neural network to classify the test documents into positive and negative classes. After experimenting with multiple configurations, we were able to achieve a maximum overall accuracy of 88% with a positive accuracy of 98%. These results are comparable to results obtained by other researchers using a variety of text classification approaches. To the best of our knowledge this is the only work providing semantic classification of multi-dialectal Arabic social media. The data set and source code are provided on GitHub at <https://github.com/therishel/ArabLeader>.

## APPENDIX A - Arabic Stop Words

Table A.1 shows the list of Arabic stop words used in this work. This is a standard list of MSA words.

Table A.1 *List of Arabic stop words*

أى؟	تسعمائة	رابع	ف	مكانكم	،
الا	تسعون	راح	فإن	مكانكما	ء
التي	تسعين	رجع	فاء	مكانكن	ء
ابتدا	نشرين	رزق	فان	مكانك	آ
ابين	تعا	رويدك	فانه	مليار	آب
اتخذ	تعلم	ريال	فيراير	مليم	أذار
اثر	تفعلان	ريث	فرادى	مليون	أض
اثنا	تفعلون	رُبَّ	فضلا	مما	آل
اثنان	تفعلين	ز	فقد	من	آمين
اثنى	تكون	زاي	فقط	منذ	آناء
اثنين	تلقاء	زعم	فكان	منه	أنفا
اجل	تلك	زود	فلان	منها	آه
احد	تم	زيارة	فلس	مه	آها
اخرى	تموز	س	فهو	مهما	آه
اخلوق	تينك	ساء	فو	ميم	آه
اذا	تئين	سابع	فوق	ن	أ
اربعة	ته	سادس	فى	نا	أبدا
اربعون	تي	سبت	في	نبأ	أبريل
اربعين	ث	سبتمبر	فيفري	نحن	أبو
ارتد	ثاء	سبحان	فيه	نحو	أب
استحال	ثالث	سبع	فيها	نعم	أجل
اصبح	ثامن	سبعة	ق	نفس	أجمع
اضحى	ثان	سبعمة	قاطبة	نفسه	أحد
اطار	ثاني	سبعمة	قاف	نهاية	أخبر
اعادة	ثلاث	سبعون	قال	نوفمبر	أخذ
اعلنت	ثلاثاء	سبعين	قام	نون	أخو

Table A1 continued

اف	ثلاثة	ست	قبل	نيسان	أخ
اكثر	ثلاثمئة	سنة	قد	نيف	أربع
اكد	ثلاثمائة	سكون	قرش	نخ	أربعاء
الان	ثلاثون	ستمئة	قط	ن	أربعة
الالاء	ثلاثين	ستمائة	قلما	هـ	أربعمئة
الألى	ثم	ستون	قوة	هؤلاء	أربعمائة
الا	ثمان	ستين	ك	ها	أرى
الاخيرة	ثمانمئة	سحقا	كان	هاء	أسكن
الان	ثمانون	سرا	كان	هاك	أصبح
الاول	ثماني	سرعان	كأي	هـب	أصلا
الاولى	ثمانية	سقى	كأين	هذا	أضحى
التى	ثمانين	سمعا	كاد	هذه	أطعم
التي	ثمانمئة	سنة	كاف	هل	أعطى
الثاني	ثم	سنتين	كان	هالة	أعلم
الثانية	ثم	سنوات	كانت	هلم	أغسطس
الحالي	ثمة	سوف	كانون	هلا	أفريل
الذاتي	ج	سوى	كثيرا	هم	أفعل به
الذى	جانفي	سين	كذا	هما	أف
الذي	جدا	ش	كذلك	همزة	أقبل
الذين	جعل	شباط	كرب	هن	أكتوبر
السابق	جل	شبه	كسا	هنا	أل
الف	جمعة	شتان	كل	هناك	ألا
اللاتي	جميع	شخصا	كلتا	هنالك	ألف
اللتان	جنيه	شرع	كلم	هو	ألفى
اللتيا	جوان	شمال	كلا	هي	أم
اللتين	جويلية	شيكل	كلما	هيا	أما
اللذان	جير	شين	كم	هيهات	أمام
الذين	جيم	شتان	كما	هيا	أمامك
اللواتي	ح	ص	كن	هؤلاء	أمامك
الماضي	حاء	صاد	كى	هاتان	أمد
المقبل	حادي	صار	كيت	هاتين	أمس
الوقت	حار	صباح	كيف	هاتيه	أمسى
الى	حاشا	صبر	كيفما	هاتي	أما
الي	حاليا	صبرا	كخ	هـج	أن
اليه	حاي	صدقا	ل	هذا	أنا
اليها	حبذا	صراحة	لأن	هذان	أنبا
اليوم	حبيب	صفر	لا	هذين	أننت
اما	حتى	صه	لا سيما	هذه	أنتم
امام	حجا	صه	لات	هذي	أنتما

Table A1 continued

امس	حدّث	ض	لازال	هَيَّهَات	أَنْتَن
امسى	حرى	ضاد	لاسيما	و	أَنْتَب
ان	حزيران	ضحوة	لام	و6	أَنْشَأْ
انبرى	حسب	ضد	لايزال	وأبو	أنه
انقلب	حقا	ضمن	لبيك	وأن	أَنْ
انه	حمدا	ط	لدى	وا	أَتَى
انها	حمو	طاء	لدى	واحد	أَهْلَا
او	حَمَّ	طاق	لدى	واضاف	أو
اول	حوالى	طالما	لذلك	واضافت	أوت
اي	حول	طرا	لعل	واكد	أوشك
ايار	حيث	طفق	لعلَّ	والتي	أول
ايام	حيثما	طُق	لعمري	والذي	أولئك
ايضا	حين	ظ	لقاء	وان	أولاء
ب	حيَّ	ظاء	لكن	واهاً	أولالك
بؤسا	حَذَار	ظل	لكنه	واو	أَوْه
بان	خ	ظَلَّ	لكنَّ	واوضح	أى
بئس	خاء	ظَنَّ	للامم	وبين	أي
باء	خاصة	ع	لم	وثي	أيا
بات	خال	عاد	لما	وجد	أيار
باسم	خامس	عاشر	لما	وراءك	أيضا
بان	خَبَّر	عام	لن	ورد	أيلول
بخ	خلا	عاما	له	وعلى	أين
بد	خلافاً	عامّة	لها	وفي	أيّ
بدلاً	خلال	عجبا	لهذا	وقال	أَيَّان
برس	خلف	عدا	لهم	وقالت	أَفِ
بسبب	خمس	عدة	لو	وقد	ؤ
بسّ	خمسة	عدد	لوكالة	وقف	إحدى
بشكل	خمسمئة	عدم	لولا	وكان	إذ
بضع	خمسمائة	عدّ	لوما	وكانت	إذا
بطآن	خمسون	عسى	ليت	ولا	إذاً
بعد	خمسين	عشر	ليرة	ولايزال	إنما
بعدا	خميس	عشرة	ليس	ولكن	إنن
بعض	د	عشرون	ليسب	ولم	إزاء
بغثة	دال	عشرين	م	وله	إلى
بل	درهم	عل	مئة	وليس	إلى
بلى	درى	علق	مئتان	ومع	إليكُم
بن	دواليك	علم	ما	ومن	إليكما
به	دولار	على	ما أفعله	وهب	إلَيْكَ
بها	دون	علي	ما انفك	وهذا	إليك
بهذا	دونك	عليك	ما برح	وهو	إلَيْكَ
بيد	ديسمبر	عليه	مائة	وهي	إِلَّا

Table A1 continued

بين	دينار	عليها	ما انفك	وي	إما
يس	ذ	عل	ما برح	وشكان	إن
بله	ذا	عن	مادم	ى	إن
ة	ذات	عند	ماذا	ي	إى
ت	ذاك	عندما	مارس	ياء	إياك
تاء	ذال	عنه	ما زال	يفعلان	إياكم
تارة	ذاتك	عنها	مافتئ	يفعلون	إياكما
تاسع	ذان	عوض	ماي	يكون	إياكن
تان	ذلك	عيانا	مايزال	يلي	إيانا
تانيك	ذهب	عين	مايو	يمكن	إياه
تبذل	ذو	عَدَسْ	متى	يمين	إياها
تجاه	ذيت	غ	مثل	ين	إياهم
تحت	ذبتك	غادر	مذ	يناير	إياهما
تحول	ذَّين	غالبا	مزة	يوان	إياهن
تخذ	ذه	غدا	مساء	يورو	إياي
ترك	ذي	غداة	مع	يوليو	إيه
تسع	ر	غير	معاذ	يوم	ئ
تسعة	راى	غين	معه	يونيو	ا
تسعمئة	راء	-	مقابل	آيان	ا؟

## APPENDIX B – Arabic query text

The following text was assembled by the author as lists of terms associated with each topic. The translation to Arabic was done by Google Translate. A query document provided by one or more native speakers of dialectal Arabic including multiple dialects as well as typical syntactic usage and structure for social media posts would very likely improve the accuracy of the semantic rankings used for this research.

Table B.1 *Arabic terrorism query terms*

إرهاب	عنف	إنتهاك	مواد خطرة	منظمة التحرير الفلسطينية
الإرهاب	إنتحار	خرق	الخطرة	حركة طالبان
إرهابي	إستشهاد	تهديد	منشآت حيوية	الشباب
الإرهابي	شهيد	خطر	حرب	الجهاد
قام بعملية	منفذ عملية إنتحارية	طواري	حروب	هدف
هجوم	الأمن	حالة طوارئ	محارب	يستهدف
هجمات	أمن	مواجهة	معركة	إنتحار
ضربات	إطلاق نار	المواجهة	معارك	عملية إنتحارية
فجر	قنص	فحص	إعدام	تطرف
تفجير	قناص	فرز	قتل	يتطرف
دمر	مطلق نار	غلق	خطف	متطرف
تدمير	مصاب بطلق ناري	إغلاق	إختطاف	قومي
إغتيال	مقتل	تحطم	القاعدة	متطرف
قتل	موت دموي	إصطدام	داعش	متعصب
منفذ عملية	قتلى	حادث	دولة الخلافة	أصولية
إغتيال	رهينة	حادثة	حماس	أصولي
إعتداء	مختطف	حوادث	أبو سيف	فيروس
معندي	تفجير عبوة ناسفة	سلاح	حزب الله	
عنيف	إنفجار	أسلحة	جبهة التحرير الفلسطينية	

Table B.2 *Arabic religion query terms*

سني	غفران	حج	مزار	رب
شيعة	غفران	عمرة	مذابح	الشياطين
شيوعي	رحمة	حلال	مذبح	الشيطان
علوي	أقدار	حرام	معابد	مسيح

Table B2 continued

مخلص	معبد	قداديس	قدر	المسيحية
عيسى	أديار	قداس	أرواح	مسيحي
يسوع	دير	مزامير	روح	أورثودوكس
ملائكة	رهبان	مزمور	مقدس	اليهودية
ملاك	راهب	ترانيم	اسلام	يهود
الجنة	قسس	ترنيمه	الحاد	يهودي
النار	قسيس	أعياد	ارتداد	اليهودية الإصلاحية
الجحيم	أساقف	عيد	تضحية	اليهودية المحافظة
الجهنم	أسقف	عيد الفطر	اعتراف	اليهودية الأرثودوكسية
نجمة داوود	البابا	عيد الأضحى	ضحية	الهندوسية
جوامع	حاخامات	ليلة القدر	تجديف	البوذية
جامع	حاخام	عيد الغطاس	يوم القيامة	الصوفية
مساجد	أنبياء	الاسبوع العظيم	متدين	وثني
مسجد	نبي	عيد الصعود المجيد	ذنوب	وثنية
مآذن	الكريم	إيمان	خطيئة	ملحد
مأذنة	القرآن	أمن – يؤمن	معجزات	الإلحاد
اذان	سور	عبادة	معجزة	علمانية

Table B.3 Arabic soccer query terms

طاح – يطيح	ضربة ركنية	حارس المرمى	وسط الملعب	رياضات
للنهاية	ضربة مرمى	مهاجم	خط الوسط	رياضة
تسوية	ضربة رأس	مدافع	خط المرمى	ألعاب
تعادل	غطس	أمامي	خط التماس	لعبة
بطاقة صفراء	كسب	ضارب	تسلل	كرة القدم
بطاقة حمراء	فوز	ضارب	قائم	كور
الشوط الأول	خسارة	مدرب	عارضة	كرة
الشوط الثاني	تسجيل	ضربة	أهداف	فريق
الاستراحة	احراز	ضربة حرة	هدف	منتخب
كأس العالم	تمرير	غير	لاعبون	مبارات
لاتحاد الدولي لكرة القدم	تمريرة	مباشرة	لاعب	مباراة
	أخطاء	ترجيح	قائد الفريق	استاد



Table B3 continued

ملعب	لاعب الوسط	ضربة جزاء	خطاء	
------	------------	-----------	------	--

## APPENDIX C - Example of document tagged as semantically related to terrorism

This appendix contains an example of a document from our corpus that was tagged as semantically related to terrorism together with the translation to English from Google Translate. The first several pages provide the original Arabic and are followed by the English translation. This translation has not been validated.

يحذر من خطورة القاعدة ! لكن الخشية أن تعتقل الدكتور فوزية الدريع بتهمة الانتماء للقاعدة -2

فجأة الطيران اليمني يقصف مركزا لعمليات القاعدة ! والجزائر تعتقل خلايا للقاعدة واتهام الثورة الليبية بانها -1  
مدعومة من القاعدة ! والعراق

تقبل الله طاعتكم وكل عام وانتم بخير وعيدكم مبارك

كلامي لا يبرئ القبائل إطلاقا, فعدد لا بأس به من حاشية الفساد من القبائل والدليل أنهم أهينوا بعنف ومع ذلك -4  
استمروا مؤيدين وكأن الأمر لا يعينهم

بل كان بعض الشيوخ والتجار داعمين لهذه الحملة و مشجعين لها, وأرشف الأحداث شاهد على ذلك, والله من -3  
وراء القصد وعيدكم مبارك شعبا وشيوخا وتجارا

مع العلم أنهم لا تجار ولا شيوخ ! بل أن أكبر حملة شنت على القبائل وتعمدت إهانتهم والطعن بولائهم و -2  
الإستهزاء بهم لم ينتصر لهم لاشيخ ولاتاجر

التفاعل السياسي في الكويت هو نتيجة مباشرة لمعادلة الكويت = تجار + شيوخ , لكن المستغرب في هذه - 1  
المعادلة أن القبائل دائما في واجهة الأحداث

العفو أتى في محله تصحيحا لوضع غير قانوني في الأساس, أما تحميل الخطأ للطرفين فهو تحول ملحوظ في  
الموقف الرسمي, وندعوا للبحرين وأهلها بالأمن.

قريبا سنقول وداعا بشار وعندها سيدفع هو ونظامه وكل من ساندتهم في إجرامهم ثمن كل قطرة دم أريقَت دون وجه  
حق وسيعلم اللذين ظلموا اي منقلب ينقلبون

وزير الخارجية الإيراني علي أكبر صالحى: على الحكومة السورية أن تلبي المطالبات المشروعة للشعب السوري فهل  
لبيت المطالبات المشروعة للشعب الإيراني !؟

على الرغم من وضوح الجريمة والمساس الواضح بالمقدسات والحرمات والإستهداف الموثق للمدنيين, وبشاعة - 3  
الإجرام والسلمية المطلقة للحراك الشعبي

وفي سوريا يريدون حل سياسي بجدول زمني سيعطي للنظام السوري مساحة زمنية للإبادة الجماعية وتجاهلوا - 2  
تصريحات ايران المؤيدة دون إشارة لها

الجامعة العربية دعمت الحراك الثوري المسلح للشعب الليبي ببيانات واضحة ونددت بالقمع الوحشي وايدت - 1  
التدخل العسكري للناطو وشاركت به فعلا

لذلك من يعقد آمالا على نتائج محددة بشأن الحسابات المليونية لا شك أنه يحلم فلم يسبق ولن يحدث سابقة - 4  
لمحاسبة فاسد واحد

ولم يسبق في كل مما سبق أن حوسب فاسد واحد أو أجري تحقيق جدي واحد يطمئن الناس أن شعارات تطبيق 3- القانون و دولة المؤسسات لم تكن دعاية فارغة

مر الغزو وتبعته المديونيات الصعبة ثم انهيار سوق الأوراق المالية ثم عقود مئات الملايين فالشركات الورقية و 2- سرقة أموال الناس والأغذية الفاسدة

الفساد ظاهرة توجد في كل المجتمعات وبإختلاف النظم لكن ما يحدث في الكويت مسألة تجاوزت الفساد المجرى 1- إلى شرعته والتعايش معه لذلك

من سأل عن موقف في الحالتين لا شك انه غير متابع: فموقعي في الحالتين كان منسجما ومعيارى كان واحدا أما 6- ان لا اكون مرضيا للجانبين فهو أمر منطقي

الخلاصة أن المواقف الطائفية كانت مفضوحة على الجانبين ومتناقضة فمن عارض مطالبات البحرين أيد و 5- بشدة مطالبات الشعب السوري والعكس صحيح

لذلك كان تناقضا أن يؤيد السيد نصر الله ثورة الشعب البحريني و يباركها ثم يناصر علنا النظام السوري و 4- يعارض ثورة شعبه وهو تناقض مضحك

ديمقراطي و لذلك اتفقت مصالح ايران مع مصالح الخليج لكن الفرق الوحيد بينهم في هذا الخصوص أن ايران 3- صرحت بتأييدها أما الخليج فصمت بتأييده

سيضعف النفوذ الإيراني و بالتالي قدرة حزب الله على التواجد الفاعل في معادلة القوى في تلك المنطقة أما 2-

الصمت الخليجي فهو الخشية من نشوء نظام

خطابات نصرالله لن تخدم النظام السوري إطلاقا لكنها حتما ستضعف خطاب حزب الله أمام 1-

جمهور "المقاومة" وهو موقف سياسي بحت لأن سقوط النظام السوري

للتوضيح حول موضوع ال 25 مليون : انظر رأي الدكتور ثقل العجمي المنشور في سبر

<http://t.co/8wbhRHI>

كل رؤساء الحكومات يقطعون إجازاتهم في أي أزمة سياسية إلاسموالرئيس لاتحلو له الإجازات الخاصة إلا

بالأزمات وحتى باستجواباته وكأنه غير معني بها!؟

ليست المرة الأولى التي يغادر فيها رئيس الوزراء في إجازة خاصة في ظل أزمة تستوجب موقفا حكوميا واضحا

فهل الأزمة لاتستحق؟ أم أن الشعب لا يستحق!؟

ليس مقبولا أن يقتصر الموقف الخليجي حيال ما يحدث من جرائم إبادة وحشية في سوريا على مطالبة خجولة بإجراء

إصلاحات وحقن للدماء! ألايعلما سبب ذلك؟

النظام السوري يستهدف مواقع مدنية بشكل عشوائي, واستمرار الصمت العربي حيال هذه الممارسات الإجرامية

يعطي الكيان الصهيوني ذريعة لأي عمل في غزة

التباطؤ الحكومي في التعامل مع أزمة بهذا الحجم ليس له إلامبرر واحد: البحث عن منفذ أجنبي للتغطية على جريمة

تواطأت فيها الحكومة وحاشية الفساد

ستلتزم به مثل هذه الحكومة و أي مصالح شعب ستحميها!!؟-4

أيضا حكومة تجاهر في دفع الرشاوى نقدية "بشنت" و "شيكات" و "تحويلات" وعينية "بتعيينات" ومناصب ولجان -3  
و"رحلات علاج" يتحملها المال العام فأى قانون

فكيف يمكن لمستنقع الجريمة هذا.. أن يعبر عن إرادة الأمة!!؟-2

واقع سياسي غريب:برلمان وصل اعضاؤه"بتزوير"الإرادة الشعبية ويعمل عدد من اعضاؤه"بالرشوة" ويقدم -1  
بعضهم الآخر"إفادات كاذبة"على حساب"النظام العام

وكنت سعيدا أن يكون منزلي محطة من محطات التهنة بأن أشارك الأشقاء في ليبيا فرحة الخلاص من نظام -2  
القذافي متمنيا أن نفرح كذلك مع الشعب السوري

للتو غادر الأستاذين آدم رقيق رئيس تحرير صحيفة ليبيا اليوم والأستاذ حسين بوهدمة و مضيفهم الأستاذ فالج -1  
المطيري للتهنة بانتصار الشعب الليبي

وهو أمر متصور في نطاق أي حرية من الحريات العامة, وسبق للنائب العام السابق أن صرح بذلك لجريدة -8  
,القبس و كذلك اشارت اليه أحكام قضائية حديثة

ضوابط التجريم و نطاقه ويحدد فيه الأدلة المعتبرة و المشروعة في سبيل ذلك فالحريات الشخصية لا يجوز -7  
اهدارها لمجرد احتمال الإساءة أو الانحراف

وإذا كان لدى الحكومة قناعة بأن الاستخدام غير المسؤول لوسائل التواصل الإلكتروني يشكل خطراً فليس هناك -6- ما يمنع من تبنيها مشروع قانون يحدد فيه

و محدداً لأن ذلك ما يجعل منه ضابطاً للسلوك ومقياساً على سلامته أو خطاه، وأما القول بخلاف ذلك فهو -5- تفسيرات "أمنية" وليست تفسيرات قانونية معتبرة

يعتبر بطبيعته قيداً على الحرية الشخصية ومن ثم فإن أي تفسير له يجب أن لا يخرج عن مفترضات التجريم -4- الأساسية وهي أن يكون التجريم صريحاً وواضحاً

مبدأً عاماً استقر باعتباره أصلاً من أصول القانون طالما كانت تطبيقاته قائمة في جميع النظم اللاتينية والانجلو -3- أمريكية على السواء فنص التجريم

النص الجزائي لا يتوسع في تفسير نطاق انطباقه كما أن القياس عليه لا يجوز مطلقاً، ولذلك توجيهات النائب -2- العام وحتى الأحكام خلاف ذلك لا تلغي

تجريم النشر الإلكتروني باعتبار صلته في المكان العام والقواعد العامة للتجريم في السب والقذف تتعارض مع -1- مبدأً عاماً من مبادئ القانون وهو أن

سننتظر نهاية الدورة الرضائية للرأي ونطلب اللقاء بعد رمضان مباشرة فتقديري الشخصي أن اللقاءات المباشرة أكثر إقناعاً للمشاهد من اللقاء المكتوب

للتو كنت في اتصال هاتفي مع مسؤول بمؤسسة الرأي وبلغني اعتذار الرأي عن المقابلة بسبب خطة القناة الرضائية وموافقتهم على نشر الرأي في الجريدة

عموما ابلغت الأخ جاسم بودي برغبتي في شرح كامل للموضوع والرد على تصريحات الأجهزة الحكومية في -7  
هذا الخصوص وفي انتظار موافقة الراي أوعتذارها

بإختصار شديد لا يوجد عذر لأي جهة كانت واقعا و قانونا و أي كلام خلاف ذلك ليس له اساس قانوني.ولدي -6  
استعداد أن اثبت ذلك علنا وأمام أي جهة

وأنها قد أبلغت البنك المركزي والبنك المركزي قد تقاعس عن إبلاغ جهة الإختصاص وهوأمريشکل جريمة, -5  
والنيابة العامة أتصلت علما بالواقعة فعليا

للمعلومات: البنوك مكلفة قانونا بالإبلاغ عن "أي شبهة" في "أي حساب مصرفي" وهذا يعني أن البنوك قد -4  
ارتكبت مخالفة للقانون بعدم الإبلاغ

أما الحكومة وتصريحاتها فذكرتني بحريق الجهراء واللجنة التي شكلت حينها وتقاريرها التي لايعلمها أحد -3  
والمطلوب من الناس أن تتق دائما بالحكومة

اما تصريحات محافظ البنك المركزي فتفيد انه لا يمارس اي رقابة على البنوك إلا في مؤتمراته الصحفية,كما -2  
ان فهمه لمعنى الرقابة المصرفيه معدوما

تأكيدات القبس في الرشاوى المليونية تفيد أن إجراءات البنوك خاطئة والبنك المركزي لايمارس رقابة وتلزم النائب  
العام باستدعاء المحرر لسماع اقواله



و شماعه لعدم القيام بإصلاحات سياسية حتمية تستوجبها ظروف المرحلة وطموحات الشعوب في المشاركة و -6  
اتخاذ القرار و تكريس مفاهيم العدالة والقانون

فكيف تدعم نظاما مقاوما للطغيان وتقمع ثورة شعبها المدنية المطالبة بالإصلاح والتغيير!، لكن هذا لا يعني ان -5  
تستخدم دول الخليج ايران مبرر

دعم ايران لحماس ليس مبنيا على ايمان بدعم المقاومة بل لإستخدامها أداة ضغط والدليل تغير الموقف من حماس -4  
بسبب موقفها المعارض لقمع النظام السوري

كل من أعان مجرما على إجرامه بالفعل أو القول أو الصمت فهو شريك له فيما يصنع بالأبرياء، وهذا الأمر -3  
ينسحب على الأفراد كما ينسحب على الدول كذلك

باهظ التكاليف، والدليل أن "مجرم دمشق" لم يوقف مطالبات مدنية سلمية حتى باستخدام الدبابات والبارجات -2  
والقتل والترويع وستكون نهايته كنوع أعماله

من النتائج الحتمية أن "أي نظام سياسي" لا يستجيب لمتطلبات التحول الديمقراطي الحقيقي والمشاركة فلن يصمد -1  
إلا بحماية الدبابات ومؤقتا وهو خيار

بينما انحازت باقي دول الخليج للنظم القمعية فدعمتها حتى النهاية فانهارت الأنظمة وانتصرت الشعوب " وكل -2  
"نفس بما كسبت رهينة

نجحت قطر في الوقوف بجانب الشعوب فكسبت احترام الجميع ونجحت الجزيرة في أن تكون الناطق الرسمي -1  
باسم الثورات فشاركت فيها، فهنيئا لقطر وللجزيرة

إن ما يحدث في المحيط هو درس بالمجان , بأن الحرية والكرامة لا يستحقها إلا من يكون لديه استعدادا أن يضحي من أجلها فالقيم العظيمة لا تمنح للجبناء

إن ما يحدث في المحيط هو درس للجبناء, بأن الحرية والكرامة لا يستحقها إلا من يكون لديه استعدادا أن يضحي , من أجلها فالقيم العظيمة لا تمنح بالمجان

عام 2011 عظيم , فيه أخليت الزنزانة لينزل بها حسني و بن علي و القذافي, واعتقد انها لازالت تتسع للمزيد, يالها من سنة واتمنى أن لا تنتهي أبدا

يوم تاريخي آخر في حياة الأمة, و ستشرق شمس ليبيا غدا بلا طغيان وبلا القذافي وعصابته وسنحتفل قريبا بانتصار, الإرادة السورية, اللهم آمين

و أما غير ذلك فيعني إما أن تكون الحكومة غير محترمة فعلا, أو أنها ترى أن الشعب أصلا لا يستحق -2- الإحترام وهو الأرجح.

عندما يثار موضوع بهذه الخطورة فإن ابسط إجراءات أي حكومة محترمة أن يخرج مسئول يعرض فيها وجهة نظر الحكومة في الأزمة والإجراءات التي ستقوم بها

البنك المركزي كالعادة لن يسأل عن أي شئ والبنوك في الكويت هي الأكثر نزاهة في المنطقة والنواب جميعهم شرفاء والمتهم حتما هو "ماكينة السحب الآلي

وفى وجوده اصبحت جوازات السفر وثائق عديمة القيمة واصبحت الحدود مسألة اختيارية كما انه ساهم بتقارب -2  
عربي فشلت عن تحقيقه منظماتها واجتماعاتها

وجود برنامج كالتويتتر من ضرورات الأمن السياسي للشعوب العربية فهو مكان لا تستطيع الحكومات فيه أن -1  
تحدد للناس مايقولون أو ما يجب عليهم سماعه

أما الخليج فيعتبر انتصار الإرادة الشعبية بذاته تهديدا لأمنها السياسي لأنها ستصبح المنظومة الوحيدة في العالم -4  
التي لا يشارك الشعب بالقرار

لإعتبارات متعلقة بالأمن القومي لأنها لن تجد أمنا أفضل من هذا الجوار, وايران تعتبر النظام السوري حليف -3  
استراتيجي وناطق باسمها في العمق العربي

ومن المفارقات الغربية أن حكومات اسرائيل وايران والخليج تقف مع النظام السوري رغم اختلاف منطلقاتها -2  
في ذلك: فاسرائيل تدعم النظام لسوري

الغارة الإسرائيلية على الحدود المصرية ومقتل الجنود المصريين مثال واضح على حرص دولة إسرائيل على  
وجود النظام السوري ودعمه ويفضح كذبة المقاومة

على بشار الأسد أن يتخذ أهم قرار في حياته باللجوء لروسيا, فنهايته باتت محصورة بين أمرين: محاكمة دولية  
نتيجتها معروفة أو توزيع اشلاءه بشوارع دمشق

الحكومة الشكلية ونوابها لازالوا يعتبرون أن التجاوزات الدستورية واللائحية تكتيكاً سياسياً ! فكيف يقتنع الأفراد  
بشعارات إحترام القانون !؟

الأمر الأخطر: لدينا مجلس لا يراقب ولا يريد معظم اعضاؤه ذلك وحكومة شكلية لاتعمل وليست قادره على -3

العمل أصلا فهل الدولة مسيرة بالجن لاسمح الله

الأمر الآخر: إذا كان الإخلالات الهيكلية سببها قرارات تنفيذية وهدر مالي حكومي فكيف يمكن قبول أن تخطأ -2

الحكومة و يحاسب الناس على ذلك!!؟

إذا كان كل إشكال تشكل له لجنة لتقدم توصيات ثم تنبثق عنها لجنة أخرى فالسؤال ماهي مهام مجلس الوزراء -1

أصلا؟ من الناحيتين الدستورية والواقعية؟

الأهم من ذلك أن انتقادنا المواقف الخليجية لا يعني بالضرورة نسيان أن النظام الايراني نظام قمعي غير إنساني -7

ومواقفه مبنية على اسس طائفية بحتة

من يشير الى مصطلح العلماء الربانيين اقول له: من هؤلاء تحديدا؟ وما موقفهم من قتل مسلم وهدم مسجد ومنع -6

صلاة؟ و ماذا قالوا في ذلك؟ واين ومتى؟

فهل لا يزال الإتهام قائما؟ وهل هؤلاء من أهل الفتن والبدع؟! والسؤال الأهم من هم العلماء المعتبرين تحديدا؟ -5

و ماهو رأيهم في الأحداث؟!؟

عندما وصف البعض منهم من طالب الناس بالإعتراض والتظاهر "بدعاة الفتنة والبدع", لكن بعد تأييد الشيخ -4

عبدالرحمن عبدالخالق وعجيل النشمي وغيرهم

صدرت قرارات في معظم دول الخليج لمنع المساجد من التعرض للأحداث السياسية، ولم يعترض "العلماء" على -3- ذلك! فهل هو إقرار وقبول بفصل الدين عن السياسة؟!

صدرت فتاوى تم تسويقها إقليمياً لتحريم التظاهرات، ثم سمحت السلطة فسكتت الفتاوى! والسؤال هل تم التراجع -2- عنها؟ أم أنها لا زالت صالحة للاستعمال

سمحت حكومات الخليج لشعوبها بالتظاهر تأييداً لمطالب الشعب السوري واعتراضاً على قمع سلطته، لكن لماذا -1- منعت شعوبها للتظاهر للمطالبة بحقوقهم؟!

فطالما كان استقباله لدواعي إنسانية فيجب أن تستمر إقامته على هذا الأساس إلى أن يغادر أو يطلب اللجوء -2- السياسي وكان الواجب تنبيهه لذلك

السعودية استقبلت صالح لدواعي إنسانية كما صرح بذلك، لكن أن يستخدم السعودية كمحطة لإطلاق تصريحات سياسية للشعب اليمني فهو أمر غير ملائم إطلاقاً

استمع الآن لخطاب الشاويش صالح وسبق أن استمعتم جميعاً لخطابات القذافي لكن التساؤل المنطقي كيف يمكن لهذه الكائنات الحية الغريبة أن تحكم شعباً؟!

و الحقيقة أنني لأجد سبباً مفهوماً للإبقاء على هذه العلاقات في ظل هذا الوضع المأساوي سوى منح النظام السوري فرصة لمزيد من القتل والوحشية والإباد

لازال الشعب السوري يذبح بلا رحمة ولازال الخليج يصدر بيانات غير مفهومة في العلن ويعتذر عنها في السر ولا تزال العلاقات مع النظام القاتل قائمة

ألا يجب أن يعرف حجم إسهام القطاع الخاص في الناتج الوطني أولاً؟! حتى يعطى صلاحية المساهمة في رسم القرار الإقتصادي أصلاً!؟

الحكومة صاحبة الحلول الإرتجالية والتفيعية يجب أن تتحمل تبعه قراراتها الخاطئة، أما أن تخطئ الحكومة ويتحمل المواطن تبعه ذلك فهذا أمر غير مقبول

تمنيت لو أن نفقات "غبقات الإستعراض" قد تبرع بها اصحابها لجوعى الصومال ومشردي سوريا بدل انفاق اموال طائفة على اشخاص شيعى أتى بعضهم للفرجة

ولذلك فإن شيوع الظلم السلطوي وتحقق اليقين الشعبي بوجوده وانكار السلطة له يخلق ارضية خصبة لرفض -4- الواقع والعمل على تغييره ومن هنا تبدأ الثورة

ثارت الشعوب لأن الظروف و ضعتها أمام خيارات تساوى فيها الموت والحياة فجريت هذا النمط الجديد -3- ببساطة لأنه ليس لديها ما تخسره

فأعلى درجات الفساد السياسي أن لا يكذب الفاسد لإعلامه يتولى المهمة ولا يقتل بيده فالأمن موجود ولا يسلب -2- الناس حقوقهم فالقضاء يقوم بالدور لذلك

ثورات الربيع العربي لم تكشف النظم الفاسدة فقط بل كشفت فساد أدواتها في القمع والظلم وفي كل مرة أمن فاسد -1- قضاء فاسد, إعلام فاسد, فهل من متعظ؟

بشار يدخل التاريخ كأول رئيس دولة يعلن الحرب صراحة على شعبة الأعزل ويقصف اللاذقية برا وبحرا و التضامن الآن يعني تقديم طلب محاكمة أركان النظام

إن أي مبادرة للإصلاح السياسي لاتؤسس على قواعد واضحة للمسائلة والمحاسبة فإنه لايمكن إعتبارها من الناحية السياسية إصلاحا وأيا كان مسمها

فكل يؤخذ من كلامه و يرد, أما ترهيب الناس وتخويفهم بغطاء شرعي فهي سياسات أمنية وليست سياسات -5- شرعية أصلاو ولن يمنعنا أحد من قول الحق أيا كان

جاء الدين بالكرامة والحرية والمساواة, وأي ممارسة تتعارض مع هذه القيم ليست أصلا من الدين ويجب -4- رفضها, كما أن الإسلام لا يعرف الكهنوت

أليس من الأولى بدل الإنشغال باتهامي بالفتنة والدعوة للخروج على بشار,إنكار هدمه المساجد وقتله الأبرياء -3- ومنعه للصلاة! فأني حرص على الدين هذا

ولمن اتهمنا بالتحريض على الفتنة والخروج على الحاكم أيضا: أنظر فتاوى الشيخ الزنداني والشيخ القرضاوي -2- وعلماء الأزهر والشيخ العودة وغيرهم

ولمن اتهمنا بالتحريض على الفتنة والخروج على ولي الأمر: أيضا شيخ الأزهر "لايجوز شرعاً السكوت عن -1- "المأساة التي يتعرض لها الشعب السوري

لمن اتهمنا بالتحريض على الفتنة والخروج على ولي أمر سوريا باعتباره غير جائز شرعا هيئة علماء المسلمين

<http://t.co/gFKZQkW> <http://t.co/ElYEsKf>

ولذلك كوارث المناخ وانهيارات سوق الأوراق والمديونيات الصعبة والسهلة! وتجاوزات البنوك وشركات -3

الاستثمار وعبث المناقصات فهي مجرد اجتهادات

أما ما يقوم به اي فرد ومهما كان عظيما فهو غير وطني مالم يقرروا انه كذلك. والغريب أن السلطة تصدق -2

. وتطالبنا بالتصديق

مشكلة " النخبة الإقتصادية" في الكويت أنهم يريدون احتكار كل شئ، الإقتصاد والسياسة والرياضة -1

والشعروالأدب ,وهم اصحاب التاريخ الوطني والتضحيات

لا تنتشط إلا في البرك العفنة,لذلك امرحتمي ان يقفوا خلف كل ماهو فاسد لأن خلق بيئة نظيفة سيكون قاتلا لهذا -4

النشاط.لذلك صمتوا في ثورات التصحيح

مليارية في سوريا ومصر(قبل الثورة)ولبنان على الرغم من من أن اسواق اوربا وامريكا والدول الإسكندنافية -3

مفتوحة, ببساطة هي استثمارات بكثيرة

ولذلك عارض رأس المال الكويتي كل الثورات لذات السبب لأن ما يسمى بالقطاع الخاص لا ينشط إلا في البيئة -2

السياسية الفاسدة وغلا ما مبرر استثمارات

الصمت عن ما يحدث في سوريا من البعض ليس قبولا به ,لكنه الخشية من ردة فعل النظام على استثمارات أوجدها

الفساد ونماها الفساد وسيقتلها المحرومين



بطائفيتنا و فنويتنا لأننا لم ن تعود أن ننشغل بانتاج دواء أو تأمين غذاء أو منافسة المدنيات الرائدة وهي المتطلبات -4- الأولية لوجود أي بناء مدني

منها للدول واقتصر البناء المدني والتطوير على المباني الإسمنتية والمظاهر الشكلية المدنية دون بناء حقيقي -3- للإنسان لذلك يعد طبيعيا أن ننشغل

فتضعف قدرة المجتمعات على المشاركة الفاعلة والرقابة و المحاسبة وهو ما حدث فعلا حتى تحولت مجتمعات -2- المنطقة الى كيانات اقرب للملكيات الخاصة

وما يحدث في الكويت وغيرها من تمييز مفضوح ليس ميلا حقيقيا لفئة دون أخرى لكنه سياسة وضعها الإنجليز -1- وسارت عليها الحكومات لإشغال الناس ببعضهم

القانون قواعد عامة مجردة تخاطب سلوك الفرد لا لقبه أو معتقده , لكن أعمال السلطة هي ما يصنع التفرقة بالتمييز والانتقائية على هذه الاسس غير الشرعية

لرفض القانون وعدم الإنصياع له لأن قبوله وفقا لهذه المعطيات معناه التسليم به أداة للإبتزاز وهو غرض -4- غير مشروع وصورة للانحراف في استعماله

التي لا يعرف القائمين عليها إلا فئات دون غيرها, ويجب أن تعلم الداخلية ومن يسايرها في هذا النهج أن -3- الانتقائية في تطبيق القانون مبرر مشروع

وعن صاحبة القناة اللي قالت " ترى إحنا ملينا و انتوا...., وعن الجهة التي حجبت عرض امراميري موقع -2- ومكتوب عن جهة الاختصاص به, غريب أمر الداخلية

أين الداخلية وأمن الدولة عن النائبة التي قالت ان منحة الألف قرار خاطئ ومتسرع! وعن الكاتب الذي قال -1-

" " الشرهه على اللي عفا عنهم

لن نجامل في القانون أحدا والحق الذي لا يجد من ينطق به يتحول باطلا, ومن يسكت عنه شيطان وأما إتهامات

". الداخلية فهي كتصريحات ومؤتمرات "الرصيف

و صمتكم المريب عن "حريته في الحياة" فما قيمة "الجينز" اصلا لإنسان ميت؟! إن الحرية مفهوم اعمق بكثير -5-

من أن يختزل في الملابس والمناقصات!!؟

لذلك منطقي جدا أن نتساءل عن موقف النواب والنخب اللبيرالية كما سوقت لنفسها كيف يقبل دفاعكم الدائم عن -4-

حرية الأفراد في الملابس

والمحصلة النهائية أنه إما أن يكون فهم الحرية بالأساس كان خاطئا, أو أنها مجرد شعار يستخدم للتسويق -3-

الانتخابي دون أن تكون ثقافة يترجمها موقف

لكن غير مفهوم إطلاقا أن يعارض منع عمل المرأة مساء باعتبار ذلك إعتداء حرية عامة يجب الدفاع عنها, ثم -2-

تسكت مطلقا عن قتل يحدث صباحا ومساء

بصرف النظر عن ما يقال في مهرجانات التضامن مع سوريا, أثبت الإسلاميين موقفهم وبصرف النظر عن -1-

دوافعهم لكنه بالنهاية يلتقي مضمونا مع مفهوم الحرية

مطالبة المعارضة اليمنية بمحاكمة علي صالح في اليمن مطالبة مشروعة ومستحقة واعتبارات العدالة ومتطلبات التغيير تستدعي أن لا يفلت أحد من العقاب

جماعات ارهابية,خونة,عملاء,جرذان,خلايا القاعدة,وهابيين,صفويين,هذه بعض اوصاف الحكومات العربية لمن عارضها وفي المقابل هي حكومات عادلة و محترمة

نتمنى من المشاركين في الاعتصام الابتعاد عن الإطروحات الطائفية لأن آلة القتل والوحشية لم تفرق بينهم على هذا الأساس كما انه ليس في صالح ثورتهم

الأسد يتحدى العالم مجددا ويصمم على محاربة الجماعات الإرهابية المسلحة التي تقتل السوريين !!! فهل الكذب الغبي اختراع الحكومات العربية!؟

اللهم أجعلني مع الحق حيثما كان,وأن لأتهم أحدا بلادلل ولاأرد على أحد بغير حجة, اللهم أجعلني من الناصحين. والسلام

لأن الشئ بالشئ يذكر دعاة الحرية في الكويت أطلقوا تصريحات في حق المرأة في ارتداء ملابس السباحة -10  
وسكتوا عن حق الشعب السوري في الحياة

ومن باب أولى أن لا يكون عالما فلا تتحججوا بالسنة وتلبسوا عليها ما ليس منها لأن سنة الرسول(ص) مطهرة -9  
عن أباطيل المنافقين وفتاوى الأمن الرسمية

المسلم الذي لا يتألم لمعاناة ابرياء يقتلون بلا ذنب ولارحمة وعجوز تستنجد وطفل يبكي خوفا و مدينة تمنع فيها -8  
الصلاة بقرار رسمي ليس انسانا أصلا

أخيرا وليس آخرا :ما حكم الشرع في قرارالحكومات الخليجية في منع الخطباء من التعرض للشأن

السوري؟حلال أم حرام؟ وإذا كان حرام لماذا لم ينكروه؟

سمعنا رأيهم في التظاهر ووجوب الطاعة لولي الأمر لكننا لم نسمع رأي ذات المفتين في جوازحبس انسان

بشبهه؟ أو حبسه دون اتهام؟أو منع الخطباء؟

الإجابة على سؤال لماذا المشايخ تحديدا؟ فالإجابة ببساطة: أن أعمال القمع التي تقوم بها السلطة تستند على آراء

بعضهم ومؤيده من البعض الآخر

ورجال الدين بشر لاعصمة لهم ولا هم بمنزهين عن الخطأ و "كل ابن آدم خطاء" ومن يمارس نشاطا عاما

يعرض للنقد فليسوا بأكرم من الخلفاء الراشدين

و للمعلومات التخاذل موجود في كل ميدان فهناك سياسيين متخاذلين ورجال دين متخاذلين واساتذة قانون كذلك

والمنافيين ورد ذكرهم في القرآن

فإذا كان التظاهر حرام وهي مسألة محل خلاف كما هو واضح فهل قتل نفس بلا ذنب أو هتك عرض أو منع

صلاة أو هدم مسجد حلال!؟

لأعلم سبب الجدل القائم حول فتاوى تحريم التظاهر الآن: لكن الأكيد ان اثره صرف الأنظار عن الأمر الأساسي وهو

قطعية حرمة قتل نفس ومنع صلاة وهدم مسجد

ننتظر إبادة الشعب كاملا !!؟-5

ما يقارب 2000 شهيد وأكثر من 20000 معتقل ومحاصرة بالدبابات ومنع الصلاة وقتل أبرياء يجب أن لا نبالغ -4

بردة فعلنا ومتى يجب طرد السفراء؟ و هل

عزل النظام دوليا بتقديم طلب واضح لمحاكمة أعضاء النظام تشجيعا لهم للتبرؤ منه وتخيير السفراء بين طردهم -3

أواستقالتهم ومنحهم لجوء سياسيا

كما أنه لن يساهم في منع القتل و التشريد و المطلوب مواقف جدية تعكس حقيقة التضامن مع قضية عادلة, -2

وأولها منع المساعدات بكل أنواعها والعمل

المبالغة في الإطراء على ردة الفعل الخليجية المتواضعة والمتأخرة ليس أمرا محمودا فاستدعاء سفير إجراء -1

بروتوكولي يحدث في ماهو دون ذلك بكثير

حكومتنا الله يهديها أحيانا تخلي بعض المشايخ والليبراليين في موقف بالايخ.. الحين وش RT @Alqudaimi:

... يسوون بالفتاوى والمقالات الي تجرم المظ

و المعروف ان السكوت في معرض الحاجة للكلام يعد تصريحا وهي قاعدة أصولية, فأين كان هؤلاء في منع -2

صلاة الجمعة بحماة بقرار رسمي معلن, من منعهم؟

كل صاحب عالم سكت عن قتل مسلم وهتك عرض وهدم مسجد ومنع صلاة دون أن يكون له رأي فلا قيمة -1

لعلمه والعبارة تنسحب على كل من تتوافر فيه الشروط

إنه ليس عدلولا إنصافا أن نساوي موقف العالم الدكتور النشمي بموقف مفتى الفوائد الربوية وهل يستقيم أن -2-  
نساوي البوطي الذي أجاز القتل بمن حرمه

كل ما ذكرت لفظ علماء السلاطين انزعج البعض من الوصف! فهل كل رجل دين عالم سلطان؟! إنه أمر غريب أن  
ينفعل البعض منها وكأنه يقول.. لماذا تشتمني!؟

حتى ولو كانت خطوة متأخرة: على دول الخليج تقديم طلب لإحالة مجرمي النظام السوري للمحكمة الجنائية الدولية  
كأقل ما يمكن تقديمه لشعبها وطرد سفرائه

وطبيعي أن هذا الكلام لا ينسحب إلا على فئة واحدة وهم السادة دعاة الأجهزة الأمنية لأن اجتهاداتهم مرتبطة  
بموافقة السلطة الأمنية عليها سلفا

رغم تأخر إعلان المملكة لموقفها كثيرا إلا أن من نتائجه المباشرة غير السياسية انه فتح الباب إعتبارا من غدا لخطب  
نصرة سوريا وفتاوى الجهاد والتكفير

النتيجة الحتمية لهذا الخطاب وضع الحكومة في مأزق وغالبا سيستدعى السفير السوري غدا على "غبقة" و ربما يقال  
له بالكويتي "تري اللي تسوونه كلش موشي

بابا الفاتيكان يندد بالقمع السوري وانجلينا جولي تقود حملة تبرعات للسوريين ومي سكاف تعترض في دمشق: ماذا  
تبقى لنا نحن ! الحديث في آداب الصوم!؟

عندما زار الأسد الكويت امتلأت صفحات الجرائد بإعلانات الترحيب التي دفع فيها مئات الآلاف فهل تبرع المعلنون  
للأجنبي سوريا بأعشار هذه المبالغ

من الحقائق المؤلمة أن استثمارات القطاع الخاص لا تنشط إلا في كل بيئة فاسدة ولذلك صمت الوطني عن جرائم الأسد اسبابه اقتصادية في المقام الأول

داعية الشيطان سكت لقتل الأبرياء وهدم المساجد ومنع الصلاة وفجأة تحدث لبيان رأي الشرع في "مشروع" فتوى إهدار دم السفير! قبح الله امثالك يا مدلس

أما توصية الحث على حل مشكلة قبول الطلبة فيبدو ان رئيس الوزراء لا يعرف الفرق بين مجلس الوزراء وهو اعلى جهاز تنفيذي و بين مجلس اولياء الأمور

الدبابات تقتحم دير الزور وتقتل 50 مدنيا في الإقحام ومجلس الوزراء الكويتي يبقي على التمثيل الدبلوماسي ويدعو لتفعيل الحوار! ياله من موقف عظيم

فتاوى إهدار الدم ليست ثقافة لأحد دون أحد و"هو أمر مرفوض بالنسبة لي" لكن سبق ان صدرت فتوى بإهدار دم سلمان رشدي فهل يجرؤ القلاف على استنكارها!؟

والغريب أن يصدر البيان "غير المحترم" في أعقاب صدور قرار النظام السوري بمنع صلاة الجمعة في حماة وقصف! المساجد: لاحول ولا قوة إلا بالله

بيان التعاون بشأن سوريا عبارات لاتحمل مضمونا أو دلالة, كما يوحي بأن الشعب السوري شريك بإراقة الدماء, عموما أوقفوا دعمكم للنظام وشكر الله سعيكم

شاركت في تجمع السفارة لقناعاتي بجدوى رمزية التجمع هناك لأن الرسالة ستكون أكثر وضوحا أما الحديث -6

فليس مهما دائما, فليس لدي أكثر من ما أقوله هنا

السفير فدعوة مرفوضة شكلا ومضمونا ولا يجب أن تؤخذ كمبرر لإختزال كل إيجابيات الحراك وتفسير لتخاذل -5

,الآخرين عن قول كلمة الحق و نصرة المظلوم

و من ينتصر لسوريا وشعبها عليه أن لا يضعف ثورتها بإعطائها بعدا طائفيا لأن ذلك ما يروج له النظام القاتل -4

وما يريده :أما استفتاء استباحة دم

الصمت الذي يقتلهم, وأنا متأكد بأن الكويت التي صمتت في أمس لن تصمت الآن حيال ما يحدث من قتل وقمع -3

وبطش وستكون الكويت شعبا وحكومة صوتا للحق

كما أن إضفاء البعض طابعا طائفيا على الإعتراض لايعكس حقيقة القمع السوري الذي لايميز بين دين أو -2

طائفة, لكننا في أقل تقدير كسرنا جزء من

رسالة التضامن الصريحة التي بعثها الشعب الكويتي لشعب سوريا "السجين المذبوح"لايجوز اختزالها بعبارة -1

مندفعة متأكد أن من قالها لا يدرك أبعادها

أرى ان السفارة السورية هي المكان المناسب للإحتجاج والإعتراض على اعمال النظام في سوريا, مالم يكن -2

الغرض مجرد اثبات موقف إبراء للذمة

لا أملك صلاحية توجيه الناس للإعتراض بشكل معين لكن اعتقادي ان ساحة الارادة ليست المكان المناسب -1

لتوجيه رسالة احتجاج من هذا النوع لذلك



وانكار الحقوق والوحشية جريمة في كل الأديان والملل والشرائع وليس من الإنسانية أن تسعد بألم مظلوم أو -4-  
موت برئ لإختلافه معك في الراي أوالمعتقد

عندما نتعامل من منظور الإنسانية كقيمة,نستطيع ان نزن الامور بعدالة,أما المواقف الإنتقائية فتعكس قصورا -3-  
في الحس الإنساني وقتل الابرياء

ولذلك اتى رأبي في سوريا منسجما مع ما ذكرته في البحرين واليمن,وهو موقف مبدأي لايتأثر بمحيط الراي -2-  
الطائفي جدا أما مواقف غيري فيسألون عنها

أنا أنحاز للمطالبات المشروعة لأي شعب ولذلك كان رأبي واضحا وصريحا في موضوع البحرين مؤيدا لحقهم في  
مطالبات مشروعة وتنديد بكل صور الوحشية والقمع

في ساحة الإرادة فهو إبراء ذمة لا يتناسب مع حجم الجريمة وبشاعتها علاوة على ذلك جرت العادة ان يكون -2-  
التواجد الأمني بها اكبر من عدد المشاركين

من يريد نصره الشعب السوري عليه التوجه للمكان الذي تصل فيه الرسالة للأسد مؤثرة وهوسفارة سوريا أو -1-  
وزارة الخارجية وبمطالب محددة اما التجمع

التيارات والكتاب والمؤسسات التي تحدثت في الشأن السوري بعد إدانة مجلس الأمن وتصريح الخارجية الأمريكية:  
فهل تنبهوا للوضع الآن أم اخذوا إذننا

وفي سوريا يقتل الأطفال والنساء في شهر رمضان وتهدم المساجد وتعتبرون مجرد التعرض للموضوع فتنة ! -2  
فأي فتنة أكبر من تدليسكم على الدين في رمضان

رسالة اخيرة لمشايخ الجهاز الأمني في احداث 11 سبتمبر جندت المنابر لإنكار الجريمة وصدرت الفتاوى 1-  
بتحريم ذلك وتباريتم على الظهور والكتابة و في

باعتبار ما يحدث في سوريا من تصفيات يمثل جريمة إبادة جماعية وهي جريمة تعطي الحق لأي عضو في 3-  
الأمم المتحدة في طلب عرضها على جهة الاختصاص بها

العمل على بلورة مشروع يقدم لمجلس الأمن و يعرض على دول التعاون الخليجي بطلب تفعيل ميثاق روما 2-  
ومحاكمة المتسببين في قتل المدنيين السوريين

فأولا:مندوب قاتل الأطفال والنساء يجب أن لايبقى في الكويت يوما إضافيا,ثانيا:بيان الحكومة بشأن موقف 1-  
الكويت الرسمي يجب ان يكون واضحا وصريحا

حتى لايتحول مهرجان التضامن مع الشعب السوري إلى مهرجان استعراضى لا يحمل تصورا أو رؤية تخفف معاناة  
الشعب السوري يجب أن تكون المطالب واضحة ومحددة

لهذا السبب لم يكن قتل الأفراد أو انكار حقوقهم المشروعة دافعا كافيا لإستنطاق أي نظام عربي فالقتل و انكار 2-  
الحقوق وسائل تستعملها كل السلطات

الثورات العربية اثبتت انحياز السلطة للسلطة والشعوب للشعوب, ولفهم هذه المعادلة يجب أن تعلم ان السلطات 1-  
العربية لم تكن في اي بلد جزء من الشعب

سقوط نظام الأسد أصبح امرا حتميا, لكن توقيت هذا السقوط وآلية محاسبة أركان هذا النظام هو الأمر الذي سيحدده

#Syria الشعب السوري قريبا

حزن البعض على مجرد محاكمة حسني مبارك واعتبروها إهانة له رغم الضمانات الواضحة, لكنهم تناسوا إهانة  
ثمانين مليون انسان والـ ألف سنة من تاريخ مصر

و أن الإهانات و الآلام لا تنسى مهما طال عليها الأمد -2-

رمزية محاكمة الرئيس المصري السابق ومثوله وأركان حكومته أمام قضاء بلاده رسالة بالغة تترجم حراك -1-  
الشارع العربي وطموحه للحرية والكرامة

اتصال مواطن سوري بقناة الجزيرة جاء فيه " عندما تهدم المساجد وترمل النساء وتنتهك الأعراض في رمضان فإن  
صيام المسلمين يصبح كجوع الكلاب". لا تعليق

إن العلم الذي لا يترجمه عمل أو موقف ليس علما والعالم الذي لا ينكر هذه الجريمة قولاً وعملاً ليس عالماً أما من -5-  
يستخدم الدين لتبرير الإجرام فهو مجرم

هل من الدين أن يقتل الناس بلا ذنب؟ وما حكم قتل مسلم في المسجد؟ وفي رمضان؟ وما حكم قصف المساجد؟ -4-  
سؤال لمن سمتهم السلطة علماء الأمة

ما يحدث في سوريا الآن جريمة إبادة جماعية والسكوت عن ما يحدث تواطؤ على قتل شعب أعزل لم تراعي -3-  
حكومته حرمة رمضان أو قداسة المساجد

هل لا يزال بشار ونظامه ولاة أمر لا يجوز الخروج أو حتى الدعاء عليهم؟ أم أن التوجيهات الأمنية هي المصدر -2-  
!المباشر لنشر فتاوى التدليس على الدين

بعد قصف الدبابات لمسجد الحميدية وقتل المصلين فيه, لا أعلم ماهي التبريرات الجديدة لهيئات التدليس و دعاة -1-  
الشياطين وائمة السلاطين, والسؤال

2- He warns of the danger of Al-Qaeda! However, the fear is that Dr. Fawzia Al-Durai will be arrested on charges of belonging to Al-Qaeda

1- Suddenly the Yemeni air force bombed an Al-Qaeda operations center! And Algeria arrests al-Qaeda cells and accuses the Libyan revolution of being supported by al-Qaeda! And Iraq

May God accept your obedience, and every year you are well and your Eid is blessed

4- My words do not absolve the tribes at all, for there is a good number of the corrupt entourage from the tribes, and the evidence is that they were violently insulted. Nevertheless, they continued in support as if the matter did not concern them.

3- Rather, some sheikhs and merchants were supportive of and encouraging this campaign, and the archive of events is a witness to that, and God is behind the intention, and your Eid is blessed by the people, the sheikhs, and merchants.

2- Knowing that they are neither merchants nor sheikhs! Rather, the biggest campaign was launched against the tribes and deliberately insulted them, challenged their loyalty, and mocked them. No sheikh did not win them and did not trade

1- Political interaction in Kuwait is a direct result of the equation Kuwait = merchants + sheikhs, but what is surprising in this equation is that the tribes are always at the forefront of events

The amnesty was in order to correct an essentially illegal situation. As for blaming the mistake on both sides, it is a noticeable shift in the official position, and we call on Bahrain and its people for security.

Soon we will say goodbye to Bashar, and then he and his regime and all those who supported them in their crimes will pay the price of every drop of blood that has been spilled unjustly, and those who have wronged will know which turn they are upset.

Iranian Foreign Minister Ali Akbar Salehi: The Syrian government must meet the legitimate demands of the Syrian people, so does it meet the legitimate demands of the Iranian people !?

3- Despite the clarity of the crime, the clear prejudice to sanctities and sanctities, the documented targeting of civilians, and the ugliness of the criminality and the absolute peace of the popular movement

2- In Syria, they want a political solution with a timetable that will give the Syrian regime a time frame for genocide, and they have ignored Iran's supportive statements without any reference to it.

1- The Arab League supported the armed revolutionary movement of the Libyan people with clear statements and condemned the brutal repression and supported the military intervention of NATO and actually participated in it.

4- Therefore, whoever holds hopes for specific results regarding million-dollar accounts, there is no doubt that he dreams. He has never, nor will there ever be a precedent for accounting for one corrupt

3 - Never in all of the above has been a single corrupt computer or a single serious investigation that reassures people that the slogans of law enforcement and the state of institutions were not empty propaganda.

2- The invasion passed, and difficult debts followed, then the stock market collapsed, then hundreds of millions of contracts and paper companies, and the theft of people's money and spoiled food.

1- Corruption is a phenomenon that exists in all societies and with different systems, but what is happening in Kuwait is a matter that goes beyond mere corruption to legitimacy and coexistence with it.

6- Whoever asks about my position in both cases is undoubtedly not following up: My position in both cases was consistent and my standard was the same. As for not being satisfactory to both sides, it is logical.

5- In conclusion, the sectarian positions were exposed on both sides and contradictory. Whoever opposes the demands of Bahrain strongly supports the demands of the Syrian people and vice versa

4- That is why it was a contradiction that Sayyed Nasrallah supported and blesses the Bahraini people's revolution, then fights for the Syrian regime and opposes the revolution of its people, which is a funny contradiction.

3- Democratic, and therefore Iran's interests agreed with the interests of the Gulf, but the only difference between them in this regard is that Iran declared its support, while the Gulf remained silent in support of it.

2- The Iranian influence will weaken and thus the ability of Hezbollah to effectively exist in the power equation in that region. As for the Gulf silence, it is the fear of the emergence of a regime.

1- Nasrallah's speeches will not serve the Syrian regime at all, but they will inevitably weaken Hezbollah's rhetoric in front of the "resistance" audience, which is a purely political position because the fall of the Syrian regime

For clarity on the topic of the 25 million: See Dr. Thaqal Al-Ajami's opinion published in Saber <http://t.co/8wbhRHI>

All heads of government cut off their vacations in any political crisis. The name of the president does not allow him to leave for him except for crises and even with his interrogations, as if he is not concerned with them !?

It is not the first time that the prime minister leaves on a special vacation, in light of a crisis that requires a clear government position, so is the crisis unworthy? Or are the people unworthy !?

It is not acceptable for the Gulf position to be limited to the brutal crimes of genocide in Syria that are limited to a timid demand for reforms and bloodshed! Do they not know the reason for this?



The Syrian regime is targeting civilian sites indiscriminately, and the continued Arab silence regarding these criminal practices gives the Zionist entity an excuse for any action in Gaza.

The government's slowdown in dealing with a crisis of this scale has no single justification: the search for a foreign outlet to cover up a crime in which the government colluded and the periphery of corruption

4- Such a government will adhere to it, and what interests of the people will it protect !!?

3- Also a government that publicly pays cash bribes with “bags”, “checks”, “transfers” and in-kind “appointments”, positions, committees, and “treatment trips” that are borne by public money, what law

2- How can this crime swamp .. express the will of the nation !!?

1- A strange political reality: a parliament whose members arrived at the "falsification" of the popular will, and a number of its members work "for bribery" and some of its members provide "false statements" at the expense of the public order

2- I was happy that my home was one of the congratulatory stations that I shared with the brothers in Libya the joy of being saved from the Gaddafi regime, wishing that we would also be happy with the Syrian people.

1- Professors Adam Rakik, the editor-in-chief of Libya Al-Youm newspaper, and Mr. Hussein Bouhdma, and their host, Mr. Faleh Al-Mutairi, have just left to congratulate the victory of the Libyan people

8- It is an envisaged matter within the scope of any freedom of public freedoms, and the former Attorney General had previously stated this to Al-Qabas newspaper, as well as recent judicial rulings referred to it,

7 - The controls and scope of criminalization, and it specifies the considered and legitimate evidence for that. Personal freedoms may not be wasted solely because of the possibility of offense or delinquency.

6- If the government is convinced that the irresponsible use of electronic means of communication constitutes a danger, then there is nothing to prevent it from adopting a draft law defining it.

5- And specific because that is what makes it a control of behavior and a measure of TBHis ummah or his mistake, and to say otherwise, they are "security" explanations, not considered legal interpretations.

4- It is considered by its nature a restriction on personal freedom, and therefore any interpretation of it must not deviate from the basic assumptions of criminalization, which is that the criminalization is explicit and clear.

3- A general principle that was established as a basic principle of law as long as its applications exist in all the Latin and Anglo-American systems alike, so the criminalization text

2- The penal text does not elaborate the scope of its applicability, just as analogy to it is not permissible at all, and therefore the directions of the Attorney General and even rulings otherwise are not nullified

1- The criminalization of electronic publishing, considering its link in the public place, and the general rules of criminalization of insulting and defamation contradict a general principle of law, which is that

We will wait for the end of the Ramadan cycle for the rai and ask for a meeting immediately after Ramadan. My personal assessment is that direct meetings are more convincing to the viewer than a written one.

I was just in a phone call with an official at the Al-Rai Foundation and he informed me that Al-Rai apologized for the interview due to the Ramadan channel's plan and their agreement to publish the opinion in the newspaper.

7- In general, I informed Brother Jassim Bodi of my desire to fully explain the matter and respond to the statements of government agencies in this regard, and awaiting approval or apology for the opinion.

6- Very briefly, there is no excuse for any party that was factual and legal, and any words to the contrary have no legal basis, and I am willing to prove this publicly and in front of any party

5- Or that it has notified the Central Bank and the Central Bank that it has failed to inform the competent authority, which constitutes a crime, and the Public Prosecution has contacted the actual incident.

4- Information: Banks are legally mandated to report "any suspicion" in "any bank account" and this means that banks have committed a violation of the law by not reporting

3- As for the government and its statements, they reminded me of the Jahra fire, the committee that was formed at the time, and its reports that no one knows about. People are required to always trust the government.

2- As for the statements of the Central Bank Governor, it is stated that he does not exercise any supervision over banks except in his press conferences, and that his understanding of the meaning of banking supervision is non-existent.

Al-Qabas's assertions regarding the million-dollar bribes indicate that the bank's procedures are wrong and the Central Bank does not exercise oversight and obliges the Attorney General to summon the editor to hear his statements.

6- And a peg for not carrying out the inevitable political reforms required by the circumstances of the stage and the aspirations of the peoples to participate and make decisions and to perpetuate the concepts of justice and law

5- How can you support a regime resisting tyranny and suppress the civil revolution of its people calling for reform and change !, but this does not mean that the Gulf states use Iran as justification

4- Iran's support for Hamas is not based on a belief in supporting the resistance, but rather on its use as a tool of pressure. The evidence is that the position towards Hamas has changed due to its opposition to the suppression of the Syrian regime.

3- Whoever helps a criminal to commit his crime, verbally, or silently, then he is a partner for him in what is done with the innocent, and this matter applies to individuals as it applies to states as well

2 - Exorbitant costs, and the evidence is that the "Damascus criminal" did not stop peaceful civilian demands, even by using tanks and barges, killing and intimidation, and his end would be like his type of actions.

1- One of the inevitable results is that "any political system" does not respond to the requirements of real democratic transformation and participation, and it will only survive with the protection of tanks, and temporarily, which is an option.

2- While the rest of the Gulf states sided with the oppressive regimes and supported them until the end, the regimes collapsed and the people triumphed, "and every soul has what it won as a hostage."

1- Qatar succeeded in standing by the peoples and gained the respect of everyone. Al Jazeera succeeded in being the official spokesman for the revolutions, so it participated in them. Congratulations to Qatar and the island.

What is happening in the ocean is a free lesson, that freedom and dignity are not deserving of anything but one who is willing to sacrifice for it. Great values are not given to cowards.

, What is happening in the ocean is a lesson for cowards, that freedom and dignity are not deserving of anyone but he who is willing to sacrifice for it. Great values are not given for free.

2011 is a great year, in which the cell was cleared for Hosni, Ben Ali and Gaddafi to go down to it, and I think it still accommodates more, what a year it is and I hope it never ends

Another historic day in the life of the nation, and the sun of Libya will rise tomorrow without tyranny, without Gaddafi and his gang, and we will soon celebrate the victory of the Syrian will. Oh God, amen.

2- As for otherwise, it means either that the government is really disrespectful, or that it believes that the people do not deserve respect in the first place, which is more likely.

When an issue of this seriousness is raised, the simplest procedure of any respectable government is for an official to present in it the government's viewpoint on the crisis and the measures it will take.

The Central Bank, as usual, will not ask about anything. The banks in Kuwait are the most honest in the region, and the representatives are all honest, and the accused is definitely the "ATM machine."

2- In his presence, passports became worthless documents and borders became an optional issue, as he contributed to an Arab rapprochement that its organizations and meetings failed to achieve.

1- Having a program like Twitter is a necessity for the political security of the Arab peoples, as it is a place where governments cannot determine what people say or what they should hear.

4- As for the Gulf, the victory of the popular will by itself is considered a threat to its political security, because it will become the only system in the world that does not participate in the people's decision.

3- Due to considerations related to national security because it will not find better security than this neighborhood, and Iran considers the Syrian regime a strategic ally and a mouthpiece in the Arab world.

2- One of the strange paradoxes is that the governments of Israel, Iran and the Gulf stand with the Syrian regime despite their different starting points in this regard: Israel supports the Syrian regime



The Israeli raid on the Egyptian border and the killing of Egyptian soldiers is a clear example of the State of Israel's keenness on the presence and support of the Syrian regime and exposes the lie of resistance

Bashar Al-Assad has to make the most important decision in his life by resorting to Russia, as his end has become confined between two things: an international trial whose outcome is known or the distribution of his body parts on the streets of Damascus

The formal government and its deputies still consider the excesses Constitutional and regulatory a political tactic! How can individuals be convinced of the slogans of respecting the law !?

3- The most dangerous matter: We have a council that does not monitor and most of its members do not want that, and a formal government that does not work and is unable to function at all. Is the state marching by the jinn, God forbid

2- The other matter: If the structural deficiencies were caused by executive decisions and government financial waste, then how can it be accepted for the government to make mistakes and hold people accountable for that !!?

1- If every problem is formed for him a committee to present recommendations and then another committee emerges from it, then the question is what are the tasks of the Council of Ministers in the first place? From a constitutional and realistic point of view?

7 - Most importantly, our criticism of the Gulf stances does not necessarily mean forgetting that the Iranian regime is a repressive and inhumane regime and its positions are based on purely sectarian foundations.

6- Whoever refers to the term rabbinic scholars, I say to him: Exactly who are these? What is their position on killing a Muslim, demolishing a mosque, and preventing prayer? And what did they say about that? Where and when?

5- Does the accusation still exist? Are these people of sedition and heresy !? The most important question is who are the specific scholars? And what do they think of the events !?

4- When some of them described those who demanded people to protest and demonstrate as "advocates of sedition and heresy," but after the support of Sheikh Abdul-Rahman Abdul-Khaleq, Ajil Al-Nashmi and others

3- Decisions were issued in most of the Gulf countries to prevent mosques from being exposed to political events, and the "scholars" did not object to that! Is it an endorsement and acceptance of separating religion from politics !?

2- Fatwas were issued that were marketed regionally to prohibit demonstrations, then the authority allowed it, and the fatwas were silent! The question is, has it been withdrawn? Or is it still usable?

1- Gulf governments allowed their people to demonstrate in support of the demands of the Syrian people and objecting to suppressing their authority, but why did they prevent their people from demonstrating to demand their rights !?

2- As long as his reception was for humanitarian reasons, his residence should continue on this basis until he leaves or seeks political asylum, and he must be alerted to that.

Saudi Arabia received Saleh for humanitarian reasons, as he stated, but for him to use Saudi Arabia as a platform to launch political statements for the Yemeni people, it is totally inappropriate.

Now listen to the speech of the Shawish Saleh, and you have all heard the speeches of Gaddafi, but the logical question is: How can these strange living creatures rule a people !?

The truth is that I find no understandable reason to maintain these relations in light of this tragic situation other than to give the Syrian regime an opportunity for more killing, brutality and extermination.

The Syrian people are still being slaughtered without mercy, and the Gulf continues to issue statements that are not understood in public and to apologize for them in secret, and relations with the murderous regime still exist.

Shouldn't the size of the private sector's contribution to the national product be known first ?! So that he is given the authority to contribute to the economic decision-making in the first place !?

The government that has extemporaneous and extemporaneous solutions must bear the consequences of its wrong decisions, but for the government to make a mistake and the citizen to bear the consequences of that, this is unacceptable.

I wished that the expenses of the "show-offs" had been donated by their owners to the hungry Somalia and the displaced in Syria, instead of spending huge sums of money on satiated people, some of whom came to watch

4- Therefore, the prevalence of authoritarian injustice, the popular certainty of its existence, and the denial of authority to it creates fertile grounds for rejecting reality and working to change it, and from here the revolution begins

3- Peoples revolted because the circumstances put them in front of options equal to death and life, so they tried this new pattern simply because they had nothing to lose.

2 - The highest degree of political corruption is that the corrupt does not lie, for his informer takes on the task and does not kill with his hand. Security is present and people do not rob their rights, so the judiciary plays the role for that.

1- The Arab Spring revolutions not only exposed the corrupt regimes, but rather the corruption of their tools of oppression and injustice, and every time, corrupt security, corrupt judiciary, corrupt media, is there any preacher?

Bashar enters history as the first head of state to openly declare war on the defenseless division and bomb Lattakia by land and sea, and solidarity now means submitting a request to try the pillars of the regime

Any political reform initiative is not based on clear rules of accountability and accountability, as it cannot be considered politically as reform, whatever its name.

5- Everyone is taken from his words and responded, as for intimidating people and intimidating them with a legitimate cover, they are security policies and not inherently legitimate policies, and no one will prevent us from saying the truth, whatever it is.

4- Religion came with dignity, freedom and equality, and any practice that contradicts these values is not originally from religion and should be rejected, just as Islam does not define priesthood.

3- Wouldn't it be better for me to instead be preoccupied with accusing me of sedition and calling out to Bashar, denying him demolishing mosques, killing innocent people and preventing him from praying! What concern for this religion?

2- And for those who have accused us of inciting sedition and disobeying the ruler as well: See the fatwas of Sheikh Al-Zindani, Sheikh Al-Qaradawi, Al-Azhar scholars, Sheikh Al-Awda, and others.

1- And for those who have accused us of inciting sedition and disobeying the ruler: also the Sheikh of Al-Azhar, "It is not permissible according to Sharia to remain silent about the tragedy that the Syrian people are exposed to."

For those who have accused us of inciting sedition and disobeying the ruler of Syria as it is not legally permissible. Association of Muslim Scholars

<http://t.co/gFKZQkW> <http://t.co/EIYEskf>

3- Therefore, climate disasters, paper market crashes, and hard and easy debts!

The excesses of banks and investment companies and the tampering of tenders are only diligence

2- As for what anyone does, no matter how great, it is patriotic unless they decide that it is. It is strange that the authority approves and demands ratification.

1- The problem of the "economic elite" in Kuwait is that they want to monopolize everything, economy, politics, sports, poetry and literature, and they are the owners of national history and sacrifices.

4- Do not activate except in rotten puddles, so I have a duty to stand behind everything that is rotten because creating a clean environment will kill this activity.

3- billion in Syria, Egypt (before the revolution) and Lebanon Although the markets of Europe, America and Scandinavia are open, they are simply bacterial investments

2- Therefore, Kuwaiti capital has opposed all revolutions for the same reason, because the so-called private sectorIt is only active in a corrupt political environment, and investments are often justified

Silence about what is happening in Syria from some is not an acceptance of it, but it is the fear of the regime's reaction to investments created by corruption, grown by corruption, and will be uprooted by the deprived.

4- Our sectarianism and sectarianism because we are not used to being preoccupied with producing medicine, providing food, or competing with leading civilians, which are the primary requirements for the existence of any civil building.

3- It is for states, and civil construction and development is limited to concrete buildings and urban formalities without real building for the human being, so it is considered natural for us to be preoccupied.

2- The ability of societies to participate effectively, control and accountability is weakened, which actually happened until the region's societies turned into entities closer to private ownership.

1- What happens in Kuwait and other blatant discrimination is not a real tendency for one group without another, but it is a policy developed by the British and followed by governments to keep people busy with each other.

The law is general and abstract rules that address the behavior of the individual, not his title or belief, but the acts of authority are what makes discrimination with discrimination and selectivity on these illegal grounds.



4- To reject the law and not to obey it because accepting it according to these data means recognizing it as a tool for extortion, which is an illegal purpose and a picture of deviation in its use

3- Those in charge of it are known only by groups without others, and the Ministry of Interior and those who agree with it in this approach must know that selectivity in the application of the law is a legitimate justification.

2- On the authority of the owner of the channel who said, "You see, we are melancholy and you are ....., and on the authority that withheld the presentation of my orders, a site and written on the authority of its jurisdiction, it is strange for the Ministry of Interior."

1- Where is the Ministry of Interior and State Security, on the authority of the deputy who said that the thousand grant is a wrong and hasty decision! And about the writer who said, "The greed is for the one who pardoned them."

We will not compliment anyone in the law, and the truth that no one can utter will be rendered void, and whoever is silent by the devil. As for the accusations of the interior, they are like declarations and conferences of "the sidewalk."

5- And your suspicious silence about his "freedom in life", what is the value of "jeans" originally for a dead person !? Freedom is a much deeper concept than being reduced to clothes and bids !!?

4- Therefore, it is very logical to ask about the position of the parliamentarians and the liberal elites as they marketed to themselves how to accept your permanent defense of the freedom of clothing for individuals

3- The bottom line is that either the understanding of freedom was fundamentally wrong, or it is just a slogan used for election marketing without being a culture that is translated into a position.

2- But it is completely incomprehensible that he opposes preventing a woman from working in the evening, considering that an assault on a general freedom that must be defended, and then she is absolutely silent about killing that takes place in the morning and evening.

1 - Regardless of what is said in the festivals of solidarity with Syria, the Islamists have proven their position, regardless of their motives, but in the end, the content coincides with the concept of freedom.

The Yemeni opposition's demand for Ali Saleh to be tried in Yemen is a legitimate and deserved claim, and considerations of justice and the requirements for change require that no one escapes with punishment.

Terrorist groups, traitors, agents, rats, Al Qaeda cells, Wahhabis, Safavids. These are some of the Arab governments' descriptions of those who opposed them, and in return they are just and respectful governments.

We hope that the participants in the sit-in will stay away from sectarian narratives because the machine of murder and brutality did not differentiate them on this basis, and it is not in the interest of their revolution.

Assad again challenges the world and is determined to fight the armed terrorist groups that are killing Syrians !!! Is stupid lie an invention of Arab governments !?

Oh God, make me with the truth wherever it is, and that I do not accuse anyone of evidence and I will not respond to anyone without an excuse, O God, make me one of the counselors. And peace

10- Because the thing reminds the advocates of freedom in Kuwait who made statements about the right of women to wear swimwear and remained silent about the right of the Syrian people to life

9 - It is more appropriate for him not to be a scholar, so do not invoke the Sunnah and wear something that is not of it, because the Sunnah of the Messenger (PBUH) is purified from the falsehood of the hypocrites and official security fatwas

8- A Muslim who does not suffer because of the suffering of innocent people who are killed without guilt or mercy and an old woman for help, a child crying out for fear, and a city in which prayer is forbidden by an official decision that is not a human being at all

7 - Last but not least: What is the Islamic ruling in the Gulf governments' decision to prevent preachers from touching on the Syrian issue, permissible or forbidden? And if it is forbidden, why did they not deny it?

6- We heard their opinion about demonstrating and the necessity of obedience to the guardian, but we did not hear the opinion of the muftis about the permissibility of imprisoning a person like him? Or imprisoning him without charge? Or banning preachers?

5- The answer to the question of why the sheikhs specifically? The answer is simple: that the acts of repression carried out by the authority are based on the opinions of some and their supporters from the others

4- The clerics are human beings who are not infallible, nor are they impervious to error. “Every son of Adam is wrong.” And whoever practices a public activity is exposed to criticism, then they are not the most generous of the rightly-guided caliphs.

3- And the information is lethargic in every field, there are lethargic politicians, lethargic clerics and law professors as well, and hypocrites mentioned in the Qur’an.

2- If the protest is forbidden, and it is an obviously disputed issue, then is killing a soul without guilt, indecent assault, forbidding prayer, or demolishing a mosque, is permissible !?

I do not know the reason for the controversy surrounding fatwas prohibiting demonstrations now: But it is certain that its effect distracted attention from the basic matter, which is definitively the sanctity of killing a soul, preventing prayer and demolishing a mosque

5- We are waiting for the extermination of the entire people !!?

4- Approximately 2000 martyrs and more than 20000 detainees surrounded by tanks, forbidding prayer and killing innocent people. Should we not exaggerate our reaction and when should ambassadors be expelled? And does

3- Isolating the regime internationally by submitting a clear request to prosecute regime members to encourage them to disavow it and give ambassadors the choice between expelling them or resigning them and granting them political asylum

2- It will also not contribute to preventing killing and displacement, and what is required is serious attitudes that reflect the truth of solidarity with a just cause, the first of which is preventing aid of all kinds and work.

1 - Exaggeration in flattery of the modest and late Gulf reaction is not a welcome matter, as summoning an ambassador is a protocol procedure that takes place in what is far below that.

RT @Alqudaimi: Our government, may God guide it, sometimes gives up some sheikhs and liberals in the position of a Sheikh... Now, what do they do fatwas and articles criminalizing misrepresentation ...

2- It is well known that silence in the event of the need to speak is a statement and it is a fundamental rule, so where were those who prevented Friday prayers in Hama by an official announced decision, who prevented them?

1- Every scholar who has kept silent about killing a Muslim, defilement of honor, demolishing a mosque and preventing prayer without having an opinion has no value for his knowledge and the phrase applies to everyone who meets the conditions

2- It is not fair or just to equate the position of the scholar, Dr. Al-Nashmi, with the position of the Mufti of interest.

All that you mentioned the term of the Sultans' scholars, some of the description were disturbed So is every cleric a scholar of authority !? It is strange that some of them become irritated, as if to say .. Why are you insulting me !?

Even if it was a late step: the Gulf states should submit a request to refer the criminals of the Syrian regime to the International Criminal Court as the least that can be offered to their people and to expel its ambassadors.

Naturally, this talk does not apply only to one group, namely the gentlemen advocates of the security services, because their jurisprudence is linked to the approval of the security authority in advance.

Although the Kingdom's announcement of its position was long overdue, one of its non-political direct results is that, starting tomorrow, it opened the door to sermons in support of Syria and fatwas of jihad and atonement.

The inevitable result of this speech has put the government in trouble, and most likely the Syrian ambassador will be summoned tomorrow on "Ghabga."

The Pope of the Vatican denounces the Syrian oppression, and Angelina Jolie leads a campaign for donations for the Syrians and May Skaf protests in Damascus: What do we have left! Talking about the etiquette of fasting !?,

When al-Assad visited Kuwait, newspaper pages were filled with welcoming announcements in which he paid hundreds of thousands. Did advertisers donate tenths of these sums to Syrian refugees?

One of the painful facts is that private sector investments are only active in every corrupt environment, and therefore the national silence about Assad's crimes is primarily for economic reasons.

The devil's advocate was silent to kill innocents, demolish mosques and prevent prayer, and suddenly he spoke to express the Sharia's opinion on the "draft" fatwa on wasting the ambassador's blood! God abuses those like you, O Medals

As for the recommendation to urge a solution to the problem of student admission, it seems that the prime minister does not know the difference between the cabinet, which is the highest executive body, and the parents 'council.

Tanks storm Deir Ezzor and kill 50 civilians in the storming, and the Kuwaiti Cabinet maintains diplomatic representation and calls for activating dialogue! What a great position



Wasting fatwas are not a culture for anyone without anyone, and “it is unacceptable to me,” but a fatwa was already issued against Salman Rushdie's blood, so would the Qallaf dare to denounce her !?

It is strange that the "disrespectful" statement was issued in the wake of the Syrian regime's decision to ban Friday prayers in Hama and the bombing of mosques: There is no strength and no power except in God!

The statement of cooperation on Syria are expressions that do not carry any implication or significance, as it suggests that the Syrian people are a partner in the bloodshed. In general, stop your support for the regime and thank God for your efforts.

6- I participated in the embassy gathering because I am convinced of the usefulness of the symbolism of the gathering there, because the message will be clearer. As for the conversation, it is not always important. I have nothing more than what I say here.

5- The ambassador is an invitation that is rejected in form and substance and should not be taken as a justification to reduce all the positive aspects of the movement and an explanation for the failure of others to say the word of truth and support the oppressed,

4- Whoever is victorious for Syria and its people must not weaken its revolution by giving it a sectarian dimension, because that is what the murderous regime promotes and what it wants: As for the referendum of blood permitting

3 - The silence that kills them, and I am sure that Kuwait, which was silent yesterday, will not be silent now about what is happening in terms of killing, oppression and oppression, and Kuwait, the people and the government, will be a voice for truth

2- The imposition of a sectarian character by some on the objection does not reflect the reality of the Syrian oppression, which does not distinguish between a religion or a sect, but at the very least we broke a part of

1- The explicit message of solidarity that the Kuwaiti people sent to the people of Syria, the "slaughtered prisoner", cannot be reduced to an impulsive expression. It is certain that whoever said it does not realize its dimensions.

2- I think that the Syrian embassy is the appropriate place to protest and object to the regime's actions in Syria, unless the purpose is merely to establish a position of discharge.

1 - I do not have the authority to direct people to protest in a specific way, but my belief is that the Will Square is not the appropriate place to send a protest message of this kind for that

4- Denying rights and brutality is a crime in all religions, boredom and laws. It is not humane to be happy with wronged pain or innocent death because it disagrees with you in opinion or belief.

3- When we deal with humanity as a value, we can weigh things fairly, while selective attitudes reflect shortcomings in the human sense and the killing of innocent people.

2- Therefore, my opinion about Syria came in line with what I mentioned in Bahrain and Yemen, and it is a principled position that is not affected by the very sectarian sphere of opinion. As for the positions of others, they are asked about it.

I am aligned with the legitimate demands of any people, and therefore my opinion was clear and frank about Bahrain, supporting their right to legitimate claims and denouncing all forms of brutality and oppression.

2- In Will Square, it is a clearance that is not commensurate with the scale of the crime and its ugliness. Moreover, it is customary for the security presence in it to be greater than the number of participants.

1- Whoever wants to support the Syrian people, he must go to the place where the message reaches to the influential Assad, the Syrian Embassy or the Ministry of Foreign Affairs, with specific demands, either the gathering

The currents, writers and institutions that spoke about the Syrian issue after the Security Council's condemnation and the statement of the US State Department: So did they pay attention to the situation now, or did they take permission?

2- In Syria, children and women are killed during the month of Ramadan, mosques are demolished, and you consider mere exposure to the subject a trial! What temptation is greater than deceiving you on religion in Ramadan?

1- A final message to the sheikhs of the security apparatus in the events of September 11th, when the platforms were mobilized to deny the crime.

3- Considering that the liquidations that happen in Syria constitute a crime of genocide, a crime that gives the right to any member of the United Nations to request that it be presented to the body that has jurisdiction over it.

2- Working on crystallizing a project that is presented and presented to the Security Council The Gulf Cooperation Council countries should request the activation of the Rome Statute and prosecute the perpetrators of the killing of Syrian civilians

1- First: the representative of the killer of children and women should not stay in Kuwait for an additional day, second: the government's statement regarding Kuwait's official position must be clear and explicit.

In order for the festival of solidarity with the Syrian people not to turn into a show festival that does not carry a vision or vision that alleviates the suffering of the Syrian people, the demands must be clear and specific.

2- For this reason, the killing of individuals or the denial of their legitimate rights was not a sufficient motive to interrogate any Arab regime, as murder and denial of rights are methods used by all authorities

1 - The Arab revolutions have proven the bias of power to the authority and the people towards the peoples, and to understand this equation you must know that the Arab authorities were not in any country a part of the people

The fall of the Assad regime has become inevitable, but the timing of this fall and the mechanism of accountability for the pillars of this regime is the matter that the Syrian people will determine soon #Syria

Some mourned the mere trial of Hosni Mubarak and considered it an insult to him despite the clear guarantees, but they forgot the insult of eighty million people and a thousand years of Egypt's history.

2- And that the insults and pain are not forgotten, no matter how long they last

1- The symbolism of the trial of the former Egyptian president, his appearance and the pillars of his government before the judiciary of his country, an outstanding message that translates the mobility of the Arab street and its aspiration for freedom and dignity

A Syrian citizen's call to Al-Jazeera stated, "When mosques are demolished, women are widowed, and symptoms are violated in Ramadan, the Muslim fasting becomes like dogs starvation."

5- Knowledge that is not translated by action or position is not science and the world that does not deny this crime by word or deed is not a scientist, but whoever uses religion to justify criminality is a criminal

4- Is it a religion to kill people without sin? What is the ruling on killing a Muslim in the mosque? And in Ramadan? What is the ruling on bombing mosques? Question to those whom the authority called the nation's scholars

3- What is happening in Syria now is a crime of genocide and silence about what is happening is collusion to kill an unarmed people whose government did not take into account the sanctity of Ramadan or the sanctity of mosques

2- Are Bashar and his regime still governors of a matter that it is not permissible to go out or even pray for them? Or is the security directive the direct source for spreading fatwas that deceive religion!

1 - After tanks bombed the Al-Hamidiyeh Mosque and killed the worshipers in it, I do not know what new justifications are for fraudulent bodies and advocates of demons and the imams of the Sultans, and the question

## REFERENCES

- [1] S. Boukil, M. Biniz, F. El Adnani, L. Cherrat and A. E. El Moutaouakkil, "Arabic text classification using deep learning technics," *International Journal of Grid and Distributed Computing*, vol. 11, no. 9, pp. 103-114, 2018.
- [2] W. Ao, "Siri and her AI friends are struggling with Arabic," Medium, 7 November 2020. [Online]. Available: <https://medium.com/1001-modern-nights/siri-and-her-ai-friends-are-struggling-with-arabic-93ab6f42b71d>. [Accessed 17 February 2021].
- [3] A. Mourad and K. Darwish, "Subjectivity and sentiment analysis of modern standard Arabic and Arabic microblogs," in *Proceedings of the 4th workshop on computational approaches to subjectivity, sentiment and social media analysis*, Atlanta, GA, USA, 2013.
- [4] K. Shaalan, S. Siddiqi, M. Alkhatib and A. A. Monem, "Challenges in Arabic Natural Language Processing," in *Computational Linguistics, Speech and Image Processing for Arabic Language*, World Scientific, 2018, pp. 59-83.
- [5] F. Sadat, F. Kazemi and A. Farzindar, "Automatic identification of arabic dialects in social media," in *Proceedings of the first international workshop on social media retrieval and analysis*, Gold Coast, Queensland, Australia, 2014.
- [6] P. Dasigi and M. Diab, "CODACT: Towards Identifying Orthographic Variants in Dialectal Arabic," in *Proceedings of the 5th International Joint Conference on Natural Language Processing*, Chaing Mai, Thailand, 2011.



- [7] C. C. Aggarwal and C. Zhai, "A survey of text classification algorithms," in *Mining Text Data*, Boston, MA, Springer, 2012, pp. 163-164.
- [8] W. Black, S. Elkateb, H. Rodriguez, M. Alkhalifa, P. Vossen, A. Pease, M. Bertran and C. Felbaum, "The Arabic Wordnet Project," in *Proceedings of the Fifth International Conference on Language Resources and Evaluation*, Genoa, Italy, 2006.
- [9] L. Abouenour, K. Bouzoubaa and P. Rosso, "On the evaluation and improvement of Arabic WordNet coverage and usability," *Language Resources and Evaluation*, vol. 47, no. 3, pp. 891-917, 2013.
- [10] S. R. El-Beltagy and A. Ali, "Open issues in the sentiment analysis of arabic social media: A case study," in *Ninth international conference on innovations in information technology*, 2013.
- [11] W. Zaghouani, "Critical survey of the freely available arabic corpora," in *Proceedings of the workshop on free/open-source arabic corpora and corpora processing tools workshop programme*, 2014.
- [12] M. Graja, M. Jaoua and L. Hadrich Belguith, "Lexical study of a spoken dialogue corpus in tunisian dialect," in *The international arab conference on information technology*, Benghazi, Libya, 2010.
- [13] K. Almeman and M. Lee, "Automatic building of arabic multi dialect text corpora by bootstrapping dialect words," in *Communications, signal processing, and their applications (iccspa)*, 2013.

- [14] A. AlFarasani, T. AlHarthi and S. AlHumoud, "ATAM: Arabic Traffic Analysis Model for Twitter," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 3, pp. 328-336, 2019.
- [15] I. Hmeidi, M. Al-Ayyoub, N. Abdlla, A. A. Almodawar, R. Abooraig and N. A. Mahyoub, "Automatic Arabic text categorization: A comprehensive comparative study," *Journal of Information Science*, vol. 41, no. 1, pp. 114-124, 2015.
- [16] M. S. Khorsheed and A. O. Al-Thubaity, "Comparative evaluation of text classification techniques using a large diverse Arabic dataset," *Language Resources and Evaluation*, vol. 47, no. 2, pp. 513-538, 2013.
- [17] S. Alowaidi, M. Saleh and O. Abulnaja, "Semantic sentiment analysis of Arabic texts," *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 2, pp. 256-262, 2017.
- [18] S. Yousif, I. Elkabani, V. Samawi and R. Zantout, "Enhancement of Arabic text classification using semantic relations of Arabic WordNet," *Journal of Computer Science*, vol. 11, no. 3, pp. 498-509, 2015.
- [19] S. Elkateb, W. Black, H. Rodriguez, M. Alkhalifa and P. Vossen, "Building a WordNet for Arabic," in *Proceedings of the 5th International Conference on Language Resources and Evaluation*, Genoa, Italy, 2006.
- [20] S. A. Yousif, I. Elkabani, V. W. Samawi and R. Zantout, "The effect of combining different semantic relations on Arabic text classification," *World of Computer Science and Information Technology Journal*, vol. 5, no. 6, pp. 112-118, 2015.

- [21] V. Cavalli-Sforza, H. Saddiki, K. Bouzoubaa, L. Abouenour, M. Maamouri and E. Goshey, "Bootstrapping a WordNet for an Arabic dialect from other WordNets and dictionary resources," in *10th ACS/IEEE International Conference on Computer Systems and Applications*, Fes/Ifrane, Morocco, 2013.
- [22] T. Khalil, A. Halaby, M. Hammad and S. R. El Beltagy, "Which Configuration Works Best? An Experimental Study on Supervised Arabic Twitter Sentiment Analysis," in *The First Conference on Arabic Computational Linguistics*, Cairo, Egypt, 2015.
- [23] F. S. Al-Anzi and D. AbuZeina, "Toward an enhanced Arabic text classification using cosine similarity and Latent Semantic Indexing," *Journal of King Saud University - Computer and Information Sciences*, vol. 29, no. 2, pp. 189-195, 2017.
- [24] S. Tartir and I. Abdul-Nabi, "Semantic sentiment analysis in Arabic Social Media," *Journal of King Saud University - Computer and Information Sciences*, vol. 29, no. 2, pp. 229-233, 2017.
- [25] F. Harrag and E. Al-Qawasmah, "Improving Arabic text categorization using neural network with SVD," *Journal of Digital Information Management*, vol. 8, no. 4, pp. 233-239, 2010.
- [26] F. Harrag, E. El-Qawasmah and A. M. S. Al-Salman, "Stemming as a feature reduction technique for Arabic text categorization," in *10th International Symposium on Programming and Systems*, Algiers, Algeria, 2011.
- [27] A. Elnagar, R. AlDebsi and O. Einea, "Arabic text classification using deep learning models," *Information Processing & Management*, vol. 57, no. 1, 2020.

- [28] R. Duwairi, R. Marji, N. Sha'ban and S. Rushaidat, "Sentiment analysis in arabic tweets," in *Information and communication systems*, 2014.
- [29] H. S. Ibrahim, S. M. Abdou and M. Gheith, *Sentiment analysis for modern standard arabic and colloquial*, arXiv preprint arXiv:1505:03105, 2015.
- [30] K. Darwish and W. Gao, "Simple effective microblog named entity recognition: Arabic as an example," in *Proceedings of the Language Resources and Evaluation Conference*, Reykjavik, Iceland, 2014.
- [31] A. Zirkly and M. Diab, "Named entity recognition for dialectal Arabic," in *Proceedings the the Arabic Natural Language Processing Workshop*, Doha, Qatar, 2014.
- [32] A. Zirikly and M. Diab, "Named entity recognition for Arabic social media," in *Proceedings of the North American Chapter of the Association for Computational Linguistics – Human Language Technologies*, Denver, Colorado, USA, 2015.
- [33] A. R. Hedar and M. Doss, "Mining social networks Arabic slang comments," in *IEEE Symposium on Computational Intelligence and Data Mining*, Singapore, 2013.
- [34] N. El-Fishawy, A. Hamouda, G. M. Attiya and M. Atef, "Arabic summarization in Twitter social network," *Ain Shams Engineering Journal*, vol. 5, no. 2, pp. 411-420, 2014.
- [35] A. Pasha, M. Al-Badrashiny, M. Altantawy, N. Habash, M. Pooleery, O. Rambow, R. M. Roth and M. Diab, "Dialectal Arabic information retrieval assistant," in *The*

*companion volume of the proceedings of international joint conference on natural language processing*, Nagoya, Japan, 2013.

- [36] J. Roesslein, "Tweepy: Twitter for Python!," 2020. [Online]. Available: <http://www.tweepy.org>. [Accessed 10 February 2021].
- [37] G. Song, Y. Ye, X. Du, X. Huang and S. Bie, "Short text classification: a survey," *Journal of Multimedia*, vol. 9, no. 5, pp. 635-643, 2014.
- [38] R. Rehurek and P. Sojka, "Software framework for topic modelling with large corpora," in *Proceedings of the 2010 LREC Workshop on New Challenges for NLP Frameworks*, Valletta, Malta, 2010.
- [39] G. Salton and C. Buckley, "Term weighting approaches in automatic text retrieval," *Information processing and management*, vol. 24, no. 5, pp. 513-523, 1988.
- [40] A. Kent, M. M. Berry, F. U. J. Luehrs and J. W. Perry, "Machine literature searching VIII. Operational criteria for designing information retrieval systems," *American Documentation*, vol. 6, pp. 93-101, 1955.
- [41] S. Deerwester, S. Dumais, G. W. Furnas, T. K. Landauer and R. Harshman, "Indexing by latent semantic analysis," *Journal of the American Society for Information Science*, vol. 41, no. 6, pp. 391-407, 1990.
- [42] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62-66, 1979.
- [43] H. Ng., "Automatic thresholding for defect detection," *Pattern Recognition Letters*, vol. 27, no. 14, pp. 1644-1649, 2006.

- [44] I. Goodfellow, Y. Bengio and A. Courville, "Deep feedforward networks," in *Deep Learning*, Cambridge, Massachusetts, The MIT Press, 2016, pp. 197-206.
- [45] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison and A. a. Y. E. a. D. Z. Kopf, "PyTorch: An imperative style, high-performance deep learning library," in *Advances in Neural Information Processing Systems 32*, Curran Associates, Inc., 2019, pp. 8024-8035.
- [46] C. R. Harris, K. J. Millman, S. J. v. d. Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. v. Kerkwijk, M. Brett and A, "Array programming with NumPy," *Nature*, vol. 585, no. 7825, pp. 357-362, 2020.
- [47] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman and May, "SciPy 1.0: Fundamental algorithms for scientific computing in Python," *Nature Methods*, vol. 17, pp. 261-272, 2020.
- [48] S. El Beltagy and A. A. Ali, "Open issues in the sentiment analysis of Arabic social media: A case study," in *The 9th International Conference on Innovations and Information Technology*, Al Ain, United Arab Emirates, 2013.
- [49] A. Farghaly and K. Shaalan, "Arabic natural language processing:Challenges and solutions," *ACM Transactions on Asian Language Information Processing*, vol. 8, no. 4, p. 22, 2009.