

**KNOWLEDGE DISCOVERY IN AL-HADITH USING TEXT
CLASSIFICATION ALGORITHM**

By

Khitam Mahmoud Abdalla Jbara

Supervisor

Dr. Azzam T. Sleit

Co-Supervisor

Dr. Bassam H. Hammo

**This Thesis was Submitted in Partial Fulfillment of the Requirements
for the Master's Degree of Computer Science**

**Faculty of Graduate Studies
The University of Jordan**

May, 2009

This Thesis (knowledge discovery in AL-Hadith using text classification algorithm) was successfully defended and approved on May 5, 2009.

Examination Committee

Signature

Dr. Azzam T. Sleit, Supervisor
Associate Prof. of Imaging Database

Dr. Bassam H. Hammo, Co-Supervisor
Assistant Prof. of Natural Language Processing

Dr. Ammar Huneiti, Member
Assistant Prof. Intelligent information systems

Dr. Mohammad S. Qataweh, Member
Associate Prof. of Parallel Systems and Networks

Dr. Bayan A. Abu Shawar, External
Assistant Prof. of Natural Language Processing
Faculty of Computer Studies, Arab Open University

DEDICATION

This work is dedicated to my beloved father my idol , for my mother soul who learned me how to be strong in hard times and my beloved sisters and broths who showed me unconditional love and support, I couldn't achieve my goal without you all with me .

Also my thanks for my friend Ghada who always been there for me, my thanks also for all my friends and colleagues.

ACKNOWLEDMENT

All praises are due to Allah, the lord of the worlds, who give me the ability and strength to make this thesis.

I would like to express my grateful thanks, gratitude and sincerest appreciation to my supervisor Dr. Azzam T. Sleit, for his continuous support, guidance and motivation which was needed through my work his supervision was an honor for me .

I also would thank my co-supervisor Dr. Bassam H. Hammo for his kind supervision and for the time he spent on my guidance to enhance this work. His kind assistance and guidance are extremely appreciated.

Khitam Mahmoud Jbara

Amman, 2009

TABLE OF CONTENTS

Committee Decision	II
Dedication	III
Acknowledgment	IV
Table of Contents	V
List of Tables	VII
List of Figures	VIII
List of Abbreviations	IX
List of Appendices	IX
Abstract	XI
1. Introduction	1
1.1 Problem Overview	3
1.2 Motivation	6
1.3 The Proposed Technique	8
1.4 Thesis Organization	9
2. Related Works	10
2.1 Related Work on Text Classification Methods	10
2.2 Related Work on Arabic Text Classification	16
3. The Proposed Mining Technique	25
3.1 System Overview	25
3.2 The Preprocessing Stage	27
3.2.1 Removing Al Hadith Sanad	28
3.2.2 Tokenization	29
3.2.3 Removing Punctuation and Diacritical Marks	30
3.2.4 Removing Stop Words	31
3.2.5 Stemming	33
3.3 Training Phase	34
3.3.1 Finding IDF Matrix For The Corpus	35
3.3.2 Finding Features Weights for the Training Set	39
3.4 Classification Process	41
3.4.1 Computing Query Weights	41
3.4.2 Query Expansion	43
3.4.3 Constructing the Similarity Coefficient Table	46
3.4.4 Assigning AL-Hadith Class	49
3.5 Evaluation	49
4. Discussion and Analysis of Experimental Results	50
4.1 Content of AL-Hadith Corpus	50
4.2 Hardware and Software	51
4.3 Discussion of the Experimental Results and Analysis	52
4.3.1 Classification Methods Applied On AL-Hadith Corpus	52
4.3.2 Experiments Specifications	54
4.3.3 Performance Measures	55
4.4 Results Analysis and Comparisons	56
4.4.1 Results of Our Proposed Method	56
4.4.2 Results of Word base classification (WBC)	60
4.4.3 Results of AL-Kabi Method	61
4.4.4 Comparisons and Results Analysis	62
4.4.4.1 Word based classification vs. Stem Expansion classification	62

4.4.4.2 AL-Kabi vs. Stem expansion classification.....	64
5. Conclusions and Future Works.....	67
5.1 Conclusions.....	67
5.2 Future Work.....	68
References.....	69
Appendix A.....	73
Appendix B.....	74
Appendix C.....	109
Abstract in Arabic.....	120

LIST OF TABLES

NUMBER	TABLE CAPTION	PAGE
3.1	Rules For Light Root Extraction Algorithm	34
3.2	Corpus IDF Matrix	38
3.3	Training Weights For Features In Training Set b	40
3.4	Training Weights for Features in a Training Set from The book of food	40
3.5	Query Weights For Features In AL-Hadith h	42
3.6	Query Weights Table for Mined Hadith	43
3.7	Extended Query Weights Table For Mined Hadith	46
3.8	Similarity Coefficient Table For Hadith b	48
3.9	similarity coefficient for features in the example Hadith	48
4.1	List Of Books In AL-Hadith Corpus	51
4.2	Recall And Precision Parameters	56
4.3	Our method Recall and Precision for Book of Food during Query Progress Stages	57
4.4	Stem Expansion Classification Results	59
4.5	Word Based Classification Results	60
4.6	Result of AL-Kabi's Classification Method	61

LIST OF FIGURES

NUMBER	FIGURE CAPTION	PAGE
1.1	Example of AL-Hadith in Arabic Text.	5
1.2	Example of AL-Hadith Translated To English.	5
3.1	An Overview of Proposed System Main Stages.	26
3.2	An Overview of Preprocessing Substages.	27
3.3	Example of Sanad Removal Process	29
3.4	Tokenizing process	30
3.5	Removing Diacritic Marks and Stop Words Algorithm	32
3.6	Output of Preprocessing Phase Structure	33
3.7	Structure Of The List Of Stems	34
3.8	Classification Process Steps.	41
3.9	Algorithm for Weighting in Expanded Set	45
3.10	Algorithm for Constructing the Similarity Coefficient Table	47
4.1	AL-Kabi Algorithm	53
4.2	Stem Expansion Method Results during Query Progress Phases	58
4.3	Stem Expansion Classification Results	59
4.4	Word Based Classification Results	60
4.5	Result of AL-Kabi's Classification Method	62
4.6	Word based classification vs. Stem expansion classification Precision Comparison	63
4.7	Word based classification vs. Stem expansion classification F_Measure Comparison	64
4.8	AL-Kabi's classification vs. Stem expansion classification Precision Comparison	66
4.9	AL-Kabi's classification vs. Stem expansion classification F_Measure Comparison	66

LIST OF ABBREVIATIONS

KDT	Knowledge Discovery in Text
TM	Text Mining
IR	Information Retrieval
TC	Text Classification
ATC	Automatic Text Classification
MSA	Modern Standard Arabic
NB	Naïve Bayes
SVMs	Support Vector Machines
DF	Document Frequency
CF-DF	Category Frequency-Document Frequency
TF-IDF	Term Frequency-Inverse Document Frequency
IG	Information Gain
NLP	Natural language processing
Ah	Year of the Hegira
MFSVM	Multi-subject TC Algorithm based on Fuzzy SVMs
NE	Named Entity
WBC	Word Base Classification
SEC	Stem Expansion Classification

List of Appendices

Appendix A: Stop Words list.....	73
Appendix B: IDF Matrix for Al-Hadith Corpus	74
Appendix B: Training weights for features in training set consists of 82 Ahadith from the book of food.....	109

KNOWLEDGE DISCOVERY IN AL-HADITH USING TEXT CLASSIFICATION ALGORITHM

By

Khitam Mahmoud Abdalla Jbara

Supervisor

Dr. Azzam T. Sleit

Co-Supervisor

Dr. Bassam H. Hammo

ABSTRACT

Machine Learning and Data Mining have been applied to language datasets to discover patterns: such as grammar inference, grammatical word-classifications for English and other European languages. Data mining is based on a variety of computational techniques Some of these techniques fall under the rubric of machine learning, examples include decision trees, neural networks and association rules (clustering) (Kanaan, et al., 2005). In this context, machine learning involves the acquisition of structural descriptions from examples, which can be used for prediction, explanation and understanding

AL-Hadith is the second religious source for all Muslims. AL-Hadith is the saying of Prophet Mohammed (Peace and blessings of Allah be upon him (PBUH)), because of AL-Hadith importance for Muslims all over the world we conduct this study which examine knowledge discovery from AL-Hadith to make AL-Hadith more understandable for both Muslims and non-Muslims.

In this thesis, we propose AL-Hadith knowledge discovery algorithm to classify AL-Hadith to one of predefined classes (books), the proposed mining algorithm is consists of two main phases the first one is the training phase in which we label the representative words of each chosen books, we use a training data set consists of already classified Ahadith (plural of AL-Hadith) based on Sahih Al-Bukhari which is the famous Ahadith book .The second phase, classification phase where the classification of each Hadith in the testing data set is conducted based on similarity function and class vector resulted from the training phase.

The proposed mining algorithm is considered as supervised classification system, since we use set of training Ahadith to train the classifier. The proposed algorithm use vector space model in representing Ahadith, so each Hadith is represented as a vector of its words. The weighting schema used is Term frequency X inverse document frequency (TFXIDF).

The experiments will be conducted on selected sample of Ahadith from Al-Bukhari book, where we have chose 13 books (classes) to run the experiments. The results of our algorithm indicate the importance of classification algorithm to extract knowledge from AL-Hadith. The evaluation of our algorithm is carried out by comparing our classification algorithm results to Al-Bukhari classification.

1. Introduction

Information Retrieval (IR) is the discipline that deals with retrieval of unstructured data, especially textual documents, in response to a query or topic statement, which may itself be unstructured, e.g., a sentence or even another document, or which may be structured, e.g., a Boolean expression. The need for effective methods of automated IR has grown in importance because of the tremendous explosion of the amount of unstructured data (Greengrass, 2000).

Text mining is a class of what called nontraditional (IR) strategies (Kroeze, et al., 2003). The goal of these strategies is to reduce the required effort from the users to obtain useful information from large computerized text data sources.

Text classifications(TC) which is also a subfield of Data mining refers generally to the process of deriving high quality information from text, that High quality information is typically derived through the dividing of patterns and trends through means such as statistical pattern learning.

Text classification is one of the most important topics in the field of natural language processing (NLP). The purpose of text classification algorithms is to assign each document of the text data set to one or more pre-specified classes. More formally if d_i is a document of set of documents D and $\{c_1, c_2, \dots, c_n\}$ is the set of all classes, then text classification assigns one category c_j to a document d_i and in multi-subjects classification d_i can be assign to more than one class from the classes set .

Text classification techniques are used in many applications, including e- mail filtering, mail routing, spam filtering, news monitoring, sorting through digitized paper archives, automated indexing of scientific articles, classification of news stories and searching for interesting information on the web (Khreisat, 2006).

Automatic text classification (ATC) is an important application and research topic since the inception of the digital documents. Today, automatic text classification is a necessity due to the large amount of text documents we have to deal with daily (Duwairi, 2006).

With huge growth of text documents on web and the growing number of Arabic document sources on the Internet, information retrieval becomes a crucial task to satisfy the needs of different end users. Automatic text (or document) categorization attempts to replace and save human effort required in performing manual categorization.

TC system design usually consists of three phases (Al-Mesleh, 2007):

1. Data pre-processing: in this step we try to make text documents compact and applicable to train text classifier by applying stemming and removal of stop words.
2. Text classification: in this step text classifier, which is the core of TC learning algorithm will be constructed, learned and tuned using the compact form of documents resulted from previous step.

3. Performance measurement: in this step we evaluate the effectiveness and performance of the proposed system using some of TC most common evaluation measures, such as precision, recall, etc.

1.1. Problem Overview

Arabic language is considered as one of the languages that will never distinguish. However, it is the official language of twenty Middle Eastern and African countries and is the religious language of all Muslims, regardless of their origin. Nevertheless few researches were made on Arabic corpus linguistics. Arabic differs from Indo-European languages syntactically, morphologically and semantically. It is a Semitic language whose main characteristic feature is that most of its words are built up from roots by following certain fixed patterns and adding infixes, prefixes or suffixes. It is an old language, but classical Arabic was standardized around fourteen centuries ago.

There are two main types of written Arabic language: the first one is classical Arabic, this type is used mainly in Quran, AL-Hadith (saying of the Prophet Mohammed) and classical literature. The second type is called modern standard Arabic (MSA), it is a modern form of Arabic which is used by the people in Arabic countries. This form of Arabic has many vocabularies which was developed during years and it doesn't use complicated forms of words and grammar rules that are used in classical Arabic (Khoja, 2001).

AL-Hadith is the utility of our research in text classification as a mean of text mining in Arabic documents .AL-Hadith is the second religious source for all Muslims.

AL-Hadith is the saying of the Prophet Mohammed (Peace and blessings of Allah be upon him (PBUH)); because of AL-Hadith importance for Muslims all over the world we conduct this study using Ahadith.

Bukhari who was born a couple of centuries after the Prophet's death collect AL-Hadith in a book called after him "Sahih Al- Bukhari" , he is one of the greatest scholars of Islam. He also wrote At-Tarikh and Ad-Du'afa, about the narrators of Hadith. He was the first scholar in Islam to compile a book in this manner, Al- Bukhari is well known for his restrict rules for Hadith acceptance, so his book considered the most precise collection of Prophet's Ahadith.

He undertook a long journey in 210 Ah in pursuit of AL-Hadith; he traveled to Khurasan, Iraq, Egypt and Syria and heard Ahadith from thousands of Shaiks. He collected thousands of Ahadith from which he selected those whose chains of narrations found to be correct and so he included them in his Sahih, which contains 7275 Hadith including those that are repeated. If the repeated Ahadith are omitted, the number will be 2230 which there is no doubt about their authenticity.

Sahih Al-Bukhari the most famous Ahadith book uses the feature ketab 'book' for his classification of Ahadith , ketab or book stand for subject , class and category , his classification is considered manual classification for Ahadith since he relied on the AL-Hadith meaning and content.

AL- Hadith text in Sahih Al-Bukhari consists of two parts (Al-Kabi and AL-Shalabi, 2005):

1. Sanad: it is the first part and it refers to the chain of names of those who have transmitted the saying to guarantee the authenticity and verbal accuracy of each saying. In other words, Sanad consists of all those who narrated the text, starting with the last narrator and ending with the Prophet (PBUH).
2. Matn: it is the second part and it refers to what the Prophet actually said or did and this part indicates the subject of Hadith or the reason for which it was said. In fact, this is the most important part that we are trying to classify, an example of Hadith where the two parts (Sanad and Matn) in Arabic text are shown in Figure 1.1, while Figure 1.2 show the same Hadith (Matn and Sanad) in English for clarification. Notice that the Sanad in italic letters while the Matn in normal letters.

عن عبد الله بن عمرو رضي الله عنهما عن النبي صلى الله عليه وسلم قال:
(المسلم من سلم المسلمون من لسانه ويده والمهاجر من هجر ما نهى الله عنه قال أبو عبد الله وقال
أبو معاوية حدثنا داود عن عامر قال سمعت عبد الله عن النبي صلى الله عليه وسلم وقال عبد الأعلى
عن داود عن عامر عن عبد الله عن النبي صلى الله عليه وسلم).

Figure 1.1: Example of AL-Hadith in Arabic Text.

Narrated 'Abdullah bin ' Amr The Prophet said:

"A Muslim is the one who avoids harming Muslims with his tongue and hands. And a Muhajir (emigrant) is the one who gives up (abandons) all what Allah has forbidden."

Figure 1.2: Example of AL-Hadith Translated To English.

Developments in electronic documents over web now offer an opportunity for applying the newest technologies of Knowledge Discovery in Text(KDT) for interrupting Arabic documents in both of its types ,the classical form like Quran and AL-Hadith or its MSA form like those in paper articles and newspapers . Our study exploited KDT concept, since we propose a methodology to discover knowledge from AL-Hadith using ATC Techniques. The Knowledge that is expected to be discovered is the classe (Subject) of AL-Hadith.

1.2 Motivation

The majorities of present Text classification techniques were designed to handle documents written in English language. However, they were not applicable to documents written in Arabic language, because Arabic language differs syntactically, morphologically and semantically from English language as mentioned before.

Developing text classification systems for Arabic text is a challenging task due to the complex nature of Arabic language (Al-Shalabi, et al., 2006). Arabic is a highly inflectional and derivational language which makes morphological analysis very complex (Hammo, et al., 2002). Arabic language consists of 28 letters, it is written from right to left .The majority of words have a tri- letter root, but rest of words have either quad-letter root, penta- letter root or hexa- letter root(Khreisat, 2006). For this reason, there are few text categorization techniques applied to Arabic text nowadays (Duwairi, 2006).

Most studies found in text mining literature use different document sets (Duwairi, 2006), which makes it difficult to differentiate the best text classification method. In this study, we use a data set represented by Ahadith (saying of the Prophet Mohammed (PBUH)) which considered a good choice to evaluate our proposed algorithm, since Ahadith were already classified manually by number of Islamic religion like Al-Bukhari.

The overwhelming majority of Muslims consider AL-Hadith to be essential supplements and clarifications of Quran, Islam's holy book. In Islamic jurisprudence, Quran contains many rules for the behavior expected from Muslims but there are no specific Quranic rules on many religious and practical matters.

Muslims believe that they can look at the way of life, or Sunnah, of Prophet Muhammad and his companions to discover what to imitate and what to avoid. Muslim scholars also find that it is useful to know how Prophet Muhammad or his companions explained the revelations, or on what occasion Prophet Muhammad received them. Sometimes this will clarify a passage that otherwise seems obscure. Many information about Islam can be found in Ahadith, in addition more understanding for Islam can be achieved by understanding Ahadith as apart of Sunnah ,to help all Muslims regardless of their origin and non-Muslims to understand Islam, it's Necessary to consider an automatic AL-Hadith classifier.

In this thesis a new methodology for discovering knowledge from Hadith based on ATC techniques were proposed. This research was carried out due to the following reasons:

1. Studies conducted in field of text mining and its application to Hadith are almost lacking. Up to our knowledge, one study (AL-Kabi ,et al., 2007) found as a result of investigation throw the internet, local and international Journals
2. The difficulty of dealing with AL-Hadith as a text, since AL-Hadith is written in classical Arabic language as mentioned earlier, In addition to the formation of AL-Hadith (mtan, Sanad).
3. Interpretation of Hadith is fundamental to Muslims throughout the world and to the non Islamic world in order to understand Islam.
4. Ahadith are well-suited for this study due to the fact that the contents of the book are constant, identified by Prophet Mohammed and authenticated by AL-Bukhari in his Sahih. Also, Ahadith have been analyzed by millions of people (i.e. Muslims and non-Muslims) over the past 15 centuries (Robinson, 2003).

1.3 The Proposed Technique

This research proposes a technique for knowledge discovery in AL-Hadith using ATC as IR mechanism. This technique specifically depends on a classification method to classify the selected testing set to a predefined classes or books, so each Hadith will be related to the most similar class which will help Muslims and non-Muslims who interested in AL-Hadith science to narrow their search through a huge number of Ahadith.

The proposed mechanism consists of several stages. The first stage is the preprocessing which consists of a number of steps, starting from removing AL-Hadith Sanad, tokenizing of Matn, removing punctuation marks and stop words. As a query

expansion technique we use light stemming to extract the root of the features. the second stage is the training phase which aims to find the most representative features for each class, this stage is completely automated .in this stage each Hadith is represented by the weights of its features according to the term Frequency- Inverse Document Frequency (TF- IDF) weighting schema, the third stage we use is the TF- IDF weighting and the cosine similarity coefficient to classify the testing set. In the final stage results, evaluation and comparisons were displayed.

1.4 Thesis Organization

Chapters of this thesis are constructed as follows: Chapter two reviews related work on classification techniques and text mining in foreign language (English). Also, it reviews studies on Arabic text mining and classification and work done in AL-Hadith classification. Chapter three presents the proposed methodology, its architecture, the design and implementation of the system is discussed in details. Chapter four illustrates the experimental results, the analysis and the evaluation of the proposed system. Chapter five provides a conclusion for the thesis as well as some suggestions for future approaches and ideas.

2. Related Work

This chapter reviews some of the previous studies that are concerned to text classification techniques. Also, related work on Arabic text classification is presented. Finally, related work on AL-Hadith knowledge discovery is described in details.

2.1 Related Work on Text Classification Methods

Chu (2008) offers a synopsis on how linear algebra techniques can help to carry out the task of data mining. Examples from factor analysis, cluster analysis, latent semantic indexing and link analysis are used to demonstrate how matrix factorization helps to uncover hidden connection and fasten the data mining task, the paper demonstrates the role and importance of Low rank matrix approximation which plays a fundamental role in cleaning and compressing the data, by replacing the original data by a lower dimensional representation obtained via subspace approximation. The study also demonstrates the problems of matrix factorization and suggests fundamentals of models and methods from linear algebra prospective.

Naïve Bayes (NB) is a probabilistic classifier, which is one of the most efficient and effective inductive learning algorithms for machine learning. Naïve bayes classifier system consists of two main parts; the classifier module based on Naïve bayes theorem and the learning module. The NB classifier calculates the posterior probability for classes and the estimation is based on the training documents. In the testing phase, the posterior probability for each class is computed and then the document is classified to the class that has the maximum posterior probability. In the learning module of the

document classification system, the module has to be learned from a set of documents that are pre-classified (Zhang, 2004).

Zhang (2004) proposed in his work a novel explanation on the superb classification performance of NB. He found that that how the local dependence of a node is distributed in each class, evenly or unevenly and how the local dependencies of all nodes work together, consistently (supporting a certain classification) or inconsistently (canceling each other out), plays a crucial role. Therefore, no matter how strong the dependences among attributes are, naive Bayes can still be optimal if the dependences distribute evenly in classes, or if the dependences cancel each other out.

Liu, et al. (2004) in their work propose a different approach from the traditional manual labeling for training sets. Instead of labeling a set of documents, the proposed method labels a set of representative words for each class. Then it used these words to extract a set of documents for each class from a set of unlabeled documents to form the initial training set. They also proposed a method which combines between clustering and feature selection. Their technique effectively ranks the words in the unlabeled set according to their importance, then the user can select some words from the ranked list for each class, their technique considered as highly effective and promising method that requires less effort than providing words with no help or manual labeling of documents, their technique can be summarized as follows:

1. Given a set of unlabeled documents U . Cluster each document in U .
2. Perform a feature selection on the clusters. By applying the entropy-based method (information gain). The words with higher gain values are considered as the most important words in the unlabeled set U .

3. Identify (by the expert) a set of representative words for each class. This step is performed manually.
4. Build the final classifier. The classifier that is used was combining the Naïve Bayesian and Expectation Maximization algorithms.

They used a data set consist of articles and four main categories .The experimental results show that the proposed technique is highly effective.

Lukui, et al. (2007) investigates improving of executing efficiency for classification methods. In their work they presented a classification algorithm based on nonlinear dimensionality reduction techniques and support vector machines. The algorithm consists of two steps:

1. ISOMAP algorithm is used to reduce the dimension of the high-dimensional text; ISOMAP is one of representative non-linear dimensionality reduction techniques, which generalizes multidimensional scaling (MDS) by replacing Euclidean distances with an approximation of the geodesic distances on the Manifold.
2. The resulted low-dimensional data are classified with a multi-class classifier based on a single-class SVM.

Their Experimental results demonstrate that the executing efficiency of categorization methods is greatly improved after decreasing the dimension of the text data without loss of the classification accuracy.

Ikonomakis, et al. (2005) illustrate text classification process using machine learning techniques. They discuss the ATC related issues like vector space document representation, various feature selection and transformation methods, different machine learning algorithms like VSM, Latent Semantic Indexing (LSI) and k-Nearest Neighbor (KNN); also they mention the most popular evaluation measurements.

Yu-ping, et al. (2007) proposed a multi-subject text classification algorithm based on fuzzy support vector machines (MFSVM), 1-a-1 method is used to train sub-classifiers (Krebel, 1999). For the sample to be classified, the sub-classifiers are used to obtain membership matrix and the membership degree matrix is obtained according to sub-classifiers, then subjects that the sample belongs to are confirmed according to the sum of every line of the membership degree matrix. Their strategy extends the capability of the standard SVM to solve multi-subject classification problems. Experimental results show that MFSVM realizes multi-subject text classification and has higher precision and recall.

Bellot, et al. (2003) proposed an approach that combines a Named Entity recognition system and an answer retrieval system based on Vector Space model that uses some knowledge bases. Their system consists of several steps, in the first step they used modified Named Entity (NE) Recognition system for each type of NE, a transducer has been manually developed using a test corpus for validation. The transducer vocabulary is not only based on lexical information, but also on semantic information. In the next step, the information retrieval system evaluates the classification and segmentation methods, their work based on (Bellot & El-Bèze, 2000).

They used IR system to index and rank sentences extracted from the top documents by employing some classical methods: vector space model, cosine similarity and TF-IDF weighting scheme. At last they take benefit from a set of knowledge Databases, this process was useful to extract other instances of such relations from full text collections.

Ontology learning could be defined as a set of methods and techniques used for building, enriching, or adapting an existing ontology in a semi-automatic fashion using several sources. Saidah, et al. (2007) presented a general skeletal methodology using software engineering V-process model for building ontology as a mean of structuring knowledge and building knowledge-intensive systems.

They describe a case study for mining ontologies from text using methods based on natural language text; where they try to build ontology from Islamic knowledge understanding text (English version) includes sources from AL-Quran, Hadith and tafseer .The stages in the V-process model are: identifying purpose and scope, knowledge acquisition, conceptualization, encoding, documentation and evaluation, they recommended for more investigation on what types of linguistic and heuristic output are best suited to optimize performance. Finally they integrate several existing ontologies such as WordNet as a core resource into their cyclic approach.

An important feature of neural network classifier is that it must be trained before it can be used for text categorization. Training of the neural network classifier is done by Back Propagation learning rule based on supervised learning, while the main difficulty in the application of neural network for text categorization is the high

dimensionality of the input feature space for textual data, this is essential as the number and types of the categories is taken into account in choosing the topology and the size of the neural network classifier. Lam, et al. (1999) address this problem and study the efficiency of four feature reduction techniques, document frequency (DF), category frequency-document frequency (CF-DF) , term frequency-inverse document frequency (TDF-IDF) and principle component analysis (PCA) .

Their system is a three layer feed-forward neural network , the three layers are an input layer, a hidden layer and an output layer ,where neurons are non-linear units with the sigmoid function as the activation function, the number of input units is equal to the dimensionality of the reduced feature space using one of the studied reduction method , number of hidden units in the neural depends on the size of the training set and the complexity of the classification task the network is trying to learn, while the number of output units is equal to the number of pre-defined categories in particular text categorization task which was ten in their study .

The results show that Back propagation learning in neural networks was able to give good categorization performance as measured by precision and recall. They also found that principal component analysis was the most effective reduction feature method with high reduction rate (98.9%) achieved with insignificant decrease in categorization effectiveness.

Isa, et al. (2008) presented an enhanced hybrid classification method through the utilization of the Naïve Bayes classifier and Support Vector Machine (SVM).In their

study they implemented the Bayes formula to victories a document according to probability distribution reflecting the probable categories the document may belong to.

Their system consists of two main steps; first, the Bayes formula is used to calculate the posterior probability between the document in the training set and each one of existing categories .Second, the resulted vector from first step is used as a vector representing the raw document which will be the input to SVM classifier

In their experiments they compared between the proposed hybrid NB-SVM classifier and other several classifiers like pure NB, term frequency-inverse document frequency (TF-IDF) with SVM classifier and Lsquare method. The results show that the proposed system has improved classification accuracy compared to pure NB and it showed better performance compared to other techniques.

2.2 Related Work on Arabic Text Classification

AL-Kabi, et al. (2007) presented a comparative study on the efficiency of different measures to classify Arabic documents. In their study they tried to find the optimal method to classify Arabic text among six methods (Inner Product, Cosine, Jaccard, Dice, Naïve Bayesian and Euclidean). They investigated the use of TF-IDF to represent document vector, then document vector is used to compute and compare four different associative coefficients of the Vector Space Model (VSM) based on the inner product, cosine, jaccard and dice. They found that the cosine measure outperformed the other three associative coefficients of the VSM. They also compare the results obtained from VSM with cosine measure against Naïve Bayesian and Euclidean classifiers for

Arabic text. Also their experimental results show that NB method slightly outperforms the other methods.

AL-Mesleh (2007) presented a classification system based on Support Vector Machines (SVMs). His classifier uses CHI square as a feature selection method in the pre-processing step of the text classification system procedure. He used dataset consists of Arabic articles to test the proposed system which consists of two main phases:

1. Preprocessing: this phase consists of several steps beginning at removing the digits and punctuation marks, normalizing of some letters like (hamza, alef), filtering all non-Arabic words along with stops words. Finally, he used (4 features) feature redundancy for removing infrequent features.
2. Feature selection: as mentioned before his work uses the CHI as feature selection method. He found that CHI and IG were the best selection methods.
3. Classification: he used SVM Classifier; the proposed classifier is formulated into two different cases, the separable case where the training data is linearly separable and the non-separable case where real data is not usually linearly separable.

He compared F-measure between his system, NB and k Nearest Neighbor (K-NN) classifiers, his system achieved high classification effectiveness for Arabic data set with F-measure (F=88.11) and outperforms both of them.

El-Halees (2007) built a system called ArabCat based on maximum entropy model to classify Arabic documents. Maximum entropy is a technique for modeling probability distributions, it is supervised learning algorithm that needs a training set. His system used the posterior probability and can be described as follows: first, both the training and testing document must pass through the preprocessing step by removing the stop words and diacritics above the letters, then the root is extracted for the remaining words.

Second, in the training step, the maximum entropy is applied to train the documents. After that, the categorizer takes the processed document as an input and classifies it to the class with a maximum posterior probability. The experiments of their system were performed with and without stemming approach and using part-of-speech (POS). The results show that using POS besides preprocessing (remove stop words) increased the performance of the system. Finally, the ArabCat system is evaluated by performing the recall and the precision measures.

The experimental results showed that the ArabCat system overcomes the other existing systems that deals with Arabic text classification (El-Kourdi et al., 2004) (El-Halees, 2006) in precision. But (Sawaf et al., 2001) had a better recall value.

Saleem et al. (2007) approach combined shallow parsing and information extraction techniques with conventional information retrieval ,by that they add some artificial intelligence to build more richer database of events to improve the IR .

Their IR system has the following architecture:

1. Event extraction module: The event in their system has three main types: keywords, dates and proper nouns. The event extraction module looks for events in the text, marks the paragraph that contains events, parses it and extracts some information to save it in the database.
2. Event classification module: this module determines to which category any old extracted event belongs to and to which category a new event should be added in the database.
3. Information retrieval module: this module uses the information about the events saved in the database, in order to build menus that will be used by the users to request a query, then the query processed to find related events and to retrieve the related identifiers from the database.

AL-Kabi, et al. (2005) built an automatic classifier to classify the verses of Fatiha and Yaseen Chapters to predefined themes. Their system is based on linear classification function (score function).

Their methodology of Quran classifier system can be explained as follow:

After selecting the desired Sura (Yaseen or Fatiha) as mentioned before, the user must select a verse from that Sura. Then normalization phase start which consists of removing stop words, diacritics marks above the letters and punctuations. The remaining words of the verse are considered as features (keywords). For each feature extracted from the verse, the system finds the recurrences of it in Yaseen and Fatiha verses. Then the system will determine the subject of each verse that contains that

feature. The subject of the verses of Yaseen and Fatiha Chapters are pre-determined and built manually. At the end of this process, the themes-frequency which is the frequency of all themes for each feature is calculated. Then the final weight of the feature is calculated. When this process is accomplished, two dimensional matrixes are built where columns represent the subjects (15 subjects), the rows represent the features and the entries represent the weight of features that are relative to the subject. Then the entries of each column are added to produce the percentage of the subject to all the features of a verse. Finally, the system assigns the verse to the class with the maximum summation value.

The accuracy of their system reached 91% in classifying different verses of Yaseen and Fatiha Chapters. The authors recommended using larger Quranic corpus in future experiments.

Hammo, et al. (2002) presented a design and implementation of a Question Answering (QA) system called QARAB. The system primary source of knowledge is a collection of Arabic newspaper text extracted from AL- Ray, a newspaper published in Qatar. The main goal of the QARAB system is to identify text passages that answer a natural language question. Their proposed system can be summarized as follows:

Given a set of questions expressed in Arabic, find answers to the questions with some holding assumptions: first, an answer exists in a selected collection of Arabic newspaper. Second, an answer does not span through documents and the answer is a short passage.

The basic QA processing in QARAB is composed of three major steps:

1. Processing the input question: This step involves tokenization, stop word removal, root extraction and feature weighting.

2. Retrieving the candidate documents (paragraphs) containing answers from the IR System; the purpose of the IR system is to search the document collection to select documents containing information relevant to the user's query.
3. Processing each one of the candidate documents (paragraphs) in the same way the question was processed in returning sentences that may contain the answer.

Hammo, et al. (2008) in this paper we discuss the enhancement of Arabic passage retrieval for both diacritized and non-diacritized text. They propose a passage retrieval approach to search for diacritic and diacritic-less text through query expansion to match user's query. They applied a rule-based stemmer and compiled a huge thesaurus for this purpose; their approach was tested on scripts of the Quran as an open domain source of diacritized text using a set of 40 non-diacritical words obtained from testers.

They also discuss the query expansion methods, query expansion can be defined as the process of reformulating the query's bag- of- words to overcome the problem of mismatching potential documents and improving the performance of a search engine by including documents that are more relevant, they specifically test the effectiveness of two query expansions approaches; Query Expansion through Stemming and Query Expansion through Thesaurus. Their study presented a good improvement in recall when using the expansion technique by stemming and thesaurus.

EL-Kourdi, et al. (2004) built an Arabic document classification system to classify non-vocalized Arabic web documents (after their words have been transformed, i.e., root extraction) based on Naïve Bayes algorithm. Their NB classifier consists of two parts; the classifier module and the learning module. In the classifier module, NB classifier calculates the posterior probability of the classes using estimates obtained from a training set of labeled documents. When unlabeled document is presented the posterior probability is computed for each category, then the document will be assigned to the category with the highest posterior probability.

They investigated the benefit of using a feature selection method in the classifier, also their experiments showed that the root extraction method used for preprocessing phase is not sensitive to the performance of the NB algorithm

Khreisat (2007) presented a comprehensive study for the behavior of the N-Gram Frequency Statistics technique for classifying Arabic text documents, An N- gram is an N- character slice of a string, this study compared between two measurements, dissimilarity measure called the “Manhattan distance” and Dice’s measure of similarity, for the purposes of classification.

Khreisat used a corpus of Arabic text documents that was collected from online Arabic newspapers, the documents spited into a training set and a testing set. Her study can be summarized as following:

1. All text documents (training set and testing set) passed through a preprocessing stage, which include, remove punctuation marks, diacritics, non-letters, stop words, etc.
2. Then the N-gram profile was generated, generating the N-gram profile consisted of many steps beginning with tokenizing and removing (digits are removed), then find all possible N- grams (N=3), after that compute the frequency of occurrence for each N-gram. Finally, sort the N- grams according to their frequencies from most frequent to least frequent.
3. The N-gram profile of each text document was compared with the profiles of all documents in the training classes in features of similarity.

They found that N-gram text classification using the Dice measure gives better classification results compared to the Manhattan measure.

Baarah, A.H (2007) proposed a Quran Mining technique (QMT) to classify the Suras of Quran based on pre-defined themes. The proposed QMT is consists of two phases. The first phase is to label a set of representative words for each of the pre-defined themes, which was accomplished manually. In the second phase, the classification of each Sura is done based on the classification function (Score function).

The proposed algorithm is applied to a collection of 14 Sura (Chapter) from Quran .The evaluation of QMT is carried out by comparing QMT results with experts' results. The average accuracy of the QMT reached approximately 79%.

3. The Proposed Mining Technique

The proposed method aims to discover knowledge in AL-Hadith .Knowledge discovery in our system depends on text Classification algorithm to relate each Hadith to the nearest class, depending on AL-Hadith content and the knowledge discovered from the training data set.

In this chapter we give a general overview for the structure of the proposed system, followed by an illustration for each component in the system in details.

3.1. System Overview

Our classification system consists of four main stages as shown in Figure 3.1. First, the preprocessing phase. Second, the training phases where we build the learning database which contains the weights of features representing a class, a set of pre-classified documents are the input for this stage. Third, the classification stage in this stage we use the resulted training database from previous stage with the classification method to classify AL_Hadith, in this stage the query expansion occur and the output of this stage will be the class (book) of AL-Hadith need to be classified. Finally, data analyzing and evaluation. Proposed classification System stages will be discussed in details in this chapter.

We can define the corpus that contains a set of Ahadith as in definition 1.

Definition 1: Corpus Definition

Suppose corpus $C = \{H_1, H_2, H_3, \dots, H_n\}$. Where H_i represents the i th tested Hadith in C , n is the number of tested Hadith in the C and $i: 1 \dots n$.

Suppose $H_j = \{w_1, w_2, w_3, \dots, w_m\}$. Where w_d represents the d th word in Hadith H_j , m is the number of words in H_j and $d: 1 \dots m$.

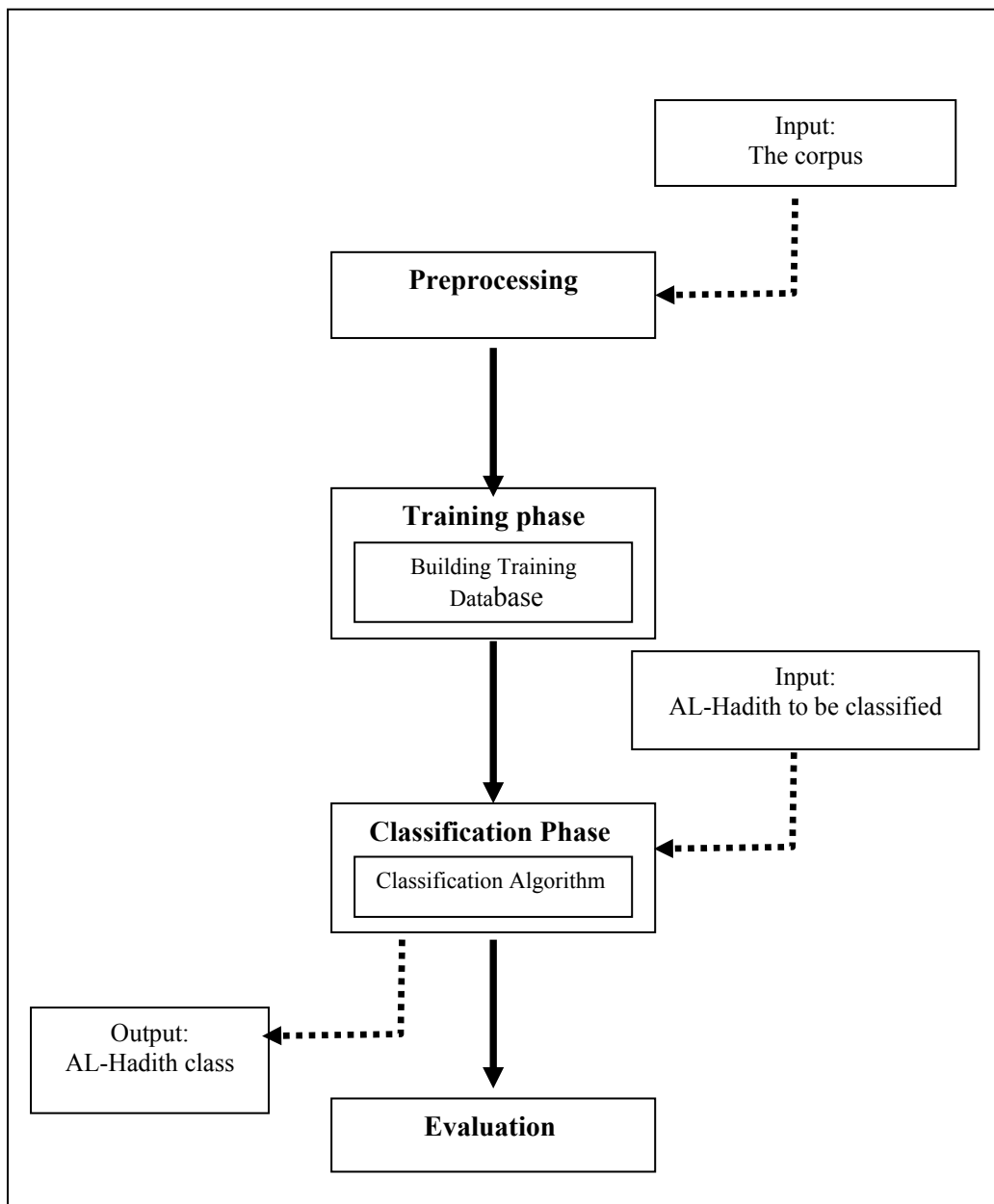


Figure 3.1: An Overview of Proposed System Main Stages.

3.2. The Preprocessing Stage

In this section we introduce the preprocessing techniques which will be conducted on each Hadith used in the training and testing sets. This stage is necessary before the classification phase can be applied to discover knowledge from AL-Hadith. Preprocessing stage consists of several Substages as shown in Figure3.2. In the following sections the details of each Substage will be discussed.

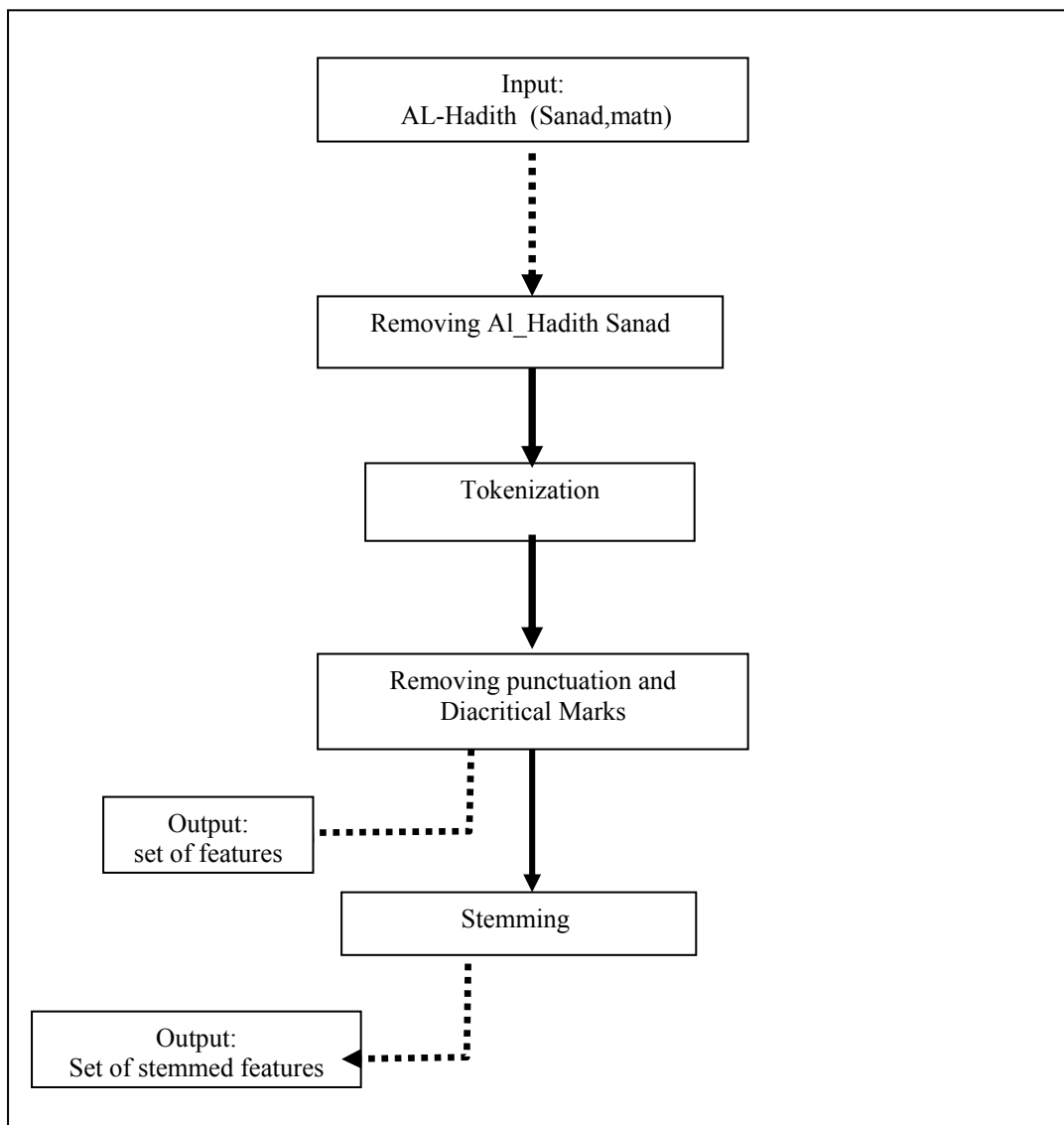


Figure 3.2: An Overview of Preprocessing Substages.

3.2.1 Removing AL-Hadith Sanad

As mentioned before Sanad is the part of AL-Hadith which refers to the chain of names of those who had transmitted the saying, the Sanad mainly used to guarantee the authenticity and verbal accuracy of each saying, in fact the same Hadith can have two or more different Sanad. The chain of AL-Hadith narrators can not affect or imply the subject of AL-Hadith, because of that we found that it is necessary to remove the Sanad and not to consider it in the classification stage.

The process of Sanad removal was conducted manually, this is because there is no consistence pattern for Sanad, as we notice in our corpus the Sanad can appear at the beginning or at the end of AL-Hadith and it not always begins or ends with static words, because of that we choose to remove the Sanad manually to assure the correctness of the removal process, so the main body (Matn) of AL-Hadith will not be affected. Sanad removal was a difficult and tedious process but it has a big effect in the testing accuracy.

Example:

In this example we demonstrate preprocessing phase where we used Hadith form book of food .Figure 3.3 shows example of AL-Hadith before and after Sanad removal. AL-Hadith after removing the sanad consists of 26 words and some of the words are vocalized.

حدثني إسحاق بن إبراهيم: أخبرنا رُوْح بن عبادة: حدثنا ابن أبي نَُنب، عن سعيد المَقْبُري، عن أبي هريرة رضي الله عنه:
أنه مر بقوم بين أيديهم شاة مَصْلِيَّة، فدعوه، فأبى أن يأكل وقال: خرج رسول الله صلى الله عليه وسلم من الدنيا ولم يشبع من خبز الشعير.

(a) Al-Hadith before Sanad Removal

أنه مر بقوم بين أيديهم شاة مَصْلِيَّة، فدعوه، فأبى أن يأكل وقال: خرج رسول الله صلى الله عليه وسلم من الدنيا ولم يشبع من خبز الشعير.

(b) Al-Hadith after Sanad Removal

Figure 3.3: Example of Sanad Removal Process

3.2.2 Tokenization

To induct the classification process in our algorithm which depends on the TF-IDF weighting schema, the Matn of tested AL-Hadith must be divided into tokens (words), in this stage some of tokens (words) are vocalized.

AL-Hadith segmentation was easily resolved since each word (Token) can be identified as a string of letters between white spaces (Thabet, 2005). Figure 3.4 explains the tokenization algorithm for AL_Hadith where the input is Ahadith in the corpus and the output is the Table containing those Ahadith tokens.

This stage includes also normalizing of some characters to its common form, for example alef (ا) which has many forms like {أ، آ، إ} in Arabic language specially when it appear in the beginning of the word.

For the same example demonstrated in previous sections, the result of removing Punctuation and Diacritical Marks will be the following set of features:

{ انه , مر, بقوم, بين, ايديهم, شاة, مصلية, فدعوه, فايى, ان, ياكل, وقال, خرج, رسول, الله, صلى, الله, عليه, وسلم, من, الدنيا, ولم, يشبع, من, خبز, الشعير }

3.2.4 Removing Stop Words

Stop words are words that found in AL-Hadith and have no discriminative meaning .Stop words are considered as the most frequent words in the language so these words are worthless and not useful in classification process (AL-Kabi, et al., 2005).

In our system we built manually a list of stop words which consists of Arabic pronouns, prepositions, names of people (companions of Prophet Mohammed) and places were mentioned in AL-Hadith corpus. After removing stop words from AL_Hadith, the remaining words (terms) are considered as features. Appendix A contains the list of stop words used in our study. The algorithm for removing the diacritic marks and stop words is described in Figure 3.5.

Algorithm 3.2: Removing Diacritic marks and Stop Words.

Input: Table of words resulted from tokenization process.

Output: A Table consists of all features of C

for Each Hi in C Do

 Begin

 For Each word in Hj Do

 Begin

 Word=remove diacritic (word)

 IF word not an entry in Stop Word Table then

 Feature = word

 else

 Remove word

 End

 End

 End

End

Figure 3.5: Removing Diacritic Marks and Stop Words Algorithm.

For the same example Hadith demonstrated in previous sections, the result of removing stop words will be the following set of features:

{ مر، بقوم، ايديهم، شاة، مصلية، فدعوه، فابي، ياكل، خرج ، الله ، صلى ، عليه، وسلم ، الدنيا ، يشبع ، خبز،
الشعير }

The out put of preprocessing phase after removing diacritic marks and stop words is a list of features, Figure 3.6 shows the structure of the resulted list of features, where AL_Hadith ID is the identification number for AL_Hadith in the specific class and feature ID is a sequence number of the feature in mined Hadith .

CLASS_NAME	AL_Hadith ID	Feature ID	Feature
The Book of Food	10	1	الدنيا
The Book of Food	10	2	الشعير
The Book of Food	10	3	الله
The Book of Food	10	4	ايديهم
The Book of Food	10	5	بقوم
The Book of Food	10	6	خبز
The Book of Food	10	7	خرج
The Book of Food	10	8	شاة
The Book of Food	10	9	صلى
The Book of Food	10	10	فابي
The Book of Food	10	11	فدعوه
The Book of Food	10	12	مر
The Book of Food	10	13	مصلية
The Book of Food	10	14	وسلم
The Book of Food	10	15	ياكل
The Book of Food	10	16	يشبع

Figure 3.6: Output of Preprocessing Phase Structure

3.2.5 Stemming

In this step we extract the stems of the features, stem extraction we implement is considered as light stem extraction which depends on removing some prefixes or suffixes from the word to relate the word to its stems. We used the algorithm proposed by (Al-Serhan, et al., 2003), the result of the stem extraction was filtered to eliminate the incorrect roots and roots with less than three characters. The stems of AL-Hadith features will be used in the query expansion process which will be discussed in details in section 3.4.2. Table 3.1 shows some rules used to extract roots of features and an example for each rule.

Table 3.1: Rules for Light Root Extraction Algorithm.

Rule	Characters to remove	Example	
		Before stemming	After stemming
Remove prefix	ال، بب، و، ال، كال	الطعام، والطعام، كالطعام	طعام
Remove suffices	ين، كم، ون، ان، ها، هن، هم، ات	لباسهن، لباسكم، لباسها	لباس

For the same example demonstrated in previous sections, the result of light stemming described above will be the set of stems { ايدي، دنيا، شعير } and the structure of the stems list produced by the stemming stage is shown in Figure 3.7, where feature ID is the identification number for the original word from which the stem was extracted and stem ID is a sequence number for the stem in AL_Hadith .

AL_Hadith ID	Feature ID	Stem_id	Stem
10	1	1	دنيا
10	2	2	شعير
10	4	3	ايدي

Figure 3.7: structure of the list of stems

3.3 Training Phase

Traditional text classification depends mainly on choosing a set of words to represent a class or category, the dilemma in traditional classification is how to select the set of representing words for a class. The labeling of words in this technique is done manually by experts of the selected field, so choosing of these representative words will be a hard and time consuming process.

Supervised classification exploited the predefined training documents that belong to specific class to extract the features that representing a class. So, every class will have a feature vector representing it, and then those features will be reduced using one of the feature selection techniques. Finally, the classification techniques are applied to the test document to assign it to one of the pre-defined class.

Supervised classification has its difficulties; one main problem is how to be sure that trained document actually belongs to a specific class. In our study this problem does not exists, since we conduct the study on set of Ahadith which were classified by the famous AL-Hadith scientist AL-Bukhari , that gave us a good base to evaluate the proposed classification algorithm.

This step consists of two main stages; the first step was executed once to produce the matrix of inverse document frequency (IDF) for the corpus . The second step was executed for each training set, sections 3.3.1 and 3.3.2 explain those stages in details.

3.3.1 Finding IDF Matrix for the Corpus

After conducting the preprocessing phase we got a list of features for each Hadith in the corpus which will be used in the classification process. Building the IDF matrix for AL-Hadith corpus is a one time process and this matrix will be used in the classification process every time we need the IDF value of a feature.

One common approach for document representation and indexing for statistical purposes is to represent each textual document as a set of terms; this approach is known as Vector Space model. The terms are words extracted automatically from the documents themselves. This set of terms defines a “space” such that each distinct term represents one dimension in that space (Greengrass, 2000). Since in vector space model each document is represented as a set of terms, then each term can be assigned a numeric weight in a given document. Term weight represents an estimation of the usefulness of the given term as a descriptor of a given document. A given term may receive a different Weight in each document in which it occurs; that mean a term may be a better descriptor of one document than of another.

According to Vector Space Model AL-Hadith in our system is represented as a vector of features, one of the most successful and widely used scheme for automatic weighting is the “term frequency×inverse document frequency” (TF×IDF) Weighting scheme, which is the product of two concepts (TF and IDF), where TF is a document specific statistic refers to a feature occurrence time; since feature that appears many times within a training document can be assumed to be more important than that appears only once, while IDF is a global statistic refers to that feature which occurs in a few training documents is assumed to be a better discriminator than a feature that appears in most of training documents or all of them.

idf for specific feature (i) is calculated as follows:

$$= \left(\frac{1}{N} \right)$$

Where

N: number of Ahadith in the corpus.

dfi: Number of Ahadith in the corpus containing feature i.

Fewer documents containing a given feature will produce a large IDF value and if every document in the collection contains the given feature, the IDF will be zero which means that the feature which occurs in every document in a given collection is not likely to be useful for distinguishing relevant from non-relevant documents.

Computing the weight of a given feature in a given document as (TF×IDF) is the best descriptor, because this weighting scheme combines the importance of TF and IDF in the same time. In other words, a feature of a given document will be the feature that occurs many times in the given document and fewer times in other documents also a feature that occurs a moderate number of times in a moderate proportion of the documents in the given collection will be a good descriptor (Greengrass, 2000).

As illustrated before feature IDF is a global measurement, it means that we will explore all ahadith in the corpus to find each feature IDF. Table 3.2 shows the IDF sparse matrix of features for all classes; we call it sparse matrix since some features will not appear in some classes.

In our system a feature will be considered in the classification algorithm only if its feature redundancy value is less than 45, according to empirical studies threshold 45 give the best classification results (AL-Kabi ,et al., 2007)(Forman,2003).

3.3.2 Finding Features Weights for the Training Sets

Our system depends on using a set of Ahadith as training documents to extract set of representative words for each book (class) then compute their weights. For finding features weights we will use the IDF matrix and the features training weight is computed as shown in the next formula:

$$TW_{bi} = TF_{bi} \times IDF_i$$

Where:

TW_{bi} : feature i training weight in training set b.

TF_{bi} : feature i frequency in training set b.

IDF_i : feature i inverse document frequency calculated earlier in IDF matrix.

Features that will be considered to calculate their training weights must satisfy the feature redundancy condition mentioned earlier (feature redundancy <45). Table 3.3 shows training features weights which will be calculated for each training set, notice that we will view features in Ahadith as one bag-of-words. In each book (class) a specific portion of Ahadith will be used as a training set, this process will be discussed in details in chapter 4.

Table 3.3: Training Weights for Features in Training Set b.

		Book_b	
Feature	IDF	TF	TW
Feature1	IDF ₁	TF _{b1}	TW _{b1} = TF _{b1} * IDF ₁
Feature2	IDF ₂	TF _{b2}	TW _{b2} = TF _{b2} * IDF ₂
Feature3	IDF ₃	TF _{b3}	TW _{b3} = TF _{b3} * IDF ₃
Feature4	IDF ₄	TF _{b4}	TW _{b4} = TF _{b4} * IDF ₄
Feature5	IDF ₅	TF _{b5}	TW _{b5} = TF _{b5} * IDF ₅
.....			
.....			
.....			
.....			
.....			
.....			
.....			
.....			
Feature n			TW _{bn} = TF _{bn} * IDF _n

Appendix C shows the training weights for feature in a training set consisting of 82 Ahadith from the book of food (90% of the book of food Ahadith), and table 3.4 shows the training weights for the features included in the example Hadith illustrated in the previous sections.

Table 3.4: Training Weights for Features in a Training Set from The book of food.

The book of food (training set No.1)			
Feature	IDF	TF	TW
ياكل	1.80	8	14.39
الدنيا	1.84	1	1.84
شاة	1.97	4	7.90
مر	2.12	1	2.12
ايديهم	2.22	1	2.22
خبز	2.34	4	9.37
فابى	2.42	1	2.42
الشعير	2.52	2	5.04

3.4 Classification Process

The classification process consists of four steps as shown in Figure 3.8. First, is computing query weights where we find feature's weight in AL-Hadith to be classified. Second, is the expansion process where we use the stems to expand our query. Third, is calculating the similarity coefficient for each feature in AL-Hadith to be classified. The final step is finding the cumulative similarity for AL-Hadith over the predefined classes (books).

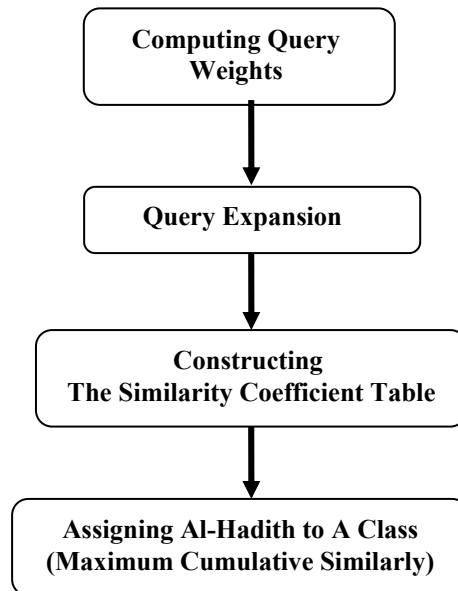


Figure 3.8: Classification Process Steps

.

3.4.1 Computing Query Weights

A feature weight in the query (specific Hadith) is calculated as follows :

$$Q_h W_i = TF_{hi} \times IDF_i$$

Where:

$Q_h W_i$: feature i weight in AL-Hadith h (Hadith to be classified).

TF_{hi} : feature i frequency in AL-Hadith h

IDF_i : inverse document frequency calculated earlier in IDF matrix.

Query weights will be computed for each feature in AL-Hadith to be classified, TF depend on AL-Hadith features occurrences while the IDF is a global value referenced from IDF matrix.

Table 3.5 shows query weights calculation for AL-Hadith where the feature column represents the set of all features belongs to mined Hadith, while Table 3.6 shows the query weights for example Hadith illustrated in precious sections.

Table 3.5: Query Weights for Features in AL-Hadith h.

Feature	IDF	Hadith h	
		TF	QW
Feature1	IDF_1	TF_{h1}	$Q_h W_1 = TF_{h1} * IDF_1$
Feature2	IDF_2	TF_{h2}	$Q_h W_2 = TF_{h2} * IDF_2$
Feature3	IDF_3	TF_{h3}	$Q_h W_3 = TF_{h3} * IDF_3$
.....			
.....			
.....			
.....			
.....			
.....			
.....			
.....			
Feature n			$QW_{hn} = TF_{hn} * IDF_n$

Table 3.6: Query Weights Table for Mined Hadith.

Feature No.	Feature	IDF	Feature Redundancy	TF	QW
1	الدنيا	1.84	1.44	1	1.84
2	الشعير	2.52	0.30	1	2.52
3	الله	Feature redundancy >45	80.02	2	0
4	ايديهم	2.22	0.61	1	2.22
5	يقوم	2.64	0.23	1	2.64
6	خبز	2.34	0.45	1	2.34
7	خرج	1.53	2.95	1	1.53
8	شاة	1.97	1.06	1	1.97
9	صلى	Feature redundancy >45	68.89	1	0
10	فابى	2.42	0.38	1	2.42
11	فدعوه	3.12	0.08	1	3.12
12	مر	2.12	0.76	1	2.12
13	مصلية	3.12	0.08	1	3.12
14	وسلم	Feature redundancy >45	68.58	1	0
15	ياكل	1.80	1.59	1	1.80
16	يشيع	2.64	0.23	1	2.64

3.4.2 Query Expansion

The process of query expansion depends mainly on using the stems of features to expand the searching area. The stems for all features in the training and the AL-Hadith to be classified were produced in the preprocessing phase, and the stem of a feature is going to have the same weight of its origin feature.

Definition 2 illustrates the meaning of the expanded set for a training set or a query set, as shown in definition 2 the features of expanded training set will be the origin features from the training set union the set of all valid stems produced for that

training set using the light stemming mentioned in the preprocessing phase , the same goes for expanding the query set which will be the set of all features of the query Hadith (AL-Hadith to be classified) union the set of all valid stems for those features .

Definition 2: Expanded Set (training set (TS) or query set (QS)).

Let G is a training set or a query set :

Where $G = \{feature_1, feature_2, feature_3, \dots, feature_n\}$ where n is the number of features in TS or QS separately.

Let W is all valid stems of every feature belong to TS, QS

Where $W = \{stem_1, stem_2, stem_3, \dots, stem_n\}$ where n is the number of valid stems found in TS or QS separately

Let ES is the expanded set (training or query):

Where $ES = G \cup W$

After the expansion process the newly added stems in the expanded training set will have the same weights for its origin feature, in other words if we have the couple $\{(W,S), (W, TW)\}$ where S is the stem of word W and TW is the training weight for W from the training weights Table then the weight for stem S will be the same weight of W , the same procedure is applied to stems in expanding the query set where stem S will have the same weight of its origin word W from the query weight Table .

Expansion process widens the searching area by using the stems of features; in this case we are searching not only for the existence of the word but also for the existence of its stem, which will affect the classification performance. Figure 3.9 explains the weighting process for the expanded set.

Algorithm 3.3: algorithm for expanding process

Input: expanded training set (ETS), expanded query set (EQS), training set weights Table (TWT), query Hadith weights Table (QWT).

Output: Expanded Training Set Weights Table (ETWT), Expanded Query Set Weights Table (EQWT)

```

begin
For Each stem S in EQS Do
    begin
        stem S weight(SW) =weight of the feature in QWT which is the origin of stem S
    end
EQWT=QWT  $\cup$  { (S1,SW1), (S2,SW2),....., (Sn,SWn)} for n stems
end
begin
For each stem S in ETS
    Begin
        stem S weight(SW) = weight of the feature in TWT which is the origin of stem S
    end
ETWT=TWT  $\cup$  { (S1,SW1), (S2,SW2),....., (Sn,SWn)} for n stems
end

```

Figure 3.9: Algorithm for Weighting in Expanded Set.

For the example illustrated previously the expanded query Table will be as shown in Table 3.7, as illustrated in table 3.7 the stems { ايدي، دنيا، شعير } were added along with its weights to the original query feature .

Table 3.7: Extended Query Weights Table for Mined Hadith

Feature No.	Feature	IDF	Feature Redundancy	TF	QW
1	الدنيا	1.84	1.44	1	1.84
2	الشعير	2.52	0.30	1	2.52
3	الله	Feature redundancy >.45	80.02	2	0
4	ايديهم	2.22	0.61	1	2.22
5	يقوم	2.64	0.23	1	2.64
6	خيز	2.34	0.45	1	2.34
7	خرج	1.53	2.95	1	1.53
8	شاة	1.97	1.06	1	1.97
9	صلى	Feature redundancy >.45	68.89	1	0
10	فابى	2.42	0.38	1	2.42
11	فدعوه	3.12	0.08	1	3.12
12	مر	2.12	0.76	1	2.12
13	مصلىة	3.12	0.08	1	3.12
14	وسلم	Feature redundancy >.45	68.58	1	0
15	ياكل	1.80	1.59	1	1.80
16	يشبع	2.64	0.23	1	2.64
17	دنيا	Feature No. 1			1.84
18	شعير	Feature No. 2			2.52
19	ايدي	Feature No. 4			2.22

3.4.3 Constructing the Similarity Coefficient Table

In our system we use the cosine similarity coefficient where the similarity between two documents (document (D) & query (Q)) is actually the cosine of the angle (in N-dimensions) between the 2 vectors and can be calculated as follows (Baarah, 2007):

$$\text{sim}(D, Q) = \sum_{i=1}^n (w_{di} \times w_{qi})$$

Where i denote the query feature and n is the number of feature in the query Hadith.

Figure 3.10 explains the process of constructing the similarity coefficient Table. Table 3.8 shows the calculation and the entries in the similarity coefficient Table, where the columns represent the predefined book (classes), and the rows represent the mined Hadith features after expanding process. Last row represents the cumulative weights for all features for the mined Hadith, and table 3.9 shows similarity coefficient for features in the example Hadith illustrated previously against the training set from the book of food shown in section 3.3.2.

Algorithm 3.4: the algorithm for constructing the similarity coefficient Table

Input: extended training sets (ETS), extended query set (EQS), extended training set weights Table(ETWT), extended query set weights Table(EQWT).

Out put: similarity coefficient Table

For Each H_b in EQS Do

Begin

For each ETS J

Begin

Cumulative_similarity (ETS_J, H_b)=0

For Each word I in EQS b Do

Begin

$SIM(TS_J, EQS_b, Word_I) = T_J W_I * Q_b W_I$

Cumulative_similarity (ETS_J, H_b)= Cumulative_similarity (ETS_J, H_b)+SIM(ETS_J, EQS_b, Word_I)

where $Q_b W_{I=0}$ if the word have no occurrence in H_b

end

end

end

Figure 3.10: Algorithm for Constructing the Similarity Coefficient Table.

Table 3.8: Similarity Coefficient Table for Hadith b.

Feature	Pre-defined Themes					
	Book1	Book2	Book3	Book13
Feature1	$Sim1=Q_bW_1 * T_1W_1$					$Q_bW_1 * T_{13}W_1$
Feature2	$Sim2=Q_bW_2 * T_1W_2$					$Q_bW_2 * T_{13}W_2$
Feature3	$Sim3=Q_bW_3 * T_1W_3$					$Q_bW_3 * T_{13}W_3$
Feature4	$Sim4=Q_bW_4 * T_1W_4$					$Q_bW_4 * T_{13}W_4$
Feature5	$Sim5=Q_bW_5 * T_1W_5$					$Q_bW_5 * T_{13}W_5$
Feature n	$Sim n=Q_bW_n * T_1W_n$					
	$\sum_{i=1}^n Sim$					

Table 3.9: similarity coefficient for features in the example Hadith .

Feature No.	Feature	The Book of Food Similarity coefficient
1	الدنيا	3.39
2	الشعير	12.69
3	الله	0.00
4	ايديهم	4.92
5	يقوم	0.00
6	خبز	21.93
7	خرج	0.00
8	شاة	1.97
9	صلى	0.00
10	فابي	2.42
11	فدعوه	0.00
12	مر	2.12
13	مصلية	0.00
14	وسلم	0.00
15	ياكل	1.80
16	يشبع	0.00
17	دنيا	3.39
18	شعير	12.69
19	ايدي	4.92
cumulative similarty		72.26

3.4 Assigning AL_Hadith to A class

After Constructing the Similarity Coefficient Table for AL-Hadith to be classified against the predefined classes we will find the cumulative similarity weights for mined Hadith against each of those classes, the a cumulative similarity values indicate how much there are a common features between AL_Hadith to be classified and the predefined books.

After finding the cumulative weight for the mined Hadith corresponding to each predefined book (class), AL-Hadith will be assigned to the book with the maximum cumulative weight .Since maximum cumulative weight is an indication of larger common features between the training set and the mined Hadith feature set.

3.5 Evaluation

The evaluation of proposed classification algorithm results will be based on Sahih AL-Bukhari classification for each tested Hadith, as mentioned before Sahih AL-Bukhari is considered as the most authenticated book for prophet Mohamed saying in the Islamic world .by comparing AL-Hadith class resulted from our algorithm with the actual class for AL-Hadith from Sahih AL-Bukhari we can evaluate the performance of our algorithm.

4. Discussion and Analysis of Experimental Results

In this part, we give an overview of AL-Hadith corpus content which we used to run our experiments on the proposed classifying algorithm .The detailed experiments conducted are also illustrated in this part.

Next we discuss the results and the measurements we used to evaluate our system's performance.

Finally, the final results of our system are compared with other's work in AL-Hadith classification field. This step is essential to evaluate the accuracy and the effectiveness of our methodology.

4.1 Content of AL-Hadith Corpus

AL-Hadith corpus used in running the experiments consists of 13 books (classes). Ahadith were taken from Sahih AL-Bukhari, the famous Hadith book that is well known all over the Islamic world and the most trusted Hadith book to researchers in this field,12 of those books were included in AL_Kabi(2007) study , we added the Book of the Virtues of the Prophet and His Companions in our experiment with 143 additional Ahadith.

Table 4.1 shows statistics information for each book along with its name in English and Arabic as it was used by AL-Bukhari in his Sahih. The testing corpus has 1321 Hadith distributed in 13 books as shown in Table 4.1.

As we mentioned in our experiments, each Hadith is represented as a vector consists of the set of AL-Hadith features. The number of the distinct words found in each class is shown also in Table 4.1.

Table 4.1: List of Books in AL-Hadith Corpus.

Book (Chapter)Name	اسم الكتاب	No. of documents	No. of distinct features after stop words removal
The Book of Faith	كتاب الايمان	38	938
The Book of Knowledge	كتاب العلم	76	1946
The Book of Praying	كتاب الصلاة	115	2137
The Book of Call to Praying	كتاب الأذان	38	574
The Book of the Eclipse Prayer	كتاب الكسوف	24	715
The Book of Almsgiving	كتاب الزكاة	91	2267
The Book of Good Manners	كتاب الأدب	225	5258
The Book of Fasting	كتاب الصوم	107	1905
The Book of medicine	كتاب الطب	92	1895
The Book of Food	كتاب الطعام	91	1894
The Book of Pilgrimage (Hajj)	كتاب الحج	231	4885
The Book of Grievance	كتاب المظالم	40	906
The Book of the Virtues of the Prophet and His Companions	كتاب المناقب	143	3410

4.2 Hardware and Software

Our proposed classification system is implemented using the Microsoft Visual Basic .Net 2003 and the Access data base to store the result of the preprocessing phase and through the classification steps. Visual basic .Net was used because of its characteristics to handle Unicode languages such as Arabic .Since AL-Hadith is written in Arabic language, the software proved suitable for our task.

All experiments were carried on 1.46 GHz, 504 MB of RAM laptop running under Microsoft Windows XP Professional 2002, Service Pack 2.

4.3 Discussion of the Experimental Results and Analysis

In this section, the results of applying our proposed algorithm on the experimented corpus, which contains 13 books (classes), will be introduced. First of all we should demonstrate three main subjects:

1. Explanation of the three main methods we applied in our experiments on AL-Hadith corpus.
2. The experiments specification conducted on AL-Hadith corpus for all methods.
3. The performance measures used to compare our classification method, and other methods.

4.3.1 Classification Methods Applied On AL-Hadith Corpus

In our study, it was important to try other methods used by researchers in AL-Hadith classification filed. We could find one paper in this field done by (AL-Kabi, et al., 2007). In this paper AL-Kabi did not mention an accurate description of AL-Hadith corpus or the stop words list used in their experiments. Therefore , it was necessarily to reimplement their classification algorithm on the corpus used in our study. The following is a detailed explanation of the three methods , that we have been implemented.

1. AL-Kabi method : this method was proposed by AL-Kabi and his Colleagues on AL-Hadith classification (AL-Kabi, et al., 2007) , their method based mainly on using the stems of Ahadith words to calculate the IDF , wieghting the training and query set

and the matching process through their system , Of course if there are no valid stems of the word then the word is considered as its own stem ,Figure 4.1 shows the classification algorithm used by AL_Kabi in his study .

```

Read AL-Hadith from a file
Filtering AL-Hadith by
  Begin
  Removing AL-Hadith Sanad.
  Removing stop words
  End
Initialize the Count
For each term in AL-Hadith
  Begin
  Count=Count+1
  find term stem
  end
For each book
begin
F=frequency of each stem.(F values stored in term frequency table )
  For each entry in term frequency table
  begin
  Compute feature redundancy=F / No. of books
  If feature redundancy>=45 or feature redundancy= 0 then
  mark the term as uninteresting
  Else
  Stem weight=term frequency *log2 (N\DF)
  End If
  // DF: number of documents that include the stem.
  // N: Number of documents in the corpus.
  End
Calculate cumulative weights as
for j=1 to # of books
Cumulative weight =  $\sum_{i=1}^{\text{count}} (w_{ij})$ 
Rank the results in descending order.
Result =book with the highest rank
Display the result.
end

```

Figure 4.1: AL-Kabi Algorithm

2. Word based classification (WBC): this method uses the words of the AL-Hadith after going through the preprocessing phase except, for the stemming stage. Words after preprocessing are used in the calculation of the IDF and the weighting process. Stems of the words are not used in this method neither in building the training database nor in applying the classification algorithm.

3. Stem expansion classification (SEC): this is a new method that we are proposing in our system. Here we are using words after going through the preprocessing stage including the stemming, words are used in the IDF and weighting calculation for both the training and query set, but the stems are used in query expanding process, the expansion process is discussed in details in section 3.4.2, and the whole classification system is explained in details in chapter 3.

4.3.2 Experiments Specifications

Our Hadith corpus consists of 1321 Hadith distributed over 13 books (classes). The system is considered as supervised classification since we use training sets to apply learning algorithm to build the leaning database which will be used for the classification algorithm.

In our experiments we used (90%) of each class Ahadith as training set while the rest (10%) of each class is used as testing set for our classification system. Of course for each class the (10%) Ahadith in the testing set are not included in the training set Ahadith and the training phase calculation .

For each class we choice five testing – training set combination to run our algorithm, that mean that for each class we will run five separable experiments, this give us variation for system testing.

It is important to mention that the same training set - testing set combination was used with the three classification methods mentioned in section 4.3.1. This was important to insure fair comparison among the different methods in our study.

4.3.3 Performance Measures

In order to articulate the efficiency of any classification algorithm we must have measurements, to compare our system's outcome with others. In our study, we used the most popular measurement in text classification algorithm: recall, precision and F-measure.

Recall and precision both based on the concept of relevance, precision is defined as the ratio of relevant documents retrieved to all documents retrieved while Recall is defined as the ratio of relevant items retrieved to all relevant items in a corpus.

There are obvious trade-off between recall and Precision, if the system retrieves all the documents in a corpus then the system will surely retrieve all the relevant documents in the corpus, in this case the recall will be perfect, on the other hand, since commonly there are only a small proportion of the documents in a collection are truly relevant to the given query, retrieving everything will give a very low precision (Greengrass, 2000).

Because of this trade-off between recall and precision, we need the best achievable combination of good precision and good recall, in the best case we would like to retrieve all the relevant documents and no non-relevant documents, this combination of recall and precision found in F-measure (Harmonic mean).

Table 4.2 shows the meaning of the parameter used in recall, precision calculation shown bellow (Al-Mesleh, 2007):

Precision (P) = $A / (A + B)$.

Recall(R) = $A / (A + C)$.

F-measure (Harmonic mean) = $(2 * P * R) / (P + R)$

Table 4.2: Recall and Precision Parameters.

System says...	<i>In reality, the document is...</i>	
	<i>Relevant</i>	<i>Irrelevant</i>
document is relevant	A	B
document is irrelevant	C	D

4.4 Results Analysis and Comparisons:

In this section we introduce the result of our experiments on the Hadith corpus for the three method mentioned before in section 4.3.1, then a comparisons between the three methods is conducted in order to show the preference of our classification algorithm.

4.4.1 Results of Our Proposed Method

As mentioned before we have conducted five trails on each class .In each trail we have chosen different 10% of the class documents as testing set to test our classification system.

For example, for the book of food we have 91 Ahadith in our corpus belong to this class as shown in Table 4.1, so the 10% of those documents will be 9 documents for each trail, in each trail we choice different training – testing combination from the 91 Ahadith we have .

In the experiments stage for each trail we added new 9 Ahadith to the query process, that mean that in the first 20% (first trail) of the query process we have 9 Ahadith belong to book of food and when the query process reach to 40 % (second trail) we have 18 Ahadith belong to the book of food .

Table 4.3 shows the query progress results for the book of food, where positive retrieved documents are Ahadith correctly classified to book of food when reaching the specified progress percentage while negative retrieved documents are Ahadith wrongly classified to book of food while they belong in reality to other books.

Table 4.3: Our method Recall and Precision for Book of Food during Query Progress Stages

Query Progress	positive Retrieved Documents	Negative Retrieved Documents	Recall	precision
20%	6	2	0.66	0.75
40%	5	1	0.61	0.79
60%	6	2	0.63	0.77
80%	5	1	0.61	0.79
100%	4	3	0.58	0.74

Calculating recall and precision is an accumulative process. For example, to calculate the recall and precision at 20% and 40% of query progress (depending on the formulas shown in section 4.3.3), provided that at 20% we have 9 documents belong to

book food in real, while at 40% we added new 9 documents to the query process, so the total documents belong in reality to the book of food at 40% will be 18 documents.

When reaching 20% of query progress recall= (6/9) and precision= (6/8).

When reaching 40% of query progress recall= (11/18) and precision= (11/14).

And so at the end of query progress (reaching 100%) we have 45 documents belong to book of food with recall = (26/45) and precision = (26/35).

Figure 4.2 shows the recall and precision values for the 5 progress phases for the book of food.

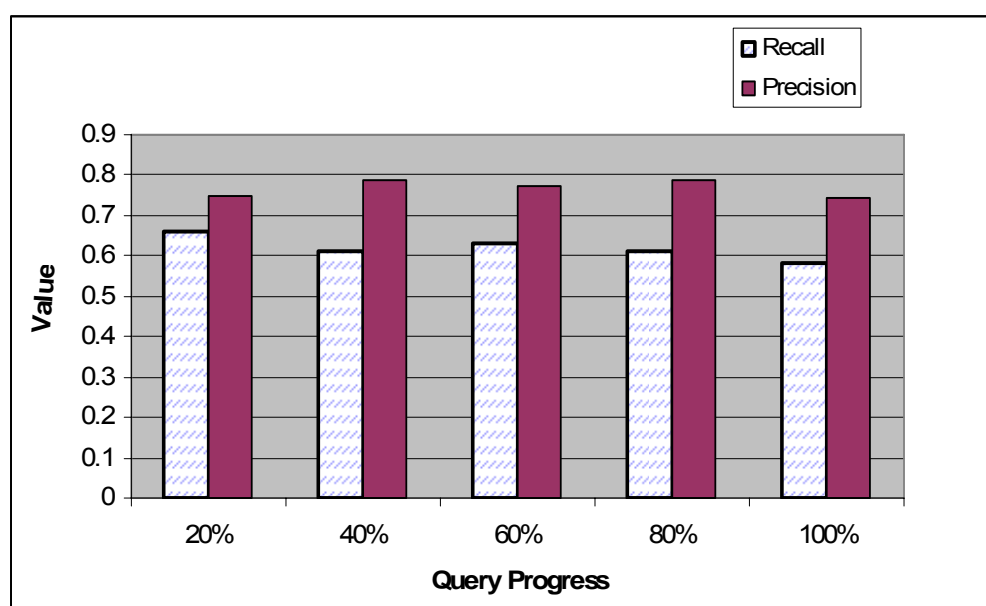


Figure 4.2: Stem Expansion Method Results during Query Progress Phases.

Our proposed system which is based on stem expansion classification method was implemented and tested with the same testing and training set used in WBC and AL-Kabi to insure the fairness of the comparisons between our work and other two methods. Table 4.4 shows our method results when reaching 100% of query progress.

As shown in the results the best precision achieved by our method was for The Book of the Eclipse Prayer (.94) and the best recall was with same value also for the same book, Figure 4.3 shows recall and precision values for each book.

Table 4.4: Stem Expansion Classification Results.

Book No	Book Name	Tested Ahadith No	Precision	Recall	F- measure
1	The Book of Food	45	0.74	0.58	0.65
2	The Book of Praying	48	0.60	0.65	0.62
3	The Book of Knowledge	38	0.43	0.32	0.36
4	The Book of Good Manners	99	0.39	0.61	0.47
5	The Book of Call to Praying	17	1.00	0.41	0.58
6	The Book of Faith	29	0.80	0.28	0.41
7	The Book of Almsgiving	45	0.63	0.67	0.65
8	The Book of Pilgrimage (Haj)	133	0.67	0.83	0.74
9	The Book of Fasting	50	0.78	0.78	0.78
10	The Book of medicine	53	0.90	0.53	0.67
11	The Book of the Eclipse Prayer	18	0.94	0.94	0.94
12	The Book of Grievance	23	0.59	0.43	0.50
13	The Book of the Virtues of the Prophet and His Companions	67	0.46	0.45	0.45

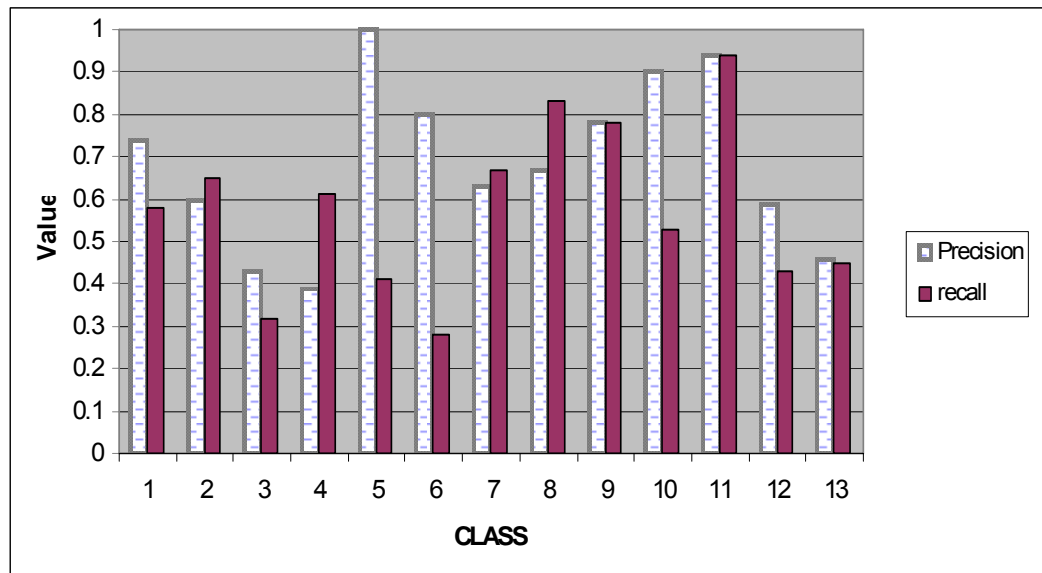


Figure 4.3: Stem Expansion Classification Results.

4.4.2 Results of Word based classification (WBC)

As mentioned previously this method depends on matching words only without using the stems of the words. The results of this method are shown in Table 4.5.

Table 4.5: Word Based Classification Results.

Book No	Book Name	Tested Ahadith No	Precision	Recall	F- measure
1	The Book of Food	45	0.64	0.47	0.54
2	The Book of Praying	48	0.43	0.48	0.45
3	The Book of Knowledge	38	0.48	0.29	0.36
4	The Book of Good Manners	99	0.25	0.37	0.30
5	The Book of Call to Praying	17	0.80	0.24	0.36
6	The Book of Faith	29	0.43	0.21	0.28
7	The Book of Almsgiving	45	0.46	0.42	0.44
8	The Book of Pilgrimage (Haj)	133	0.57	0.79	0.66
9	The Book of Fasting	50	0.62	0.62	0.62
10	The Book of medicine	53	0.81	0.49	0.61
11	The Book of the Eclipse Prayer	18	0.73	0.89	0.80
12	The Book of Grievance	23	0.80	0.35	0.48
13	The Book of the Virtues of the Prophet and His Companions	67	0.18	0.18	0.18

Figure 4.4 shows the recall and precision values at this experiment .as shown in Figure 4.4 the WBC achieved the best values at recall for The Book of the Eclipse Prayer with (0.89), where the best precision value was for The Book of medicine with (.81).

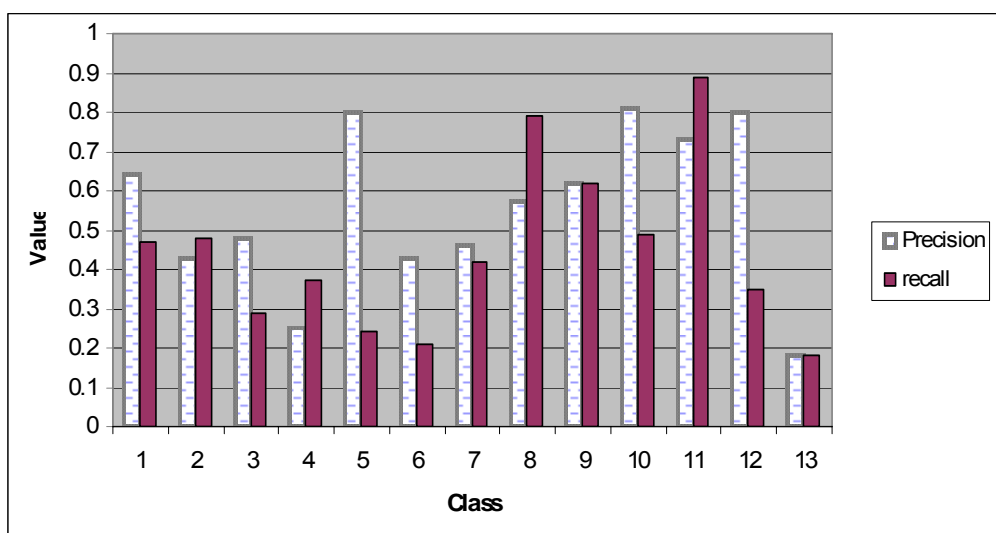


Figure 4.4: Word Based Classification Results.

4.4.3 Results of AL-Kabi Method

Next we implemented the classification algorithm proposed by AL-Kabi (AL-Kabi, et al., 2007) which was the only work in AL-Hadith classification field we found after extensive searching over the web. We ran their algorithm on the same training-testing set combination which has been used to get the result of our method and Word based classification shown in Tables 4.4 and 4.5 respectively.

As shown in Table 4.6 AL-Kabi's method achieved its best precision value for The Book of Grievance (0.83) while its best recall value was (0.77) for the Book of Pilgrimage (Haj) . Figure 4.5 shows the recall and precision values for reapplying this method.

Table 4.6: Result of AL-Kabi's Classification Method.

Book No	Book Name	Tested Ahadith No	Precision	Recall	F- measure
1	The Book of Food	45	0.42	0.18	0.25
2	The Book of Praying	48	0.30	0.06	0.10
3	The Book of Knowledge	38	0.62	0.21	0.31
4	The Book of Good Manners	99	0.09	0.12	0.10
5	The Book of Call to Praying	17	0.50	0.24	0.32
6	The Book of Faith	29	0.27	0.10	0.15
7	The Book of Almsgiving	45	0.28	0.11	0.16
8	The Book of Pilgrimage (Haj)	133	0.19	0.77	0.30
9	The Book of Fasting	50	0.43	0.20	0.27
10	The Book of medicine	53	0.48	0.21	0.29
11	The Book of the Eclipse Prayer	18	0.50	0.17	0.25
12	The Book of Grievance	23	0.83	0.22	0.34
13	The Book of the Virtues of the Prophet and His Companions	67	0.17	0.04	0.07

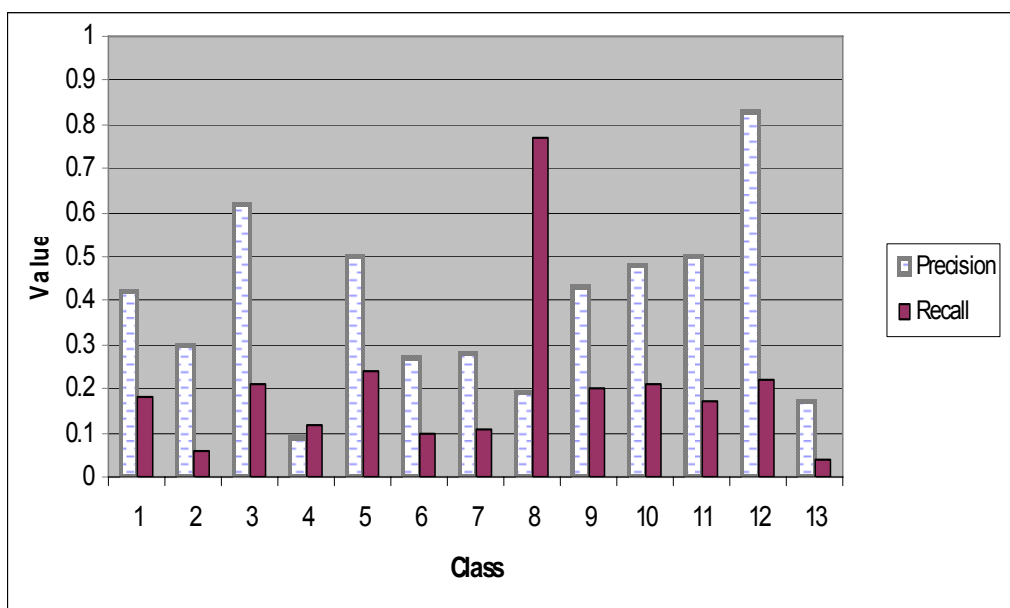


Figure 4.5: Result of AL-Kabi's Classification Method.

4.4.4 Comparisons and Results Analysis

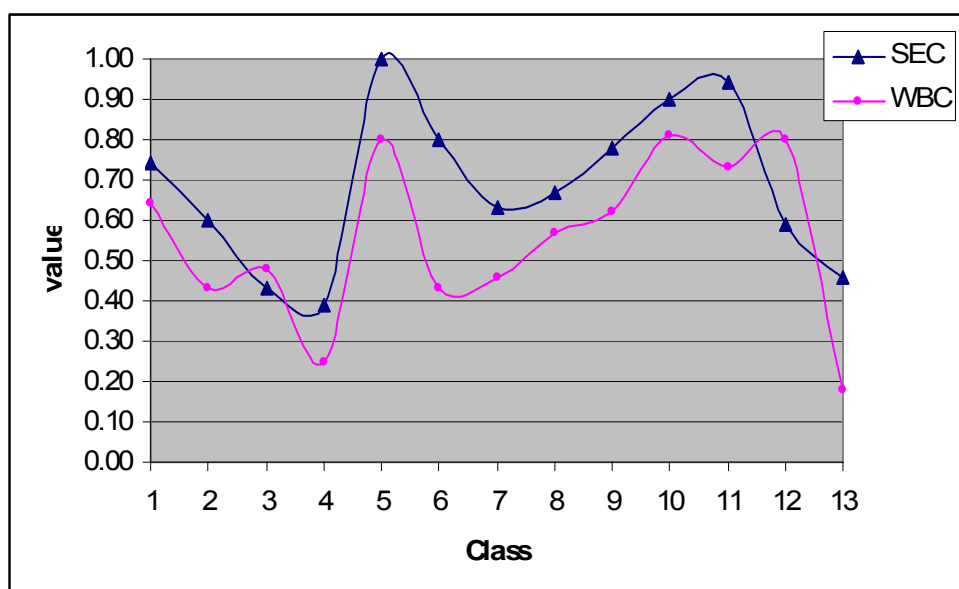
In this section we are introducing the comparisons between our proposed system and the other two methods, in order to show the preference of our work over the other methods.

4.4.4.1 Word based classification vs. Stem Expansion classification.

Figure 4.6 shows Precision curves, as shown our system (SEC) outperform WBC for 11 out of 13 books, WBS achieve better precision for The Book of Grievance and the book of Knowledge, this result is predicted since using the stem expansion phase gives a large morphology variation for the words which can result in retrieving more documents that belong in reality to other classes.

On the other hand our SEC overcome this side effect of expanding the query using stemming , by the weighting strategy we adopt in our system, since stems in the expansion phase were giving the same weights of its original word in AL-Hadith .

SEC achieved precision value of 1 for The Book of Call to Praying and enhances the precision of 11 classes by 45% in average.



**Figure 4.6: Word based classification vs. Stem expansion classification
Precision Comparison.**

F_Measure (harmonic measure) combine the recall and precision in one measure, the best classification method will be the one that achieve the best harmony between the recall and the precision.

Figure 4.7 shows comparison between SEC and the WBC, SEC achieved better F_Measure values for all classes against WBS method and enhances the F_Measure by 49% in average.

Since our SEC achieves better F_Measure, it means that SEC has a better harmony between the recall and the precision performance scales, which is considered a good indicator for the performance of the SEC as a classification method.

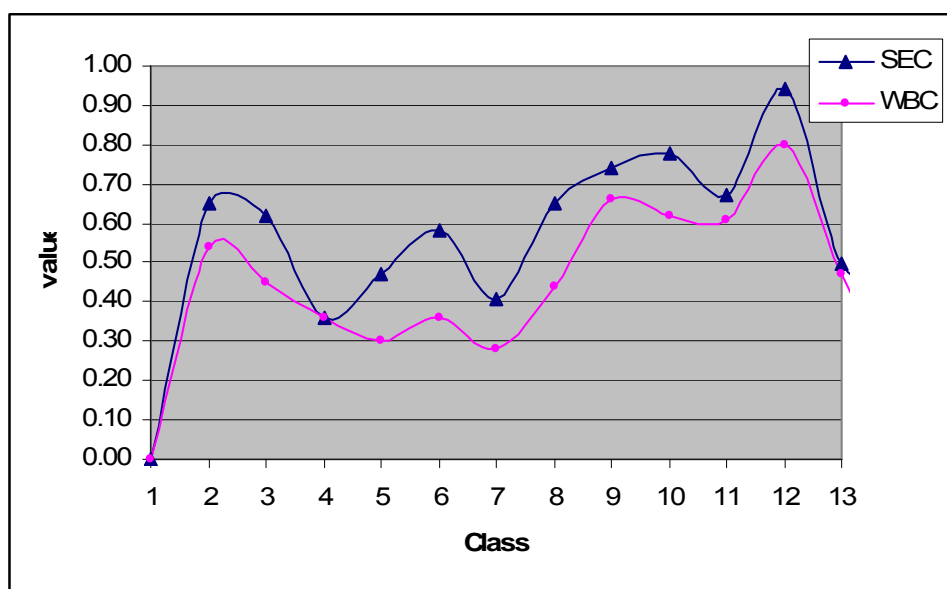


Figure 4.7: Word based classification vs. Stem expansion classification F_Measure Comparison.

4.4.4.2 AL-Kabi vs. Stem expansion classification:

After implementing method proposed by AL-Kabi and his colleagues for the same training and testing sets used to examine our proposed method, we conduct the following comparisons.

Figure 4.8 shows the precision graph of the two methods. AL-Kabi's method achieved better precision for The Book of Knowledge and The Book of Grievance, while our method out performed AL-Kabi's method in 11 out of 13 classes.

That can be justified since those two book have small number of Ahadith since our experiments is conducted on Ahadith in closed domain (Sahih AL_Bukhari) we believe if those classes have bigger number of Ahadith the superb of our method will appear , in addition AL-Kabi used the stem of words for term weighting in the preprocessing phase ,that mean term frequency used in the training process will be the number of the stems occurrences on the training set, in contrary in our classification method we use the words occurrences for the weighting process while the stems were used for the expansion process in training and testing sets, where the stems were given the same weight for the original word.

Since we use stem expansion for both the training and testing set, that will cause more non -related documents to be retrieved, that justify that AL-Kabi's method achieved better precision for two classes.

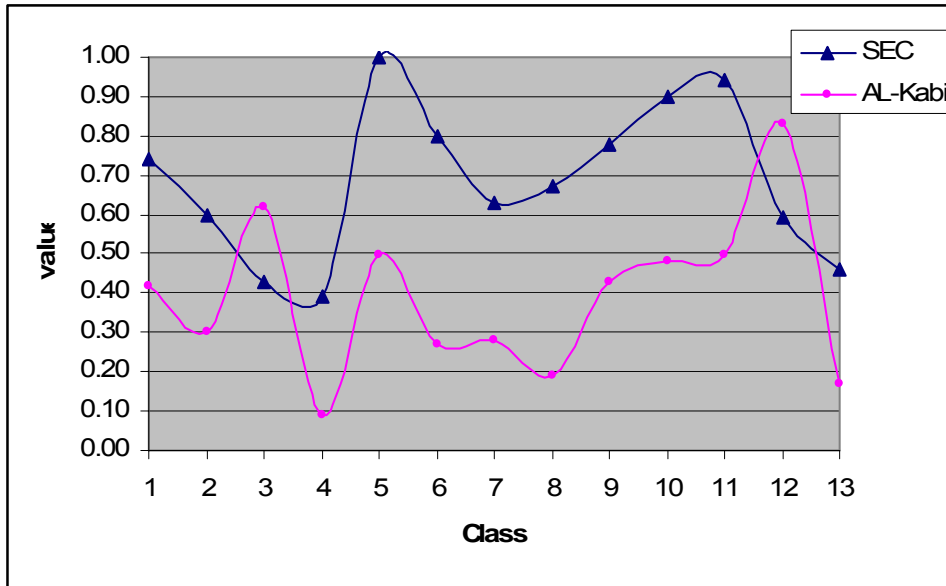


Figure 4.8: AL-Kabi's classification vs. Stem expansion classification Precision Comparison.

As mentioned before F_Measure is a good indicator about the recall and precision harmony, as shown in Figure 4.9 our method outperformed AL-Kabi F_Measure in all the 13 classes.

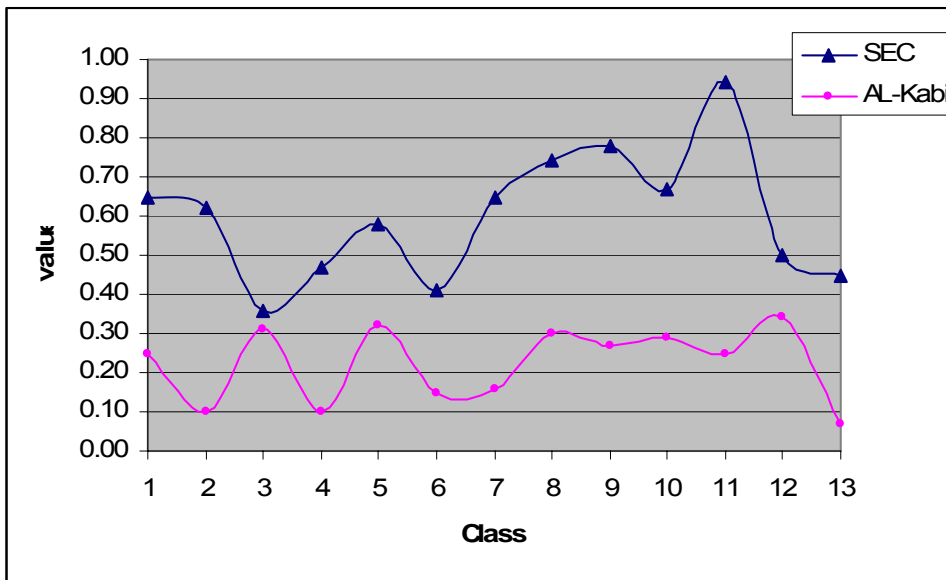


Figure 4.9: AL-Kabi's classification vs. Stem Expansion Classification F_Measure Comparison.

5. Conclusions and Future Works

5.1 Conclusions

In this thesis, we proposed classification method Stem Expansion Classification in order to discover knowledge from AL-Hadith by assign each Hadith to one book (class) of predefined classes. The Stem Expansion Classification is classified as supervised classification method since we use training data set to train our classifier.

We had selected a corpus containing 1321 Hadith from 13 books from Sahih AL-Bukhari and each Hadith assigned to one class, we use Sahih AL-Bukhari as our base line for the correctness of each Hadith classification.

The Final results of the Stem Expansion Classification were compared with the results of the two other methods; one proposed by AL-Kabi and the other is word based classification technique (WBC). The final results show that our method was better than WBC in recall for all classes. WBC achieved better precision only for two class but for the rest 11 classes Stem Expansion Classification achieved better precision and our Stem Expansion Classification had better harmony measure (F_Measure) for all the 13 classes .

The final results show that our Stem Expansion Classification had a better recall for all classes against AL-Kabi method. AL-Kabi got a better precision for two classes while Stem Expansion Classification had better performance for the rest 11 books. Stem

Expansion Classification achieved a better harmony measure (F_Measure) for all the 13 books.

We recommended, in this section, the following:

1. The results of Stem Expansion Classification and its performance (recall, precision, F_Measure) were satisfactory comparing to the results of other existing methods.
2. Discovering the class of Hadith, depends on training the classifier by using selected training sets for each predefined class is an acceptable approach.
3. Expansion process used in our method is proven to be a good modification in the classification method used since it keeps the F_Measure in its highest values comparing with the existing methods used for classifying AL-Hadith.

5.2 Future Work

Discovering the knowledge from AL-Hadith is till now a hot research. As a feature work we recommended the following:

1. Increasing the number of general classes included in the study.
2. Study the multi-subject classification for AL-Hadith.
3. enhancing the classification efforts by using the thesaurus of features
4. Experiments the effectiveness of using manual labeling for each class features against the automatic labeling we used in this study.
5. Building a comprehensive corpus containing the whole Sahih-Al-Bukhari Ahadith, This requires a great time and effort.

REFERENCES

- AL-Kabi, M. N., AL-Sinjilawi, S. I. (2007). "A Comparative Study of the Efficiency of Different Measures to Classify Arabic Text", **University of Sharjah Journal of Pure and Applied Sciences**, 4(2), pp. 13-26.
- AL-Kabi, M. N., and AL-Shalabi (2005), "AL-Hadith Text Classifier", **Journal of applied sciences**, 5(3), pp.548-587.
- Baarah, A.H (2007). "**Applying Text Mining Algorithms For Knowledge Discovery On The Holy Quran Text**". Master's Degree thesis (not published), The University of Jordan, Amman, Jordan.
- AL-Mesleh, A. M. (2007). "CHI Square Feature Extraction Based SVMs Arabic Language Text Categorization System", **Journal of Computer Science**, 3(6), pp.430-435.
- AL-Shalabi, R., Kanaan, G., and Gharaibeh, M. H. (2006). "Arabic Text Categorization Using kNN Algorithm", **Proceedings of the 4th International Multiconference on Computer Science and Information Technology**, 4, Amman, Jordan, 5-7 April.
- Al-Serhan, H., Al Shalabi, R. and Kannan, G. (2003). "New Approach for Extracting Arabic roots". **Proceedings of the 2003 Arab conference on Information Technology (ACIT'2003)**, pp 42-59.
- Bellot P., Crestan E., El-Bèze M., Gillard L., de Loupy C. (2003), "Coupling Named Entity Recognition, Vector-Space Model and Knowledge Bases for TREC-11 Question Answering Track", **In Proceedings of The Eleventh Text retrieval Conference (TREC 2002)**, NIST Special Publication.
- Bellot, P., El-Bèze, M., (2000), "Clustering by means of decision trees without learning or hierarchical and K-Means like algorithms", **in Proceedings of RIAO'2000**, Paris, France, pp. 344-363.
- Chu M. (2008), "Data Mining and Applied Linear Algebra", **International Conference on Informatics Education and Research for Knowledge-Circulating Society**, pp. 20-25.

Duwairi, R. M. (2006). "Machine Learning for Arabic Text Categorization". **Journal of the American Society for Information Science and Technology**, 57(8), pp.1005-1010.

El-Halees, A. (2006). "Mining Arabic Association Rules for Text Classification", **Proceedings of the First International Conference on Mathematical Sciences**, AL-Azhar University of Gaza, Palestine, 15-17 July.

El-Kourdi, M., Bensaid, A., and Rachidi, T. (2004). "Automatic Arabic Document Categorization Based on the Naïve Bayes Algorithm", **Workshop on Computational Approaches to Arabic Script-Based language (COLING-2004)**, University of Geneva, Geneva, Switzerland, pp. 51-58 .

Forman, G. (2003). "An Extensive Empirical Study of Feature Selection Metrics for Text Classification". **Journal of Machine Learning Research** 3 (2003), pp 1289-1305.

Greengrass Ed. (2000), . Information Retrieval: A Survey, 2000.

Hammo, B., Abu-Salem, H., Lytinen, S., and Evens, M. (2002). "A Question Answering System to Support the Arabic Language", **Workshop on Computational Approaches to Semitic Languages**, Philadelphia, PA, July, pp.55-65.

Hammo, B., Sleit, A., El- Haj, M. (2008), "Enhancing Retrieval Effectiveness of Diacritized Arabic Passages Using Stemmer and Thesaurus", **The 19th Midwest Artificial Intelligence and Cognitive Science Conference (MAICS2008)**.

Ikonomakis, M., Kotsiantis, S., Tampakas, V. (2005), "Text Classification Using Machine Learning Techniques", **wseas transactions on computers**, Issue 8, Volume 4, August 2005, pp. 966-974

Isa, D., Lee, L.H., Kallimani, V.P., RajKumar, R. (2008), "Text Document pre-processing with the Bayes formula for Classification using the Support Vector Machine", **IEEE Transactions on Knowledge and Data engineering**, Volume 20, Issue 9, pp. 1264 - 1272 .

Kanaan, G., AL-Kabi, M. N., and AL-Shalabi, R. ,2005 , "Statistical Classifier of the Holy Quran Verses (Fatiha and YaSeen Chapters)", **Journal of Applied Science**, 5(3), pp.580-583.

Khoja, S. (2001), "APT: Arabic Part-of-Speech Tagger", **Proceedings of the Student Workshop at the Second Meeting of North American Chapter of the Association for Computational Linguistics (NAACL-2001)**, Carnegie Mellon University, Pennsylvania, USA, pp 20–25.

Khreisat, L. (2006). "Arabic Text classification Using N-gram Frequency Statistics A Comparative study", **Proceedings of the 2006 International Conference on Data Mining (DMIN 2006)**, Las vegas, USA, pp. 78-82.

Krebel U. G. (1999). "Pairwise Classification and Support Vector Machines. Advances in Kernel Methods: Support Vector Learning". Cambridge, MA: **MIT press**, pp.255-268.

Kroeze, J., Matthee, M., and Bothma T. (2003), "Differentiating Data- and Text-Mining Featureinology", **University of Pretoria, Proceedings of SAICSIT 2003**, pp. 93 –101

Lam, S. L. and Lee, D. L. (1999). "Feature Reduction for Neural Network Based Text Categorization", **Proceedings of the Sixth IEEE International Conference on Database Advanced Systems for Advanced Application**, Hsinchu, Taiwan, pp.195-202.

Liu, B., Li, X., Lee, W. S., and Philip, S. Yu³ (2004), "Text Classification by Labeling Words", in **Proceedings of Nineteenth National Conference on Artificial Intelligence (AAAI-2004)**, San Jose, USA, pp.425-430.

Lukui S., Jun Z., Enhai L. Pilian H. (2007), "Text Classification Based on Nonlinear Dimensionality Reduction Techniques and Support Vector Machines", **Third International Conference on Natural Computation (ICNC 2007)**, pp. 674-677.

Saidah S., Naomi S. (2007), "Build Islamic Ontology based on Ontology Learning", **Universiti Teknologi Malaysia**, Postgraduate Annual Research Seminar.

Saleem A., Martha E. (2004), "Event Extraction and Classification for Arabic Information Retrieval Systems", **IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2004)**

Sawaf, H., Zaplo, J., and Ney, H. (2001). "Statistical Classification Methods for Arabic News Articles", **Arabic Natural Language Processing Workshop**, Toulouse, France.

Thabet, N. (2005). "Understanding the Thematic Structure of the Qur'an: an Exploratory Multivariate Approach", **Proceeding of the ACL Student Workshop**, Michigan, USA, June 2005, 7-12.

Yu-ping, Q., Qing A., Xiu-kun, W., Xiang-na, L. (2007), "Study on Classification Algorithm of Multi-subject Text", **ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing**, pp. 435-438.

Zhang J. (2004). "The Optimality of Naïve Bayes ", **Proceedings of the 17th International Florida Artificial Intelligence Research Society Conferences**, Florida, USA , pp.562-567.

Robinson C. F. (2003). "Islamic Historiography". Cambridge University Press. ISBN 0521629365.

Appendix A

Stop Words list

ومع	لي	في	به	آل	ا
ومن	ليس	في	بها	ألو	ابدا
وَمَنْ	م	فيما	بهذه	أ	ابدا
وهذا	ما	فيه	بين	أبدأ	ابدا
وهو	مَا	فيها	تلك	أبدأ	اتي
وهو	مَا	فِيهَا	تم	أبدأ	احياناً
وهي	ماذا	قَالَ	ثم	أتى	احياناً
يَا	مع	قبل	ثمَّ	أحياناً	اخر
يأتي	معهُ	قد	جداً	أحياناً	اذا
يأتي	مما	قُلْ	جدا	أذن	إذا
يتم	من	كان	حتى	أكثر	اذن
يكن	من	كَانَ	حيث	ألا	اكثر
يمكن	مِنْ	كانت	حينما	أما	الا
ابة بكر	مَنْ	كانوا	خلال	أمام	الا
عمر	مَنْ	كل	د	أن	إلا
الخطاب	مِنْ	كل	ذلك	أنت	إنا
ابن	مَنْ	كم	ذلكَ	أنتم	الآن
ابو	منهُ	كما	شيء	أنتما	التي
ابي	منها	كنا	شيءٍ	أنتن	الذي
علي	منها	كنت	على	أننا	الذي
طالب	منهم	كي	على	أنني	الَّذِينَ
عثمان	منهن	كيف	عليه	أنه	او
عفان	هذا	لا	عليها	أنها	أو
فاطمه	هذه	لا	عن	أنهما	وما
سلمه	هل	لا	عن	أهم	وَمَا
حمزه	هم	لأبد	عند	أو	لو
محمد	هن	لان	عندما	أي	لهم
مكه	هنا	لانه	عنه	أيهما	الى
قال	هو	لانها	غير	إذا	إلى
ولكن	هُوَ	لاي	فان	إذا	اليه
ولم	هي	لأ	فأن	إنن	اما
الاوزعي	و	لأن	فان	إلا	امام
اي	و	لأنه	ففي	إلى	ان
ايهما	والتي	لأنها	فقد	إليه	إنَّ
أتِ	والتي	لأي	فلا	إما	أن
أتى	والذي	لإن	فلايد	إن	إن
أثية	وان	لإنها	فلم	إننت	اننت
آخر	وأن	لدى	فلماذا	إننا	انتم
بكل	وإن	لذا	فلن	إنني	انتما
بل	وإنما	لعل	فما	إنه	انتن
بلا	وفي	لكا	فماذا	إنها	اننا
بمّا	وقد	لكل	فمن	إنهما	انني
بن	وكان	لكم	فنحن	بان	انه
به	وكيف	لم	فهذا	بانهُ	انها
فهما	ولا	لمماذا	فهذه	بأن	انهما
فهمت	ولا	لنا	فهكذا	بأنهُ	اهم
فهن	ولّا	له	فهل	بانهُ	لَهُمْ
فهناك	ولأ	لهُ	فههم	بعد	لهن
فهو	ولعل	لها	فههي	ولعلها	لهذا

Appendix B

IDF Matrix for Al-Hadith Corpus

Feature No	feature	DF	IDF	Feature redundancy	Feature No	feature	DF	IDF	Feature redundancy
1	الله	1057	0.10	80.02	49	الانصار	26	1.71	1.97
2	صلى	910	0.16	68.89	50	راى	26	1.71	1.97
3	وسلم	906	0.16	68.58	51	قريش	25	1.72	1.89
4	رجل	111	1.08	8.40	52	الحليفة	25	1.72	1.89
5	الصلاة	99	1.13	7.49	53	يديه	25	1.72	1.89
6	اهل	76	1.24	5.75	54	واحد	25	1.72	1.89
7	الرجل	68	1.29	5.15	55	صدقة	24	1.74	1.82
8	رايت	68	1.29	5.15	56	القيامة	24	1.74	1.82
9	الشمس	63	1.32	4.77	57	الاسلام	24	1.74	1.82
10	رجلا	53	1.40	4.01	58	بشيء	24	1.74	1.82
11	جاء	49	1.43	3.71	59	الارض	24	1.74	1.82
12	قام	47	1.45	3.56	60	الساعة	23	1.76	1.74
13	صلاة	47	1.45	3.56	61	نبي	23	1.76	1.74
14	يصلى	45	1.47	3.41	62	فجاء	23	1.76	1.74
15	دخل	44	1.48	3.33	63	اصحابه	23	1.76	1.74
16	فصلى	44	1.48	3.33	64	عهد	23	1.76	1.74
17	النار	41	1.51	3.10	65	تابعه	23	1.76	1.74
18	امر	40	1.52	3.03	66	فخرج	23	1.76	1.74
19	خير	40	1.52	3.03	67	والمروة	22	1.78	1.67
20	العصر	40	1.52	3.03	68	النحر	22	1.78	1.67
21	خرج	39	1.53	2.95	69	الظهر	22	1.78	1.67
22	رمضان	38	1.54	2.88	70	نهى	22	1.78	1.67
23	قدم	36	1.56	2.73	71	الآخر	21	1.80	1.59
24	بالبيت	35	1.58	2.65	72	النساء	21	1.80	1.59
25	امراة	35	1.58	2.65	73	ياكل	21	1.80	1.59
26	ركعتين	35	1.58	2.65	74	عرفت	21	1.80	1.59
27	يعني	35	1.58	2.65	75	خمس	21	1.80	1.59
28	ليلة	34	1.59	2.57	76	ثلاث	20	1.82	1.51
29	المسجد	33	1.60	2.50	77	يجد	20	1.82	1.51
30	فجعل	33	1.60	2.50	78	راسه	20	1.82	1.51
31	الليل	32	1.62	2.42	79	الماء	20	1.82	1.51
32	احب	32	1.62	2.42	80	السلام	20	1.82	1.51
33	الحج	31	1.63	2.35	81	اله	20	1.82	1.51
34	وجهه	31	1.63	2.35	82	فامر	20	1.82	1.51
35	البيت	30	1.64	2.27	83	الصدقة	20	1.82	1.51
36	الجنة	30	1.64	2.27	84	الفجر	20	1.82	1.51
37	الرحمن	30	1.64	2.27	85	اكل	20	1.82	1.51
38	فقام	30	1.64	2.27	86	سبيل	20	1.82	1.51
39	ثلاثة	30	1.64	2.27	87	الصبح	20	1.82	1.51
40	بيده	30	1.64	2.27	88	الصفا	19	1.84	1.44
41	اهله	30	1.64	2.27	89	يحل	19	1.84	1.44
42	القوم	28	1.67	2.12	90	يطوف	19	1.84	1.44
43	اعلم	28	1.67	2.12	91	ايام	19	1.84	1.44
44	شاء	28	1.67	2.12	92	المرأة	19	1.84	1.44
45	يده	28	1.67	2.12	93	يخرج	19	1.84	1.44
46	اليوم	27	1.69	2.04	94	سال	19	1.84	1.44
47	يوما	27	1.69	2.04	95	الدنيا	19	1.84	1.44
48	فلان	26	1.71	1.97	96	المغرب	19	1.84	1.44

Feature No	feature	DF	IDF	Feature redundancy	Feature No	feature	DF	IDF	Feature redundancy
97	قوم	19	1.84	1.44	147	فدعا	15	1.94	1.14
98	قط	19	1.84	1.44	148	طاف	15	1.94	1.14
99	صائم	19	1.84	1.44	149	جهنم	15	1.94	1.14
100	الزكاة	19	1.84	1.44	150	ابيه	15	1.94	1.14
101	الحديث	18	1.87	1.36	151	المؤمنين	15	1.94	1.14
102	شهر	18	1.87	1.36	152	صوم	14	1.97	1.06
103	خرجنا	18	1.87	1.36	153	فانن	14	1.97	1.06
104	سمع	18	1.87	1.36	154	رجلان	14	1.97	1.06
105	اربع	18	1.87	1.36	155	شاة	14	1.97	1.06
106	نزل	18	1.87	1.36	156	فانطلق	14	1.97	1.06
107	الاية	17	1.89	1.29	157	خيرا	14	1.97	1.06
108	مسلم	17	1.89	1.29	158	نفسى	14	1.97	1.06
109	يريد	17	1.89	1.29	159	بعمره	14	1.97	1.06
110	الاول	17	1.89	1.29	160	الشيطان	14	1.97	1.06
111	منى	17	1.89	1.29	161	اصحاب	14	1.97	1.06
112	رفع	17	1.89	1.29	162	فعل	14	1.97	1.06
113	فدخل	17	1.89	1.29	163	لموت	14	1.97	1.06
114	انصرف	17	1.89	1.29	164	فاكل	14	1.97	1.06
115	الشهر	17	1.89	1.29	165	يدي	14	1.97	1.06
116	فزل	17	1.89	1.29	166	الجاهلية	14	1.97	1.06
117	الهدى	16	1.92	1.21	167	شعبة	14	1.97	1.06
118	ذكر	16	1.92	1.21	168	انظر	14	1.97	1.06
119	دعا	16	1.92	1.21	169	غيره	14	1.97	1.06
120	اية	16	1.92	1.21	170	الكعبة	14	1.97	1.06
121	قول	16	1.92	1.21	171	يجب	14	1.97	1.06
122	ثوب	16	1.92	1.21	172	انزل	13	2.01	0.98
123	فجعلت	16	1.92	1.21	173	عشرة	13	2.01	0.98
124	ماء	16	1.92	1.21	174	بلغت	13	2.01	0.98
125	العشاء	16	1.92	1.21	175	احدا	13	2.01	0.98
126	المسلم	16	1.92	1.21	176	القران	13	2.01	0.98
127	فطاف	16	1.92	1.21	177	عام	13	2.01	0.98
128	سالت	16	1.92	1.21	178	بيت	13	2.01	0.98
129	السماء	15	1.94	1.14	179	بغير	13	2.01	0.98
130	اراد	15	1.94	1.14	180	المؤمن	13	2.01	0.98
131	الايمان	15	1.94	1.14	181	اتى	13	2.01	0.98
132	اخبرني	15	1.94	1.14	182	اقبل	13	2.01	0.98
133	يدخل	15	1.94	1.14	183	عمرة	13	2.01	0.98
134	راسي	15	1.94	1.14	184	حرم	13	2.01	0.98
135	ثلاثا	15	1.94	1.14	185	نزلت	13	2.01	0.98
136	والعصر	15	1.94	1.14	186	يحدث	13	2.01	0.98
137	حج	15	1.94	1.14	187	حاجة	13	2.01	0.98
138	صدق	15	1.94	1.14	188	رب	13	2.01	0.98
139	التهار	15	1.94	1.14	189	يسال	13	2.01	0.98
140	فتوضا	15	1.94	1.14	190	ادم	13	2.01	0.98
141	حجة	15	1.94	1.14	191	تكون	13	2.01	0.98
142	اول	15	1.94	1.14	192	سبع	13	2.01	0.98
143	فاتي	15	1.94	1.14	193	بالصلاة	13	2.01	0.98
144	غلام	15	1.94	1.14	194	فرايت	13	2.01	0.98
145	الابل	15	1.94	1.14	195	سبعة	13	2.01	0.98
146	اليمن	15	1.94	1.14	196	سنة	13	2.01	0.98

Feature No	feature	DF	IDF	Feature redundancy	Feature No	feature	DF	IDF	Feature redundancy
197	الوداع	13	2.01	0.98	247	صنع	11	2.08	0.83
198	حرج	12	2.04	0.91	248	ايات	11	2.08	0.83
199	فرض	12	2.04	0.91	249	شفاء	11	2.08	0.83
200	فقتت	12	2.04	0.91	250	سجد	11	2.08	0.83
201	ركب	12	2.04	0.91	251	ترك	11	2.08	0.83
202	بالحج	12	2.04	0.91	252	والقمر	11	2.08	0.83
203	بمشي	12	2.04	0.91	253	شهاب	11	2.08	0.83
204	الذباء	12	2.04	0.91	254	ايتان	11	2.08	0.83
205	والعمرة	12	2.04	0.91	255	رمى	11	2.08	0.83
206	والشمس	12	2.04	0.91	256	اليهود	11	2.08	0.83
207	افعل	12	2.04	0.91	257	هدي	11	2.08	0.83
208	المسلمين	12	2.04	0.91	258	بماء	11	2.08	0.83
209	طعام	12	2.04	0.91	259	مائة	11	2.08	0.83
210	لولا	12	2.04	0.91	260	الحق	11	2.08	0.83
211	فجلس	12	2.04	0.91	261	قضى	11	2.08	0.83
212	فمر	12	2.04	0.91	262	حسنة	11	2.08	0.83
213	نجد	12	2.04	0.91	263	كتاب	11	2.08	0.83
214	عاشوراء	12	2.04	0.91	264	الدهر	10	2.12	0.76
215	عندها	12	2.04	0.91	265	هؤلاء	10	2.12	0.76
216	فذكر	12	2.04	0.91	266	مر	10	2.12	0.76
217	الشم	12	2.04	0.91	267	استاذن	10	2.12	0.76
218	شدت	12	2.04	0.91	268	ينادي	10	2.12	0.76
219	المنبر	12	2.04	0.91	269	الانصاري	10	2.12	0.76
220	كتب	12	2.04	0.91	270	الدين	10	2.12	0.76
221	مات	12	2.04	0.91	271	اشهد	10	2.12	0.76
222	نفسه	12	2.04	0.91	272	اصبح	10	2.12	0.76
223	الطعام	12	2.04	0.91	273	جعل	10	2.12	0.76
224	يمينه	12	2.04	0.91	274	بال	10	2.12	0.76
225	حرام	11	2.08	0.83	275	الحجر	10	2.12	0.76
226	سليم	11	2.08	0.83	276	اتيت	10	2.12	0.76
227	فاتي	11	2.08	0.83	277	العلم	10	2.12	0.76
228	سلم	11	2.08	0.83	278	سفر	10	2.12	0.76
229	ادري	11	2.08	0.83	279	تقوم	10	2.12	0.76
230	امراته	11	2.08	0.83	280	اجر	10	2.12	0.76
231	فسكت	11	2.08	0.83	281	راحلته	10	2.12	0.76
232	خيبر	11	2.08	0.83	282	فذهبت	10	2.12	0.76
233	الوادي	11	2.08	0.83	283	الاسود	10	2.12	0.76
234	يفعل	11	2.08	0.83	284	اخبرنا	10	2.12	0.76
235	امتي	11	2.08	0.83	285	الامر	10	2.12	0.76
236	والنبي	11	2.08	0.83	286	الليلة	10	2.12	0.76
237	رايته	11	2.08	0.83	287	سبعا	10	2.12	0.76
238	الغنم	11	2.08	0.83	288	فنظر	10	2.12	0.76
239	بلغ	11	2.08	0.83	289	رجع	10	2.12	0.76
240	جمع	11	2.08	0.83	290	الحر	10	2.12	0.76
241	مثله	11	2.08	0.83	291	كذب	10	2.12	0.76
242	افطر	11	2.08	0.83	292	خلف	10	2.12	0.76
243	ناسا	11	2.08	0.83	293	لحياته	10	2.12	0.76
244	ارابت	11	2.08	0.83	294	والعشاء	10	2.12	0.76
245	مرتين	11	2.08	0.83	295	الركن	10	2.12	0.76
246	الحسن	11	2.08	0.83	296	ربيعة	10	2.12	0.76

Feature No	feature	DF	IDF	Feature redundancy	Feature No	feature	DF	IDF	Feature redundancy
297	جاءت	10	2.12	0.76	347	حق	9	2.17	0.68
298	كثير	10	2.12	0.76	348	احسن	9	2.17	0.68
299	وصلى	10	2.12	0.76	349	يصوم	9	2.17	0.68
300	تمررة	10	2.12	0.76	350	المؤذن	9	2.17	0.68
301	صام	10	2.12	0.76	351	الحرام	9	2.17	0.68
302	بالناس	10	2.12	0.76	352	دخلت	9	2.17	0.68
303	يصلون	10	2.12	0.76	353	محمدا	9	2.17	0.68
304	نرى	10	2.12	0.76	354	بطن	9	2.17	0.68
305	الجمرة	10	2.12	0.76	355	ورسول	9	2.17	0.68
306	اخاه	9	2.17	0.68	356	فسالت	9	2.17	0.68
307	صيام	9	2.17	0.68	357	فشرب	9	2.17	0.68
308	فصم	9	2.17	0.68	358	فساله	9	2.17	0.68
309	يؤذن	9	2.17	0.68	359	المال	9	2.17	0.68
310	وجه	9	2.17	0.68	360	حديث	9	2.17	0.68
311	تصدق	9	2.17	0.68	361	ذهب	9	2.17	0.68
312	اتاه	9	2.17	0.68	362	قرن	9	2.17	0.68
313	الحمد	9	2.17	0.68	363	فصلوا	9	2.17	0.68
314	حلة	9	2.17	0.68	364	لاهل	9	2.17	0.68
315	ناس	9	2.17	0.68	365	الاولى	9	2.17	0.68
316	الرجال	9	2.17	0.68	366	الصائم	9	2.17	0.68
317	مرة	9	2.17	0.68	367	اسلم	9	2.17	0.68
318	بعث	9	2.17	0.68	368	يزل	9	2.17	0.68
319	يؤمن	9	2.17	0.68	369	تقدم	9	2.17	0.68
320	فليقل	9	2.17	0.68	370	رايتم	9	2.17	0.68
321	اذهب	9	2.17	0.68	371	رجال	9	2.17	0.68
322	منزله	9	2.17	0.68	372	بمنى	9	2.17	0.68
323	قدمنا	9	2.17	0.68	373	نصلي	9	2.17	0.68
324	يذكر	9	2.17	0.68	374	قياما	9	2.17	0.68
325	فاقبل	9	2.17	0.68	375	بمكة	9	2.17	0.68
326	شهادة	9	2.17	0.68	376	اخذ	9	2.17	0.68
327	شهرها	9	2.17	0.68	377	امي	8	2.22	0.61
328	ركع	9	2.17	0.68	378	عطاء	8	2.22	0.61
329	عز	9	2.17	0.68	379	الخير	8	2.22	0.61
330	وجل	9	2.17	0.68	380	كبير	8	2.22	0.61
331	العمل	9	2.17	0.68	381	مالا	8	2.22	0.61
332	اصلي	9	2.17	0.68	382	الباب	8	2.22	0.61
333	شديدا	9	2.17	0.68	383	سالم	8	2.22	0.61
334	واصحابه	9	2.17	0.68	384	احتجم	8	2.22	0.61
335	يسمع	9	2.17	0.68	385	الجمعة	8	2.22	0.61
336	شان	9	2.17	0.68	386	الكتاب	8	2.22	0.61
337	خمسة	9	2.17	0.68	387	المزدلفة	8	2.22	0.61
338	ارى	9	2.17	0.68	388	وقع	8	2.22	0.61
339	قوما	9	2.17	0.68	389	الطواف	8	2.22	0.61
340	وايتاء	9	2.17	0.68	390	انطلق	8	2.22	0.61
341	سمعته	9	2.17	0.68	391	حديثا	8	2.22	0.61
342	يرجع	9	2.17	0.68	392	السفر	8	2.22	0.61
343	اقيمت	9	2.17	0.68	393	بات	8	2.22	0.61
344	اشد	9	2.17	0.68	394	فانزل	8	2.22	0.61
345	فراى	9	2.17	0.68	395	اربعاء	8	2.22	0.61
346	عمله	9	2.17	0.68	396	البحر	8	2.22	0.61

Feature No	feature	DF	IDF	Feature redundancy	Feature No	feature	DF	IDF	Feature redundancy
397	وخرج	8	2.22	0.61	447	تعلم	7	2.28	0.53
398	نافع	8	2.22	0.61	448	كمثل	7	2.28	0.53
399	لاي	8	2.22	0.61	449	اربعين	7	2.28	0.53
400	يمر	8	2.22	0.61	450	خسفت	7	2.28	0.53
401	الطريق	8	2.22	0.61	451	يلبس	7	2.28	0.53
402	طعاما	8	2.22	0.61	452	محرم	7	2.28	0.53
403	ال	8	2.22	0.61	453	العمره	7	2.28	0.53
404	اصابعه	8	2.22	0.61	454	وعشرين	7	2.28	0.53
405	غفر	8	2.22	0.61	455	بدا	7	2.28	0.53
406	اعرابي	8	2.22	0.61	456	طويلا	7	2.28	0.53
407	شنت	8	2.22	0.61	457	لست	7	2.28	0.53
408	فيح	8	2.22	0.61	458	كلمة	7	2.28	0.53
409	صلاته	8	2.22	0.61	459	شر	7	2.28	0.53
410	سنين	8	2.22	0.61	460	الفتح	7	2.28	0.53
411	الحجة	8	2.22	0.61	461	الفضل	7	2.28	0.53
412	تغرب	8	2.22	0.61	462	يرى	7	2.28	0.53
413	واقام	8	2.22	0.61	463	فحمد	7	2.28	0.53
414	طيرة	8	2.22	0.61	464	يتوضا	7	2.28	0.53
415	حدث	8	2.22	0.61	465	احق	7	2.28	0.53
416	الحي	8	2.22	0.61	466	واقطر	7	2.28	0.53
417	باب	8	2.22	0.61	467	وراء	7	2.28	0.53
418	امير	8	2.22	0.61	468	تمرات	7	2.28	0.53
419	عشر	8	2.22	0.61	469	قطعة	7	2.28	0.53
420	فذكرت	8	2.22	0.61	470	ارمي	7	2.28	0.53
421	فوضع	8	2.22	0.61	471	فضحك	7	2.28	0.53
422	ينظر	8	2.22	0.61	472	فغضب	7	2.28	0.53
423	الصيام	8	2.22	0.61	473	ركعات	7	2.28	0.53
424	القبلة	8	2.22	0.61	474	تصوم	7	2.28	0.53
425	واهل	8	2.22	0.61	475	عدوى	7	2.28	0.53
426	فبلغ	8	2.22	0.61	476	واليوم	7	2.28	0.53
427	ايديهم	8	2.22	0.61	477	نسانه	7	2.28	0.53
428	جنت	8	2.22	0.61	478	لرجل	7	2.28	0.53
429	يلبي	8	2.22	0.61	479	قدموا	7	2.28	0.53
430	فدخلت	8	2.22	0.61	480	المريض	7	2.28	0.53
431	جاءه	8	2.22	0.61	481	خرجت	7	2.28	0.53
432	سم	8	2.22	0.61	482	التتعيم	7	2.28	0.53
433	نفر	8	2.22	0.61	483	واحدا	7	2.28	0.53
434	مضر	8	2.22	0.61	484	ليبك	7	2.28	0.53
435	انزلت	8	2.22	0.61	485	قائم	7	2.28	0.53
436	حلقت	8	2.22	0.61	486	ربنا	7	2.28	0.53
437	العقبة	8	2.22	0.61	487	توفي	7	2.28	0.53
438	فخرجت	8	2.22	0.61	488	الشعر	7	2.28	0.53
439	شجرة	8	2.22	0.61	489	الشجر	7	2.28	0.53
440	زمان	8	2.22	0.61	490	حقه	7	2.28	0.53
441	فرجع	8	2.22	0.61	491	اسوة	7	2.28	0.53
442	وامر	8	2.22	0.61	492	شك	7	2.28	0.53
443	الفتن	8	2.22	0.61	493	نحوه	7	2.28	0.53
444	قتل	8	2.22	0.61	494	المرء	7	2.28	0.53
445	ادرك	8	2.22	0.61	495	اسمعه	7	2.28	0.53
446	يبلغ	8	2.22	0.61	496	السائل	7	2.28	0.53

Feature No	feature	DF	IDF	Feature redundancy	Feature No	feature	DF	IDF	Feature redundancy
595	فيصلي	6	2.34	0.45	644	اصاب	6	2.34	0.45
596	الاذان	6	2.34	0.45	645	الكفر	6	2.34	0.45
597	فاتاه	6	2.34	0.45	646	مرحبا	6	2.34	0.45
598	اسود	6	2.34	0.45	647	بعير	6	2.34	0.45
599	اشتد	6	2.34	0.45	648	مال	6	2.34	0.45
600	قائمة	6	2.34	0.45	649	منا	6	2.34	0.45
601	فقرا	6	2.34	0.45	650	تطلع	6	2.34	0.45
602	فارسلت	6	2.34	0.45	651	بليل	6	2.34	0.45
603	شهركم	6	2.34	0.45	652	بقية	6	2.34	0.45
604	ظهره	6	2.34	0.45	653	اجرا	6	2.34	0.45
605	فوقع	6	2.34	0.45	654	تواصل	6	2.34	0.45
606	الجحفة	6	2.34	0.45	655	يقدم	6	2.34	0.45
607	والمزفت	6	2.34	0.45	656	ذهبت	6	2.34	0.45
608	الغداة	6	2.34	0.45	657	ستين	6	2.34	0.45
609	وغفار	6	2.34	0.45	658	لرسول	6	2.34	0.45
610	نار	6	2.34	0.45	659	يقبل	6	2.34	0.45
611	اعينهم	6	2.34	0.45	660	تكلم	6	2.34	0.45
612	قلبه	6	2.34	0.45	661	اناس	6	2.34	0.45
613	وقت	6	2.34	0.45	662	اخرج	6	2.34	0.45
614	السفلى	6	2.34	0.45	663	انسا	6	2.34	0.45
615	وليلة	6	2.34	0.45	664	اثر	6	2.34	0.45
616	الركوع	6	2.34	0.45	665	المنام	6	2.34	0.45
617	فجاءه	6	2.34	0.45	666	للناس	6	2.34	0.45
618	والحنتم	6	2.34	0.45	667	امثالها	6	2.34	0.45
619	اطعم	6	2.34	0.45	668	رحمه	6	2.34	0.45
620	العليا	6	2.34	0.45	669	دماءكم	6	2.34	0.45
621	فنزلت	6	2.34	0.45	670	يتتبع	6	2.34	0.45
622	اعرف	6	2.34	0.45	671	حي	6	2.34	0.45
623	الخمس	6	2.34	0.45	672	خبز	6	2.34	0.45
624	اعرابيا	6	2.34	0.45	673	يرفع	6	2.34	0.45
625	اعتمر	6	2.34	0.45	674	رات	6	2.34	0.45
626	مولى	6	2.34	0.45	675	يسوق	6	2.34	0.45
627	وماله	6	2.34	0.45	676	كعب	6	2.34	0.45
628	فليتبوا	6	2.34	0.45	677	يكبر	6	2.34	0.45
629	قبض	6	2.34	0.45	678	كحرمة	6	2.34	0.45
630	الشاة	6	2.34	0.45	679	اناخ	6	2.34	0.45
631	صاحب	6	2.34	0.45	680	حدثته	6	2.34	0.45
632	بالمعروف	6	2.34	0.45	681	ابل	6	2.34	0.45
633	مقعه	6	2.34	0.45	682	سورة	6	2.34	0.45
634	مسجد	6	2.34	0.45	683	حيي	6	2.34	0.45
635	صاحبه	6	2.34	0.45	684	احل	6	2.34	0.45
636	قم	6	2.34	0.45	685	لقي	6	2.34	0.45
637	ايمن	6	2.34	0.45	686	استوت	6	2.34	0.45
638	الركعة	6	2.34	0.45	687	زوج	6	2.34	0.45
639	فامرني	6	2.34	0.45	688	ربها	5	2.42	0.38
640	والمغرب	6	2.34	0.45	689	انذن	5	2.42	0.38
641	التروية	6	2.34	0.45	690	فاكلوا	5	2.42	0.38
642	واموالكم	6	2.34	0.45	691	ليال	5	2.42	0.38
643	فليصم	6	2.34	0.45	692	يفرا	5	2.42	0.38

Feature No	feature	DF	IDF	Feature redundancy	Feature No	feature	DF	IDF	Feature redundancy
693	خرجوا	5	2.42	0.38	743	ليبيد	5	2.42	0.38
694	غدا	5	2.42	0.38	744	حلف	5	2.42	0.38
695	قومك	5	2.42	0.38	745	البر	5	2.42	0.38
696	اوسق	5	2.42	0.38	746	الشياطين	5	2.42	0.38
697	معشر	5	2.42	0.38	747	قسم	5	2.42	0.38
698	ركوعا	5	2.42	0.38	748	ولي	5	2.42	0.38
699	قالها	5	2.42	0.38	749	اخبر	5	2.42	0.38
700	افضل	5	2.42	0.38	750	متى	5	2.42	0.38
701	بنر	5	2.42	0.38	751	والذين	5	2.42	0.38
702	وكبر	5	2.42	0.38	752	ضيفه	5	2.42	0.38
703	ظهر	5	2.42	0.38	753	بلدكم	5	2.42	0.38
704	امسيت	5	2.42	0.38	754	الفرس	5	2.42	0.38
705	يعلم	5	2.42	0.38	755	وعشرون	5	2.42	0.38
706	العرة	5	2.42	0.38	756	جمرة	5	2.42	0.38
707	يقولون	5	2.42	0.38	757	يفعله	5	2.42	0.38
708	اترون	5	2.42	0.38	758	فاردت	5	2.42	0.38
709	زال	5	2.42	0.38	759	توضا	5	2.42	0.38
710	خمسين	5	2.42	0.38	760	كلوا	5	2.42	0.38
711	صوت	5	2.42	0.38	761	يفطر	5	2.42	0.38
712	ميرور	5	2.42	0.38	762	اتينا	5	2.42	0.38
713	ودعا	5	2.42	0.38	763	اخاف	5	2.42	0.38
714	فابي	5	2.42	0.38	764	البانها	5	2.42	0.38
715	رقبة	5	2.42	0.38	765	فخطب	5	2.42	0.38
716	وضع	5	2.42	0.38	766	جاره	5	2.42	0.38
717	غيركم	5	2.42	0.38	767	فنام	5	2.42	0.38
718	علم	5	2.42	0.38	768	يصل	5	2.42	0.38
719	رجلي	5	2.42	0.38	769	اياه	5	2.42	0.38
720	الاعمش	5	2.42	0.38	770	خمسا	5	2.42	0.38
721	واحدة	5	2.42	0.38	771	وعمره	5	2.42	0.38
722	فاخيره	5	2.42	0.38	772	عمرتك	5	2.42	0.38
723	ابوك	5	2.42	0.38	773	بابا	5	2.42	0.38
724	علمت	5	2.42	0.38	774	باربع	5	2.42	0.38
725	فسجد	5	2.42	0.38	775	امرنا	5	2.42	0.38
726	واثنى	5	2.42	0.38	776	كلها	5	2.42	0.38
727	ورجل	5	2.42	0.38	777	النداء	5	2.42	0.38
728	امرني	5	2.42	0.38	778	وصلوا	5	2.42	0.38
729	يخطب	5	2.42	0.38	779	اخرى	5	2.42	0.38
730	نذر	5	2.42	0.38	780	فجعلوا	5	2.42	0.38
731	اعلى	5	2.42	0.38	781	الكلمة	5	2.42	0.38
732	وبني	5	2.42	0.38	782	بصدقته	5	2.42	0.38
733	شرا	5	2.42	0.38	783	داء	5	2.42	0.38
734	نم	5	2.42	0.38	784	تسمع	5	2.42	0.38
735	بالصدقة	5	2.42	0.38	785	فنظرت	5	2.42	0.38
736	غربت	5	2.42	0.38	786	احداهن	5	2.42	0.38
737	بالعمره	5	2.42	0.38	787	يقتل	5	2.42	0.38
738	الشر	5	2.42	0.38	788	ربي	5	2.42	0.38
739	شعبة	5	2.42	0.38	789	بعيره	5	2.42	0.38
740	تركه	5	2.42	0.38	790	بعشر	5	2.42	0.38
741	البيداء	5	2.42	0.38	791	ترى	5	2.42	0.38
742	الثوب	5	2.42	0.38	792	اذبح	5	2.42	0.38

Feature No	feature	DF	IDF	Feature redundancy	Feature No	feature	DF	IDF	Feature redundancy
793	اختي	5	2.42	0.38	843	فصنع	5	2.42	0.38
794	اجد	5	2.42	0.38	844	نتوضا	5	2.42	0.38
795	وجعل	5	2.42	0.38	845	المسلمون	5	2.42	0.38
796	الغلام	5	2.42	0.38	846	يساره	5	2.42	0.38
797	يكتب	5	2.42	0.38	847	بالماء	5	2.42	0.38
798	اعظم	5	2.42	0.38	848	كتابه	5	2.42	0.38
799	والناس	5	2.42	0.38	849	ينكسفان	5	2.42	0.38
800	المهاجرين	5	2.42	0.38	850	وراءنا	5	2.42	0.38
801	تريد	5	2.42	0.38	851	نفس	5	2.42	0.38
802	حائض	5	2.42	0.38	852	انسان	5	2.42	0.38
803	ذنبه	5	2.42	0.38	853	يترك	5	2.42	0.38
804	سالوا	5	2.42	0.38	854	شرطة	5	2.42	0.38
805	لطعام	5	2.42	0.38	855	حمد	5	2.42	0.38
806	امشي	5	2.42	0.38	856	صفر	5	2.42	0.38
807	يعرف	5	2.42	0.38	857	شانك	5	2.42	0.38
808	محجم	5	2.42	0.38	858	وراء	5	2.42	0.38
809	امعاء	5	2.42	0.38	859	فرع	5	2.42	0.38
810	والنقير	5	2.42	0.38	860	فليكرم	5	2.42	0.38
811	اعطيت	5	2.42	0.38	861	امة	5	2.42	0.38
812	عطس	5	2.42	0.38	862	بامر	5	2.42	0.38
813	الجهاد	5	2.42	0.38	863	عمل	5	2.42	0.38
814	ثلاثين	5	2.42	0.38	864	سمعتم	5	2.42	0.38
815	اطيب	5	2.42	0.38	865	السجود	5	2.42	0.38
816	عظيم	5	2.42	0.38	866	هلكت	5	2.42	0.38
817	والصدقة	5	2.42	0.38	867	يزعم	5	2.42	0.38
818	عبيد	5	2.42	0.38	868	السييل	5	2.42	0.38
819	قريبا	5	2.42	0.38	869	حية	5	2.42	0.38
820	الحاجة	5	2.42	0.38	870	حضرت	5	2.42	0.38
821	الغائب	5	2.42	0.38	871	اليهم	5	2.42	0.38
822	بعثني	5	2.42	0.38	872	واحتسابا	5	2.42	0.38
823	ولدها	5	2.42	0.38	873	جالسا	5	2.42	0.38
824	ايماننا	5	2.42	0.38	874	فركب	5	2.42	0.38
825	طوى	5	2.42	0.38	875	وزاد	5	2.42	0.38
826	بقي	5	2.42	0.38	876	حول	5	2.42	0.38
827	وقام	5	2.42	0.38	877	يامرنا	5	2.42	0.38
828	الجنب	5	2.42	0.38	878	تمر	5	2.42	0.38
829	دفع	5	2.42	0.38	879	فسلم	5	2.42	0.38
830	وانزل	5	2.42	0.38	880	فعله	5	2.42	0.38
831	ولاهل	5	2.42	0.38	881	بالبطحاء	5	2.42	0.38
832	فامرهم	5	2.42	0.38	882	الشاهد	5	2.42	0.38
833	الغد	5	2.42	0.38	883	قومي	5	2.42	0.38
834	فاخير	5	2.42	0.38	884	عينية	5	2.42	0.38
835	طلع	5	2.42	0.38	885	يومين	5	2.42	0.38
836	ينبغي	5	2.42	0.38	886	فمرنا	5	2.42	0.38
837	يهل	5	2.42	0.38	887	استيقظ	5	2.42	0.38
838	حصيات	5	2.42	0.38	888	فذكروا	5	2.42	0.38
839	شعبان	5	2.42	0.38	889	صليت	5	2.42	0.38
840	نفسها	5	2.42	0.38	890	رقاب	5	2.42	0.38
841	مسدد	5	2.42	0.38	891	فقضى	5	2.42	0.38
842	المقام	5	2.42	0.38	892	الخيال	5	2.42	0.38

Feature No	feature	DF	IDF	Feature redundancy	Feature No	feature	DF	IDF	Feature redundancy
893	واشربوا	5	2.42	0.38	943	حمده	4	2.52	0.30
894	لبن	5	2.42	0.38	944	افتاني	4	2.52	0.30
895	اصنع	5	2.42	0.38	945	حلقى	4	2.52	0.30
896	فرقع	5	2.42	0.38	946	بالصهباء	4	2.52	0.30
897	فذكرته	5	2.42	0.38	947	الصخرة	4	2.52	0.30
898	المحرم	5	2.42	0.38	948	اهدى	4	2.52	0.30
899	ينزل	5	2.42	0.38	949	بالكعبة	4	2.52	0.30
900	شهد	5	2.42	0.38	950	الفتنة	4	2.52	0.30
901	واعراضكم	5	2.42	0.38	951	فاكلنا	4	2.52	0.30
902	يزال	5	2.42	0.38	952	نقاعة	4	2.52	0.30
903	هديا	5	2.42	0.38	953	بكرة	4	2.52	0.30
904	الأخرى	5	2.42	0.38	954	اخبركم	4	2.52	0.30
905	شانه	5	2.42	0.38	955	نخلها	4	2.52	0.30
906	عيناه	5	2.42	0.38	956	جناح	4	2.52	0.30
907	انحر	5	2.42	0.38	957	قيام	4	2.52	0.30
908	الذهب	5	2.42	0.38	958	رؤوس	4	2.52	0.30
909	ونحر	5	2.42	0.38	959	شديد	4	2.52	0.30
910	قائد	5	2.42	0.38	960	اتاني	4	2.52	0.30
911	المكتوبة	5	2.42	0.38	961	الهندي	4	2.52	0.30
912	ازواجه	5	2.42	0.38	962	العام	4	2.52	0.30
913	بالعشاء	5	2.42	0.38	963	طعامه	4	2.52	0.30
914	ناحية	5	2.42	0.38	964	حاضت	4	2.52	0.30
915	قبلة	5	2.42	0.38	965	تطوف	4	2.52	0.30
916	امرت	5	2.42	0.38	966	طوافا	4	2.52	0.30
917	شهرين	5	2.42	0.38	967	الدجال	4	2.52	0.30
918	فاجدح	4	2.52	0.30	968	عقرى	4	2.52	0.30
919	اركيها	4	2.52	0.30	969	القدر	4	2.52	0.30
920	عباد	4	2.52	0.30	970	الحرير	4	2.52	0.30
921	تحاسدوا	4	2.52	0.30	971	قائل	4	2.52	0.30
922	اخوانا	4	2.52	0.30	972	زريق	4	2.52	0.30
923	شبعوا	4	2.52	0.30	973	فشربت	4	2.52	0.30
924	اسمه	4	2.52	0.30	974	مطبوب	4	2.52	0.30
925	تدابروا	4	2.52	0.30	975	تسعا	4	2.52	0.30
926	يهجر	4	2.52	0.30	976	ماءها	4	2.52	0.30
927	تباغضوا	4	2.52	0.30	977	رجلين	4	2.52	0.30
928	اطيق	4	2.52	0.30	978	ناخذ	4	2.52	0.30
929	بلد	4	2.52	0.30	979	تعيد	4	2.52	0.30
930	مشط	4	2.52	0.30	980	نام	4	2.52	0.30
931	الفال	4	2.52	0.30	981	الركنين	4	2.52	0.30
932	ابرء	4	2.52	0.30	982	طبه	4	2.52	0.30
933	اهللت	4	2.52	0.30	983	العود	4	2.52	0.30
934	الزور	4	2.52	0.30	984	النخل	4	2.52	0.30
935	عباده	4	2.52	0.30	985	ذر	4	2.52	0.30
936	ادخل	4	2.52	0.30	986	عملا	4	2.52	0.30
937	السام	4	2.52	0.30	987	فضحكت	4	2.52	0.30
938	اياكم	4	2.52	0.30	988	سهل	4	2.52	0.30
939	يعمل	4	2.52	0.30	989	تعني	4	2.52	0.30
940	كسرى	4	2.52	0.30	990	فمضى	4	2.52	0.30
941	نحر	4	2.52	0.30	991	والنار	4	2.52	0.30
942	استقتيته	4	2.52	0.30	992	واتبعنا	4	2.52	0.30

Feature No	feature	DF	IDF	Feature redundancy	Feature No	feature	DF	IDF	Feature redundancy
993	فخذ	4	2.52	0.30	1043	القراءة	4	2.52	0.30
994	فائتي	4	2.52	0.30	1044	بيضاء	4	2.52	0.30
995	دينار	4	2.52	0.30	1045	غنمتم	4	2.52	0.30
996	كراهية	4	2.52	0.30	1046	وجوههم	4	2.52	0.30
997	اخو	4	2.52	0.30	1047	النظر	4	2.52	0.30
998	الحناء	4	2.52	0.30	1048	العزير	4	2.52	0.30
999	ربك	4	2.52	0.30	1049	واشهد	4	2.52	0.30
1000	طبيب	4	2.52	0.30	1050	يلعب	4	2.52	0.30
1001	اخذت	4	2.52	0.30	1051	فرج	4	2.52	0.30
1002	يسقط	4	2.52	0.30	1052	بالايطح	4	2.52	0.30
1003	يسب	4	2.52	0.30	1053	واسلم	4	2.52	0.30
1004	قائما	4	2.52	0.30	1054	فرسه	4	2.52	0.30
1005	ربكم	4	2.52	0.30	1055	ناكل	4	2.52	0.30
1006	رجالا	4	2.52	0.30	1056	ناخذه	4	2.52	0.30
1007	جريح	4	2.52	0.30	1057	نائم	4	2.52	0.30
1008	يرمي	4	2.52	0.30	1058	الرحم	4	2.52	0.30
1009	وتقيم	4	2.52	0.30	1059	العرب	4	2.52	0.30
1010	لقيه	4	2.52	0.30	1060	ات	4	2.52	0.30
1011	القطر	4	2.52	0.30	1061	بخير	4	2.52	0.30
1012	مسكينا	4	2.52	0.30	1062	فسار	4	2.52	0.30
1013	فاتيته	4	2.52	0.30	1063	ابيت	4	2.52	0.30
1014	القعدة	4	2.52	0.30	1064	يغلق	4	2.52	0.30
1015	تجدون	4	2.52	0.30	1065	واكثر	4	2.52	0.30
1016	فابوا	4	2.52	0.30	1066	يقف	4	2.52	0.30
1017	فابردوا	4	2.52	0.30	1067	بشر	4	2.52	0.30
1018	وجاره	4	2.52	0.30	1068	فحدثته	4	2.52	0.30
1019	وجهي	4	2.52	0.30	1069	الحمى	4	2.52	0.30
1020	وجهها	4	2.52	0.30	1070	مواقع	4	2.52	0.30
1021	اخبره	4	2.52	0.30	1071	فخذه	4	2.52	0.30
1022	بيديه	4	2.52	0.30	1072	فخذوا	4	2.52	0.30
1023	اموالهم	4	2.52	0.30	1073	الدار	4	2.52	0.30
1024	غداة	4	2.52	0.30	1074	باعلى	4	2.52	0.30
1025	غسل	4	2.52	0.30	1075	بر	4	2.52	0.30
1026	وجهينة	4	2.52	0.30	1076	فدعاه	4	2.52	0.30
1027	غلاما	4	2.52	0.30	1077	فمضض	4	2.52	0.30
1028	الابيض	4	2.52	0.30	1078	يكسر	4	2.52	0.30
1029	شع	4	2.52	0.30	1079	الخطبة	4	2.52	0.30
1030	بيتي	4	2.52	0.30	1080	فمسح	4	2.52	0.30
1031	اخرك	4	2.52	0.30	1081	فذهب	4	2.52	0.30
1032	القيام	4	2.52	0.30	1082	فنادى	4	2.52	0.30
1033	غزا	4	2.52	0.30	1083	نزلنا	4	2.52	0.30
1034	الكافر	4	2.52	0.30	1084	بالخير	4	2.52	0.30
1035	يسالني	4	2.52	0.30	1085	نخل	4	2.52	0.30
1036	وراي	4	2.52	0.30	1086	بالسكينة	4	2.52	0.30
1037	معروف	4	2.52	0.30	1087	بالشام	4	2.52	0.30
1038	الاقامة	4	2.52	0.30	1088	الشرك	4	2.52	0.30
1039	بياض	4	2.52	0.30	1089	بالامس	4	2.52	0.30
1040	فادعوا	4	2.52	0.30	1090	الشعير	4	2.52	0.30
1041	وجعلت	4	2.52	0.30	1091	فطفت	4	2.52	0.30
1042	كنانة	4	2.52	0.30	1092	حل	4	2.52	0.30

Feature No	feature	DF	IDF	Feature redundancy	Feature No	feature	DF	IDF	Feature redundancy
1093	الصماء	4	2.52	0.30	1144	بلالا	4	2.52	0.30
1094	فقدم	4	2.52	0.30	1145	لعطاء	4	2.52	0.30
1095	فقدمت	4	2.52	0.30	1146	لاحد	4	2.52	0.30
1096	فكبر	4	2.52	0.30	1147	يصومه	4	2.52	0.30
1097	حمر	4	2.52	0.30	1148	بطريق	4	2.52	0.30
1098	حصاة	4	2.52	0.30	1149	فاين	4	2.52	0.30
1099	الرهط	4	2.52	0.30	1150	العين	4	2.52	0.30
1100	ابواب	4	2.52	0.30	1151	فوجدته	4	2.52	0.30
1101	فليصل	4	2.52	0.30	1152	فوجد	4	2.52	0.30
1102	بثلاث	4	2.52	0.30	1153	اثره	4	2.52	0.30
1103	نبيه	4	2.52	0.30	1154	الحجارة	4	2.52	0.30
1104	فليتيم	4	2.52	0.30	1155	الوالدين	4	2.52	0.30
1105	امامك	4	2.52	0.30	1156	دباء	4	2.52	0.30
1106	بالمهاجرة	4	2.52	0.30	1157	بعثه	4	2.52	0.30
1107	هدي	4	2.52	0.30	1158	بطحان	4	2.52	0.30
1108	نودي	4	2.52	0.30	1159	بصيامه	4	2.52	0.30
1109	بايحت	4	2.52	0.30	1160	حجر	4	2.52	0.30
1110	فضرب	4	2.52	0.30	1161	اتوا	4	2.52	0.30
1111	فضل	4	2.52	0.30	1162	اتقوا	4	2.52	0.30
1112	الطيب	4	2.52	0.30	1163	يقسم	4	2.52	0.30
1113	هامة	4	2.52	0.30	1164	بطنه	4	2.52	0.30
1114	ينصرف	4	2.52	0.30	1165	اخيك	4	2.52	0.30
1115	خياطا	4	2.52	0.30	1166	يقضي	4	2.52	0.30
1116	ابنته	4	2.52	0.30	1167	حاجته	4	2.52	0.30
1117	بفدح	4	2.52	0.30	1168	فيال	4	2.52	0.30
1118	لعبد	4	2.52	0.30	1169	يغتسل	4	2.52	0.30
1119	جنب	4	2.52	0.30	1170	فوعظهن	4	2.52	0.30
1120	النوم	4	2.52	0.30	1171	الحبة	4	2.52	0.30
1121	نود	4	2.52	0.30	1172	بعرق	4	2.52	0.30
1122	بشق	4	2.52	0.30	1173	الجبل	4	2.52	0.30
1123	ملا	4	2.52	0.30	1174	حبط	4	2.52	0.30
1124	ملك	4	2.52	0.30	1175	يفتح	4	2.52	0.30
1125	الثالثة	4	2.52	0.30	1176	فتح	4	2.52	0.30
1126	يعطي	4	2.52	0.30	1177	الجوع	4	2.52	0.30
1127	والنصح	4	2.52	0.30	1178	يفر	4	2.52	0.30
1128	الثلاثة	4	2.52	0.30	1179	فتحت	4	2.52	0.30
1129	دينا	4	2.52	0.30	1180	فبعث	4	2.52	0.30
1130	وامرهن	4	2.52	0.30	1181	يجيء	4	2.52	0.30
1131	احرم	4	2.52	0.30	1182	اطواف	4	2.52	0.30
1132	الباس	4	2.52	0.30	1183	عمران	4	2.52	0.30
1133	ربا	4	2.52	0.30	1184	جاؤوا	4	2.52	0.30
1134	يصنع	4	2.52	0.30	1185	ارم	4	2.52	0.30
1135	احسنت	4	2.52	0.30	1186	سجدتين	4	2.52	0.30
1136	فاعتمرت	4	2.52	0.30	1187	طرفيه	4	2.52	0.30
1137	راس	4	2.52	0.30	1188	سترا	4	2.52	0.30
1138	الفطر	4	2.52	0.30	1189	ولد	4	2.52	0.30
1139	يطعمني	4	2.52	0.30	1190	اهلوا	4	2.52	0.30
1140	راني	4	2.52	0.30	1191	كفارا	4	2.52	0.30
1141	فافز عوا	4	2.52	0.30	1192	طوافه	4	2.52	0.30
1142	امنا	4	2.52	0.30	1193	سعي	4	2.52	0.30
1143	يضره	4	2.52	0.30	1194	عمدت	4	2.52	0.30

Feature No	feature	DF	IDF	Feature redundancy	Feature No	feature	DF	IDF	Feature redundancy
1195	صنعه	4	2.52	0.30	1245	ماندته	4	2.52	0.30
1196	المكان	4	2.52	0.30	1246	ر هطأ	4	2.52	0.30
1197	متتابعين	4	2.52	0.30	1247	اشهدكم	4	2.52	0.30
1198	كفر	4	2.52	0.30	1248	قبلها	4	2.52	0.30
1199	زاد	4	2.52	0.30	1249	اقل	4	2.52	0.30
1200	اريد	4	2.52	0.30	1250	جانب	4	2.52	0.30
1201	اطعمه	4	2.52	0.30	1251	سمعه	4	2.52	0.30
1202	ارسل	4	2.52	0.30	1252	تكفرها	4	2.52	0.30
1203	ومزينة	4	2.52	0.30	1253	تحب	4	2.52	0.30
1204	قتلوا	4	2.52	0.30	1254	ليصمت	4	2.52	0.30
1205	اقرا	4	2.52	0.30	1255	اشفية	4	2.52	0.30
1206	عمير	4	2.52	0.30	1256	سيد	4	2.52	0.30
1207	ضحى	4	2.52	0.30	1257	يتبع	4	2.52	0.30
1208	تشارك	4	2.52	0.30	1258	ويصوم	4	2.52	0.30
1209	اواق	4	2.52	0.30	1259	ويصلي	4	2.52	0.30
1210	ومشى	4	2.52	0.30	1260	صالح	4	2.52	0.30
1211	ثوبه	4	2.52	0.30	1261	اقول	4	2.52	0.30
1212	الموت	4	2.52	0.30	1262	ليالي	4	2.52	0.30
1213	عجوة	4	2.52	0.30	1263	اشرف	4	2.52	0.30
1214	قرنا	4	2.52	0.30	1264	ويقرا	4	2.52	0.30
1215	تخرج	4	2.52	0.30	1265	المشرق	4	2.52	0.30
1216	ترجعوا	4	2.52	0.30	1266	بيسط	4	2.52	0.30
1217	كسفت	4	2.52	0.30	1267	صلوات	4	2.52	0.30
1218	تربة	4	2.52	0.30	1268	استطاع	4	2.52	0.30
1219	ثقل	4	2.52	0.30	1269	عنقه	4	2.52	0.30
1220	سانر	4	2.52	0.30	1270	اسفل	4	2.52	0.30
1221	الكهان	4	2.52	0.30	1271	اسفلها	4	2.52	0.30
1222	افاض	4	2.52	0.30	1272	وولده	4	2.52	0.30
1223	عسل	4	2.52	0.30	1273	يبقى	4	2.52	0.30
1224	يد	4	2.52	0.30	1274	شربة	4	2.52	0.30
1225	زهير	4	2.52	0.30	1275	ونم	4	2.52	0.30
1226	عرضت	4	2.52	0.30	1276	تموج	4	2.52	0.30
1227	يحفظ	4	2.52	0.30	1277	ونسيت	4	2.52	0.30
1228	يدعى	4	2.52	0.30	1278	عيد	4	2.52	0.30
1229	زمزم	4	2.52	0.30	1279	صدقته	4	2.52	0.30
1230	اردف	4	2.52	0.30	1280	يرد	4	2.52	0.30
1231	اقام	4	2.52	0.30	1281	تمرا	4	2.52	0.30
1232	اعطاه	4	2.52	0.30	1282	ويفطر	4	2.52	0.30
1233	ظننت	4	2.52	0.30	1283	ر هط	4	2.52	0.30
1234	اراه	4	2.52	0.30	1284	كداء	3	2.64	0.23
1235	اعطى	4	2.52	0.30	1285	قصرت	3	2.64	0.23
1236	ضرب	4	2.52	0.30	1286	لمسلم	3	2.64	0.23
1237	وقعت	4	2.52	0.30	1287	مؤمنا	3	2.64	0.23
1238	سالنا	4	2.52	0.30	1288	مسلمنا	3	2.64	0.23
1239	عاصم	4	2.52	0.30	1289	اتحنث	3	2.64	0.23
1240	وقديد	4	2.52	0.30	1290	شائين	3	2.64	0.23
1241	وقتها	4	2.52	0.30	1291	الهجرة	3	2.64	0.23
1242	ترون	4	2.52	0.30	1292	طويلة	3	2.64	0.23
1243	يلحوا	4	2.52	0.30	1293	يخفف	3	2.64	0.23
1244	مريم	4	2.52	0.30	1294	قراءة	3	2.64	0.23

Feature No	feature	DF	IDF	Feature redundanc	Feature No	feature	DF	IDF	Feature redundanc
1295	دور	3	2.64	0.23	1345	صم	3	2.64	0.23
1296	فعلت	3	2.64	0.23	1346	بئس	3	2.64	0.23
1297	امك	3	2.64	0.23	1347	فوضعها	3	2.64	0.23
1298	فلانا	3	2.64	0.23	1348	بياب	3	2.64	0.23
1299	وانهاكم	3	2.64	0.23	1349	تسعة	3	2.64	0.23
1300	ستر	3	2.64	0.23	1350	المصدق	3	2.64	0.23
1301	كبير	3	2.64	0.23	1351	مثلها	3	2.64	0.23
1302	صوتك	3	2.64	0.23	1352	يطوفون	3	2.64	0.23
1303	ماله	3	2.64	0.23	1353	فخرجنا	3	2.64	0.23
1304	امرؤ	3	2.64	0.23	1354	تنفر	3	2.64	0.23
1305	فاعتزل	3	2.64	0.23	1355	الحياة	3	2.64	0.23
1306	ليون	3	2.64	0.23	1356	الحوت	3	2.64	0.23
1307	العشر	3	2.64	0.23	1357	فشددت	3	2.64	0.23
1308	بالقوارير	3	2.64	0.23	1358	زهرة	3	2.64	0.23
1309	فطفق	3	2.64	0.23	1359	تعرف	3	2.64	0.23
1310	احسب	3	2.64	0.23	1360	بالليل	3	2.64	0.23
1311	لحم	3	2.64	0.23	1361	حجرتها	3	2.64	0.23
1312	يضحك	3	2.64	0.23	1362	سلوني	3	2.64	0.23
1313	ثوبا	3	2.64	0.23	1363	طاقوا	3	2.64	0.23
1314	مقامي	3	2.64	0.23	1364	ستره	3	2.64	0.23
1315	يضع	3	2.64	0.23	1365	ثلث	3	2.64	0.23
1316	اصبحت	3	2.64	0.23	1366	بشاة	3	2.64	0.23
1317	يكشف	3	2.64	0.23	1367	مشرك	3	2.64	0.23
1318	ايديكم	3	2.64	0.23	1368	رواحة	3	2.64	0.23
1319	احمد	3	2.64	0.23	1369	القصة	3	2.64	0.23
1320	فارسل	3	2.64	0.23	1370	افقر	3	2.64	0.23
1321	فراشه	3	2.64	0.23	1371	يدع	3	2.64	0.23
1322	جعلهم	3	2.64	0.23	1372	اقربهما	3	2.64	0.23
1323	يومه	3	2.64	0.23	1373	رحم	3	2.64	0.23
1324	تذكر	3	2.64	0.23	1374	اكلت	3	2.64	0.23
1325	جلست	3	2.64	0.23	1375	مضى	3	2.64	0.23
1326	اسد	3	2.64	0.23	1376	ورب	3	2.64	0.23
1327	جلبابها	3	2.64	0.23	1377	وشرايه	3	2.64	0.23
1328	مرات	3	2.64	0.23	1378	اقسم	3	2.64	0.23
1329	البدن	3	2.64	0.23	1379	بيد	3	2.64	0.23
1330	لاعلم	3	2.64	0.23	1380	فاخرج	3	2.64	0.23
1331	شديدة	3	2.64	0.23	1381	غلبني	3	2.64	0.23
1332	انفق	3	2.64	0.23	1382	جاءنا	3	2.64	0.23
1333	فقلن	3	2.64	0.23	1383	الامة	3	2.64	0.23
1334	اتقاء	3	2.64	0.23	1384	عندهم	3	2.64	0.23
1335	اطول	3	2.64	0.23	1385	فيك	3	2.64	0.23
1336	دخلنا	3	2.64	0.23	1386	غلامه	3	2.64	0.23
1337	بعثت	3	2.64	0.23	1387	بيع	3	2.64	0.23
1338	فلقيت	3	2.64	0.23	1388	يمرفون	3	2.64	0.23
1339	يحج	3	2.64	0.23	1389	فريضة	3	2.64	0.23
1340	اظنه	3	2.64	0.23	1390	الرجال	3	2.64	0.23
1341	ومشاقة	3	2.64	0.23	1391	وارجلهم	3	2.64	0.23
1342	السوداء	3	2.64	0.23	1392	فم	3	2.64	0.23
1343	ذروان	3	2.64	0.23	1393	خزايا	3	2.64	0.23
1344	طفت	3	2.64	0.23	1394	العرق	3	2.64	0.23

Feature No	feature	DF	IDF	Feature redundan	Feature No	feature	DF	IDF	Feature redundan
1395	وندعو	3	2.64	0.23	1445	الرمية	3	2.64	0.23
1396	لنفسه	3	2.64	0.23	1446	صدقتك	3	2.64	0.23
1397	نبغي	3	2.64	0.23	1447	بالارض	3	2.64	0.23
1398	حراما	3	2.64	0.23	1448	فسقاه	3	2.64	0.23
1399	حرمت	3	2.64	0.23	1449	الوقوف	3	2.64	0.23
1400	صغار	3	2.64	0.23	1450	فايليس	3	2.64	0.23
1401	تطبيقون	3	2.64	0.23	1451	اشياء	3	2.64	0.23
1402	يسار	3	2.64	0.23	1452	فسماه	3	2.64	0.23
1403	امته	3	2.64	0.23	1453	غم	3	2.64	0.23
1404	سل	3	2.64	0.23	1454	سلف	3	2.64	0.23
1405	قلدها	3	2.64	0.23	1455	فرخص	3	2.64	0.23
1406	اسال	3	2.64	0.23	1456	غنثر	3	2.64	0.23
1407	كثر	3	2.64	0.23	1457	غنى	3	2.64	0.23
1408	بدرا	3	2.64	0.23	1458	الرطب	3	2.64	0.23
1409	يناجي	3	2.64	0.23	1459	الاحسان	3	2.64	0.23
1410	بين	3	2.64	0.23	1460	اشهر	3	2.64	0.23
1411	بدر	3	2.64	0.23	1461	فسال	3	2.64	0.23
1412	تامرنا	3	2.64	0.23	1462	روحة	3	2.64	0.23
1413	سقما	3	2.64	0.23	1463	خربت	3	2.64	0.23
1414	الذود	3	2.64	0.23	1464	فساء	3	2.64	0.23
1415	فرفعت	3	2.64	0.23	1465	فسلمت	3	2.64	0.23
1416	بدت	3	2.64	0.23	1466	الاذى	3	2.64	0.23
1417	الزعفران	3	2.64	0.23	1467	تصنع	3	2.64	0.23
1418	الراس	3	2.64	0.23	1468	اريت	3	2.64	0.23
1419	الراعي	3	2.64	0.23	1469	خلقا	3	2.64	0.23
1420	واستاقوا	3	2.64	0.23	1470	واكره	3	2.64	0.23
1421	بثوبه	3	2.64	0.23	1471	صومه	3	2.64	0.23
1422	كفف	3	2.64	0.23	1472	الخزائن	3	2.64	0.23
1423	يليك	3	2.64	0.23	1473	تصوموا	3	2.64	0.23
1424	اتاذن	3	2.64	0.23	1474	امراتي	3	2.64	0.23
1425	يلي	3	2.64	0.23	1475	العشير	3	2.64	0.23
1426	صدر	3	2.64	0.23	1476	واقف	3	2.64	0.23
1427	الركعتين	3	2.64	0.23	1477	قاتلهم	3	2.64	0.23
1428	لجريء	3	2.64	0.23	1478	الاذن	3	2.64	0.23
1429	يتحرى	3	2.64	0.23	1479	اوجبت	3	2.64	0.23
1430	خذ	3	2.64	0.23	1480	برزة	3	2.64	0.23
1431	حزن	3	2.64	0.23	1481	فدعي	3	2.64	0.23
1432	بالبركة	3	2.64	0.23	1482	المكتل	3	2.64	0.23
1433	يمنع	3	2.64	0.23	1483	يكذب	3	2.64	0.23
1434	وادعوا	3	2.64	0.23	1484	لخمس	3	2.64	0.23
1435	اقم	3	2.64	0.23	1485	قلاية	3	2.64	0.23
1436	فسالته	3	2.64	0.23	1486	فرس	3	2.64	0.23
1437	العبد	3	2.64	0.23	1487	حدثت	3	2.64	0.23
1438	وسمر	3	2.64	0.23	1488	باصبعه	3	2.64	0.23
1439	اشفعوا	3	2.64	0.23	1489	فنز لنا	3	2.64	0.23
1440	يتبين	3	2.64	0.23	1490	فاتته	3	2.64	0.23
1441	العنمة	3	2.64	0.23	1491	عمه	3	2.64	0.23
1442	بحجة	3	2.64	0.23	1492	تنفق	3	2.64	0.23
1443	بحج	3	2.64	0.23	1493	امرکم	3	2.64	0.23
1444	بجمع	3	2.64	0.23	1494	ودخلت	3	2.64	0.23

Feature No	feature	DF	IDF	Feature redundan	Feature No	feature	DF	IDF	Feature redundan
1495	مقحشا	3	2.64	0.23	1545	لحومها	3	2.64	0.23
1496	لذعة	3	2.64	0.23	1546	المحصب	3	2.64	0.23
1497	مستقبل	3	2.64	0.23	1547	كفضل	3	2.64	0.23
1498	وددت	3	2.64	0.23	1548	يسالونه	3	2.64	0.23
1499	الوحي	3	2.64	0.23	1549	بالكفر	3	2.64	0.23
1500	الحياء	3	2.64	0.23	1550	اسلمت	3	2.64	0.23
1501	اذكره	3	2.64	0.23	1551	شنتم	3	2.64	0.23
1502	والامر	3	2.64	0.23	1552	ينهى	3	2.64	0.23
1503	فترك	3	2.64	0.23	1553	نفسك	3	2.64	0.23
1504	حجنا	3	2.64	0.23	1554	قاموا	3	2.64	0.23
1505	بساحة	3	2.64	0.23	1555	اكبركم	3	2.64	0.23
1506	تصدقوا	3	2.64	0.23	1556	شريك	3	2.64	0.23
1507	انكسفت	3	2.64	0.23	1557	تجلت	3	2.64	0.23
1508	فانتت	3	2.64	0.23	1558	وسبح	3	2.64	0.23
1509	تاكل	3	2.64	0.23	1559	الشفاء	3	2.64	0.23
1510	فرايته	3	2.64	0.23	1560	ويظهر	3	2.64	0.23
1511	تنام	3	2.64	0.23	1561	فعفر	3	2.64	0.23
1512	اخلف	3	2.64	0.23	1562	اباه	3	2.64	0.23
1513	مالي	3	2.64	0.23	1563	بالقوم	3	2.64	0.23
1514	واشجع	3	2.64	0.23	1564	يهدى	3	2.64	0.23
1515	اقصى	3	2.64	0.23	1565	حمار	3	2.64	0.23
1516	امرا	3	2.64	0.23	1566	لبدت	3	2.64	0.23
1517	بدنه	3	2.64	0.23	1567	الشيء	3	2.64	0.23
1518	مشربة	3	2.64	0.23	1568	يهلون	3	2.64	0.23
1519	لحوم	3	2.64	0.23	1569	نبيا	3	2.64	0.23
1520	لحام	3	2.64	0.23	1570	ويسقين	3	2.64	0.23
1521	يلقاه	3	2.64	0.23	1571	ندامى	3	2.64	0.23
1522	فرحا	3	2.64	0.23	1572	استلام	3	2.64	0.23
1523	بدن	3	2.64	0.23	1573	الشافي	3	2.64	0.23
1524	حذافة	3	2.64	0.23	1574	اليسرى	3	2.64	0.23
1525	صلاطين	3	2.64	0.23	1575	فلانة	3	2.64	0.23
1526	صباح	3	2.64	0.23	1576	غزوة	3	2.64	0.23
1527	انفه	3	2.64	0.23	1577	قامت	3	2.64	0.23
1528	تحدث	3	2.64	0.23	1578	الشعب	3	2.64	0.23
1529	واشار	3	2.64	0.23	1579	فكيف	3	2.64	0.23
1530	خفيه	3	2.64	0.23	1580	باتيني	3	2.64	0.23
1531	حدثتني	3	2.64	0.23	1581	يرفت	3	2.64	0.23
1532	الخلق	3	2.64	0.23	1582	حقها	3	2.64	0.23
1533	الخمير	3	2.64	0.23	1583	كلب	3	2.64	0.23
1534	تنزل	3	2.64	0.23	1584	ورواه	3	2.64	0.23
1535	تاخذ	3	2.64	0.23	1585	شعرة	3	2.64	0.23
1536	الخنق	3	2.64	0.23	1586	المسك	3	2.64	0.23
1537	قلبي	3	2.64	0.23	1587	تبكي	3	2.64	0.23
1538	اتت	3	2.64	0.23	1588	المسكين	3	2.64	0.23
1539	يرحم	3	2.64	0.23	1589	حكمة	3	2.64	0.23
1540	امراوك	3	2.64	0.23	1590	حمى	3	2.64	0.23
1541	رويدك	3	2.64	0.23	1591	كافر	3	2.64	0.23
1542	الخيظ	3	2.64	0.23	1592	الصبيان	3	2.64	0.23
1543	صلينا	3	2.64	0.23	1593	فقسمتها	3	2.64	0.23
1544	خفيفتين	3	2.64	0.23	1594	الصحفة	3	2.64	0.23

Feature No	feature	DF	IDF	Feature redundan	Feature No	feature	DF	IDF	Feature redundan
1595	يوشك	3	2.64	0.23	1645	اليد	3	2.64	0.23
1596	فقطع	3	2.64	0.23	1646	ينحر	3	2.64	0.23
1597	فقعد	3	2.64	0.23	1647	غضبا	3	2.64	0.23
1598	اكتب	3	2.64	0.23	1648	سمعنا	3	2.64	0.23
1599	يوضع	3	2.64	0.23	1649	هديه	3	2.64	0.23
1600	انشق	3	2.64	0.23	1650	تمتع	3	2.64	0.23
1601	وسادة	3	2.64	0.23	1651	ويتصدق	3	2.64	0.23
1602	نسي	3	2.64	0.23	1652	صائما	3	2.64	0.23
1603	يوقظ	3	2.64	0.23	1653	خامس	3	2.64	0.23
1604	قبة	3	2.64	0.23	1654	وابوالها	3	2.64	0.23
1605	وسالنا	3	2.64	0.23	1655	خيات	3	2.64	0.23
1606	بالطويل	3	2.64	0.23	1656	امتك	3	2.64	0.23
1607	حلق	3	2.64	0.23	1657	يزد	3	2.64	0.23
1608	الكبائر	3	2.64	0.23	1658	فحدثني	3	2.64	0.23
1609	اسحق	3	2.64	0.23	1659	قليل	3	2.64	0.23
1610	شجر	3	2.64	0.23	1660	المجان	3	2.64	0.23
1611	فقتل	3	2.64	0.23	1661	خان	3	2.64	0.23
1612	بالقصير	3	2.64	0.23	1662	ينام	3	2.64	0.23
1613	عد	3	2.64	0.23	1663	هلك	3	2.64	0.23
1614	اكتوي	3	2.64	0.23	1664	فشربه	3	2.64	0.23
1615	نعاليهم	3	2.64	0.23	1665	ركعة	3	2.64	0.23
1616	بالفتاء	3	2.64	0.23	1666	الساعي	3	2.64	0.23
1617	يؤدي	3	2.64	0.23	1667	بيبيك	3	2.64	0.23
1618	ويقوم	3	2.64	0.23	1668	كبشين	3	2.64	0.23
1619	حلالا	3	2.64	0.23	1669	فشق	3	2.64	0.23
1620	يوتر	3	2.64	0.23	1670	فشكر	3	2.64	0.23
1621	أوئمن	3	2.64	0.23	1671	السبع	3	2.64	0.23
1622	فكرهت	3	2.64	0.23	1672	خالقهم	3	2.64	0.23
1623	فقرانهم	3	2.64	0.23	1673	لتلبسها	3	2.64	0.23
1624	الصلوات	3	2.64	0.23	1674	تلقي	3	2.64	0.23
1625	حلاوة	3	2.64	0.23	1675	رواه	3	2.64	0.23
1626	يؤتى	3	2.64	0.23	1676	حوله	3	2.64	0.23
1627	انصر	3	2.64	0.23	1677	المظلوم	3	2.64	0.23
1628	نساءك	3	2.64	0.23	1678	بالوفد	3	2.64	0.23
1629	يوحى	3	2.64	0.23	1679	شماله	3	2.64	0.23
1630	ويل	3	2.64	0.23	1680	تغلبوا	3	2.64	0.23
1631	شانهما	3	2.64	0.23	1681	شكوت	3	2.64	0.23
1632	انظروا	3	2.64	0.23	1682	حضر	3	2.64	0.23
1633	رمت	3	2.64	0.23	1683	قتال	3	2.64	0.23
1634	خالتي	3	2.64	0.23	1684	ياكلون	3	2.64	0.23
1635	فصام	3	2.64	0.23	1685	قومه	3	2.64	0.23
1636	فصبيت	3	2.64	0.23	1686	بام	3	2.64	0.23
1637	يبعث	3	2.64	0.23	1687	ويرفع	3	2.64	0.23
1638	ينبع	3	2.64	0.23	1688	النسهم	3	2.64	0.23
1639	ركعتان	3	2.64	0.23	1689	نخلة	3	2.64	0.23
1640	عهدهم	3	2.64	0.23	1690	اكون	3	2.64	0.23
1641	مسه	3	2.64	0.23	1691	فعرقت	3	2.64	0.23
1642	فصل	3	2.64	0.23	1692	المطرقة	3	2.64	0.23
1643	قوة	3	2.64	0.23	1693	ورقها	3	2.64	0.23
1644	ابنة	3	2.64	0.23	1694	نهار	3	2.64	0.23

Feature No	feature	DF	IDF	Feature redundan	Feature No	feature	DF	IDF	Feature redundan
1695	الطاعون	3	2.64	0.23	1745	بكنيتي	3	2.64	0.23
1696	وسعي	3	2.64	0.23	1746	وفضل	3	2.64	0.23
1697	بيبيت	3	2.64	0.23	1747	مخافة	3	2.64	0.23
1698	بيرون	3	2.64	0.23	1748	مخاض	3	2.64	0.23
1699	تلا	3	2.64	0.23	1749	للصلاة	3	2.64	0.23
1700	الضيف	3	2.64	0.23	1750	بكفر	3	2.64	0.23
1701	استطيع	3	2.64	0.23	1751	يطلع	3	2.64	0.23
1702	شهيد	3	2.64	0.23	1752	يطوفوا	3	2.64	0.23
1703	سواد	3	2.64	0.23	1753	فاخذه	3	2.64	0.23
1704	فليوزن	3	2.64	0.23	1754	احتلمت	3	2.64	0.23
1705	سواهما	3	2.64	0.23	1755	املك	3	2.64	0.23
1706	سودة	3	2.64	0.23	1756	قريشا	3	2.64	0.23
1707	شهدت	3	2.64	0.23	1757	فالتمس	3	2.64	0.23
1708	اشار	3	2.64	0.23	1758	بكلمة	3	2.64	0.23
1709	بالحق	3	2.64	0.23	1759	بغلس	3	2.64	0.23
1710	نحرت	3	2.64	0.23	1760	يدعو	3	2.64	0.23
1711	خاتم	3	2.64	0.23	1761	والنهي	3	2.64	0.23
1712	ابعد	3	2.64	0.23	1762	والنهار	3	2.64	0.23
1713	ويحيى	3	2.64	0.23	1763	وقم	3	2.64	0.23
1714	هات	3	2.64	0.23	1764	ارجعوا	3	2.64	0.23
1715	حصير	3	2.64	0.23	1765	اجعل	3	2.64	0.23
1716	ابطيه	3	2.64	0.23	1766	عاتقيه	3	2.64	0.23
1717	يامر	3	2.64	0.23	1767	عاتقه	3	2.64	0.23
1718	ويخرج	3	2.64	0.23	1768	وحده	3	2.64	0.23
1719	حيسا	3	2.64	0.23	1769	ساله	3	2.64	0.23
1720	غسله	3	2.64	0.23	1770	انتصف	3	2.64	0.23
1721	افلح	3	2.64	0.23	1771	التريد	3	2.64	0.23
1722	فامروا	3	2.64	0.23	1772	المنزرين	3	2.64	0.23
1723	بكتاب	3	2.64	0.23	1773	تحل	3	2.64	0.23
1724	عكة	3	2.64	0.23	1774	جننا	3	2.64	0.23
1725	راحة	3	2.64	0.23	1775	يحررم	3	2.64	0.23
1726	بقين	3	2.64	0.23	1776	اجرت	3	2.64	0.23
1727	وانطلق	3	2.64	0.23	1777	فاخيرني	3	2.64	0.23
1728	ارايتم	3	2.64	0.23	1778	لصاحبه	3	2.64	0.23
1729	ارب	3	2.64	0.23	1779	مريض	3	2.64	0.23
1730	املحين	3	2.64	0.23	1780	وقرا	3	2.64	0.23
1731	بقوم	3	2.64	0.23	1781	التلول	3	2.64	0.23
1732	اولادكن	3	2.64	0.23	1782	وحتى	3	2.64	0.23
1733	مركبا	3	2.64	0.23	1783	لايتيها	3	2.64	0.23
1734	التراب	3	2.64	0.23	1784	بقر	3	2.64	0.23
1735	فامرہ	3	2.64	0.23	1785	المنكر	3	2.64	0.23
1736	فاخذت	3	2.64	0.23	1786	الغني	3	2.64	0.23
1737	رحمة	3	2.64	0.23	1787	ثدي	3	2.64	0.23
1738	محلہ	3	2.64	0.23	1788	فاخرجت	3	2.64	0.23
1739	الثنية	3	2.64	0.23	1789	ملتحفا	3	2.64	0.23
1740	احداكن	3	2.64	0.23	1790	يعضد	3	2.64	0.23
1741	فاخذها	3	2.64	0.23	1791	فياتي	3	2.64	0.23
1742	فاقمنا	3	2.64	0.23	1792	فانطلقت	3	2.64	0.23
1743	اراك	3	2.64	0.23	1793	الف	3	2.64	0.23
1744	تركته	3	2.64	0.23	1794	اعطي	3	2.64	0.23

Feature No	feature	DF	IDF	Feature redundancy	Feature No	feature	DF	IDF	Feature redundancy
1795	وحجة	3	2.64	0.23	1845	البعير	3	2.64	0.23
1796	اويننا	3	2.64	0.23	1846	جلوسا	3	2.64	0.23
1797	بلحم	3	2.64	0.23	1847	تخرجوا	3	2.64	0.23
1798	رجله	3	2.64	0.23	1848	اذرع	3	2.64	0.23
1799	وتصوم	3	2.64	0.23	1849	جليسه	3	2.64	0.23
1800	امنوا	3	2.64	0.23	1850	معادن	3	2.64	0.23
1801	فاستسقى	3	2.64	0.23	1851	فافطر	3	2.64	0.23
1802	معمر	3	2.64	0.23	1852	الفريقين	3	2.64	0.23
1803	احمر	3	2.64	0.23	1853	جماعة	3	2.64	0.23
1804	بنى	3	2.64	0.23	1854	الامام	3	2.64	0.23
1805	وعد	3	2.64	0.23	1855	امن	3	2.64	0.23
1806	افاضت	3	2.64	0.23	1856	كسوف	3	2.64	0.23
1807	الاناء	3	2.64	0.23	1857	ثمانون	3	2.64	0.23
1808	لعينك	3	2.64	0.23	1858	اهديت	3	2.64	0.23
1809	عطية	3	2.64	0.23	1859	وحدثني	3	2.64	0.23
1810	مرضه	3	2.64	0.23	1860	عذب	3	2.64	0.23
1811	البحرين	3	2.64	0.23	1861	البنات	3	2.64	0.23
1812	فاشار	3	2.64	0.23	1862	فاصبح	3	2.64	0.23
1813	فاشترى	3	2.64	0.23	1863	راينا	3	2.64	0.23
1814	بنار	3	2.64	0.23	1864	فادر كنا	3	2.64	0.23
1815	يشير	3	2.64	0.23	1865	كهينكم	3	2.64	0.23
1816	تخلف	3	2.64	0.23	1866	فاطال	3	2.64	0.23
1817	تدخل	3	2.64	0.23	1867	يسلم	3	2.64	0.23
1818	اخاك	3	2.64	0.23	1868	فادخل	3	2.64	0.23
1819	يسير	3	2.64	0.23	1869	احسبه	3	2.64	0.23
1820	اخبرتني	3	2.64	0.23	1870	عرفها	3	2.64	0.23
1821	القتل	3	2.64	0.23	1871	اهلك	3	2.64	0.23
1822	رجلا	3	2.64	0.23	1872	وجع	3	2.64	0.23
1823	فارتدا	3	2.64	0.23	1873	قرا	3	2.64	0.23
1824	يشبع	3	2.64	0.23	1874	تراه	3	2.64	0.23
1825	قدمي	3	2.64	0.23	1875	الكعبيين	3	2.64	0.23
1826	تدغرن	3	2.64	0.23	1876	مغلقا	3	2.64	0.23
1827	عسى	3	2.64	0.23	1877	اخذت	3	2.64	0.23
1828	الكلب	3	2.64	0.23	1878	فيشربوا	3	2.64	0.23
1829	النفاق	3	2.64	0.23	1879	الفيء	3	2.64	0.23
1830	يدري	3	2.64	0.23	1880	البراء	3	2.64	0.23
1831	وعصرت	3	2.64	0.23	1881	الامم	3	2.64	0.23
1832	الانبياء	3	2.64	0.23	1882	ادع	3	2.64	0.23
1833	وجاء	3	2.64	0.23	1883	فوجدت	3	2.64	0.23
1834	الايام	3	2.64	0.23	1884	طرف	3	2.64	0.23
1835	زوجها	3	2.64	0.23	1885	وصم	3	2.64	0.23
1836	رهم	3	2.64	0.23	1886	الكراع	3	2.64	0.23
1837	يشفع	3	2.64	0.23	1887	انني	3	2.64	0.23
1838	كف	3	2.64	0.23	1888	ادعو	3	2.64	0.23
1839	يدركه	3	2.64	0.23	1889	اطم	3	2.64	0.23
1840	فافرغ	3	2.64	0.23	1890	فقلت	3	2.64	0.23
1841	البيت	3	2.64	0.23	1891	والصبح	3	2.64	0.23
1842	بمثل	3	2.64	0.23	1892	اثارهما	3	2.64	0.23
1843	مقام	3	2.64	0.23	1893	حجتك	3	2.64	0.23
1844	جنته	3	2.64	0.23	1894	لقيت	3	2.64	0.23

Feature No	feature	DF	IDF	Feature redundancy	Feature No	feature	DF	IDF	Feature redundancy
1895	حجته	3	2.64	0.23	1945	فهينا	3	2.64	0.23
1896	بضع	3	2.64	0.23	1946	يقرب	3	2.64	0.23
1897	بطنها	3	2.64	0.23	1947	والدار	3	2.64	0.23
1898	حجاب	3	2.64	0.23	1948	الحديبية	3	2.64	0.23
1899	الجهل	3	2.64	0.23	1949	والخميس	3	2.64	0.23
1900	المنازل	3	2.64	0.23	1950	انهى	3	2.64	0.23
1901	سنة	3	2.64	0.23	1951	بصري	3	2.64	0.23
1902	تسع	3	2.64	0.23	1952	فجدح	3	2.64	0.23
1903	ادويتكم	3	2.64	0.23	1953	حجرة	3	2.64	0.23
1904	امري	3	2.64	0.23	1954	سجودا	3	2.64	0.23
1905	عيني	3	2.64	0.23	1955	الحسنة	3	2.64	0.23
1906	لقمة	3	2.64	0.23	1956	وليقطعهما	3	2.64	0.23
1907	وجبت	3	2.64	0.23	1957	الحرم	3	2.64	0.23
1908	وقلدت	3	2.64	0.23	1958	دم	3	2.64	0.23
1909	زعم	3	2.64	0.23	1959	الثياب	3	2.64	0.23
1910	فتصدق	3	2.64	0.23	1960	فيكلمة	3	2.64	0.23
1911	بيوتكم	3	2.64	0.23	1961	قصصا	3	2.64	0.23
1912	قضي	3	2.64	0.23	1962	والكافر	3	2.64	0.23
1913	يجر	3	2.64	0.23	1963	والقوم	3	2.64	0.23
1914	طريق	3	2.64	0.23	1964	فاحشا	3	2.64	0.23
1915	وكثرة	3	2.64	0.23	1965	ارسلني	3	2.64	0.23
1916	يقبض	3	2.64	0.23	1966	مجاهد	3	2.64	0.23
1917	زعموا	3	2.64	0.23	1967	دين	3	2.64	0.23
1918	تؤدوا	3	2.64	0.23	1968	يعود	3	2.64	0.23
1919	ارضين	3	2.64	0.23	1969	يجمع	3	2.64	0.23
1920	الملك	3	2.64	0.23	1970	ظل	3	2.64	0.23
1921	قاتل	3	2.64	0.23	1971	يحتز	3	2.64	0.23
1922	منكنا	3	2.64	0.23	1972	سبعون	3	2.64	0.23
1923	فجعله	3	2.64	0.23	1973	ارسلت	3	2.64	0.23
1924	المتعة	3	2.64	0.23	1974	جاءك	3	2.64	0.23
1925	ضعف	3	2.64	0.23	1975	يغادر	3	2.64	0.23
1926	الوجه	3	2.64	0.23	1976	والمرأة	3	2.64	0.23
1927	ارني	3	2.64	0.23	1977	مضطجع	3	2.64	0.23
1928	يقصروا	3	2.64	0.23	1978	ارسلك	3	2.64	0.23
1929	قطع	3	2.64	0.23	1979	الاعمال	3	2.64	0.23
1930	اطاعوا	3	2.64	0.23	1980	اعجبهم	3	2.64	0.23
1931	اضيافاك	3	2.64	0.23	1981	منافقا	3	2.64	0.23
1932	يقطر	3	2.64	0.23	1982	فيرا	3	2.64	0.23
1933	خيارهم	3	2.64	0.23	1983	اقبلت	3	2.64	0.23
1934	بشراب	3	2.64	0.23	1984	يحيه	3	2.64	0.23
1935	وليه	3	2.64	0.23	1985	يحتني	3	2.64	0.23
1936	باسمي	3	2.64	0.23	1986	وكل	3	2.64	0.23
1937	كدت	3	2.64	0.23	1987	فتنبعهم	3	2.64	0.23
1938	والحديث	3	2.64	0.23	1988	فاوحى	3	2.64	0.23
1939	يقذف	3	2.64	0.23	1989	مثلي	3	2.64	0.23
1940	طافت	3	2.64	0.23	1990	سبق	3	2.64	0.23
1941	اطف	3	2.64	0.23	1991	المائة	3	2.64	0.23
1942	يقبله	3	2.64	0.23	1992	دعه	3	2.64	0.23
1943	يقبلها	3	2.64	0.23	1993	منزل	3	2.64	0.23
1944	جحيفة	3	2.64	0.23	1994	اعتم	3	2.64	0.23

Feature No	feature	DF	IDF	Feature redundancy	Feature No	feature	DF	IDF	Feature redundancy
1995	اعتق	3	2.64	0.23	2045	فضالة	2	2.82	0.15
1996	لزورك	3	2.64	0.23	2046	فرجة	2	2.82	0.15
1997	ردف	3	2.64	0.23	2047	بعذبان	2	2.82	0.15
1998	لسان	3	2.64	0.23	2048	عذاب	2	2.82	0.15
1999	بعيد	3	2.64	0.23	2049	ازواج	2	2.82	0.15
2000	لاخيه	3	2.64	0.23	2050	والظن	2	2.82	0.15
2001	والفضة	3	2.64	0.23	2051	بييسا	2	2.82	0.15
2002	اردنا	3	2.64	0.23	2052	يكفرن	2	2.82	0.15
2003	حبله	3	2.64	0.23	2053	عملت	2	2.82	0.15
2004	لزوجك	3	2.64	0.23	2054	بالنميمة	2	2.82	0.15
2005	رحيما	3	2.64	0.23	2055	تحسسوا	2	2.82	0.15
2006	لعشيرة	2	2.82	0.15	2056	بخوف	2	2.82	0.15
2007	ظننا	2	2.82	0.15	2057	كسرت	2	2.82	0.15
2008	درهما	2	2.82	0.15	2058	العلاء	2	2.82	0.15
2009	اعصم	2	2.82	0.15	2059	الوجهين	2	2.82	0.15
2010	تطوع	2	2.82	0.15	2060	الوصال	2	2.82	0.15
2011	اسقه	2	2.82	0.15	2061	انسية	2	2.82	0.15
2012	حكيم	2	2.82	0.15	2062	كنفه	2	2.82	0.15
2013	صياد	2	2.82	0.15	2063	النجار	2	2.82	0.15
2014	تقبل	2	2.82	0.15	2064	الموقف	2	2.82	0.15
2015	سيسمييه	2	2.82	0.15	2065	قريب	2	2.82	0.15
2016	لاراه	2	2.82	0.15	2066	القتال	2	2.82	0.15
2017	عسلا	2	2.82	0.15	2067	المنذر	2	2.82	0.15
2018	طالت	2	2.82	0.15	2068	متقبلة	2	2.82	0.15
2019	الهرج	2	2.82	0.15	2069	فليطمعه	2	2.82	0.15
2020	يشاء	2	2.82	0.15	2070	قاله	2	2.82	0.15
2021	اكذب	2	2.82	0.15	2071	الكلام	2	2.82	0.15
2022	بالصفا	2	2.82	0.15	2072	باسا	2	2.82	0.15
2023	تجسسوا	2	2.82	0.15	2073	مكفي	2	2.82	0.15
2024	ونهاهم	2	2.82	0.15	2074	كلام	2	2.82	0.15
2025	القبر	2	2.82	0.15	2075	لاخبركم	2	2.82	0.15
2026	فاشارت	2	2.82	0.15	2076	العشيرة	2	2.82	0.15
2027	شعائر	2	2.82	0.15	2077	قسمة	2	2.82	0.15
2028	امنهم	2	2.82	0.15	2078	اوذي	2	2.82	0.15
2029	الحكم	2	2.82	0.15	2079	اهاجر	2	2.82	0.15
2030	الخبز	2	2.82	0.15	2080	بالبيئات	2	2.82	0.15
2031	ففرج	2	2.82	0.15	2081	ليهود	2	2.82	0.15
2032	اذنت	2	2.82	0.15	2082	المرتاب	2	2.82	0.15
2033	ابوي	2	2.82	0.15	2083	المسيح	2	2.82	0.15
2034	بريرة	2	2.82	0.15	2084	مظلوما	2	2.82	0.15
2035	بيستتر	2	2.82	0.15	2085	قلد	2	2.82	0.15
2036	اشعر	2	2.82	0.15	2086	مرارا	2	2.82	0.15
2037	امرك	2	2.82	0.15	2087	النجوى	2	2.82	0.15
2038	كدا	2	2.82	0.15	2088	كلمته	2	2.82	0.15
2039	عكاشة	2	2.82	0.15	2089	مجالسنا	2	2.82	0.15
2040	عدي	2	2.82	0.15	2090	الغشى	2	2.82	0.15
2041	اليهودي	2	2.82	0.15	2091	مقدم	2	2.82	0.15
2042	ينفت	2	2.82	0.15	2092	مثقال	2	2.82	0.15
2043	الدرداء	2	2.82	0.15	2093	لاباس	2	2.82	0.15
2044	الظن	2	2.82	0.15	2094	متعة	2	2.82	0.15

Feature No	feature	DF	IDF	Feature redundanc	Feature No	feature	DF	IDF	Feature redundanc
2095	باكثر	2	2.82	0.15	2145	فاصحوا	2	2.82	0.15
2096	لموقنا	2	2.82	0.15	2146	فاستشفع	2	2.82	0.15
2097	مخرمة	2	2.82	0.15	2147	البارحة	2	2.82	0.15
2098	كفه	2	2.82	0.15	2148	وتغير	2	2.82	0.15
2099	الظهيرة	2	2.82	0.15	2149	بوجه	2	2.82	0.15
2100	عرفات	2	2.82	0.15	2150	بوله	2	2.82	0.15
2101	ورحمة	2	2.82	0.15	2151	فادع	2	2.82	0.15
2102	غفار	2	2.82	0.15	2152	اصب	2	2.82	0.15
2103	تبوك	2	2.82	0.15	2153	يعمل	2	2.82	0.15
2104	تجلاني	2	2.82	0.15	2154	يغوث	2	2.82	0.15
2105	وسعديك	2	2.82	0.15	2155	ونحبه	2	2.82	0.15
2106	علمنا	2	2.82	0.15	2156	جاهلية	2	2.82	0.15
2107	علمك	2	2.82	0.15	2157	اخراانكم	2	2.82	0.15
2108	وعسى	2	2.82	0.15	2158	رفعت	2	2.82	0.15
2109	اقتدرون	2	2.82	0.15	2159	رعوفة	2	2.82	0.15
2110	فرضها	2	2.82	0.15	2160	يسنزه	2	2.82	0.15
2111	عريان	2	2.82	0.15	2161	رطب	2	2.82	0.15
2112	وجعه	2	2.82	0.15	2162	يستلمه	2	2.82	0.15
2113	اغرها	2	2.82	0.15	2163	رخص	2	2.82	0.15
2114	ترفع	2	2.82	0.15	2164	جف	2	2.82	0.15
2115	وفيم	2	2.82	0.15	2165	جهل	2	2.82	0.15
2116	عبدة	2	2.82	0.15	2166	ذرة	2	2.82	0.15
2117	وكثر	2	2.82	0.15	2167	جائزته	2	2.82	0.15
2118	ظالما	2	2.82	0.15	2168	اثير	2	2.82	0.15
2119	وكسرت	2	2.82	0.15	2169	يدخلني	2	2.82	0.15
2120	طلق	2	2.82	0.15	2170	دخلوا	2	2.82	0.15
2121	اظن	2	2.82	0.15	2171	دحية	2	2.82	0.15
2122	طرفي	2	2.82	0.15	2172	انق	2	2.82	0.15
2123	وايلبسه	2	2.82	0.15	2173	خلاق	2	2.82	0.15
2124	ومضمضنا	2	2.82	0.15	2174	خشبة	2	2.82	0.15
2125	افتح	2	2.82	0.15	2175	خذ	2	2.82	0.15
2126	والهدى	2	2.82	0.15	2176	ابوا	2	2.82	0.15
2127	واعتقت	2	2.82	0.15	2177	حيطان	2	2.82	0.15
2128	بردا	2	2.82	0.15	2178	حياة	2	2.82	0.15
2129	بسويق	2	2.82	0.15	2179	حوالينا	2	2.82	0.15
2130	الحمس	2	2.82	0.15	2180	حليف	2	2.82	0.15
2131	فجلست	2	2.82	0.15	2181	انذونا	2	2.82	0.15
2132	فجاءت	2	2.82	0.15	2182	يغلبه	2	2.82	0.15
2133	قتلاحي	2	2.82	0.15	2183	استاذنا	2	2.82	0.15
2134	فتكلموا	2	2.82	0.15	2184	عد	2	2.82	0.15
2135	فتزوجها	2	2.82	0.15	2185	اشق	2	2.82	0.15
2136	والعمل	2	2.82	0.15	2186	وهيب	2	2.82	0.15
2137	فيدا	2	2.82	0.15	2187	صالحا	2	2.82	0.15
2138	الازار	2	2.82	0.15	2188	ووضع	2	2.82	0.15
2139	فاوحي	2	2.82	0.15	2189	تفتنون	2	2.82	0.15
2140	فاجبنا	2	2.82	0.15	2190	شفاني	2	2.82	0.15
2141	فانفري	2	2.82	0.15	2191	اسمك	2	2.82	0.15
2142	واليهود	2	2.82	0.15	2192	ويعطيه	2	2.82	0.15
2143	بقول	2	2.82	0.15	2193	ويهل	2	2.82	0.15
2144	بليلة	2	2.82	0.15	2194	سوقك	2	2.82	0.15

Feature No	feature	DF	IDF	Feature redundan	Feature No	feature	DF	IDF	Feature redundan
2195	سلول	2	2.82	0.15	2245	كالمجاهد	2	2.82	0.15
2196	جارية	2	2.82	0.15	2246	تكتنوا	2	2.82	0.15
2197	يتحدثون	2	2.82	0.15	2247	تكرهون	2	2.82	0.15
2198	صل	2	2.82	0.15	2248	كاليوم	2	2.82	0.15
2199	اساس	2	2.82	0.15	2249	شي	2	2.82	0.15
2200	سقت	2	2.82	0.15	2250	كاملا	2	2.82	0.15
2201	ازاري	2	2.82	0.15	2251	سيورثه	2	2.82	0.15
2202	تنشرت	2	2.82	0.15	2252	كاهلال	2	2.82	0.15
2203	اريتها	2	2.82	0.15	2253	كانك	2	2.82	0.15
2204	سجوده	2	2.82	0.15	2254	كبد	2	2.82	0.15
2205	سجدت	2	2.82	0.15	2255	سيده	2	2.82	0.15
2206	سنرت	2	2.82	0.15	2256	سيئه	2	2.82	0.15
2207	يحبنا	2	2.82	0.15	2257	بالنبي	2	2.82	0.15
2208	ارحم	2	2.82	0.15	2258	تلبسوا	2	2.82	0.15
2209	ساعتي	2	2.82	0.15	2259	سهما	2	2.82	0.15
2210	يخيل	2	2.82	0.15	2260	تنتهك	2	2.82	0.15
2211	سلمان	2	2.82	0.15	2261	تمارى	2	2.82	0.15
2212	نسبية	2	2.82	0.15	2262	تملوا	2	2.82	0.15
2213	واحب	2	2.82	0.15	2263	كانا	2	2.82	0.15
2214	السبي	2	2.82	0.15	2264	انتظر	2	2.82	0.15
2215	الزناد	2	2.82	0.15	2265	انتقم	2	2.82	0.15
2216	فصبر	2	2.82	0.15	2266	سبقك	2	2.82	0.15
2217	النس	2	2.82	0.15	2267	كذبة	2	2.82	0.15
2218	الثان	2	2.82	0.15	2268	كذبتم	2	2.82	0.15
2219	فشدد	2	2.82	0.15	2269	سبعين	2	2.82	0.15
2220	بحلة	2	2.82	0.15	2270	سبعمانه	2	2.82	0.15
2221	الربيع	2	2.82	0.15	2271	كرامة	2	2.82	0.15
2222	فقلته	2	2.82	0.15	2272	فعلى	2	2.82	0.15
2223	السحاب	2	2.82	0.15	2273	تنتب	2	2.82	0.15
2224	فسمع	2	2.82	0.15	2274	سياب	2	2.82	0.15
2225	بيبطن	2	2.82	0.15	2275	انتهى	2	2.82	0.15
2226	كتابا	2	2.82	0.15	2276	ثابتا	2	2.82	0.15
2227	سموا	2	2.82	0.15	2277	فعليكم	2	2.82	0.15
2228	تماريت	2	2.82	0.15	2278	ثبير	2	2.82	0.15
2229	سمعا	2	2.82	0.15	2279	كرسي	2	2.82	0.15
2230	سما	2	2.82	0.15	2280	كرهوا	2	2.82	0.15
2231	فعرض	2	2.82	0.15	2281	انينكم	2	2.82	0.15
2232	سله	2	2.82	0.15	2282	سالماها	2	2.82	0.15
2233	بالموعظة	2	2.82	0.15	2283	انينت	2	2.82	0.15
2234	سلام	2	2.82	0.15	2284	كسب	2	2.82	0.15
2235	فعرف	2	2.82	0.15	2285	ثائر	2	2.82	0.15
2236	فلقد	2	2.82	0.15	2286	انجشة	2	2.82	0.15
2237	كنفيه	2	2.82	0.15	2287	سقاينه	2	2.82	0.15
2238	كئفه	2	2.82	0.15	2288	سفرة	2	2.82	0.15
2239	سكرجة	2	2.82	0.15	2289	انسانيه	2	2.82	0.15
2240	تمسك	2	2.82	0.15	2290	فعلموا	2	2.82	0.15
2241	سقيما	2	2.82	0.15	2291	شبعنا	2	2.82	0.15
2242	بالمعوذات	2	2.82	0.15	2292	انجلت	2	2.82	0.15
2243	سيف	2	2.82	0.15	2293	تتحيت	2	2.82	0.15
2244	تقوموا	2	2.82	0.15	2294	فلتاجروا	2	2.82	0.15

Feature No	feature	DF	IDF	Feature redundan	Feature No	feature	DF	IDF	Feature redundan
2295	سريت	2	2.82	0.15	2345	صنعتة	2	2.82	0.15
2296	سبيله	2	2.82	0.15	2346	صنعت	2	2.82	0.15
2297	سره	2	2.82	0.15	2347	صليتها	2	2.82	0.15
2298	سترون	2	2.82	0.15	2348	تطوه	2	2.82	0.15
2299	سرنا	2	2.82	0.15	2349	صلحت	2	2.82	0.15
2300	تنفروا	2	2.82	0.15	2350	تطمع	2	2.82	0.15
2301	تنفع	2	2.82	0.15	2351	صلاتي	2	2.82	0.15
2302	سراقة	2	2.82	0.15	2352	صلاتهم	2	2.82	0.15
2303	تتكسر	2	2.82	0.15	2353	فضمته	2	2.82	0.15
2304	فعلوا	2	2.82	0.15	2354	انقذه	2	2.82	0.15
2305	فلبث	2	2.82	0.15	2355	تقاسموا	2	2.82	0.15
2306	تواصلوا	2	2.82	0.15	2356	شطر	2	2.82	0.15
2307	فعلم	2	2.82	0.15	2357	تفعلوا	2	2.82	0.15
2308	بالمزلفة	2	2.82	0.15	2358	تفور	2	2.82	0.15
2309	تعدو	2	2.82	0.15	2359	تقاتلوا	2	2.82	0.15
2310	انشا	2	2.82	0.15	2360	تقاتلون	2	2.82	0.15
2311	صفراء	2	2.82	0.15	2361	انطلقى	2	2.82	0.15
2312	انفقت	2	2.82	0.15	2362	كائن	2	2.82	0.15
2313	صدقته	2	2.82	0.15	2363	انطلقت	2	2.82	0.15
2314	نظهر	2	2.82	0.15	2364	شرح	2	2.82	0.15
2315	صدري	2	2.82	0.15	2365	صار	2	2.82	0.15
2316	بانا	2	2.82	0.15	2366	كادت	2	2.82	0.15
2317	صخرة	2	2.82	0.15	2367	شعرا	2	2.82	0.15
2318	قلوبكما	2	2.82	0.15	2368	تقام	2	2.82	0.15
2319	انفكت	2	2.82	0.15	2369	كاسية	2	2.82	0.15
2320	صحب	2	2.82	0.15	2370	تقبلون	2	2.82	0.15
2321	فضمها	2	2.82	0.15	2371	تقدموا	2	2.82	0.15
2322	قلوبهم	2	2.82	0.15	2372	ساق	2	2.82	0.15
2323	تعرض	2	2.82	0.15	2373	كالبنين	2	2.82	0.15
2324	صبيا	2	2.82	0.15	2374	انام	2	2.82	0.15
2325	باليمن	2	2.82	0.15	2375	شبية	2	2.82	0.15
2326	تعسروا	2	2.82	0.15	2376	فلقيني	2	2.82	0.15
2327	فظن	2	2.82	0.15	2377	شرب	2	2.82	0.15
2328	صاع	2	2.82	0.15	2378	شق	2	2.82	0.15
2329	قمت	2	2.82	0.15	2379	شاهدا	2	2.82	0.15
2330	قميص	2	2.82	0.15	2380	صاحبتها	2	2.82	0.15
2331	تعد	2	2.82	0.15	2381	قولا	2	2.82	0.15
2332	تضر	2	2.82	0.15	2382	صابرا	2	2.82	0.15
2333	تصدقت	2	2.82	0.15	2383	تعلمون	2	2.82	0.15
2334	تصدقن	2	2.82	0.15	2384	شيوخا	2	2.82	0.15
2335	صياهم	2	2.82	0.15	2385	شبيئا	2	2.82	0.15
2336	صورتى	2	2.82	0.15	2386	تعول	2	2.82	0.15
2337	تصبيه	2	2.82	0.15	2387	شممت	2	2.82	0.15
2338	بالحجاب	2	2.82	0.15	2388	فلكناه	2	2.82	0.15
2339	صوتا	2	2.82	0.15	2389	فعاد	2	2.82	0.15
2340	انقص	2	2.82	0.15	2390	تفعل	2	2.82	0.15
2341	تضامون	2	2.82	0.15	2391	فعاذه	2	2.82	0.15
2342	بالحديث	2	2.82	0.15	2392	انفت	2	2.82	0.15
2343	صواحب	2	2.82	0.15	2393	فعدت	2	2.82	0.15
2344	انفري	2	2.82	0.15	2394	شفاؤك	2	2.82	0.15

Feature No	feature	DF	IDF	Feature redundan	Feature No	feature	DF	IDF	Feature redundan
2395	انفا	2	2.82	0.15	2445	حجاج	2	2.82	0.15
2396	تفسير	2	2.82	0.15	2446	حجارة	2	2.82	0.15
2397	شعف	2	2.82	0.15	2447	امرین	2	2.82	0.15
2398	شعره	2	2.82	0.15	2448	حمید	2	2.82	0.15
2399	قیصر	2	2.82	0.15	2449	خالصا	2	2.82	0.15
2400	صاحبي	2	2.82	0.15	2450	خالدا	2	2.82	0.15
2401	شكوا	2	2.82	0.15	2451	خاصم	2	2.82	0.15
2402	خيلبي	2	2.82	0.15	2452	حشفة	2	2.82	0.15
2403	ثريد	2	2.82	0.15	2453	حشوها	2	2.82	0.15
2404	داوم	2	2.82	0.15	2454	حصن	2	2.82	0.15
2405	داره	2	2.82	0.15	2455	خابوا	2	2.82	0.15
2406	حجبت	2	2.82	0.15	2456	فقل	2	2.82	0.15
2407	خبيبة	2	2.82	0.15	2457	حياء	2	2.82	0.15
2408	خياط	2	2.82	0.15	2458	خلا	2	2.82	0.15
2409	فكره	2	2.82	0.15	2459	حوالي	2	2.82	0.15
2410	خياركم	2	2.82	0.15	2460	حسد	2	2.82	0.15
2411	حجرتي	2	2.82	0.15	2461	حقة	2	2.82	0.15
2412	لارى	2	2.82	0.15	2462	حمراء	2	2.82	0.15
2413	خمارا	2	2.82	0.15	2463	حمت	2	2.82	0.15
2414	فقدت	2	2.82	0.15	2464	حماد	2	2.82	0.15
2415	خلفه	2	2.82	0.15	2465	لبركتها	2	2.82	0.15
2416	لاسلانك	2	2.82	0.15	2466	لبس	2	2.82	0.15
2417	لاصحابه	2	2.82	0.15	2467	حلوة	2	2.82	0.15
2418	خلق	2	2.82	0.15	2468	الين	2	2.82	0.15
2419	امراتين	2	2.82	0.15	2469	فقهوا	2	2.82	0.15
2420	خلفه	2	2.82	0.15	2470	بالصوم	2	2.82	0.15
2421	لاعطى	2	2.82	0.15	2471	فقرب	2	2.82	0.15
2422	خلفا	2	2.82	0.15	2472	جوف	2	2.82	0.15
2423	فقدوا	2	2.82	0.15	2473	خفه	2	2.82	0.15
2424	خوان	2	2.82	0.15	2474	حديثه	2	2.82	0.15
2425	لاخيك	2	2.82	0.15	2475	خطبنا	2	2.82	0.15
2426	ذاكر	2	2.82	0.15	2476	خطب	2	2.82	0.15
2427	حائط	2	2.82	0.15	2477	خضرة	2	2.82	0.15
2428	حابس	2	2.82	0.15	2478	خضرا	2	2.82	0.15
2429	دونهن	2	2.82	0.15	2479	خضر	2	2.82	0.15
2430	امسى	2	2.82	0.15	2480	خصوصة	2	2.82	0.15
2431	دونها	2	2.82	0.15	2481	حسناات	2	2.82	0.15
2432	لاحلت	2	2.82	0.15	2482	لاقى	2	2.82	0.15
2433	دنبا	2	2.82	0.15	2483	خب	2	2.82	0.15
2434	دنونا	2	2.82	0.15	2484	حرمه	2	2.82	0.15
2435	فقدنا	2	2.82	0.15	2485	خذوا	2	2.82	0.15
2436	دناتير	2	2.82	0.15	2486	فقسماها	2	2.82	0.15
2437	حدثوني	2	2.82	0.15	2487	لام	2	2.82	0.15
2438	حالت	2	2.82	0.15	2488	حريصا	2	2.82	0.15
2439	دعوة	2	2.82	0.15	2489	خدرها	2	2.82	0.15
2440	فقامت	2	2.82	0.15	2490	خيزا	2	2.82	0.15
2441	فقاموا	2	2.82	0.15	2491	فقطعه	2	2.82	0.15
2442	فقتلوا	2	2.82	0.15	2492	حساب	2	2.82	0.15
2443	دسته	2	2.82	0.15	2493	حدثه	2	2.82	0.15
2444	حجا	2	2.82	0.15	2494	خصلة	2	2.82	0.15

Feature No	feature	DF	IDF	Feature redundanc	Feature No	feature	DF	IDF	Feature redundanc
2495	ففت	2	2.82	0.15	2545	راق	2	2.82	0.15
2496	فعلوا	2	2.82	0.15	2546	راغبة	2	2.82	0.15
2497	ركوعه	2	2.82	0.15	2547	راعي	2	2.82	0.15
2498	ركوع	2	2.82	0.15	2548	جمعوا	2	2.82	0.15
2499	ركوة	2	2.82	0.15	2549	جمل	2	2.82	0.15
2500	فغرف	2	2.82	0.15	2550	رؤيته	2	2.82	0.15
2501	ركبته	2	2.82	0.15	2551	جنازة	2	2.82	0.15
2502	ركبت	2	2.82	0.15	2552	رسولا	2	2.82	0.15
2503	رقدة	2	2.82	0.15	2553	رؤوسهم	2	2.82	0.15
2504	رقد	2	2.82	0.15	2554	امنت	2	2.82	0.15
2505	امورا	2	2.82	0.15	2555	جنة	2	2.82	0.15
2506	فغلبته	2	2.82	0.15	2556	ذنوبا	2	2.82	0.15
2507	جامعة	2	2.82	0.15	2557	ذكروا	2	2.82	0.15
2508	فكلم	2	2.82	0.15	2558	ففعت	2	2.82	0.15
2509	رطبة	2	2.82	0.15	2559	جهينة	2	2.82	0.15
2510	فكلاهما	2	2.82	0.15	2560	لابراهيم	2	2.82	0.15
2511	رضينا	2	2.82	0.15	2561	ذكرت	2	2.82	0.15
2512	رصافه	2	2.82	0.15	2562	فعلوا	2	2.82	0.15
2513	جذع	2	2.82	0.15	2563	سالته	2	2.82	0.15
2514	جذعة	2	2.82	0.15	2564	رؤوسهما	2	2.82	0.15
2515	فكل	2	2.82	0.15	2565	جعل	2	2.82	0.15
2516	ذاهبا	2	2.82	0.15	2566	رزقه	2	2.82	0.15
2517	رفيفا	2	2.82	0.15	2567	رديف	2	2.82	0.15
2518	بالشاء	2	2.82	0.15	2568	رداءه	2	2.82	0.15
2519	تصبح	2	2.82	0.15	2569	رداءك	2	2.82	0.15
2520	ثعلبة	2	2.82	0.15	2570	جزارتها	2	2.82	0.15
2521	قطعك	2	2.82	0.15	2571	ففرحنا	2	2.82	0.15
2522	ثلاثا	2	2.82	0.15	2572	كليهما	2	2.82	0.15
2523	فعلبه	2	2.82	0.15	2573	رحله	2	2.82	0.15
2524	ثمامة	2	2.82	0.15	2574	جسده	2	2.82	0.15
2525	ثمانى	2	2.82	0.15	2575	راوا	2	2.82	0.15
2526	ثوبان	2	2.82	0.15	2576	رجليها	2	2.82	0.15
2527	بالمرأة	2	2.82	0.15	2577	رايتموني	2	2.82	0.15
2528	كقتله	2	2.82	0.15	2578	جعلت	2	2.82	0.15
2529	بالمحصب	2	2.82	0.15	2579	رجعوا	2	2.82	0.15
2530	رميت	2	2.82	0.15	2580	رجعت	2	2.82	0.15
2531	جاءني	2	2.82	0.15	2581	ففزعت	2	2.82	0.15
2532	كفارة	2	2.82	0.15	2582	كنوزهما	2	2.82	0.15
2533	اناء	2	2.82	0.15	2583	ربط	2	2.82	0.15
2534	زرعة	2	2.82	0.15	2584	جلياب	2	2.82	0.15
2535	فكلوا	2	2.82	0.15	2585	جلد	2	2.82	0.15
2536	زاغت	2	2.82	0.15	2586	رايتني	2	2.82	0.15
2537	روينا	2	2.82	0.15	2587	ذبحت	2	2.82	0.15
2538	ان	2	2.82	0.15	2588	كموج	2	2.82	0.15
2539	جاع	2	2.82	0.15	2589	فاقول	2	2.82	0.15
2540	جنوع	2	2.82	0.15	2590	فيدخل	2	2.82	0.15
2541	زيت	2	2.82	0.15	2591	فسالته	2	2.82	0.15
2542	جنبا	2	2.82	0.15	2592	ايمانهم	2	2.82	0.15
2543	راهم	2	2.82	0.15	2593	فامرنا	2	2.82	0.15
2544	راكبة	2	2.82	0.15	2594	فالقها	2	2.82	0.15

Feature	feature	DF	IDF	Feature	Feature	feature	DF	IDF	Feature
2595	فالتفت	2	2.82	0.15	2645	فيضع	2	2.82	0.15
2596	فاكملوا	2	2.82	0.15	2646	فسقى	2	2.82	0.15
2597	بكروا	2	2.82	0.15	2647	فاسلم	2	2.82	0.15
2598	بكير	2	2.82	0.15	2648	فار تحلنا	2	2.82	0.15
2599	بقطرات	2	2.82	0.15	2649	فيقولون	2	2.82	0.15
2600	فيجربها	2	2.82	0.15	2650	فيقوم	2	2.82	0.15
2601	بقصعة	2	2.82	0.15	2651	فأذنت	2	2.82	0.15
2602	فيحطب	2	2.82	0.15	2652	فأدها	2	2.82	0.15
2603	فاقصر	2	2.82	0.15	2653	فأدمته	2	2.82	0.15
2604	ايسرهما	2	2.82	0.15	2654	فيكسر	2	2.82	0.15
2605	فاقدروا	2	2.82	0.15	2655	بيتنا	2	2.82	0.15
2606	فاقبلوا	2	2.82	0.15	2656	فأدخلت	2	2.82	0.15
2607	فيخاطون	2	2.82	0.15	2657	بوجهه	2	2.82	0.15
2608	فاقام	2	2.82	0.15	2658	بمحمد	2	2.82	0.15
2609	ايديهما	2	2.82	0.15	2659	بحقه	2	2.82	0.15
2610	بالا غاليط	2	2.82	0.15	2660	بم	2	2.82	0.15
2611	ايغذب	2	2.82	0.15	2661	فاعمل	2	2.82	0.15
2612	فيء	2	2.82	0.15	2662	بمثله	2	2.82	0.15
2613	قطيفة	2	2.82	0.15	2663	بمثلها	2	2.82	0.15
2614	بعود	2	2.82	0.15	2664	فاعطيتها	2	2.82	0.15
2615	فوقف	2	2.82	0.15	2665	فاعطوا	2	2.82	0.15
2616	بحمد	2	2.82	0.15	2666	فيسهل	2	2.82	0.15
2617	فاهوى	2	2.82	0.15	2667	فاعطاني	2	2.82	0.15
2618	فوقها	2	2.82	0.15	2668	بنسبي	2	2.82	0.15
2619	فاهلي	2	2.82	0.15	2669	اياها	2	2.82	0.15
2620	فاهلانا	2	2.82	0.15	2670	فايكم	2	2.82	0.15
2621	بعيرا	2	2.82	0.15	2671	فاطعم	2	2.82	0.15
2622	فامسك	2	2.82	0.15	2672	بمن	2	2.82	0.15
2623	فساررته	2	2.82	0.15	2673	فيشهدن	2	2.82	0.15
2624	بلغنا	2	2.82	0.15	2674	بمنزلة	2	2.82	0.15
2625	اينا	2	2.82	0.15	2675	بحق	2	2.82	0.15
2626	فانطلقوا	2	2.82	0.15	2676	فاصابني	2	2.82	0.15
2627	فيايى	2	2.82	0.15	2677	بناء	2	2.82	0.15
2628	بالايمان	2	2.82	0.15	2678	فاصاب	2	2.82	0.15
2629	فانصرف	2	2.82	0.15	2679	فاسلمت	2	2.82	0.15
2630	فيايتهم	2	2.82	0.15	2680	فاعط	2	2.82	0.15
2631	بقدر	2	2.82	0.15	2681	بسم	2	2.82	0.15
2632	بقرا	2	2.82	0.15	2682	فجمعوا	2	2.82	0.15
2633	بقرونها	2	2.82	0.15	2683	فممن	2	2.82	0.15
2634	فانكم	2	2.82	0.15	2684	فدخلنا	2	2.82	0.15
2635	فارجع	2	2.82	0.15	2685	فداك	2	2.82	0.15
2636	فافعلوا	2	2.82	0.15	2686	بركة	2	2.82	0.15
2637	فاسحقوها	2	2.82	0.15	2687	برمة	2	2.82	0.15
2638	فاستيقظ	2	2.82	0.15	2688	فحملته	2	2.82	0.15
2639	بحجر	2	2.82	0.15	2689	فحلفت	2	2.82	0.15
2640	فاستحييت	2	2.82	0.15	2690	فحلف	2	2.82	0.15
2641	بهز	2	2.82	0.15	2691	فمكثنا	2	2.82	0.15
2642	فيفتح	2	2.82	0.15	2692	فزرعه	2	2.82	0.15
2643	بجمار	2	2.82	0.15	2693	فمكث	2	2.82	0.15
2644	فليصمه	2	2.82	0.15	2694	فحضرت	2	2.82	0.15

Feature No	feature	DF	IDF	Feature redundancy	Feature No	feature	DF	IDF	Feature redundancy
2695	بسنة	2	2.82	0.15	2745	بئرا	2	2.82	0.15
2696	بسهم	2	2.82	0.15	2746	فيذلك	2	2.82	0.15
2697	فرق	2	2.82	0.15	2747	فزعم	2	2.82	0.15
2698	فليله	2	2.82	0.15	2748	بعضنا	2	2.82	0.15
2699	فحدثت	2	2.82	0.15	2749	بعرقة	2	2.82	0.15
2700	فحاضت	2	2.82	0.15	2750	فجاءته	2	2.82	0.15
2701	فحاشا	2	2.82	0.15	2751	فرقة	2	2.82	0.15
2702	فسئل	2	2.82	0.15	2752	بصائم	2	2.82	0.15
2703	فنجرت	2	2.82	0.15	2753	فجعلنا	2	2.82	0.15
2704	برة	2	2.82	0.15	2754	بصدقة	2	2.82	0.15
2705	فرسا	2	2.82	0.15	2755	فجر	2	2.82	0.15
2706	بدنته	2	2.82	0.15	2756	فنعم	2	2.82	0.15
2707	بدينه	2	2.82	0.15	2757	بصفية	2	2.82	0.15
2708	فرب	2	2.82	0.15	2758	فجاءني	2	2.82	0.15
2709	فمرض	2	2.82	0.15	2759	بانن	2	2.82	0.15
2710	فراش	2	2.82	0.15	2760	فتقدم	2	2.82	0.15
2711	فراس	2	2.82	0.15	2761	باخفافها	2	2.82	0.15
2712	فرارا	2	2.82	0.15	2762	فيكون	2	2.82	0.15
2713	براسه	2	2.82	0.15	2763	فرقد	2	2.82	0.15
2714	فملا	2	2.82	0.15	2764	فهتكه	2	2.82	0.15
2715	فذكره	2	2.82	0.15	2765	باختك	2	2.82	0.15
2716	فجمع	2	2.82	0.15	2766	بصوم	2	2.82	0.15
2717	فدفنت	2	2.82	0.15	2767	بابين	2	2.82	0.15
2718	باكل	2	2.82	0.15	2768	بطل	2	2.82	0.15
2719	برد	2	2.82	0.15	2769	فرمي	2	2.82	0.15
2720	بخيف	2	2.82	0.15	2770	بخبر	2	2.82	0.15
2721	بردائه	2	2.82	0.15	2771	بطني	2	2.82	0.15
2722	فمشى	2	2.82	0.15	2772	ففنزل	2	2.82	0.15
2723	باكير	2	2.82	0.15	2773	تركنا	2	2.82	0.15
2724	فدعوت	2	2.82	0.15	2774	تروه	2	2.82	0.15
2725	فدعاهم	2	2.82	0.15	2775	قرام	2	2.82	0.15
2726	بد	2	2.82	0.15	2776	اهلت	2	2.82	0.15
2727	فتاه	2	2.82	0.15	2777	ترجمان	2	2.82	0.15
2728	فجىء	2	2.82	0.15	2778	قرية	2	2.82	0.15
2729	بطواقه	2	2.82	0.15	2779	ترده	2	2.82	0.15
2730	فوجده	2	2.82	0.15	2780	ترضين	2	2.82	0.15
2731	بظاهر	2	2.82	0.15	2781	فقولوا	2	2.82	0.15
2732	بعيد	2	2.82	0.15	2782	بيعضه	2	2.82	0.15
2733	فزار	2	2.82	0.15	2783	قراءته	2	2.82	0.15
2734	بعثك	2	2.82	0.15	2784	عجل	2	2.82	0.15
2735	فرع	2	2.82	0.15	2785	بالجاهلية	2	2.82	0.15
2736	فتبسم	2	2.82	0.15	2786	عنيق	2	2.82	0.15
2737	فوجدنا	2	2.82	0.15	2787	عبده	2	2.82	0.15
2738	باء	2	2.82	0.15	2788	اهدت	2	2.82	0.15
2739	فتقول	2	2.82	0.15	2789	عبدنا	2	2.82	0.15
2740	فبم	2	2.82	0.15	2790	عبدا	2	2.82	0.15
2741	فبكيت	2	2.82	0.15	2791	عبادي	2	2.82	0.15
2742	فبكت	2	2.82	0.15	2792	عاهد	2	2.82	0.15
2743	فبسطته	2	2.82	0.15	2793	قسما	2	2.82	0.15
2744	فبرك	2	2.82	0.15	2794	فسمعت	2	2.82	0.15

Feature No	feature	DF	IDF	Feature redundancy	Feature No	feature	DF	IDF	Feature redundancy
2795	فادبر	2	2.82	0.15	2844	عارية	2	2.82	0.15
2796	عشيتهم	2	2.82	0.15	2845	قص	2	2.82	0.15
2797	عقال	2	2.82	0.15	2846	تزال	2	2.82	0.15
2798	عظيمة	2	2.82	0.15	2847	عاب	2	2.82	0.15
2799	قدمه	2	2.82	0.15	2848	عائذا	2	2.82	0.15
2800	تدحض	2	2.82	0.15	2849	قصر	2	2.82	0.15
2801	تدردر	2	2.82	0.15	2850	ظهراني	2	2.82	0.15
2802	تدري	2	2.82	0.15	2851	ظن	2	2.82	0.15
2803	تدعو	2	2.82	0.15	2852	تسحروا	2	2.82	0.15
2804	تدعونها	2	2.82	0.15	2853	ظلم	2	2.82	0.15
2805	عصب	2	2.82	0.15	2854	علقت	2	2.82	0.15
2806	عرض	2	2.82	0.15	2855	تسالها	2	2.82	0.15
2807	بالتمر	2	2.82	0.15	2856	تسبه	2	2.82	0.15
2808	عاقبه	2	2.82	0.15	2857	فليبغ	2	2.82	0.15
2809	بت	2	2.82	0.15	2858	بالحشية	2	2.82	0.15
2810	فدذه	2	2.82	0.15	2859	تستره	2	2.82	0.15
2811	عشرون	2	2.82	0.15	2860	طول	2	2.82	0.15
2812	عسفان	2	2.82	0.15	2861	طوقه	2	2.82	0.15
2813	تذهب	2	2.82	0.15	2862	طوفي	2	2.82	0.15
2814	عريانا	2	2.82	0.15	2863	تستطيع	2	2.82	0.15
2815	اهلنا	2	2.82	0.15	2864	تسال	2	2.82	0.15
2816	بالجار	2	2.82	0.15	2865	ناتي	2	2.82	0.15
2817	عروسا	2	2.82	0.15	2866	عقلته	2	2.82	0.15
2818	فصاح	2	2.82	0.15	2867	فاتموا	2	2.82	0.15
2819	تشربوا	2	2.82	0.15	2868	فاتكم	2	2.82	0.15
2820	تروني	2	2.82	0.15	2869	تؤخذ	2	2.82	0.15
2821	فضريت	2	2.82	0.15	2870	فاتخذ	2	2.82	0.15
2822	انبابه	2	2.82	0.15	2871	قاتله	2	2.82	0.15
2823	تسل	2	2.82	0.15	2872	فسوق	2	2.82	0.15
2824	فضلي	2	2.82	0.15	2873	تؤدي	2	2.82	0.15
2825	تسمعه	2	2.82	0.15	2874	فابردوها	2	2.82	0.15
2826	تسموا	2	2.82	0.15	2875	فاتينا	2	2.82	0.15
2827	ضيفك	2	2.82	0.15	2876	فابدؤوا	2	2.82	0.15
2828	قضينا	2	2.82	0.15	2877	بثلاثة	2	2.82	0.15
2829	طلعة	2	2.82	0.15	2878	غيم	2	2.82	0.15
2830	ضمرة	2	2.82	0.15	2879	غيرهم	2	2.82	0.15
2831	فضة	2	2.82	0.15	2880	تاخر	2	2.82	0.15
2832	ضعيفا	2	2.82	0.15	2881	غني	2	2.82	0.15
2833	ضعيت	2	2.82	0.15	2882	غنمك	2	2.82	0.15
2834	ضعفة	2	2.82	0.15	2883	تاكلون	2	2.82	0.15
2835	فرغت	2	2.82	0.15	2884	غنم	2	2.82	0.15
2836	تشرکوا	2	2.82	0.15	2885	تامر	2	2.82	0.15
2837	اهدي	2	2.82	0.15	2886	غلبه	2	2.82	0.15
2838	ضربا	2	2.82	0.15	2887	قاسم	2	2.82	0.15
2839	ضحك	2	2.82	0.15	2888	فاحسن	2	2.82	0.15
2840	قعد	2	2.82	0.15	2889	قعدت	2	2.82	0.15
2841	ضمه	2	2.82	0.15	2890	فاخذتها	2	2.82	0.15
2842	فصليا	2	2.82	0.15	2891	فاختلفوا	2	2.82	0.15
2843	عافاني	2	2.82	0.15	2892	فاختبات	2	2.82	0.15

Feature No	feature	DF	IDF	Feature redundancy	Feature No	feature	DF	IDF	Feature redundancy
2893	فاخبروه	2	2.82	0.15	2943	تجعل	2	2.82	0.15
2894	فاخبرهم	2	2.82	0.15	2944	قبورهم	2	2.82	0.15
2895	بيعتين	2	2.82	0.15	2945	تجعلوا	2	2.82	0.15
2896	فاخبرناه	2	2.82	0.15	2946	قبلي	2	2.82	0.15
2897	فاخبرت	2	2.82	0.15	2947	عين	2	2.82	0.15
2898	فاتوا	2	2.82	0.15	2948	قبلك	2	2.82	0.15
2899	بيمينه	2	2.82	0.15	2949	قبلكه	2	2.82	0.15
2900	غزونا	2	2.82	0.15	2950	قبلك	2	2.82	0.15
2901	فاجعل	2	2.82	0.15	2951	قبلكم	2	2.82	0.15
2902	فينتقم	2	2.82	0.15	2952	قبله	2	2.82	0.15
2903	فيظنر	2	2.82	0.15	2953	فليرده	2	2.82	0.15
2904	فيضع	2	2.82	0.15	2954	عيدا	2	2.82	0.15
2905	فاجدني	2	2.82	0.15	2955	عياش	2	2.82	0.15
2906	فيهن	2	2.82	0.15	2956	تبسم	2	2.82	0.15
2907	اوفي	2	2.82	0.15	2957	عيني	2	2.82	0.15
2908	اوعى	2	2.82	0.15	2958	ارجلنا	2	2.82	0.15
2909	بيوتهن	2	2.82	0.15	2959	اذا	2	2.82	0.15
2910	فاخاف	2	2.82	0.15	2960	الزمان	2	2.82	0.15
2911	عم	2	2.82	0.15	2961	وجدناه	2	2.82	0.15
2912	فشكوت	2	2.82	0.15	2962	الزنا	2	2.82	0.15
2913	عوالي	2	2.82	0.15	2963	وابوه	2	2.82	0.15
2914	عزرة	2	2.82	0.15	2964	الزيانب	2	2.82	0.15
2915	تحرروا	2	2.82	0.15	2965	ارسلتني	2	2.82	0.15
2916	عناقا	2	2.82	0.15	2966	وابدا	2	2.82	0.15
2917	عملك	2	2.82	0.15	2967	يجمعون	2	2.82	0.15
2918	اهليكم	2	2.82	0.15	2968	السامة	2	2.82	0.15
2919	قحط	2	2.82	0.15	2969	هنتاه	2	2.82	0.15
2920	فقحطان	2	2.82	0.15	2970	همام	2	2.82	0.15
2921	فشربنا	2	2.82	0.15	2971	هلمي	2	2.82	0.15
2922	عمارة	2	2.82	0.15	2972	يحقر	2	2.82	0.15
2923	اوتوا	2	2.82	0.15	2973	هلكته	2	2.82	0.15
2924	عليكن	2	2.82	0.15	2974	السؤال	2	2.82	0.15
2925	تحلب	2	2.82	0.15	2975	واجود	2	2.82	0.15
2926	عليا	2	2.82	0.15	2976	الرفث	2	2.82	0.15
2927	علمه	2	2.82	0.15	2977	الرقية	2	2.82	0.15
2928	فشمت	2	2.82	0.15	2978	يجاوز	2	2.82	0.15
2929	قدمتم	2	2.82	0.15	2979	الرمل	2	2.82	0.15
2930	بالتمام	2	2.82	0.15	2980	الرواح	2	2.82	0.15
2931	تحول	2	2.82	0.15	2981	يجدوا	2	2.82	0.15
2932	بيدها	2	2.82	0.15	2982	اردت	2	2.82	0.15
2933	عمرتي	2	2.82	0.15	2983	ارضعتني	2	2.82	0.15
2934	عيناى	2	2.82	0.15	2984	الستين	2	2.82	0.15
2935	اهلين	2	2.82	0.15	2985	يجعلوها	2	2.82	0.15
2936	تبعتنا	2	2.82	0.15	2986	الريان	2	2.82	0.15
2937	تبغنا	2	2.82	0.15	2987	الريح	2	2.82	0.15
2938	غرة	2	2.82	0.15	2988	يجلس	2	2.82	0.15
2939	فشربوا	2	2.82	0.15	2989	واجابة	2	2.82	0.15
2940	غدوت	2	2.82	0.15	2990	ارض	2	2.82	0.15
2941	عدر	2	2.82	0.15	2991	ارضنا	2	2.82	0.15
2942	غالب	2	2.82	0.15	2992	اذناى	2	2.82	0.15

Feature No	feature	DF	IDF	Feature redundancy	Feature No	feature	DF	IDF	Feature redundancy
2993	نيرانا	2	2.82	0.15	3043	الخروج	2	2.82	0.15
2994	يخشى	2	2.82	0.15	3044	والثياب	2	2.82	0.15
2995	يخطفها	2	2.82	0.15	3045	ياخذه	2	2.82	0.15
2996	نوبته	2	2.82	0.15	3046	ياخذون	2	2.82	0.15
2997	نواصبها	2	2.82	0.15	3047	والحر	2	2.82	0.15
2998	نواحي	2	2.82	0.15	3048	استقبلت	2	2.82	0.15
2999	هلا	2	2.82	0.15	3049	استقبل	2	2.82	0.15
3000	نهر يقها	2	2.82	0.15	3050	استغفر	2	2.82	0.15
3001	هاتين	2	2.82	0.15	3051	بيبعته	2	2.82	0.15
3002	نهر	2	2.82	0.15	3052	الحمار	2	2.82	0.15
3003	يدعها	2	2.82	0.15	3053	والاقامة	2	2.82	0.15
3004	اذانين	2	2.82	0.15	3054	الحر	2	2.82	0.15
3005	نهارا	2	2.82	0.15	3055	الحميد	2	2.82	0.15
3006	النشاء	2	2.82	0.15	3056	يبندرون	2	2.82	0.15
3007	الغلطان	2	2.82	0.15	3057	استطعم	2	2.82	0.15
3008	نهل	2	2.82	0.15	3058	والاول	2	2.82	0.15
3009	هجر	2	2.82	0.15	3059	الحيض	2	2.82	0.15
3010	يتوكلون	2	2.82	0.15	3060	يترجم	2	2.82	0.15
3011	يحلل	2	2.82	0.15	3061	الحلواء	2	2.82	0.15
3012	السكر	2	2.82	0.15	3062	وارجع	2	2.82	0.15
3013	يحمه	2	2.82	0.15	3063	ازارك	2	2.82	0.15
3014	السرراويل	2	2.82	0.15	3064	ازار	2	2.82	0.15
3015	يخاف	2	2.82	0.15	3065	واسقى	2	2.82	0.15
3016	يخرجون	2	2.82	0.15	3066	واسعا	2	2.82	0.15
3017	يخب	2	2.82	0.15	3067	الذاهب	2	2.82	0.15
3018	يخرجن	2	2.82	0.15	3068	يتكلوا	2	2.82	0.15
3019	هجا	2	2.82	0.15	3069	يتحولنا	2	2.82	0.15
3020	هبيرة	2	2.82	0.15	3070	يتمثل	2	2.82	0.15
3021	السفينة	2	2.82	0.15	3071	يتقدم	2	2.82	0.15
3022	يختلف	2	2.82	0.15	3072	الرحمة	2	2.82	0.15
3023	يخدمني	2	2.82	0.15	3073	واربع	2	2.82	0.15
3024	ارادت	2	2.82	0.15	3074	واراد	2	2.82	0.15
3025	ارتفعت	2	2.82	0.15	3075	يتوخي	2	2.82	0.15
3026	هذان	2	2.82	0.15	3076	وادخل	2	2.82	0.15
3027	يتبعه	2	2.82	0.15	3077	نقية	2	2.82	0.15
3028	استزده	2	2.82	0.15	3078	يتلو	2	2.82	0.15
3029	بيكي	2	2.82	0.15	3079	يتصدقن	2	2.82	0.15
3030	والارض	2	2.82	0.15	3080	الخفين	2	2.82	0.15
3031	والا	2	2.82	0.15	3081	واعطوا	2	2.82	0.15
3032	الخدور	2	2.82	0.15	3082	الخمرة	2	2.82	0.15
3033	واكلوا	2	2.82	0.15	3083	الخويصرة	2	2.82	0.15
3034	اسب	2	2.82	0.15	3084	واضعا	2	2.82	0.15
3035	الخزرج	2	2.82	0.15	3085	واضربوا	2	2.82	0.15
3036	بيدو	2	2.82	0.15	3086	ازاره	2	2.82	0.15
3037	استاذنت	2	2.82	0.15	3087	واصل	2	2.82	0.15
3038	واقطا	2	2.82	0.15	3088	واسماعيل	2	2.82	0.15
3039	الخصم	2	2.82	0.15	3089	الدابة	2	2.82	0.15
3040	واقيل	2	2.82	0.15	3090	ازيد	2	2.82	0.15
3041	واقامة	2	2.82	0.15	3091	يتطبرون	2	2.82	0.15
3042	واختها	2	2.82	0.15	3092	الدخ	2	2.82	0.15

Feature No	feature	DF	IDF	Feature redundancy	Feature No	feature	DF	IDF	Feature redundancy
3093	يتعوزوا	2	2.82	0.15	3143	منافق	2	2.82	0.15
3094	الدم	2	2.82	0.15	3144	يطيل	2	2.82	0.15
3095	يثرب	2	2.82	0.15	3145	الغضب	2	2.82	0.15
3096	واضاعة	2	2.82	0.15	3146	مناديا	2	2.82	0.15
3097	بشرب	2	2.82	0.15	3147	يطعم	2	2.82	0.15
3098	نمسخ	2	2.82	0.15	3148	منهبط	2	2.82	0.15
3099	نائما	2	2.82	0.15	3149	نتناب	2	2.82	0.15
3100	العتاء	2	2.82	0.15	3150	احمرت	2	2.82	0.15
3101	العطاس	2	2.82	0.15	3151	يشمت	2	2.82	0.15
3102	موظة	2	2.82	0.15	3152	موليه	2	2.82	0.15
3103	مومك	2	2.82	0.15	3153	مولى	2	2.82	0.15
3104	ناتيك	2	2.82	0.15	3154	يشهوا	2	2.82	0.15
3105	يشد	2	2.82	0.15	3155	منعوني	2	2.82	0.15
3106	يشبهه	2	2.82	0.15	3156	منهبطة	2	2.82	0.15
3107	العفو	2	2.82	0.15	3157	يضر	2	2.82	0.15
3108	اخا	2	2.82	0.15	3158	يصدوك	2	2.82	0.15
3109	موضع	2	2.82	0.15	3159	يصرف	2	2.82	0.15
3110	اخ	2	2.82	0.15	3160	العنزة	2	2.82	0.15
3111	احياء	2	2.82	0.15	3161	منكيه	2	2.82	0.15
3112	يشفى	2	2.82	0.15	3162	يصلى	2	2.82	0.15
3113	العطش	2	2.82	0.15	3163	يصلها	2	2.82	0.15
3114	العزراء	2	2.82	0.15	3164	احوج	2	2.82	0.15
3115	يسروا	2	2.82	0.15	3165	مهل	2	2.82	0.15
3116	العب	2	2.82	0.15	3166	نصبا	2	2.82	0.15
3117	العباد	2	2.82	0.15	3167	ادركتم	2	2.82	0.15
3118	العجين	2	2.82	0.15	3168	ادركت	2	2.82	0.15
3119	يسعى	2	2.82	0.15	3169	يرحمك	2	2.82	0.15
3120	نال	2	2.82	0.15	3170	نضيه	2	2.82	0.15
3121	ناتي	2	2.82	0.15	3171	الصالحة	2	2.82	0.15
3122	ناقوسا	2	2.82	0.15	3172	نصله	2	2.82	0.15
3123	العلاق	2	2.82	0.15	3173	نستطيع	2	2.82	0.15
3124	ناقة	2	2.82	0.15	3174	الصبي	2	2.82	0.15
3125	يسلموا	2	2.82	0.15	3175	نطعمه	2	2.82	0.15
3126	بسمعها	2	2.82	0.15	3176	ادبر	2	2.82	0.15
3127	العراق	2	2.82	0.15	3177	نشرب	2	2.82	0.15
3128	اخبرته	2	2.82	0.15	3178	نشا	2	2.82	0.15
3129	يسيرا	2	2.82	0.15	3179	الصدق	2	2.82	0.15
3130	اختها	2	2.82	0.15	3180	يركب	2	2.82	0.15
3131	منزلنا	2	2.82	0.15	3181	اخذه	2	2.82	0.15
3132	منعها	2	2.82	0.15	3182	نصل	2	2.82	0.15
3133	العوام	2	2.82	0.15	3183	نعلي	2	2.82	0.15
3134	منع	2	2.82	0.15	3184	نقص	2	2.82	0.15
3135	يضرهم	2	2.82	0.15	3185	الشجرة	2	2.82	0.15
3136	منصور	2	2.82	0.15	3186	يدني	2	2.82	0.15
3137	احرام	2	2.82	0.15	3187	الشعرة	2	2.82	0.15
3138	مودع	2	2.82	0.15	3188	يديها	2	2.82	0.15
3139	احدانا	2	2.82	0.15	3189	النشغل	2	2.82	0.15
3140	احسنكم	2	2.82	0.15	3190	نطع	2	2.82	0.15
3141	يطول	2	2.82	0.15	3191	الشمال	2	2.82	0.15
3142	منحر	2	2.82	0.15	3192	ادركته	2	2.82	0.15

Feature No	feature	DF	IDF	Feature redundancy	Feature No	feature	DF	IDF	Feature redundancy
3193	يراني	2	2.82	0.15	3243	ورد	2	2.82	0.15
3194	نعت	2	2.82	0.15	3244	وراءكم	2	2.82	0.15
3195	نظرت	2	2.82	0.15	3245	ولكنه	2	2.82	0.15
3196	نظر	2	2.82	0.15	3246	وذو	2	2.82	0.15
3197	نطق	2	2.82	0.15	3247	الادوية	2	2.82	0.15
3198	نطعمهم	2	2.82	0.15	3248	اطعموا	2	2.82	0.15
3199	اخوه	2	2.82	0.15	3249	وذا	2	2.82	0.15
3200	يذهب	2	2.82	0.15	3250	وليقيض	2	2.82	0.15
3201	الطالب	2	2.82	0.15	3251	الارملة	2	2.82	0.15
3202	نخلص	2	2.82	0.15	3252	ورائي	2	2.82	0.15
3203	اخرين	2	2.82	0.15	3253	اعانك	2	2.82	0.15
3204	نخبر	2	2.82	0.15	3254	اعتقها	2	2.82	0.15
3205	يستعفف	2	2.82	0.15	3255	ولحله	2	2.82	0.15
3206	يستغن	2	2.82	0.15	3256	ولحيته	2	2.82	0.15
3207	اخرهم	2	2.82	0.15	3257	ولدت	2	2.82	0.15
3208	اخي	2	2.82	0.15	3258	وركع	2	2.82	0.15
3209	يستلم	2	2.82	0.15	3259	اكننت	2	2.82	0.15
3210	يسترقون	2	2.82	0.15	3260	ولنفسك	2	2.82	0.15
3211	نحب	2	2.82	0.15	3261	اعافه	2	2.82	0.15
3212	اخرجوا	2	2.82	0.15	3262	الاسد	2	2.82	0.15
3213	الطباء	2	2.82	0.15	3263	اعادها	2	2.82	0.15
3214	يستيقظ	2	2.82	0.15	3264	ولسنا	2	2.82	0.15
3215	اخرجت	2	2.82	0.15	3265	اظل	2	2.82	0.15
3216	ياتيهم	2	2.82	0.15	3266	ورسله	2	2.82	0.15
3217	نحو	2	2.82	0.15	3267	ورس	2	2.82	0.15
3218	اخطا	2	2.82	0.15	3268	ولكل	2	2.82	0.15
3219	اخوان	2	2.82	0.15	3269	اكنتم	2	2.82	0.15
3220	نساله	2	2.82	0.15	3270	ومعاذ	2	2.82	0.15
3221	يريهم	2	2.82	0.15	3271	الاكبر	2	2.82	0.15
3222	نسال	2	2.82	0.15	3272	ومسح	2	2.82	0.15
3223	الصلت	2	2.82	0.15	3273	وحاضت	2	2.82	0.15
3224	نزلتم	2	2.82	0.15	3274	وجنتاه	2	2.82	0.15
3225	يستطع	2	2.82	0.15	3275	ومشاة	2	2.82	0.15
3226	اخلاقا	2	2.82	0.15	3276	وجلودها	2	2.82	0.15
3227	اخسا	2	2.82	0.15	3277	وخيرها	2	2.82	0.15
3228	يستازن	2	2.82	0.15	3278	اصلها	2	2.82	0.15
3229	يستبين	2	2.82	0.15	3279	الاقدام	2	2.82	0.15
3230	يستحب	2	2.82	0.15	3280	اليمنى	2	2.82	0.15
3231	الضب	2	2.82	0.15	3281	ومعنا	2	2.82	0.15
3232	يستحي	2	2.82	0.15	3282	وجدت	2	2.82	0.15
3233	ندخل	2	2.82	0.15	3283	اصغر	2	2.82	0.15
3234	اخذاها	2	2.82	0.15	3284	وجار	2	2.82	0.15
3235	يزني	2	2.82	0.15	3285	وثمانيا	2	2.82	0.15
3236	وذوات	2	2.82	0.15	3286	وجف	2	2.82	0.15
3237	الحصبة	2	2.82	0.15	3287	الاعصم	2	2.82	0.15
3238	الاجرب	2	2.82	0.15	3288	وساق	2	2.82	0.15
3239	ولها	2	2.82	0.15	3289	وماجوج	2	2.82	0.15
3240	اطلقت	2	2.82	0.15	3290	وخسروا	2	2.82	0.15
3241	ولي	2	2.82	0.15	3291	وخسرت	2	2.82	0.15
3242	وليؤمكم	2	2.82	0.15	3292	الاسودين	2	2.82	0.15

Feature No	feature	DF	IDF	Feature redundancy	Feature No	feature	DF	IDF	Feature redundancy
3293	وخالد	2	2.82	0.15	3343	اقيموا	2	2.82	0.15
3294	الاقرع	2	2.82	0.15	3344	وسعيد	2	2.82	0.15
3295	وحلق	2	2.82	0.15	3345	اعرض	2	2.82	0.15
3296	الاقربين	2	2.82	0.15	3346	وسع	2	2.82	0.15
3297	ومروهم	2	2.82	0.15	3347	وصوم	2	2.82	0.15
3298	الاعلى	2	2.82	0.15	3348	وستون	2	2.82	0.15
3299	الاعين	2	2.82	0.15	3349	اقطروها	2	2.82	0.15
3300	اصول	2	2.82	0.15	3350	اعدى	2	2.82	0.15
3301	الافطار	2	2.82	0.15	3351	وسبعة	2	2.82	0.15
3302	الافق	2	2.82	0.15	3352	وسالته	2	2.82	0.15
3303	وخير	2	2.82	0.15	3353	اعدل	2	2.82	0.15
3304	ومرقا	2	2.82	0.15	3354	اكره	2	2.82	0.15
3305	اقاتل	2	2.82	0.15	3355	وتوضانا	2	2.82	0.15
3306	اعلمه	2	2.82	0.15	3356	وسط	2	2.82	0.15
3307	اعلقت	2	2.82	0.15	3357	اقراصا	2	2.82	0.15
3308	افعلوا	2	2.82	0.15	3358	وصلاة	2	2.82	0.15
3309	وقتاله	2	2.82	0.15	3359	وقومه	2	2.82	0.15
3310	وقتلوا	2	2.82	0.15	3360	وصارت	2	2.82	0.15
3311	افي	2	2.82	0.15	3361	وصاحبى	2	2.82	0.15
3312	وصلك	2	2.82	0.15	3362	وقيل	2	2.82	0.15
3313	وضوء	2	2.82	0.15	3363	وشيء	2	2.82	0.15
3314	وعباس	2	2.82	0.15	3364	وكدا	2	2.82	0.15
3315	وقرة	2	2.82	0.15	3365	وشهادة	2	2.82	0.15
3316	اقال	2	2.82	0.15	3366	وكتب	2	2.82	0.15
3317	وضحى	2	2.82	0.15	3367	وشرب	2	2.82	0.15
3318	وصيامه	2	2.82	0.15	3368	اقرنين	2	2.82	0.15
3319	وصيام	2	2.82	0.15	3369	وشبعوا	2	2.82	0.15
3320	اكله	2	2.82	0.15	3370	اقسمت	2	2.82	0.15
3321	افيض	2	2.82	0.15	3371	وكتاب	2	2.82	0.15
3322	وعقوق	2	2.82	0.15	3372	اقضى	2	2.82	0.15
3323	وعندها	2	2.82	0.15	3373	ولاهلك	2	2.82	0.15
3324	اغنياهم	2	2.82	0.15	3374	وكاءها	2	2.82	0.15
3325	وغطفان	2	2.82	0.15	3375	والمسكين	2	2.82	0.15
3326	اغرم	2	2.82	0.15	3376	ويعلمها	2	2.82	0.15
3327	اغير	2	2.82	0.15	3377	والملاسة	2	2.82	0.15
3328	اغتسل	2	2.82	0.15	3378	والمقير	2	2.82	0.15
3329	وفيه	2	2.82	0.15	3379	اسلامه	2	2.82	0.15
3330	افاق	2	2.82	0.15	3380	الجامعة	2	2.82	0.15
3331	وفينا	2	2.82	0.15	3381	ويفيض	2	2.82	0.15
3332	وفاطمة	2	2.82	0.15	3382	ويكثر	2	2.82	0.15
3333	وعفاصها	2	2.82	0.15	3383	الجبال	2	2.82	0.15
3334	اعملوا	2	2.82	0.15	3384	الثلاثين	2	2.82	0.15
3335	اقتل	2	2.82	0.15	3385	الجدار	2	2.82	0.15
3336	وعدلت	2	2.82	0.15	3386	الجدع	2	2.82	0.15
3337	وعجل	2	2.82	0.15	3387	ويقل	2	2.82	0.15
3338	اعطني	2	2.82	0.15	3388	والماء	2	2.82	0.15
3339	وعلموهم	2	2.82	0.15	3389	والليلة	2	2.82	0.15
3340	وكلوا	2	2.82	0.15	3390	ويدخل	2	2.82	0.15
3341	وكذب	2	2.82	0.15	3391	الجانب	2	2.82	0.15
3342	وسمعت	2	2.82	0.15	3392	والنعمة	2	2.82	0.15

Feature No	feature	DF	IDF	Feature redundancy	Feature No	feature	DF	IDF	Feature redundancy
3393	اصطحب	2	2.82	0.15	3443	اشبه	2	2.82	0.15
3394	وامرهم	2	2.82	0.15	3444	والعسل	2	2.82	0.15
3395	ويده	2	2.82	0.15	3445	ونفسه	2	2.82	0.15
3396	ويذكر	2	2.82	0.15	3446	وترد	2	2.82	0.15
3397	التميمي	2	2.82	0.15	3447	البدر	2	2.82	0.15
3398	التوراة	2	2.82	0.15	3448	وتر	2	2.82	0.15
3399	ويعجبي	2	2.82	0.15	3449	وتجدون	2	2.82	0.15
3400	والوقار	2	2.82	0.15	3450	البرنس	2	2.82	0.15
3401	ويعتزل	2	2.82	0.15	3451	وتابعه	2	2.82	0.15
3402	ويشرب	2	2.82	0.15	3452	وهب	2	2.82	0.15
3403	الثبت	2	2.82	0.15	3453	وتؤدي	2	2.82	0.15
3404	والنصاري	2	2.82	0.15	3454	وتزودوا	2	2.82	0.15
3405	الثرى	2	2.82	0.15	3455	وتؤتي	2	2.82	0.15
3406	والنحر	2	2.82	0.15	3456	وتقوم	2	2.82	0.15
3407	اسم	2	2.82	0.15	3457	اشك	2	2.82	0.15
3408	ويكره	2	2.82	0.15	3458	ونهي	2	2.82	0.15
3409	واليتيم	2	2.82	0.15	3459	البصر	2	2.82	0.15
3410	يؤذ	2	2.82	0.15	3460	التلبينة	2	2.82	0.15
3411	يؤخر	2	2.82	0.15	3461	ونغسلها	2	2.82	0.15
3412	الحبيبة	2	2.82	0.15	3462	الايسر	2	2.82	0.15
3413	الحجام	2	2.82	0.15	3463	ياجوج	2	2.82	0.15
3414	والسكينة	2	2.82	0.15	3464	الانوف	2	2.82	0.15
3415	والسكين	2	2.82	0.15	3465	وتنطحه	2	2.82	0.15
3416	الحجامة	2	2.82	0.15	3466	اصحابا	2	2.82	0.15
3417	الجسد	2	2.82	0.15	3467	ومنهم	2	2.82	0.15
3418	الحجفة	2	2.82	0.15	3468	ومنى	2	2.82	0.15
3419	الحبشة	2	2.82	0.15	3469	البحار	2	2.82	0.15
3420	استهل	2	2.82	0.15	3470	الايات	2	2.82	0.15
3421	والزبير	2	2.82	0.15	3471	ونزل	2	2.82	0.15
3422	الحره	2	2.82	0.15	3472	البائن	2	2.82	0.15
3423	والدم	2	2.82	0.15	3473	البيادية	2	2.82	0.15
3424	الحزن	2	2.82	0.15	3474	اصبحنا	2	2.82	0.15
3425	يحدو	2	2.82	0.15	3475	وتصل	2	2.82	0.15
3426	يؤدونها	2	2.82	0.15	3476	وتسعين	2	2.82	0.15
3427	اسرد	2	2.82	0.15	3477	اصابني	2	2.82	0.15
3428	اسفاره	2	2.82	0.15	3478	البضعة	2	2.82	0.15
3429	ويكفرن	2	2.82	0.15	3479	وتقوم	2	2.82	0.15
3430	ويلذ	2	2.82	0.15	3480	اشتقنا	2	2.82	0.15
3431	الجمل	2	2.82	0.15	3481	وايكم	2	2.82	0.15
3432	الجنزة	2	2.82	0.15	3482	اشتمال	2	2.82	0.15
3433	الجنى	2	2.82	0.15	3483	اشتكى	2	2.82	0.15
3434	والصبيان	2	2.82	0.15	3484	اشتكى	2	2.82	0.15
3435	اسرع	2	2.82	0.15	3485	واهدى	2	2.82	0.15
3436	استوى	2	2.82	0.15	3486	واهدى	2	2.82	0.15
3437	الحائض	2	2.82	0.15	3487	وبيص	2	2.82	0.15
3438	الحائط	2	2.82	0.15	3488	وانهى	2	2.82	0.15
3439	والضيافة	2	2.82	0.15	3489	واين	2	2.82	0.15
3440	الحاج	2	2.82	0.15	3490	وانك	2	2.82	0.15
3441	والصيام	2	2.82	0.15	3491	وانطلقت	2	2.82	0.15
3442	وينظر	2	2.82	0.15	3492	ويتوضا	2	2.82	0.15

Feature No	feature	DF	IDF	Feature redundancy	Feature No	feature	DF	IDF	Feature redundancy
3493	التثاؤب	2	2.82	0.15	3543	النبوّة	2	2.82	0.15
3494	ويجعل	2	2.82	0.15	3544	لزادني	2	2.82	0.15
3495	اشتراه	2	2.82	0.15	3545	مادوم	2	2.82	0.15
3496	وانى	2	2.82	0.15	3546	اتسحر	2	2.82	0.15
3497	وبالمدينة	2	2.82	0.15	3547	بينهاكم	2	2.82	0.15
3498	وبمحمد	2	2.82	0.15	3548	للاعقاب	2	2.82	0.15
3499	اشعرت	2	2.82	0.15	3549	ينبخ	2	2.82	0.15
3500	وهذه	2	2.82	0.15	3550	القلب	2	2.82	0.15
3501	وبلال	2	2.82	0.15	3551	معتمرا	2	2.82	0.15
3502	وبقي	2	2.82	0.15	3552	ابنه	2	2.82	0.15
3503	وبسطت	2	2.82	0.15	3553	القدح	2	2.82	0.15
3504	اشتهاه	2	2.82	0.15	3554	اجرته	2	2.82	0.15
3505	اشراط	2	2.82	0.15	3555	ينافح	2	2.82	0.15
3506	البوادي	2	2.82	0.15	3556	الوفد	2	2.82	0.15
3507	ووقع	2	2.82	0.15	3557	يكنزون	2	2.82	0.15
3508	وولدا	2	2.82	0.15	3558	مظلمة	2	2.82	0.15
3509	وبالاسلام	2	2.82	0.15	3559	يهودي	2	2.82	0.15
3510	وبات	2	2.82	0.15	3560	اجران	2	2.82	0.15
3511	البناني	2	2.82	0.15	3561	لحابستنا	2	2.82	0.15
3512	البهائم	2	2.82	0.15	3562	المنزل	2	2.82	0.15
3513	وثابت	2	2.82	0.15	3563	للقاسم	2	2.82	0.15
3514	وبدا	2	2.82	0.15	3564	ليسوا	2	2.82	0.15
3515	الموقف	2	2.82	0.15	3565	ليست	2	2.82	0.15
3516	الورق	2	2.82	0.15	3566	المسيب	2	2.82	0.15
3517	لربك	2	2.82	0.15	3567	لجسدك	2	2.82	0.15
3518	الوجوه	2	2.82	0.15	3568	القتلى	2	2.82	0.15
3519	لدينه	2	2.82	0.15	3569	يهالك	2	2.82	0.15
3520	الميت	2	2.82	0.15	3570	يكرهه	2	2.82	0.15
3521	ينتظرها	2	2.82	0.15	3571	للصائم	2	2.82	0.15
3522	الوير	2	2.82	0.15	3572	للضيف	2	2.82	0.15
3523	مصعدة	2	2.82	0.15	3573	اجتروا	2	2.82	0.15
3524	ماندة	2	2.82	0.15	3574	يعوذ	2	2.82	0.15
3525	ابنتيها	2	2.82	0.15	3575	الوسطى	2	2.82	0.15
3526	لدغ	2	2.82	0.15	3576	ليف	2	2.82	0.15
3527	يكتونون	2	2.82	0.15	3577	ليلا	2	2.82	0.15
3528	ابان	2	2.82	0.15	3578	المجلس	2	2.82	0.15
3529	ابنك	2	2.82	0.15	3579	لخلوف	2	2.82	0.15
3530	اللين	2	2.82	0.15	3580	الفا	2	2.82	0.15
3531	اجرنا	2	2.82	0.15	3581	لحي	2	2.82	0.15
3532	للذئب	2	2.82	0.15	3582	مطعم	2	2.82	0.15
3533	لكنه	2	2.82	0.15	3583	اتخذ	2	2.82	0.15
3534	لكما	2	2.82	0.15	3584	المسالة	2	2.82	0.15
3535	القميص	2	2.82	0.15	3585	لحق	2	2.82	0.15
3536	اتصدق	2	2.82	0.15	3586	ليلتكم	2	2.82	0.15
3537	القلوب	2	2.82	0.15	3587	الكريم	2	2.82	0.15
3538	مبلغ	2	2.82	0.15	3588	متبذلة	2	2.82	0.15
3539	لسانه	2	2.82	0.15	3589	اثنين	2	2.82	0.15
3540	ابنتان	2	2.82	0.15	3590	اثر	2	2.82	0.15
3541	ينكر	2	2.82	0.15	3591	اللقة	2	2.82	0.15
3542	المرسلة	2	2.82	0.15	3592	الكست	2	2.82	0.15

APPENDIX C

**Training weights for features in training set consists of 82 Ahadith
form the book of food**

Feature NO.	feature	TF	DF	Training weight	Feature NO.	feature	TF	DF	Training weight
1	ااذن	1	2.42193	2.421933	50	استقر اذك	1	3.120903	3.120903
2	ابتعنه	1	3.12090	3.120903	51	استنظره	1	3.120903	3.120903
3	ابصرته	1	3.12090	3.120903	52	استوى	1	2.819873	2.819873
4	ابيع	1	3.12090	3.120903	53	استيقظ	1	2.421933	2.421933
5	اتاه	1	2.16666	2.16666	54	اسد	1	2.643782	2.643782
6	اتي	7	2.00695	14.04872	55	اسلم	1	2.16666	2.16666
7	اثلثا	1	3.12090	3.120903	56	اسمعه	1	2.275805	2.275805
8	اجد	1	2.42193	2.421933	57	اشتياه	1	2.819873	2.819873
9	اجمع	1	3.12090	3.120903	58	اشدهن	1	3.120903	3.120903
10	اجمعه	1	3.12090	3.120903	59	اشرف	1	2.518843	2.518843
11	اجمعون	1	2.81987	2.819873	60	اشهد	1	2.120903	2.120903
12	احب	5	1.61575	8.078764	61	اصابنا	1	3.120903	3.120903
13	احبه	1	3.12090	3.120903	62	اصابني	1	2.819873	2.819873
14	احداهن	2	2.42193	4.843866	63	اصبح	1	2.120903	2.120903
15	احدثهم	1	3.12090	3.120903	64	اصبحت	1	2.643782	2.643782
16	احرام	2	2.81987	5.639746	65	اصحاب	2	1.974775	3.94955
17	احرم	1	2.51884	2.518843	66	اصحابه	4	1.759175	7.0367
18	احق	1	2.27580	2.275805	67	اصلي	1	2.16666	2.16666
19	اخاه	1	2.16666	2.16666	68	اصنع	2	2.421933	4.843866
20	اخبرن	1	3.12090	3.120903	69	اصول	1	2.819873	2.819873
21	اخبرنا	1	2.12090	2.120903	70	اضطر كم	1	3.120903	3.120903
22	اخبرني	1	1.94481	1.944812	71	اطعموا	1	2.819873	2.819873
23	اختها	1	2.81987	2.819873	72	اعافه	2	2.819873	5.639746
24	اخدم	1	3.12090	3.120903	73	اعتق	1	2.643782	2.643782
25	اخدمه	1	3.12090	3.120903	74	اعجب	1	3.120903	3.120903
26	اخرجت	1	2.81987	2.819873	75	اعرف	1	2.342752	2.342752
27	اخرى	2	2.42193	4.843866	76	اعطاه	1	2.518843	2.518843
28	اخصف	1	3.12090	3.120903	77	اعلم	3	1.673745	5.021234
29	ادخل	1	2.51884	2.518843	78	اعوذ	1	3.120903	3.120903
30	ادخلتك	1	3.12090	3.120903	79	افرش	1	3.120903	3.120903
31	ادري	1	2.07951	2.07951	80	افعل	1	2.041722	2.041722
32	ادعو	2	2.64378	5.287563	81	اقال	1	2.819873	2.819873
33	ادم	1	2.00695	2.006959	82	اقبل	1	2.006959	2.006959
34	اذا	1	2.81987	2.819873	83	اقبلنا	1	3.120903	3.120903
35	اذنت	2	2.81987	5.639746	84	اقرا	1	2.518843	2.518843
36	اراد	1	1.94481	1.944812	85	اقراصا	1	2.819873	2.819873
37	ارادت	1	2.81987	2.819873	86	اقول	1	2.518843	2.518843
38	اراه	1	2.51884	2.518843	87	اقبمت	1	2.16666	2.16666
39	اربع	1	1.86563	1.86563	88	اكفنا	1	3.120903	3.120903
40	اربعين	1	2.27580	2.275805	89	اكل	11	1.819873	20.0186
41	ارتفاع	1	3.12090	3.120903	90	اكلا	1	3.120903	3.120903
42	ارتفع	1	3.12090	3.120903	91	اكلت	2	2.643782	5.287563
43	ارسلتي	1	2.81987	2.819873	92	اكلتين	1	3.120903	3.120903
44	ارسلك	1	2.64378	2.643782	93	اكلن	1	3.120903	3.120903
45	ارسلي	1	2.64378	2.643782	94	اكله	1	2.819873	2.819873
46	ازال	1	3.12090	3.120903	95	اكلهم	1	3.120903	3.120903
47	ازل	5	2.27580	11.37902	96	اكنت	1	2.819873	2.819873
48	استطاع	1	2.51884	2.518843	97	اكون	1	2.643782	2.643782
49	استطع	1	3.12090	3.120903	98	الاترجة	1	3.120903	3.120903

Feature NO.	feature	TF	DF	Training weight	Feature NO.	feature	TF	DF	Training weight
99	الآخرة	1	2.34275	2.3427515	148	الدين	1	2.120903	2.1209028
100	الآذن	1	2.64378	2.6437815	149	الذهب	1	2.421933	2.4219328
101	الآربعة	1	3.12090	3.1209028	150	الرجال	1	2.16666	2.1666603
102	الآرض	1	1.74069	1.7406915	151	الرجل	2	1.288394	2.5767878
103	الآسلام	1	1.74069	1.7406915	152	الرحمن	1	1.643782	1.6437815
104	الآسودين	2	2.81987	5.6397456	153	الرطاب	1	3.120903	3.1209028
105	الآضاحي	1	3.12090	3.1209028	154	الرطب	3	2.643782	7.9313446
106	الآظ	1	3.12090	3.1209028	155	الريحانة	1	3.120903	3.1209028
107	الآمطار	1	3.12090	3.1209028	156	الزمر	1	3.120903	3.1209028
108	الآنصار	3	1.70592	5.1177884	157	السفر	2	2.217813	4.4356256
109	الآنصاري	1	2.12090	2.1209028	158	الساق	1	3.120903	3.1209028
110	الآية	2	1.89045	3.7809077	159	السوط	1	3.120903	3.1209028
111	الآبس	1	3.12090	3.1209028	160	الشاة	1	2.342752	2.3427515
112	الآبطن	1	3.12090	3.1209028	161	الشام	1	2.041722	2.0417215
113	الآبعر	1	2.64378	2.6437815	162	الشجر	2	2.275805	4.5516095
114	الآبيت	2	1.64378	3.2875631	163	الشعير	2	2.518843	5.0376856
115	الآتفت	1	3.12090	3.1209028	164	الصحفة	3	2.643782	7.9313446
116	الآتمر	3	2.27580	6.8274143	165	الصلاة	6	1.125268	6.7516057
117	الآتمره	1	3.12090	3.1209028	166	الضرب	2	2.819873	5.6397456
118	الآتمس	1	3.12090	3.1209028	167	الطعام	4	2.041722	8.1668862
119	الآتيمن	1	3.12090	3.1209028	168	الظهران	1	3.120903	3.1209028
120	الآثانية	2	2.27580	4.5516095	169	العاني	1	3.120903	3.1209028
121	الآثريد	1	2.64378	2.6437815	170	العذاب	1	3.120903	3.1209028
122	الآثلاثة	1	2.51884	2.5188428	171	العشاء	2	1.916783	3.8335656
123	الآثلاثين	1	2.81987	2.8198728	172	العضد	1	3.120903	3.1209028
124	الآثوم	1	3.12090	3.1209028	173	العكة	1	3.120903	3.1209028
125	الآجانع	1	3.12090	3.1209028	174	الغلام	1	2.421933	2.4219328
126	الآجداد	1	3.12090	3.1209028	175	الغنم	1	2.07951	2.0795101
127	الآجمعة	1	2.21781	2.2178128	176	الغني	2	2.643782	5.2875631
128	الآجهد	1	3.12090	3.1209028	177	الفرس	1	2.421933	2.4219328
129	الآجوع	2	2.51884	5.0376856	178	الفقير	2	2.819873	5.6397456
130	الآحارث	2	2.34275	4.6855031	179	القدح	1	2.819873	2.8198728
131	الآحبلة	1	3.12090	3.1209028	180	القران	2	2.006959	4.0139189
132	الآحجاب	1	2.34275	2.3427515	181	القصة	2	2.643782	5.2875631
133	الآحرير	2	2.51884	5.0376856	182	القصعين	1	3.120903	3.1209028
134	الآحشفة	1	3.12090	3.1209028	183	القوم	3	1.673745	5.0212343
135	الآحصين	1	3.12090	3.1209028	184	الكافر	2	2.518843	5.0376856
136	الآحضور	1	3.12090	3.1209028	185	الكبات	1	3.120903	3.1209028
137	الآحلواء	1	2.81987	2.8198728	186	الكراع	2	2.643782	5.2875631
138	الآحمار	1	2.81987	2.8198728	187	اللبن	1	2.819873	2.8198728
139	الآحمد	2	2.16666	4.3333206	188	اللحم	1	3.120903	3.1209028
140	الآحنظلة	1	3.12090	3.1209028	189	اللبل	1	1.615753	1.6157528
141	الآخبز	1	2.81987	2.8198728	190	المؤمن	4	2.006959	8.0278378
142	الآخمير	1	3.12090	3.1209028	191	المائدة	1	3.120903	3.1209028
143	الآدار	1	2.51884	2.5188428	192	المريض	1	2.275805	2.2758047
144	الآدياء	5	2.04172	10.208607	193	المسجد	1	1.602389	1.6023888
145	الآدخشن	1	3.12090	3.1209028	194	المسلم	3	1.916783	5.7503485
146	الآدنيا	1	1.84214	1.8421492	195	المسلمين	1	2.041722	2.0417215
147	الآديباج	1	3.12090	3.1209028	196	المغرب	2	1.842149	3.6842984

Feature NO.	feature	TF	DF	Training weight	Feature NO.	feature	TF	DF	Training weight
197	المنافق	2	2.34275	4.6855031	246	بارك	1	3.120903	3.1209028
198	المنافقين	1	3.12090	3.1209028	247	باكلهن	1	3.120903	3.1209028
199	النار	3	1.50811	4.5243568	248	بالاسود	1	3.120903	3.1209028
200	النخل	1	2.51884	2.5188428	249	بالانطاع	1	3.120903	3.1209028
201	النخلة	2	2.27580	4.5516095	250	بالحجاب	1	2.819873	2.8198728
202	النساء	1	1.79868	1.7986835	251	بالحصباء	1	3.120903	3.1209028
203	النسوة	1	3.12090	3.1209028	252	بالصهباء	4	2.518843	10.075371
204	النطاقان	1	3.12090	3.1209028	253	بالعشاء	2	2.421933	4.8438656
205	النطاقين	1	3.12090	3.1209028	254	بالغداء	1	3.120903	3.1209028
206	النعم	1	3.12090	3.1209028	255	بالفتاء	3	2.643782	7.9313446
207	النقي	2	2.81987	5.6397456	256	بالناس	1	2.120903	2.1209028
208	النهار	2	1.94481	3.8896231	257	بالنطاقين	1	3.120903	3.1209028
209	اله	1	1.81987	1.8198728	258	ببعضه	1	2.819873	2.8198728
210	الهدى	1	1.91678	1.9167828	259	بجمار	1	2.819873	2.8198728
211	الهم	1	3.12090	3.1209028	260	بخيز	1	3.120903	3.1209028
212	الواحد	1	2.27580	2.2758047	261	بدا	1	2.275805	2.2758047
213	الوادي	1	2.07951	2.0795101	262	بدباء	1	3.120903	3.1209028
214	الوضوء	1	2.27580	2.2758047	263	بدرا	1	2.643782	2.6437815
215	الولاء	1	2.81987	2.8198728	264	بر	2	2.518843	5.0376856
216	اليهودي	1	2.81987	2.8198728	265	بركته	1	3.120903	3.1209028
217	اليوم	1	1.68953	1.6895390	266	برمة	1	2.819873	2.8198728
218	امامنا	1	3.12090	3.1209028	267	بريرة	1	2.819873	2.8198728
219	امر	2	1.51884	3.0376856	268	بزئنب	1	3.120903	3.1209028
220	امراة	1	1.57683	1.5768347	269	بسواد	1	3.120903	3.1209028
221	امري	1	2.64378	2.6437815	270	بسويق	2	2.819873	5.6397456
222	امشوا	1	3.12090	3.1209028	271	بشيء	1	1.740692	1.7406915
223	امشي	1	2.42193	2.4219328	272	بشيرا	1	3.120903	3.1209028
224	امعاء	5	2.42193	12.109664	273	بصري	1	2.643782	2.6437815
225	امه	1	2.34275	2.3427515	274	بصفية	2	2.819873	5.6397456
226	انتشل	1	3.12090	3.1209028	275	بصلا	1	3.120903	3.1209028
227	انسان	1	2.42193	2.4219328	276	بضب	1	3.120903	3.1209028
228	انظر	1	1.97477	1.9747747	277	بطريق	1	2.518843	2.5188428
229	انظره	1	3.12090	3.1209028	278	بطعام	5	2.342752	11.713757
230	انكرت	1	3.12090	3.1209028	279	بطعامه	1	3.120903	3.1209028
231	انهى	1	2.64378	2.6437815	280	بطنها	1	2.643782	2.6437815
232	انية	1	3.12090	3.1209028	281	بطني	2	2.819873	5.6397456
233	اهدت	2	2.81987	5.6397456	282	بعياءة	1	3.120903	3.1209028
234	اهل	2	1.24008	2.4801784	283	بعثتني	1	3.120903	3.1209028
235	اهله	1	1.64378	1.6437815	284	بعس	1	3.120903	3.1209028
236	اهلها	1	2.27580	2.2758047	285	بعيد	1	2.643782	2.6437815
237	اومن	1	3.12090	3.1209028	286	بعغم	1	3.120903	3.1209028
238	ايام	2	1.84214	3.6842984	287	بقبضة	1	3.120903	3.1209028
239	اياه	2	2.42193	4.8438656	288	بقصعة	1	2.819873	2.8198728
240	اية	1	1.91678	1.9167828	289	بقليل	1	3.120903	3.1209028
241	ايديهم	1	2.21781	2.2178128	290	بقي	1	2.421933	2.4219328
242	ايطب	1	3.12090	3.1209028	291	بكساء	1	3.120903	3.1209028
243	باب	1	2.21781	2.2178128	292	بكير	1	2.819873	2.8198728
244	باحدهما	1	3.12090	3.1209028	293	بلغ	1	2.07951	2.0795101
245	بارض	2	2.34275	4.6855031	294	بماء	3	2.07951	6.2385303

Feature NO.	feature	TF	DF	Training weight	Feature NO.	feature	TF	DF	Training weight
295	بمثله	1	2.81987	2.8198728	344	تمررات	4	2.275805	9.1032191
296	بمر	1	3.12090	3.1209028	345	تمررة	1	2.120903	2.1209028
297	بمرقة	1	3.12090	3.1209028	346	تمرري	1	3.120903	3.1209028
298	بمسكين	1	3.12090	3.1209028	347	تخلون	1	3.120903	3.1209028
299	بناءه	1	2.81987	2.8198728	348	توفي	2	2.275805	4.5516095
300	بنى	1	2.64378	2.6437815	349	ثريناه	1	3.120903	3.1209028
301	بواسط	1	3.12090	3.1209028	350	ثلاث	2	1.819873	3.6397456
302	بيت	1	2.00695	2.0069594	351	ثلاثا	1	1.944812	1.9448115
303	بيتك	1	3.12090	3.1209028	352	ثلاثة	2	1.643782	3.2875631
304	بيته	1	2.27580	2.2758047	353	ثلاثين	1	2.421933	2.4219328
305	بيتي	1	2.51884	2.5188428	354	ثمامة	1	2.819873	2.8198728
306	بيدي	1	2.27580	2.2758047	355	ثمانون	1	2.643782	2.6437815
307	بيع	1	2.64378	2.6437815	356	ثوبي	1	3.120903	3.1209028
308	ببمبلك	1	3.12090	3.1209028	357	ثوما	1	3.120903	3.1209028
309	بيوم	1	3.12090	3.1209028	358	جنت	3	2.217813	6.6534384
310	توكل	1	3.12090	3.1209028	359	جنتا	1	2.643782	2.6437815
311	تابعه	1	1.75917	1.7591749	360	جاء	2	1.430707	2.8614134
312	تاتي	1	2.81987	2.8198728	361	جاءه	1	2.217813	2.2178128
313	تاخذ	1	2.64378	2.6437815	362	جاع	2	2.819873	5.6397456
314	تاكلوا	1	3.12090	3.1209028	363	جالسا	1	2.421933	2.4219328
315	تاكلون	1	2.81987	2.8198728	364	جبل	1	2.275805	2.2758047
316	تبعنا	2	2.81987	5.6397456	365	جبليها	1	3.120903	3.1209028
317	تحب	1	2.51884	2.5188428	366	ججش	1	3.120903	3.1209028
318	تدخل	1	2.64378	2.6437815	367	جد	1	3.120903	3.1209028
319	تدري	1	2.81987	2.8198728	368	جريج	1	2.518843	2.5188428
320	تراه	1	2.64378	2.6437815	369	جشته	1	3.120903	3.1209028
321	ترعى	1	3.12090	3.1209028	370	جعل	1	2.120903	2.1209028
322	تركته	2	2.64378	5.2875631	371	جعلت	1	2.819873	2.8198728
323	تركه	1	2.42193	2.4219328	372	جلوس	2	2.342752	4.6855031
324	تزوجها	1	3.12090	3.1209028	373	جهد	1	3.120903	3.1209028
325	تسمعه	1	2.81987	2.8198728	374	حازها	1	3.120903	3.1209028
326	تشتريها	1	3.12090	3.1209028	375	حيات	1	3.120903	3.1209028
327	تشرىوا	1	2.81987	2.8198728	376	حجر	1	2.518843	2.5188428
328	تصبح	1	2.81987	2.8198728	377	حجرة	1	2.643782	2.6437815
329	تصدق	1	2.16666	2.1666603	378	حديث	1	2.16666	2.1666603
330	تضع	1	3.12090	3.1209028	379	حراما	2	2.643782	5.2875631
331	تضيفت	1	3.12090	3.1209028	380	حرم	3	2.006959	6.0208783
332	تطيش	1	3.12090	3.1209028	381	حره	1	3.120903	3.1209028
333	تعرق	1	3.12090	3.1209028	382	حز	1	3.120903	3.1209028
334	تعرقها	1	3.12090	3.1209028	383	حزة	1	3.120903	3.1209028
335	تعزني	1	3.12090	3.1209028	384	حزن	1	2.643782	2.6437815
336	تفارقة	1	3.12090	3.1209028	385	حشفة	2	2.819873	5.6397456
337	تفور	1	2.81987	2.8198728	386	حفيد	1	3.120903	3.1209028
338	تقارنوا	1	3.12090	3.1209028	387	حفيدة	1	3.120903	3.1209028
339	تقر	1	3.12090	3.1209028	388	حلو	1	3.120903	3.1209028
340	تقل	1	3.12090	3.1209028	389	حماما	1	3.120903	3.1209028
341	تكون	1	2.00695	2.0069594	390	حمر	1	2.518843	2.5188428
342	تلبسوا	1	2.81987	2.8198728	391	حوالي	2	2.819873	5.6397456
343	تمرا	4	2.51884	10.075371	392	حول	1	2.421933	2.4219328

Feature NO.	feature	TF	DF	Training weight	Feature NO.	Feature	TF	DF	Training weight
393	حبسا	2	2.64378	5.2875631	442	رماه	1	3.120903	3.1209028
394	حبي	1	2.34275	2.3427515	443	روحة	3	2.643782	7.9313446
395	خادمه	1	3.12090	3.1209028	444	رومة	1	3.120903	3.1209028
396	خالة	1	3.12090	3.1209028	445	ريح	1	2.342752	2.3427515
397	خالته	1	3.12090	3.1209028	446	ريحها	1	3.120903	3.1209028
398	خالتي	1	2.64378	2.6437815	447	زالت	1	2.275805	2.2758047
399	خامس	2	2.64378	5.2875631	448	زرناها	1	3.120903	3.1209028
400	خباز	1	3.12090	3.1209028	449	زمان	2	2.217813	4.4356256
401	خباها	1	3.12090	3.1209028	450	زوجها	1	2.643782	2.6437815
402	خبز	4	2.34275	9.3710062	451	سانر	1	2.518843	2.5188428
403	خبزا	2	2.81987	5.6397456	452	سابع	1	3.120903	3.1209028
404	خرجنا	2	1.86563	3.7312606	453	سافعل	1	3.120903	3.1209028
405	خرجوا	2	2.42193	4.8438656	454	سال	2	1.842149	3.6842984
406	خزير	1	3.12090	3.1209028	455	سالت	1	1.916783	1.9167828
407	خسرت	1	3.12090	3.1209028	456	سالم	1	2.217813	2.2178128
408	خطيفة	1	3.12090	3.1209028	457	ساله	1	2.643782	2.6437815
409	خمارا	1	2.81987	2.8198728	458	سبع	3	2.006959	6.0208783
410	خمس	3	1.79868	5.3960505	459	سبعاً	1	2.120903	2.1209028
411	خمسة	2	2.16666	4.3333206	460	سبعة	6	2.006959	12.041756
412	خوان	1	2.81987	2.8198728	461	سترا	1	2.518843	2.5188428
413	خياط	1	2.81987	2.8198728	462	سحر	1	2.275805	2.2758047
414	خياطا	4	2.51884	10.075371	463	سراتهم	1	3.120903	3.1209028
415	خيبر	4	2.07951	8.3180405	464	سعيي	1	3.120903	3.1209028
416	داره	1	2.81987	2.8198728	465	سفرته	1	3.120903	3.1209028
417	دباء	4	2.51884	10.075371	466	سكرجة	1	2.819873	2.8198728
418	دخل	2	1.47745	2.9549002	467	سلم	1	2.07951	2.0795101
419	دخلا	1	3.12090	3.1209028	468	سليم	2	2.07951	4.1590202
420	دسته	1	2.81987	2.8198728	469	سم	3	2.217813	6.6534384
421	دعا	6	1.91678	11.500697	470	سمعته	1	2.16666	2.1666603
422	دعوتنا	1	3.12090	3.1209028	471	سمنا	1	3.120903	3.1209028
423	نوو	1	3.12090	3.1209028	472	سنة	1	2.006959	2.0069594
424	راسي	1	1.94481	1.9448115	473	سنن	1	3.120903	3.1209028
425	راى	4	1.70592	6.8237178	474	سهل	1	2.518843	2.5188428
426	رايت	7	1.28839	9.0187573	475	سهلا	1	3.120903	3.1209028
427	رايتم	1	2.16666	2.1666603	476	سواد	1	2.643782	2.6437815
428	رايتني	1	2.81987	2.8198728	477	سويد	1	3.120903	3.1209028
429	ربنا	2	2.27580	4.5516095	478	سويقا	1	3.120903	3.1209028
430	ربييه	1	3.12090	3.1209028	479	شنت	3	2.217813	6.6534384
431	ربيع	1	3.12090	3.1209028	480	شاء	2	1.673745	3.3474895
432	رجال	3	2.16666	6.4999809	481	شاة	4	1.974775	7.8990991
433	رجالا	1	2.51884	2.5188428	482	شانه	1	2.421933	2.4219328
434	رجل	3	1.07558	3.2267395	483	شاهدا	1	2.819873	2.8198728
435	رجلا	4	1.39662	5.5865077	484	شبيع	3	2.518843	7.5565284
436	رحله	1	2.81987	2.8198728	485	شبعنا	2	2.819873	5.6397456
437	رطب	1	2.81987	2.8198728	486	شبعوا	2	2.518843	5.0376856
438	رهاها	1	3.12090	3.1209028	487	شجرة	1	2.217813	2.2178128
439	رفع	2	1.89045	3.7809077	488	شحم	1	3.120903	3.1209028
440	ركبت	1	2.81987	2.8198728	489	شدت	1	2.041722	2.0417215
441	ركعتين	1	1.57683	1.5768347	490	شديد	1	2.518843	2.5188428

Feature NO.	feature	TF	DF	Training weight	Feature NO.	feature	TF	DF	Training weight
491	شرطية	1	3.12090	3.1209028	540	عشرة	4	2.006959	8.0278378
492	شعبة	1	1.97477	1.9747747	541	عطاء	1	2.217813	2.2178128
493	شعير	5	2.34275	11.713757	542	عطية	1	2.643782	2.6437815
494	شققته	1	3.12090	3.1209028	543	عكة	2	2.643782	5.2875631
495	شكاة	1	3.12090	3.1209028	544	علمت	1	2.421933	2.4219328
496	شكوا	1	2.81987	2.8198728	545	عمدت	1	2.518843	2.5188428
497	شهاب	2	2.07951	4.1590202	546	عمله	1	2.166666	2.1666603
498	شهد	1	2.42193	2.4219328	547	عندها	2	2.041722	4.0834431
499	صاع	1	2.81987	2.8198728	548	عهد	2	1.759175	3.5183499
500	صحافها	1	3.12090	3.1209028	549	عودا	1	3.120903	3.1209028
501	صدقة	1	1.74069	1.7406915	550	عبروه	1	3.120903	3.1209028
502	صلينا	1	2.64378	2.6437815	551	عبينة	1	2.421933	2.4219328
503	صنع	3	2.07951	6.2385303	552	غانبا	1	3.120903	3.1209028
504	صنعتة	1	2.81987	2.8198728	553	غلام	4	1.944812	7.7792462
505	صنعناه	1	3.12090	3.1209028	554	غلاما	3	2.518843	7.5565284
506	صنعه	3	2.51884	7.5565284	555	غلامه	1	2.643782	2.6437815
507	صوت	1	2.42193	2.4219328	556	علمانكم	1	3.120903	3.1209028
508	ضب	1	3.12090	3.1209028	557	فابدؤوا	2	2.819873	5.6397456
509	ضبا	1	3.12090	3.1209028	558	فابصرته	1	3.120903	3.1209028
510	ضبابا	1	3.12090	3.1209028	559	فابصروا	1	3.120903	3.1209028
511	ضعيفا	1	2.81987	2.8198728	560	فابي	1	2.421933	2.4219328
512	طار	1	3.12090	3.1209028	561	فاتاه	1	2.342752	2.3427515
513	طريق	1	2.64378	2.6437815	562	فاتت	1	2.643782	2.6437815
514	طعام	5	2.04172	10.208607	563	فاتخذة	1	3.120903	3.1209028
515	طعاما	4	2.21781	8.8712513	564	فاتي	1	2.07951	2.0795101
516	طعامه	1	2.51884	2.5188428	565	فاتي	2	1.944812	3.8896231
517	طعمتي	1	3.12090	3.1209028	566	فاتينه	1	2.518843	2.5188428
518	طعيما	1	3.12090	3.1209028	567	فاجتررتة	1	3.120903	3.1209028
519	طهوره	1	3.12090	3.1209028	568	فاجتمعوا	1	3.120903	3.1209028
520	طويل	1	3.12090	3.1209028	569	فاجدني	2	2.819873	5.6397456
521	طيب	1	2.51884	2.5188428	570	فاخير	1	2.421933	2.4219328
522	طييا	1	3.12090	3.1209028	571	فاخيرته	1	2.342752	2.3427515
523	ظاهر	1	3.12090	3.1209028	572	فاخذ	1	2.275805	2.2758047
524	ظن	1	2.81987	2.8198728	573	فاخذتهما	1	3.120903	3.1209028
525	عاب	1	2.81987	2.8198728	574	فاخرجت	1	2.643782	2.6437815
526	عابس	1	3.12090	3.1209028	575	فادخلت	1	2.819873	2.8198728
527	عارها	1	3.12090	3.1209028	576	فادرکنا	1	2.643782	2.6437815
528	عاشر	1	3.12090	3.1209028	577	فادمته	1	2.819873	2.8198728
529	عام	3	2.00695	6.0208783	578	فانن	1	1.974775	1.9747747
530	عاما	1	3.12090	3.1209028	579	فاننت	1	2.819873	2.8198728
531	عبيد	1	2.42193	2.4219328	580	فاراد	1	2.342752	2.3427515
532	عتبان	1	3.12090	3.1209028	581	فاردت	1	2.421933	2.4219328
533	عجوة	1	2.51884	2.5188428	582	فاستادن	1	3.120903	3.1209028
534	عجوز	1	3.12090	3.1209028	583	فاستسقى	1	2.643782	2.6437815
535	عد	2	2.81987	5.6397456	584	فاستقراته	1	3.120903	3.1209028
536	عدد	1	3.12090	3.1209028	585	فاسرجه	1	3.120903	3.1209028
537	عرفا	1	3.12090	3.1209028	586	فاسلم	1	2.819873	2.8198728
538	عروسا	1	2.81987	2.8198728	587	فاشترى	1	2.643782	2.6437815
539	عريشك	1	3.12090	3.1209028	588	فاشرب	1	3.120903	3.1209028

Feature NO.	feature	TF	DF	Training weight	Feature NO.	feature	TF	DF	Training weight
589	فاشرت	1	3.12090	3.1209028	664	فصفنا	1	3.120903	3.1209028
590	فاصابني	2	2.81987	5.6397456	665	فصلى	5	1.47745	7.3872507
591	فاعطاني	1	2.81987	2.8198728	666	فصنع	1	2.421933	2.4219328
592	فاعطي	1	3.12090	3.1209028	667	فصنعت	1	3.120903	3.1209028
593	فاقمني	1	3.12090	3.1209028	668	فضحكت	1	2.518843	2.5188428
594	فاقبل	2	2.16666	4.3333206	669	فضرب	1	2.518843	2.5188428
595	فاكل	5	1.97477	9.8738739	670	فضل	1	2.518843	2.5188428
596	فاكلته	1	3.12090	3.1209028	671	فطاف	1	1.916783	1.9167828
597	فاكلن	1	3.12090	3.1209028	672	فظننت	1	3.120903	3.1209028
598	فاكلنا	3	2.51884	7.5565284	673	ففعجن	1	3.120903	3.1209028
599	فاكلناه	1	3.12090	3.1209028	674	فعدت	1	2.819873	2.8198728
600	فاكلها	1	3.12090	3.1209028	675	فعر ف	1	2.819873	2.8198728
601	فاكلوا	3	2.42193	7.2657984	676	فعقرته	1	3.120903	3.1209028
602	فالتفت	1	2.81987	2.8198728	677	فعله	2	2.421933	4.8438656
603	فالقاها	2	2.81987	5.6397456	678	فعلى	1	2.819873	2.8198728
604	فالقي	1	3.12090	3.1209028	679	فغدا	1	3.120903	3.1209028
605	فامر	3	1.81987	5.4596184	680	فغضبت	1	3.120903	3.1209028
606	فامسك	1	2.81987	2.8198728	681	ففت	1	2.819873	2.8198728
607	فانطلق	2	1.97477	3.9495495	682	ففر شته	1	3.120903	3.1209028
608	فاهدته	1	3.12090	3.1209028	683	فقام	3	1.643782	4.9313446
609	فاهوى	2	2.81987	5.6397456	684	فقدمت	1	2.518843	2.5188428
610	فاو كيت	1	3.12090	3.1209028	685	فقرب	2	2.819873	5.6397456
611	فبسطت	1	3.12090	3.1209028	686	فقربته	1	3.120903	3.1209028
612	فبشرته	1	3.12090	3.1209028	687	فقمت	3	2.041722	6.1251647
613	فنبعهم	2	2.64378	5.2875631	688	فكبر	1	2.518843	2.5188428
614	فتجعل	1	3.12090	3.1209028	689	فكلم	1	2.819873	2.8198728
615	فتجعله	1	3.12090	3.1209028	690	فكلمه	1	3.120903	3.1209028
616	فتصلي	1	3.12090	3.1209028	691	فلاك	1	3.120903	3.1209028
617	فتعتقها	1	3.12090	3.1209028	692	فلان	1	1.705929	1.7059294
618	فتمضمض	1	3.12090	3.1209028	693	فلانة	1	2.643782	2.6437815
619	فثاب	1	3.12090	3.1209028	694	فلفت	1	3.120903	3.1209028
620	فجنت	2	2.27580	4.5516095	695	فلقيت	1	2.643782	2.6437815
621	فجنته	1	3.12090	3.1209028	696	فلكنا	1	3.120903	3.1209028
622	فجاءني	1	2.81987	2.8198728	697	فلكناه	2	2.819873	5.6397456
623	فجاؤوني	1	3.12090	3.1209028	698	فليعتزلنا	1	3.120903	3.1209028
624	فجددت	1	3.12090	3.1209028	699	فليعجل	1	3.120903	3.1209028
625	فجعل	3	1.60238	4.8071666	700	فليناوله	1	3.120903	3.1209028
626	فجعلت	4	1.91678	7.6671313	701	فمشى	1	2.819873	2.8198728
627	فجلس	1	2.04172	2.0417215	702	فمشيت	1	3.120903	3.1209028
628	فجلست	1	2.81987	2.8198728	703	فمضمض	3	2.518843	7.5565284
629	فجيء	1	2.81987	2.8198728	704	فناكله	1	3.120903	3.1209028
630	فحضرت	1	2.81987	2.8198728	705	فناولته	1	3.120903	3.1209028
631	فحملته	1	2.81987	2.8198728	706	فنزلت	1	2.342752	2.3427515
632	فخرج	2	1.75917	3.5183499	707	فنشقتها	1	3.120903	3.1209028
633	فخرجت	1	2.21781	2.2178128	708	فنلحق	1	3.120903	3.1209028
634	فخررت	1	3.12090	3.1209028	709	فوجد	1	2.518843	2.5188428
635	فخلا	1	3.12090	3.1209028	710	فوجدت	1	2.643782	2.6437815
636	فخيرت	1	3.12090	3.1209028	711	فوددت	1	3.120903	3.1209028
637	فدخل	4	1.89045	7.5618155	712	فوضع	1	2.217813	2.2178128

Feature NO.	feature	TF	DF	Training weight	Feature NO.	feature	TF	DF	Training weight
713	فوضعتنه	1	3.12090	3.1209028	762	لحام	2	2.643782	5.2875631
714	فوقعوا	1	3.12090	3.1209028	763	لحق	1	2.819873	2.8198728
715	فوقف	1	2.81987	2.8198728	764	لحم	1	2.643782	2.6437815
716	فولى	1	3.12090	3.1209028	765	لحما	1	3.120903	3.1209028
717	فيابى	1	2.81987	2.8198728	766	لحوم	2	2.643782	5.2875631
718	فيطعمنا	1	3.12090	3.1209028	767	لرجل	1	2.275805	2.2758047
719	فيطعمنى	1	3.12090	3.1209028	768	لرسول	2	2.342752	4.6855031
720	فيطير	1	3.12090	3.1209028	769	لشبع	1	3.120903	3.1209028
721	فيهن	1	2.81987	2.8198728	770	لضرسى	1	3.120903	3.1209028
722	قائل	1	2.51884	2.5188428	771	لطعام	4	2.421933	9.6877312
723	قائم	1	2.27580	2.2758047	772	لعشرة	1	2.819873	2.8198728
724	قابل	1	3.12090	3.1209028	773	لعطاء	1	2.518843	2.5188428
725	قام	7	1.44880	10.141634	774	لقمة	1	2.643782	2.6437815
726	قاموا	1	2.64378	2.6437815	775	لقمتين	1	3.120903	3.1209028
727	قبض	1	2.34275	2.3427515	776	لقومى	1	3.120903	3.1209028
728	قبضه	1	3.12090	3.1209028	777	لقى	2	2.342752	4.6855031
729	قدر	2	2.27580	4.5516095	778	لكنه	1	2.819873	2.8198728
730	قدمت	1	2.27580	2.2758047	779	للطعام	1	3.120903	3.1209028
731	قدمتن	1	3.12090	3.1209028	780	للمساكين	1	3.120903	3.1209028
732	قربة	1	2.81987	2.8198728	781	لنرفع	2	2.819873	5.6397456
733	قسم	3	2.42193	7.2657984	782	لنفرح	1	3.120903	3.1209028
734	قصعتين	1	3.12090	3.1209028	783	لهن	1	3.120903	3.1209028
735	قضى	1	2.07951	2.0795101	784	لوجهي	1	3.120903	3.1209028
736	قضينه	1	3.12090	3.1209028	785	لولا	1	2.041722	2.0417215
737	قط	2	1.84214	3.6842984	786	لياكل	1	3.120903	3.1209028
738	قطعة	1	2.27580	2.2758047	787	ليخرج	1	2.819873	2.8198728
739	قلما	1	3.12090	3.1209028	788	ليعتزل	1	3.120903	3.1209028
740	قول	1	1.91678	1.9167828	789	مائدة	2	2.819873	5.6397456
741	قوموا	1	2.34275	2.3427515	790	مائدته	4	2.518843	10.075371
742	قومى	2	2.42193	4.8438656	791	مات	1	2.041722	2.0417215
743	كافي	1	3.12090	3.1209028	792	مادوم	2	2.819873	5.6397456
744	كالفدح	1	3.12090	3.1209028	793	مباركا	1	3.120903	3.1209028
745	كالمستفذر	1	3.12090	3.1209028	794	متكنا	1	2.643782	2.6437815
746	كانك	1	2.81987	2.8198728	795	متكىء	1	2.819873	2.8198728
747	كبركة	1	3.12090	3.1209028	796	مثله	2	2.07951	4.1590202
748	كتاب	1	2.07951	2.0795101	797	مجوسى	1	3.120903	3.1209028
749	كتف	2	2.64378	5.2875631	798	محرم	1	2.275805	2.2758047
750	كتفا	1	3.12090	3.1209028	799	محرمون	1	3.120903	3.1209028
751	كثير	1	2.12090	2.1209028	800	محمود	1	3.120903	3.1209028
752	كرهه	1	3.12090	3.1209028	801	محنوذ	1	3.120903	3.1209028
753	كعب	1	2.34275	2.3427515	802	محنوذا	1	3.120903	3.1209028
754	كفانا	1	3.12090	3.1209028	803	مد	1	2.819873	2.8198728
755	كفضل	1	2.64378	2.6437815	804	مدهم	1	3.120903	3.1209028
756	كمثل	1	2.27580	2.2758047	805	مر	1	2.120903	2.1209028
757	لاصحابه	1	2.81987	2.8198728	806	مرة	2	2.16666	4.3333206
758	لاصلي	1	3.12090	3.1209028	807	مرتين	1	2.07951	2.0795101
759	لين	1	2.42193	2.4219328	808	مرفق	1	2.819873	2.8198728
760	لييك	1	2.27580	2.2758047	809	مرققا	1	2.819873	2.8198728
761	لجابر	1	3.12090	3.1209028	810	مست	1	2.819873	2.8198728

Feature NO.	feature	TF	DF	Training weight	Feature NO.	feature	TF	DF	Training weight
811	مستغنى	2	2.81987	5.6397456	860	ننفخه	1	3.120903	3.1209028
812	مسجدنا	2	2.81987	5.6397456	861	نهمته	1	3.120903	3.1209028
813	مسجدهم	1	3.12090	3.1209028	862	نهى	1	1.77848	1.7784801
814	مسموطة	1	3.12090	3.1209028	863	نهيته	1	3.120903	3.1209028
815	مشرك	1	2.64378	2.6437815	864	نواحي	1	2.819873	2.8198728
816	مشعان	1	3.12090	3.1209028	865	نومه	1	3.120903	3.1209028
817	مشغول	1	3.12090	3.1209028	866	هبة	1	3.120903	3.1209028
818	مشوي	1	3.12090	3.1209028	867	هر	1	3.120903	3.1209028
819	مصلى	1	3.12090	3.1209028	868	هلمي	1	2.819873	2.8198728
820	مضاغى	1	3.12090	3.1209028	869	واحبا	1	3.120903	3.1209028
821	مكاهم	1	3.12090	3.1209028	870	واحد	4	1.722963	6.8918512
822	مكفور	1	3.12090	3.1209028	871	وادم	1	3.120903	3.1209028
823	مكفى	2	2.81987	5.6397456	872	واروانا	1	3.120903	3.1209028
824	مناخل	1	3.12090	3.1209028	873	واستقرئ	1	3.120903	3.1209028
825	مناديل	1	3.12090	3.1209028	874	واضبا	1	3.120903	3.1209028
826	مناقق	1	2.81987	2.8198728	875	واعتقت	1	2.819873	2.8198728
827	منخلا	1	3.12090	3.1209028	876	واقبل	1	2.819873	2.8198728
828	منخول	1	3.12090	3.1209028	877	واقدامنا	1	3.120903	3.1209028
829	منزل	1	2.64378	2.6437815	878	واقض	1	3.120903	3.1209028
830	مودع	2	2.81987	5.6397456	879	واقطا	2	2.819873	5.6397456
831	مولى	1	2.34275	2.3427515	880	واقيمت	1	3.120903	3.1209028
832	ناحية	1	2.42193	2.4219328	881	واكل	1	3.120903	3.1209028
833	نازل	1	3.12090	3.1209028	882	والاقط	1	3.120903	3.1209028
834	نافع	2	2.21781	4.4356256	883	والاله	1	3.120903	3.1209028
835	ناكل	1	2.51884	2.5188428	884	والبخل	1	3.120903	3.1209028
836	ناولوني	1	3.12090	3.1209028	885	والجين	1	3.120903	3.1209028
837	نبي	2	1.75917	3.5183499	886	والجوع	1	3.120903	3.1209028
838	ننزود	1	3.12090	3.1209028	887	والحزن	1	3.120903	3.1209028
839	ننغدى	1	3.12090	3.1209028	888	والرمح	1	3.120903	3.1209028
840	ننوضا	1	2.42193	2.4219328	889	والسكين	2	2.819873	5.6397456
841	نجد	2	2.04172	4.0834431	890	والسمن	1	3.120903	3.1209028
842	نجني	1	3.12090	3.1209028	891	والصق	1	3.120903	3.1209028
843	نحوه	1	2.27580	2.2758047	892	والعجز	1	3.120903	3.1209028
844	نخلة	1	2.64378	2.6437815	893	والعسل	1	2.819873	2.8198728
845	نخلي	1	3.12090	3.1209028	894	والفضة	1	2.643782	2.6437815
846	نرى	1	2.12090	2.1209028	895	والقوم	2	2.643782	5.2875631
847	نزل	1	1.86563	1.8656303	896	والكافر	3	2.643782	7.9313446
848	نستتظر	1	3.12090	3.1209028	897	والكسل	1	3.120903	3.1209028
849	نصفين	1	3.12090	3.1209028	898	والماء	2	2.819873	5.6397456
850	نصلي	1	2.16666	2.1666603	899	وامراته	1	3.120903	3.1209028
851	نطاي	1	3.12090	3.1209028	900	وانزل	1	2.421933	2.4219328
852	نطحنه	1	3.12090	3.1209028	901	وانطلقت	1	2.819873	2.8198728
853	نطع	2	2.81987	5.6397456	902	وبدءا	1	3.120903	3.1209028
854	نطعمهم	1	2.81987	2.8198728	903	وبينهم	1	3.120903	3.1209028
855	نعلي	1	2.81987	2.8198728	904	وترجله	1	3.120903	3.1209028
856	نعينك	1	3.12090	3.1209028	905	وتركهن	1	3.120903	3.1209028
857	نفرح	1	3.12090	3.1209028	906	وتتعله	1	3.120903	3.1209028
858	نقص	1	2.81987	2.8198728	907	وجدناه	1	2.819873	2.8198728
859	نقيل	1	3.12090	3.1209028	908	وجعلت	2	2.518843	5.0376856

Feature NO.	feature	TF	DF	Training weight	Feature NO.	Feature	TF	DF	Training weight
909	وجلس	1	3.12090	3.1209028	958	وليمته	1	3.120903	3.1209028
910	وجه	2	2.16666	4.3333206	959	ومرقا	2	2.819873	5.6397456
911	وجهه	2	1.62954	3.2590822	960	ومشيت	1	3.120903	3.1209028
912	وحبسانه	1	3.12090	3.1209028	961	ومضمضنا	2	2.819873	5.6397456
913	وحدثي	1	2.64378	2.6437815	962	ونحبه	1	2.819873	2.8198728
914	وحشفة	1	3.12090	3.1209028	963	ونسيت	1	2.518843	2.5188428
915	وحشيا	1	3.12090	3.1209028	964	ونصيحته	1	3.120903	3.1209028
916	وحضر	1	3.12090	3.1209028	965	وننفخه	1	3.120903	3.1209028
917	وخادمه	1	3.12090	3.1209028	966	وهديه	1	3.120903	3.1209028
918	وخالة	1	3.12090	3.1209028	967	ويسمي	1	3.120903	3.1209028
919	وخبات	1	3.12090	3.1209028	968	ويقول	1	2.342752	2.3427515
920	وخير	1	2.81987	2.8198728	969	يوثي	1	2.643782	2.6437815
921	ودخل	1	3.12090	3.1209028	970	يؤذوني	1	3.120903	3.1209028
922	ودك	1	3.12090	3.1209028	971	ياكل	8	1.798684	14.389468
923	وذكرت	1	3.12090	3.1209028	972	ياكله	1	3.120903	3.1209028
924	وراءه	1	2.27580	2.2758047	973	ياكلها	1	3.120903	3.1209028
925	ورجعت	1	3.12090	3.1209028	974	ياكلون	1	2.643782	2.6437815
926	وردتي	1	3.12090	3.1209028	975	ياكلونه	1	3.120903	3.1209028
927	ورسول	4	2.16666	8.6666412	976	يبتغي	1	3.120903	3.1209028
928	ورق	1	3.12090	3.1209028	977	يبني	1	3.120903	3.1209028
929	وسعديك	1	2.81987	2.8198728	978	يتتبع	5	2.342752	11.713757
930	وسمعه	1	3.12090	3.1209028	979	يتوضا	6	2.275805	13.654828
931	وسواعنا	1	3.12090	3.1209028	980	يجده	1	3.120903	3.1209028
932	وشبعنا	1	3.12090	3.1209028	981	يجلس	1	2.819873	2.8198728
933	وشبعوا	1	2.81987	2.8198728	982	يجلسه	1	3.120903	3.1209028
934	وشرب	1	2.81987	2.8198728	983	يحب	3	1.974775	5.9243243
935	وصاعهم	1	3.12090	3.1209028	958	وليمته	1	3.120903	3.1209028
936	وصلينا	1	3.12090	3.1209028	959	ومرقا	2	2.819873	5.6397456
937	وضع	2	2.42193	4.8438656	960	ومشيت	1	3.120903	3.1209028
938	وضل	1	3.12090	3.1209028	961	ومضمضنا	2	2.819873	5.6397456
939	وضلع	1	3.12090	3.1209028	962	ونحبه	1	2.819873	2.8198728
940	وطعام	1	3.12090	3.1209028	963	ونسيت	1	2.518843	2.5188428
941	وطعامه	1	3.12090	3.1209028	964	ونصيحته	1	3.120903	3.1209028
942	وطعمها	1	3.12090	3.1209028	965	وننفخه	1	3.120903	3.1209028
943	وطعمها	1	3.12090	3.1209028	966	وهديه	1	3.120903	3.1209028
944	وعرف	1	3.12090	3.1209028	967	ويسمي	1	3.120903	3.1209028
945	وعصرت	2	2.64378	5.2875631	968	ويقول	1	2.342752	2.3427515
946	وعلاجه	1	3.12090	3.1209028	969	يوثي	1	2.643782	2.6437815
947	وعودوا	1	3.12090	3.1209028	970	يؤذوني	1	3.120903	3.1209028
948	وغلبة	1	3.12090	3.1209028	971	ياكل	8	1.798684	14.389468
949	وفتحها	1	3.12090	3.1209028	972	ياكله	1	3.120903	3.1209028
950	وفضل	2	2.64378	5.2875631	973	ياكلها	1	3.120903	3.1209028
951	وفكوا	1	3.12090	3.1209028	974	ياكلون	1	2.643782	2.6437815
952	وقديد	3	2.51884	7.5565284	975	ياكلونه	1	3.120903	3.1209028
953	وكل	2	2.64378	5.2875631	976	يبتغي	1	3.120903	3.1209028
954	ولانا	1	3.12090	3.1209028	977	يبني	1	3.120903	3.1209028
955	ولبنا	1	3.12090	3.1209028	978	يتتبع	5	2.342752	11.713757
956	ولكنه	1	2.81987	2.8198728	979	يتوضا	6	2.275805	13.654828
957	ولي	1	2.42193	2.4219328	980	يجده	1	3.120903	3.1209028

Feature NO.	feature	TF	DF	Training weight	Feature NO.	Feature	TF	DF	Training weight
981	يجلس	1	2.81987	2.8198728	1030	يوما	4	1.689539	6.7581562
982	يجلسه	1	3.12090	3.1209028					
983	يحب	3	1.97477	5.9243243					
984	يحبنا	1	2.81987	2.8198728					
985	يحتز	2	2.64378	5.2875631					
986	يحدث	1	2.00695	2.0069594					
987	يحيوي	1	3.12090	3.1209028					
988	يخدمني	2	2.81987	5.6397456					
989	يدعوا	1	3.12090	3.1209028					
990	يده	6	1.67374	10.042468					
991	يدي	2	1.97477	3.9495495					
992	يديه	2	1.72296	3.4459256					
993	يردني	1	3.12090	3.1209028					
994	يردفا	1	3.12090	3.1209028					
995	يريد	1	1.89045	1.8904538					
996	يسار	1	2.64378	2.6437815					
997	يسالني	1	2.51884	2.5188428					
998	يستازن	1	2.81987	2.8198728					
999	يسلفني	1	3.12090	3.1209028					
1000	يسوقها	1	3.12090	3.1209028					
1001	يشوي	1	3.12090	3.1209028					
1002	يصلي	1	1.46769	1.4676903					
1003	يضره	1	2.51884	2.5188428					
1004	يضع	1	2.64378	2.6437815					
1005	يطعم	2	2.81987	5.6397456					
1006	يعتقبون	1	3.12090	3.1209028					
1007	يعني	1	1.57683	1.5768347					
1008	يعبرون	1	3.12090	3.1209028					
1009	يعبرونك	1	3.12090	3.1209028					
1010	يقدم	1	2.34275	2.3427515					
1011	يقرا	1	2.42193	2.4219328					
1012	يقربن	1	3.12090	3.1209028					
1013	يقولون	1	2.42193	2.4219328					
1014	يكثر	1	2.81987	2.8198728					
1015	يكفي	1	3.12090	3.1209028					
1016	يكلم	1	3.12090	3.1209028					
1017	يكني	1	3.12090	3.1209028					
1018	يلعقها	1	3.12090	3.1209028					
1019	يليك	3	2.64378	7.9313446					
1020	يمر	1	2.21781	2.2178128					
1021	يمسح	1	2.27580	2.2758047					
1022	يمنع	1	2.64378	2.6437815					
1023	يناول	1	3.12090	3.1209028					
1024	يناولوا	1	3.12090	3.1209028					
1025	ينظر	2	2.21781	4.4356256					
1026	ينقلب	1	3.12090	3.1209028					
1027	يهودي	1	2.81987	2.8198728					
1028	يوضع	1	2.64378	2.6437815					
1029	يوفظ	1	2.64378	2.6437815					

استكشاف المعلومات في الحديث الشريف باستخدام خوارزميه تصنيف النصوص

إعداد

ختام محمود عبد الله جباره

المشرف

الدكتور عزام طلال سليط

المشرف المشارك

الدكتور بسام حسن حمو

ملخص

هنالك حاجة متزايدة في مختلف المجالات لتحليل وفهم الوثائق النصية من خلال استخلاص المعلومات والمعرفة، وتبرز تلك الحاجة بشكل كبير عند التعامل مع النصوص البيانية ذات الأحجام الكبيرة. ولهذا السبب فإنه من الضروري استغلال العمليات المتوفرة في مجال التنقيب في النصوص لاكتشاف واستخلاص المعرفة من البيانات غير المنظمة (أي النصوص).

ولأن الحديث الشريف هو المصدر الثاني للتشريع لكافة المسلمين في جميع أنحاء العالم بعد القرآن الكريم، جاءت فكرة فكره تطبيق خوارزميه تصنيف اوتوماتيكيه لمعرفة الكتاب او التصنيف الذي ينتمي له الحديث ليسهل الوصول الى معلومه تابعه لتصنيف محدد من قبل المسلمين وغير المسلمين.

في هذه الرسالة البحثية يتم اقتراح تطبيق خوارزمية تصنيف نصوص لتصنيف الأحاديث إلى كتب محده مسبقاً. تقنية تصنيف النصوص تعتبر أحد العمليات التي تقع ضمن عملية تنقيب

النصوص. والتقنية المقترحة تقسم إلى مرحلتين: في المرحلة الأولى يتم فيها استخدام مجموعه من النصوص المصنفة مسبقاً لعملية تدريب خوارزمية التصنيف المستخدمة و الحصول على قاعدة بيانات سوف تستخدم في مرحلة التصنيف، وفي المرحلة الثانية يتم تصنيف كل حديث اعتماداً على خوارزمية تصنيف تم التحسين عليها من خلال استخدام توسيع البحث باستخدام جذور الكلمات.

لقد تم تطبيق الخوارزمية المقترحة على 13 كتاباً من كتب الحديث الوارده في صحيح البخاري . وأظهرت النتائج أن الخوارزمية المقترحة للتقيب في الحديث كانت مقبولة من أجل تصنيف الحديث ضمن الكتاب الاقرب لموضوع الحديث.

ولتقييم فعالية التقنية المقترحة تم مقارنة نتائجها مع النتائج التي تم الحصول عليها من تطبيق خوارزميه تصنيف تعتمد على الكلمات دون مرحله التوسيع للبحث و من تطبيق خوارزمية الكعبي المنشوره في بحثه ، حيث كانت الخوارزميه المقترحه تحقق نتائج افضل من الطريقتين السابقتين التي تم استخدامهما في مجل التقيب في الحديث الشريف لمعظم الكتب التي استخدمت في التجربة .