CSE 454 Data Mining Final Project Report

Ahmed Semih Özmekik 171044039

January 24, 2021

Abstract

Detailed explanation of design choices along with the experimental results in the homework.

1 Project Definition

Our main goal in the project was to use many data mining algorithms in the literature that can be used at the point where we will specialize by applying many preprocessing, postprocessing, methods related to data mining.

Based on this, we first determined a paper. [1] Then, we focused on improving the results obtained there by examining the experimental study in the paper. This project was firstly an implementation of a paper, and then by doing different studies on the dataset suggested in this paper, it was aimed and achieved better scores than the scores obtained in this paper.

The subject is text classification. There are many different and successful methods in the current literature on this subject. We present the results of 8 different methods. For each method, there are 6 different preprocessing parameters, 2 different feature vector method parameters, and 5 different feature number definition parameters for the feature selection method. In other words, 8x2x5 = 80 different results were obtained for each method, hence 80x8 = 640 different results obtained in total.

For these 480 different results, by making some comparisons between them, the results of all methods and independent variables in the experiment were analyzed, and the dependent variable f1 score was observed. As a result, all data were recorded and the most successful parameter selection was determined.

2 Text Classification

The number of electronic documents produced as a result of the transition to the electronic world is increasing day by day. Manual processing or classification of these electronic documents in text format, whose number is rapidly growing, has become almost impossible today. Today, text classification is carried out by machine learning or deep learning methods. Text classification, sorting the content of the text according to the specified categories is the process.[2, 3]

There are many noise words in a text that are not specific to that text and do not represent an attribute for the classification problem. [4] That's why preprocessing is important in the TC problem.

After applying the necessary pre-operations to the text, a feature must be extracted from the text, in other words, the text must be converted into a feature vector. For this, there are many vectorization studies such as word2vec, bag of words.

After obtaining this vector, it is a classification method that will be applied. For example, there are many statistics based methods such as naive bayes, support vector machine.

3 Dataset

The name of the dataset is TTC-3600.

The dataset includes a total of 3600 documents, 600 of them from each class (economy, culture-arts, health, politics, sports and technology), all collected from well-known and known news portals.

In the reference study, only 3 variances were produced from this dataset, and there were 4 different datasets in total, including the raw dataset.

The first of these is the raw dataset that we will call Original-DS throughout this study, on which no preprocess has been applied (except for the correction of html, css tags during the data set collection stage).

The second is the dataset, named F5-DS, processed using FPS-5 (mentioned in the preprocessing section) as stemmer. FPS-7 was used as stemmer in the third one, F7-DS, and Zemberek was used as stemmer in the last one, Zemb-DS.

Here, we further diversified our experiment based on this study. By processing the raw dataset in different ways, we had 8 different datasets in total. Now let's talk in detail about preprocessing processes that we implement.

4 Preprocessing

Preprocessing is the most important process of TC subject. At this stage, we have applied several methods on our data set, usually based on removal, that is, to remove noise on the data.

First of all, we extracted the numbers from the whole document. This work was not done in the article. Then, we performed a normalization by converting all words to lower cases, and operations such as removing punctuation marks, separators, operators, or meaningless characters were performed. Finally, all words were tokenized. These were common pre-processing operations across all datasets (including the raw set).

Then here, we first created 3 more datasets by applying 3 different stemming methods. These are the processes using the FPS-5, FPS-7, and Zemberek method.

The FPS method is a simple stemming method where only the first 'n' letter of the word is kept. [5] In FPS-5, only the first 5 letters are kept, and in FPS-7, only the first 7 letters are

kept and other letters are deleted.

Zemberek, on the other hand, is a Turkish dictionary-based tool for stemming. [6] Here, too, we see that besides the simple assumption in FPS, a more in-depth and accurate root separation is applied. But we will see how it affects the overall result with the experiments we have done.

In the reference study, words called stopwords, which do not represent any meaning and feature in the content of the text, were removed from the whole data set, such as conjunctions and prepositions.

In our research, we did not accept this assumption, and we divided our data set, which was 4 in total, into 2, with 3 different datasets created by raw and the stemming processes mentioned above, one with stopwords removed and no stopwords removed, and in total we had 8 different preprocess product datasets.

Below, you can examine these datasets, along with their names, in more detail.

Dataset name	Stop word filtering	Stemmer
Original-DS	No	No stemmer
F5-DS	No	FPS-5
F7-DS	No	FPS-7
Zemb-DS	No	Zemberek
OriginalSW-DS	Yes	No Stemmer
F5SW-DS	Yes	FPS-5
F7SW-DS	Yes	FPS-7
ZembSW-DS	Yes	Zemberek

5 Postprocessing

5.1 Feature Extraction

After preprocessing the datasets, we prepared independent variables for all our experiments, where another processing parameters can be specified.

The first category represents the different methods that can be applied in extracting a feature vector from the texts we have. We chose two approaches in the literature to use in our experiments: Bag of words and TFIDF vectorization methods.

These methods are methods that infer a feature vector from a text that represents that text. Bag of words method, as its name also hints, defines a feature vector by approaching texts as a word bag, losing the spatial data of the words, that is, the sequence information, using only the word histogram in that text. Compared to being a simple method, it has yielded successful results in the literature.

TFIDF (term frequency-inverse document frequency), on the other hand, is a relatively slightly more complex numerical statistic that aims to reflect how important a word is to a document in a collection or collection.

In both feature extraction processes, the maximum number of feature can be obtained is limited to 8000. The reference study recorded the highest feature number in the raw dataset, 7508.

5.2 Feature Selection

After the feature is obtained from the texts, another important independent variable is the parameter that specifies 'n' features selected by chi2 feature selection method among these features. The user can give the desired number and parameter. In the context of our experiments, we conducted this experiment through discrete values, namely 500, 1000, 2000, 5000 or 'all' as the feature number, with options to select all features.

These two parameters, the method to be applied for feature extraction and the feature number to be selected in feature selection, indicate the last parameters in the data process. From now on, the only parameter that can be changed is the classification method to be applied on the ready data. Let's examine them now.

6 Methods

6.1 Naive Bayes

The Naive Bayes classifier is based on Bayes' theorem. It is a lazy learning algorithm, it can also work on unstable datasets. The way the algorithm works calculates the probability of each state for an element and classifies it according to the one with the highest probability value. With a little training data, he can do very successful jobs. If a value in the test set has an unobservable value in the training set, it gives 0 as a probability value, which means it cannot predict. This condition is commonly known as Zero Frequency. Correction techniques can be used to resolve this situation. One of the simplest correction techniques is known as Laplace estimation. Examples of usage areas are real-time prediction, multi-class prediction, text classification, spam filtering, sentiment analysis and suggestion systems.

6.2 Random Forest

Random Forest is one of the popular machine learning models because it can be applied to both regression and classification problems, giving good results without hyperparameter estimation. To understand the random forest, it is necessary to first understand the decision trees, which is the basic blog of this model. We have devoted the 3rd lesson to this subject, a successful decision tree can be compared to people who ask questions and make accurate predictions that will increase their knowledge gain in daily life.

However, one of the biggest problems of decision trees, which is one of the traditional methods, is over-learning-overfitting. In order to solve this problem, the random forest model randomly selects 10s and 100s of different subsets from both the data set and the feature set, and trains them. With this method, hundreds of decision trees are created and each decision tree makes individual predictions. At the end of the day, if our problem is

regression, if our problem is classifying the average of the estimates of the decision trees, we choose the most votes among the predictions.

6.3 Support Vector Machine (Linear)

Suppose, in classification with support vector machines, samples belonging to two classes are linearly distributed. In this case, it is aimed to distinguish these two classes with the help of a decision function obtained using training data. It is called the correct decision line that divides the data set into two. Although it is possible to draw infinite decision lines, the important thing is to determine the optimal decision line. In order for the decision line to be resistant to the newly added data, the border line must be at the closest distance to the border lines of the two classes. The points closest to this border line are called support points. Class labels in the form of (-1, +1) are generally used in classification with support vector machines.

6.4 K-Nearest Neighbour

The K-NN (K-Nearest Neighbor) algorithm is one of the simplest and most used classification algorithm. K-NN is a non-parametric (non-parametric), lazy (lazy) learning algorithm. If we try to understand the concept of lazy, unlike eager learning, lazy learning does not have a training stage. It does not learn the training data but instead "memorizes" the training data set. When we want to make a guess, it looks for the closest neighbors in the entire data set. In the operation of the algorithm, a K value is determined. The meaning of this K value is the number of elements to look at. When a value comes, the distance between the value is calculated by taking the nearest K number of elements. The Euclidean function is generally used in the distance calculation. Manhattan, Minkowski and Hamming functions can also be used as an alternative to the Euclidean function. After the distance is calculated, it is sorted and the incoming value is assigned to the appropriate class.

6.5 CART

Classification and Decision tree (CART) learning is one of the predictive modeling approaches used in statistics, data mining and machine learning. Uses a decision tree to navigate from observations about an item to conclusions about the item's target value.

6.6 Rocchio

The Rocchio algorithm is based on a conformity feedback method found in information access systems originating from the SMART Information Retrieval System developed between 1960 and 1964. Like many other retrieval systems, the Rocchio feedback approach has been developed using the Vector Space Model.

6.7 LogisticRegression

Logistic Regression is a regression method for classification. It is used to classify categorical or numerical data. It is widely used in linear classification problems. For this reason, it is very similar to Linear Regression.

Logistic Regression is often known as binary classifications, and is only true / false, positive / negative, etc. used in binary classifications. But of course, it can turn every multi class problem into a binary classification when desired. We also used this feature of logistic regression in our multi class classification.

7 Experiment

While applying experiments, we cross validated with 10 folds.

First, the results of how each method works on different parameters will be shown. Then, a comparison of all methods will be made with their highest values.

7.1 Naive Bayes

			D	
	1	1	Dataset	Score 0.8903
			OrgDS ZembDS	0.8903
		500	F5DS	0.9144
			F7DS	0.9158
			OrgDS	0.9089
		1000	ZembDS	0.9264
		1000	F5DS	0.9225
			F7DS	0.9233
			OrgDS	0.9208
	TFIDF	2000	ZembDS	0.9300
	IFIDE	2000	F5DS	0.9303
			F7DS	0.9297
			OrgDS	0.9261
		5000	ZembDS	0.9339
		3000	F5DS	0.9319
			F7DS	0.9322
			OrgDS	0.9256
		8000 (all)	ZembDS	0.9342
			F5DS	0.9306
S			F7DS	0.9328
Z			OrgDS	0.8914
Ş		500	ZembDS	0.9164
2			F5DS	0.9136
2		<u> </u>	F7DS	0.9150
S			OrgDS ZembDS	0.9083 0.9267
WITH STOPWORDS		1000	F5DS	0.9267
M			F7DS	0.9211
_		<u> </u>	OrgDS	0.9174
	l		ZembDS	0.9294
	BOW	2000	F5DS	0.9267
			F7DS	0.9283
			OrgDS	0.9289
			ZembDS	0.9386
		5000		
			F5DS	0.9322
			F7DS	0.9292
			OrgDS	0.9297
		8000 (all)	ZembDS F5DS	0.9378 0.9322
			F7DS	0.9314
			OrgDS	0.7511
			ZembDS	0.8783
		500	F5DS	0.8697
			F7DS	0.8581
			OrgDS	0.8106
		4000	ZembDS	0.8997
		1000	F5DS	0.8958
			F7DS	0.8850
			OrgDS	0.8608
	TFIDF	2000	ZembDS	0.9117
	TITLE	2000	F5DS	
				0.9142
			F7DS	0.9047
			F7DS OrgDS	0.9047 0.8942
		5000	F7DS OrgDS ZembDS	0.9047 0.8942 0.9267
		5000	F7DS OrgDS ZembDS F5DS	0.9047 0.8942 0.9267 0.9281
		5000	F7DS OrgDS ZembDS F5DS F7DS	0.9047 0.8942 0.9267 0.9281 0.9189
		5000	F7DS OrgDS ZembDS F5DS F7DS OrgDS	0.9047 0.8942 0.9267 0.9281 0.9189 0.9042
		5000 8000 (all)	F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS	0.9047 0.8942 0.9267 0.9281 0.9189 0.9042 0.9267
			F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS	0.9047 0.8942 0.9267 0.9281 0.9189 0.9042 0.9267 0.9264
92			F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS	0.9047 0.8942 0.9267 0.9281 0.9189 0.9042 0.9267 0.9264 0.9272
~			F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS F7DS OrgDS	0.9047 0.8942 0.9267 0.9281 0.9189 0.9042 0.9267 0.9264 0.9272 0.7222
~			F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS	0.9047 0.8942 0.9267 0.9281 0.9189 0.9042 0.9267 0.9264 0.9272 0.7222 0.8711
~		8000 (all)	F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F7DS OrgDS ZembDS F7DS OrgDS ZembDS F5DS	0.9047 0.8942 0.9267 0.9281 0.9189 0.9042 0.9267 0.9264 0.9272 0.7222 0.8711 0.8636
~		8000 (all)	F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F7DS F7DS F7DS F7DS F7DS F7DS F7DS F7	0.9047 0.8942 0.9267 0.9281 0.9189 0.9042 0.9267 0.9264 0.9272 0.7222 0.8711 0.8636 0.8439
~		8000 (all)	F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS	0.9047 0.8942 0.9267 0.9281 0.9189 0.9042 0.9267 0.9264 0.9272 0.7222 0.8711 0.8636 0.8439 0.8036
~		8000 (all)	F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F7DS OrgDS ZembDS F5DS F7DS ZembDS F5DS F5DS F5DS F5DS F5DS F5DS F5DS F5	0.9047 0.8942 0.9267 0.9281 0.9189 0.9042 0.9267 0.9264 0.9272 0.7222 0.8711 0.8636 0.8439 0.8036 0.8992
~		8000 (all)	F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS F7DS OrgDS ZembDS F5DS F7DS F7DS F7DS F7DS F7DS F7DS F7DS F7	0.9047 0.8942 0.9267 0.9281 0.9189 0.9042 0.9267 0.9264 0.9272 0.77222 0.8711 0.8636 0.8439 0.8036 0.8992
~		8000 (all)	F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS F7DS F7DS F7DS F7DS F7DS F7DS F7	0.9047 0.8942 0.9267 0.9281 0.9189 0.9042 0.9264 0.9272 0.7222 0.8713 0.8636 0.8439 0.8036 0.8922 0.8803
WITHOUT STOPWORDS		8000 (all) 500 1000	F7DS OrgDS ZembDS F5DS F7DS OrgDS	0.9047 0.8942 0.9267 0.9281 0.9189 0.9042 0.9264 0.9272 0.7222 0.8711 0.8636 0.8439 0.8036 0.8992 0.8922 0.8803 0.8803
~	BOW	8000 (all)	F7DS OrgDS ZembDS F5DS F5DS F5DS F5DS F5DS F5DS F5DS F5	0.9047 0.8942 0.9267 0.9281 0.9189 0.9042 0.9264 0.9272 0.7222 0.8711 0.8636 0.8439 0.8036 0.8992 0.8922 0.8921 0.8544
~	BOW	8000 (all) 500 1000	F7DS OrgDS ZembDS F5DS F7DS F5DS F7DS F7DS F7DS F7DS F7DS F7DS F7DS F7	0.9047 0.8942 0.9267 0.9281 0.9189 0.9042 0.9267 0.9264 0.9272 0.8711 0.8639 0.8439 0.8922 0.8922 0.8922 0.8923 0.8928 0.99119
~	BOW	8000 (all) 500 1000	F7DS OrgDS ZembDS F5DS F7DS F7DS F7DS F7DS F7DS F7DS F7DS F7	0.9047 0.8942 0.9267 0.9281 0.9189 0.9042 0.9264 0.9272 0.8711 0.8636 0.8992 0.8922 0.8803 0.8544 0.9119 0.9081
~	BOW	8000 (all) 500 1000 2000	F7DS OrgDS ZembDS F5DS F7DS OrgDS	0.9047 0.8942 0.9267 0.9281 0.9189 0.90467 0.9264 0.9272 0.7222 0.8711 0.8636 0.8493 0.8036 0.8992 0.8922 0.8924 0.9081 0.9081 0.8997
~	BOW	8000 (all) 500 1000	F7DS OrgDS ZembDS F5DS F5DS F5DS F5DS F5DS F5DS F5DS F5	0.9047 0.8942 0.9267 0.9281 0.9189 0.9042 0.9267 0.9264 0.9272 0.8711 0.8639 0.8439 0.8922 0.8922 0.8921 0.9081 0.9081 0.9081
~	BOW	8000 (all) 500 1000 2000	F7DS OrgDS ZembDS F5DS F7DS F7DS F7DS F5DS F7DS F5DS F7DS F5DS F7DS F5DS F7DS F7DS F7DS F7DS F7DS F7DS F7DS F7	0.9047 0.8942 0.9267 0.9281 0.9189 0.9042 0.9267 0.9264 0.9272 0.8711 0.8636 0.8439 0.8036 0.8992 0.8803 0.8544 0.9119 0.9086 0.9066 0.9256 0.9256
~	BOW	8000 (all) 500 1000 2000	F7DS OrgDS ZembDS F5DS F7DS F7DS F7DS F7DS F7DS F7DS F7DS F7	0.9047 0.8942 0.9267 0.9281 0.9189 0.90427 0.9264 0.9272 0.8711 0.8636 0.8992 0.8922 0.8803 0.8944 0.9119 0.9096 0.9294 0.9294
~	BOW	8000 (all) 500 1000 2000 5000	F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS F7DS F5DS F7DS F5DS F7DS OrgDS ZembDS F5DS F5DS F5DS F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS	0.9047 0.8942 0.9267 0.9281 0.9189 0.9264 0.9272 0.7222 0.8711 0.8639 0.8036 0.8922 0.8803 0.8941 0.9019 0.9081 0.9081 0.9090 0.9000 0.9000 0.9000 0.9000 0.900
~	BOW	8000 (all) 500 1000 2000	F7DS OrgDS ZembDS F5DS F7DS F7DS F7DS F7DS F7DS F7DS F7DS F7	0.9047 0.8942 0.9267 0.9281 0.9189 0.90427 0.9264 0.9272 0.8711 0.8636 0.8992 0.8922 0.8803 0.8944 0.9119 0.9096 0.9294 0.9294

7.2 Random Forest

SCOPE SCOP				D	
TFIDF 500 ZembDS 0.9033 F5DS 0.8825 0.9034 0.9035				Dataset	Score
TFIDF					
TFIDF 1000 F7DS 0.8950 0.8950 0.8950 0.8950 0.8955 0.9104 0.8950			500		
### TFIDF 1000 F5DS 0.8825 ZembDS 0.9080 0.9089 0.90					
TFIDF 1000 ZembDS 0.9089 0.9089 0.9081 0.9089					
TFIDF 2000 F5DS 0.9114 F7DS 0.9007 0.8875 2embDS 0.9108 F5DS 0.9114 F7DS 0.9017 0rgDS 0.8958 2embDS 0.9158			1000		
TFIDF 2000 Corg DS Comb DS Com			1000		
TFIDF				F7DS	0.9000
Part	İ			OrgDS	0.8875
SOUD		TEIDE	2000	ZembDS	0.9108
SOUTH SOUT		IFIDE	2000	F5DS	0.9114
SOUDE SOUD					0.9017
SOUD					
SO			5000	ZembDS	0.9150
SOU			3000	F5DS	0.9164
SOUTH SOUT				F7DS	0.9081
SCOOL SCOO					
SOUND SOUN			0000 (11)		0.9133
BOW 2000 F7DS 0.8994 OrgDS 0.8856 ZembDS 0.9003 F5DS 0.9100 F7DS 0.9093 F5DS 0.99100 F7DS 0.90953 OrgDS 0.8994 ZembDS 0.9114 F5DS 0.9136 F7DS 0.9086 OrgDS 0.8950 ZembDS 0.9122 F5DS 0.9136 F7DS 0.9086 OrgDS 0.8950 ZembDS 0.8569 F5DS 0.8334 OrgDS 0.8569 F5DS 0.8334 OrgDS 0.8569 F5DS 0.8533 F7DS 0.85817 OrgDS 0.7839 OrgDS 0.7839 OrgDS 0.88517 OrgDS 0.85817 OrgDS 0.85818 F7DS 0.8581 F7DS 0.8581 F7DS 0.8581 OrgDS 0.8781 F5DS 0.8588 F7DS 0.8581 OrgDS 0.8781 F5DS 0.8588 F7DS 0.8581 OrgDS 0.8781 F5DS 0.8584 F5DS 0.8584 F5DS 0.8581 OrgDS 0.8081 ZembDS 0.8764 F5DS 0.8656 OrgDS 0.8078 ZembDS 0.8764 F5DS 0.8650 OrgDS 0.8078 ZembDS 0.8764 F5DS 0.8581 F7DS 0.8650 OrgDS 0.8078 ZembDS 0.8764 F5DS 0.8650 OrgDS 0.8083 ZembDS 0.8714 F5DS 0.8696 F7DS 0.8593 F7DS 0.8596 F7DS 0.8593 F7DS 0.8772			8000 (all)	F5DS	0.9122
BOW 2000 F7DS 0.8994 OrgDS 0.8856 ZembDS 0.9003 F5DS 0.9100 F7DS 0.9093 F5DS 0.99100 F7DS 0.90953 OrgDS 0.8994 ZembDS 0.9114 F5DS 0.9136 F7DS 0.9086 OrgDS 0.8950 ZembDS 0.9122 F5DS 0.9136 F7DS 0.9086 OrgDS 0.8950 ZembDS 0.8569 F5DS 0.8334 OrgDS 0.8569 F5DS 0.8334 OrgDS 0.8569 F5DS 0.8533 F7DS 0.85817 OrgDS 0.7839 OrgDS 0.7839 OrgDS 0.88517 OrgDS 0.85817 OrgDS 0.85818 F7DS 0.8581 F7DS 0.8581 F7DS 0.8581 OrgDS 0.8781 F5DS 0.8588 F7DS 0.8581 OrgDS 0.8781 F5DS 0.8588 F7DS 0.8581 OrgDS 0.8781 F5DS 0.8584 F5DS 0.8584 F5DS 0.8581 OrgDS 0.8081 ZembDS 0.8764 F5DS 0.8656 OrgDS 0.8078 ZembDS 0.8764 F5DS 0.8650 OrgDS 0.8078 ZembDS 0.8764 F5DS 0.8581 F7DS 0.8650 OrgDS 0.8078 ZembDS 0.8764 F5DS 0.8650 OrgDS 0.8083 ZembDS 0.8714 F5DS 0.8696 F7DS 0.8593 F7DS 0.8596 F7DS 0.8593 F7DS 0.8772	25			F7DS	0.9108
BOW 2000 F7DS 0.8994 OrgDS 0.8856 ZembDS 0.9003 F5DS 0.9100 F7DS 0.9093 F5DS 0.99100 F7DS 0.90953 OrgDS 0.8994 ZembDS 0.9114 F5DS 0.9136 F7DS 0.9086 OrgDS 0.8950 ZembDS 0.9122 F5DS 0.9136 F7DS 0.9086 OrgDS 0.8950 ZembDS 0.8569 F5DS 0.8334 OrgDS 0.8569 F5DS 0.8334 OrgDS 0.8569 F5DS 0.8533 F7DS 0.85817 OrgDS 0.7839 OrgDS 0.7839 OrgDS 0.88517 OrgDS 0.85817 OrgDS 0.85818 F7DS 0.8581 F7DS 0.8581 F7DS 0.8581 OrgDS 0.8781 F5DS 0.8588 F7DS 0.8581 OrgDS 0.8781 F5DS 0.8588 F7DS 0.8581 OrgDS 0.8781 F5DS 0.8584 F5DS 0.8584 F5DS 0.8581 OrgDS 0.8081 ZembDS 0.8764 F5DS 0.8656 OrgDS 0.8078 ZembDS 0.8764 F5DS 0.8650 OrgDS 0.8078 ZembDS 0.8764 F5DS 0.8581 F7DS 0.8650 OrgDS 0.8078 ZembDS 0.8764 F5DS 0.8650 OrgDS 0.8083 ZembDS 0.8714 F5DS 0.8696 F7DS 0.8593 F7DS 0.8596 F7DS 0.8593 F7DS 0.8772	N			OrgDS	0.8658
BOW 2000 F7DS 0.8994 OrgDS 0.8856 ZembDS 0.9003 F5DS 0.9100 F7DS 0.9093 F5DS 0.99100 F7DS 0.90953 OrgDS 0.8994 ZembDS 0.9114 F5DS 0.9136 F7DS 0.9086 OrgDS 0.8950 ZembDS 0.9122 F5DS 0.9136 F7DS 0.9086 OrgDS 0.8950 ZembDS 0.8569 F5DS 0.8334 OrgDS 0.8569 F5DS 0.8334 OrgDS 0.8569 F5DS 0.8533 F7DS 0.85817 OrgDS 0.7839 OrgDS 0.7839 OrgDS 0.88517 OrgDS 0.85817 OrgDS 0.85818 F7DS 0.8581 F7DS 0.8581 F7DS 0.8581 OrgDS 0.8781 F5DS 0.8588 F7DS 0.8581 OrgDS 0.8781 F5DS 0.8588 F7DS 0.8581 OrgDS 0.8781 F5DS 0.8584 F5DS 0.8584 F5DS 0.8581 OrgDS 0.8081 ZembDS 0.8764 F5DS 0.8656 OrgDS 0.8078 ZembDS 0.8764 F5DS 0.8650 OrgDS 0.8078 ZembDS 0.8764 F5DS 0.8581 F7DS 0.8650 OrgDS 0.8078 ZembDS 0.8764 F5DS 0.8650 OrgDS 0.8083 ZembDS 0.8714 F5DS 0.8696 F7DS 0.8593 F7DS 0.8596 F7DS 0.8593 F7DS 0.8772	≥		F00	ZembDS	0.9025
BOW 2000 F7DS 0.8994 OrgDS 0.8856 ZembDS 0.9003 F5DS 0.9100 F7DS 0.9093 F5DS 0.99100 F7DS 0.90953 OrgDS 0.8994 ZembDS 0.9114 F5DS 0.9136 F7DS 0.9086 OrgDS 0.8950 ZembDS 0.9122 F5DS 0.9136 F7DS 0.9086 OrgDS 0.8950 ZembDS 0.8569 F5DS 0.8334 OrgDS 0.8569 F5DS 0.8334 OrgDS 0.8569 F5DS 0.8533 F7DS 0.85817 OrgDS 0.7839 OrgDS 0.7839 OrgDS 0.88517 OrgDS 0.85817 OrgDS 0.85818 F7DS 0.8581 F7DS 0.8581 F7DS 0.8581 OrgDS 0.8781 F5DS 0.8588 F7DS 0.8581 OrgDS 0.8781 F5DS 0.8588 F7DS 0.8581 OrgDS 0.8781 F5DS 0.8584 F5DS 0.8584 F5DS 0.8581 OrgDS 0.8081 ZembDS 0.8764 F5DS 0.8656 OrgDS 0.8078 ZembDS 0.8764 F5DS 0.8650 OrgDS 0.8078 ZembDS 0.8764 F5DS 0.8581 F7DS 0.8650 OrgDS 0.8078 ZembDS 0.8764 F5DS 0.8650 OrgDS 0.8083 ZembDS 0.8714 F5DS 0.8696 F7DS 0.8593 F7DS 0.8596 F7DS 0.8593 F7DS 0.8772	ΡV		300		0.9042
BOW 2000 F7DS 0.8994 OrgDS 0.8856 ZembDS 0.9003 F5DS 0.9100 F7DS 0.9093 F5DS 0.99100 F7DS 0.90953 OrgDS 0.8994 ZembDS 0.9114 F5DS 0.9136 F7DS 0.9086 OrgDS 0.8950 ZembDS 0.9122 F5DS 0.9136 F7DS 0.9086 OrgDS 0.8950 ZembDS 0.8569 F5DS 0.8334 OrgDS 0.8569 F5DS 0.8334 OrgDS 0.8569 F5DS 0.8533 F7DS 0.85817 OrgDS 0.7839 OrgDS 0.7839 OrgDS 0.88517 OrgDS 0.85817 OrgDS 0.85818 F7DS 0.8581 F7DS 0.8581 F7DS 0.8581 OrgDS 0.8781 F5DS 0.8588 F7DS 0.8581 OrgDS 0.8781 F5DS 0.8588 F7DS 0.8581 OrgDS 0.8781 F5DS 0.8584 F5DS 0.8584 F5DS 0.8581 OrgDS 0.8081 ZembDS 0.8764 F5DS 0.8656 OrgDS 0.8078 ZembDS 0.8764 F5DS 0.8650 OrgDS 0.8078 ZembDS 0.8764 F5DS 0.8581 F7DS 0.8650 OrgDS 0.8078 ZembDS 0.8764 F5DS 0.8650 OrgDS 0.8083 ZembDS 0.8714 F5DS 0.8696 F7DS 0.8593 F7DS 0.8596 F7DS 0.8593 F7DS 0.8772					
BOW 2000 F7DS 0.8994 OrgDS 0.8856 ZembDS 0.9003 F5DS 0.9100 F7DS 0.9093 F5DS 0.99100 F7DS 0.90953 OrgDS 0.8994 ZembDS 0.9114 F5DS 0.9136 F7DS 0.9086 OrgDS 0.8950 ZembDS 0.9122 F5DS 0.9136 F7DS 0.9086 OrgDS 0.8950 ZembDS 0.8569 F5DS 0.8334 OrgDS 0.8569 F5DS 0.8334 OrgDS 0.8569 F5DS 0.8533 F7DS 0.85817 OrgDS 0.7839 OrgDS 0.7839 OrgDS 0.88517 OrgDS 0.85817 OrgDS 0.85818 F7DS 0.8581 F7DS 0.8581 F7DS 0.8581 OrgDS 0.8781 F5DS 0.8588 F7DS 0.8581 OrgDS 0.8781 F5DS 0.8581 OrgDS 0.8781 F5DS 0.8584 F5DS 0.8581 OrgDS 0.8781 F5DS 0.8666 OrgDS 0.8078 ZembDS 0.8764 F5DS 0.8650 OrgDS 0.8078 ZembDS 0.8761 F5DS 0.8696 F7DS 0.8593 F7DS 0.8772 F7DS 0.	S			OrgDS	0.8769
BOW 2000 F7DS 0.8994 OrgDS 0.8856 ZembDS 0.9003 F5DS 0.9100 F7DS 0.9093 F5DS 0.99100 F7DS 0.90953 OrgDS 0.8994 ZembDS 0.9114 F5DS 0.9136 F7DS 0.9086 OrgDS 0.8950 ZembDS 0.9122 F5DS 0.9136 F7DS 0.9086 OrgDS 0.8950 ZembDS 0.8569 F5DS 0.8334 OrgDS 0.8569 F5DS 0.8334 OrgDS 0.8569 F5DS 0.8533 F7DS 0.85817 OrgDS 0.7839 OrgDS 0.7839 OrgDS 0.88517 OrgDS 0.85817 OrgDS 0.85818 F7DS 0.8581 F7DS 0.8581 F7DS 0.8581 OrgDS 0.8781 F5DS 0.8588 F7DS 0.8581 OrgDS 0.8781 F5DS 0.8586 F7DS 0.8581 OrgDS 0.8081 ZembDS 0.8781 F5DS 0.8586 F7DS 0.8581 F7DS 0.8581 OrgDS 0.8081 ZembDS 0.8764 F5DS 0.8656 OrgDS 0.8078 ZembDS 0.8764 F5DS 0.8650 OrgDS 0.8083 ZembDS 0.8714 F5DS 0.8696 F7DS 0.8593 F7DS 0.8772 F7DS 0.87			1000		
BOW 2000 Crg DS Cemb D	×		1000		
BOW 2000 ZembDS 0.9083 F5DS 0.9100 F7DS 0.9053 C7DS 0.9053 C7DS 0.9053 C7DS 0.9054 C7DS 0.9054 C7DS 0.9054 C7DS 0.9054 C7DS 0.9054 C7DS 0.9054 C7DS 0.9078 C7DS 0.9078 C7DS 0.9078 C7DS 0.9078 C7DS 0.9078 C7DS 0.9078 C7DS 0.8545 C7DS 0.8555 C7DS 0.8565 C7DS 0.8565 C7DS 0.8656	_			1	
BOW 2000 F5DS 0.9100 F7DS 0.9053 0.8904 ZembDS 0.9114 F5DS 0.9131 F5DS 0.9086 OrgDS 0.8950 ZembDS 0.9122 F5DS 0.9136 F7DS 0.9086 OrgDS 0.8704 C7DS 0.8772 F7DS 0.8517 OrgDS 0.8581 OrgDS 0.8081 OrgDS OrgDS 0.8081 OrgDS 0.8081 OrgDS 0.8081 OrgDS OrgDS 0.8081 OrgDS OrgDS 0.8081 OrgDS OrgDS 0.8081 OrgDS					
TFIDF 2000 F5DS 0.9053 0.9054 2embDS 0.9114 F5DS 0.9131 F7DS 0.9086 0.9131 F7DS 0.9086 0.9125 0.9136 F7DS 0.9136 F7DS 0.9078 0.		BOW	2000		
TFIDF 2000 South Properties					
TFIDF 2000 ZembDS				1	
TFIDF 2000 F5DS 0.9131 F7DS 0.9086 O7gDS 0.8950 O7gDS 0.9132 F5DS 0.9136 F7DS 0.9078 O7gDS 0.9136 F7DS 0.9078 O7gDS 0.7411 ZembDS 0.8569 F5DS 0.8372 O7gDS 0.8581 F7DS 0.8581 O7gDS 0.8581 F7DS 0.8581 O7gDS 0.8581 F7DS 0.8581 O7gDS 0.8686 F5DS 0.8581 F7DS 0.8581 O7gDS 0.8081 F5DS 0.86686 F7DS 0.86686 F7DS 0.8581 O7gDS 0.8081 O7gDS 0.8818 O7gDS 0.8831 O7gDS					
SOUD COLUMN F7DS 0.9086 0.9086 0.9122 0.9123 0.9124 0.9124 0.9124 0.9124 0.9124 0.9124 0.9124 0.924 0.925 0.925 0.9372 0.925 0.9372 0.925 0.9372 0.925 0			5000		
Note					
Sound Sound Semb Sound Semb Sound Semb Sound Semb Sound Semb S					
TFIDF 500 (all) F5DS 0.9136 F7DS 0.9078 0.79DS 0.7411 ZembDS 0.8569 F5DS 0.8372 F7DS 0.8364 0.79DS 0.7605 0.79DS 0.7605 0.8513 0.79DS 0.8513 F7DS 0.8581 0.79DS 0.8081 0.79DS 0.8666 0.79DS 0.8078 0.8666 0.79DS 0.8078 0.8650 0.79DS 0.8078 0.8755 0.79DS 0.8294 0.79DS 0.8414 0.79DS 0.8414 0.79DS 0.8650 0.79DS 0.8650 0.79DS 0.8414 0.79DS 0.8650 0.					
TFIDF 2000 FDS 0.9078 OrgDS 0.7411 ZembDS 0.8569 F5DS 0.8333 F7DS 0.8517 OrgDS 0.7656 ZembDS 0.8669 F5DS 0.8533 F7DS 0.8517 OrgDS 0.8581 F7DS 0.8566 OrgDS 0.8081 ZembDS 0.8781 F5DS 0.8666 F7DS 0.8666 F7DS 0.8666 OrgDS 0.8078 ZembDS 0.8744 F5DS 0.8593 OrgDS 0.7647 F5DS 0.8253 OrgDS 0.7647 ZembDS 0.8617 F5DS 0.8494 F7DS 0.8253 OrgDS 0.7647 ZembDS 0.8714 F5DS 0.8494 F7DS 0.8253 OrgDS 0.7881 ZembDS 0.8714 F5DS 0.8666 F7DS 0.8666 F7DS 0.8666 F7DS 0.8593 OrgDS 0.7881 ZembDS 0.8714 F5DS 0.8606 F7DS 0.8593 OrgDS 0.7881 ZembDS 0.8714 F5DS 0.8597 OrgDS 0.8813 F7DS 0.8597 OrgDS 0.8128 ZembDS 0.8772 F5DS 0.8731 F7DS 0.8722 OrgDS 0.8083 ZembDS 0.8772 F5DS 0.8731 F7DS 0.8722 OrgDS 0.8083 ZembDS 0.8776 F5DS 0.8769			8000 (all)		
TFIDF 2000 F5DS 0.7411 2embDS 0.8569 F5DS 0.8372 F7DS 0.8344 OrgDS 0.7656 ZembDS 0.8669 F5DS 0.8533 F7DS 0.8517 OrgDS 0.7839 ZembDS 0.8678 F5DS 0.8581 F7DS 0.86686 F7DS 0.8686 F7DS 0.8081 ZembDS 0.8781 F5DS 0.8686 F7DS 0.8686 F7DS 0.86686 F7DS 0.86686 F7DS 0.8650 OrgDS 0.8078 ZembDDS 0.8764 F5DS 0.8650 OrgDS 0.7647 F5DS 0.8294 F7DS 0.8253 OrgDS 0.7647 ZembDS 0.8714 F5DS 0.8488 OrgDS 0.7881 F5DS 0.8494 F7DS 0.8253 OrgDS 0.7647 ZembDS 0.8714 F5DS 0.8488 OrgDS 0.7881 F5DS 0.8494 F7DS 0.8253 OrgDS 0.7881 F5DS 0.8494 F7DS 0.8253 OrgDS 0.7881 F5DS 0.8494 F7DS 0.8597 F7DS 0.8772 F7					
TFIDF 2000 F5DS 0.8372 F7DS 0.8374 P5DS 0.8569 P5DS 0.8372 F7DS 0.8344 P5DS 0.8551 OrgDS 0.8666 P5DS 0.8531 P7DS 0.8517 OrgDS 0.8581 P7DS 0.8581 P7DS 0.8581 P7DS 0.8581 P7DS 0.8581 P7DS 0.8666 P5DS 0.8666 P5DS 0.8666 P7DS 0.8666 P7DS 0.8666 P7DS 0.8666 P7DS 0.8665 OrgDS 0.8078 ZembDS 0.8775 P7DS 0.8650 OrgDS 0.8078 ZembDS 0.8765 OrgDS 0.8078 ZembDS 0.8765 OrgDS 0.8081 ZembDS 0.8744 P5DS 0.8253 OrgDS 0.7049 P7DS 0.8253 OrgDS 0.7049 P7DS 0.8253 OrgDS 0.7049 P7DS 0.8253 OrgDS 0.7049 P7DS 0.8253 OrgDS 0.7047 ZembDS 0.8714 P5DS 0.8458 OrgDS 0.7049 P7DS 0.8253 OrgDS 0.7049 P7DS 0.8597 OrgDS 0.8597 OrgDS 0.8714 P5DS 0.8697 OrgDS 0.8712 ZembDS 0.8772 P5DS 0.8731 P7DS 0.8772 P5DS 0.8772 P5DS 0.8772 OrgDS 0.8772 P5DS 0.8772 P5DS 0.8772 P5DS 0.8772 P5DS 0.8772 P5DS 0.87676 P5DS 0.87676					
TFIDF 2000 F5DS 0.8372 F7DS 0.8344 OrgDS 0.7656 ZembDS 0.8669 F5DS 0.8533 F7DS 0.85833 F7DS 0.8581 F7DS 0.8586 F5DS 0.8666 F5DS 0.8666 F5DS 0.8765 OrgDS 0.8078 ZembDS 0.8775 F7DS 0.8650 OrgDS 0.8078 ZembDS 0.8764 F5DS 0.8758 F7DS 0.8650 OrgDS 0.7219 ZembDS 0.8458 F7DS 0.8294 F7DS 0.8253 OrgDS 0.7647 F5DS 0.8458 OrgDS 0.7647 F5DS 0.8458 P7DS 0.8458 OrgDS 0.7814 F5DS 0.8458 OrgDS 0.7814 F5DS 0.8458 OrgDS 0.7814 F5DS 0.8458 OrgDS 0.7814 F5DS 0.8458 OrgDS 0.8714 F5DS 0.8458 OrgDS 0.8714 F5DS 0.8597 OrgDS 0.8813 F7DS 0.8597 OrgDS 0.8813 F7DS 0.8597 OrgDS 0.8714 F5DS 0.8597 OrgDS 0.8714 F5DS 0.8606 F7DS 0.8597 OrgDS 0.8713 F7DS 0.8597 OrgDS 0.8712 OrgDS 0.8712 OrgDS 0.8712 OrgDS 0.8713 F7DS 0.8722 OrgDS 0.8731 F7DS 0.8722 OrgDS 0.8731 F7DS 0.8722 OrgDS 0.8769 OrgDS 0.8769					
TFIDF 2000 Proposition of the pr			500		
TFIDF 2000 ZembDS				F7DS	0.8344
TFIDF 2000 F5DS 0.8533 F7DS 0.8517 OrgDS 0.7839 ZembDS 0.8678 F5DS 0.8581 F7DS 0.8581 F7DS 0.8581 F7DS 0.8581 F7DS 0.8581 F7DS 0.8581 F7DS 0.8666 F7DS 0.8666 F7DS 0.8666 F7DS 0.8666 F7DS 0.8666 F7DS 0.86650 OrgDS 0.8078 ZembDS 0.8764 F5DS 0.8650 OrgDS 0.7647 F7DS 0.8253 OrgDS 0.7647 F7DS 0.8253 OrgDS 0.7647 F7DS 0.8253 OrgDS 0.8017 F5DS 0.8494 F7DS 0.8253 OrgDS 0.7647 F5DS 0.8494 F7DS 0.8253 OrgDS 0.7647 F5DS 0.8666 F7DS 0.8617 F5DS 0.8666 F7DS 0.8813 ZembDS 0.8714 F5DS 0.8606 F7DS 0.8597 OrgDS 0.8128 ZembDS 0.8772 F7DS 0.8597 OrgDS 0.8128 ZembDDS 0.8772 F7DS 0.8731 F7DS 0.8722 OrgDS 0.8083 ZembDS 0.8772 F7DS 0.8722 OrgDS 0.8083 ZembDS 0.8766 F5DS 0.8769				OrgDS	0.7656
TFIDF 2000 F7DS 0.8513 F7DS 0.8518 OrgDS 0.7839 ZembDS 0.8678 F5DS 0.8581 OrgDS 0.8081 ZembDS 0.8686 F7DS 0.8686 F7DS 0.8686 F7DS 0.8686 F7DS 0.8686 F7DS 0.8650 OrgDS 0.8775 F7DS 0.8650 OrgDS 0.7219 ZembDS 0.7219 ZembDS 0.7219 ZembDS 0.8764 F5DS 0.8294 F7DS 0.8253 OrgDS 0.7647 F7DS 0.8650 OrgDS 0.7647 F7DS 0.8253 OrgDS 0.7647 F7DS 0.8253 OrgDS 0.7647 F7DS 0.85597 OrgDS 0.8714 F5DS 0.8458 OrgDS 0.8714 F5DS 0.8597 OrgDS 0.8597 OrgDS 0.8597 OrgDS 0.8597 OrgDS 0.8714 F7DS 0.8597 OrgDS 0.8718 F7DS 0.8597 OrgDS 0.8718 F7DS 0.8708 D772 F7DS 0.8772 F7DS 0.8772 F7DS 0.8772 F7DS 0.8772 F7DS 0.8772 F7DS 0.8772 OrgDS 0.8786 P7DS 0.8769			1000	ZembDS	0.8669
TFIDF 2000 OrgDS 0.7839 ZembDS 0.8678 F5DS 0.8581 OrgDS 0.8081 ZembDS 0.8686 F7DS 0.8686 F7DS 0.8686 F7DS 0.8656 OrgDS 0.8078 ZembDS 0.8775 F7DS 0.8650 OrgDS 0.7219 ZembDS 0.7219 ZembDS 0.7219 ZembDS 0.8764 F5DS 0.8294 F7DS 0.8253 OrgDS 0.7647 ZembDS 0.8617 F5DS 0.8458 OrgDS 0.7647 F5DS 0.8597 OrgDS 0.8714 F5DS 0.8597 OrgDS 0.8714 F5DS 0.8597 OrgDS 0.8712 Or			1000	F5DS	0.8533
TFIDF 2000 ZembDS 0.8678 F5DS 0.8581 F7DS 0.8581 F7DS 0.8581 F7DS 0.8581 F7DS 0.8581 F5DS 0.8686 F7DS 0.8666 OrgDS 0.8078 ZembDS 0.8775 F7DS 0.8650 OrgDS 0.8775 F7DS 0.8650 OrgDS 0.7219 ZembDDS 0.8744 F5DS 0.8294 F7DS 0.8253 OrgDS 0.7647 ZembDS 0.8617 F5DS 0.8458 P7DS 0.8458 OrgDS 0.7804 ZembDS 0.8714 F5DS 0.8458 OrgDS 0.7881 ZembDS 0.8714 F5DS 0.8606 F7DS 0.8458 ZembDS 0.8714 F5DS 0.8606 F7DS 0.8597 OrgDS 0.8128 ZembDS 0.8772 F5DS 0.8731 F5DS 0.8731 F5DS 0.8702 OrgDS 0.8712 ZembDDS 0.8772 F5DS 0.8731 F5DS 0.8731 F5DS 0.8731 F5DS 0.8732 OrgDS 0.8722 OrgDS 0.8732 ZembDS 0.8772 F5DS 0.8731 F5DS 0.8732 OrgDS 0.8732 OrgDS 0.8732 OrgDS 0.8736 F5DS 0.8769					
SCOOL F5DS 0.8581 F7DS 0.8581 F7DS 0.8581 OrgDS 0.8081 ZembDS 0.8781 F5DS 0.86686 F7DS 0.86686 F7DS 0.8656 OrgDS 0.80764 F5DS 0.8650 OrgDS 0.8764 F5DS 0.8650 OrgDS 0.7219 ZembDS 0.8253 OrgDS 0.8253 OrgDS 0.8253 OrgDS 0.8261 F5DS 0.8641 F5DS 0.8641 F5DS 0.8641 F5DS 0.8641 F5DS 0.8641 F5DS 0.8641 F5DS 0.8606 F7DS 0.8591 OrgDS 0.7881 ZembDS 0.8714 F5DS 0.8606 F7DS 0.8597 OrgDS 0.8591 F7DS 0.8597 OrgDS 0.8772 F5DS 0.8772 F5DS 0.8772 F5DS 0.8722 OrgDS 0.8083 ZembDS 0.8766 OrgDS 0.8					
SOURCE S		TEIDE	2000		
SOUD		11101	2000		
SOUD SembDS 0.8781					
SCOO F5DS					
SOUD PTDS 0.8656 OrgDS 0.8078 ZembDS 0.8764 P5DS 0.8775 P7DS 0.8650 OrgDS 0.7219 OrgDS 0.7219 OrgDS 0.8294 P7DS 0.8253 OrgDS 0.7647 P5DS 0.8454 P5DS 0.8454 P5DS 0.8454 P5DS 0.8458 OrgDS 0.7647 P5DS 0.8458 OrgDS 0.7647 P5DS 0.85697 OrgDS 0.8714 P5DS 0.8597 OrgDS 0.8597 OrgDS 0.8597 OrgDS 0.8712 OrgDS 0.8712 OrgDS 0.8712 OrgDS 0.8712 OrgDS 0.8712 OrgDS 0.8722 OrgDS 0.8769 OrgDS 0.8769			5000		
8000 (all) OrgDS					
8000 (all) ZembDS 0.8764 F5DS 0.8775 F7DS 0.8650			<u> </u>		
BOW (all) F5DS 0.8775 F7DS 0.8650 O7gDS 0.7219 P5DS 0.8253 O7gDS 0.7647 P5DS 0.8253 O7gDS 0.7647 P5DS 0.8253 O7gDS 0.7647 P5DS 0.8458 O7gDS 0.7647 P5DS 0.8458 O7gDS 0.7687 P5DS 0.8458 P5DS 0.8597 O7gDS 0.8597 O7gDS 0.8597 O7gDS 0.8597 O7gDS 0.8772 P5DS 0.8772 P5DS 0.8772 P5DS 0.8772 P5DS 0.8772 P5DS 0.8772 P5DS 0.8782 O7gDS 0.8782 O7gDS 0.8782 O7gDS 0.8769 P5DS 0.8769			1		
F7DS			8000 (all)		
BOW 2000 F7DS 0.7219 BOW 2000 F7DS 0.8294 BOW 2000 F7DS 0.8444 F5DS 0.8294 F7DS 0.8253 OrgDS 0.7647 ZembDS 0.8617 F5DS 0.8494 F7DS 0.8458 OrgDS 0.7881 ZembDS 0.7812 ZembDS 0.8714 F5DS 0.8597 OrgDS 0.8128 ZembDS 0.8731 F7DS 0.8731 F7DS 0.8722 OrgDS 0.8083 ZembDS 0.8769					
BOW 2000 F5DS 0.8606 F7DS 0.8597 OrgDS 0.8128 ZembDS 0.8772 F5DS 0.8731 F7DS 0.8722 OrgDS 0.803 8000 (all) ZembDS 0.8767 F5DS 0.8767 F5DS 0.8767	Ğ				
BOW 2000 F5DS 0.8606 F7DS 0.8597 OrgDS 0.8128 ZembDS 0.8772 F5DS 0.8731 F7DS 0.8722 OrgDS 0.803 8000 (all) ZembDS 0.8767 F5DS 0.8767 F5DS 0.8767	l K		F00		
BOW 2000 F5DS 0.8606 F7DS 0.8597 OrgDS 0.8128 ZembDS 0.8772 F5DS 0.8731 F7DS 0.8722 OrgDS 0.803 8000 (all) ZembDS 0.8767 F5DS 0.8767 F5DS 0.8767	Ž		500		
BOW 2000 F5DS 0.8606 F7DS 0.8597 OrgDS 0.8128 ZembDS 0.8772 F5DS 0.8731 F7DS 0.8722 OrgDS 0.803 8000 (all) ZembDS 0.8767 F5DS 0.8767 F5DS 0.8767	J.				0.8253
BOW 2000 F5DS 0.8606 F7DS 0.8597 OrgDS 0.8128 ZembDS 0.8772 F5DS 0.8731 F7DS 0.8722 OrgDS 0.803 8000 (all) ZembDS 0.8767 F5DS 0.8767 F5DS 0.8767	l Ĕ				
BOW 2000 F5DS 0.8606 F7DS 0.8597 OrgDS 0.8128 ZembDS 0.8772 F5DS 0.8731 F7DS 0.8722 OrgDS 0.803 8000 (all) ZembDS 0.8767 F5DS 0.8767 F5DS 0.8767	Ě		1000		
BOW 2000 F5DS 0.8606 F7DS 0.8597 OrgDS 0.8128 ZembDS 0.8772 F5DS 0.8731 F7DS 0.8722 OrgDS 0.803 8000 (all) ZembDS 0.8767 F5DS 0.8767 F5DS 0.8767	75		1000		
BOW 2000 F5DS 0.8606 F7DS 0.8597 OrgDS 0.8128 ZembDS 0.8772 F5DS 0.8731 F7DS 0.8722 OrgDS 0.803 8000 (all) ZembDS 0.8767 F5DS 0.8767 F5DS 0.8767	Įμ				
BOW 2000 F5DS 0.8606 F7DS 0.8597 OrgDS 0.8128 ZembDS 0.8772 F5DS 0.8731 F7DS 0.8722 OrgDS 0.803 8000 (all) ZembDS 0.8767 F5DS 0.8767 F5DS 0.8767	Ę				
F5DS 0.8606 F7DS 0.8597 OrgDS 0.8128 ZembDS 0.8772 F5DS 0.8731 F7DS 0.8731 F7DS 0.8722 OrgDS 0.8083 ZembDS 0.8766 8000 (all) 2 mbDS 0.8769		BOW	2000		
5000 PS 0.8128 ZembDS 0.8772 F5DS 0.8731 F7DS 0.8722 OrgDS 0.8083 ZembDS 0.8769 8000 (all) S000 0.8769					
5000 ZembDS 0.8772 F5DS 0.8731 F7DS 0.8732 OrgDS 0.8083 ZembDS 0.8767 F5DS 0.8769					
S000 F5DS					
F5DS 0.8731 F7DS 0.8722 OrgDS 0.8083 ZembDS 0.8767 F5DS 0.8769			5000		
8000 (all) OrgDS 0.8083 ZembDS 0.8767 F5DS 0.8769					
8000 (all) ZembDS 0.8767 F5DS 0.8769					
8000 (all) F5DS 0.8769					
13D3 0:8709			8000 (all)		
			L	F/D3	0.0723

7.3 SVM

SOUD Dataset Score Cypp Sc					
THIDF 1000 2cmbDS 0.9226 F5DS 0.9221 0rgDS 0.9211 0rgDS 0.9211 0rgDS 0.9361 F5DS 0.9361 F5DS 0.9361 F5DS 0.9361 F7DS 0.9366 0rgDS 0.9300 2cmbDS 0.9442 F5DS 0.9438 F7DS 0.9438 F7DS 0.9458 0rgDS 0.9458 0rgDS 0.9458 0rgDS 0.9458 0rgDS 0.9506 F7DS 0.9483 0rgDS 0.9506 F7DS 0.9483 0rgDS 0.9506 F7DS 0.9483 0rgDS 0.9506 F7DS 0.9483 0rgDS 0.8731 2cmbDS 0.8844 E7DS 0.8902 0rgDS 0.8975 E7DS 0.8903 0rgDS 0.8904 0rgDS 0.8904 0rgDS 0.8904 0rgDS 0.9094 E7DS 0.9095 E7DS 0.9005 0.9005 E7DS 0.9005 0.9005 E7DS 0.9005 0.9005 0.9005 E7DS 0.9005 0.9005 E7DS 0.9005 0.9005 0.9005 E7DS 0.9005 0.9005 0.9005 E7DS 0.9005 0.9005 0.9005 E7DS 0.9005		1		Dataset	Score
Prince					
TFIDF 1000			500		
TFIDF 1000 1					
TFIDF 2000 2cmbDS 0,9361 F5DS 0,9361 F7DS 0,9361 F7DS 0,9361 F7DS 0,9361 F7DS 0,9360 C7gDS 0,9300 C7gDS 0,9300 C7gDS 0,9433 F7DS 0,9433 F7DS 0,9438 C7gDS 0,9438 C7gDS 0,9452 F5DS 0,9458 C7gDS 0,9458 C7gDS 0,9458 C7gDS 0,9458 C7gDS 0,9458 C7gDS 0,9389 C7gDS 0,9389 C7gDS 0,9389 C7gDS 0,9389 C7gDS 0,9389 C7gDS 0,9458 C7gDS 0,8444 C7gDS 0,88778 C7gDS 0,8844 C7gDS 0,8817					
TFIDF 2000 F5DS 0.9361 F7DS 0.9336 OrgDS 0.9300 ZembDS 0.9442 F5DS 0.9433 F7DS 0.9386 OrgDS 0.9386 OrgDS 0.9386 OrgDS 0.9389 OrgDS 0.9488 OrgDS 0.9488 OrgDS 0.9488 OrgDS 0.9488 OrgDS 0.9506 F7DS 0.9483 OrgDS 0.9506 F7DS 0.9483 OrgDS 0.8704 OrgDS 0.8871 ZembDS 0.8844 F5DS 0.8849 F7DS 0.8849 F7DS 0.8893 OrgDS 0.8772 ZembDS 0.9011 F5DS 0.8905 F7DS 0.8905 OrgDS 0.8905 OrgDS 0.8905 OrgDS 0.8905 OrgDS 0.8906 OrgDS 0.8906 OrgDS 0.8906 OrgDS 0.8906 OrgDS 0.8906 OrgDS 0.8906 OrgDS 0.9006 OrgDS 0.9006 OrgDS 0.9006 OrgDS 0.9006 OrgDS 0.8906 OrgDS 0.9006 OrgDS 0.8906 OrgDS 0.9006 OrgDS 0.8906					
TFIDF 2000 F5DS 0,9306 0,9300 2,2,2,2,2,3,3,4,2,4,2,4,2,4,2,4,2,4,2,4,			1000		
TFIDF 2000			1000		
TFIDF					
SCHOOL SECTION SECTI					
SOUND FISTON 19386 19386 19386 19386 19386 19386 19386 19386 19386 19386 19386 19386 19386 193888 193888 193888 193888 19388 19388 19388 19388 193888 19388 19388 19388		TEIDE	2000	ZembDS	0.9442
SCHOOL SOURCE S		ITIDE	2000	F5DS	0.9433
SOUND SET SOUND SET			i	F7DS	0.9386
SOUND SET SOUND SET				OrgDS	0.9356
SOUD	i				
SQUO SQUEEN SQU			5000		
SOUTH SOUT					
SOUD SOUD SET SOUD SET SOUD SET SOUD SET SOUD SET SOUD SET				1	
SOUND SOUN					
SOUND SOUN			8000 (all)		
BOW South				F5DS	0.9506
BOW 2000 F7DS 0.8925	S			F7DS	0.9483
BOW 2000 F7DS 0.8925	- 2			OrgDS	0.8731
BOW 2000 F7DS 0.8925	Ş		500	ZembDS	0.8844
BOW 2000 F7DS 0.8925	ΡV		300		
BOW 2000 F7DS 0.8925	Į į		1		
BOW 2000 F7DS 0.8925	SI				
BOW 2000 F7DS 0.8925	Į		1		
BOW 2000 F7DS 0.8925	H		1000		
BOW 2000 CrgDS 0.8806 ZembDS 0.9094 FSDS 0.9094 CrgDS 0.8942 ZembDS 0.9192 FSDS 0.9050 CrgDS 0.9050 CrgDS 0.9050 CrgDS 0.9050 CrgDS 0.9156 CrgDS 0.9156 CrgDS 0.9156 CrgDS 0.9186 F7DS 0.8878 FSDS 0.88772 F7DS 0.8810 CrgDS 0.8806 CrgDS 0.8806 CrgDS 0.8806 CrgDS 0.8903 CrgDS 0.8903 CrgDS 0.8903 CrgDS 0.9147 F7DS 0.9042 CrgDS 0.9217 F7DS 0.9246 CrgDS 0.9217 F7DS 0.9246 CrgDS 0.8961 CrgDS 0.8508 FSDS 0.85	≤				
BOW 2000 ZembDS 0,9094 F5DS 0,9081 F7DS 0,8994 OrgDS 0,8942 ZembDS 0,9192 F5DS 0,9050 ZembDS 0,9050 ZembDS 0,9050 ZembDS 0,9156 ZembDS 0,9208 ZembDS 0,9209 ZembDS 0,9022 ZembDS 0,9022 ZembDS 0,9022 ZembDS 0,9022 ZembDS 0,9022 ZembDS 0,9022 ZembDS 0,9147 ZembDS 0,9267 ZembDS 0,9267 ZembDS 0,9267 ZembDS 0,9267 ZembDS 0,9268 ZembDS 0,9258 ZembDS 0,8508 ZembDS 0,8508 ZembDS 0,8508 ZembDS 0,8508 ZembDS 0,8556 Zemb					
SOW 2000					
SOUD STORE SOUD		BOW	2000		
TFIDF 2000 September Septe					
TFIDF 2000 ZembDS 0.9192 FSDS 0.9122 FSDS 0.9025 0.9050 0.9050 0.9050 0.9050 0.9050 0.9050 0.9050 0.9050 0.9050 0.9050 0.9050 0.9050 0.9050 0.9050 0.9050 0.9050 0.9050 0.9156 0.9156 0.9156 0.9156 0.9156 0.9156 0.9156 0.9156 0.9156 0.9156 0.9156 0.9155 0.9156 0.9156 0.9155 0.9156 0.9155 0.9156 0.9150 0.9					
SOUD					
SOUD COLUMN SOUD CARE SO			5000		
SOUD (all) OrgDS			0000		
SOUD SOUD SENT SOUD SENT SOUD SENT SENT SOUD SENT					
SOUD (all) F5DS					
TFIDF 2000 F5DS 0.9156 OrgDS 0.7881 ZembDS 0.8858 F5DS 0.8772 F7DS 0.8611 OrgDS 0.8306 ZembDDS 0.9022 F5DS 0.8903 OrgDS 0.8903 OrgDS 0.8903 OrgDS 0.8903 OrgDS 0.9147 F7DS 0.9042 OrgDS 0.8944 ZembDS 0.9217 F7DS 0.9206 OrgDS 0.8941 ZembDS 0.9217 F7DS 0.9206 OrgDS 0.8941 ZembDS 0.9217 F7DS 0.9206 OrgDS 0.8931 ZembDS 0.9217 F7DS 0.9206 OrgDS 0.8931 F5DS 0.9218 F5DS 0.9218 F5DS 0.9231 OrgDS 0.8344 OrgDS 0.8344 OrgDS 0.8344 OrgDS 0.8508 F5DS 0.8311 F7DS 0.8508 F5DS 0.8508 F5DS 0.8556 F7DS 0.8508 F5DS 0.8556 F7DS 0.8592 F5DS 0.8608 OrgDS 0.8475 OrgDS 0.8778 OrgDS 0.8778 OrgDS 0.8778 OrgDS 0.8778 OrgDS 0.8778 OrgDS 0.8778 OrgDS 0.8444 F5DS 0.8475 OrgDS 0.8778 OrgDS 0.8444 F5DS 0.8475 OrgDS 0.8778 OrgDS 0.8444 F5DS 0.8841			9000 (211)	ZembDS	0.9208
TFIDF 2000 FSDS 0.7881 2cmbDS 0.8858 FSDS 0.8772 F7DS 0.8611 OrgDS 0.8306 ZembDS 0.9022 F5DS 0.8903 OrgDS 0.8725 ZembDS 0.9912 F7DS 0.9912 OrgDS 0.8903 OrgDS 0.8725 ZembDS 0.9147 F7DS 0.9217 F5DS 0.9217 F5DS 0.9217 F7DS 0.9226 OrgDS 0.8961 ZembDS 0.9267 F5DS 0.9217 F7DS 0.9231 OrgDS 0.8961 ZembDS 0.9278 F5DS 0.9217 F7DS 0.9231 OrgDS 0.8736 F5DS 0.9231 OrgDS 0.8736 F5DS 0.8311 F7DS 0.8311 F7DS 0.8311 F7DS 0.8344 OrgDS 0.7886 ZembDS 0.8508 F5DS 0.8508 F5DS 0.8508 F5DS 0.8556 F7DS 0.8475 OrgDS 0.8608 OrgDS 0.8603 F7DS 0.8608 OrgDS 0.8778 OrgDS 0.8414 ZembDS 0.8778 OrgDS 0.8475 OrgDS 0.8778 OrgDS 0.8475 OrgDS 0.8778 OrgDS 0.8444 ZembDS 0.8878 OrgDS 0.8444 ZembDS 0.8878 OrgDS 0.8444 ZembDS 0.8814 F5DS 0.8814			0000 (all)	F5DS	0.9186
TFIDF 2000 ZembDS 0.8858 F5DS 0.8772 F7DS 0.8611 OrgDS 0.8306 ZembDS 0.9022 F5DS 0.8903 OrgDS 0.8903 OrgDS 0.9169 F5DS 0.9147 F7DS 0.9042 OrgDS 0.8944 ZembDS 0.9217 F7DS 0.9216 OrgDS 0.8944 ZembDS 0.9217 F7DS 0.9216 OrgDS 0.8941 OrgDS 0.8961 ZembDS 0.9217 F7DS 0.9216 OrgDS 0.8961 ZembDS 0.8961 ZembDS 0.8961 ZembDS 0.8961 ZembDS 0.8914 OrgDS 0.7356 ZembDS 0.8508 F5DS 0.8311 F7DS 0.8508 F5DS 0.8314 OrgDS 0.7886 ZembDS 0.8508 F5DS 0.8508 F5DS 0.8514 OrgDS 0.8508 F5DS 0.8514 OrgDS 0.8508 F7DS 0.8508 F7DS 0.8608 OrgDS 0.8608 OrgDS 0.8608 OrgDS 0.8778 OrgDS 0.8778 OrgDS 0.8778 OrgDS 0.8778 OrgDS 0.8444 ZembDS 0.8878 F5DS 0.8878 F5DS 0.8778 OrgDS 0.8444 DembDS 0.8814 F5DS 0.8817 OrgDS 0.8814 OrgDS 0.8814 F5DS 0.8817 OrgDS 0.8814 OrgDS 0				F7DS	0.9156
TFIDF 2000 ZembDS 0.8858 F5DS 0.8772 F7DS 0.8611 OrgDS 0.8306 ZembDS 0.9022 F5DS 0.8903 OrgDS 0.8903 OrgDS 0.9169 F5DS 0.9147 F7DS 0.9042 OrgDS 0.8944 ZembDS 0.9217 F7DS 0.9216 OrgDS 0.8944 ZembDS 0.9217 F7DS 0.9216 OrgDS 0.8941 OrgDS 0.8961 ZembDS 0.9217 F7DS 0.9216 OrgDS 0.8961 ZembDS 0.8961 ZembDS 0.8961 ZembDS 0.8961 ZembDS 0.8914 OrgDS 0.7356 ZembDS 0.8508 F5DS 0.8311 F7DS 0.8508 F5DS 0.8314 OrgDS 0.7886 ZembDS 0.8508 F5DS 0.8508 F5DS 0.8514 OrgDS 0.8508 F5DS 0.8514 OrgDS 0.8508 F7DS 0.8508 F7DS 0.8608 OrgDS 0.8608 OrgDS 0.8608 OrgDS 0.8778 OrgDS 0.8778 OrgDS 0.8778 OrgDS 0.8778 OrgDS 0.8444 ZembDS 0.8878 F5DS 0.8878 F5DS 0.8778 OrgDS 0.8444 DembDS 0.8814 F5DS 0.8817 OrgDS 0.8814 OrgDS 0.8814 F5DS 0.8817 OrgDS 0.8814 OrgDS 0				OrgDS	0.7881
TFIDF 2000 F5DS 0.8772 F7DS 0.8611 0.8611 0.7 0.7 0.800 0.800 0.8012 0.8012 0.8012 0.8012 0.8012 0.8012 0.8012 0.8012 0.8012 0.8012 0.8012 0.8012 0.8012 0.8012 0.9147 0.9206 0.9256					0.8858
TFIDF 1000 F7DS 0.8611 OrgDS 0.8306 0.8306 0.9022 F5DS 0.8902 F5DS 0.8903 0.972 F7DS 0.8903 0.9169 F5DS 0.9147 F7DS 0.9206 0.9267 F5DS 0.9217 F7DS 0.9206 0.9267 F5DS 0.9217 F7DS 0.9206 0.9278 F5DS 0.9218 F5DS 0.9278 F5DS 0.9231 0.928 F5DS 0.9231 0.928 F5DS 0.8311 F7DS 0.8344 0.928 0.8508 F5DS 0.8311 F7DS 0.8344 0.928 0.8508 F5DS 0.8508	i		500		0.8772
TFIDF 2000 CrgDS					
TFIDF 2000 ZembDS 0.9022 F5DS 0.8972 F7DS 0.8903 OrgDS 0.8725 ZembDS 0.9169 F5DS 0.9169 F5DS 0.9042 OrgDS 0.8944 ZembDS 0.9267 F5DS 0.9217 F7DS 0.9206 OrgDS 0.8961 ZembDS 0.9278 F5DS 0.9256 F7DS 0.9256 F7DS 0.9256 F7DS 0.9256 F7DS 0.9256 F7DS 0.9256 F7DS 0.8508 F5DS 0.8341 F7DS 0.8508 F5DS 0.8341 F7DS 0.8508 F5DS 0.8560 F7DS 0.8560 F7DS 0.8560 F7DS 0.8560 F7DS 0.8603 F7DS 0.8603 F7DS 0.8603 F7DS 0.8608 OrgDS 0.8778 OrgDS 0.8778 OrgDS 0.8444 F5DS 0.8778 OrgDS 0.8444 OrgDS 0.8778 OrgDS 0.8778 OrgDS 0.8778 OrgDS 0.8444 OrgDS 0.8814			1		
TFIDF 2000 F5DS 0.8972 F7DS 0.8903 OrgDS 0.8725 ZembDS 0.9169 F5DS 0.9147 F7DS 0.9042 OrgDS 0.8944 ZembDS 0.9267 F5DS 0.9217 F7DS 0.9206 OrgDS 0.8941 ZembDS 0.9278 F5DS 0.9278 F5DS 0.9256 F7DS 0.9231 OrgDS 0.7356 ZembDS 0.8508 F5DS 0.8311 F7DS 0.8344 OrgDS 0.7356 ZembDS 0.8508 F5DS 0.8311 F7DS 0.8508 F5DS 0.8556 F7DS 0.85608 OrgDS 0.8147 ZembDS 0.8603 F7DS 0.8608 OrgDS 0.8778 F5DS 0.8778 OrgDS 0.8444 SembDS 0.8878 F5DS 0.8778 OrgDS 0.8444 F5DS 0.8814 F5DS 0.8814					
TFIDF 2000 F7DS			1000		
TFIDF 2000 OrgDS					
TFIDF 2000					
SOUND STORE STOR			1		
SOUD F7DS		TFIDF	2000		
SOUD					
SOUD					
SOUD					
SOUD			5000		
BOW 2000 (all) OrgDS					
BOW 2000 (all) ZembDS					
BOW 2000 (all) F5DS			1		
BOW 2000 FSDS 0.89231 BOW 2000 FSDS 0.8445 FSDS 0.8475 FSDS 0.8344 OrgDS 0.8344 OrgDS 0.8344 OrgDS 0.8344 OrgDS 0.892 FSDS 0.892 FSDS 0.8475 OrgDS 0.892 FSDS 0.8603 F7DS 0.8608 OrgDS 0.8394 ZembDS 0.8778 FSDS 0.8778 FSDS 0.8778 FSDS 0.8778 FSDS 0.8444 ZembDS 0.8444 ZembDS 0.8444 ZembDS 0.8444 ZembDS 0.8444 ZembDS 0.8444 ZembDS 0.8814			8000 (all)		
BOW 2000 FSDS 0.7356 The column Foliage FSDS FSDS					
BOW 2000 F5DS 0.8603 F7DS 0.8608 OrgDS 0.8394 ZembDS 0.8778 F5DS 0.8736 F7DS 0.8778 OrgDS 0.8444 ZembDS 0.8444 S000 (all) ZembDS 0.8814 F5DS 0.8817	δ				0.000
BOW 2000 F5DS 0.8603 F7DS 0.8608 OrgDS 0.8394 ZembDS 0.8778 F5DS 0.8736 F7DS 0.8778 OrgDS 0.8444 ZembDS 0.8444 S000 (all) ZembDS 0.8814 F5DS 0.8817					
BOW 2000 F5DS 0.8603 F7DS 0.8608 OrgDS 0.8394 ZembDS 0.8778 F5DS 0.8736 F7DS 0.8778 OrgDS 0.8444 ZembDS 0.8444 S000 (all) ZembDS 0.8814 F5DS 0.8817	Ō		500		
BOW 2000 F5DS 0.8603 F7DS 0.8608 OrgDS 0.8394 ZembDS 0.8778 F5DS 0.8736 F7DS 0.8778 OrgDS 0.8444 ZembDS 0.8444 S000 (all) ZembDS 0.8814 F5DS 0.8817	≥		300		0.8311
BOW 2000 F5DS 0.8603 F7DS 0.8608 OrgDS 0.8394 ZembDS 0.8778 F5DS 0.8736 F7DS 0.8778 OrgDS 0.8444 ZembDS 0.8444 S000 (all) ZembDS 0.8814 F5DS 0.8817	l G				0.8344
BOW 2000 F5DS 0.8603 F7DS 0.8608 OrgDS 0.8394 ZembDS 0.8778 F5DS 0.8736 F7DS 0.8778 OrgDS 0.8444 ZembDS 0.8444 S000 (all) ZembDS 0.8814 F5DS 0.8817	Ę			OrgDS	0.7886
BOW 2000 F5DS 0.8603 F7DS 0.8608 OrgDS 0.8394 ZembDS 0.8778 F5DS 0.8736 F7DS 0.8778 OrgDS 0.8444 ZembDS 0.8444 S000 (all) ZembDS 0.8814 F5DS 0.8817	i i		1000	ZembDS	0.8592
BOW 2000 F5DS 0.8603 F7DS 0.8608 OrgDS 0.8394 ZembDS 0.8778 F5DS 0.8736 F7DS 0.8778 OrgDS 0.8444 ZembDS 0.8444 S000 (all) ZembDS 0.8814 F5DS 0.8817		I	1000		0.8556
BOW 2000 F5DS 0.8603 F7DS 0.8608 OrgDS 0.8394 ZembDS 0.8778 F5DS 0.8736 F7DS 0.8778 OrgDS 0.8444 ZembDS 0.8444 S000 (all) ZembDS 0.8814 F5DS 0.8817					
BOW 2000 F5DS 0.8603 F7DS 0.8608 OrgDS 0.8394 ZembDS 0.8778 F5DS 0.8736 F7DS 0.8778 OrgDS 0.8444 ZembDS 0.8444 S000 (all) ZembDS 0.8814 F5DS 0.8817	HOH				
BOW 2000 F5DS 0.8603 F7DS 0.8608 OrgDS 0.8394 ZembDS 0.8778 F5DS 0.8736 F7DS 0.8778 OrgDS 0.8444 ZembDS 0.8444 S000 (all) ZembDS 0.8814 F5DS 0.8817	THO			OrgDS	0.8147
5000 F7DS 0.8608 OrgDS 0.8394 ZembDS 0.8778 F5DS 0.8736 F7DS 0.8736 F7DS 0.8444 OrgDS 0.8414 F5DS 0.8817	WITHOU				
5000 PS 0.8394 ZembDS 0.8778 F5DS 0.8736 F7DS 0.8778 OrgDS 0.8444 ZembDS 0.8814 F5DS 0.8814	WITHOU	BOW	2000	ZembDS	0.8692
5000 ZembDS 0.8778 F5DS 0.8736 F7DS 0.8778 OrgDS 0.8444 ZembDS 0.8814 F5DS 0.8817	WITHOU	BOW	2000	ZembDS F5DS	0.8692 0.8603
5000 F5DS	WITHOU	BOW	2000	ZembDS F5DS F7DS	0.8692 0.8603 0.8608
F5DS 0.8758 F7DS 0.8778 OrgDS 0.8444 ZembDS 0.8814 F5DS 0.8817	WITHOU	BOW	2000	ZembDS F5DS F7DS OrgDS	0.8692 0.8603 0.8608 0.8394
8000 (all) OrgDS 0.8444 ZembDS 0.8814 F5DS 0.8817	MITHO	BOW		ZembDS F5DS F7DS OrgDS ZembDS	0.8692 0.8603 0.8608 0.8394 0.8778
8000 (all) ZembDS	WITHOU	BOW		ZembDS F5DS F7DS OrgDS ZembDS F5DS	0.8692 0.8603 0.8608 0.8394 0.8778 0.8736
8000 (all) F5DS 0.8817	WITHOU	BOW		ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS	0.8692 0.8603 0.8608 0.8394 0.8778 0.8736 0.8778
1505 0.0017	ОНТИ	BOW		ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS	0.8692 0.8603 0.8608 0.8394 0.8778 0.8736 0.8778
F7DS 0.8900	WITHOU	BOW	5000	ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS	0.8692 0.8603 0.8608 0.8394 0.8778 0.8736 0.8778 0.8444
	WITHOU	BOW	5000	ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS	0.8692 0.8603 0.8608 0.8394 0.8778 0.8736 0.8778 0.8444 0.8814

7.4 KNN

SOUNDARD				Dataset	Саоно
TFIDF 1000 2embDS 0.8367 F5DS 0.8261 F7DS 0.8189 OrgDS 0.7744 F5DS 0.7864 F7DS 0.7567 OrgDS 0.7565 OrgDS 0.7565 OrgDS 0.6992 ZembDS 0.6992 ZembDS 0.692 ZembDS 0.692 ZembDS 0.692 ZembDS 0.6936 ZembDS 0.7447 ZembDS 0.6914 ZembDS 0.6922 ZembDS 0.6938 ZembDS 0.6222 ZembDS 0.7838 ZembDS 0.7838 ZembDS 0.7838 ZembDS 0.7838 ZembDS 0.7669 ZembDS 0.7267 ZembDS 0.7286 ZembDS 0.7267 ZembDS 0.7269 ZembDS 0			1		Score 0.7861
TFIDF 1000					
TFIDF 1000			500		
TFIDF 1000 1					
TFIDF 2000 ZembDS 0.7744 F5DS 0.7864 F7DS 0.7567 OrgDS 0.6092 ZembDDS 0.6886 F5DS 0.6783 F7DS 0.6442 OrgDS 0.6986 F5DS 0.6783 F7DS 0.6442 OrgDS 0.7356 OrgDS 0.7356 OrgDS 0.7997 OrgDS 0.8569 F5DS 0.8244 F7DS 0.7997 OrgDS 0.8767 ZembDS 0.8767 ZembDS 0.9006 OrgDS 0.6311 ZembDS 0.7447 F5DS 0.7544 F7DS 0.7544 F7DS 0.7100 OrgDS 0.5311 OrgDS 0.5311 OrgDS 0.5311 OrgDS 0.5744 ZembDS 0.7242 F5DS 0.7339 F7DS 0.6431 OrgDS 0.5744 ZembDS 0.6978 F7DS 0.6431 OrgDS 0.6302 F5DS 0.6441 F7DS 0.5586 OrgDS 0.5386 OrgDS 0.7838 F7DS 0.5386 OrgDS 0.7838 F7DS 0.7799 OrgDS 0.7783 F7DS 0.7790 OrgDS 0.7502 ZembDS 0.5636 F5DS 0.7783 F7DS 0.7790 OrgDS 0.7502 ZembDS 0.5636 F5DS 0.7286 F7DS 0.5503 OrgDS 0.7502 ZembDS 0.5636 F5DS 0.5506 E7DS 0.5506 OrgDS 0.7356 OrgDS 0.7267 F5DS 0.7356 OrgDS 0.7267 F5DS 0.7356 OrgDS 0.4452 OrgDS 0.44522 OrgDS 0.44522 OrgDS 0.44522 OrgDS					
TFIDF 2000 F5DS 0.5042			1000		
TFIDF 2000			1000	F5DS	0.7864
TFIDF 2000 ZembDS 0.6886 F5DS 0.6783 F7DS 0.6442 OrgDS 0.7356 ZembDS 0.8569 F5DS 0.7997 OrgDS 0.7908 0.7908 0.7908 0.7908 0.7908 0.7447 F7DS 0.7908 0.6781 F7DS 0.7447 F5DS 0.7447 F5DS 0.7447 F5DS 0.7447 F5DS 0.7447 F5DS 0.7594 F7DS 0.6808 0rgDS 0.7747 F5DS 0.7740 0.7908 0.7740 0.7908 0.7741 0.7908 0.7908 0.7741 0.7908 0.7908 0.7741 0.7908 0.7739 0.7908 0.7739 0.7739 0.7739 0.7739 0.7739 0.7739 0.7739 0.7739 0.7739 0.77				F7DS	0.7567
SCO					
SOUTH STATE SOUTH SOUT		TEIDE	2000	ZembDS	0.6886
SOUTH SOUT		IIIDI	2000		0.6783
SOUD SEPTION SESSON SE					
SOUD					
SOUTH SOUT			5000		
SOUTH SOUT			3000		
SOUD SOUD SENDS 0.9003 SEDS 0.9004 SEDS 0.9006 OrgDS 0.6311 ZembDS 0.7447 SEDS 0.7594 ZembDS 0.7594 ZembDS 0.7594 ZembDS 0.7594 ZembDS 0.7242 SEDS 0.7319 SEDS 0.6808 OrgDS 0.5744 ZembDS 0.6914 SEDS 0.6200 ZembDS 0.6200 ZembDS 0.6200 ZembDS 0.6200 ZembDS 0.75910 ZembDS 0.6628 SEDS 0.5169 SEDS 0.5169 SEDS 0.5169 SEDS 0.5958 OrgDS 0.3461 ZembDS 0.6506 ZembDS 0.7261 OrgDS 0.6506 ZembDS 0.7261 OrgDS 0.6904 ZembDS 0.7261 OrgDS 0.6904 ZembDS 0.7267 SEDS 0.6980 SEDS 0.7350 OrgDS 0.6904 ZembDS 0.7350 OrgDS 0.6904 ZembDS 0.7350 OrgDS 0.6904 ZembDS 0.7350 OrgDS 0.6904 ZembDS 0.7267 SEDS 0.6910 OrgDS 0.6910 OrgDS 0.6910 OrgDS 0.6910 OrgDS 0.6910 OrgDS 0.6950 OrgDS 0.6950 OrgDS 0.6950 OrgDS 0.6950 OrgDS 0.6056 ZembDS 0.7356 OrgDS 0.6056 ZembDS 0.7356 OrgDS 0.6056 ZembDS 0.7356 OrgDS 0.6056 ZembDS 0.6056 ZembDS 0.5167 SEDS 0.6056 ZembDS					
SOUD (all) F5DS 0.9017 F7DS 0.9006 0.07gDS 0.6311 2.07594 F7DS 0.7594 F7DS 0.6808 0.7319 F7DS 0.6608 0.6978 F7DS 0.6978 F7DS 0.6978 F7DS 0.6392 F7DS 0.6392 F7DS 0.7581 0.720 0.7793 0.6222 F5DS 0.7783 F7DS 0.7783 F7DS 0.7783 F7DS 0.7783 F7DS 0.7786 F7DS 0.7786 F7DS 0.7786 F7DS 0.7789 0.7669 F5DS 0.7669 F5DS 0.5169 F7DS 0.5636 F7DS 0.5636 F7DS 0.5303 0.6444 F5DS 0.6444 F5DS 0.6444 F5DS 0.6444 F5DS 0.6444 F5DS 0.6444 F5DS 0.6506 0.6506 0.6506 0.6506 0.6506 0.6506 0.6506 0.6989 F7DS 0.7267 F5DS 0.6980 0.6980 0.7267 F5DS 0.6980 0.6980 0.6980 0.7267 0.7260 0.					
SCON FEDS 0.9001			8000 (all)	ZembDS	0.9003
BOW 2000 F/DS 0.6808 OrgDS 0.5744 ZembDS 0.6914 FSDS 0.6978 F7DS 0.6978 F7DS 0.6978 F7DS 0.5200 ZembDS 0.6392 FSDS 0.6224 FSDS 0.6244 F7DS 0.5581 OrgDS 0.4853 ZembDS 0.6222 FSDS 0.6222 FSDS 0.6222 FSDS 0.5386 OrgDS 0.7383 F7DS 0.7386 OrgDS 0.7783 F7DS 0.7790 OrgDS 0.7790 OrgDS 0.7790 OrgDS 0.7550 ZembDS 0.5636 FSDS 0.7550 ZembDS 0.5636 FSDS 0.5169 FSDS 0.5636 FSDS 0.5169 FSDS 0.5303 OrgDS 0.4969 ZembDS 0.6506 OrgDS 0.6506 OrgDS 0.6980 FSDS 0.7261 OrgDS 0.6980 FSDS 0.7261 OrgDS 0.6980 FSDS 0.7261 OrgDS 0.6980 FSDS 0.7267 FSDS 0.6980 OrgDS 0.6964 ZembDS 0.7267 FSDS 0.6980 OrgDS 0.6980 OrgDS 0.6980 OrgDS 0.6980 OrgDS 0.6964 ZembDS 0.7267 FSDS 0.6717 FSDS 0.6950 OrgDS 0.6056 OrgDS 0				F5DS	0.9017
BOW 2000 F/DS 0.6808 OrgDS 0.5744 ZembDS 0.6914 FSDS 0.6978 F7DS 0.6978 F7DS 0.6978 F7DS 0.5200 ZembDS 0.6392 FSDS 0.6224 FSDS 0.6244 F7DS 0.5581 OrgDS 0.4853 ZembDS 0.6222 FSDS 0.6222 FSDS 0.6222 FSDS 0.5386 OrgDS 0.7383 F7DS 0.7386 OrgDS 0.7783 F7DS 0.7790 OrgDS 0.7790 OrgDS 0.7790 OrgDS 0.7550 ZembDS 0.5636 FSDS 0.7550 ZembDS 0.5636 FSDS 0.5169 FSDS 0.5636 FSDS 0.5169 FSDS 0.5303 OrgDS 0.4969 ZembDS 0.6506 OrgDS 0.6506 OrgDS 0.6980 FSDS 0.7261 OrgDS 0.6980 FSDS 0.7261 OrgDS 0.6980 FSDS 0.7261 OrgDS 0.6980 FSDS 0.7267 FSDS 0.6980 OrgDS 0.6964 ZembDS 0.7267 FSDS 0.6980 OrgDS 0.6980 OrgDS 0.6980 OrgDS 0.6980 OrgDS 0.6964 ZembDS 0.7267 FSDS 0.6717 FSDS 0.6950 OrgDS 0.6056 OrgDS 0	S			F7DS	0.9006
BOW 2000 F/DS 0.6808 OrgDS 0.5744 ZembDS 0.6914 FSDS 0.6978 F7DS 0.6978 F7DS 0.6978 F7DS 0.5200 ZembDS 0.6392 FSDS 0.6224 FSDS 0.6244 F7DS 0.5581 OrgDS 0.4853 ZembDS 0.6222 FSDS 0.6222 FSDS 0.6222 FSDS 0.5386 OrgDS 0.7383 F7DS 0.7386 OrgDS 0.7783 F7DS 0.7790 OrgDS 0.7790 OrgDS 0.7790 OrgDS 0.7550 ZembDS 0.5636 FSDS 0.7550 ZembDS 0.5636 FSDS 0.5169 FSDS 0.5636 FSDS 0.5169 FSDS 0.5303 OrgDS 0.4969 ZembDS 0.6506 OrgDS 0.6506 OrgDS 0.6980 FSDS 0.7261 OrgDS 0.6980 FSDS 0.7261 OrgDS 0.6980 FSDS 0.7261 OrgDS 0.6980 FSDS 0.7267 FSDS 0.6980 OrgDS 0.6964 ZembDS 0.7267 FSDS 0.6980 OrgDS 0.6980 OrgDS 0.6980 OrgDS 0.6980 OrgDS 0.6964 ZembDS 0.7267 FSDS 0.6717 FSDS 0.6950 OrgDS 0.6056 OrgDS 0	N			OrgDS	
BOW 2000 F/DS 0.6808 OrgDS 0.5744 ZembDS 0.6914 FSDS 0.6978 F7DS 0.6978 F7DS 0.6978 F7DS 0.5200 ZembDS 0.6392 FSDS 0.6224 FSDS 0.6244 F7DS 0.5581 OrgDS 0.4853 ZembDS 0.6222 FSDS 0.6222 FSDS 0.6222 FSDS 0.5386 OrgDS 0.7383 F7DS 0.7386 OrgDS 0.7783 F7DS 0.7790 OrgDS 0.7790 OrgDS 0.7790 OrgDS 0.7550 ZembDS 0.5636 FSDS 0.7550 ZembDS 0.5636 FSDS 0.5169 FSDS 0.5636 FSDS 0.5169 FSDS 0.5303 OrgDS 0.4969 ZembDS 0.6506 OrgDS 0.6506 OrgDS 0.6980 FSDS 0.7261 OrgDS 0.6980 FSDS 0.7261 OrgDS 0.6980 FSDS 0.7261 OrgDS 0.6980 FSDS 0.7267 FSDS 0.6980 OrgDS 0.6964 ZembDS 0.7267 FSDS 0.6980 OrgDS 0.6980 OrgDS 0.6980 OrgDS 0.6980 OrgDS 0.6964 ZembDS 0.7267 FSDS 0.6717 FSDS 0.6950 OrgDS 0.6056 OrgDS 0	₽		F00		
BOW 2000 F/DS 0.6808 OrgDS 0.5744 ZembDS 0.6914 FSDS 0.6978 F7DS 0.6978 F7DS 0.6978 F7DS 0.5200 ZembDS 0.6392 FSDS 0.6224 FSDS 0.6244 F7DS 0.5581 OrgDS 0.4853 ZembDS 0.6222 FSDS 0.6222 FSDS 0.6222 FSDS 0.5386 OrgDS 0.7383 F7DS 0.7386 OrgDS 0.7783 F7DS 0.7790 OrgDS 0.7790 OrgDS 0.7790 OrgDS 0.7550 ZembDS 0.5636 FSDS 0.7550 ZembDS 0.5636 FSDS 0.5169 FSDS 0.5636 FSDS 0.5169 FSDS 0.5303 OrgDS 0.4969 ZembDS 0.6506 OrgDS 0.6506 OrgDS 0.6980 FSDS 0.7261 OrgDS 0.6980 FSDS 0.7261 OrgDS 0.6980 FSDS 0.7261 OrgDS 0.6980 FSDS 0.7267 FSDS 0.6980 OrgDS 0.6964 ZembDS 0.7267 FSDS 0.6980 OrgDS 0.6980 OrgDS 0.6980 OrgDS 0.6980 OrgDS 0.6964 ZembDS 0.7267 FSDS 0.6717 FSDS 0.6950 OrgDS 0.6056 OrgDS 0	PV		500		
BOW 2000 F/DS 0.6808 OrgDS 0.5744 ZembDS 0.6914 FSDS 0.6978 F7DS 0.6978 F7DS 0.6978 F7DS 0.5200 ZembDS 0.6392 FSDS 0.6224 FSDS 0.6244 F7DS 0.5581 OrgDS 0.4853 ZembDS 0.6222 FSDS 0.6222 FSDS 0.6222 FSDS 0.5386 OrgDS 0.7383 F7DS 0.7386 OrgDS 0.7783 F7DS 0.7790 OrgDS 0.7790 OrgDS 0.7790 OrgDS 0.7550 ZembDS 0.5636 FSDS 0.7550 ZembDS 0.5636 FSDS 0.5169 FSDS 0.5636 FSDS 0.5169 FSDS 0.5303 OrgDS 0.4969 ZembDS 0.6506 OrgDS 0.6506 OrgDS 0.6980 FSDS 0.7261 OrgDS 0.6980 FSDS 0.7261 OrgDS 0.6980 FSDS 0.7261 OrgDS 0.6980 FSDS 0.7267 FSDS 0.6980 OrgDS 0.6964 ZembDS 0.7267 FSDS 0.6980 OrgDS 0.6980 OrgDS 0.6980 OrgDS 0.6980 OrgDS 0.6964 ZembDS 0.7267 FSDS 0.6717 FSDS 0.6950 OrgDS 0.6056 OrgDS 0	[2		1	F7DS	
BOW 2000 F/DS 0.6808 OrgDS 0.5744 ZembDS 0.6914 FSDS 0.6978 F7DS 0.6978 F7DS 0.6978 F7DS 0.5200 ZembDS 0.6392 FSDS 0.6224 FSDS 0.6244 F7DS 0.5581 OrgDS 0.4853 ZembDS 0.6222 FSDS 0.6222 FSDS 0.6222 FSDS 0.5386 OrgDS 0.7383 F7DS 0.7386 OrgDS 0.7783 F7DS 0.7790 OrgDS 0.7790 OrgDS 0.7790 OrgDS 0.7550 ZembDS 0.5636 FSDS 0.7550 ZembDS 0.5636 FSDS 0.5169 FSDS 0.5636 FSDS 0.5169 FSDS 0.5303 OrgDS 0.4969 ZembDS 0.6506 OrgDS 0.6506 OrgDS 0.6980 FSDS 0.7261 OrgDS 0.6980 FSDS 0.7261 OrgDS 0.6980 FSDS 0.7261 OrgDS 0.6980 FSDS 0.7267 FSDS 0.6980 OrgDS 0.6964 ZembDS 0.7267 FSDS 0.6980 OrgDS 0.6980 OrgDS 0.6980 OrgDS 0.6980 OrgDS 0.6964 ZembDS 0.7267 FSDS 0.6717 FSDS 0.6950 OrgDS 0.6056 OrgDS 0	S			OrgDS	
BOW 2000 F/DS 0.6808 OrgDS 0.5744 ZembDS 0.6914 FSDS 0.6978 F7DS 0.6978 F7DS 0.6978 F7DS 0.5200 ZembDS 0.6392 FSDS 0.6224 FSDS 0.6244 F7DS 0.5581 OrgDS 0.4853 ZembDS 0.6222 FSDS 0.6222 FSDS 0.6222 FSDS 0.5386 OrgDS 0.7383 F7DS 0.7386 OrgDS 0.7783 F7DS 0.7790 OrgDS 0.7790 OrgDS 0.7790 OrgDS 0.7550 ZembDS 0.5636 FSDS 0.7550 ZembDS 0.5636 FSDS 0.5169 FSDS 0.5636 FSDS 0.5169 FSDS 0.5303 OrgDS 0.4969 ZembDS 0.6506 OrgDS 0.6506 OrgDS 0.6980 FSDS 0.7261 OrgDS 0.6980 FSDS 0.7261 OrgDS 0.6980 FSDS 0.7261 OrgDS 0.6980 FSDS 0.7267 FSDS 0.6980 OrgDS 0.6964 ZembDS 0.7267 FSDS 0.6980 OrgDS 0.6980 OrgDS 0.6980 OrgDS 0.6980 OrgDS 0.6964 ZembDS 0.7267 FSDS 0.6717 FSDS 0.6950 OrgDS 0.6056 OrgDS 0	lΕ		1000		0.7242
BOW 2000 F/DS 0.6808 OrgDS 0.5744 ZembDS 0.6914 FSDS 0.6978 F7DS 0.6978 F7DS 0.6978 F7DS 0.5200 ZembDS 0.6392 FSDS 0.6224 FSDS 0.6244 F7DS 0.5581 OrgDS 0.4853 ZembDS 0.6222 FSDS 0.6222 FSDS 0.6222 FSDS 0.5386 OrgDS 0.7383 F7DS 0.7386 OrgDS 0.7783 F7DS 0.7790 OrgDS 0.7790 OrgDS 0.7790 OrgDS 0.7550 ZembDS 0.5636 FSDS 0.7550 ZembDS 0.5636 FSDS 0.5169 FSDS 0.5636 FSDS 0.5169 FSDS 0.5303 OrgDS 0.4969 ZembDS 0.6506 OrgDS 0.6506 OrgDS 0.6980 FSDS 0.7261 OrgDS 0.6980 FSDS 0.7261 OrgDS 0.6980 FSDS 0.7261 OrgDS 0.6980 FSDS 0.7267 FSDS 0.6980 OrgDS 0.6964 ZembDS 0.7267 FSDS 0.6980 OrgDS 0.6980 OrgDS 0.6980 OrgDS 0.6980 OrgDS 0.6964 ZembDS 0.7267 FSDS 0.6717 FSDS 0.6950 OrgDS 0.6056 OrgDS 0	ΙV		1000		
BOW 2000 ZembDS 0.6914 FSDS 0.6978 F7DS 0.6431 OrgDS 0.5200 ZembDS 0.6392 FSDS 0.6224 FSDS 0.6224 FSDS 0.6224 FSDS 0.6224 FSDS 0.6225 FSDS 0.6225 FSDS 0.6225 FSDS 0.6225 FSDS 0.5386 OrgDS 0.7383 FSDS 0.7783 FSDS 0.7786 FSDS 0.7786 FSDS 0.7786 FSDS 0.7786 FSDS 0.7789 OrgDS 0.7550 ZembDS 0.5636 FSDS 0.5169 FSDS 0.5169 FSDS 0.5303 OrgDS 0.4969 ZembDS 0.8639 FSDS 0.8508 FSDS 0.6506 OrgDS 0.7267 FSDS 0.7267 OrgDS 0.7267 FSDS 0.7267 FSDS 0.6989 FSDS 0.6717 FSDS 0.6980 OrgDS 0.6964 ZembDS 0.6444 FSDS 0.7267 OrgDS 0.7267 FSDS 0.7267 OrgDS 0.7267 FSDS 0.6980 OrgDS 0.6980 OrgDS 0.6980 OrgDS 0.6980 OrgDS 0.6970 OrgDS 0.6950 OrgDS 0.6950 OrgDS 0.6950 OrgDS 0.6950 OrgDS 0.6950 OrgDS 0.4997 OrgDS 0.4991 OrgDS 0.4997 OrgDS 0.4991 OrgDS					
SOW 2000					
TFIDF 2000 Solution	BOW	2000			
SOUND SOUN		BOW	2000		
SOUD SECTION SOUD SECTION SOUD SECTION SEC				F7DS	
SOUD					
SOUD			5000		
SOUD (all)			3000		
SOUD				1	
SOUD (all) F5DS 0.5858 F7DS 0.5386 F7DS 0.7383 CrembDS 0.7919 F5DS 0.7783 F7DS 0.7783 F7DS 0.7906 OrgDS 0.7763 F7DS 0.7906 OrgDS 0.7550 CrembDS 0.6636 F5DS 0.6536 F5DS 0.5166 F5DS 0.6444 F5DS 0.6444 F5DS 0.6536 F7DS 0.8556 OrgDS 0.3461 CrembDS 0.6308 F7DS 0.8508 F7DS 0.8508 F7DS 0.8508 F7DS 0.8556 OrgDS 0.4969 CrembDS 0.7269 F5DS 0.7261 OrgDS 0.7261 OrgDS 0.7261 OrgDS 0.7267 OrgDS 0.6989 F7DS 0.7267 OrgDS 0.6717 F7DS 0.6950 OrgDS 0.6717 F7DS 0.6950 OrgDS 0.6056 CrembDS 0.6167 F5DS 0.6567 OrgDS 0.4997 CrembDS 0.5167 F5DS 0.4997 CrembDS 0.4517 F5DS 0.4512 CrembDS 0.4517 F5DS 0.4522 CrembDS 0.4522 Crem					
SOUND SOUN			8000 (all)		
TFIDF 2000 F5DS 0.7383 ZembDS 0.7919 F5DS 0.7906 OrgDS 0.7772 ZembDS 0.7572 ZembDS 0.7669 F5DS 0.7526 F7DS 0.7739 OrgDS 0.77550 ZembDS 0.5636 F5DS 0.5169 F7DS 0.5958 OrgDS 0.3461 ZembDS 0.6636 F5DS 0.5169 F7DS 0.5958 OrgDS 0.3461 ZembDS 0.6528 F7DS 0.5303 OrgDS 0.4969 ZembDS 0.8639 F5DS 0.6528 F7DS 0.8508 F7DS 0.8506 F7DS 0.8556 OrgDS 0.4969 ZembDS 0.7261 OrgDS 0.7261 OrgDS 0.6964 ZembDS 0.7267 F5DS 0.7360 OrgDS 0.7267 F5DS 0.6989 F7DS 0.7356 OrgDS 0.7356 OrgDS 0.7366 OrgDS 0.7267 F5DS 0.6989 F7DS 0.7356 OrgDS 0.69650 OrgDS 0.7267 F5DS 0.6980 F7DS 0.6980 F7DS 0.7356 OrgDS 0.7567 OrgDS 0.4997 ZembDS 0.5567 OrgDS 0.4997 ZembDS 0.4517 F5DS 0.4517			0000 (411)		
TFIDF 2000 F5DS 0.7919 TFIDF 2000 F5DS 0.7572 ZembDS 0.7572 ZembDS 0.7669 F5DS 0.7286 F5DS 0.7286 F5DS 0.7286 F5DS 0.7550 ZembDS 0.6366 F5DS 0.5636 F5DS 0.5636 F5DS 0.3461 ZembDS 0.3461 ZembDS 0.4444 F5DS 0.6528 F7DS 0.5303 OrgDS 0.4969 ZembDS 0.8508 F7DS 0.8508 F7DS 0.8508 F7DS 0.8508 F7DS 0.8508 F7DS 0.7267 F5DS 0.7267 F5DS 0.6989 F5DS 0.7267 OrgDS 0.6950 OrgDS 0.6950 OrgDS 0.6950 OrgDS 0.6950 OrgDS 0.6957 OrgDS 0.6957 OrgDS 0.6957 OrgDS 0.6967 OrgDS 0.6957 OrgDS 0.6967 OrgDS 0.6957 OrgDS 0.6957 OrgDS 0.6957 OrgDS 0.4957 OrgDS 0.4517 F5DS 0.4522					
TFIDF 1000 F5DS 0.7783 F7DS 0.7790 0.7906 0.7906 0.7906 0.7669 F5DS 0.7286 F7DS 0.7550 0.7695 0.7550 0.7695 0.5636 F5DS 0.5169 F7DS 0.5958 0.6444 F5DS 0.6444 F5DS 0.6303 0.6440 0.6303 0.6405 0.6306 0.6306 0.6306 0.6306 0.6306 0.6306 0.6306 0.6306 0.6306 0.6306 0.7260 0.72					
TFIDF 1000 F7DS 0.7906 0.7572 2.7572 0.7572 2.7572 0.7572 0.7572 0.7586 F7DS 0.7286 F7DS 0.7739 0.7739 0.7550 2.7550 0.5636 F5DS 0.5169 F7DS 0.5303 0.769DS 0.4614 F5DS 0.6528 F7DS 0.6528 F7DS 0.5303 0.769DS 0.4969 0.7590 0.8556 0.769DS 0.8508 F7DS 0.8508 F7DS 0.8508 F7DS 0.8508 F7DS 0.7261 0.7509 0.7261 0.7269 F5DS 0.7261 0.7267 F5DS 0.6989 F7DS 0.7356 0.7356 0.7509			500		
TFIDF 2000 F5DS 0.7572 ZembDS 0.7669 F5DS 0.7286 F5DS 0.7739 OrgDS 0.7550 ZembDS 0.5636 F5DS 0.5636 F5DS 0.5169 F7DS 0.3958 OrgDS 0.4969 ZembDS 0.6528 F7DS 0.6528 F7DS 0.6528 F7DS 0.8508 F7DS 0.8508 F7DS 0.8508 F7DS 0.8508 F7DS 0.7267 OrgDS 0.69964 ZembDS 0.7267 F5DS 0.7360 OrgDS 0.69964 ZembDS 0.7267 F5DS 0.7366 OrgDS 0.6995 OrgDS 0.6995 OrgDS 0.6995 OrgDS 0.6528 F7DS 0.7550 OrgDS 0.6995 OrgDS 0.6995 OrgDS 0.6950 OrgDS 0.6950 OrgDS 0.6950 OrgDS 0.6957 OrgDS 0.4957 OrgDS 0.4517 F5DS 0.4522					
TFIDF 2000 ZembDS 0.7669 F5DS 0.7739 0.7286 F7DS 0.7739 0.7550 ZembDS 0.5636 F5DS 0.5636 F5DS 0.5958 0.5958 0.6444 F5DS 0.6528 F7DS 0.5303 0.6444 F5DS 0.6506 ZembDS 0.4969 ZembDS 0.4969 0.4969 0.8639 F5DS 0.8508 F7DS 0.8556 0.7261					
TFIDF 2000 F5DS 0.7286 F7DS 0.7739 OrgDS 0.7759 OrgDS 0.5169 F5DS 0.5169 F5DS 0.5169 F5DS 0.5169 F5DS 0.5303 OrgDS 0.4969 F5DS 0.8508 F7DS 0.8508 F7DS 0.8508 F7DS 0.8508 F7DS 0.8508 F7DS 0.8508 F7DS 0.8508 F5DS 0.6506 OrgDS 0.7261 OrgDS 0.6989 F5DS 0.6989 F7DS 0.7356 OrgDS 0.6989 F7DS 0.7356 OrgDS 0.6989 F7DS 0.7356 OrgDS 0.6989 F7DS 0.6980 OrgDS 0.6980 OrgDS 0.6056 OrgDS 0.6980 OrgDS 0.6980 OrgDS 0.6980 OrgDS 0.6980 OrgDS 0.6056 OrgDS 0.6980 OrgDS 0.6980 OrgDS 0.6980 OrgDS 0.6980 OrgDS 0.6056 OrgDS 0.4997 OrgDS 0.4517 F5DS 0.4517 F5DS 0.4517					
TFIDF 2000 F7DS 0.7739 OrgDS 0.7550 ZembDS 0.5636 F5DS 0.5169 F7DS 0.5958 OrgDS 0.6528 F7DS 0.5303 OrgDS 0.4946 ZembDS 0.8639 F5DS 0.8556 OrgDS 0.8556 OrgDS 0.6508 F7DS 0.8556 OrgDS 0.6508 F7DS 0.8556 OrgDS 0.6950 OrgDS 0.6964 ZembDS 0.7267 F5DS 0.7261 OrgDS 0.6989 F7DS 0.7356 OrgDS 0.6989 F7DS 0.7356 OrgDS 0.6989 F7DS 0.7356 OrgDS 0.6944 F5DS 0.6950 OrgDS 0.6950 Or			1000		
TFIDF 2000 OrgDS					
TFIDF 2000 ZembDS 0.5636 F5DS 0.5169 F5DS 0.5169 F7DS 0.5998 OrgDS 0.3461 ZembDS 0.6444 F5DS 0.6508 F7DS 0.8508 F7DS 0.7261 OrgDS 0.7261 OrgDS 0.7261 OrgDS 0.7267 F5DS 0.6989 F7DS 0.7297 ZembDS 0.6444 F5DS 0.6717 F7DS 0.6989 OrgDS 0.6940 ZembDS 0.6717 F7DS 0.6980 OrgDS 0.6717 F7DS 0.6950 OrgDS 0.6056 ZembDS 0.5167 OrgDS 0.5567 OrgDS 0.4997 ZembDS 0.4997 ZembDS 0.4997 ZembDS 0.4997 ZembDS 0.4997 ZembDS 0.4517 F5DS 0.4517 F5DS 0.4522 ZembDS 0.4517 F5DS 0.4522					
SCOON F5DS 0.5169 F7DS 0.5958 OrgDS 0.3461 ZembDS 0.6444 F5DS 0.6528 F7DS 0.6528 F7DS 0.6528 F7DS 0.8639 F5DS 0.8556 OrgDS 0.4969 ZembDS 0.7269 F5DS 0.7122 F7DS 0.7261 OrgDS 0.6964 ZembDS 0.7267 F5DS 0.6989 F7DS 0.7356 OrgDS 0.6948 F7DS 0.7356 OrgDS 0.6717 ZembDS 0.6717 F7DS 0.6716 F7DS 0.6716 F7DS 0.6950 OrgDS 0.6056 ZembDS 0.5167 OrgDS 0.5567 OrgDS 0.4997 OrgDS 0.4997 OrgDS 0.4997 OrgDS 0.4997 OrgDS 0.4997 OrgDS 0.4997 OrgDS 0.4517 F5DS 0.4522 OrgDS					
SOUD		TFIDF	2000		
SOUTH SOUT			1		
SOUD				1	
SGOOM F5DS 0.6528 F7DS 0.5303 OrgDS 0.4969 ZembDS 0.8568 F7DS 0.8556 OrgDS 0.6506 ZembDS 0.7269 F5DS 0.7261 OrgDS 0.6964 ZembDS 0.7261 OrgDS 0.6989 F7DS 0.7356 OrgDS 0.6989 F7DS 0.7356 OrgDS 0.6940 ZembDS 0.7267 F5DS 0.6989 F7DS 0.7356 OrgDS 0.6940 ZembDS 0.7267 F5DS 0.6950 OrgDS 0.6056 ZembDS 0.5167 F5DS 0.6950 OrgDS 0.5167 F5DS 0.5567 OrgDS 0.4997 ZembDS 0.4997 ZembDS 0.4997 ZembDS 0.4997 ZembDS 0.4517 F5DS 0.4522			F000		
SOUD (all) F7DS			5000		
8000 (all) OrgDS			1		
BOW 2000 (all) ZembDS				OrgDS	
BOW 2000 F5DS 0.8556 OrgDS 0.6506 ZembDS 0.7269 F5DS 0.7269 F5DS 0.7269 F5DS 0.7269 F5DS 0.7261 OrgDS 0.6964 ZembDS 0.7269 F5DS 0.7356 OrgDS 0.6944 F5DS 0.6444 F5DS 0.6950 OrgDS 0.6056 ZembDS 0.5167 F5DS 0.5064 F7DS 0.5567 OrgDS 0.4997 ZembDS 0.4497 ZembDS 0.4497 ZembDS 0.4517 F5DS 0.4522			8000 (211)		
BOW 2000 F5DS 0.6506 ZembDS 0.7269 F5DS 0.7122 F7DS 0.7261 OrgDS 0.6964 ZembDS 0.7267 F5DS 0.7356 OrgDS 0.6964 ZembDS 0.7267 F5DS 0.7356 OrgDS 0.6989 F7DS 0.7356 OrgDS 0.6444 F5DS 0.6444 F5DS 0.6950 OrgDS 0.6950 OrgDS 0.6056 ZembDS 0.5167 F5DS 0.5064 F7DS 0.5567 OrgDS 0.4997 ZembDS 0.4497 ZembDS 0.4517 F5DS 0.4522			oud (all)		
Second	S				
Solution				OrgDS	
BOW 2000 F5DS 0.6717 F7DS 0.6950 OrgDS 0.6056 ZembDS 0.5167 F5DS 0.5064 F7DS 0.5567 OrgDS 0.4997 ZembDS 0.4517 F5DS 0.4522	1 14		1	ZembDS	
BOW 2000 F5DS 0.6717 F7DS 0.6950 OrgDS 0.6056 ZembDS 0.5167 F5DS 0.5064 F7DS 0.5567 OrgDS 0.4997 ZembDS 0.4517 F5DS 0.4522	0		500		
BOW 2000 F5DS 0.6717 F7DS 0.6950 OrgDS 0.6056 ZembDS 0.5167 F5DS 0.5064 F7DS 0.5567 OrgDS 0.4997 ZembDS 0.4517 F5DS 0.4522	ΟMα		500	F5DS	
BOW 2000 F5DS 0.6717 F7DS 0.6950 OrgDS 0.6056 ZembDS 0.5167 F5DS 0.5064 F7DS 0.5567 OrgDS 0.4997 ZembDS 0.4517 F5DS 0.4522	OPWO		500	F5DS F7DS	0.7261
BOW 2000 F5DS 0.6717 F7DS 0.6950 OrgDS 0.6056 ZembDS 0.5167 F5DS 0.5064 F7DS 0.5567 OrgDS 0.4997 ZembDS 0.4517 F5DS 0.4522	STOPWO		500	F5DS F7DS OrgDS	0.7261 0.6964
BOW 2000 F5DS 0.6717 F7DS 0.6950 OrgDS 0.6056 ZembDS 0.5167 F5DS 0.5064 F7DS 0.5567 OrgDS 0.4997 ZembDS 0.4517 F5DS 0.4522	JT STOPWO			F5DS F7DS OrgDS ZembDS	0.7261 0.6964 0.7267
BOW 2000 F5DS 0.6717 F7DS 0.6950 OrgDS 0.6056 ZembDS 0.5167 F5DS 0.5064 F7DS 0.5567 OrgDS 0.4997 ZembDS 0.4517 F5DS 0.4522	OUT STOPWO			F5DS F7DS OrgDS ZembDS F5DS	0.7261 0.6964 0.7267 0.6989
BOW 2000 F5DS 0.6717 F7DS 0.6950 OrgDS 0.6056 ZembDS 0.5167 F5DS 0.5064 F7DS 0.5567 OrgDS 0.4997 ZembDS 0.4517 F5DS 0.4522	THOUT STOPWO			F5DS F7DS OrgDS ZembDS F5DS F7DS	0.7261 0.6964 0.7267 0.6989 0.7356
F5DS 0.6971 F7DS 0.6950 OrgDS 0.6056 ZembDS 0.5167 F5DS 0.5064 F7DS 0.5567 OrgDS 0.4997 ZembDS 0.4517 F5DS 0.4512	VITHOUT STOPWO			F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS	0.7261 0.6964 0.7267 0.6989 0.7356 0.7297
OrgDS 0.6056 ZembDS 0.5167 F5DS 0.5064 F7DS 0.5567 OrgDS 0.4997 ZembDS 0.4517 F5DS 0.4522	WITHOUT STOPWO	BOW	1000	F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS	0.7261 0.6964 0.7267 0.6989 0.7356 0.7297 0.6444
ZembDS	WITHOUT STOPWO	BOW	1000	F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS	0.7261 0.6964 0.7267 0.6989 0.7356 0.7297 0.6444 0.6717
S000 F5DS	WITHOUT STOPWO	BOW	1000	F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS	0.7261 0.6964 0.7267 0.6989 0.7356 0.7297 0.6444 0.6717 0.6950
F5DS	WITHOUT STOPWO	BOW	1000	F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS	0.7261 0.6964 0.7267 0.6989 0.7356 0.7297 0.6444 0.6717 0.6950 0.6056
8000 (all) OrgDS 0.4997 ZembDS 0.4517 F5DS 0.4522	WITHOUT STOPWO	BOW	1000	F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS	0.7261 0.6964 0.7267 0.6989 0.7356 0.7297 0.6444 0.6717 0.6950 0.6056 0.5167
8000 (all) ZembDS 0.4517 F5DS 0.4522	WITHOUT STOPWO	BOW	1000	F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS	0.7261 0.6964 0.7267 0.6989 0.7356 0.7297 0.6444 0.6717 0.6950 0.6056 0.5167
8000 (all) F5DS 0.4522	WITHOUT STOPWO	BOW	1000	F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS F7DS	0.7261 0.6964 0.7267 0.6989 0.7356 0.7297 0.6444 0.6717 0.6950 0.6056 0.5167 0.5064
1303 0.4322	WITHOUT STOPWO	BOW	1000	F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS F5DS F5DS F5DS F5DS F5DS F5DS F5	0.7261 0.6964 0.7267 0.6989 0.7356 0.7297 0.6444 0.6717 0.6950 0.6056 0.5167 0.5064 0.5567 0.4997
F/D5 0.429/	WITHOUT STOPWO	BOW	1000 2000 5000	F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS	0.7261 0.6964 0.7267 0.6989 0.7356 0.7297 0.6444 0.6717 0.6950 0.6056 0.5167 0.5064 0.5567 0.4997
	WITHOUT STOPWO	BOW	1000 2000 5000	F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS	0.7261 0.6964 0.7267 0.6989 0.7356 0.7297 0.6444 0.6717 0.6950 0.5064 0.5167 0.5064 0.5567 0.4517 0.4522

7.5 CART

TFIDF South Sout				Dataset	Score
TFIDF 1000					
F5DS 0.8061			500	"	
THIDF 2000 THIDF			300	1	
TFIDF 2000 EmbDS 0.7897 F5DS 0.8044 F7DS 0.7944 OrgDS 0.7472 ZembDS 0.7894 F5DS 0.7894 F5DS 0.7896 F7DS 0.7869 OrgDS 0.7244 ZembDS 0.7806 F5DS 0.7866 F5DS 0.7866 F5DS 0.7867 OrgDS 0.7439 ZembDS 0.7808 F5DS 0.7763 OrgDS 0.7763 OrgDS 0.7763 OrgDS 0.7769 F7DS 0.7769 F7DS 0.7769 F7DS 0.7763 OrgDS 0.7503 ZembDS 0.7808 F5DS 0.7769 F7DS 0.7769 F7DS 0.7769 F7DS 0.7868 F5DS 0.7768 F7DS 0.7868 F7DS 0.7869 F7DS 0.7869 F7DS 0.7869 F7DS 0.7567 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0					
TFIDF 2000 FSDS 0.8044 F7DS 0.7944 0.79DS 0.7472 0.7944 0.79DS 0.7944 0.79DS 0.7945 0.7905 0.7806 0.79DS 0.7806 0.79DS 0.7806 0.7805 0.7806 0.7805 0.7805 0.7806 0.7805 0.					
TFIDF 2000			1000		
TFIDF 2000					
SCOON					0.7472
SOUND STATE STAT		TFIDF	2000		
SOUTH SOUT			2000		
SOUD SOUD SEMBLES CONTROL				1	
SOUD			E000		
SOUD			3000		
SOUD					
SOUD (all)					
BOW South			8000 (all)		
BOW 2000 F7DS 0.7853 0.7853 0.7794 F5DS 0.7794 F5DS 0.7794 F5DS 0.7795 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.7815	28			F7DS	0.7753
BOW 2000 F7DS 0.7853 0.7853 0.7794 F5DS 0.7794 F5DS 0.7794 F5DS 0.7795 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.7815	N N				
BOW 2000 F7DS 0.7853 0.7853 0.7794 F5DS 0.7794 F5DS 0.7794 F5DS 0.7795 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.7815	M		500		
BOW 2000 F7DS 0.7853 0.7853 0.7794 F5DS 0.7794 F5DS 0.7794 F5DS 0.7795 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.7815	P				
BOW 2000 F7DS 0.7853 0.7853 0.7794 F5DS 0.7794 F5DS 0.7794 F5DS 0.7795 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.7815	ST				
BOW 2000 F7DS 0.7853 0.7853 0.7794 F5DS 0.7794 F5DS 0.7794 F5DS 0.7795 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.77850 0.7815	Ξ		1000		
BOW 2000 CrgDS 0.7592 ZembDS 0.7794 F5DS 0.7783 F7DS 0.7850 0.7850 0.7850 0.7850 0.7860 0.7860 0.7761 F7DS 0.7760 0.7803 F5DS 0.7761 F7DS 0.7803 F5DS 0.7811 0.7811 0.7815 0.7811 0.7815 0.7815 0.7815 0.7815 0.7815 0.7815 0.7815 0.7815 0.7693 0.7425 0.7505 0.7425 0.7505 0.7425 0.7505 0.7425 0.7505 0.7425 0.7505 0.7425 0.7505 0.7425 0.7505 0	MI		1000	1	
BOW 2000 ZembDS 0.7794 F5DS 0.7785 0.7785 0.7850 0.7504 ZembDS 0.7504 ZembDS 0.7803 F5DS 0.7761 0.7805 0.7803 F5DS 0.7769 0.7805 0.7801 0.7805 0.7801 0.7805 0.7801 0.7805 0.7801 0.7805 0.7801 0.7805 0.7801 0.7805 0.7801 0.7805 0.7801 0.7805 0.7801 0.7805 0.7801 0.7805 0.7801 0.7805 0.7811 0.7805 0.7425 0.7905 0.7425 0.7905 0.7425 0.7505 0.7425 0.7505 0.7425 0.7505 0.7425 0.7505 0.7669 0.7505 0.7669 0.7505 0.7669 0.7505 0.7669 0.7505 0.7669 0.7505 0.7669 0.7505 0.7669 0.7505 0.7669 0.7505 0.7660 0.7	_			4	
SOW 2000					
SOUD F7DS 0.7850 0.7544 2.7840 0.7544 2.7840 0.7544 2.7840 0.7540 0.7540 0.7540 0.7690 0.7690 0.7690 0.7811 0.7690 0.7690 0.7425 0.7553 0.7550 0.7550 0.7550 0.7550 0.7669 0.7669 0.7669 0.7669 0.7550 0.7550 0.7550 0.7550 0.7550 0.7550 0.7550 0.7550 0.7669 0.7550 0.755		BOW	2000		
SOUD					
S000					
SOUD (all)			5000		
SOUD CASE SOUD					
Semble					
SOUND STATE STOPS STATE STOPS STATE STOPS STATE STOPS STATE STOPS STATE STAT			9000 (-11)		
TFIDF 2000 TFIDS 0.7114 2embDS 0.7633 F5DS 0.7425 F7DS 0.7558 0.			8000 (all)		
TFIDF 2000 ZembDS 0.7653 F5DS 0.7256 0.7255 0.7250				1	
TFIDF 1000 F5DS 0.7425 F7DS 0.7558 0.7250 2embDS 0.7269 F5DS 0.7436 F7DS 0.7366 F7DS 0.7366 F7DS 0.7366 E7DS 0.7347 F7DS 0.7567 0.7567 0.7567 0.7567 0.7567 0.7567 0.7567 0.7567 0.7567 0.7567 0.7567 0.7567 0.7567 0.7567 0.7567 0.7567 0.7568 0.7534 F5DS 0.7347 F7DS 0.7606 0.7508					
TFIDF 2000 CrgDS 0.7250 CrembDS 0.7669 F5DS 0.7436 F7DS 0.7567 OrgDS 0.7356 CrembDS 0.7569 CrembDS 0.7569 CrembDS 0.7569 CrembDS 0.7536 CrembDS 0.7539 CrembDS 0.7489 CrembDS 0.7533 CrembDS 0.7489 CrembDS 0.7533 CrembDS 0.7535 Cre			500		
TFIDF 2000 ZembDS 0.7669 F5DS 0.7436 F7DS 0.7567 OrgDS 0.7356 ZembDS 0.7669 F5DS 0.7695 OrgDS 0.7356 ZembDS 0.7525 OrgDS 0.7572 OrgDS 0.7572 OrgDS 0.7525 ZembDS 0.7534 F5DS 0.7606 OrgDS 0.7060 OrgDS 0.7060 OrgDS 0.7256 F7DS 0.7606 OrgDS 0.7256 F7DS 0.7478 OrgDS 0.7489 F5DS 0.7590 OrgDS 0.7530 F5DS 0.7590 OrgDS 0.7530 F5DS 0.7590 OrgDS 0.7530					
TFIDF 2000 F5DS 0.7436 F7DS 0.7567 OrgDS 0.7356 ZembDS 0.7669 F5DS 0.7442 F7DS 0.7669 F5DS 0.7442 F7DS 0.7572 OrgDS 0.7225 ZembDS 0.7536 F5DS 0.7347 F7DS 0.7606 OrgDS 0.7106 ZembDS 0.7539 F5DS 0.7256 F7DS 0.7408 P5DS 0.7478 OrgDS 0.7478 OrgDS 0.7478 OrgDS 0.7499 F7DS 0.7500 OrgDS 0.7136 ZembDS 0.7539 F5DS 0.7459 F7DS 0.7617 OrgDS 0.7531 F5DS 0.7459 F7DS 0.7617 OrgDS 0.7303 ZembDS 0.7544 F5DS 0.7417 F7DS 0.7661 OrgDS 0.7303 ZembDS 0.7530 F5DS 0.7417 F7DS 0.7661 OrgDS 0.7303 ZembDS 0.7530 F5DS 0.7386 F5DS 0.7386 F5DS 0.7386 F5DS 0.7386 F5DS 0.7583 F5DS 0.7594 OrgDS 0.7539 ZembDS 0.7553 F5DS 0.7386 F5DS 0.7594 OrgDS 0.7537 ZembDS 0.7553 F5DS 0.7584 OrgDS 0.7555 F5DS 0.7594 OrgDS 0.7555 F5DS 0.7587 ZembDS 0.7525 F5DS 0.7587 ZembDS 0.7525 F5DS 0.7587 ZembDS 0.7552 F5DS 0.7587 ZembDS 0.7525 F5					
TFIDF 2000 F7DS 0.7567 OrgDS 0.7356 ZembDS 0.7669 F5DS 0.7442 F7DS 0.7572 OrgDS 0.7252 ZembDS 0.7252 ZembDS 0.7252 ZembDS 0.7347 F7DS 0.7606 OrgDS 0.7106 ZembDS 0.7539 F5DS 0.7478 OrgDS 0.7499 F7DS 0.7500 OrgDS 0.7303 ZembDS 0.7533 F5DS 0.7489 F7DS 0.7617 OrgDS 0.7303 ZembDS 0.7533 F5DS 0.7489 F7DS 0.7661 OrgDS 0.7303 ZembDS 0.7533 F5DS 0.7499 F7DS 0.7661 OrgDS 0.7330 ZembDS 0.7583 F5DS 0.7584 F5DS 0.7583 F5DS 0.7583 F5DS 0.7583 F5DS 0.7583 F5DS 0.7583 F5DS 0.7584 OrgDS 0.7594 OrgDS 0.7594 OrgDS 0.7594 OrgDS 0.7594 OrgDS 0.7595 F5DS 0.7594 OrgDS 0.7595 OrgDS 0.7594 OrgDS 0.7595 OrgDS 0.7583 F5DS 0.7583 F5DS 0.7583 F5DS 0.7583 F5DS 0.7584 OrgDS 0.7595 OrgDS 0.7594 OrgDS 0.7595 OrgDS 0.7595 OrgDS 0.7583 F5DS 0.7583 F5DS 0.7583 F5DS 0.7583 F5DS 0.7583			1000		
TFIDF 2000 OrgDS 0.7356 ZembDS 0.7669 F5DS 0.7442 F7DS 0.7572 OrgDS 0.7525 ZembDS 0.7536 F5DS 0.7347 F7DS 0.7606 OrgDS 0.7160 OrgDS 0.7255 F7DS 0.7256 F7DS 0.7256 F7DS 0.7489 F5DS 0.7489 F5DS 0.7489 F7DS 0.7590 OrgDS 0.7136 ZembDS 0.7539 F7DS 0.7617 OrgDS 0.7330 ZembDS 0.7534 F5DS 0.7439 F7DS 0.7617 OrgDS 0.7330 ZembDS 0.7533 F5DS 0.7489 F7DS 0.7661 OrgDS 0.7330 ZembDS 0.7544 F5DS 0.7417 F7DS 0.7661 OrgDS 0.7350 OrgDS 0.7583 F5DS 0.7584 OrgDS 0.7594 OrgDS 0.77525 OrgDS 0.77397 OrgDS 0.73397				1	
BOW 2000 F5DS 0.7442 F7DS 0.7572					
SOUND SOUN		TEIDE	2000		0.7669
SOUND SOUN		111111	2000		
SOUD ZembDS 0.7534					
SOUO			Food		
SOUD			5000	F5DS	0.7347
BOW 2000 (all) ZembDS 0.7539 F5DS 0.7256 F7DS 0.7478 OrgDS 0.6778 ZembDS 0.7449 F5DS 0.7439 F7DS 0.7500 OrgDS 0.7136 ZembDS 0.7533 F5DS 0.7549 F7DS 0.7617 OrgDS 0.7501 OrgDS 0.7501 OrgDS 0.7501 OrgDS 0.7502 OrgDS 0.7503 F5DS 0.7617 OrgDS 0.7504 F5DS 0.7504 F5DS 0.7504 F5DS 0.7504 OrgDS 0.7350 ZembDS 0.7583 F5DS 0.7583 F5DS 0.7583 F5DS 0.7594 OrgDS 0.7178 ZembDS 0.7525 F5DS 0.7397 OrgDS 0.7178 OrgDS 0.718 OrgDS 0.7					
BOW 2000 (all) F5DS 0.7256 F7DS 0.7478 OrgDS 0.6778 ZembDS 0.7489 F5DS 0.7489 F7DS 0.7489 F7DS 0.7489 F7DS 0.7617 OrgDS 0.7330 ZembDS 0.7504 F5DS 0.7489 F7DS 0.7661 OrgDS 0.7330 ZembDS 0.7548 F5DS 0.7549 F7DS 0.7661 OrgDS 0.7350 ZembDS 0.7583 F5DS 0.7584 OrgDS 0.7594 OrgDS 0.7525 F5DS 0.7397					
SOUND SOUN			8000 (all)		
BOW 2000 F7DS 0.6778 75DS 0.7464 75DS 0.7469 75DS 0.7469 75DS 0.7489 75DS 0.7617 0.750 0.7617 0.7610 0.7617 0.7617 0.7617 0.7617 0.7617 0.7617 0.7617 0.7617 0.7618 0.7617 0.7618	S				
BOW 2000 F5DS 0.7417 F7DS 0.7661 OrgDS 0.7350 ZembDS 0.7583 F5DS 0.7386 F7DS 0.7594 OrgDS 0.7178 ZembDS 0.7525 F5DS 0.7397	RD.			OrgDS	0.6778
BOW 2000 F5DS 0.7417 F7DS 0.7661 OrgDS 0.7350 ZembDS 0.7583 F5DS 0.7386 F7DS 0.7594 OrgDS 0.7178 ZembDS 0.7525 F5DS 0.7397	Į.		500		
BOW 2000 F5DS 0.7417 F7DS 0.7661 OrgDS 0.7350 ZembDS 0.7583 F5DS 0.7386 F7DS 0.7594 OrgDS 0.7178 ZembDS 0.7525 F5DS 0.7397	ΡW				
BOW 2000 F5DS 0.7417 F7DS 0.7661 OrgDS 0.7350 ZembDS 0.7583 F5DS 0.7386 F7DS 0.7594 OrgDS 0.7178 ZembDS 0.7525 F5DS 0.7397) JC				
BOW 2000 F5DS 0.7417 F7DS 0.7661 OrgDS 0.7350 ZembDS 0.7583 F5DS 0.7386 F7DS 0.7594 OrgDS 0.7178 ZembDS 0.7525 F5DS 0.7397	T S		1000	ZembDS	0.7533
BOW 2000 F5DS 0.7417 F7DS 0.7661 OrgDS 0.7350 ZembDS 0.7583 F5DS 0.7386 F7DS 0.7594 OrgDS 0.7178 ZembDS 0.7525 F5DS 0.7397	00		1000		
BOW 2000 F5DS 0.7417 F7DS 0.7661 OrgDS 0.7350 ZembDS 0.7583 F5DS 0.7386 F7DS 0.7594 OrgDS 0.7178 ZembDS 0.7525 F5DS 0.7397	ΙÉ				
BOW 2000 F5DS 0.7417 F7DS 0.7661 OrgDS 0.7350 ZembDS 0.7583 F5DS 0.7386 F7DS 0.7594 OrgDS 0.7178 ZembDS 0.7525 F5DS 0.7397	WI				
F7DS 0.7661 OrgDS 0.7350 ZembDS 0.7583 F5DS 0.7386 F7DS 0.7594 OrgDS 0.7178 ZembDS 0.7525 F5DS 0.7397		BOW	2000		
5000 ZembDS 0.7583 F5DS 0.7386 F7DS 0.7594 OrgDS 0.7178 ZembDS 0.7525 F5DS 0.7397					
5000 F5DS 0.7386 F7DS 0.7594 OrgDS 0.7178 ZembDS 0.7525 F5DS 0.7397					
F7DS 0.7368 F7DS 0.7594 OrgDS 0.7178 ZembDS 0.7525 F5DS 0.7397			5000		
8000 (all) OrgDS 0.7178 ZembDS 0.7525 F5DS 0.7397					
8000 (all) ZembDS 0.7525 F5DS 0.7397					
8000 (all) F5DS 0.7397			8000 (-11)		0.7525
F7DS 0.7589			ouou (aii)		
				F7DS	0.7589

7.6 Rocchio

SOUTH SOUT				Dataset	Saara
TFIDF 500 ZembDS 0.8872 F5DS 0.8700 OrgDS 0.8608 CrembDS 0.8914 F7DS 0.8931 ZembDS 0.9031 E7DS 0.8981 OrgDS 0.8981 OrgDS 0.8981 OrgDS 0.8981 OrgDS 0.8981 OrgDS 0.9092 CrembDS 0.9016 E7DS 0.8021 E7DS 0.8031				Dataset OrgDS	Score 0.8342
### PROPRIES 1000 F7DS 0.870			500		
### PROPRIES 1000 1			500	F5DS	0.8761
TFIDF 2000					
TFIDF 2000 F5DS 0.8944 F7DS 0.8892 0.9025 0.8781 0.9025 F7DS 0.9025 F7DS 0.8981 0.9025 F7DS 0.8981 0.9025 F7DS 0.9026 0.8981 0.9026 0.					
TFIDF 2000			1000		
TFIDE 2000					
Part				1	
SOUTH SOUT		TEIDE	2000		0.9053
SOUTH SOUT		IFIDE	2000		
SOUR					
SOUD					
SOUTH SOUT			5000		
SCOO SCOOL					
SOU (all)					
SOUND FOLS Composition			9000 (211)	ZembDS	0.9106
BOW Depth			8000 (all)	F5DS	0.9106
BOW South	ω			F7DS	0.9111
BOW 2000 F/DS 0.5275	2			OrgDS	
BOW 2000 F/DS 0.5275	Š		500		0.5375
BOW 2000 F/DS 0.5275	ŀΡV		500	F5DS	0.6003
BOW 2000 F/DS 0.5275	T _C			1	
BOW 2000 F/DS 0.5275	HS				
BOW 2000 F/DS 0.5275	ΙĘ		1000		
BOW 2000 CrgDS CystDS CystD	≥				
BOW 2000 ZembDS 0.5450 F5DS 0.6053 F7DS 0.5358 OrgDS 0.4661 ZembDS 0.5472 F5DS 0.6083 F7DS 0.5439 OrgDS 0.4667 ZembDS 0.5439 OrgDS 0.6086 F7DS 0.6086 F7DS 0.6086 F7DS 0.6542 ZembDS 0.8297 F5DS 0.8044 F7DS 0.7808 OrgDS 0.7203 ZembDS 0.8297 F5DS 0.8364 F7DS 0.8581 F5DS 0.8364 F7DS 0.8742 OrgDS 0.7203 ZembDS 0.8789 F7DS 0.8628 F7DS 0.8628 F7DS 0.8628 F7DS 0.8628 F7DS 0.8644 OrgDS 0.8789 OrgDS 0.8278 ZembDS 0.8956 F5DS 0.8814 F7DS 0.8664 OrgDS 0.8333 ZembDS 0.8742 OrgDS 0.8742 OrgDS 0.8742 OrgDS 0.8742 OrgDS 0.8742 OrgDS 0.8742 OrgDS 0.5714 ZembDS 0.7047 F5DS 0.6439 F7DS 0.6618 OrgDS 0.6086 ZembDS 0.7339 F7DS 0.6608 ZembDS 0.7728 F7DS 0.66467 OrgDS 0.6497 Org					
TFIDF 2000 F5DS 0.6035 F7DS 0.5358 OrgDS 0.4661 ZembDS 0.5472 E7DS 0.6088 F7DS 0.5439 OrgDS 0.4667 ZembDS 0.5439 OrgDS 0.6086 F7DS 0.6086 F7DS 0.6086 F7DS 0.6542 ZembDS 0.8297 F5DS 0.8044 F7DS 0.7808 OrgDS 0.7203 ZembDS 0.8298 F5DS 0.8364 F7DS 0.8581 F5DS 0.8364 F7DS 0.8581 F5DS 0.8628 F7DS 0.8628 F7DS 0.8628 F7DS 0.8628 F7DS 0.8628 F7DS 0.8628 F7DS 0.8644 F7DS 0.8644 F7DS 0.8664 OrgDS 0.8278 ZembDS 0.8956 F5DS 0.8814 F7DS 0.8664 OrgDS 0.8333 ZembDS 0.9006 F5DS 0.8628 F7DS 0.6439 F7DS 0.6618 OrgDS 0.5725 ZembDS 0.7339 F7DS 0.6608 ZembDS 0.7583 F7DS 0.6608 ZembDS 0.7583 F7DS 0.66467 OrgDS 0.6497 OrgDS 0.6497 OrgDS 0.6619 OrgDS 0.6619 OrgDS 0.6497 OrgDS 0.6883 F7DS 0.6688 OrgDS 0.6497 OrgDS 0.6883 OrgDS 0.68		DOM	2000		
SOUND SOUN		BOW	2000		
TFIDF 2000 ZembDS 0.5472 F5DS 0.6083 F7DS 0.5439					
SOUD					
SOUD (all)			5000		
SOUD (all)					
SOUD					
TFIDF 2000 (all) F5DS					
TFIDF 2000 S0,06542 ZembDS 0.8297 S0,0614 F7DS 0.7808 OrgDS 0.7203 ZembDS 0.8581 F5DS 0.8364 F7DS 0.8158 OrgDS 0.7781 ZembDS 0.8789 F5DS 0.8628 F7DS 0.8628 F7DS 0.8628 F7DS 0.8422 OrgDS 0.8278 ZembDS 0.8956 F5DS 0.8814 F7DS 0.8664 OrgDS 0.8333 ZembDS 0.9006 F5DS 0.8814 F7DS 0.8664 OrgDS 0.8333 ZembDS 0.9006 F5DS 0.8814 F7DS 0.8664 OrgDS 0.8333 ZembDS 0.9006 F5DS 0.8742 OrgDS 0.5314 ZembDS 0.7047 F5DS 0.6439 F7DS 0.6449 F7DS 0.6449 F7DS 0.6449 F7DS 0.6618 OrgDS 0.6739 F7DS 0.6608 ZembDS 0.7583 F7DS 0.6618 OrgDS 0.6739 F7DS 0.6467 OrgDS 0.6739 F7DS 0.6467 OrgDS 0.6497 ZembDS 0.7728 F5DS 0.6619 OrgDS 0.6497 ZembDS 0.7728 F5DS 0.6619 OrgDS 0.6497 ZembDS 0.7778 F5DS 0.6619 OrgDS 0.6497 ZembDS 0.7778 F5DS 0.6619 OrgDS 0.6497 ZembDS 0.7778 F5DS 0.6497 ZembDS 0.7778 F5DS 0.6619 OrgDS 0.6497 ZembDS 0.7778 F5DS 0.6683 F7DS 0.6619 OrgDS 0.6497 ZembDS 0.7778 F5DS 0.6683 F7DS 0.6619 OrgDS 0.6497 ZembDS 0.7778 F5DS 0.6883 F5DS			8000 (all)		
TFIDF 2000 ZembDS 0.8297 F5DS 0.8044 F7DS 0.7808 OrgDS 0.7203 ZembDS 0.8581 F5DS 0.8664 F7DS 0.8158 OrgDS 0.7781 ZembDS 0.8789 F5DS 0.8628 F7DS 0.8628 F7DS 0.8422 OrgDS 0.8278 ZembDS 0.8664 OrgDS 0.8814 F7DS 0.8664 OrgDS 0.8333 ZembDDS 0.9006 F5DS 0.8742 OrgDS 0.5314 ZembDS 0.7047 F5DS 0.6439 F7DS 0.6439 F7DS 0.6439 F7DS 0.6608 F7DS 0.6608 F7DS 0.6608 F7DS 0.6608 OrgDS 0.6467 OrgDS 0.6467 OrgDS 0.6467 OrgDS 0.6467 OrgDS 0.6619 OrgDS 0.6699 OrgDS 0.6619 OrgDS 0.6497 OrgDS				F7DS	0.5450
TFIDF 1000 F5DS 0.8044 F7DS 0.7808 0.7203 ZembDS 0.7203 ZembDS 0.8581 F5DS 0.8364 F7DS 0.8789 E7DS 0.8628 F7DS 0.8628 F7DS 0.8422 OrgDS 0.8278 ZembDS 0.8956 E7DS 0.8864 OrgDS 0.8333 ZembDS 0.9966 E7DS 0.8842 F7DS 0.8664 OrgDS 0.8333 ZembDS 0.9006 E7DS 0.8742 OrgDS 0.8742 OrgDS 0.5314 ZembDS 0.7047 F5DS 0.6439 F7DS 0.6618 OrgDS 0.5725 ZembDS 0.7339 F5DS 0.6608 ZembDS 0.7583 F7DS 0.6608 ZembDS 0.7583 F7DS 0.6467 OrgDS 0.6739 F7DS 0.6467 OrgDS 0.6739 F7DS 0.6619 OrgDS 0.6497 ZembDS 0.7778 F5DS 0.6883 F7DS 0.6619 OrgDS 0.6497 OrgDS 0					
TFIDF 1000 F7DS 0.7808 0.7203 2.7205 0.7203 2.7205 0.7203 2.7205 0.8364 F7DS 0.8158 0.8789 F5DS 0.8628 F7DS 0.8628 F7DS 0.8278 2.7205 0.8956 0.7205 0.8956 0.9006 0.8956 0.9006 0.8956 0.9006			500		
TFIDF 2000 CrgDS 0.7203					
TFIDF 2000 ZembDS 0.8581 F5DS 0.8364 F7DS 0.8158 OrgDS 0.7781 ZembDS 0.8789 F5DS 0.8628 F7DS 0.8422 OrgDS 0.8278 ZembDS 0.8956 F5DS 0.8864 F7DS 0.8664 OrgDS 0.8333 ZembDS 0.9006 F5DS 0.8822 F7DS 0.8742 OrgDS 0.5314 ZembDS 0.7047 F5DS 0.6439 F7DS 0.6439 F7DS 0.6418 OrgDS 0.5725 ZembDS 0.7339 F7DS 0.6608 ZembDS 0.7583 F7DS 0.6467 OrgDS 0.6467 OrgDS 0.6467 OrgDS 0.6467 OrgDS 0.6467 OrgDS 0.6619 OrgDS 0.6619 OrgDS 0.6619 OrgDS 0.6497 OrgDS					
TFIDF 2000 F5DS 0.8364 F7DS 0.8158 OrgDS 0.7781 ZembDS 0.8789 F5DS 0.8628 F7DS 0.8628 F7DS 0.8956 F5DS 0.8956 F5DS 0.8814 F7DS 0.8664 OrgDS 0.8933 ZembDS 0.9006 F5DS 0.8814 F7DS 0.8664 OrgDS 0.8333 ZembDS 0.9006 F5DS 0.8822 F7DS 0.8742 OrgDS 0.5314 ZembDD 0.7047 F5DS 0.6439 F7DS 0.6443 F7DS 0.6181 OrgDS 0.5725 ZembDS 0.7339 F5DS 0.6608 ZembDS 0.7339 F5DS 0.6608 ZembDS 0.7583 F5DS 0.6636 OrgDS 0.6739 F7DS 0.6467 OrgDS 0.6739 F7DS 0.6467 OrgDS 0.6467 OrgDS 0.6619 OrgDS 0.6619 OrgDS 0.6619 OrgDS 0.6619 OrgDS 0.6619 OrgDS 0.6619 OrgDS 0.6497 ZembDS 0.7728 F5DS 0.66619 OrgDS 0.6497 ZembDS 0.7778 F5DS 0.66619 OrgDS 0.6497 ZembDS 0.7778 F5DS 0.66853 F7DS 0.66619 OrgDS 0.6497 ZembDS 0.7778 F5DS 0.66883					
TFIDF 2000 OrgDS 0.7781 ZembDS 0.8789 F5DS 0.8628 F7DS 0.8422 OrgDS 0.8278 ZembDS 0.8956 F5DS 0.8814 F7DS 0.8664 OrgDS 0.8333 ZembDDS 0.8333 ZembDS 0.9006 F5DS 0.8822 F7DS 0.8742 OrgDS 0.5314 ZembDS 0.7047 F5DS 0.6439 F7DS 0.6336 OrgDS 0.5725 ZembDS 0.7339 F7DS 0.6336 OrgDS 0.5725 ZembDS 0.7339 F7DS 0.6608 F7DS 0.6608 ZembDS 0.7583 F7DS 0.66467 OrgDS 0.6739 F7DS 0.66467 OrgDS 0.6619 OrgDS 0.6619 OrgDS 0.6619 OrgDS 0.6698 F7DS 0.66619 OrgDS 0.66883 F7DS 0.6619 OrgDS 0.6697 ZembDS 0.7728 F5DS 0.66883 F7DS 0.6619 OrgDS 0.6497 OrgDS 0.6778 F5DS 0.6883 F7DS 0.6883			1000		
TFIDF 2000 ZembDS 0.8789 F5DS 0.8628 F7DS 0.8628 F7DS 0.8628 F7DS 0.8628 F7DS 0.8956 P5DS 0.8956 P5DS 0.8956 P5DS 0.8864 P7DS 0.8664 P7DS 0.8664 P5DS 0.8956 P5DS 0.8814 P7DS 0.8664 P5DS 0.8956 P5DS 0.8682 P7DS 0.8742 P7DS 0.8742 P7DS 0.8742 P7DS 0.6439 P7DS 0.6439 P7DS 0.6439 P7DS 0.6439 P7DS 0.6608 P7DS 0.6336 P7DS 0.66086 P7DS 0.6636 P7DS 0.6636 P7DS 0.6636 P7DS 0.6636 P7DS 0.66467 P7DS 0.66619 P7DS 0.66619 P7DS 0.66619 P7DS 0.66619 P7DS 0.66619 P7DS 0.66883 P5DS 0.6883				F7DS	0.8158
SCOOK STOCK STOC					
SOUND SOUN		TEIDE	2000		
SQUENCE SQUE					
SOUD SembDS 0.8956 F5DS 0.8814 F7DS 0.8664 OrgDS 0.8333 SembDS 0.9006 F5DS 0.8812 F7DS 0.8664 OrgDS 0.8333 SembDS 0.9006 SF5DS 0.8822 F7DS 0.8742 OrgDS 0.5314 SembDS 0.7047 SF5DS 0.6439 F7DS 0.6439 F7DS 0.6181 OrgDS 0.5725 SembDS 0.7339 F5DS 0.6608 F7DS 0.6366 OrgDS 0.6086 SF7DS 0.6366 OrgDS 0.6739 SF5DS 0.6739 SF5DS 0.6739 SF7DS 0.64467 OrgDS 0.6417 SembDS 0.7728 SF5DS 0.6619 OrgDS 0.6497			-	1	
SOUU F5DS 0.8814 F7DS 0.8664 P7DS 0.8664 OrgDS 0.8333 ZembDS 0.9006 F5DS 0.8822 F7DS 0.8642 OrgDS 0.5314 ZembDS 0.7047 F5DS 0.6439 F7DS 0.64181 OrgDS 0.5725 ZembDS 0.7339 F5DS 0.6608 F7DS 0.6336 OrgDS 0.6608 F7DS 0.6336 OrgDS 0.6608 F7DS 0.6336 OrgDS 0.6467 OrgDS 0.6467 OrgDS 0.6467 OrgDS 0.6467 OrgDS 0.6467 OrgDS 0.6467 OrgDS 0.6469 ZembDS 0.7728 F5DS 0.6883 F7DS 0.6619 OrgDS 0.6497 ZembDS 0.7778 F5DS 0.6619 OrgDS 0.6497 ZembDS 0.7778 F5DS 0.6883					
BOW 2000 (all) F7DS 0.8664 OrgDS 0.8333 SembDS 0.9006 F5DS 0.8822 F7DS 0.8742 OrgDS 0.5314 ZembDS 0.7047 F5DS 0.6439 F7DS 0.6181 OrgDS 0.5725 ZembDS 0.7339 F5DS 0.6608 F7DS 0.6336 OrgDS 0.6086 ZembDS 0.7583 F5DS 0.6467 OrgDS 0.6417 OrgDS 0.6467 OrgDS 0.6417 OrgDS 0.6619 OrgDS 0.6619 OrgDS 0.6619 OrgDS 0.6619 OrgDS 0.6619 OrgDS 0.6619 OrgDS 0.6497 OrgDS 0.6497 OrgDS 0.6497 OrgDS 0.6497 OrgDS 0.6883 F5DS 0.6883 OrgDS 0.6883 OrgDS 0.6883 OrgDS 0.6497 OrgDS 0.6497 OrgDS 0.6497 OrgDS 0.6883 OrgDS			5000		
SOUD				F7DS	0.8664
BOW 2000 (all) F5DS					
BOW 2000 F7DS 0.8742 OrgDS 0.5314 ZembDS 0.7047 F5DS 0.6439 F7DS 0.6181 OrgDS 0.7339 F5DS 0.6608 F7DS 0.6336 OrgDS 0.6986 ZembDS 0.7583 F5DS 0.6467 OrgDS 0.6417 ZembDS 0.7728 F5DS 0.6619 OrgDS 0.6417 ZembDS 0.7728 F5DS 0.6619 OrgDS 0.6497 ZembDS 0.7778 F5DS 0.66497 ZembDS 0.7778 F5DS 0.6497 ZembDS 0.7778 F5DS 0.6497 ZembDS 0.7778 F5DS 0.6497 ZembDS 0.7778 F5DS 0.6883			8000 (all)		
BOW 2000 F5DS 0.5314 BOW 2000 F5DS 0.6489 F7DS 0.6181 OrgDS 0.5725 ZembDS 0.7339 F5DS 0.6608 F7DS 0.6336 OrgDS 0.5725 ZembDS 0.7339 F5DS 0.6608 ZembDS 0.7583 F5DS 0.6467 OrgDS 0.6467 OrgDS 0.6467 OrgDS 0.6467 OrgDS 0.6467 ZembDS 0.7728 F5DS 0.6883 F7DS 0.6619 ZembDS 0.7778 F5DS 0.6497 ZembDS 0.7778 F5DS 0.6497 ZembDS 0.7778 F5DS 0.6883					
BOW 2000 F5DS 0.6739 F7DS 0.6467 OrgDS 0.6417 ZembDS 0.7728 F5DS 0.6853 F7DS 0.6619 OrgDS 0.6497 ZembDS 0.7778 8000 (all) F5DS 0.6883	DS		-		
BOW 2000 F5DS 0.6739 F7DS 0.6467 OrgDS 0.6417 ZembDS 0.7728 F5DS 0.6853 F7DS 0.6619 OrgDS 0.6497 ZembDS 0.7778 8000 (all) F5DS 0.6883)RI				
BOW 2000 F5DS 0.6739 F7DS 0.6467 OrgDS 0.6417 ZembDS 0.7728 F5DS 0.6853 F7DS 0.6619 OrgDS 0.6497 ZembDS 0.7778 8000 (all) F5DS 0.6883	Μ		500		
BOW 2000 F5DS 0.6739 F7DS 0.6467 OrgDS 0.6417 ZembDS 0.7728 F5DS 0.6853 F7DS 0.6619 OrgDS 0.6497 ZembDS 0.7778 8000 (all) F5DS 0.6883	C			F7DS	0.6181
BOW 2000 F5DS 0.6739 F7DS 0.6467 OrgDS 0.6417 ZembDS 0.7728 F5DS 0.6853 F7DS 0.6619 OrgDS 0.6497 ZembDS 0.7778 8000 (all) F5DS 0.6883	ST				
BOW 2000 F5DS 0.6739 F7DS 0.6467 OrgDS 0.6417 ZembDS 0.7728 F5DS 0.6853 F7DS 0.6619 OrgDS 0.6497 ZembDS 0.7778 8000 (all) F5DS 0.6883	E		1000		
BOW 2000 F5DS 0.6739 F7DS 0.6467 OrgDS 0.6417 ZembDS 0.7728 F5DS 0.6853 F7DS 0.6619 OrgDS 0.6497 ZembDS 0.7778 8000 (all) F5DS 0.6883	10				
BOW 2000 F5DS 0.6739 F7DS 0.6467 OrgDS 0.6417 ZembDS 0.7728 F5DS 0.6853 F7DS 0.6619 OrgDS 0.6497 ZembDS 0.7778 8000 (all) F5DS 0.6883	E				
BOW 2000 F5DS 0.6739 F7DS 0.6467 OrgDS 0.6417 ZembDS 0.7728 F5DS 0.6853 F7DS 0.6619 OrgDS 0.6497 ZembDS 0.7778 8000 (all) F5DS 0.6883	×				
F7DS 0.6467 OrgDS 0.6417 ZembDS 0.7728 F5DS 0.6853 F7DS 0.6619 OrgDS 0.6497 ZembDS 0.7778 F5DS 0.6883		BOW	2000		
OrgDS 0.6417 ZembDS 0.7728 F5DS 0.6853 F7DS 0.6619 OrgDS 0.6619 OrgDS 0.6497 ZembDS 0.7778 F5DS 0.6883					
S000 F5DS				OrgDS	0.6417
F5DS 0.6853 F7DS 0.6619 OrgDS 0.6497 ZembDS 0.7778 F5DS 0.6883			5000		
OrgDS 0.6497 ZembDS 0.7778 F5DS 0.6883			3000		
8000 (all) ZembDS 0.7778 F5DS 0.6883					
8000 (all) F5DS 0.6883					
			8000 (all)		
1.25 0.0007					
		-			

7.7 Logistic Regression

	TFIDF	500	Dataset OrgDS ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS	Score 0.8861 0.9158 0.9142 0.9086 0.9081 0.9297 0.9253 0.9231
	TFIDF		ZembDS F5DS F7DS OrgDS ZembDS F5DS F7DS	0.9158 0.9142 0.9086 0.9081 0.9297 0.9253
	TFIDF		F5DS F7DS OrgDS ZembDS F5DS F7DS	0.9142 0.9086 0.9081 0.9297 0.9253
	TFIDF	1000	F7DS OrgDS ZembDS F5DS F7DS	0.9086 0.9081 0.9297 0.9253
	TFIDF	1000	OrgDS ZembDS F5DS F7DS	0.9081 0.9297 0.9253
	TFIDF	1000	ZembDS F5DS F7DS	0.9297 0.9253
	TFIDF	1000	F5DS F7DS	0.9253
	TFIDF		F7DS	
	TFIDF			
	TFIDF		OrgDS	0.9219
	TFIDF		ZembDS	0.9367
		2000	F5DS	0.9342
			F7DS	0.9328
			OrgDS	0.9278
		F000	ZembDS	0.9428
		5000	F5DS	0.9392
			F7DS	0.9367
			OrgDS	0.9300
		0000 (-11)	ZembDS	0.9433
		8000 (all)	F5DS	0.9417
(A)			F7DS	0.9389
l ĝ -			OrgDS	0.9389
Ō			ZembDS	0.9011
≥		500	F5DS	0.8994
Q.			F7DS	0.8964
ST			OrgDS	0.8950
WITH STOPWORDS			ZembDS	0.8930
<u> </u>		1000	F5DS	0.9119
≤			F7DS	0.9119
			OrgDS	0.9003
	no		ZembDS	0.9189
	BOW	2000	F5DS	0.9203
			F7DS	0.9178
			OrgDS	0.9094
			ZembDS	0.9247
		5000	F5DS	0.9244
			F7DS	0.9225
			OrgDS	0.9111
		0000 (11)	ZembDS	0.9275
		8000 (all)	F5DS	0.9269
			F7DS	0.9247
			OrgDS	0.7786
		500	ZembDS	0.8750
		300	F5DS	0.8647
			F7DS	0.8511
			OrgDS	0.8186
		1000	ZembDS	0.8967
		1000	F5DS	0.8892
			F7DS	0.8797
			OrgDS	0.8636
	TFIDF	2000	ZembDS	0.9128
			F5DS	0.9086
			F7DS	0.8997
			OrgDS	0.8933
		5000	ZembDS F5DS	0.9261
				0.9208
			F7DS	0.9219
			OrgDS ZembDS	0.9006 0.9297
		8000 (all)	F5DS	0.9297
			F7DS	0.9239
22			OrgDS	0.9242
K			ZembDS	0.7361
γ		500	F5DS	0.8461
ΡV			F7DS	0.8322
TC			OrgDS	0.8039
S			ZembDS	0.8819
Ε'		1000	F5DS	0.8731
\supset			F7DS	0.8561
100		<u> </u>	OrgDS	0.8431
THOU			ZembDS	0.8906
WITHOUT STOPWORE			F5DS	
WITHOU	BOW	2000		
WITHOU	BOW	2000		0.8858 0.8847
WITHOU	BOW	2000	F7DS	0.8847
WITHOU	BOW		F7DS OrgDS	0.8847 0.8706
WITHOU	BOW	5000	F7DS OrgDS ZembDS	0.8847 0.8706 0.8981
WITHOU	BOW		F7DS OrgDS ZembDS F5DS	0.8847 0.8706 0.8981 0.8942
MITHOU	BOW		F7DS OrgDS ZembDS F5DS F7DS	0.8847 0.8706 0.8981 0.8942 0.9014
WITHOU	BOW	5000	F7DS OrgDS ZembDS F5DS F7DS OrgDS	0.8847 0.8706 0.8981 0.8942 0.9014 0.8725
WITHOU	BOW		F7DS OrgDS ZembDS F5DS F7DS	0.8847 0.8706 0.8981 0.8942 0.9014

8 Conclusion

First of all, based on the results of the above deep and extensive experiments, we tried to determine what is the best parameter for a method, and then to determine the common point between these parameters.

All of them featured a certain pattern. When we examine the tables, the datasets where the removal of stop-words are made, certainly performed worse. When we compare the lines, stop-word, i.e. conjunction, etc. texts from which words were removed always showed less successful results. We have seen in all experiments without exception that removing the stop-words lowers the scores.

A second pattern is that the TFIDF vector generally performs better than the BOW vector. TFIDF lines gave more successful results than BOW lines in the scores of the parameter experiments we performed for all methods. This does not apply to the Naive Bayes alone. Naive Bayes provided the best score for all parameters with the BOW vector. But still, we can consider the success of TFIDF.

Among the best 7 scores, 4 are the result obtained on Zemb-DS. Two of them were obtained on F5-DS and the last one on F7-DS.

We can see from the general pattern analysis we made about the results that the dataset processed by stemming with the spring wire has generally yielded more successful results.

The best results, together with these common patterns we have caught with; using the TFIDF in the attribute vector, the preeprocessed data set is stuck in the resulting range, with the stop-words not extracted. So now, in order to gain a more distant view of the picture, we can draw the following table, which only takes this range into our perspective.

	Feature: 500				Feature: 5000				Feature: 8000 (All)			
	OrgDS	ZembDS	F5DS	F7DS	OrgDS	ZembDS	F5DS	F7DS	OrgDS	ZembDS	F5DS	F7DS
NB	0.8903	0.9139	0.9144	0.9158	0.9261	0.9339	0.9319	0.9322	0.9256	0.9342	0.9306	0.9328
RF	0.8783	0.9033	0.9083	0.8950	0.8958	0.9150	0.9164	0.9081	0.8903	0.9133	0.9122	0.9108
SVM	0.8969	0.9236	0.9222	0.9211	0.9356	0.9492	0.9475	0.9458	0.9389	0.9508	0.9506	0.9483
KNN	0.7861	0.8367	0.8261	0.8189	0.7356	0.8569	0.8244	0.7997	0.8767	0.9003	0.9017	0.9006
CART	0.7850	0.8072	0.8061	0.7944	0.7244	0.7806	0.7856	0.7747	0.7339	0.7842	0.7867	0.7753
Rocchio	0.8342	0.8872	0.8761	0.8700	0.8931	0.9106	0.9081	0.9092	0.8958	0.9106	0.9106	0.9111
LR	0.8861	0.9158	0.9142	0.9086	0.9278	0.9428	0.9392	0.9367	0.9300	0.9433	0.9417	0.9389

In the table below, the best scores obtained in certain data sets in the reference study in all methods and the best scores obtained in the specific dataset in our study are compared.

	[1]				Ours			
	OrgDS	ZembDS	F5DS	F7DS	OrgDS	ZembDS	F5DS	F7DS
NB	0.8294	0.8719	0.8222	0.8403	0.9256	0.9342	0.9319	0.9328
RF	0.8887	0.9103	0.8828	0.8859	0.8903	0.9150	0.9122	0.9108
SVM	0.8603	0.8497	0.8239	0.8356	0.9389	0.9492	0.9506	0.9483
KNN	0.7311	0.7497	0.6944	0.7256	0.8767	0.9003	0.9017	0.9006
CART	0.7897	0.7939	0.7736	0.7597	0.7850	0.8072	0.8061	0.7944

Better results were obtained in our study in all methods and in all datasets.

9 References

- 1 Kilinç, Deniz & Ozcift, Akin & Bozyiğit, Fatma & Yildirim, Pelin & Yucalar, Fatih & Borandağ, Emin. (2017). TTC-3600: A new benchmark dataset for Turkish text categorization. Journal of Information Science. 43. 174-185.
- 2 A. Fallis, "Text Categorisation: A Survey," J. Chem. Inf. Model., vol. 53, no. 9, pp. 1689–1699, 1999.
- 3 M. U. SALUR and İ. AYDIN, "The Impact of Preprocessing on Classification Performance in Convolutional Neural Networks for Turkish Text," 2018 International Conference on Artificial Intelligence and Data Processing (IDAP), Malatya, Turkey, 2018, pp. 1-4.
- 4 İ. Aydın, M. U. Salur and F. Başkaya "Duygu Analizi için Çoklu Popülasyon Tabanlı Parçacık Sürü Optimizasyonu," Türkiye Bilişim Vakfı Bilgi. Bilim. ve Mühendisliği Derg., vol. 11, no. 1, pp. 52–64, 2018.
- 5 Can F, Kocberber S, Balcik E, Kaynak C, Ocalan HC and Vursavas OM. Information retrieval on Turkish texts. Journal of the American Society for Information Science and Technology 2008; 59(3): 407–421.
- 6 Akin AA and Akin MD. Zemberek, an open source NLP framework for Turkic Languages. Structure 2007; 10: 1–5.