**Syracuse University**

**IST-707 Assignment 5**

ThulasiRam RuppaKrishnan

IST 707Section: 35

Professor Ami Gates

Contents

[Introduction 3](#_Toc8258624)

[Analysis and Models 5](#_Toc8258625)

[**About the data** 5](#_Toc8258626)

[**Models** 12](#_Toc8258627)

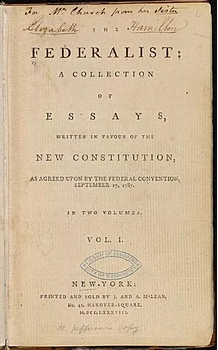
[Results 18](#_Toc8258628)

[Conclusion 25](#_Toc8258629)

## 

## **Introduction**

The Federalist (later known as The Federalist Papers) is a collection of 85 articles and essays written by Alexander Hamilton, James Madison, and John Jay under the pseudonym "Publius" to promote the ratification of the United States Constitution. The first 77 of these essays were published serially in the Independent Journal, the New York Packet, and The Daily Advertiser between October 1787 and April 1788. A two-volume compilation of these 77 essays and eight others was published as "The Federalist: A Collection of Essays", Written in Favor of the New Constitution, as Agreed upon by the Federal Convention, September 17, 1787, by publishing firm J. & A. McLean in March and May 1788.



The authors of The Federalist intended to influence the voters to ratify the Constitution. In "Federalist No. 1", they explicitly set that debate in broad political terms:

It has been frequently remarked, that it seems to have been reserved to the people of this country, by their conduct and example, to decide the important question, whether societies of men are capable or not, of establishing good government from reflection and choice, or whether they are forever destined to depend, for their political constitutions, on accident and force.

"Federalist No. 10" is generally regarded as the most important of the 85 articles from a philosophical perspective. In it, Madison discusses the means of preventing rule by majority faction and advocates a large, commercial republic. This is complemented by "Federalist No. 14", in which Madison takes the measure of the United States, declares it appropriate for an extended republic, and concludes with a great defense of the constitutional and political creativity of the Federal Convention. In "Federalist No. 84", Hamilton makes the case that there is no need to amend the Constitution by adding a Bill of Rights, insisting that the various provisions in the proposed Constitution protecting liberty amount to a "bill of rights." "Federalist No. 78", also written by Hamilton, lays the groundwork for the doctrine of judicial review by federal courts of federal legislation or executive acts. "Federalist No. 70" presents Hamilton's case for a one-person chief executive. In "Federalist No. 39", Madison presents the clearest exposition of what has come to be called "Federalism." In "Federalist No. 51", Madison distills arguments for checks and balances in an essay often quoted for its justification of government as "the greatest of all reflections on human nature."

According to historian Richard B. Morris, the essays that make up The Federalist Papers are an "incomparable exposition of the Constitution, a classic in political science unsurpassed in both breadth and depth by the product of any later American writer."

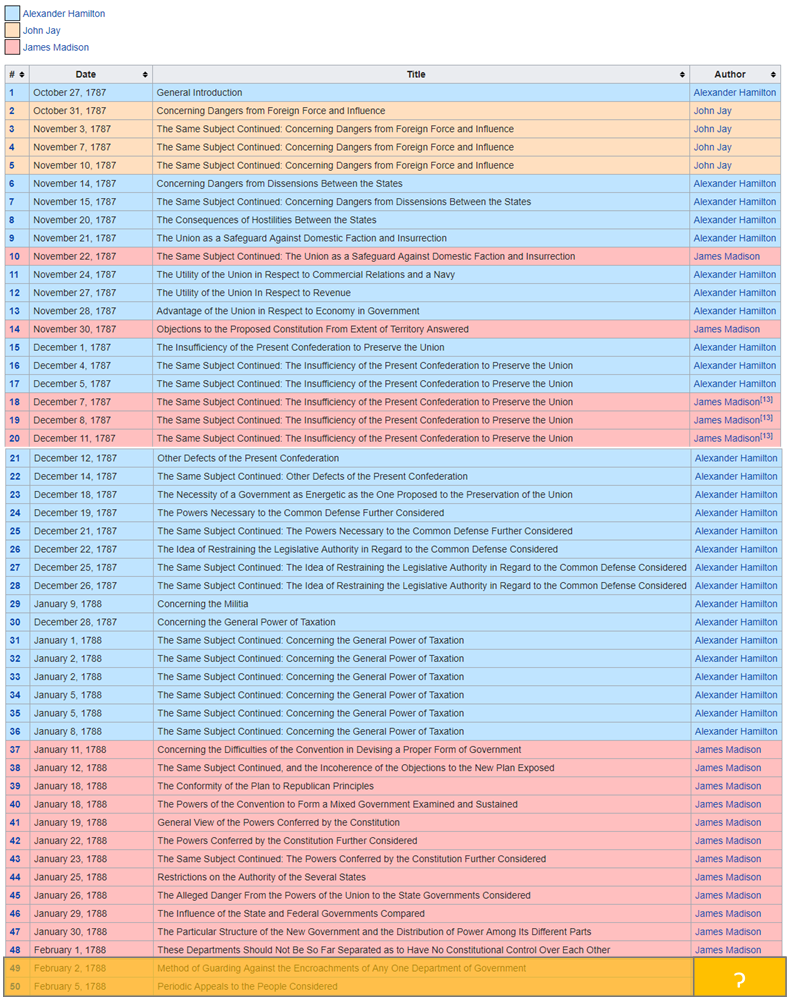
The authorship of certain of the Federalist essays was disputed from the beginning. Both Hamilton and Madison produced lists that claimed some of the same papers. There followed a series of lists, some claiming authorship for Madison and some for Hamilton.

The consensus of traditional scholarship, seconded by Mosteller and Wallace, allocates the papers: Hamilton 51 (1, 6-9, 11-13, 15-17, 21-36, 59-61, 65-85); Madison 29 (10, 14, 18-20, 37-58, 62, 63); Jay 5 (2-5, 64). Mosteller and Wallace set the boundary conditions for the subsequent non-traditional work – e.g., not using the Jay articles as a control. Most of these later practitioners do not select or prepare the input text as carefully as Mosteller and Wallace – and their selection and preparation were not as rigorous and complete as it should have been – as we will see.

## **Analysis and Models**

### **About the data**

The list of federalist papers is shown in **Table 1.1** and are highlighted with colors for each author.



****

**Table 1.1 Federalist papers by Author**

In the above list of papers, the ones which are highlighted is disputed between two authors Madison and Hamilton. Clustering techniques are applied here to solve the puzzle.

**Summary of the imported corpus data**

(summary(EssayCorpus))

Length Class Mode

dispt\_fed\_49.txt 2 PlainTextDocument list

dispt\_fed\_50.txt 2 PlainTextDocument list

dispt\_fed\_51.txt 2 PlainTextDocument list

dispt\_fed\_52.txt 2 PlainTextDocument list

dispt\_fed\_53.txt 2 PlainTextDocument list

dispt\_fed\_54.txt 2 PlainTextDocument list

dispt\_fed\_55.txt 2 PlainTextDocument list

dispt\_fed\_56.txt 2 PlainTextDocument list

dispt\_fed\_57.txt 2 PlainTextDocument list

dispt\_fed\_62.txt 2 PlainTextDocument list

dispt\_fed\_63.txt 2 PlainTextDocument list

Hamilton\_fed\_1.txt 2 PlainTextDocument list

Hamilton\_fed\_11.txt 2 PlainTextDocument list

Hamilton\_fed\_12.txt 2 PlainTextDocument list

Hamilton\_fed\_13.txt 2 PlainTextDocument list

Hamilton\_fed\_15.txt 2 PlainTextDocument list

Hamilton\_fed\_16.txt 2 PlainTextDocument list

Hamilton\_fed\_17.txt 2 PlainTextDocument list

Hamilton\_fed\_21.txt 2 PlainTextDocument list

Hamilton\_fed\_22.txt 2 PlainTextDocument list

Hamilton\_fed\_23.txt 2 PlainTextDocument list

Hamilton\_fed\_24.txt 2 PlainTextDocument list

Hamilton\_fed\_25.txt 2 PlainTextDocument list

Hamilton\_fed\_26.txt 2 PlainTextDocument list

Hamilton\_fed\_27.txt 2 PlainTextDocument list

Hamilton\_fed\_28.txt 2 PlainTextDocument list

Hamilton\_fed\_29.txt 2 PlainTextDocument list

Hamilton\_fed\_30.txt 2 PlainTextDocument list

Hamilton\_fed\_31.txt 2 PlainTextDocument list

Hamilton\_fed\_32.txt 2 PlainTextDocument list

Hamilton\_fed\_33.txt 2 PlainTextDocument list

Hamilton\_fed\_34.txt 2 PlainTextDocument list

Hamilton\_fed\_35.txt 2 PlainTextDocument list

Hamilton\_fed\_36.txt 2 PlainTextDocument list

Hamilton\_fed\_59.txt 2 PlainTextDocument list

Hamilton\_fed\_6.txt 2 PlainTextDocument list

Hamilton\_fed\_60.txt 2 PlainTextDocument list

Hamilton\_fed\_61.txt 2 PlainTextDocument list

Hamilton\_fed\_65.txt 2 PlainTextDocument list

Hamilton\_fed\_66.txt 2 PlainTextDocument list

Hamilton\_fed\_67.txt 2 PlainTextDocument list

Hamilton\_fed\_68.txt 2 PlainTextDocument list

Hamilton\_fed\_69.txt 2 PlainTextDocument list

Hamilton\_fed\_7.txt 2 PlainTextDocument list

Hamilton\_fed\_70.txt 2 PlainTextDocument list

Hamilton\_fed\_71.txt 2 PlainTextDocument list

Hamilton\_fed\_72.txt 2 PlainTextDocument list

Hamilton\_fed\_73.txt 2 PlainTextDocument list

Hamilton\_fed\_74.txt 2 PlainTextDocument list

Hamilton\_fed\_75.txt 2 PlainTextDocument list

Hamilton\_fed\_76.txt 2 PlainTextDocument list

Hamilton\_fed\_77.txt 2 PlainTextDocument list

Hamilton\_fed\_78.txt 2 PlainTextDocument list

Hamilton\_fed\_79.txt 2 PlainTextDocument list

Hamilton\_fed\_8.txt 2 PlainTextDocument list

Hamilton\_fed\_80.txt 2 PlainTextDocument list

Hamilton\_fed\_81.txt 2 PlainTextDocument list

Hamilton\_fed\_82.txt 2 PlainTextDocument list

Hamilton\_fed\_83.txt 2 PlainTextDocument list

Hamilton\_fed\_84.txt 2 PlainTextDocument list

Hamilton\_fed\_85.txt 2 PlainTextDocument list

Hamilton\_fed\_9.txt 2 PlainTextDocument list

HM\_fed\_18.txt 2 PlainTextDocument list

HM\_fed\_19.txt 2 PlainTextDocument list

HM\_fed\_20.txt 2 PlainTextDocument list

Jay\_fed\_2.txt 2 PlainTextDocument list

Jay\_fed\_3.txt 2 PlainTextDocument list

Jay\_fed\_4.txt 2 PlainTextDocument list

Jay\_fed\_5.txt 2 PlainTextDocument list

Jay\_fed\_64.txt 2 PlainTextDocument list

Madison\_fed\_10.txt 2 PlainTextDocument list

Madison\_fed\_14.txt 2 PlainTextDocument list

Madison\_fed\_37.txt 2 PlainTextDocument list

Madison\_fed\_38.txt 2 PlainTextDocument list

Madison\_fed\_39.txt 2 PlainTextDocument list

Madison\_fed\_40.txt 2 PlainTextDocument list

Madison\_fed\_41.txt 2 PlainTextDocument list

Madison\_fed\_42.txt 2 PlainTextDocument list

Madison\_fed\_43.txt 2 PlainTextDocument list

Madison\_fed\_44.txt 2 PlainTextDocument list

Madison\_fed\_45.txt 2 PlainTextDocument list

Madison\_fed\_46.txt 2 PlainTextDocument list

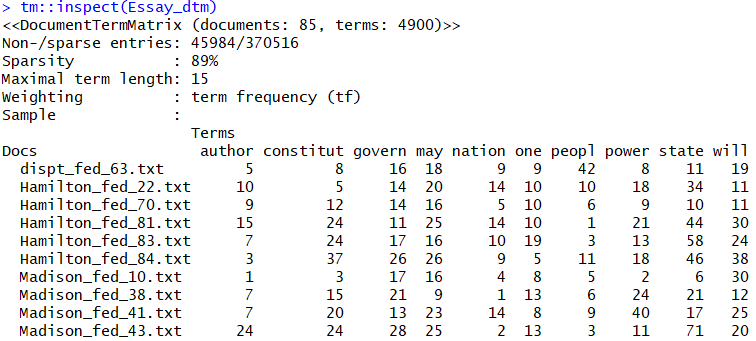
Madison\_fed\_47.txt 2 PlainTextDocument list

Madison\_fed\_48.txt 2 PlainTextDocument list

Madison\_fed\_58.txt 2 PlainTextDocument list

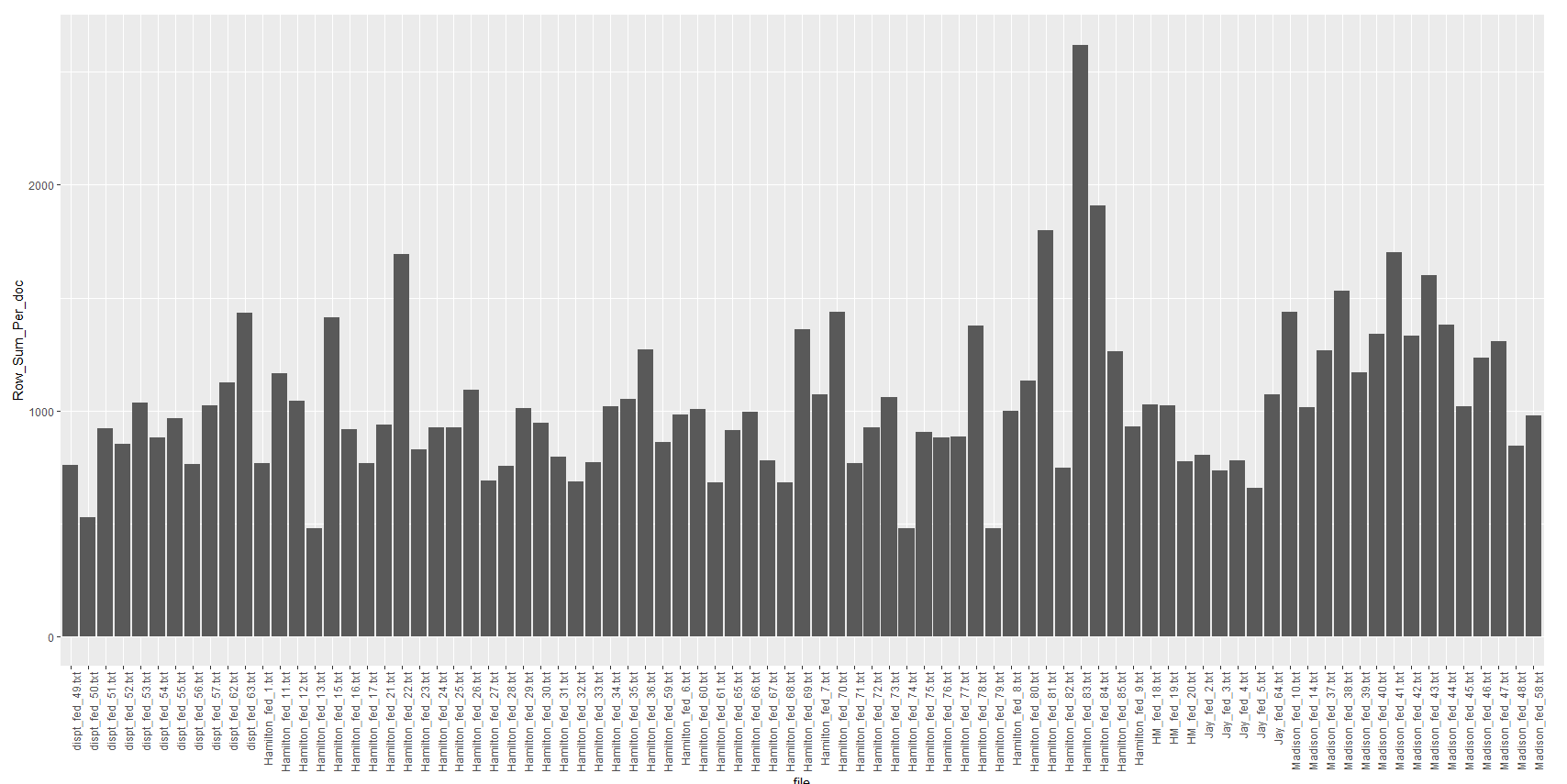
**Table 1.2 Papers**

**Inspect the Document Term Matrix**



**Table 1.3 Document Term Matrix**

**Word count by papers**



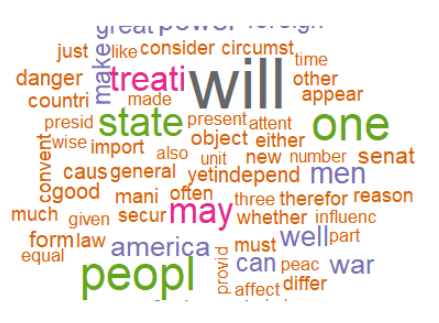
**Fig 1.1 Word count by papers**

**Word Clouds**

Fig 1.2,1.3,1.4 and 1.5 shows the word cloud of federalist papers by authors



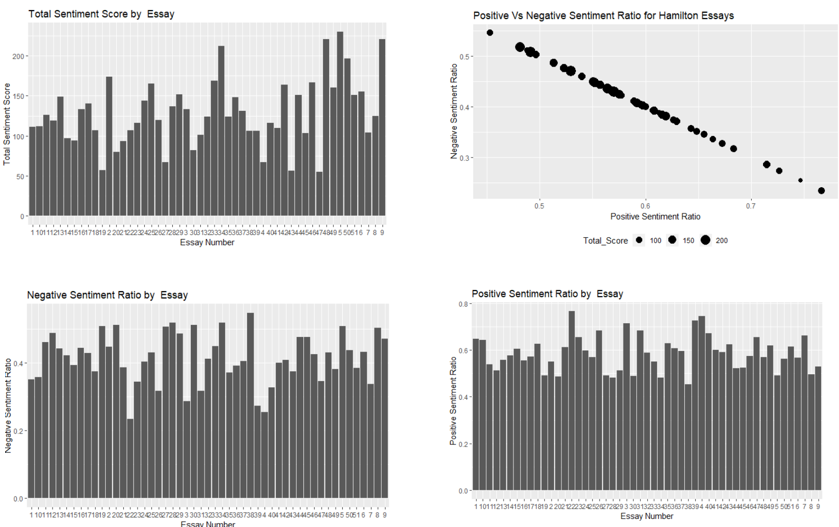
**Fig 1.2 Word Cloud on Disputed Papers Fig 1.3 Word Cloud on Hamilton Papers**



**Fig 1.4 Word Cloud on Jay Papers Fig 1.5 Word Cloud of Madison Papers**

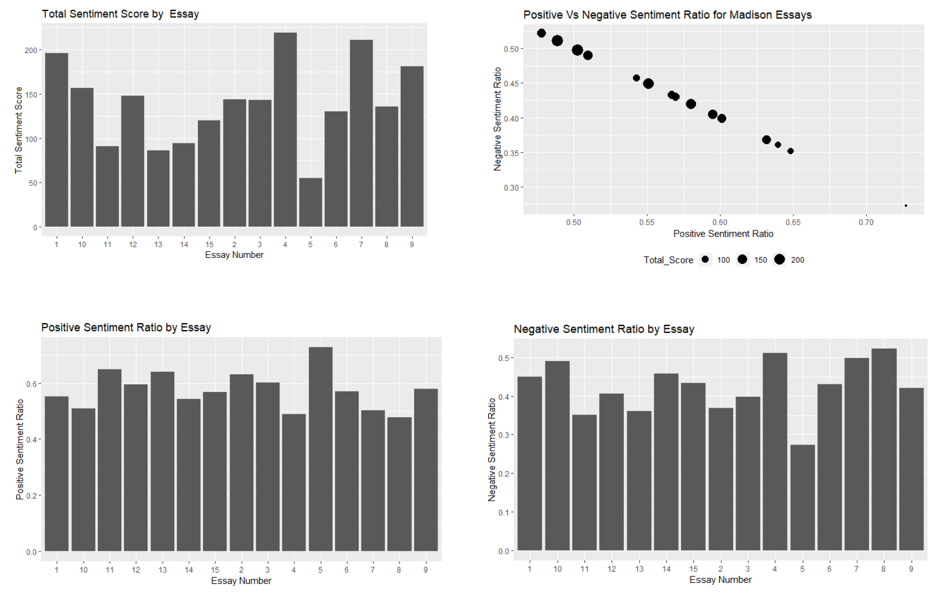
**Sentiment Analysis**

**Fig 1.6** shows the distribution of positive sentiment score, negative sentiment score and their plot to each other for all Hamilton papers



**Fig 1.6 Sentiment Analysis on Hamilton papers**

**Fig 1.7** shows the distribution of positive sentiment score, negative sentiment score and their plot to each other for all Madison papers



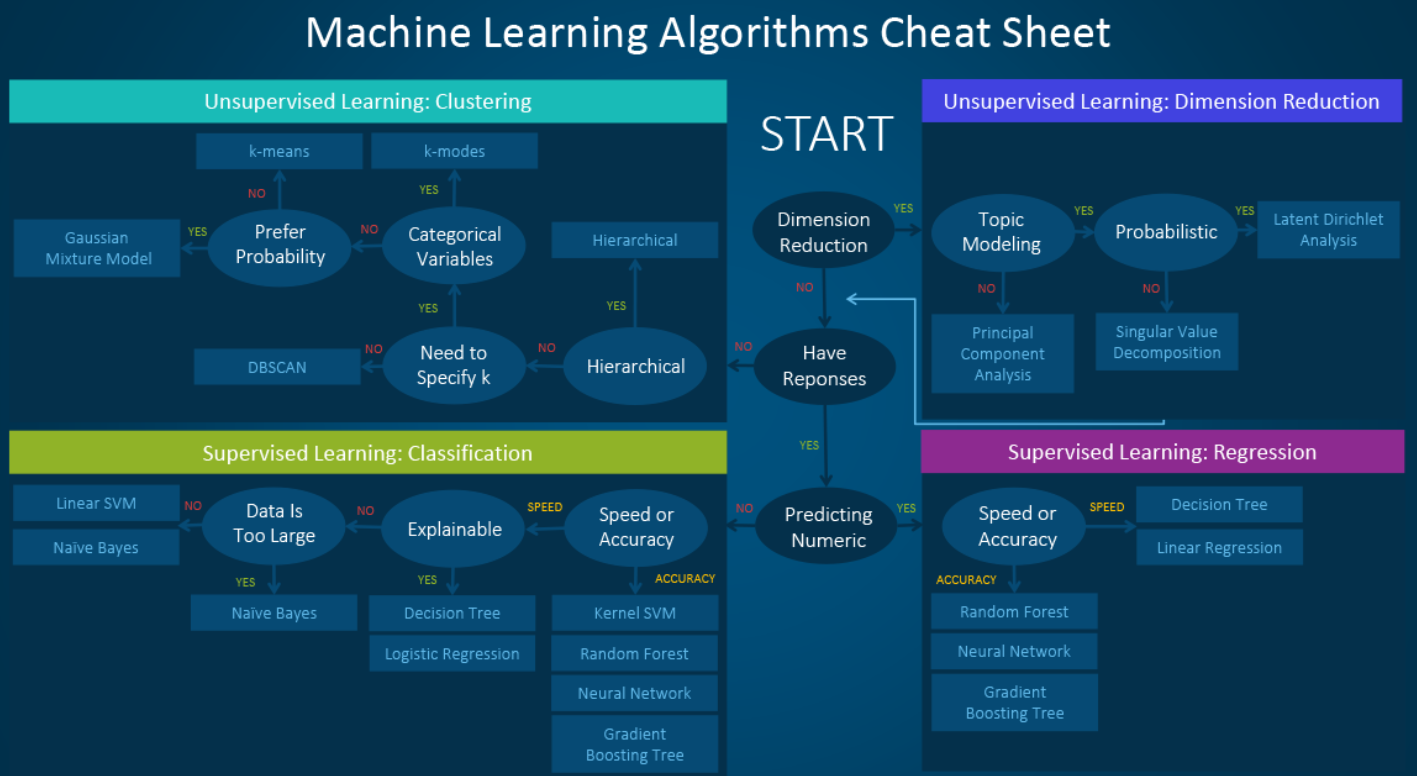
**Fig 1.7 Sentiment Analysis on Madison papers**

### **Models**

Supervised Machine Learning: Most of the practical machine learning uses supervised learning. Supervised learning is where you have input variables (x) and an output variable (Y) and you use an algorithm to learn the mapping function from the input to the output Y = f(X). The goal is to approximate the mapping function so well that when the output variables (Y) for that data can be predicted using the new input data (X).

Techniques of Supervised Machine Learning algorithms include linear and logistic regression, multi-class classification, Decision Trees and support vector machines. Supervised learning requires that the data used to train the algorithm is already labeled with correct answers. For example, a classification algorithm will learn to identify animals after being trained on a dataset of images that are properly labeled with the species of the animal and some identifying characteristics.

Supervised learning problems can be further grouped into Regression and Classification problems. Both problems have as goal the construction of a succinct model that can predict the value of the dependent attribute from the attribute variables. The difference between the two tasks is the fact that the dependent attribute is numerical for regression and categorical for classification.



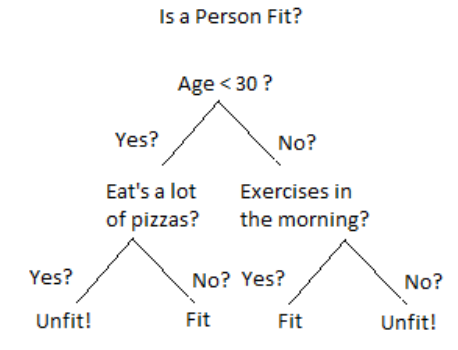
**Fig 2.1 Machine Learning Algorithm Cheat Sheet**

#### **Classification**

Classification is a technique for determining class of the dependent variable based on the one or more independent variables. Classification is used for predicting discrete responses.

#### **Decision Tree Classification**

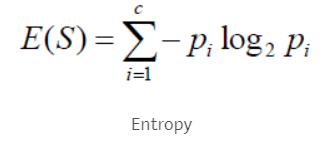
Decision tree builds classification or regression models in the form of a tree structure. It breaks down a dataset into smaller and smaller subsets while at the same time an associated decision tree is incrementally developed. The result is a tree with decision nodes and leaf nodes. It follows Iterative Dichotomiser 3(ID3) algorithm structure for determining the split. Decision Tree uses Entropy and Information Gain to construct a decision tree. **Fig 2.2** shows the simple illustration on how decision tree algorithm works



**Fig 2.2 Decision Tree**

#### **Entropy**

Entropy is the degree or amount of uncertainty in the randomness of elements or in other words it is a measure of impurity. Intuitively, it tells us about the predictability of a certain event. Entropy calculates the homogeneity of a sample. If the sample is completely homogeneous the entropy is zero and if the sample is an equally divided it has an entropy of one.



#### **Information Gain**

It measures the relative change in entropy with respect to the independent attribute. It tries to estimate the information contained by each attribute. Constructing a decision tree is all about finding the attribute that returns the highest information gain (i.e., the most homogeneous branches). Where Gain (T, X) is the information gain by applying feature X. Entropy(T) is the Entropy of the entire set, while the second term calculates the Entropy after applying the feature X.

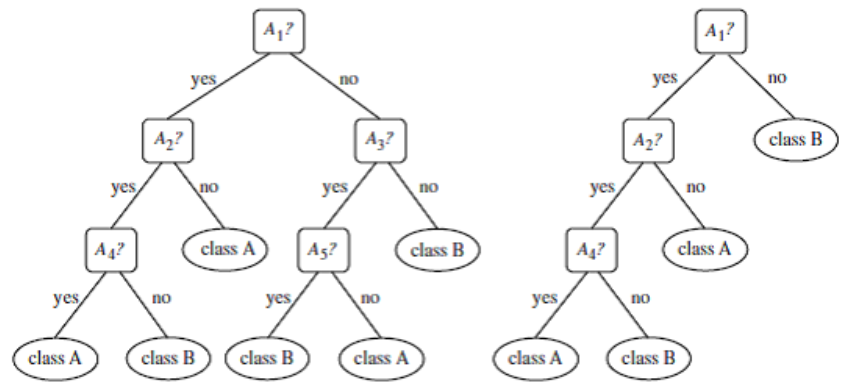


Information Gain ranks attribute for filtering at a given node in the tree. The ranking is based on the highest information gain entropy in each split.

The disadvantage of a Decision Tree Model is overfitting as it tries to fit the model by going deeper in the training set and thereby reducing test accuracy.

#### **Overfitting**

Overfitting in Decision Trees can be minimized by pruning nodes. Tree pruning is illustrated in the **Fig 2.3**

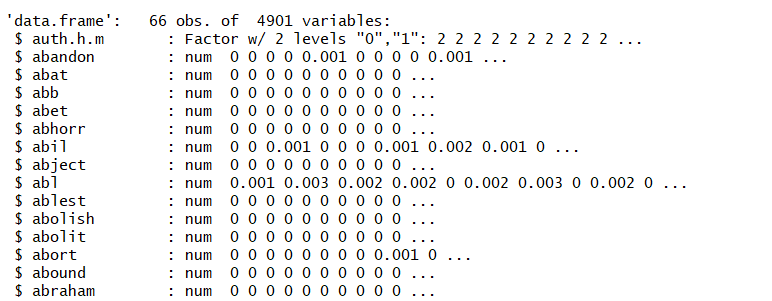


**Fig 2.3 Decision Tree Pruning**

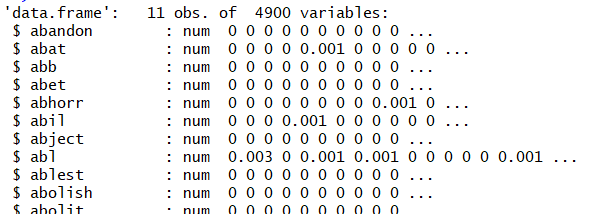
#### **Model 1: Decision tree using word frequencies of all words**

All the Hamilton and Madison papers along with word frequencies which are normalized and stored in the training set. All the dispute papers with the same word frequencies with no labels are normalized and stored in a test set

**Training set**

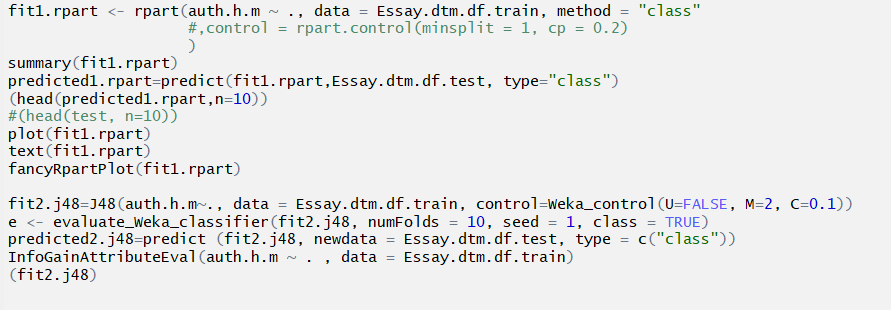


**Test set**



**Algorithm**

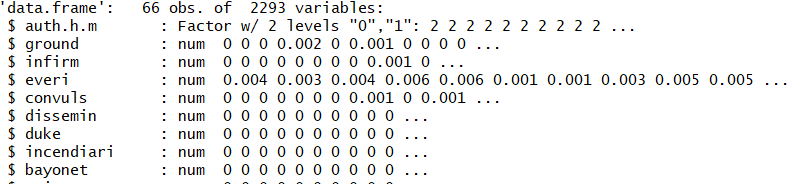
Decision trees are constructed using rpart and also with Rweka and the below code shows the construction of the model using the training set and predicted the results of the test set



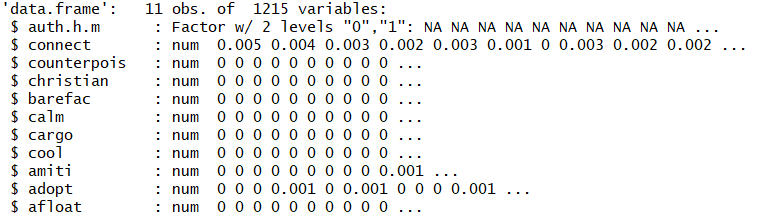
#### **Model 2: Decision tree using word frequencies of top 20% of the words**

All the Hamilton and Madison papers along with word frequencies which are normalized and stored in the training set with only top 20% of frequently used words. All the dispute papers with the same word frequencies with no labels are normalized and stored in a test set

**Training set**

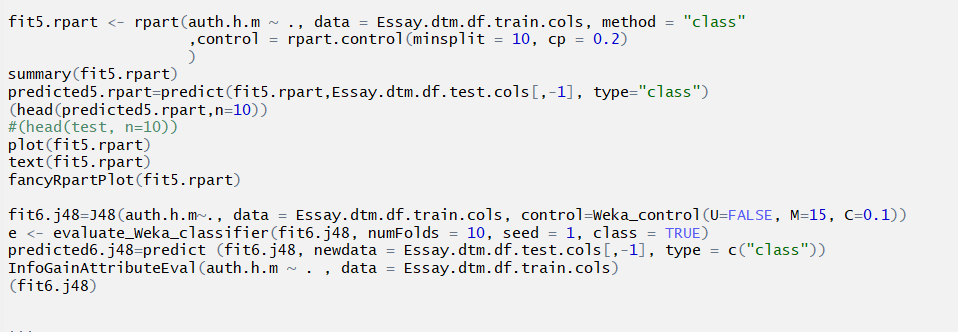


**Test set**



**Algorithm**

Decision trees are constructed using rpart and also with Rweka and the below code shows the construction of the model using the training set and predicted the results of the test set

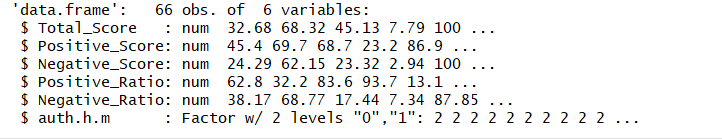


#### **Model 3: Decision tree using sentiment score**

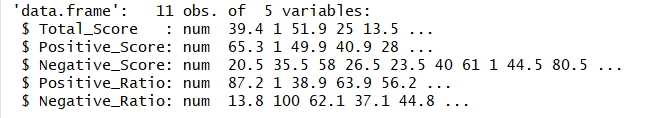
Sentiment scores are derived for all Hamilton and Madison papers and the below variables are used for determining the decision trees

"Total\_Score","Positive\_Score","Negative\_Score","Positive\_Ratio","Negative\_Ratio"

**Training set**

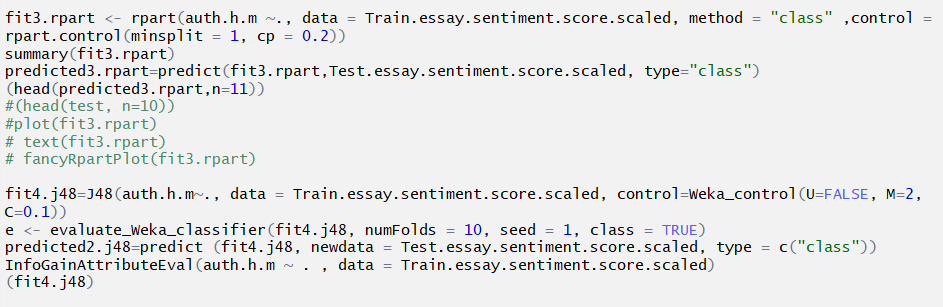


**Test set**



**Algorithm**

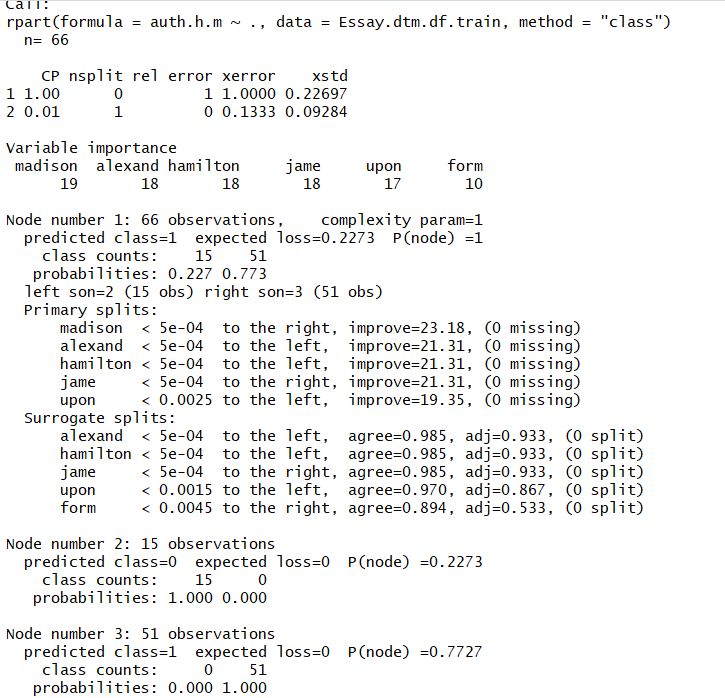
Decision trees are constructed using rpart and also with Rweka and the below code shows the construction of the model using the training set and predicted the results of the test set



## **Results**

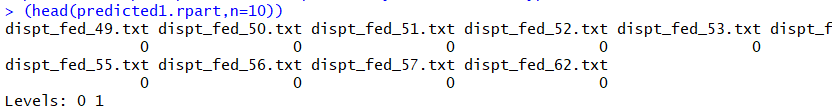
#### **Model 1: Results of decision tree using word frequencies of all words**

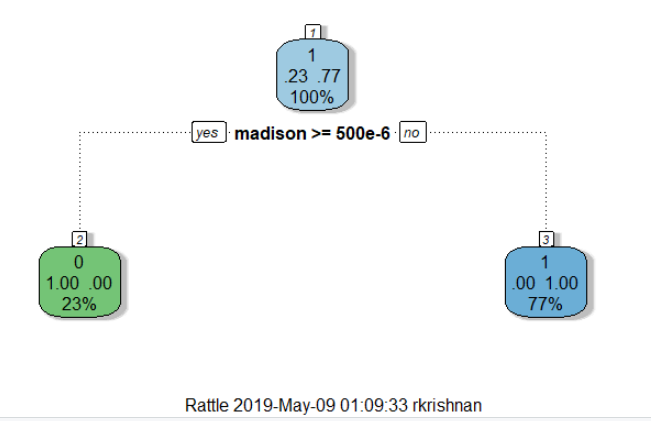
Built model rpart is given below



**Predicted Results**

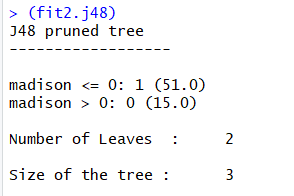
All the disputed essays are classified as 1 as shown in the **Fig.3.1**



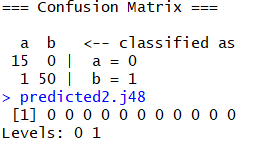


**Fig 3.1 Decision Tree**

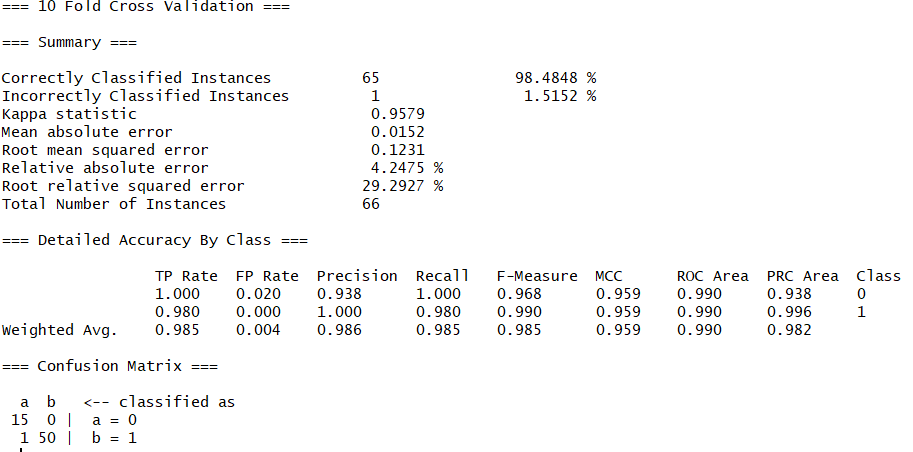
Built model j48 is given below



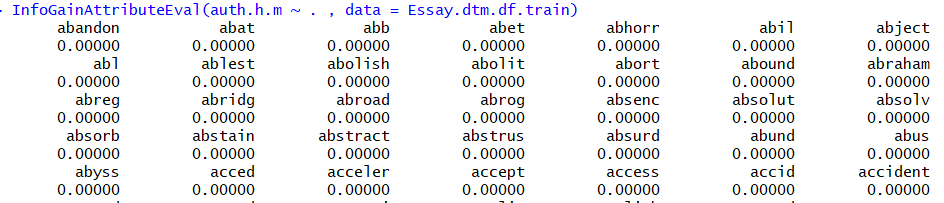
**Confusion Matrix**



**Cross Validation**

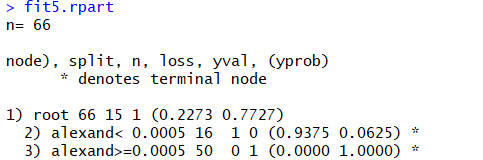


**Information Gain**



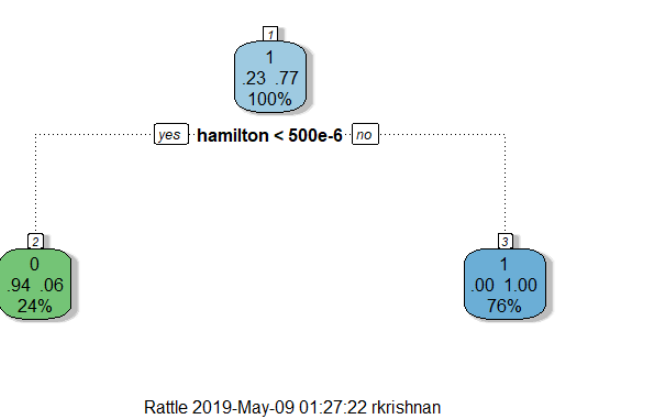
#### **Model 2: Results of decision tree using word frequencies of top 20% of the words**

Built model is given below



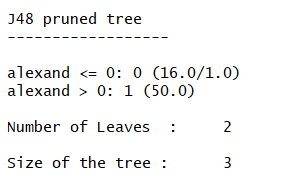
**Predicted Results**

All the disputed essays are classified as 1 as shown in the **Fig.3.1**



**Fig 3.3 Decision Tree**

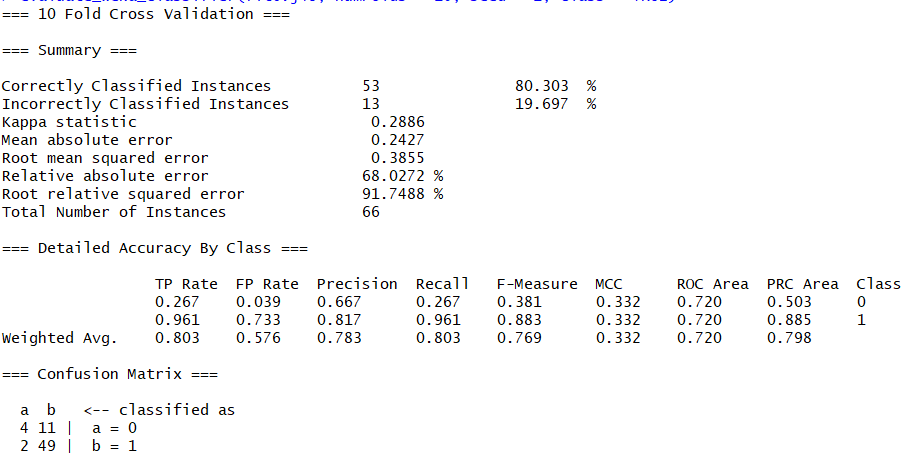
Built model j48 is given below



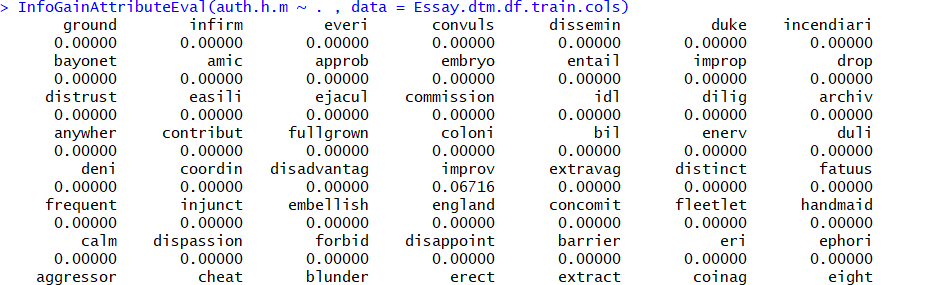
**Confusion Matrix**



**Cross Validation**

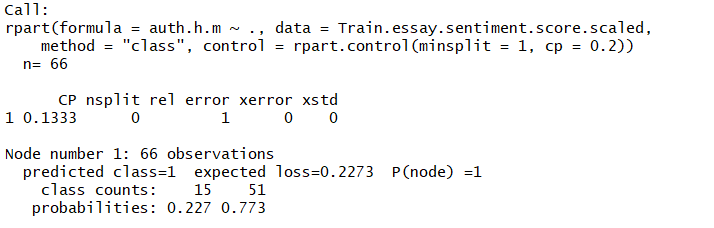


**Information Gain**



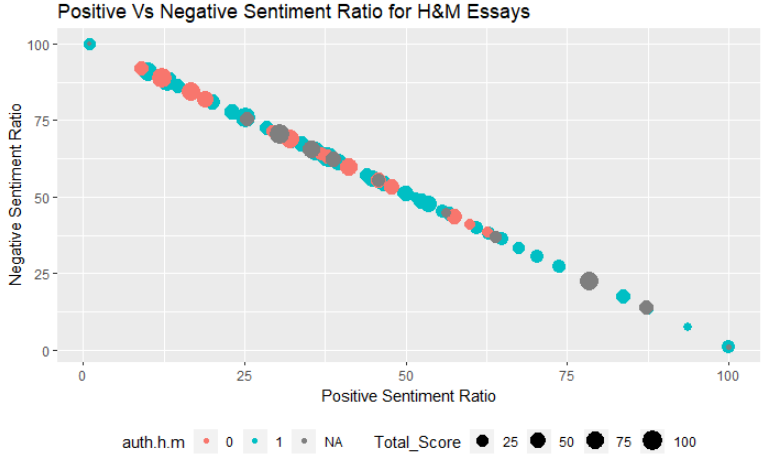
#### **Model 3: Results of decision tree using sentiment score**

Built model is given below



**Predicted Results**

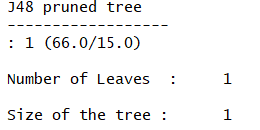
All the disputed essays are classified as 1 as shown in the **Fig 3.4**



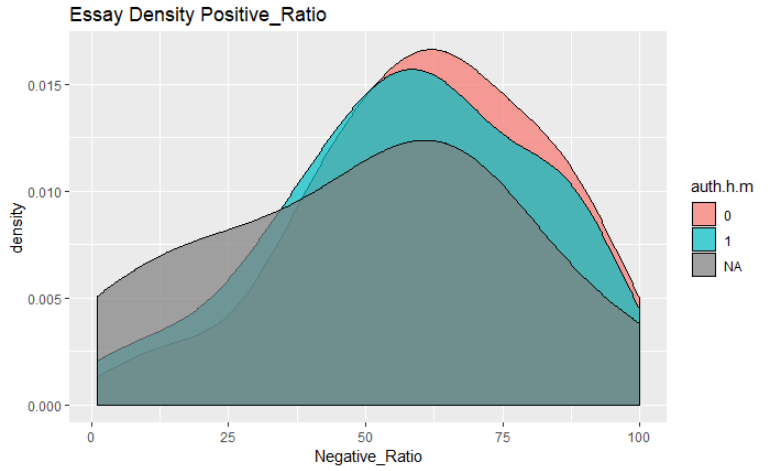
**Fig.3.4 Sentiment ratio plot**



Built model j48 is given below

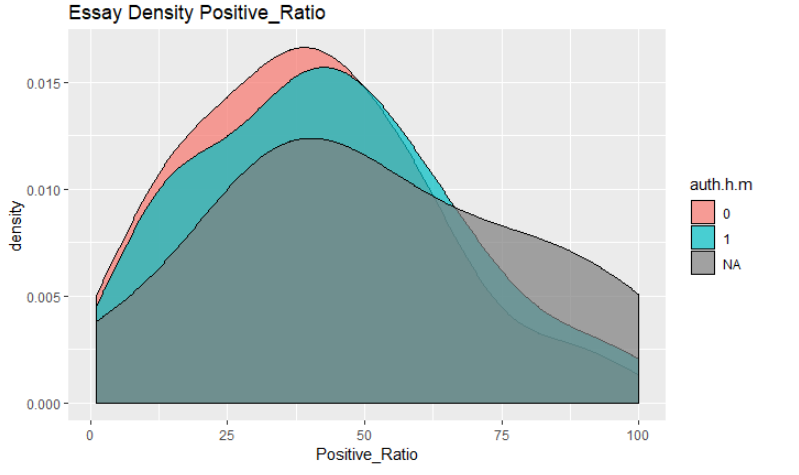


**Positive ratio density curve**



**Fig.3.5 Positive ratio density**

**Negative ratio density curve**



**Fig.3.6 Negative ratio density**

## **Conclusion**

**Hamilton papers**

Hamilton papers are significant in numbers when compared with other authors. Most of the Hamilton papers exhibit the same pattern except for few that are distinct from the general pattern. This shows that Hamilton has got a similar writing pattern and the variation on his essays are not very large.

**Madison papers**

Madison papers are not very large in number like Hamilton. With fewer papers compared to Hamilton, Madison papers are falling into two to three patterns. This indicates that Madison had a very different writing pattern in each of his essay showing more considerable variations when compared with all his essays.

**Dispute papers**

Dispute paper shows a similar pattern like Hamilton papers. All the dispute papers when compared with each other shows fewer variations and can be grouped with the writing style of Hamilton. It didn’t show more significant differences like Madison papers within its group. Also, the writing style of dispute papers are very different from Madison style and couldn’t group them under the same bucket of Madison papers. This suggests more inclination towards the authorship of Hamilton and not Madison