

Computer Engineering Department

A.P. Shah Institute of Technology

— G.B.Road, Kasarvadavli, Thane(W), Mumbai-400615

UNIVERSITY OF MUMBAI

Academic Year 2019-2020

A Project Report on
Artificial Teaching Assistant
Submitted in partial fulfillment of the degree of
Bachelor of Engineering(Sem-7)
in

Computer Engineering

By

Nirish Samant(16102008)

Akshay Sumbhe(16102062)

Tarakeshwarnath Sahani(16102033)

Under the Guidance of
Prof.Sachin Takmare

1. Project Conception and Initiation

1.1 Abstract

- It is an Artificial Teaching Assistant that has been developed with main aim to assist teachers in their teaching process.
- It is designed to work for english language although it might work on some other language depending on the nature of the language, the structure, grammar and semantics.
- This project i.e. Artificial Teaching Assistant is as the name implies is the program developed for teaching assistance by using various techniques from artificial intelligence and image processing in order to reduce the teaching efforts in many auxiliary but time consuming teaching tasks so that teachers can focus more on the core knowledge teaching aspects.

1.2 Objectives

- As educational AI progresses and becomes more sophisticated, it's essential to include teachers in the process. In an ideal world, teachers and AI will create an immersive learning experience for students, together.
- Develop the AI programs to do many auxiliary teaching tasks like improving the quality of available images and notes, automatically corrects the spellings in the document, grammar correction, auto answering the key questions on the given contents, sentiment analysis, text summarization etc.

1.3 Literature Review

We referred various books and papers to understand various concepts related to Artificial Intelligence, Image Processing , Signal Processing , Machine Learning etc

Some of the books are :

- 1) Digital Image Processing by Rafael C. Gonzalez and Richard E. Woods
- 2) Digital Image Processing by S. Sridhar
- 3) Speech and Language Processing by Daniel Jurafsky and James H. Martin
- 4) Foundations of Statistical Natural Language Processing by Christopher D. Manning and Hinrich Schütze
- 5) Natural Language Processing with Python by Steven Bird and Ewan Klein
- 6) Artificial Intelligence A Modern Approach , Stuart J. Russell and Peter Norvig
- 7) Artificial Intelligence and Intelligent Systems , N.P. Padhy
- 8) Neural Networks, Fuzzy Logic and Genetic Algorithms , S. Rajasekaran and G.A. Pai
- 9) Fuzzy Set Theory and its Applications , Zimmermann
- 10) Digital Signal Processing John G. Proakis, Dimitris and G. Manolakis
- 11) Digital Signal Processing , A. Anand Kumar
- 12) Artificial Intelligence , Elaine Rich and Kevin Knight

1.5 Scope

- The different AI techniques available like natural language processing, image processing will be used to perform these tasks.
- Image processing technique is used for improving image quality in the available teaching material.
- Teaching notes prepared can be checked for spelling and grammar. NLP techniques are used for developing the program code to these tasks.
- The feedbacks received from students for any lecture can be analysed by the sentiment analysis program. Attempt has been made to develop an sentiment analysis program to perform aspect level of teaching staff feedback.
- Text summarisation program will automatically condense the important information in the form of summary which would benefit the teachers.

1.6 Technology stack

- Python

Python is an interpreted, high-level, general-purpose programming language. Its language constructs and object-oriented approach aim to help programmers write clear, logical code for small and large scale projects.

- Open CV

OpenCV (Open source computer vision) is a library of programming functions mainly aimed at real time computer vision. The library is cross-platform and free for use under the open source BSD license.

- NLTK

The Natural Language Toolkit, or more commonly NLTK, is a suite of libraries and programs for symbolic and statistical natural language processing for English written in the Python programming language. NLTK is a leading platform for building Python programs to work with human language data..

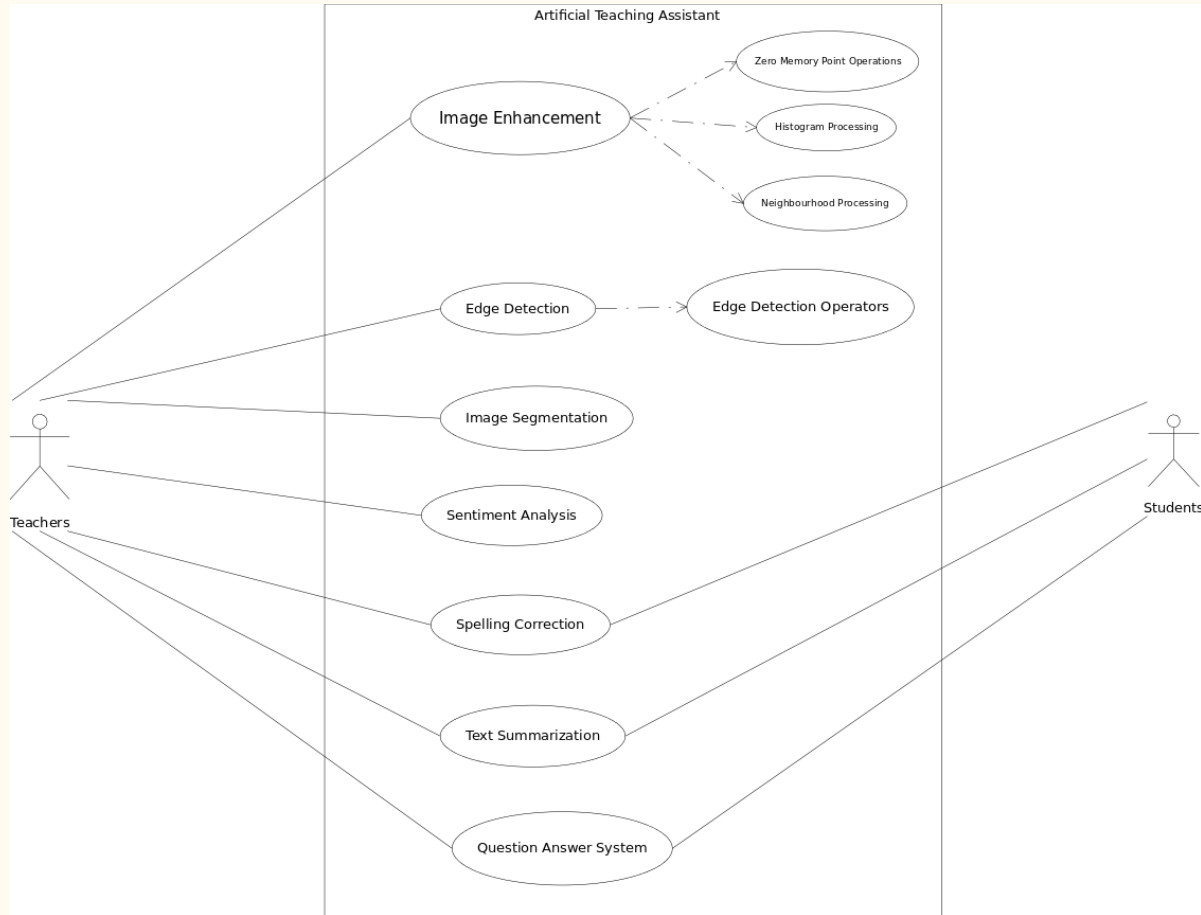
1.7 Benefits for environment & Society

- Reduced Carbon Footprint
- Access to Documents from Everywhere
- Keeps Things Simple
- Reduced Costs
- Increased output and productivity
- Increased quality
- Reliability

2. Project Design

—

Use Case Diagram



Digital Negative / Inversion

In "Preprocess" panel, click "open file" and choose the ARFF file format.

Pre Process Panel

- 1) Current relation pane: Specification of data file
- 2) Attributes and Remove button to possibly remove some attributes from experiment
- 3) Statistics about Selected attribute
- 4) Histogram shows the distribution of class as the function of selected attribute.
- 5) Edit button to edit input data on a separate window.

Build a Naive Bayes classifier

- 1) Switch to "classify" tab.
- 2) Select "Bayes" - by clicking "choose" button - Select classifier >> Bayes >> Naive Bayes.

- Invoke classifier by clicking start button

Result Classifier Output

- 1) Summary of the data set
- 2) Correctly classified Instances
- 3) Incorrectly classified Instances
- 4) Total Number of Instances
- 5) Detailed accuracy by class
- 6) Confusion matrix
- 7) Some other statistics

Conclusion: we have successfully implemented Naive Bayes Algorithm.

In "Preprocess" panel, click "open file" and choose the ARFF file format.

Pre Process Panel

- 1) Current relation pane: Specification of data file
- 2) Attributes and Remove button to possibly remove some attributes from experiment
- 3) Statistics about Selected attribute
- 4) Histogram shows the distribution of class as the function of selected attribute.
- 5) Edit button to edit input data on a separate window.

Build a Naive Bayes classifier

- 1) Switch to "classify" tab.
- 2) Select "Bayes" - by clicking "choose" button - Select classifier >> Bayes >> Naive Bayes.

- Invoke classifier by clicking start button

Result Classifier Output

- 1) Summary of the data set
- 2) Correctly classified Instances
- 3) Incorrectly classified Instances
- 4) Total Number of Instances
- 5) Detailed accuracy by class
- 6) Confusion matrix
- 7) Some other statistics

Conclusion: we have successfully implemented Naive Bayes Algorithm.

Thresholding

In "Preprocess" panel, click "open file" and choose the ARFF file format.

Pre Process Panel

- 1) Current relation pane: Specification of data file
- 2) Attributes and Remove button to possibly remove some attributes from experiment
- 3) Statistics about Selected attribute
- 4) Histogram shows the distribution of class as the function of selected attribute.
- 5) Edit button to edit input data on a separate window.

Build a Naive Bayes classifier

- 1) Switch to "classify" tab
- 2) Select "Bayes" - by clicking "choose" button - Select classifiers >> Bayes >> Naive Bayes.

- Invoke classifier by clicking start button

Result Classifier Output

- 1) Summary of the data set
- 2) Correctly classified Instances
- 3) Incorrectly classified Instances
- 4) Total Number of Instances
- 5) Detailed accuracy by class
- 6) Confusion Matrix
- 7) Some other statistics

Conclusion: we have successfully implemented Naive Bayes Algorithm.

In "Preprocess" panel, click "open file" and choose the ARFF file format.

Pre Process Panel

- 1) Current relation pane: Specification of data file
- 2) Attributes and Remove button to possibly remove some attributes from experiment
- 3) Statistics about Selected attribute
- 4) Histogram shows the distribution of class as the function of selected attribute.
- 5) Edit button to edit input data on a separate window.

Build a Naive Bayes classifier

- 1) Switch to "classify" tab
- 2) Select "Bayes" - by clicking "choose" button - Select classifiers >> Bayes >> Naive Bayes.

- Invoke classifier by clicking start button

Result Classifier Output

- 1) Summary of the data set
- 2) Correctly classified Instances
- 3) Incorrectly classified Instances
- 4) Total Number of Instances
- 5) Detailed accuracy by class
- 6) Confusion Matrix
- 7) Some other statistics

Conclusion: we have successfully implemented Naive Bayes Algorithm.

Intensity Level Slicing without Background

In "Preprocess" panel, click "open file" and choose the ARFF file format.

Pre Process Panel

- 1) Current relation pane: Specification of data file
- 2) Attributes and Remove button to possibly remove some attributes from experiment
- 3) Statistics about Selected attribute
- 4) Histogram shows the distribution of class as the function of selected attribute.
- 5) Edit button to edit input data on a separate window.

Build a Naive Bayes classifier

- 1) Switch On to "classify" tab.
- 2) Select "Bayes" - by clicking "choose" button - Select classifiers >> Bayes >> Naive Bayes.

- Invoke classifier by clicking start button

Result Classifier output

- 1) Summary of the dat set
- 2) Correctly classified Instances
- 3) Incorrectly classified Instances
- 4) Total Number of Instances
- 5) Detailed accuracy by class
- 6) Confusion Matrix
- 7) Some other Statistics

Conclusion: we have successfully implemented Naive Bayes Algorithm.

In "Preprocess" panel, click "open file" and choose the ARFF file format.

Pre Process Panel

- 1) Current relation pane: Specification of data file.
- 2) Attributes and Remove button to possibly remove some attributes from experiment
- 3) Statistics about Selected attribute
- 4) Histogram shows the distribution of class as the function of selected attribute.
- 5) Edit button to edit input data on a separate window.

Build a Naive Bayes classifier

- 1) Switch On to "classify" tab.
- 2) Select "Bayes" - by clicking "choose" button - Select classifiers >> Bayes >> Naive Bayes.

- Invoke classifier by clicking start button

Result Classifier output

- 1) Summary of the dat set
- 2) Correctly classified Instances
- 3) Incorrectly classified Instances
- 4) Total Number of Instances
- 5) Detailed accuracy by class
- 6) Confusion Matrix
- 7) Some other Statistics

Conclusion: we have successfully implemented Naive Bayes Algorithm.

Intensity Level Slicing with Background

In "Preprocess" panel, click "open file" and choose the ARFF file format.

Pre Process Panel

- 1) Current relation pane: Specification of data file
- 2) Attributes and Remove button to possibly remove some attributes from experiment
- 3) Statistics about Selected attribute
- 4) Histogram shows the distribution of class as the function of selected attribute.
- 5) Edit button to edit input data on a separate window.

Build a Naive Bayes classifier

- 1) Switch to "classify" tab
 - 2) Select "Bayes" - by clicking "choose" button - Select classifiers >> Bayes >> Naive Bayes
- Invoke classifier by clicking start button

Result Classifier output

- 1) Summary of the dat set
- 2) Correctly classified Instances
- 3) Incorrectly classified Instances
- 4) Total Number of Instances
- 5) Detailed accuracy by class
- 6) Confusion Matrix
- 7) Some other statistics

Conclusion: we have successfully implemented Naive Bayes Algorithm.

In "Preprocess" panel, click "open file" and choose the ARFF file format.

Pre Process Panel

- 1) Current relation pane: Specification of data file
- 2) Attributes and Remove button to possibly remove some attributes from experiment
- 3) Statistics about Selected attribute
- 4) Histogram shows the distribution of class as the function of selected attribute.
- 5) Edit button to edit input data on a separate window.

Build a Naive Bayes classifier

- 1) Switch to "classify" tab
 - 2) Select "Bayes" - by clicking "choose" button - Select classifiers >> Bayes >> Naive Bayes
- Invoke classifier by clicking start button

Result Classifier output

- 1) Summary of the dat set
- 2) Correctly classified Instances
- 3) Incorrectly classified Instances
- 4) Total Number of Instances
- 5) Detailed accuracy by class
- 6) Confusion Matrix
- 7) Some other statistics

Conclusion: we have successfully implemented Naive Bayes Algorithm.

Bit Plane Slicing

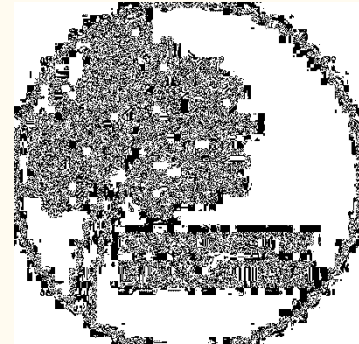
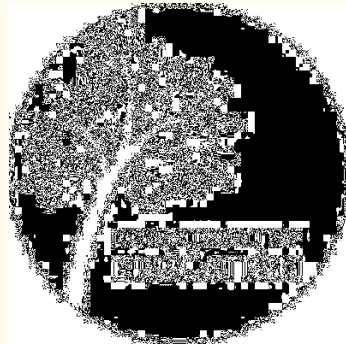
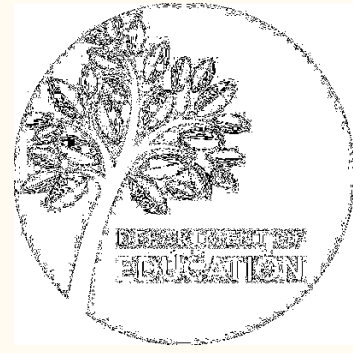


Image Text Colouration

In "Preprocess" panel, click "open file" and choose the ARFF file format.

Pre Process Panel

- 1) Current relation pane: Specification of data file
- 2) Attributes and Remove button to possibly remove some attributes from experiment
- 3) Statistics about selected attribute
- 4) Histogram shows the distribution of class as the function of selected attribute
- 5) Edit button to edit input data on a separate window.

Build a Naïve Bayes classifier

- 1) Switch to "classify" tab
- 2) Select "Bayes" - by clicking "choose" button - Selected classifiers >> Bayes >> Naive Bayes

- Invoke classifier by clicking start button

Result Classifier output

- 1) Summary of the data set
- 2) Correctly classified Instances
- 3) Incorrectly classified Instances
- 4) Total Number of Instances
- 5) Detailed accuracy by class
- 6) Confusion Matrix
- 7) Some other statistics

Conclusion: we have successfully implemented Naive Bayes Algorithm.

In "Preprocess" panel, click "open file" and choose the ARFF file format.

Pre Process Panel

- 1) Current relation pane: Specification of data file
- 2) Attributes and Remove button to possibly remove some attributes from experiment
- 3) Statistics about selected attribute
- 4) Histogram shows the distribution of class as the function of selected attribute
- 5) Edit button to edit input data on a separate window.

Build a Naïve Bayes classifier

- 1) Switch to "classify" tab
- 2) Select "Bayes" - by clicking "choose" button - Selected classifiers >> Bayes >> Naive Bayes

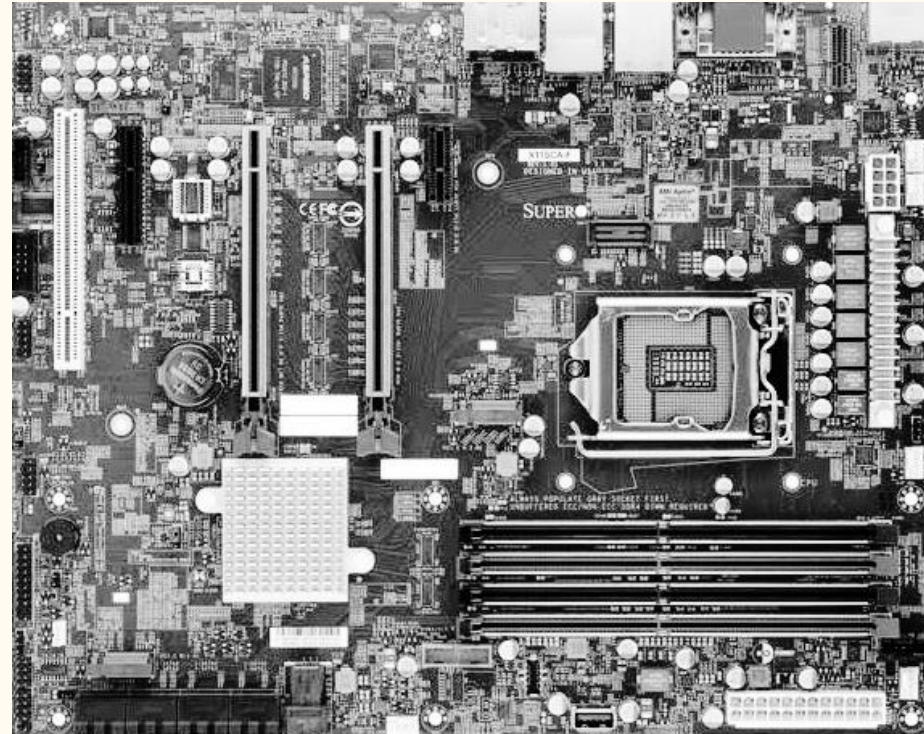
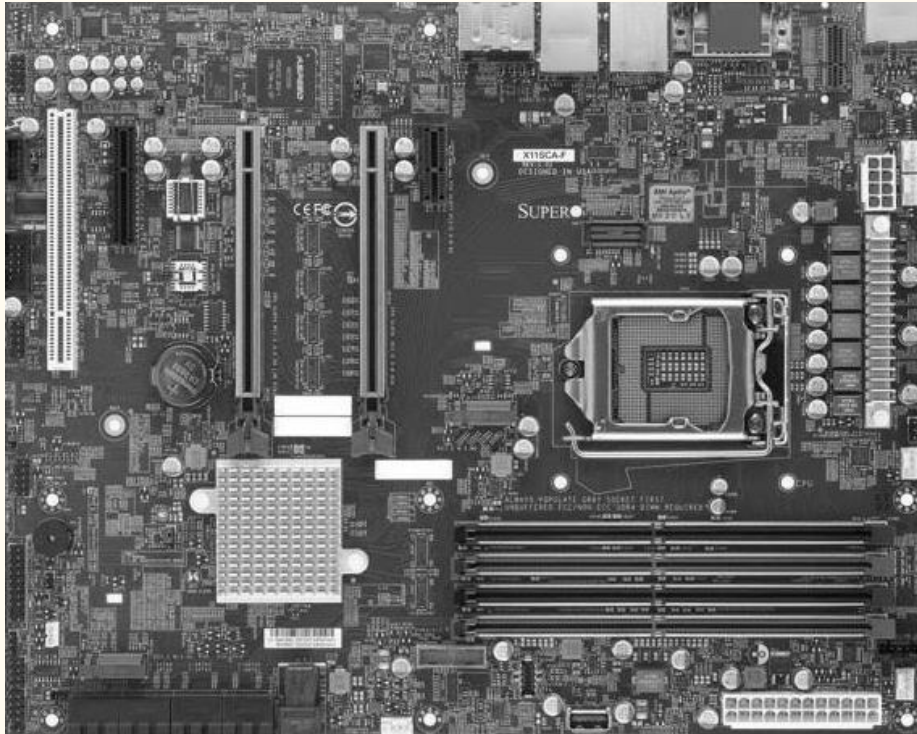
- Invoke classifier by clicking start button

Result Classifier output

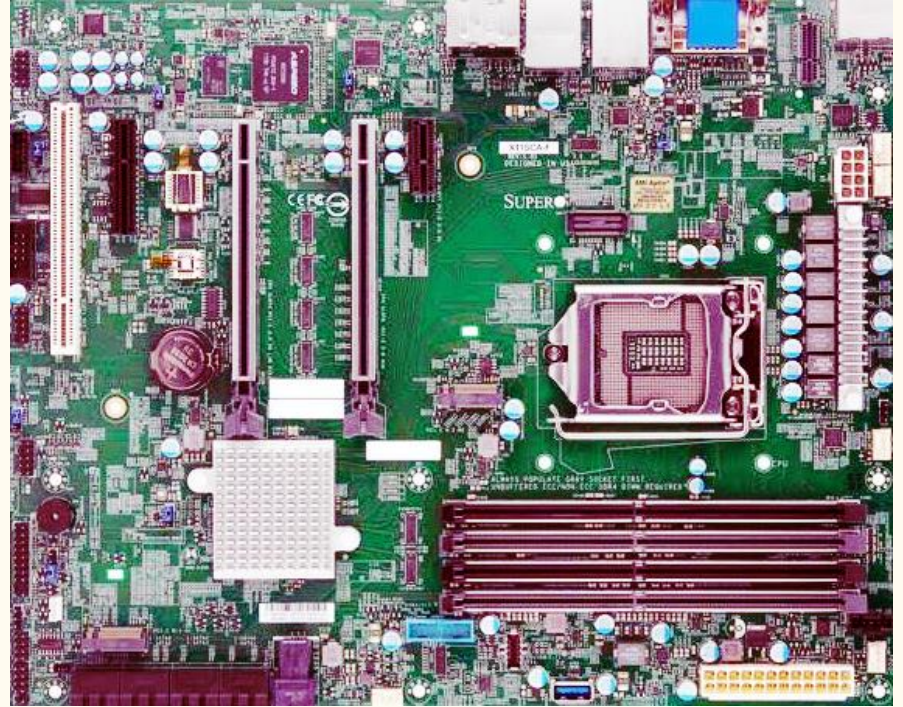
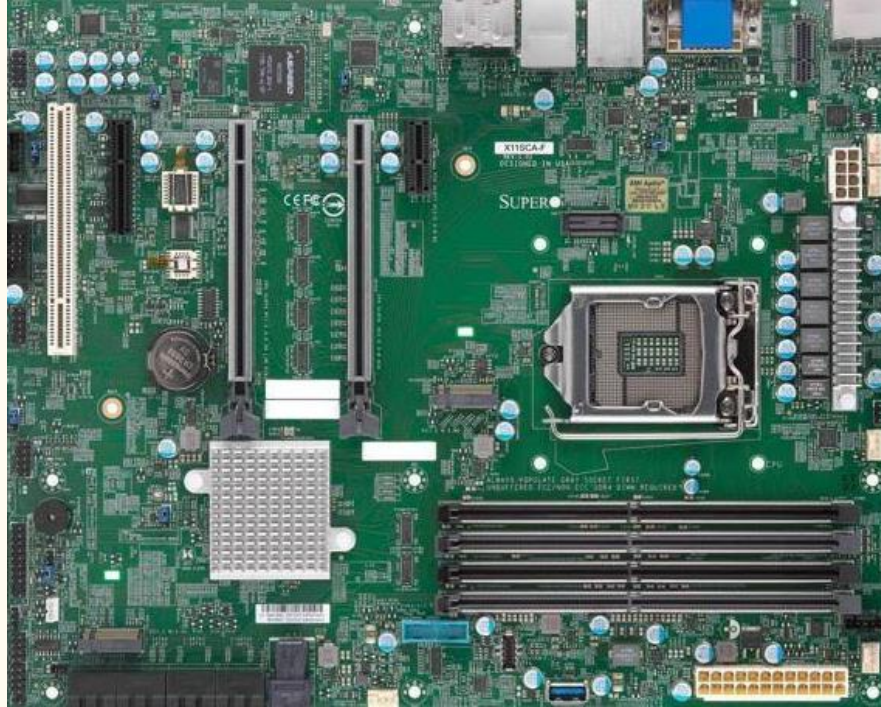
- 1) Summary of the data set
- 2) Correctly classified Instances
- 3) Incorrectly classified Instances
- 4) Total Number of Instances
- 5) Detailed accuracy by class
- 6) Confusion Matrix
- 7) Some other statistics

Conclusion: we have successfully implemented Naive Bayes Algorithm.

Histogram Equalization



Histogram Equalization



Histogram Equalization

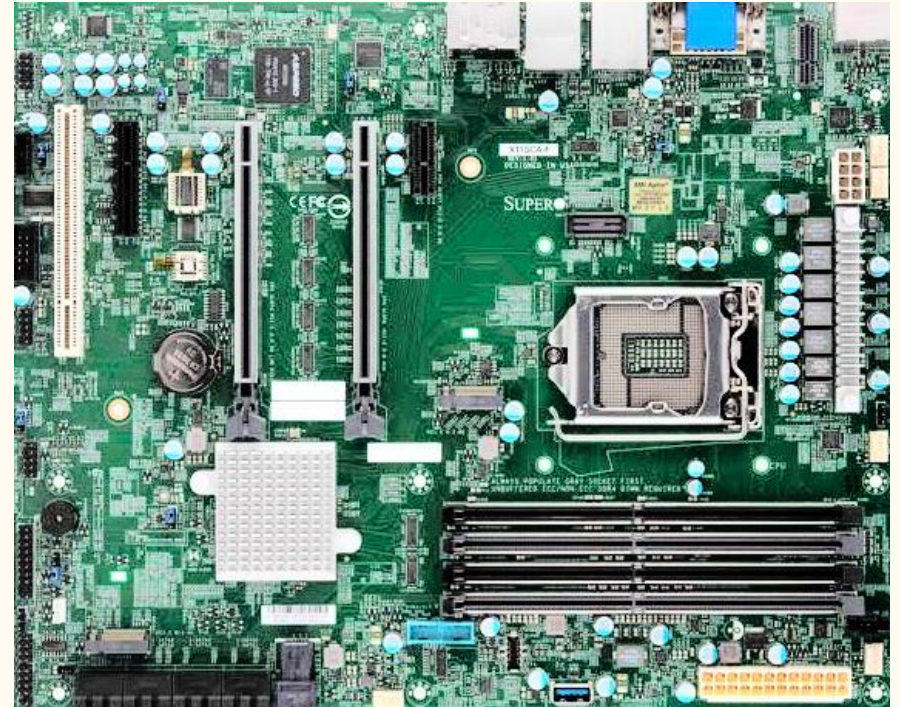


Image Smoothing

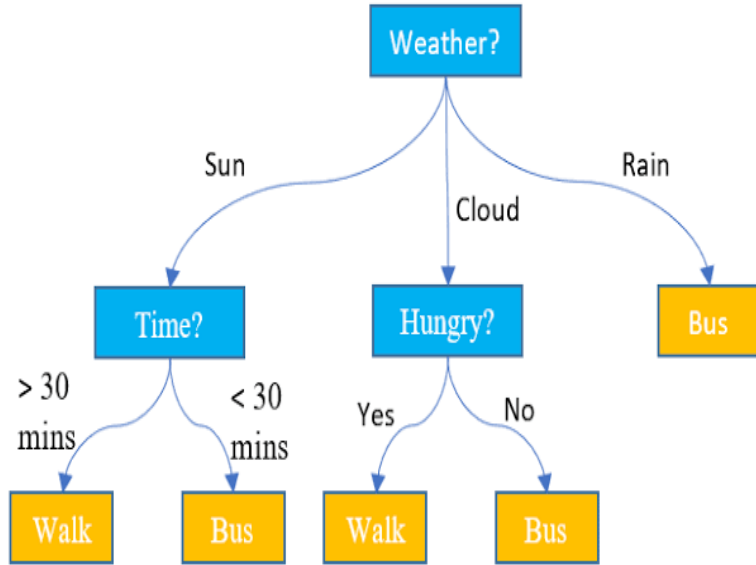


Image Sharpening

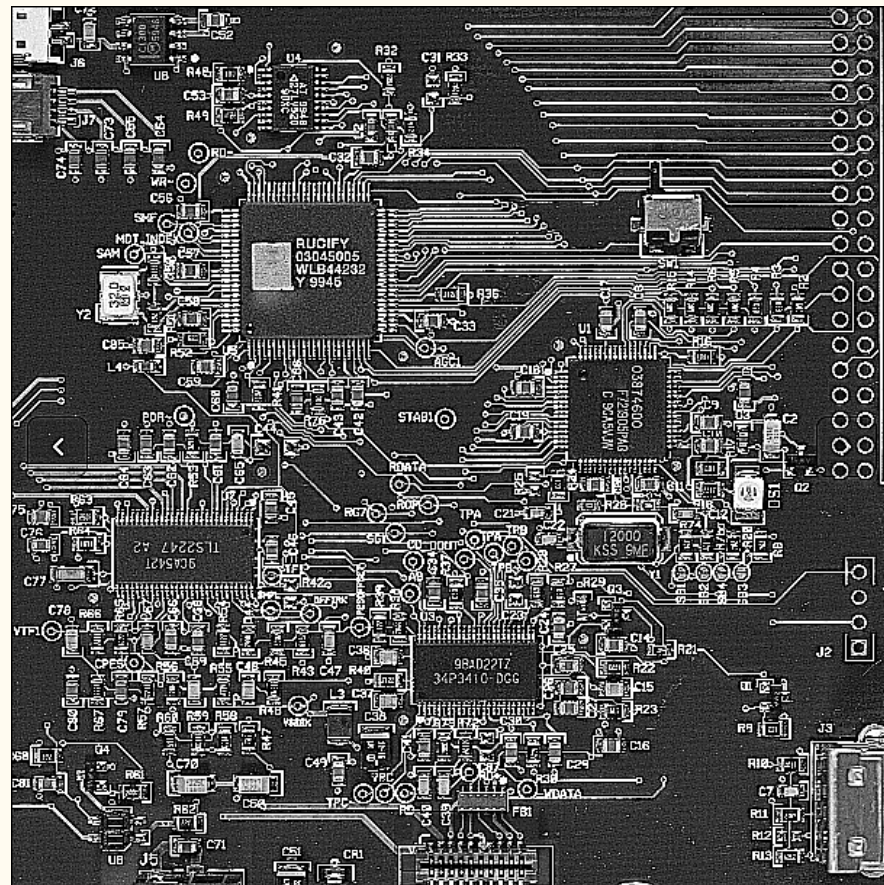
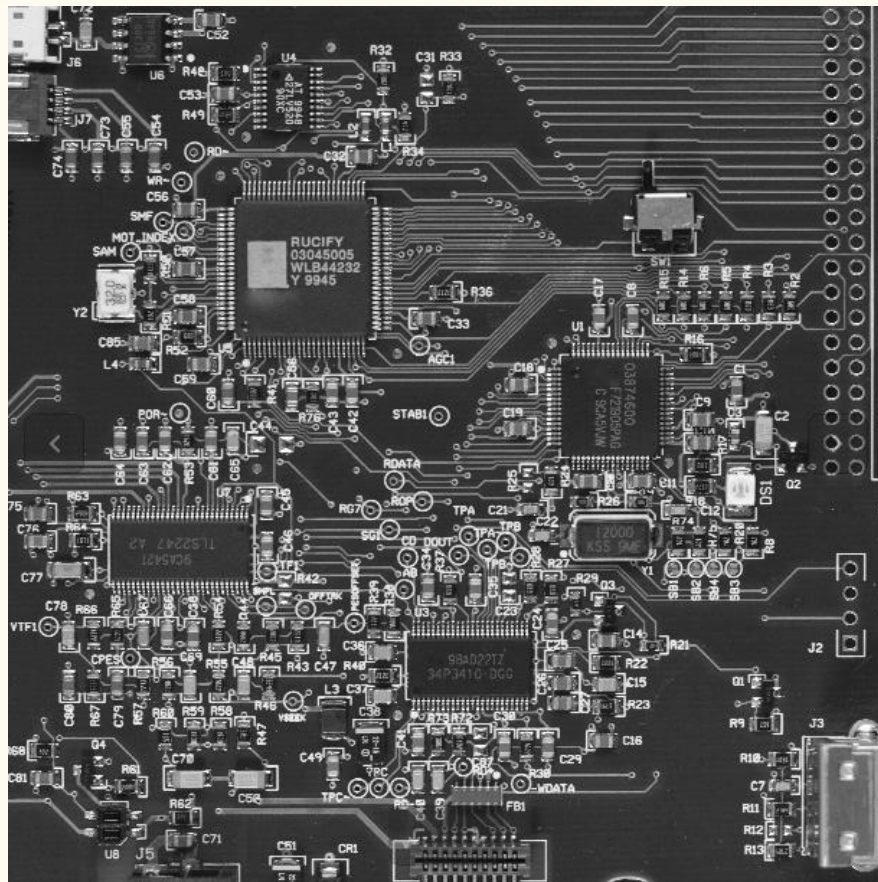


Image Sharpening

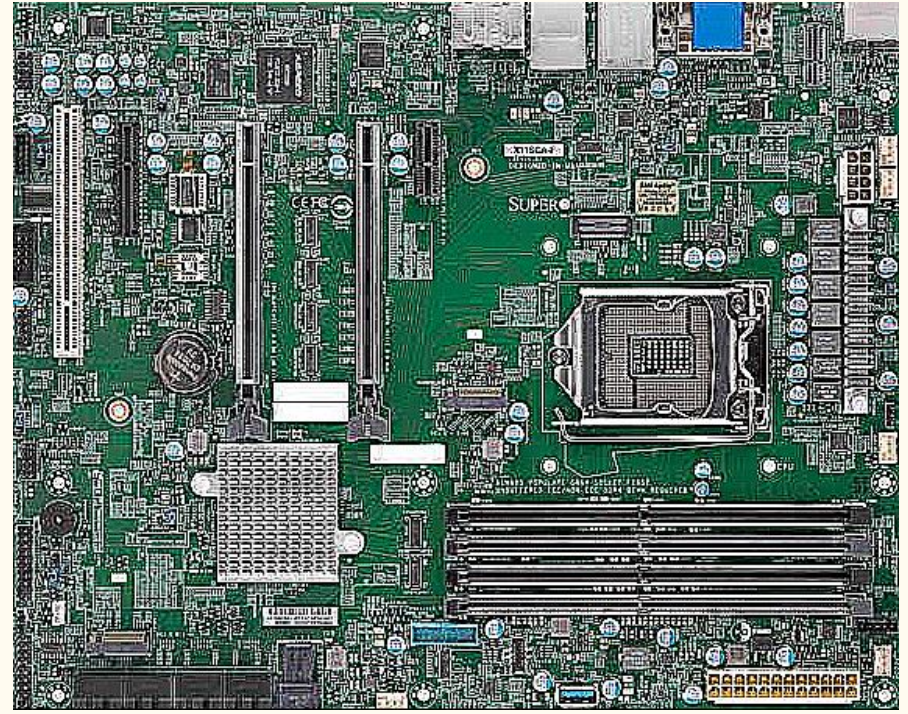


Image Sharpening

In "Preprocess" panel, click "open file" and choose the ARFF file format.

Pre Process Panel

- 1) Current relation pane : Specification of data file
- 2) Attributes and Remove button to possibly remove some attributes from experiment
- 3) Statistics about Selected attribute
- 4) Histogram shows the distribution of class as the function of Selected attribute.
- 5) Edit button to edit input data on a separate window.

Build a Naive Bayes classifier

- 1) Switch to "classify" tab
- 2) Select "Bayes" - by clicking "choose" button - Select classifiers >> Bayes >> Naive Bayes

- Invoke classifier by clicking start button

Result Classifier output

- 1) Summary of the dat set
- 2) Correctly classified Instances
- 3) Incorrectly classified Instances
- 4) Total Number of Instances
- 5) Detailed accuracy by class
- 6) Confusion Matrix
- 7) Some other statistics

Conclusion: we have successfully implemented Naive Bayes Algorithm.

In "Preprocess" panel, click "open file" and choose the ARFF file format.

Pre Process Panel

- 1) Current relation pane : Specification of data file
- 2) Attributes and Remove button to possibly remove some attributes from experiment
- 3) Statistics about Selected attribute
- 4) Histogram shows the distribution of class as the function of Selected attribute.
- 5) Edit button to edit input data on a separate window.

Build a Naive Bayes classifier

- 1) Switch to "classify" tab
- 2) Select "Bayes" - by clicking "choose" button - Select classifiers >> Bayes >> Naive Bayes

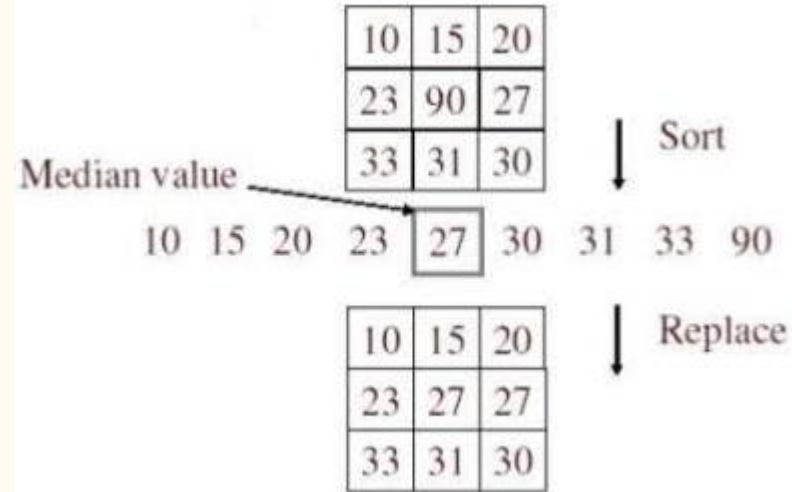
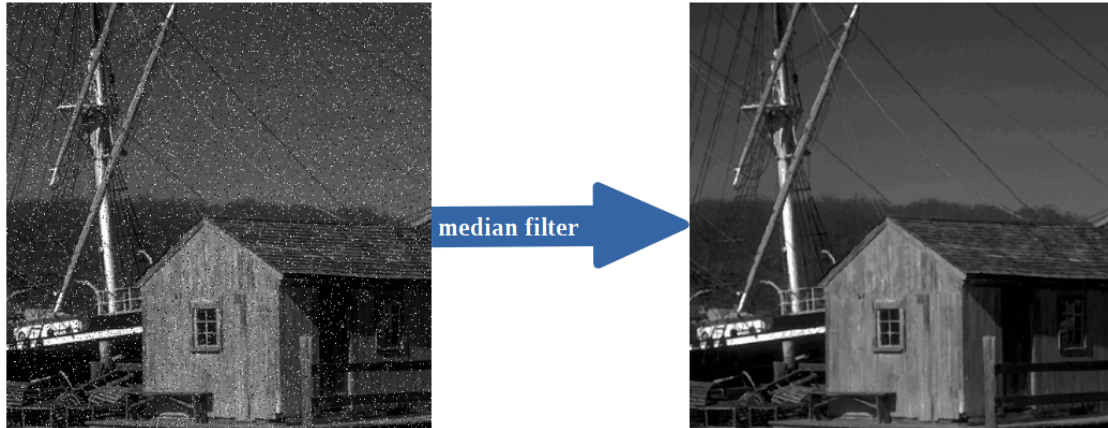
- Invoke classifier by clicking start button

Result Classifier output

- 1) Summary of the dat set
- 2) Correctly classified Instances
- 3) Incorrectly classified Instances
- 4) Total Number of Instances
- 5) Detailed accuracy by class
- 6) Confusion Matrix
- 7) Some other statistics

Conclusion: we have successfully implemented Naive Bayes Algorithm.

Median Filter



Median Filter



Median Filter



Edge Detection



Edge Detection



Preprocessing images



Sobel edge detection



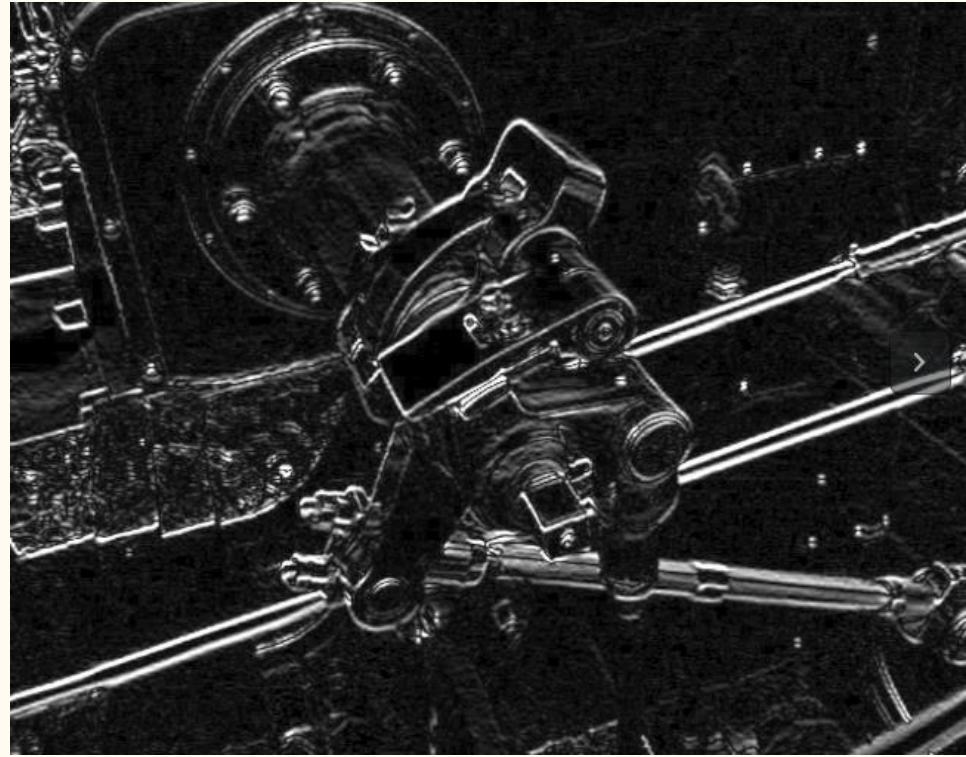
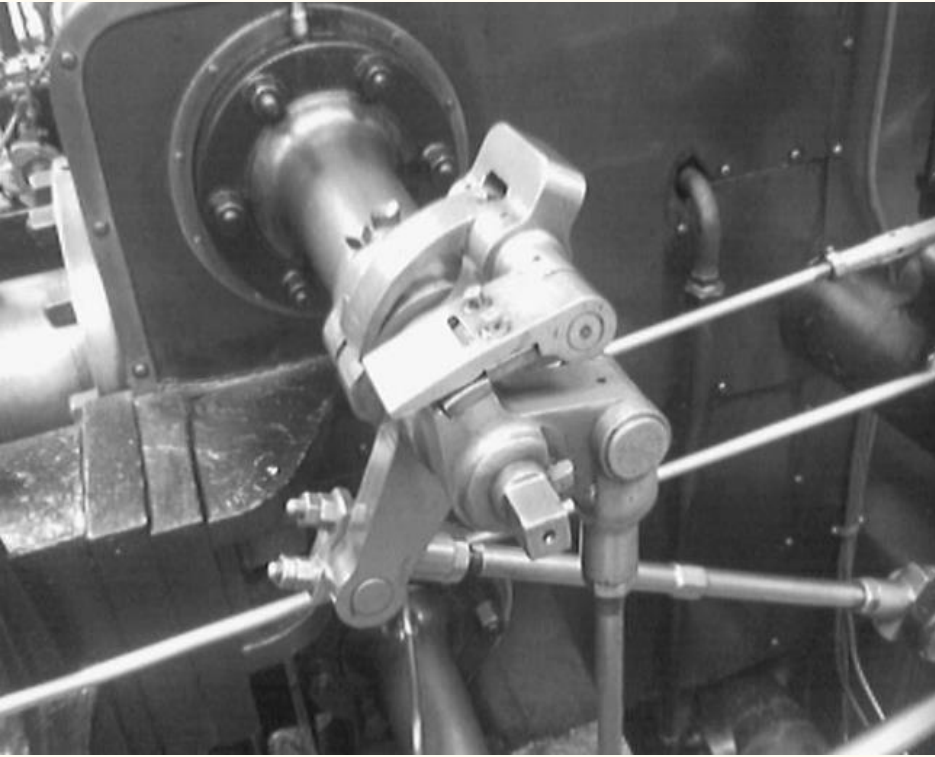
Robert edge detection



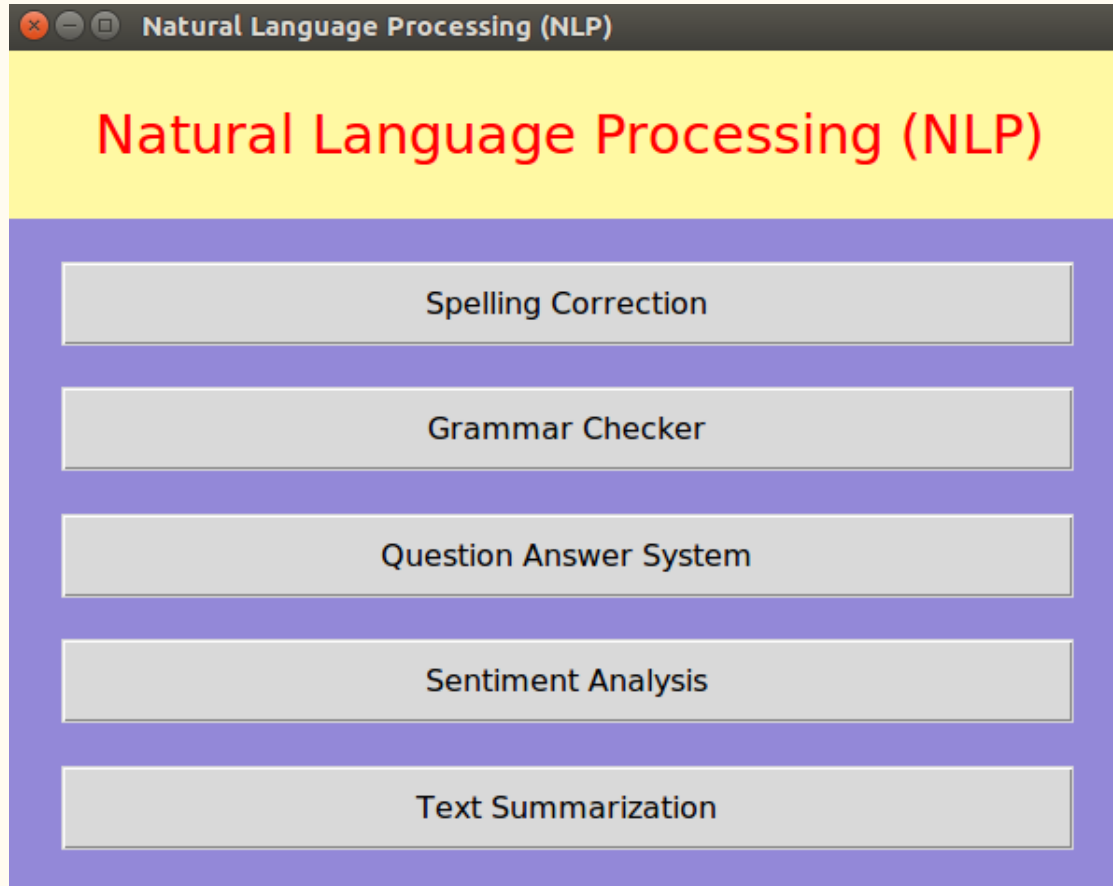
Prewitt edge detection



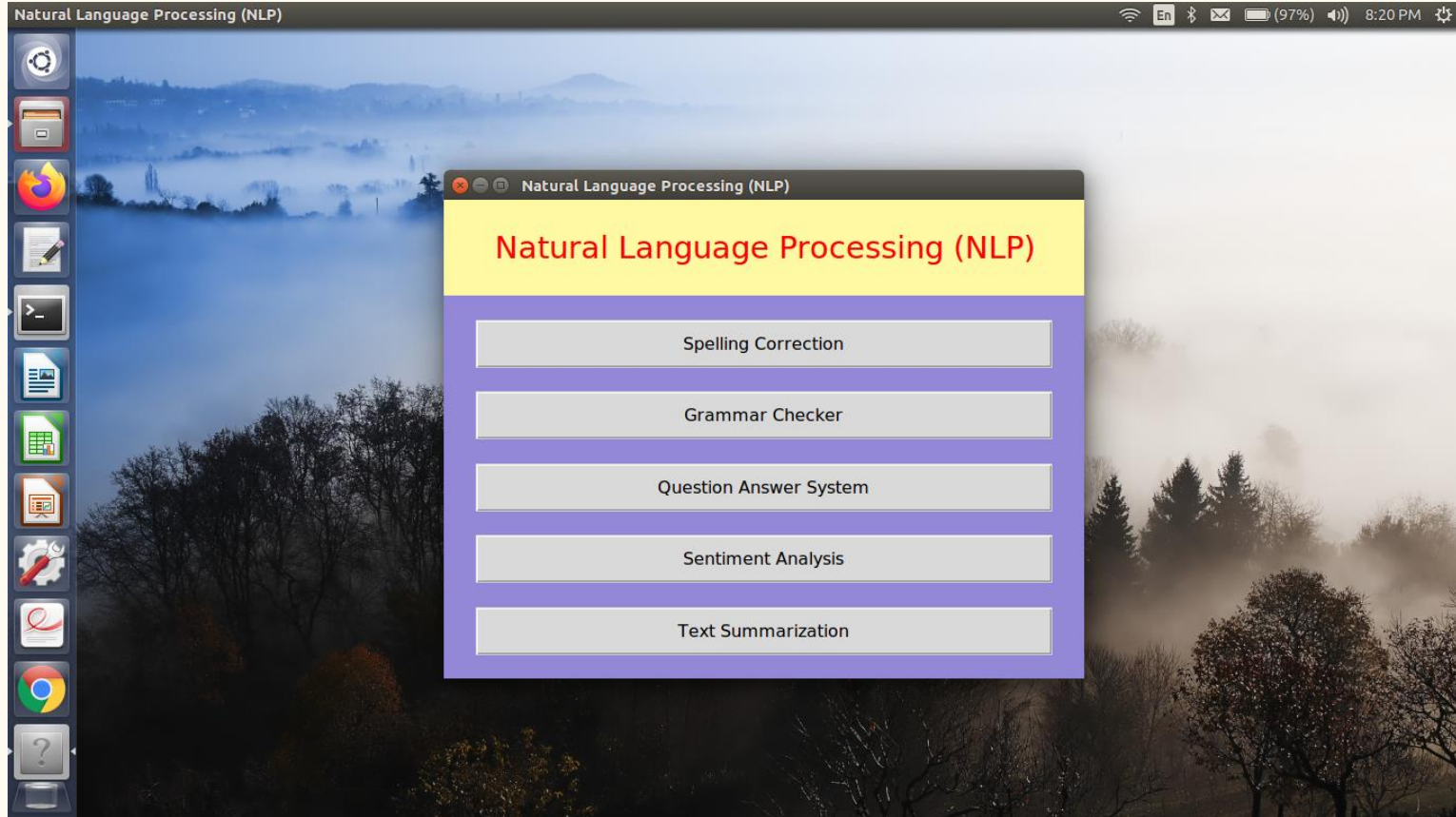
Edge Detection



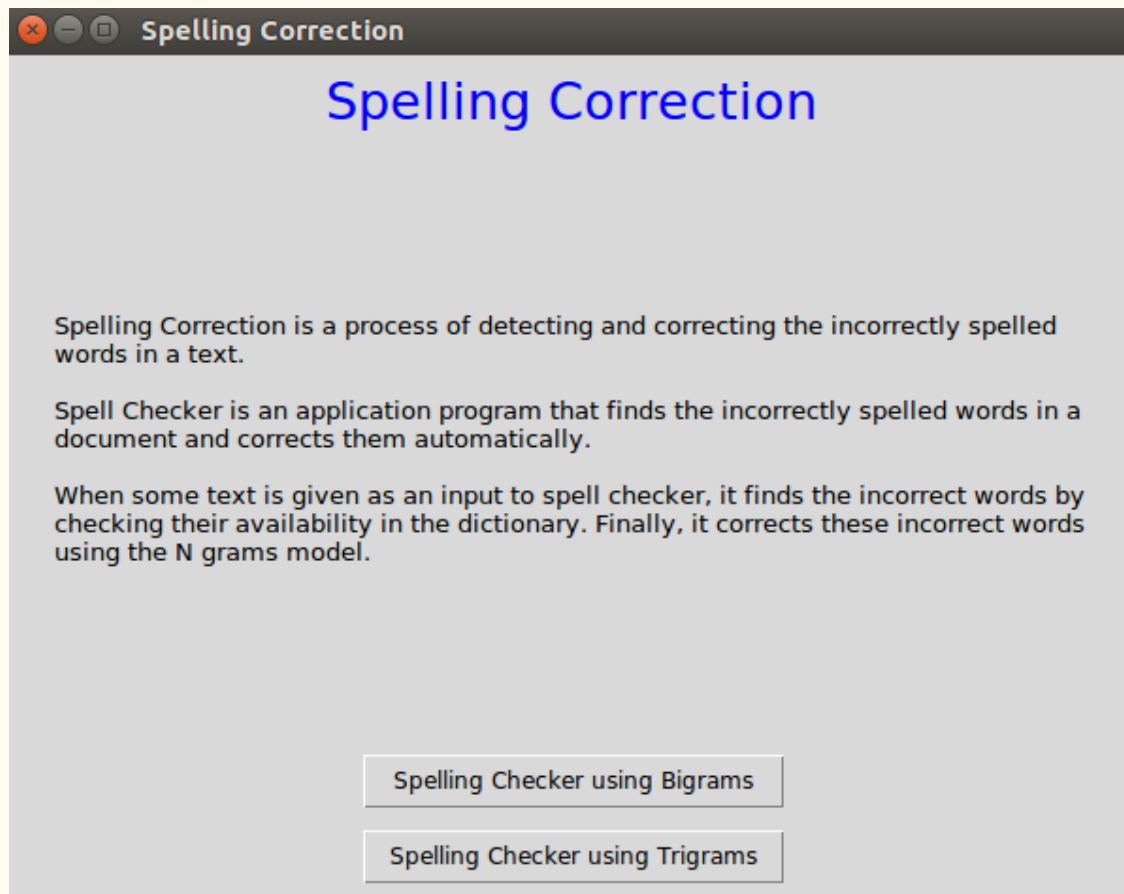
Natural Language Processing (NLP)




Natural Language Processing (NLP)




Spelling Correction



Spelling Correction

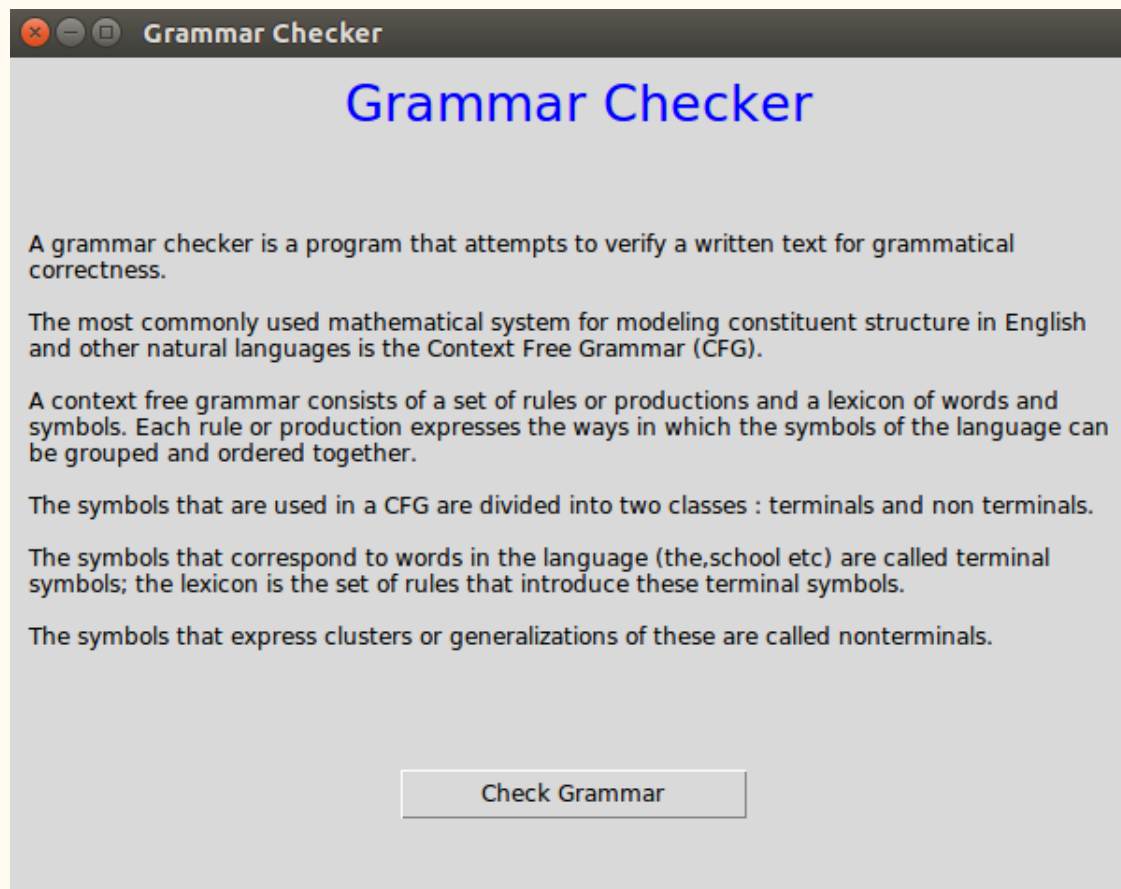
Open ▾  Save

```
1 The boy was going to scheel. The book was kept on the tablo. She was happa. He was playeng. He was helpeng.
```

Open ▾  Save

```
1 The boy was going to school . The book was kept on the table . She was happy . He was playing . He was helping .
```

Grammar Checker

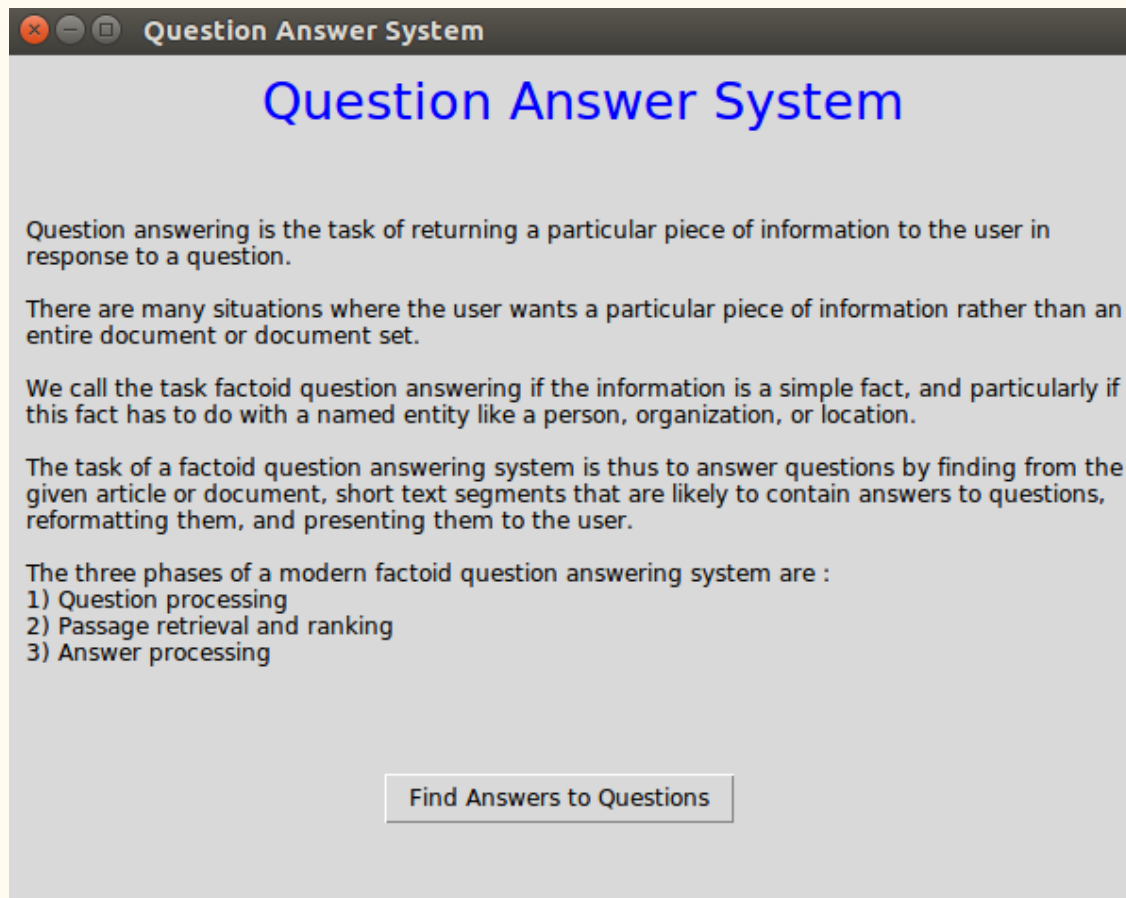


Grammar Checker


```
Open ▾ [icon] Save
1 I shot an apple in my pajamas. The man saw a dog in the park. The man a park the saw dog in. The angry bear chased the frightened little squirrel. The man a. Man dog park.
```

```
Open ▾ [icon] Save
1 *****Grammatically Correct Sentences*****
2
3 I shot an apple in my pajamas.
4 The man saw a dog in the park.
5 The angry bear chased the frightened little squirrel.
6
7 *****Grammatically Incorrect Sentences*****
8
9 The man a park the saw dog in.
10 The man a.
11 Man dog park.
12
13 *****Ambiguous Sentences*****
14
15 I shot an apple in my pajamas.
16 The man saw a dog in the park.
```

Question Answer System




Question Answer System

Open ▾ 


Save

```
1 Narendra Modi is the prime minister of India. Mango is very tasty. Mohan is the king of the jungle. Angela Merkel is the Chancellor of Germany.
```

Open ▾ 

Save

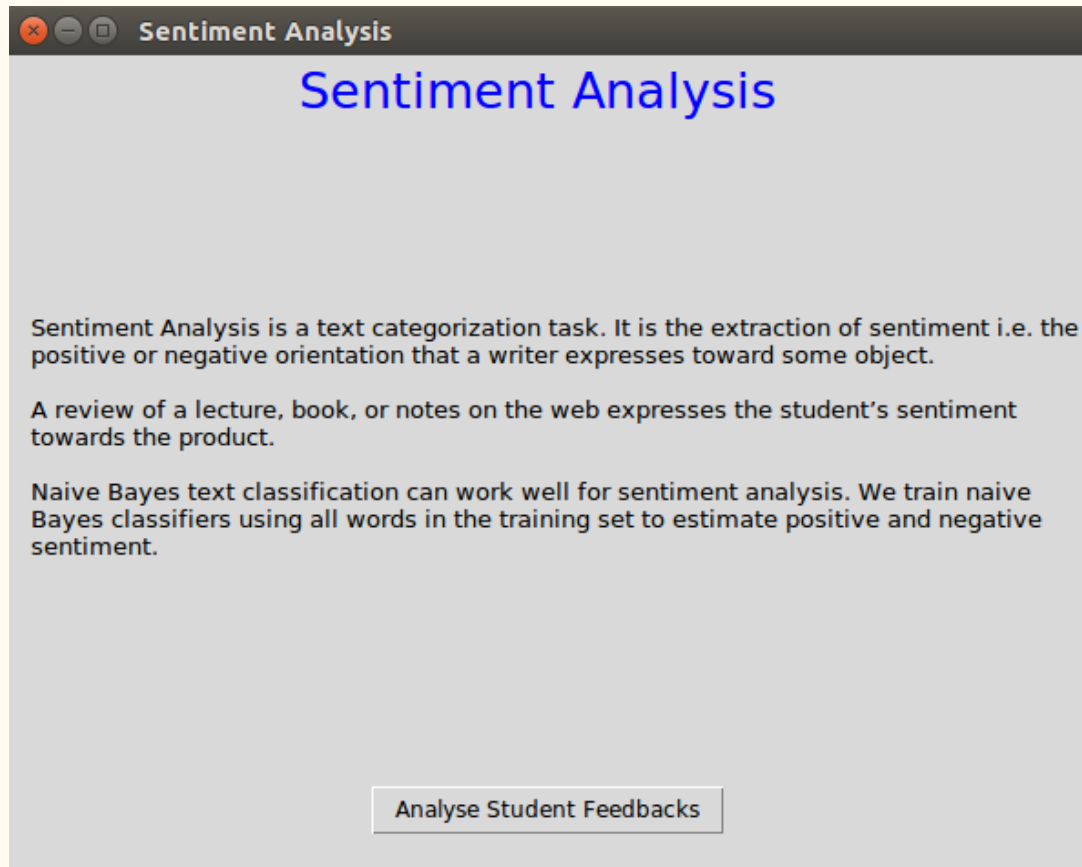
```
1 Who is the prime minister of India?
2 Who is the chancellor of Germany?
```

Open ▾ 

Save

```
1 Who is the prime minister of India?
2 Narendra Modi is
3
4 Who is the chancellor of Germany?
5 Angela Merkel
6
```

Sentiment Analysis



Sentiment Analysis

Open ▾



Save

```
1 The lecture was very good.
2 The lecture was bad.
3 I did not understand anything.
4 I liked the lecture.
5 The lecture was boring.
6 The lecture was excellent.
7 The lecture was not interactive.
```

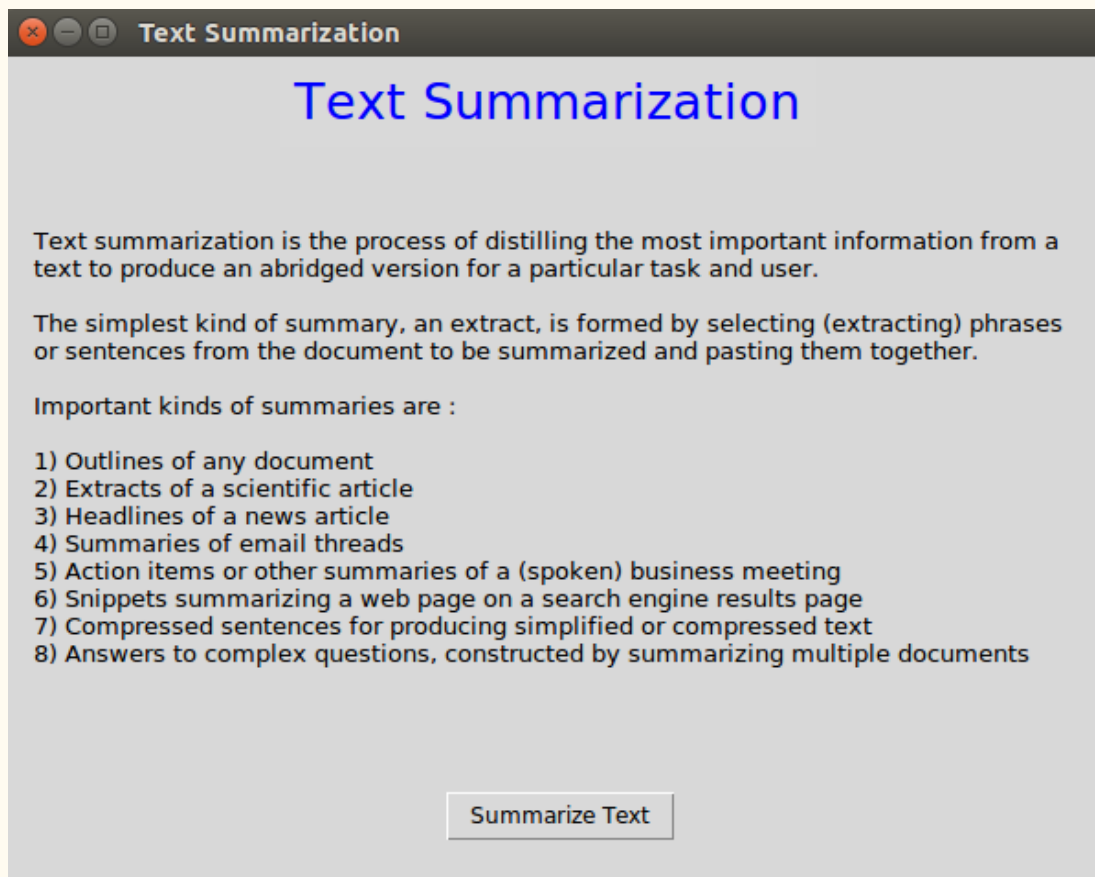
Open ▾



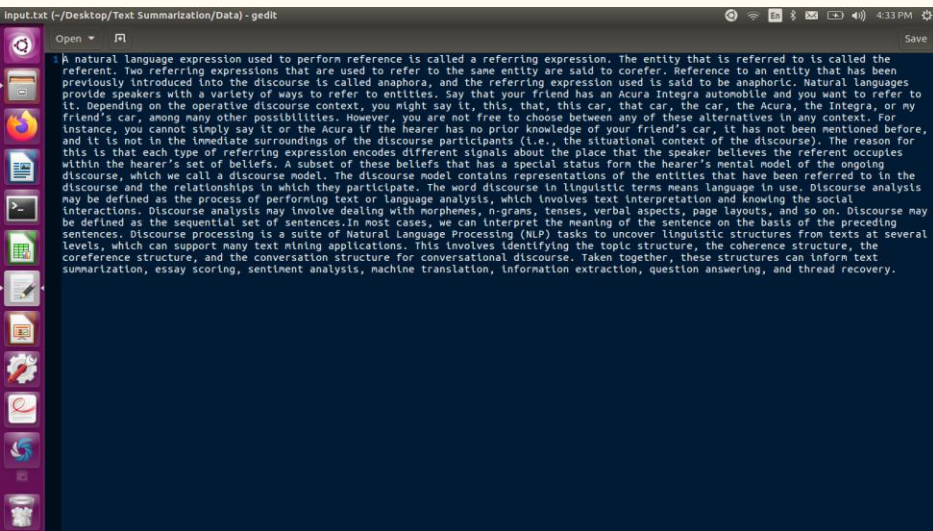
Save

```
1 Total no of Student Feedbacks : 7
2 Total no of Positive Student Feedbacks : 3
3 Total no of Negative Student Feedbacks : 4
4
5 *****Positive Student Feedbacks*****
6
7 The lecture was very good.
8 I liked the lecture.
9 The lecture was excellent.
10
11 *****Negative Student Feedbacks*****
12
13 The lecture was bad.
14 I did not understand anything.
15 The lecture was boring.
16 The lecture was not interactive.
```


Text Summarization

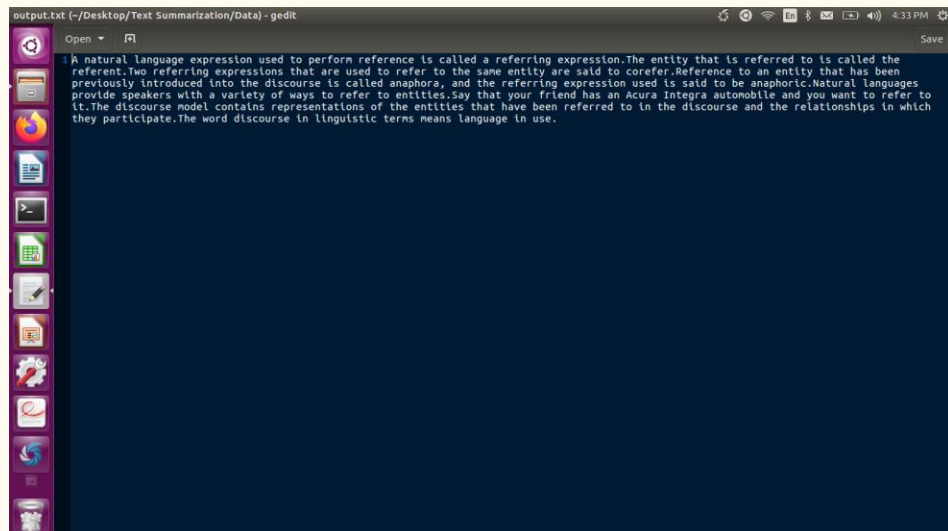


Text Summarization



```
input.txt (~/.Desktop/Text Summarization/Data) - gedit
Open  Save

A natural language expression used to perform reference is called a referring expression. The entity that is referred to is called the referent. Two referring expressions that are used to refer to the same entity are said to corefer. Reference to an entity that has been previously introduced into the discourse is called anaphora, and the referring expression used is said to be anaphoric. Natural languages provide speakers with a variety of ways to refer to entities. Say that your friend has an Acura Integra automobile and you want to refer to it. Depending on the operative discourse context, you might say it, this, that, this car, that car, the car, the Acura, the Integra, or my friend's car, among many other possibilities. However, you are not free to choose between any of these alternatives in any context. For instance, you cannot simply say it or the Acura if the hearer has no prior knowledge of your friend's car, it has not been mentioned before, and it is not in the immediate surroundings of the discourse participants (i.e., the situational context of the discourse). The reason for this is that each type of referring expression encodes different signals about the place that the speaker believes the referent occupies within the hearer's set of beliefs. A subset of these beliefs that has a special status form the hearer's mental model of the ongoing discourse, which we call a discourse model. The discourse model contains representations of the entities that have been referred to in the discourse and the relationships in which they participate. The word discourse in linguistic terms means language in use. Discourse analysts may be defined as the process of performing text or language analysis, which involves text interpretation and knowing the social interactions. Discourse analysis may involve dealing with morphemes, n-grams, tenses, verbal aspects, page layouts, and so on. Discourse may be defined as the sequential set of sentences. In most cases, we can interpret the meaning of the sentence on the basis of the preceding sentences. Discourse processing is a suite of Natural Language Processing (NLP) tasks to uncover linguistic structures from texts at several levels, which can support many text mining applications. This involves identifying the topic structure, the coherence structure, the coreference structure, and the conversation structure for conversational discourse. Taken together, these structures can inform text summarization, essay scoring, sentiment analysis, machine translation, information extraction, question answering, and thread recovery.
```



```
output.txt (~/.Desktop/Text Summarization/Data) - gedit
Open  Save

A natural language expression used to perform reference is called a referring expression. The entity that is referred to is called the referent. Two referring expressions that are used to refer to the same entity are said to corefer. Reference to an entity that has been previously introduced into the discourse is called anaphora, and the referring expression used is said to be anaphoric. Natural languages provide speakers with a variety of ways to refer to entities. Say that your friend has an Acura Integra automobile and you want to refer to it. The discourse model contains representations of the entities that have been referred to in the discourse and the relationships in which they participate. The word discourse in linguistic terms means language in use.
```

2.7 References

- 1) Digital Image Processing by Rafael C. Gonzalez and Richard E. Woods
- 2) Digital Image Processing by S. Sridhar
- 3) Speech and Language Processing by Daniel Jurafsky and James H. Martin
- 4) Foundations of Statistical Natural Language Processing by Christopher Manning and Hinrich Schutze
- 5) Natural Language Processing with Python by Steven Bird and Ewan Klein
- 6) Stephen Marsland, "Machine Learning An Algorithmic Perspective" CRC Press
- 7) Artificial Intelligence A Modern Approach , Stuart J. Russell and Peter Norvig
- 8) Artificial Intelligence and Intelligent Systems , N.P.Padhy
- 9) Artificial Intelligence , Elaine Rich and Kevin Knight
- 10) Tom M. Mitchell "Machine Learning" McGraw Hill
- 11) Fuzzy Set Theory and its Applications , Zimmermann
- 12) Introduction to soft computing , Samir Roy and Chakraborty
- 13) Principles of Soft Computing , S.N.Sivanandam, S.N.Deepa

Thank You

—