NLP Applications (AIMLCZG519)

Assignment 2 - Sentiment Analysis Application

Main Assignment Report

Group #37

Team Members

Ankit Kumar Agarwal - 2024aa05560

Chandrababu Yelamuri - 2024aa05820

Preety Gupta - 2023ac05892

Srithin Nair - 2024ab05197

Deepan KG - 2024aa05755

Table of Contents

Executive Summary

This document presents a comprehensive details related to the Assignment Part-A Task-A – “Simple application that can accurately analyze the sentiment (positive, negative, or neutral) of user-provided text.”, including:

1. Instructions for running the application locally.
2. Design choices and any challenges faced during implementation.
3. Screenshots that explain the entire flow of the application.

1 Local Execution Instructions

Follow below steps to set up and run the application on your local machine:

1. **Prerequisites**
   - Python 3.8 or higher installed.
   - Modern web browser (Chrome, Firefox, Edge).

2. **Step 1: Install Dependencies**

   Open your terminal in the project directory and run the below command:

   ```
   pip install -r requirements.txt
   ```

   ***Note**: This will install Flask, NLTK, TextBlob, and security extensions like JWT and Limiter.*

3. **Step 2: Start the Backend Server**

   Run the Flask application:

   ```
   python app.py
   ```

   Wait for the message: ** Running on http://127.0.0.1:5000*.

   ***Note**: On the first run, the app will automatically download necessary NLTK corpora (VADER, Tokenizers, WordNet).*

4. **Step 3: Launch the Frontend**

   You can launch the user interface in two ways:

   1. **Directly:** Right-click *index.html* and "Open with Browser".
   2. **Local Server (Recommended):** Visit the site *http://127.0.0.1:5000*.

2 Design Choices & Challenges

2.1 Design Choices

1. **Hybrid Sentiment Engine**

   Instead of relying on a single analyzer, we implemented a **Hybrid Decision Logic**:

   - **VADER:** Used for its strength in handling social media text, emojis, and intensity (e.g., "GREAT!!!" vs "great").
   - **TextBlob:** Used to validate results through polarity and subjectivity scoring.
   - **Custom Weighting:** We added a keyword-weighing layer (e.g., giving extra "Negative" weight to words like "disappointed" or "waste") to handle edge cases where statistical models often miss.

2. **UI Layout (Optimized 3-Column Dashboard)**

We chose a fixed-height, 3-column grid layout that is visible on a single screen:

- **Input (Left):** Context-switching between direct text and file uploads (using a tabbed layout)

- **Result Metrics (Middle):** Immediate visual feedback via color-coded badges and a **High-Contrast Bar Chart** for score distribution.
- **Detailed Analysis (Right):** Granular sentence-level breakdown and token visualization.
- **Visual Feedback:** Integrated CSS animations and an SVG loading timer to enhance user experience during heavy processing.

3. **Batch Processing Logic**

The application automatically detects paragraph breaks in uploaded files. This allows the UI to provide a "Topic Navigator," where users can see the sentiment of the entire file summary or drill down into specific sections.

2.2 Challenges Faced

1. **Fixed Threshold Bias**

- **Problem:** Default VADER thresholds often labeled mixed reviews (e.g., "It was a good attempt but ultimately a waste of time") as Positive because of the word "good."
- **Solution:** We fine-tuned the thresholds (shifting the "Neutral" zone from ±0.05 to ±0.1) and added a "negation-checker" that penalizes positive words when follow-up disappointment keywords are detected.

2. **UI Alignment & Column Synchronization**

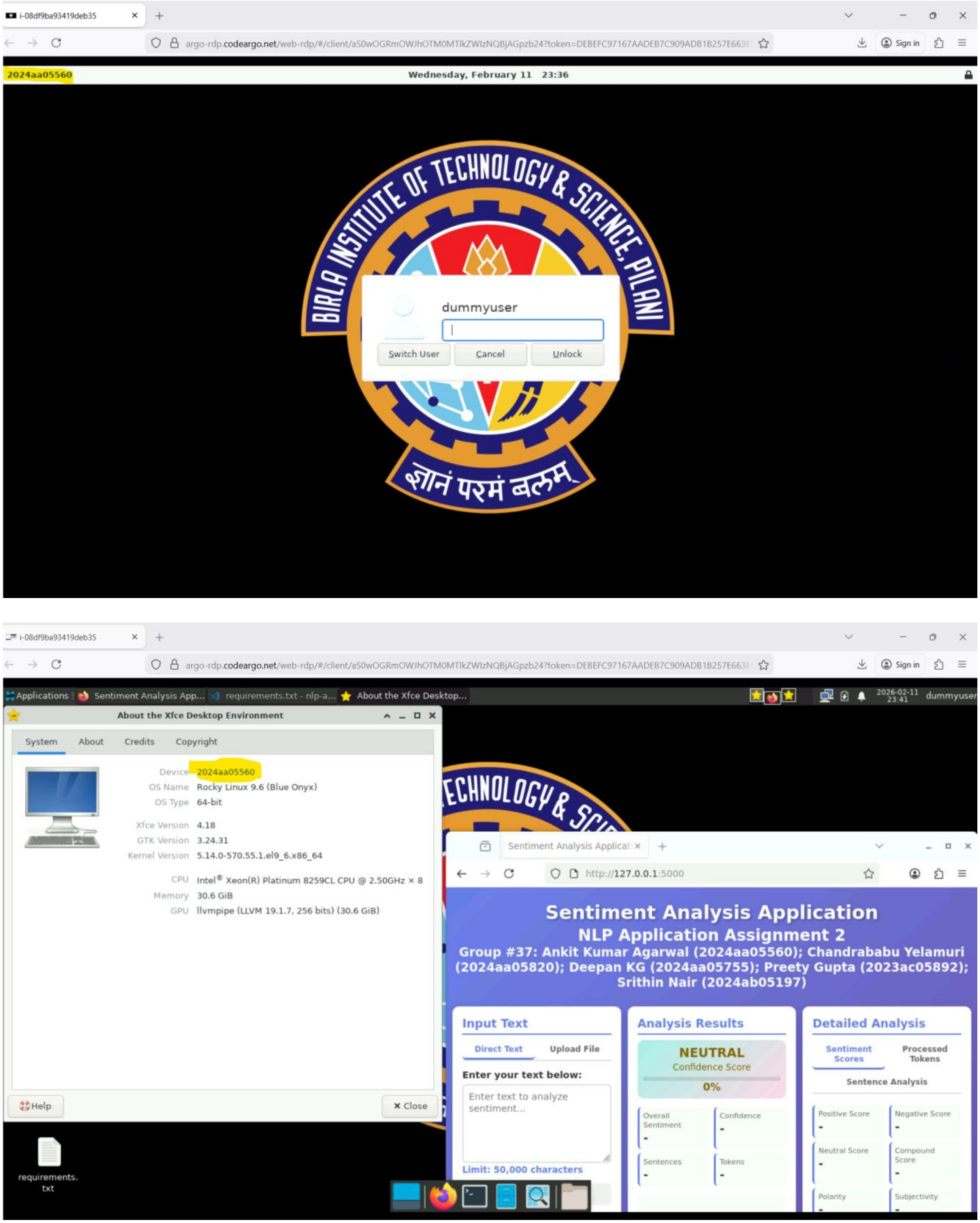
- **Problem:** With varying text lengths, the three columns often had uneven heights.
- **Solution:** Matched the height of all columns to ensure that the dashboard always looks like a professional desktop application, regardless of screen size.

3. **Processing Performance for Large Files**

- **Problem:** Large .txt documents (up to 50,000 characters) caused the browser to hang while waiting for a response.
- **Solution:** Optimized the backend with efficient regex cleaning and chunk-based processing, while adding a robust frontend loading state to prevent redundant API calls.

3 Screenshots

3.1 Proof of BITS OSHA Cloud Lab Usage



i-08df9ba93419deb35
argo-rdp.codeargo.net/web-rdp/#/client/aS0wOGRmOWJhOTM0MTlkZWIzNQBjAGpzb24?token=DEBEFC97167AADEB7C909ADB1B257E663E
Sign in
2024aa05560
Wednesday, February 11 23:36
BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE, PILANI
ज्ञानं परमं बलम्
dummyuser
Switch User
Cancel
Unlock

i-08df9ba93419deb35
argo-rdp.codeargo.net/web-rdp/#/client/aS0wOGRmOWJhOTM0MTlkZWIzNQBjAGpzb24?token=DEBEFC97167AADEB7C909ADB1B257E663E
Sign in
Applications
Sentiment Analysis App...
requirements.txt - nlp-a...
About the Xfce Desktop...
2026-02-11 23:41
dummyuser
About the Xfce Desktop Environment
System
About
Credits
Copyright

Device 2024aa05560
OS Name Rocky Linux 9.6 (Blue Onyx)
OS Type 64-bit
Xfce Version 4.18
GTK Version 3.24.31
Kernel Version 5.14.0-570.55.1.el9_6.x86_64
CPU Intel® Xeon(R) Platinum 8259CL CPU @ 2.50GHz × 8
Memory 30.6 GiB
GPU llvmpipe (LLVM 19.1.7, 256 bits) (30.6 GiB)

Help
× Close
requirements.txt
Sentiment Analysis Applicat
http://127.0.0.1:5000

Sentiment Analysis Application

NLP Application Assignment 2

Group #37: Ankit Kumar Agarwal (2024aa05560); Chandrababu Yelamuri (2024aa05820); Deepan KG (2024aa05755); Preety Gupta (2023ac05892); Srithin Nair (2024ab05197)

Input Text

Direct Text | Upload File

Enter your text below:

Enter text to analyze sentiment...

Limit: 50,000 characters

Analysis Results

NEUTRAL
Confidence Score
0%

Overall Sentiment -
Confidence -
Sentences -
Tokens -

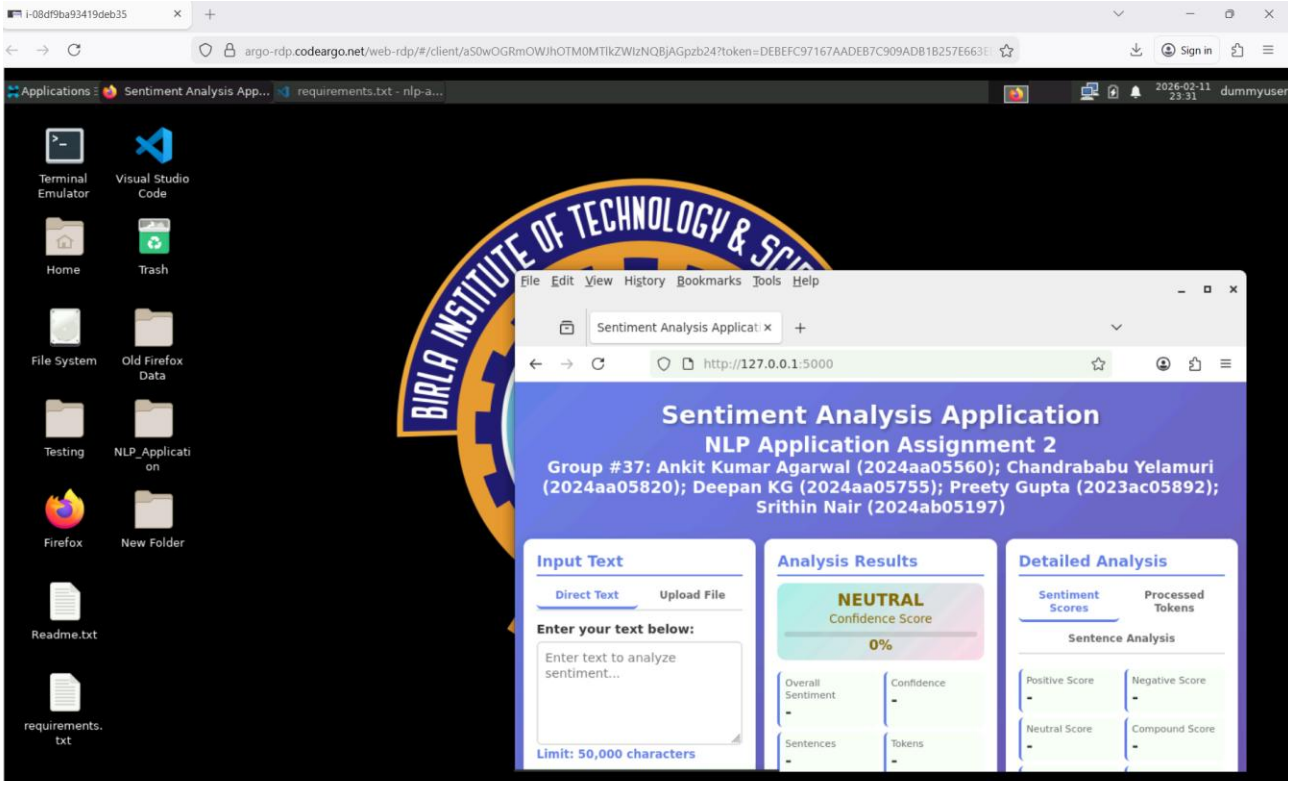
Detailed Analysis

Sentiment Scores | Processed Tokens

Sentence Analysis

Positive Score -
Negative Score -
Neutral Score -
Compound Score -
Polarity -
Subjectivity -

Figure 1: BITS OSHA Lab with BITS Logo on the RDP and Student ID highlighted



Sentiment Analysis Application

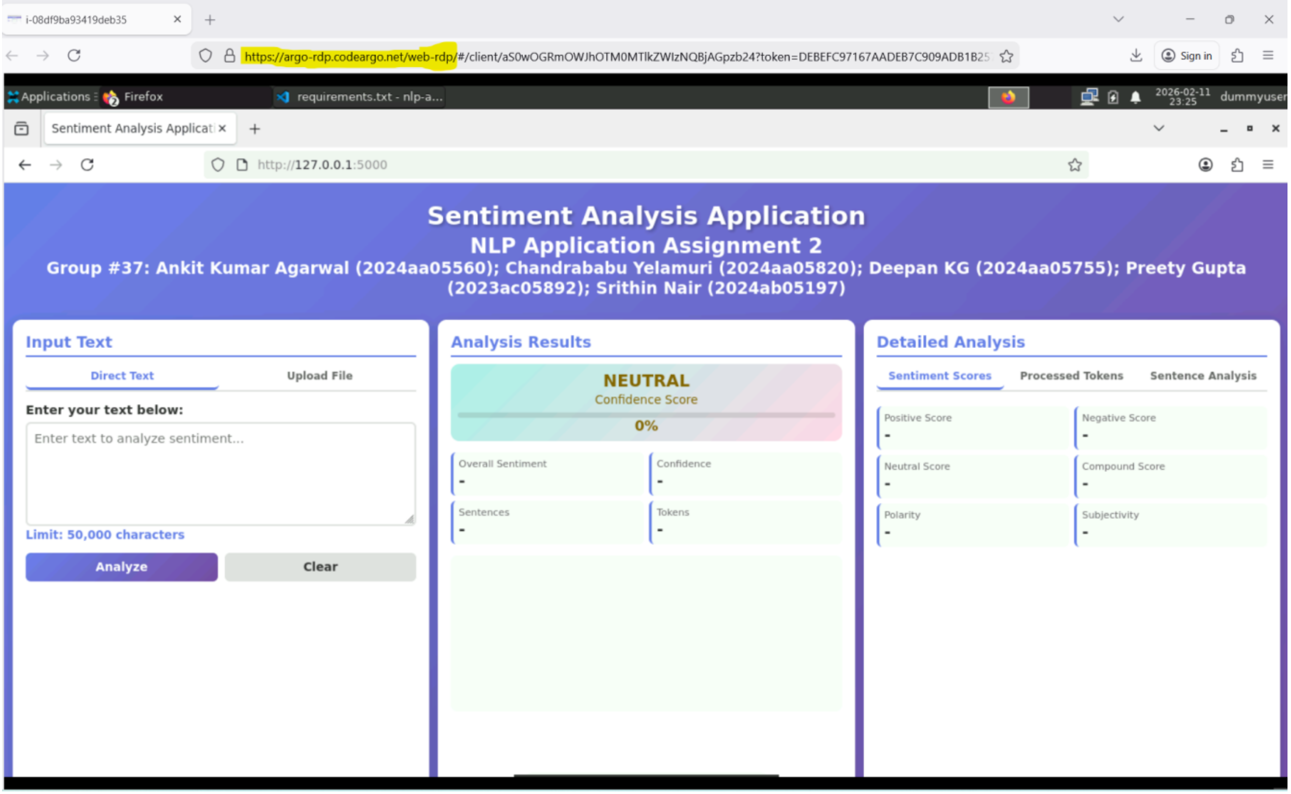
NLP Application Assignment 2

Group #37: Ankit Kumar Agarwal (2024aa05560); Chandrababu Yelamuri (2024aa05820); Deepan KG (2024aa05755); Preety Gupta (2023ac05892); Srithin Nair (2024ab05197)

Figure 2: Application GUI running in the Browser on the BITS OSHA Cloud RDP

Note: The BITS logo on the Desktop shows that the application is running on the BITS OSHA Lab portal

3.2 Application Opening Screen



i-08df9ba93419deb35
https://argo-rdp.codeargo.net/web-rdp/#/client/aS0wOGRmOWJhOTM0MTlkZWIzNQBjAGpzb24?token=DEBEFC97167AADEB7C909ADB1B25
Sign in
Applications
Firefox
requirements.txt - nlp-a...
2026-02-11 23:25
dummyuser
Sentiment Analysis Applicat
http://127.0.0.1:5000

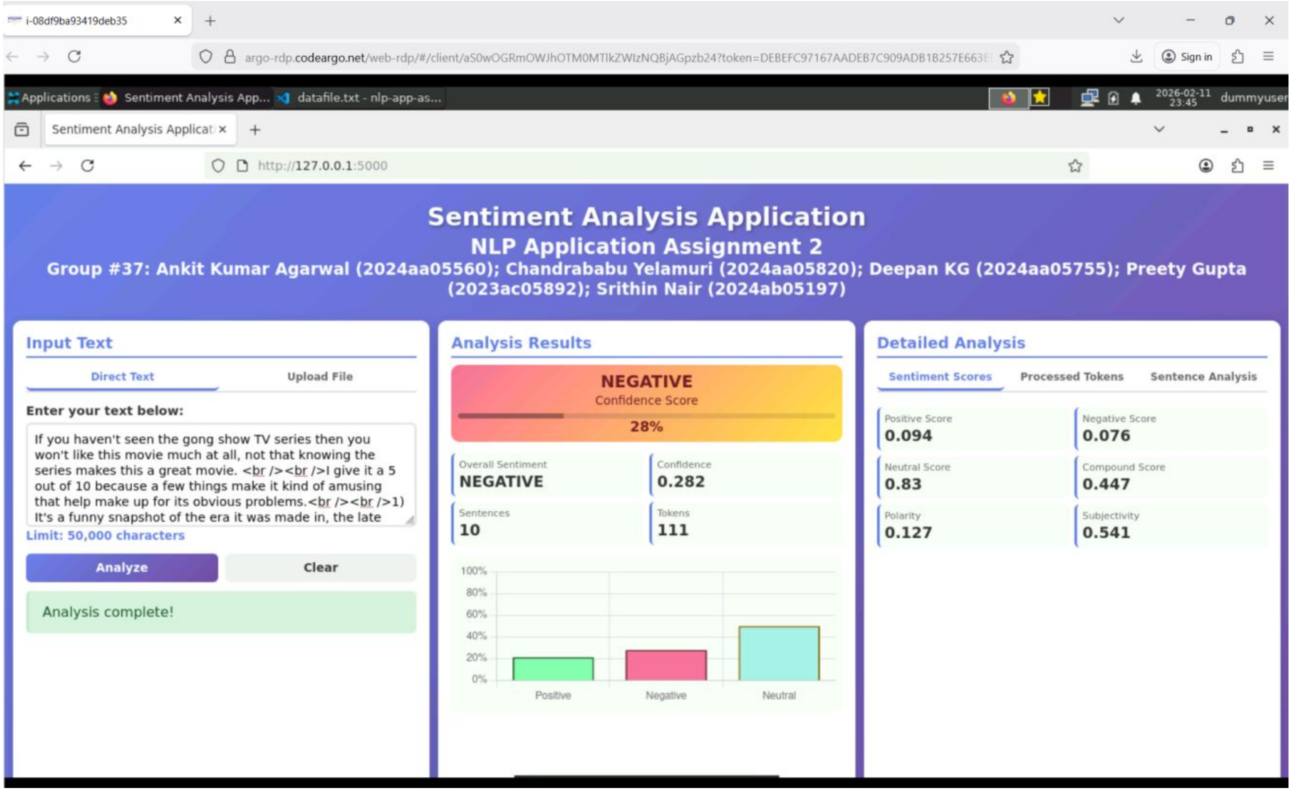
Sentiment Analysis Application

NLP Application Assignment 2

Group #37: Ankit Kumar Agarwal (2024aa05560); Chandrababu Yelamuri (2024aa05820); Deepan KG (2024aa05755); Preety Gupta (2023ac05892); Srithin Nair (2024ab05197)

Figure 3: Application GUI opening screen on the browser

3.3 Sentiment Analysis – Single chunk



Sentiment Analysis Application

NLP Application Assignment 2

Group #37: Ankit Kumar Agarwal (2024aa05560); Chandrababu Yelamuri (2024aa05820); Deepan KG (2024aa05755); Preety Gupta (2023ac05892); Srithin Nair (2024ab05197)

Input Text

Enter your text below:

If you haven't seen the gong show TV series then you won't like this movie much at all, not that knowing the series makes this a great movie. <br /><br />I give it a 5 out of 10 because a few things make it kind of amusing that help make up for its obvious problems.<br /><br />1) It's a funny snapshot of the era it was made in, the late

Limit: 50,000 characters

Analysis complete!

Analysis Results

NEGATIVE
Confidence Score
28%

Overall Sentiment: NEGATIVE

Confidence: 0.282

Sentences: 10

Tokens: 111

Detailed Analysis

Positive Score: 0.094

Negative Score: 0.076

Neutral Score: 0.83

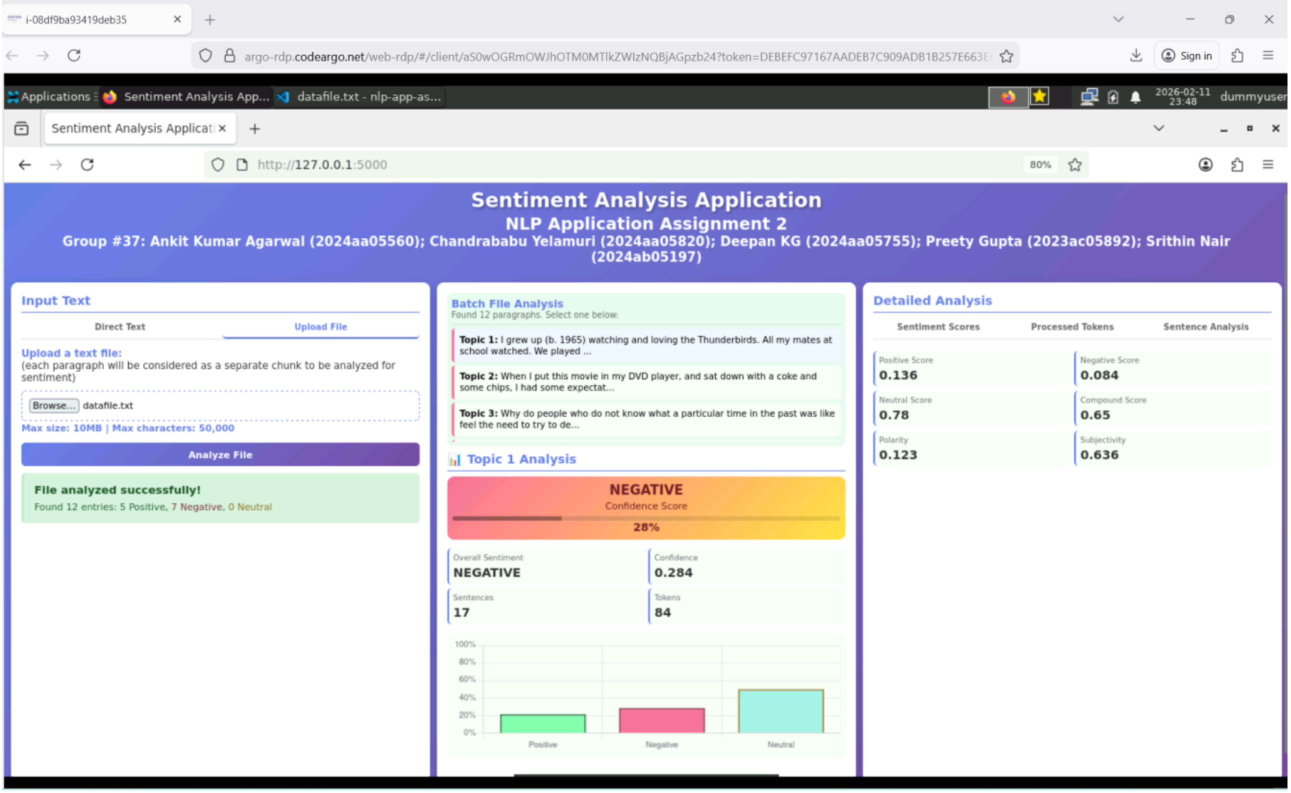
Compound Score: 0.447

Polarity: 0.127

Subjectivity: 0.541

Figure 4: Example showing Sentiment Analysis of single chunk of text

3.4 Sentiment Analysis – Multiple chunks (File Upload option)


i-08df9ba93419deb35
argo-rdp.codeargo.net/web-rdp/#/client/aS0wOGRmOWJhOTM0MTlkZWIzNQBjAGpzb24?token=DEBEFC97167AADEB7C909ADB1B257E663E
Sign in
Applications
Sentiment Analysis App...
datafile.txt - nlp-app-as...
2026-02-11 23:48
dummyuser
Sentiment Analysis Applicati
http://127.0.0.1:5000
80%

Sentiment Analysis Application

NLP Application Assignment 2

Group #37: Ankit Kumar Agarwal (2024aa05560); Chandrababu Yelamuri (2024aa05820); Deepan KG (2024aa05755); Preety Gupta (2023ac05892); Srithin Nair (2024ab05197)

Input Text

Direct Text | Upload File

Upload a text file:
(each paragraph will be considered as a separate chunk to be analyzed for sentiment)

Browse... datafile.txt

Max size: 10MB | Max characters: 50,000

Analyze File

File analyzed successfully!
Found 12 entries: 5 Positive, 7 Negative, 0 Neutral

Batch File Analysis
Found 12 paragraphs. Select one below:

Topic 1: I grew up (b. 1965) watching and loving the Thunderbirds. All my mates at school watched. We played ...

Topic 2: When I put this movie in my DVD player, and sat down with a coke and some chips, I had some expectat...

Topic 3: Why do people who do not know what a particular time in the past was like feel the need to try to de...

📊 **Topic 1 Analysis**

NEGATIVE
Confidence Score
28%

Overall Sentiment NEGATIVE	Confidence 0.284
Sentences 17	Tokens 84

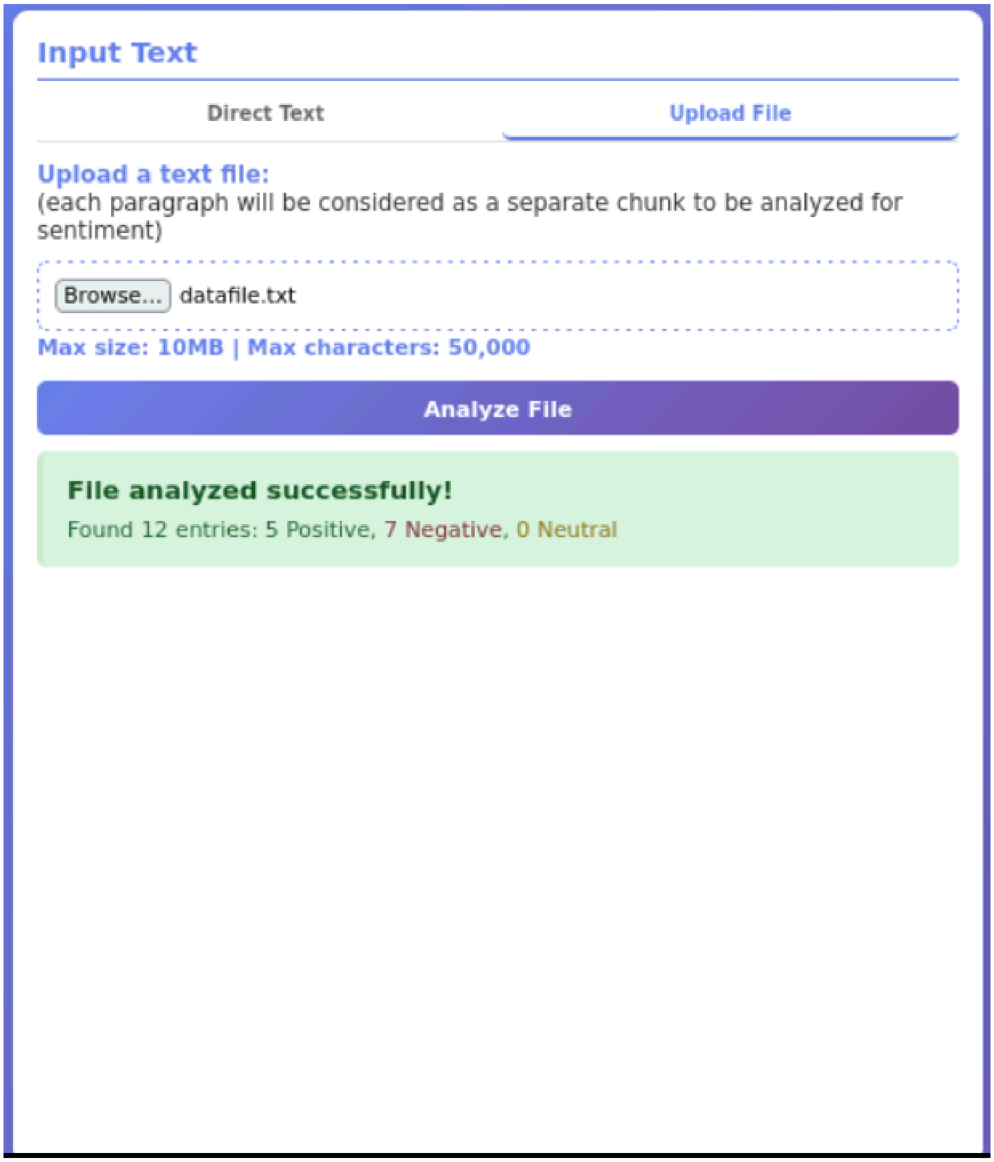
Detailed Analysis

Sentiment Scores | Processed Tokens | Sentence Analysis

Positive Score 0.136	Negative Score 0.084
Neutral Score 0.78	Compound Score 0.65
Polarity 0.123	Subjectivity 0.636

Figure 5: Example showing Sentiment Analysis of multiple chunks in an uploaded text file

3.5 Details of Multiple Chunks Analysis



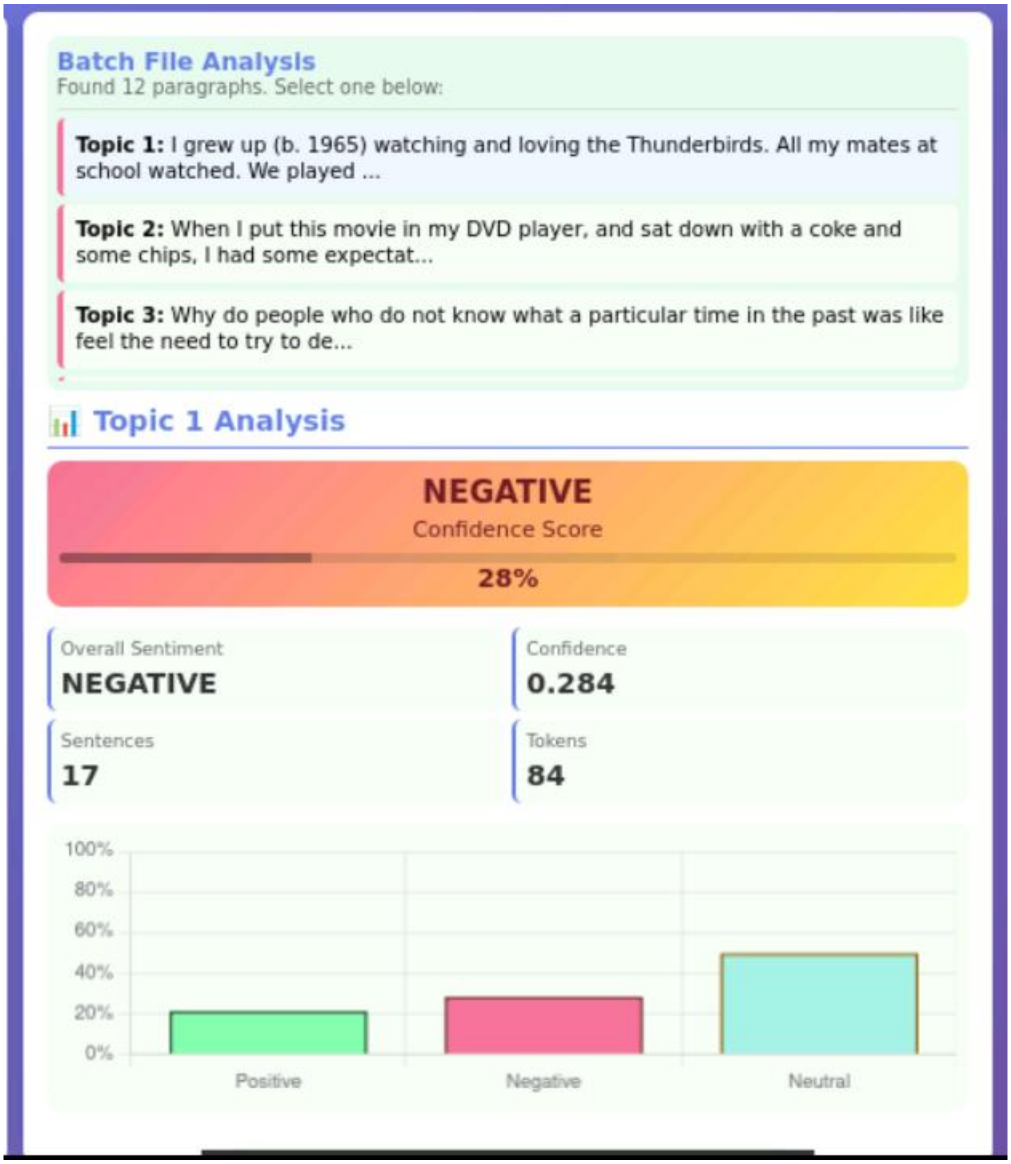
Upload a text file:
(each paragraph will be considered as a separate chunk to be analyzed for sentiment)

datafile.txt

Max size: 10MB | Max characters: 50,000

File analyzed successfully!
Found 12 entries: 5 Positive, 7 Negative, 0 Neutral

Figure 5: After analysis, "Input Text" column shows a quick summary of the content shows total number of entries, with Positive/Negative/Neutral entries



Batch File Analysis

Found 12 paragraphs. Select one below:

Topic 1: I grew up (b. 1965) watching and loving the Thunderbirds. All my mates at school watched. We played ...

Topic 2: When I put this movie in my DVD player, and sat down with a coke and some chips, I had some expectat...

Topic 3: Why do people who do not know what a particular time in the past was like feel the need to try to de...

Topic 1 Analysis

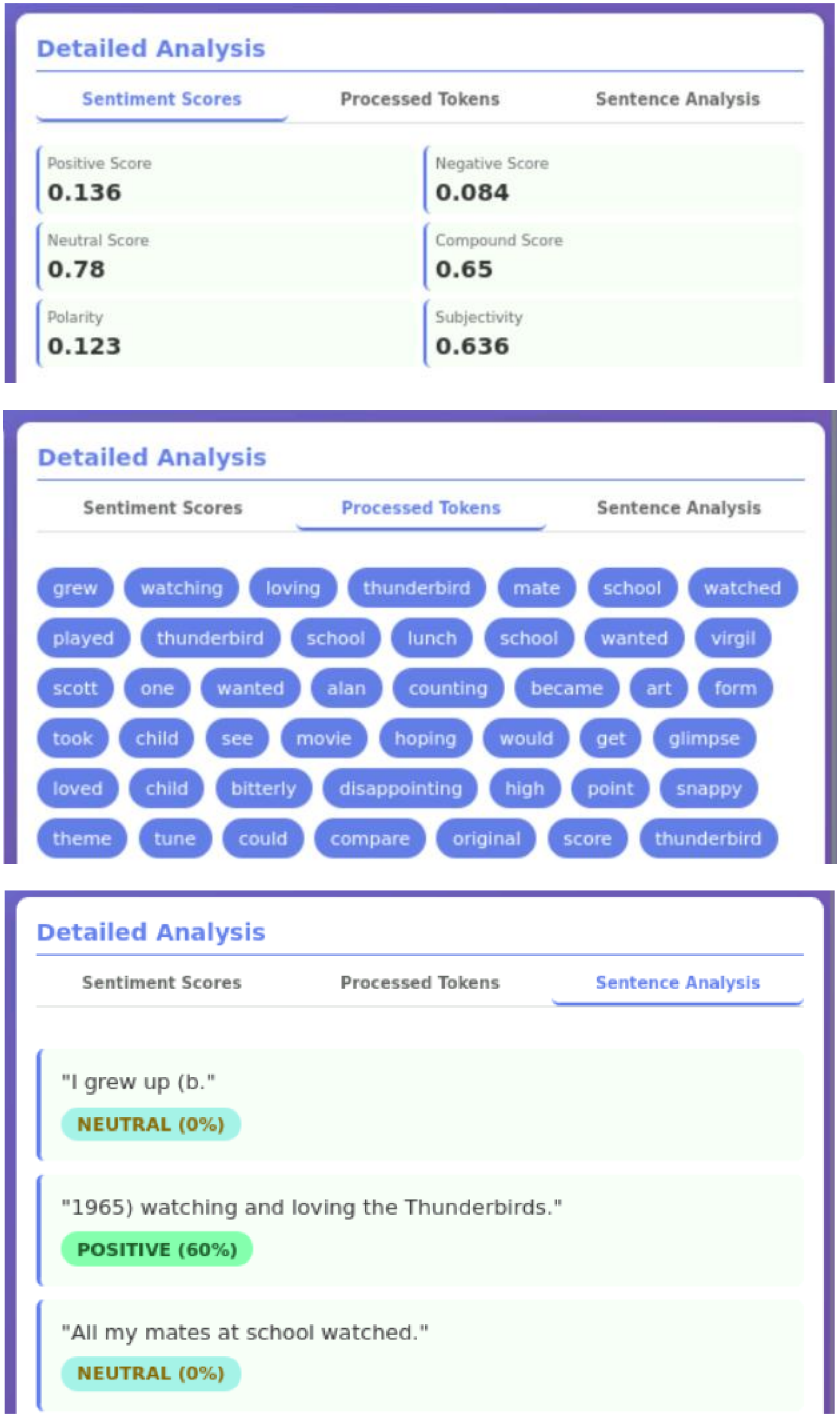
NEGATIVE

Confidence Score

28%

Overall Sentiment **NEGATIVE**	Confidence **0.284**
Sentences **17**	Tokens **84**

Figure 6: After analysis, "Batch File Analysis" column lists all the chunks at the top and shows the analysis of the selected chunk at the bottom



Detailed Analysis

Positive Score 0.136	Negative Score 0.084
Neutral Score 0.78	Compound Score 0.65
Polarity 0.123	Subjectivity 0.636

Detailed Analysis

grew watching loving thunderbird mate school watched played thunderbird school lunch school wanted virgil scott one wanted alan counting became art form took child see movie hoping would get glimpse loved child bitterly disappointing high point snappy theme tune could compare original score thunderbird

Detailed Analysis

"I grew up (b."

NEUTRAL (0%)

"1965) watching and loving the Thunderbirds."

POSITIVE (60%)

"All my mates at school watched."

NEUTRAL (0%)

Figure 7: "Detailed Analysis" column shows the Sentiment Scores, list of Processed Tokens and individual Sentence Analysis for the selected chunk