**System and Unit Test Report**

Summit.ai

Summarize Everything

March 8, 2022

**Sprint 1:**

User stories covered:

* As a backend engineer, I want at least four cleaned datasets so that we can begin training
* As a user, I want the machine learning model to be able to encapsulate a variety of inputs.
* As a user I want the summary output in bullet point or paragraph format

Scenario:

1. Clone the main branch of summarize everything
2. CD into the /data directory
3. Run “make build”
4. Check results in /datasets directory
5. Observe that data is in bullet point or paragraph format, and the data is diverse

User stories covered:

* As a user, I want to see the process status of summarizing.
* As a user, I want to have a functional interface that allows me to upload ‘.txt’ files

Scenario:

1. Launch the summarize everything web application
2. Upload a .txt
3. Press submit
4. Observe summarization message

**Sprint 2:**

User stories covered:

* As a user, I would like to be able to access the website in my web browser and be able to get a summary for simple .txt files.
* As a developer, I would like to have a model that performs accurately for most summarization tasks taken from the test set.
* As a developer I would like to avoid binary/non-text files from being submitted
* As a user, I would like to view the summarized content in a simple and easy to read format.

Scenario:

1. Launch the summarize everything web application
2. Click the upload button
3. Upload a .png
4. Observe that the image is not accepted
5. Click the upload button
6. Upload a .txt
7. Press submit
8. Observe summarization message
9. Evaluate results

User stories covered:

* As a developer, I would like to be able to visualize/gain statistical insights into my datasets to choose the most viable ones.
* As a developer, I would like to have a model that performs accurately for most summarization tasks taken from the test set.

Scenario:

* Launch tensorboard, pointed at training log files
* Observe training results

**Sprint 3**

User stories covered:

* As a developer, I would like to have qualitative performance metrics (test inputs)

Scenario:

* Launch tensorboard, pointed at training log files
* Observe validation results

User stories covered:

* As a user, I would like to be able to get summaries that take longer than http timeouts would typically enable
* As a user, I would like to have an interface that works intuitively and provides me information about my input text and the resulting summary.

Scenario:

1. Launch the summarize everything web application
2. Click the upload button
3. Upload a large .txt
4. Observe the text from the .txt is copied onto the webpage
5. Press submit
6. Wait for summary to complete
7. Evaluate results

**Sprint 4**

User stories covered:

* As a user I would like to summarize news articles using general URL links (i.e. not just NY Times articles)
* As a user I would like to choose between the top 3 performing models
* As a user, I would like to have the system automatically notify or prevent repeat queries

Scenario:

1. Launch the summarize everything web application
2. Click the url button
3. Paste a link to a news article
4. Save
5. Observe the url translated to text
6. Select a model from the dropdown
7. Press submit
8. Observe the new loading message
9. Evaluate the results