Experiential Project Plan - Team Repurpost

Yashwanth Balan Arumugam, Pragya Avinash Mishra, Shruti Sham Kotwal, Atharva Shantanu Kulkarni, Saju Chacko Rajan

1. A) Problem Statement:

Repurpost has at its core a content repository (DAM) where users can classify their content based on strategy (campaign, buyer persona, product line) or by tags. The problem is that the user frequently does not know which is the right tag, so content is not well classified, or the tags used are not useful for other platform users.

B) Desired End Result:

Creation of a real-time web application that can provide a list of tags that are based on the text content of the input. The underlying logic of web application will include a Machine Learned model that is trained on existing dataset and tags based on data from various content assets from the Repurpost platform.

C) Prior Work:

There is no prior feature on the repurpost platform for tag recommendations and suggestions based on content. Currently, the users add the tags to the content causing misclassification and incorrect categorization of content.

D) Success Criteria:

- 60% - 90% Accuracy for the ML model.
- Create Web Application / API that can integrate with existing platform with minimal changes.
- Create a working POC/Demo page to show Web API in action.

2. Communication Plan:

Date	Communication Method	Purpose of Communication
Oct 04, 2022	**Zoom Meeting, Slack**	<u>Sponsor Meeting:</u> • Team & Sponsor introduction, initial research, and objective definition.
Oct 11, 2022	**Google Meet**	<u>Internal Team Meeting:</u> • Concepts Ideation, Dataset research and development-I, Domain selection.
Oct 18, 2022 **Oct 25, 2022**	**Google Meet**	<u>Internal Team Meeting:</u> • Dataset research-II and finalize dataset. • Initial dataset collection, analysis, cleaning, and engineering process. • Research on various technologies that can be used.
Oct 31, 2022	**Zoom Meeting, Slack**	<u>Sponsor Meeting: Checkpoint I</u> • Verifying the research done and the objectives accomplished up till. • Discussion on further steps aligning with the objective and the deliverables.
Nov 01, 2022 **Nov 8, 2022**	**Google Meet**	<u>Internal Team Meeting:</u> • Building different AI models. • Comparing the accuracies and outcomes. • Choosing the right AI model.
Nov 14, 2022	**Zoom Meeting, Slack**	<u>Sponsor Meeting: Checkpoints II</u> • After reinforcing the suggested changes, demonstrating the initial results. • Getting feedback on the initial phase of the prototype.

Nov 15, 2022	**Google Meet**	Internal Team Meeting: • Testing and deployment phase I. • Initial Demo to the sponsor.
Nov 22, 2022	**Google Meet**	Internal Team Meeting: • Testing and deployment phase II. • Quality checking of the final product.
Nov 28, 2022	**Zoom Meeting**	Sponsor Meeting: • Final delivery of the project and the paper to the sponsor. • Final Demo of the product to the sponsor.
Dec 01, 2022	**Zoom Meeting**	Coursework In-class/Online Presentation: • Final project presentation to the class. • Feedback and Q&A from the professor and other teams.

3. Key Stakeholders:

Name	Title	Project Role
Alexandra Cowen	Chief Executive Officer (CEO)	• Sponsor owns the project and provides the details and high-level requirements for the project in addition to the resources and support to enable the project's success.
David F.	Chief Technology Officer (CTO)	• Co-Sponsor makes sure the initiative has the resources it needs to succeed while promoting its value. • Giving the project manager expertise and direction.

Yashwanth Balan Arumugam	Project Lead & Development	• As the project's primary coordinator, the leader keeps everything running smoothly by keeping the team and the sponsor informed of progress and ensuring that each work is completed on time. • Responsible for deciding how to approach tasks and develop a plan to accomplish them and for distributing information to team members and stakeholders. • Formulating the project's goals and outcomes, choosing how to distribute resources, and making sure the team is well-managed are some of the other duties of the team head.
Pragya Avinash Mishra	Research & Development	• The research lead outlines the context, in-depth examination of the project, and stakeholder needs. • Throughout the prototype-building process, the fundamental goal of this phase is to maintain a consistent research strategy. • Other tasks include matching methodologies to research objectives, utilizing a variety of tools to gather data and relevant information.

Shruti Sham Kotwal	Architecture & Development	• The architecture and development team in this project determines the software requirements in collaboration with the project's stakeholders and • Make high-level design and product specs papers. Supplies architectural guidelines for the development team to follow.
Atharva Shantanu Kulkarni		• Directing and helping the development team all the way through. Identify and fix coding or design problems and make sure the team follows development deadlines.
Saju Chacko Rajan	Research, Quality & Development	• The research and quality lead makes quality standards and puts them into action. • Develop dependable documentation and work together with the stakeholders to find ways to make workflow and controls better while ensuring that they go by the rules and regulations. • Verify and test procedures, and services to see if they adhere to specifications or surpass them and review prototyping-related statistics and reports. Delivers final quality checking documentation to review and access the final deliverable.

4. Desired Results and Final Deliverable(s):

<u>**Client Desired end state:**</u>

- Creation of a real-time web application that collects data from various content assets and displays a list of tags that identify the related topics.
- Provide tag suggestions to content creators and repurpost platform users.
- Provide related content references along with tag suggestions from repurpost database.

<u>**Team Desired Results:**</u>

- Provide a working web application for Repurpost team along with Demo.
- Create multiple ML models to perform NLP text analysis of content.
- Determine the best model to be used for web application/API.
- Provide comparison results and model selection criteria.
- Deploy POC (proof of concept) model.
- Provide documentation of solution design, architecture, approach, and usage guide.

<u>**Final Deliverables:**</u>

- Solution workflow diagram.
- Working ML Model for tag suggestions.
- API endpoint to invoke ML model.
- GitHub repository of model code.
- Provide Web API usage guide and documentation.

5. Potential Timeline:

Week 1: October 6th – October 12th, 2022:

Team & Sponsor introduction, initial research, and objective definition:

- Team onboards and understands the requirements.
- Team meets with sponsor and ensure any questions are asked and answered.

Week 2: October 13th – October 19th, 2022:

Concepts Ideation, Dataset research:

- The team will perform research on the sponsor provided datasets and other datasets for use in this project.
- Team will meet to discuss findings and finalize the dataset to be used for project purposes.

Week 3: October 20th – October 26th, 2022:

Dataset analysis, cleaning, and exploratory data analysis:

- The team is expected to perform initial dataset cleanup and analysis to prepare for EDA.
- Team to prepare some EDA results performed on the dataset.

Week 4: October 27th – November 2nd, 2022:

Architecture, Design and Technology Selection:

- The team is expected to create architecture and solution design chart or diagram for the project.
- Team to document the various models that are available for text and NLP Multi-label classification.

Week 5: November 3rd – November 9th, 2022:

Code Development:

- Performing text analysis using various preprocessing techniques, applying lemmatization and tokenization on the text data.
- Implementing Term-Frequency and Inverse document frequency (TF-IDF) for document search and information retrieval from the document.

Week 6: November 10th – November 16th, 2022:

Model comparison and tuning:

- Investigate, examine, and put several AI models into practice. Then, analyze the accuracies of each model.

- Determine which algorithm seems appropriate for tag suggestion.

- Improving the model's performance by introducing hyperparameter tuning

Week 7: November17th – November 23rd, 2022:

Testing, deployment, and Quality check:

- Using some performance metrics like accuracy score, F1-score, Precision, recall, etc. to test the performance of the model.

- Deploying the model and expose web API for integration.

- Testing the agility and the quality of the ML model.

- Create Help and Documentation for using and modifying the ML model

Week 8: November 24th – November 29th, 2022:

Demo and Delivery to Sponsor:

- Demonstrate the deployed product to the sponsor, deliver the final presentation and the documentation.