AudiEyes

Ayoub Maimmadi

# 1. Project Inception

## 1.1. Framing the Business Idea as an ML Problem

* Business case description

Audieyes is a groundbreaking project aimed at enhancing the independence and safety of visually impaired individuals through advanced machine learning technologies. The service utilizes image and video captioning to provide real-time, detailed descriptions of environments and objects, improving navigation and accessibility. This addresses a critical need within the underserved assistive technology market, offering both significant social impact and commercial potential. By focusing on inclusivity and technological innovation, Audieyes not only opens new markets for businesses but also profoundly enhances the quality of life for its users.

* Business value of using ML

Utilizing machine learning in the Audieyes project enables precise and adaptive recognition of visual data, greatly enhancing the service's capability to assist visually impaired users. ML algorithms optimize the accuracy and speed of image and video captioning, ensuring real-time feedback that is critical for navigation and interaction. Moreover, continuous learning from user interactions and feedback refines the system's effectiveness, ensuring that the technology evolves to meet diverse user needs. This not only fosters user dependency and satisfaction but also positions Audieyes as an innovative leader in assistive technologies, potentially increasing market share and generating sustainable revenue streams.

* Data overview

Audieyes utilizes a rich dataset comprising image-text pairs from diverse sources such as COCO, flickr30k, vqa, and nlvr. These datasets provide a wide variety of visual scenarios and associated descriptions, essential for training the ML models to recognize and articulate the content of images accurately. The data includes everyday objects, people, scenes, and activities, ensuring comprehensive coverage and relevance to real-world situations faced by visually impaired users. This extensive data foundation enables the ML models to deliver precise and contextually appropriate captions, critical for the functionality of Audieyes.

* Project archetype

Audieyes represents a hybrid of autonomous real-time systems and human-in-the-loop architectures. It combines the speed and scalability of automated machine learning for instant image and video captioning with the precision of human oversight to enhance accuracy and contextual relevance. This model ensures that the system can operate independently while still benefiting from periodic human intervention to refine outputs, making it highly effective for applications in dynamic and varied environments. The design caters specifically to the needs of the visually impaired, promoting independence through technology while maintaining the flexibility to adapt and improve with user feedback. For more details click [here](https://github.com/AyoubMaimmadi/Audieyes?tab=readme-ov-file#architectural-archetypes-for-vision-captioning).

## 1.2. Feasibility analysis

* + (2 pts) literature review
  + (0.5 pts) model choice/ specification of a baseline
  + (0.5 pts) Metrics for business goal evaluation

# 2. ML Pipeline Development - From a Monolith to a Pipeline

## 2.1. Ensuring ML Pipeline Reproducibility (milestone 2, 15%)

* (2 pts) Project structure definition and modularity
* (2 pts) Code versioning
* (3 pts) Data versioning
* (5 pts) Experiment tracking and model versioning
* (1 pts) Setting up a meta store for metadata
* (2 pts) Setting up the machine learning pipeline under an MlOps platform

## 2.2. Pipeline Components (Milestone 3 and 4, 20%)

* + 1. Setup of data pipeline within the larger ML pipeline/ MLOps Platform
  + (5 pts) Data Validation and Verification
  + (5 pts) Preprocessing and Feature Engineering
    1. (5 pts) Integration of model training and offline evaluation into the ML pipeline / MLOps Platform
    2. (5pts) Development of model behavioral tests
    3. Energy efficiency measurement (optional, 2 pts bonus)

# 3. Model Deployment (Milestones 5-6, 35%)

## 3.1 ML System Architecture

* (5 pts) Drawing with architecture highlights

## 3.2 Application development

* (5 pts) Model service development
* (5 pts) Front-end client development

## 3.3 Integration and Deployment

* (3 pts) Packaging and containerization
* (5 pts) Integration with a CI/CD Pipeline
* (3 pts) Hosting the application

## 3.4. Model Serving and online testing

* (3 pts) Model serving runtime
* (3 pts) Serving mode (batch, on demand to a human, on demand to a machine)
* (3 pts) Online testing (A/B Testing, Bandit)

# Monitoring and Continual Learning (milestone 7, 25%)

## 4.1. (2 pts) Resource Monitoring

## 4.2. (10 pts) Model Performance Monitoring or data distribution drift monitoring

## 4.3. (10 pts) Continual Learning: CT/CD pipeline

## 4.4. (3 pts) Pipeline orchestration

# 5. Responsible AI (milestone 8-optional, for later, 15% bonus)

## 5.1 Evaluation Beyond Accuracy

* (7.5 pts) Audit Model for Bias
* (7.5 pts) Model Explainability and Interpretability

# 6. Conclusion

* Summary of Achievements
* Lessons Learned
* Future Directions

# References

# Appendices (if necessary)