

Structure for Project Workflow (PACE) - Plan

Project: TikTok Claims Classification with Machine Learning

This document outlines the project workflow for developing a machine learning model to classify user comments on TikTok videos as containing claims or expressing opinions. The PACE framework will guide the project execution, ensuring a structured and efficient approach.

PACE Stages and Associated Tasks:

- **Plan (1 week)**

- Task 1.1: Data Acquisition and Exploration (Orion Rainier)
 - Milestone 1.1.1: Confirm data availability and access procedures (1 day)
 - Milestone 1.1.2: Conduct initial data exploration to understand data characteristics (2 days)
- Task 1.2: Project Workflow Definition (Data Team)
 - Milestone 1.2.1: Determine and document the project workflow using tools like Git or Jira (2 days)
- Task 1.3: Model Selection and Assumptions (Willow Jaffey)
 - Milestone 1.3.1: Research and propose suitable machine learning model options for claim classification (2 days)

- **Analyze (2 weeks)**

- Task 2.1: Data Cleaning and Preprocessing (Orion Rainier)
 - Milestone 2.1.1: Address missing data and inconsistencies within the dataset (3 days)
 - Milestone 2.1.2: Feature engineering to create new informative variables from existing data (3 days)
- Task 2.2: Model Training and Evaluation (Data Team)
 - Milestone 2.2.1: Train the chosen machine learning model on the prepared data (4 days)
 - Milestone 2.2.2: Evaluate model performance using relevant metrics (2 days)

- Construct (1 week)

- Task 3.1: Model Refinement and Optimization (Data Team)
 - Milestone 3.1.1: Refine the model based on evaluation results to improve performance (3 days)
 - Milestone 3.1.2: Test and validate the final model on a hold-out dataset (2 days)

- Execute (1 week)

- Task 4.1: Model Deployment and Integration (Data Team)
 - Milestone 4.1.1: Develop a plan for model deployment into TikTok's content moderation system (2 days)
- Task 4.2: Model Documentation (Data Team)
 - Milestone 4.2.1: Document the final model, including assumptions, limitations, and performance metrics (3 days)

Communication and Collaboration:

- Regular team meetings (weekly) to discuss progress, address challenges, and make collaborative decisions.
- Project management tools (e.g., Git, Jira) for task tracking and communication.
- Status updates are sent to relevant stakeholders, including leadership.

Benefits of Utilizing PACE:

- **Structured Approach:** Ensures a logical progression through the project lifecycle.
- **Improved Efficiency:** Minimizes rework and streamlines tasks.
- **Enhanced Communication:** Fosters collaboration and information sharing.
- **Clear Deliverables:** Defines key milestones and expected outcomes.