Course 1: Gathering Information from Stakeholder Notes within TikTok

In the latest stakeholder meeting at TikTok, key points were discussed regarding the machine learning model for classifying claims made in videos. Here are the relevant notes:

1. \*\*Data Sources:\*\*

- Identify and gather data sources related to claims in videos, considering both textual and visual content.

- Explore possibilities for incorporating user feedback and annotations to enhance the dataset.

2. \*\*Labeling and Annotation:\*\*

- Develop a strategy for labeling and annotating the dataset, ensuring accuracy in claims classification.

- Consider utilizing both automated and manual labeling processes.

3. \*\*Model Integration:\*\*

- Understand the technical requirements for integrating the machine learning model into TikTok's existing infrastructure.

- Ensure compatibility with the platform's real-time processing capabilities.

4. \*\*Ethical Considerations:\*\*

- Evaluate potential biases in the data and model, considering TikTok's commitment to inclusivity.

- Implement measures to address biases and promote fairness in claims classification.

5. \*\*User Interface Design:\*\*

- Collaborate with UX/UI teams to design an intuitive interface for users to report and provide feedback on claims.

Course 2: Assigning PACE Stages to the Requested Tasks for the Classification Project

P - Planning: Tasks 1, 2, 3

- Planning involves identifying data sources, developing labeling strategies, and understanding model integration requirements.

A - Acquisition: Task 1

- Acquisition involves gathering relevant data sources for claims classification.

C - Cleaning: Task 2

- Cleaning involves developing strategies for labeling and annotating the dataset.

E - Exploration: Task 4

- Exploration involves evaluating potential biases in data and model.

Course 3: Organizing Tasks into Milestones

\*\*Milestone 1: Data Preparation\*\*

- Task 1: Identify and gather relevant data sources.

- Task 2: Develop labeling and annotation strategies.

- Timeline: 6 weeks

\*\*Milestone 2: Model Integration\*\*

- Task 3: Understand technical requirements for model integration.

- Timeline: 4 weeks

\*\*Milestone 3: Ethical Considerations and Bias Mitigation\*\*

- Task 4: Evaluate potential biases in data and model.

- Timeline: 5 weeks

\*\*Milestone 4: User Interface Design\*\*

- Task 5: Collaborate with UX/UI teams for an intuitive reporting interface.

- Timeline: 3 weeks

Course 4: Creating a Project Proposal for the TikTok Data Team

\*\*Project Title:\*\* TikTok Claims Classification Model Development

\*\*Objective:\*\* Develop a machine learning model for classifying claims made in videos submitted to the platform.

\*\*Milestones:\*\*

1. Data Preparation

- Tasks: Identify and gather data sources, develop labeling strategies.

- Timeline: 6 weeks

2. Model Integration

- Task: Understand technical requirements for model integration.

- Timeline: 4 weeks

3. Ethical Considerations and Bias Mitigation

- Task: Evaluate potential biases in data and model.

- Timeline: 5 weeks

4. User Interface Design

- Task: Collaborate with UX/UI teams for an intuitive reporting interface.

- Timeline: 3 weeks

\*\*Overall Timeline:\*\* 18 weeks

\*\*Resources Needed:\*\*

- Data Scientists: 3

- Data Engineers: 2

- UX/UI Designers: 1

- Project Manager: 1

\*\*Budget Estimate:\*\* $1,000,000

\*\*Risk Management:\*\*

- Potential delays in data acquisition and labeling.

- Technical challenges in model integration.

\*\*Approval:\*\*

This proposal is submitted for the TikTok data team's approval. Upon approval, the project will proceed with the outlined milestones and timelines.

[Signature]

[Date]