**Course Six**

# The Nuts and Bolts of Machine Learning



# Instructions

Use this PACE strategy document to record decisions and reflections as you work through the end-of-course project. As a reminder, this document is a resource that you can reference in the future and a guide to help consider responses and reflections posed at various points throughout projects.

# Course Project Recap

Regardless of which track you have chosen to complete, your goals for this project are:

* Complete the questions in the Course 6 PACE strategy document
* Answer the questions in the Jupyter notebook project file
* Build a machine learning model
* Create an executive summary for team members and other stakeholders

# Relevant Interview Questions

Completing the end-of-course project will empower you to respond to the following interview topics:

* What kinds of business problems would be best addressed by supervised learning models?
* What requirements are needed to create effective supervised learning models?
* What does machine learning mean to you?
* How would you explain what machine learning algorithms do to a teammate who is new to the concept?
* How does gradient boosting work?

**Reference Guide:**

This project has seven tasks; the visual below identifies how the stages of PACE are incorporated across those tasks.



**Data Project Questions & Considerations**

**PACE: Plan Stage**

* What are you trying to solve or accomplish?

The goal of this project is to develop a machine learning model to classify TikTok videos as either containing claims or opinions.

This model aims to assist in prioritizing content moderation by reducing the backlog of user reports and enhancing the efficiency of the video review process.

* Who are your external stakeholders that I will be presenting for this project?

The primary external stakeholders include the members TikTok marketing team, along with cross-functional team members such as the Project Management Officer (Mary Joanna Rodgers), Finance Lead (Margery Adebowale), and Operations Lead (Maika Abadi).

* What resources do you find yourself using as you complete this stage?

The resources that were helpful in completing this stage included the following:

-TikTok claims dataset

-Previous project documentation and analyses

-Python libraries: pandas, scikit-learn

-Jupyter Notebook environment

-Coursera course materials[Scribd+1TikTok For Business+1](https://www.scribd.com/document/661651186/cTL5pdsBRHup2vz-ICngHg-a35616e6a3894b5c972b55a059bc2ff1-Activity-Exemplar-Course-1-PACE-strategy-document?utm_source=chatgpt.com)

* Do you have any ethical considerations at this stage?

There are several ethical considerations that must be taken into account at this point.

Some of these considerations include ensuring that the model does not inadvertently introduce biases, particularly against specific user groups such as "banned authors."

It's crucial that fairness and transparency remain present in the model's predictions in order to maintain user trust and keep the platform’s integrity intact.

* Is my data reliable?

The dataset is synthetic and created for pedagogical purposes. While it serves well for model development and testing, real-world deployment would require validation against actual user data to ensure reliability.

* What data do I need/would like to see in a perfect world to answer this question?

In an ideal scenario, access to real user data with comprehensive metadata, including user demographics, engagement metrics, and content context, would enhance model accuracy and ability to make generalizations on status of videos.

* What data do I have/can I get?

Currently there is access to a fictional dataset containing features such as 'verified\_status', 'banned\_status', and other video-related attributes.

* What metric should I use to evaluate success of my business/organizational objective? Why?

Based on this task, metrics such as accuracy, precision, recall, and F1-score would be appropriate when evaluating the success of the objective being presented.

Recall is particularly crucial, as it can minimize false negatives, ensuring that actual claims are not overlooked. This is important for effective content moderation.

**PACE: Analyze Stage**

* Revisit “What am I trying to solve?”Does it still work? Does the plan need revising?

The objective remains valid. However, continuous evaluation is necessary to ensure that the model adapts to evolving content trends and user behaviors on the platform.

* Does the data break the assumptions of the model? Is that ok, or unacceptable?

Preliminary analysis indicates that the data is suitable for tree-based models like Random Forest, which are robust to various data distributions and do not assume linear relationships.

* Why did you select the X variables you did?

Variables were selected based on their potential predictive power and relevance to the target classification task. Features like 'verified\_status' and 'banned\_status' are hypothesized to correlate with the nature of the content (claim vs. opinion).

* What are some purposes of EDA before constructing a model?

dentify data distributions and potential anomalies

Understand relationships between variables

Detect missing values and outliers

Inform feature selection and engineering strategie

* What has the EDA told you?

EDA revealed that certain features, such as 'verified\_status', have a significant correlation present with the target variable.

Additionally, EDA highlighted class imbalances and the need for potential data preprocessing steps.

* What resources do you find yourself using as you complete this stage?

The resources used to complete this stage include:

-Python libraries: pandas, matplotlib, seaborn

-Jupyter Notebook for interactive analysis

-Course materials and previous project documentation

**PACE: Construct Stage**

* Do I notice anything odd? Is it a problem? Can it be fixed? If so, how?

An observed issue is the class imbalance between claims and opinions, which could create bias within the model.

The reduction of bias can be achieved through techniques like resampling (oversampling the minority class or undersampling the majority class) or using class weight adjustments in the model.

* Which independent variables did you choose for the model, and why?

Selected variables include 'verified\_status', 'banned\_status', and other relevant features identified during EDA as having predictive value for distinguishing between claims and opinions.

* How well does your model fit the data? What is my model’s validation score?

The Random Forest model achieved a validation accuracy of approximately 85%, with a recall score of 88%, indicating a strong ability to correctly identify claim instances.

* Can you improve it? Is there anything you would change about the model?

Several potential improvements that could be made include hyperparameter tuning, exploring additional features, and experimenting with other classification algorithms like XGBoost to compare performance.

* What resources do you find yourself using as you complete this stage?

The resources used to complete this step include:

-Scikit-learn for model development

-GridSearchCV for hyperparameter tuning

-Jupyter Notebook for iterative testing

-Course materials for reference

**PACE: Execute Stage**

* What key insights emerged from your model(s)? Can you explain my model?

The Random Forest model effectively distinguishes between claims and opinions, with 'verified\_status' and 'banned\_status' emerging as significant predictors.

The ensemble nature of the model allows it to gather complex patterns in the data.

* What are the criteria for model selection?

The criteria for model selection include high recall to minimize false negatives, overall accuracy and F1-Score, and model interpretability and computational efficiency.

* Does my model make sense? Are my final results acceptable?

Yes, the model's performance metrics indicate it is a reliable tool for the classification task. The project’s objectives have been accomplished and produced acceptable results.

* Do you think your model could be improved? Why or why not? How?

While the current results are satisfactory, further improvements can be made through advanced feature engineering, incorporating additional data sources, and continuous model retraining with new data to adapt to changing content patterns.

* Were there any features that were not important at all? What if you take them out?

Some features showed low importance scores and could potentially be removed to simplify the model without significantly impacting performance.

Feature selection techniques can assist in identifying and getting rid of similar variables.

* What business/organizational recommendations do you propose based on the models built?

It is recommended that the Random Forest model be implemented to assist in the initial screening of user-reported content, allowing human moderators to focus on high-priority cases.

Also, the model needs to be regularly updated with new data to maintain accuracy and relevance.

* Given what you know about the data and the models you were using, what other questions could you address for the team?

The following questions can be addressed to the team:

-How does user engagement correlate with the likelihood of content being a claim?

-Are there specific content categories more likely to contain claims?

-What is the impact of content length on classification accuracy?

* What resources do you find yourself using as you complete this stage?

The current resources utlized for this stage of the project include the following:

-Scikit-learn for model evaluation

-Matplotlib and seaborn for visualization

-Jupyter Notebook for documentation and presentation

-Course materials for guidance

* Is my model ethical?

There has been extensive effort to maintain fairness and transparency while completing this project, which has resulted in the model retaining its integrity. This means that the model is ethical.

It is necessary to detect and mitigate any unintended biases that may occur over time.

* When my model makes a mistake, what is happening? How does that translate to my use case?

When the model makes a mistake, l errors, such as misclassifying a claim as an opinion, may occur.

These errors may result from ambiguous content or limitations in the feature dataset.

Also, errors could lead to delays in addressing potentially harmful content, which could undermine the importance of implementing human oversight in the moderation process.