

Machine Learning Models for TikTok: Analyzing Claims and Opinions in Content

Credit : Bhavan

Company Overview

TikTok is a popular social media platform that allows users to create, share, and discover short-form videos. With its rapidly growing user base, TikTok has become a significant player in the social media landscape, focusing on engaging content and innovative features. The platform's algorithm personalizes content for users based on their preferences, making it essential for creators and marketers to understand the dynamics of video performance and audience engagement.

Company Overview

This project aims to develop a predictive model that can classify TikTok videos as either containing claims or offering opinions. By analyzing various features of the videos and user interactions, we seek to provide insights into content trends and improve the platform's ability to manage and moderate content effectively.

Current Stage

I am currently at the **Project Proposal** stage, where I am defining the project's objectives, scope, and methodology.

Methodology: Using PACE Strategy

I am utilizing the **PACE** strategy (Plan, Analyze, Construct, Execute) to guide this project. This structured approach allows me to:

- Clearly outline project goals and deliverables.
- Systematically analyze data and derive insights.
- Construct effective models and visualizations.
- Execute the project efficiently while keeping stakeholders informed.

Next Steps

The remaining stages of the project will be completed in the coming week and will include:

- Understanding the Data
- Exploratory Data Analysis (EDA)
- Statistical Tests
- Regression Modeling
- Machine Learning Mode

PACE strategy document

Credit : Bhavan

Data Project Questions & Considerations



PACE: Plan Stage

- Who is your audience for this project?

The audience for this project will be the TikTok data team and other cross-functional teams.

- What are you trying to solve or accomplish? And, what do you anticipate the impact of this work will be on the larger needs of the client?

I am currently engaged in the development of a predictive model aimed at determining whether a video presents a claim or expresses an opinion.

- What questions need to be asked or answered?

What is the condition of the provided dataset? What variables will be the most useful? Are there trends within the data that can provide insight? What steps can I take to reduce the impact of bias? How can the model be adapted for different content types or formats? How will we collect feedback on the model's performance post-deployment?

- What resources are required to complete this project?

We require the project dataset, the Python notebook, and input from stakeholders to proceed.

- What are the deliverables that will need to be created over the course of this project?

The deliverables consist of a dataset prepared for exploratory data analysis, along with visualizations, a statistical model, regression analysis, and/or a machine learning model.

Project tasks

Following are a group of tasks your company's data team has determined need to be completed within this project. The data analysis manager has asked you to organize these tasks in preparation for the project proposal document. First, identify which stage of the PACE workflow each task would best fit under using the drop down menu. Next, give an explanation of why you selected the stage for each task. Review the following readings to help guide your selections and explanation: The PACE stages and Communicate objectives with a project proposal. You will later reorder these tasks within a project proposal.

1. Evaluating the model: **Execute**

Why did you select this stage for this task?

After the model has been constructed, data is run through to evaluate whether it meets the project's expectations and goals.

2. Conduct hypothesis testing: **Analyze** and **Construct**

Why did you select these stages for this task?

During the analyzing stage, it is determined that a statistical test will be used. During the construction phase, the test is carried out.

3. Begin exploring the data: **Analyze**

Why did you select this stage for this task?

During the analysis phase, you will gain a deeper understanding of the dataset and the information within it.

4. Data exploration and cleaning: **Plan** and **Analyze**

Why did you select these stages for this task?

Planning takes place when you first make choices about the methods needed. The cleaning process then takes place in the analyzing stage.

The PACE WORKFLOW

Credit : Bhavan

5. Establish structure for project workflow (PACE): Plan ▾

Why did you select this stage for this task?

Planning stage. Creating an initial project PACE document outlines the workflow and helps to plan how to best approach a project.

6. Communicate final insights with stakeholders: Execute ▾

Why did you select this stage for this task?

Communication is necessary at various points throughout a project. Final insights are shared with stakeholders in the execute phase of the data project workflow.

7. Compute descriptive statistics: Analyze ▾

Why did you select this stage for this task?

Investigating the statistics within data takes place during analysis.

8. Visualization building: Analyze ▾ and Construct ▾

Why did you select these stages for this task?

Visualization begins with data assessment and is created during the construction stage.

9. Write a project proposal: Plan ▾

Why did you select this stage for this task?

Planning stage. A project proposal is the initial document used to define a project.

10. Build a regression model: Analyze ▾ and Construct ▾

Why did you select this stage for this task?

During the analyzing stage, the model is examined in detail to be sure it will meet the needs of the task. The building of the regression model will take place in the construction phase.

The PACE WORKFLOW

Credit : Bhavan

11. Compile summary information about the data: **Analyze** ▾

Why did you select this stage for this task?

Inspecting a dataset to compile information would take place in the analysis phase.

12. Build machine learning model: **Construct** ▾

Why did you select this stage for this task?

The building of a data model would take place in the construct stage.

TikTok project proposal

Credit : Bhavan

Overview

TikTok seeks a way to use machine learning to extract claims or propositions within TikTok videos and comments and classify each data point as a claim/opinion.

Milestone	Tasks	Deliverables/Reports	Relevant Stakeholder (Optional)
1	<div>Establish structure for project workflow (PACE)</div> <div>Plan ▾</div>	<ul style="list-style-type: none">Global-level project document	Rosie Mae Bradshaw — DS Manager
1a	<div>Write a project proposal</div> <div>Plan ▾</div>		Mary Joanna Rodgers — PM Officer
2	<div>Compile summary information about the data</div> <div>Analyze ▾</div>	<ul style="list-style-type: none">Data files ready for EDA	Orion Rainier — DS
2a	<div>Begin exploring the data</div> <div>Analyze ▾</div>		Orion Rainier — DS
3	<div>Data exploration and cleaning</div> <div>Plan ▾ and Analyze ▾</div>	<ul style="list-style-type: none">EDA results	Orion Rainier — DS
3a	<div>Visualization building</div> <div>Analyze ▾ and Construct ▾</div>	<ul style="list-style-type: none">Tableau dashboard/visualizations	Mary Joanna Rodgers — PM Officer

TikTok project proposal

Credit : Bhavan

4	<p>Compute descriptive statistics</p> <p>Analyze ▾</p>	<ul style="list-style-type: none">Analysis of testing results between two important variables	Willow Jaffey — DS Lead
4a	<p>Conduct hypothesis testing</p> <p>Analyze ▾ and Construct ▾</p>		Orion Rainier — DS
5	<p>Build a regression model</p> <p>Analyze ▾ and Construct ▾</p>	<ul style="list-style-type: none">Regression model	Willow Jaffey — DS Lead
5a	<p>Evaluate the model</p> <p>Execute ▾</p>		Rosie Mae Bradshaw — DS Manager
6	<p>Build a machine learning model</p> <p>Construct ▾</p>	<ul style="list-style-type: none">Final model	Willow Jaffey — DS Lead
6a	<p>Communicate final insights with stakeholders</p> <p>Execute ▾</p>	<ul style="list-style-type: none">Report to all stakeholders	Willow Jaffey — DS Lead

Estimated Timeline for Project Milestones

Credit : Bhavan

Each milestone's timeline reflects an estimated completion time based on the skills and tasks required. In practice, working with real clients and data scientists often involves tight deadlines, so the timeline here is designed to be both intensive and realistic.

Milestone Timeline:

- **Milestone 1:** 1-2 days
- **Milestone 2:** 2-3 weeks
- **Milestone 3:** 1 week
- **Milestone 4:** 1 week
- **Milestone 5:** 1-2 weeks