

# Philippines' Mahalika fund

## 1a Classification of positions and arguments on the sample data (100 posts)

**PROMPT 1**  
You are a helpful research assistant that is an expert in classifying small texts such as social media posts.

Based on a dataset of Twitter posts, categorize the posts that talk about the Mahalika sovereign wealth fund bill in the Philippines based on the **position** below:

**Support:** Posts that support or defend the bill for its potential to promote development and economic growth, and capacity to invest in government projects outside of the legal frameworks. The posts may contain words such as investment, development, growth, future, profitable, and other related words.

**Opposition:** Posts that oppose or criticize the bill as another vehicle for corruption and theft of public funds for private gains. These posts also raise concerns about the mechanisms of the fund, the experts who will manage the funds, and the other issues about the implementation of the fund. The posts may contain words such as rob, scam, scandal, gamble, oligarchy, and other related words.

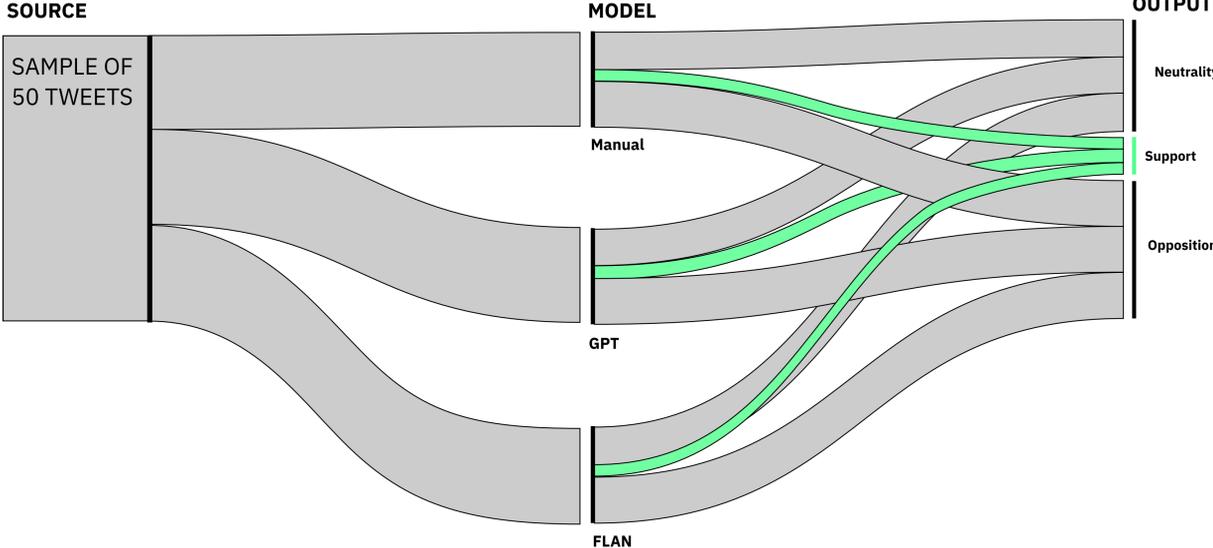
**Neutrality:** Posts that report about the bill without expressing criticisms or support. These posts also express openness to what the fund could bring about, as well as questions and suggestions about how to use the fund. These posts may contain news reports about the bill, and other unclear or ambiguous expressions.

**Others:** Posts that do not fall in the themes A to C.

Please create a new column "position" and populate it with the categories above. Please answer with A, B, C, and D.

Text: {...}

**Prompt**  
Context Persona Task Structure



## 1b Stance towards the fund (15k posts)

support (845)
others
neutral
opposition

## 1c Arguments for supportive stance (845 posts)

**PROMPT 2**  
You are a helpful research assistant that is an expert in extracting information from small texts such as social media posts.

Prompt: Select the option that best describes what this Tweet is saying:

- (A) it says that the Mahalika fund can hinder corruption in the government and lessen inefficiency in the system;
- (B) it says that the Mahalika fund can attract investors to put their money in the fund and invest in the country;
- (C) it says that the Mahalika fund can promote good governance among government officials and public sector workers;
- (D) it says that the Mahalika fund can create or generate jobs for people;
- (E) it says that the Mahalika fund will lead to supporting or financing the country's infrastructure;
- (F) it says that the Mahalika fund can boost the economy and bring wealth to the country;
- (G) None of the above, or uncertain answer.

Please answer with A, B, C, D, E, F, G or a combination

(E) it says that the Mahalika fund will lead to supporting or financing the country's infrastructure;

(B) It says that the Mahalika fund can attract investors to put their money in the fund and invest in the country;

(D) It says that the Mahalika fund can create or generate jobs for people;

(F) It says that the Mahalika fund can boost the economy and bring wealth to the country;

(G) None of the above

## Findings

### Methodology findings

- When running the stance classification prompt, Google flan (8% errors) performed slightly better than GPT 3.5 (10% errors) using the sample tweets (n = 50) based on the manually labeled categories by someone by local and issue knowledge.
- When running the argument classification prompt, the two models performed significantly different from each other: GPT 3.5 (8% errors) and Google flan (34% error), indicating that GPT 3.5 is far more adept at more conceptual categorization tasks than Google flan.
- On making decisions about classification, we observed three things: (1) Order of words in a text might matter in the labeling, (2) Tweets with longer text are more accurately labeled than the other tweets, (3) With limited textual cues, GPT 3.5 tend to overinterpret the texts
- In general, better performing prompt design involve more direct and simple-worded instructions that avoid conceptual load on the model

### Controversy findings

- Stance:** Opposition to the bill is a dominant trend (62.13 percent), while support is only X% of the Twitter data. Despite this, the state agenda still prevailed. This directed us to investigate the select tweets advancing Mahalika fund.
- Arguments:** Among the supporting tweets in the dataset, the majority of the tweets were labeled as "none of the above" (523/845 or 61.89 percent), while those tweets that were labeled as investment (270/845), allowing for multiple labels per tweet. Those labeled outside of the listed arguments are mostly news reports and tweets that identify factual information about the Mahalika fund.

## Conclusion

While the two models performed at par when classifying positions towards an issue, GPT 3.5 performed better at extracting arguments or providing conceptual summary from small texts such as tweets. For the controversy, we conclude that either the timeframe or the platform was not the target of state propaganda and exploring other timeframes or platforms might reveal more meaningful insights.