# **Objectives**

* This is to test each candidate’s unique problem solving skills which applies to the day-to-day tasks of an AI Engineer in Sprout AI Labs (SAIL)
* By doing the demo, this tests how well they can present their work with stakeholders whether in tech and non-tech.

# **Task Definition**

* We want a model which can identify the general sentiment of the given query. This can be useful for knowing how our customer feels towards a product or a situation.
* Specifications:
  + The Python script should be able to take in a user input then output the sentiment and confidence score
    - sentiments: “positive”, “neutral”, “negative”
    - confidence score up to 2 decimal places
    - example
      * Input: “I hate going to that restaurant”
      * Output: {“model\_output”:“negative”,”confidence\_score”: 98.42}
    - English queries only
  + A separate CSV file will be given to test the model’s accuracy. The output should be included in the repository to be submitted. The accuracy will be presented during the demo.
    - Input
      * CSV with ff. columns: text, expected\_sentiment
      * filename: *sentiment\_test\_cases.csv*
    - Output
      * CSV with ff. columns: text, expected\_sentiment, model\_output, confidence\_score
      * filename: *output\_sentiment\_test.csv*
  + Clean code practices will also be evaluated in the code review

# 

# **Constraints**

* Programming Language: Python only
* ML framework: Any from TensorFlow, Pytorch, HuggingFace etc. as long as it’s using Python as backend
* Models: Allowed to use open source models OR train your own (no cloud resource will be provided for this)
* Accuracy: Ideally this should be >80% so the model with the highest accuracy should be presented
* Test cases: Test is done on all the given test cases and should not be sampled. Preprocessing is done as necessary.

# **Test Cases**

* Use “sentiment\_test\_cases.csv” from this drive
* This is a modified version of the Sentiment140 - Test dataset containing 489 test cases.

# **Expected Output**

* Repository for scripts/outputs (public github repo)
  + All scripts used for training (only if applicable), inference, evaluation (taking in csv input and producing output), metrics computation
  + CSV output
* Slides for demo

Content Guidelines:

* + Model used and methodology
    - justify the use of the model
    - why is it the best for this use case
  + Quick code walkthrough
  + Demo of working inference
  + Test cases output with metrics (accuracy etc.)
  + Possible future improvements
* (optional) Additional slides/demo to showcase personal projects

# **Task Submission**

* Submission will be done through email
* Maximum of 4 calendar days to complete the task and submit the expected outputs
  + e.g. If the task was given on Monday, submission will be on Friday
  + If it falls on a weekend, an email is still expected which contains the requirements but presentation/interview will still be scheduled on a weekday
* Allowed to submit before the deadline for faster interview process