**ML Lab Exam Question:**

Analyze public sentiment on a specific niche topic (e.g., emerging technologies, niche markets, or local events) using social media data. The aim is to detect shifts in sentiment trends and uncover key drivers behind the sentiment.  
::

* Use a dataset related to tweets, Reddit posts, or forum discussions on a specific topic.
* Example dataset: Sentiment140 (contains tweets with sentiment labels).
* Alternatively, students can scrape recent social media data (if allowed).

:

* Clean the text (remove hashtags, mentions, URLs, and special characters).
* Tokenize, lemmatize, and perform sentiment analysis.

:

* Visualize trends in sentiment over time.
* Identify frequently mentioned topics or keywords using word clouds or topic modeling (LDA).

:

* Build a sentiment classifier using a machine learning algorithm like Logistic Regression, SVM, or Random Forest.
* Evaluate the model on classification metrics (precision, recall, F1-score).

:

* Cluster tweets/posts to identify latent sentiment groups or trends.
* Identify top influencers or users driving specific sentiments.

Dataset Link

<https://www.kaggle.com/datasets/milobele/sentiment140-dataset-1600000-tweets/data>

* Solution file is in this folder named as “ML\_Lab\_Exam.ipynb”
* Downloaded the zip file of dataset in local and then we proceed with the model development
* Created the model named as “Twitter\_sentiment\_model.sav”
* For the ML Ops we have some checkpoint which are given as under:
  + Checkpoint 1: Upload the whole model with dataset and development file on github for versioning.
    - Github repository link is:

<https://github.com/sury-git/Twitter_Sentiment_ML_Model>

* + Checkpoint 2: Create an .gitignore file to ignore other unnecessary files to upload on github => In Progress
  + Checkpoint 3: Develop an app.py to interact with model and predict the output => need to develop
  + Checkpoint 4: Create an repository on Docker and containerize the application on it
  + Checkpoint 5: Create the pipeline on gihub Action or Jenkins
* *I know to this point for further development need to gather knowledge.*