Project: Sentiment Analysis of Social Media Comments

Deliverable #1
MAIS Project
Team members: Hitaishi Dhoowooah, Carmin Vidé

Dataset selection
1. [Social Media Sentiments Analysis Dataset 📊](#)
2. [Social media emotion dataset](#)

Why did we choose this dataset?
<mark>Add what's good in this specific dataset</mark>
As social media influencers have become professional, advanced tools for analysing the trends and performances of their content are becoming decisive to help them boost their career. This tool will allow social media creators to draw insights from comments under their posts and better gauge the sentiments of their viewerbase. This will help them to tailor their future content to increase engagement.

Methodology
1. **Data preprocessing:**
   a. The dataset we chose is feasible as it is comprehensive. It contains detailed information about each comment and the associated sentiment analysis.
   b. Preprocessing…
2. **Machine learning model:**
   a. Using this dataset, we want to classify comments into 'positive', 'negative' and 'neutral' labels.
   b. The machine learning model we propose is logistic regression. Logistic regression offers these advantages:
      i. Simple and interpretable
      ii. Good baseline for binary classification
      iii. Works well with feature extraction like Bag of Words which suit the purposes of this project
      iv. Easy to implement and train on smaller datasets since we have a relatively small dataset

3. **Evaluation metric:**

Application:
Input: The user will copy the Instagram comment and paste the text as input to the model.
Output: The model classifies the comment and outputs 'positive', 'negative' or 'neutral'.