**COVID-19 Vaccine Tweet Sentiment Analysis**

- **1. The Problem We Solved**
  - Online chatter on COVID-19 vaccines is massive and shapes public perception.
  - Stakeholders need an automatic, real-time gauge of sentiment (negative, neutral, positive).
  - We built a classifier that labels each tweet's stance in seconds.

- **Dataset & Quick Exploration**
  - 45 k hydrated tweets covering Pfizer, Moderna, AstraZeneca, etc.
  - Manual sentiment labels: 1 = Negative, 2 = Neutral, 3 = Positive.
  - Class distribution ~ 28 % Neg, 42 % Neu, 30 % Pos.
  - Average length ≈ 150 chars → sequence length fixed at 40 tokens.

- **Machine-Learning Approach**
  - Text cleaning: remove URLs, mentions, hashtags, punctuation; lowercase.
  - Keras TextVectorization (20 k vocab).
  - Bidirectional LSTM (64 units) over 128-dim embedding.
  - Trained 5 epochs, batch 64, Adam optimizer.

- **Model Results (Test Set)**
  - Overall accuracy ≈ 82 %.
  - F1-scores – Negative 0.79 | Neutral 0.80 | Positive 0.87.
  - Confusion matrix: main confusion between Neutral & Negative.
  - Inference < 15 ms per tweet → ready for live dashboards.

- **Impact & Next Steps**
  - ✔️ Real-time monitoring for public-health agencies & vaccine manufacturers.
  - ✔️ Early detection of misinformation or reputation crises.
  - Next steps:
    - • Fine-tune BERT to push accuracy beyond 90 %.
    - • Deploy REST API & interactive dashboard.
    - • Expand to multilingual analysis for global coverage.