📌 Problem Statement

Social media platforms like Twitter and Instagram thrive on short, expressive communication, where **emojis** often serve as emotional, contextual, or emphatic cues. However, not all users explicitly use emojis, and many downstream NLP tasks (like sentiment analysis, intent detection, and user profiling) benefit from emoji-aware representations.

This research addresses the task of **predicting an appropriate emoji category from tweet text**, which can be used to:

- Autocomplete emoji suggestions while typing

- Enhance emoji-based emotion/sentiment analysis

- Provide richer embeddings for short text understanding

Given a dataset of tweets labelled with emojis, we aim to build a model that learns to predict a **semantic emoji category** (e.g., *love*, *celebrate*, *gesture*) from the tweet text alone.

🚀 Novelty

Unlike traditional emoji prediction models that:

- Treat each emoji as a separate label (leading to extreme class imbalance),

- Or ignore emojis entirely,

this approach introduces **semantic emoji grouping**, where 100+ emojis are clustered into **5 meaningful categories**:

- love, gesture, celebrate, hype, and emotion

This grouping:

- Reduces the label space from 127+ to 5

- Mitigates class imbalance using up sampling

- Makes the model more **interpretable** and **practically useful** for real-world applications

We also incorporate:

- A **BiLSTM + Attention** model to focus on important words

- Class distribution visualization and result metrics (accuracy, F1, confusion matrix)

🎯 Use Cases

1. **Emoji Recommendation Systems**:
   Enhance user experience in chat apps, keyboards, and comment boxes by suggesting relevant emoji categories during typing.

2. **Emotion Detection**:
   Emoji groups act as proxies for sentiment/emotion detection. Especially useful in **social media analytics** and **mental health monitoring**.

3. **Content Categorization**:
   Automatically tag posts with emojis for better indexing, retrieval, or filtering.

4. **Digital Marketing**:
   Predict popular or emotionally resonant emojis from campaign slogans or user posts.

5. **Assistive Technologies**:
   Help neurodivergent or visually impaired user's express emotions using AI-powered emoji suggestions.