**IMDb Movie Review Sentiment Analysis**

**Project Overview**

This project analyzes IMDb movie reviews using **text-based sentiment analysis, emotion classification, sarcasm detection, aspect-based sentiment analysis (ABSA), and emoji analysis**. The aim is to provide **multi-dimensional insights** into audience opinions, including:

* Overall sentiment (positive / negative / neutral)
* Underlying emotions (joy, anger, sadness, fear, surprise)
* Sarcasm detection (sarcastic / non-sarcastic)
* Aspect-level sentiment (acting, storyline, music, cinematography, direction)
* Emoji analysis (frequency, emotional mapping)
* Visualizations like word clouds, aspect sentiment plots, emotion trends

## ****Project Components****

### **1. Sentiment Analysis**

* **Tool:** TextBlob (fast CPU-friendly polarity scoring)
* **Output:**
  + Polarity score: -1 (very negative) to +1 (very positive)
  + Emotion label: joy, anger, neutral

### **2. Sarcasm Detection**

* **Method:** Keyword-based detection for quick analysis
* **Output:** sarcastic / not\_sarcastic

### **3. Aspect-Based Sentiment Analysis (ABSA)**

* **Aspects considered:** acting, storyline, music, cinematography, direction
* **Method:** TextBlob polarity mapped to aspects if mentioned in the review
* **Output:** Aspect sentiment dictionaries per review, e.g., {'acting':'positive','storyline':'negative'}

### **4. Emoji Analysis**

* **Method:** Extract emojis using Python emoji library
* **Output:**
  + Frequency of emojis across reviews
  + Optional mapping to basic emotions (joy, sadness, anger, fear, surprise)

### **5. Visualizations**

* Word clouds: overall, positive, negative, sarcastic, emotion-wise, aspect-wise
* Bar plots: aspect frequency, weighted sentiment
* Box plots: sarcasm vs polarity
* Count plots: emotion distribution, sarcasm distribution
* Emotion trends over time (if timestamp available)

## ****Implementation Pipeline****

1. **Load Dataset** → CSV with reviews
2. **Text Preprocessing** → Clean text, lowercase
3. **Sentiment & Emotion Analysis** → TextBlob polarity → map to emotion
4. **Sarcasm Detection** → Keyword matching
5. **Aspect Sentiment** → Extract aspects mentioned → assign polarity
6. **Emoji Analysis** → Extract emojis → count frequency → optional emotion mapping
7. **Visualization** → Word clouds, bar charts, box plots, count plots
8. **Predictions / Further Analysis**
   * Top positive & negative reviews
   * Aspect importance & weighted sentiment
   * Correlation between sarcasm & polarit

## ****Future Work / Improvements****

* Use **DistilBERT or RoBERTa** for faster and more accurate sentiment and emotion detection
* Implement **deep learning-based sarcasm detection** for higher precision
* Fine-tune **aspect sentiment extraction** using transformer models
* Incorporate **time series analysis** for release-week sentiment trends
* Enhance **emoji emotion mapping** for nuanced visual analysis
* Build an **interactive dashboard** (e.g., Streamlit or Dash) for exploring reviews