

Topic: **UCO BANK Facebook Comments Intelligence & Sentiment Mining.**

**Name of the Project:** FB Prism: Facebook Comment Insight Engine

**Objective:** To extract post descriptions and user comments from Facebook posts—whether displayed as modal pop-ups or full screen views (including video posts)—using Selenium automation with anti-detection features and layout adaptability, and then apply language translation, sentiment analysis, and topic categorisation to transform raw comments into structured, actionable insights suitable for analytics, reporting, and customer feedback mining.

**UCO Bank Relevance & Use Case:**

**UCO Bank has a strong presence on Facebook where it regularly shares product promotions, customer engagement posts, and service updates.**

**FB Prism helps UCO Bank automatically extract valuable user-generated content—such as comments and feedback—from both regular and modal-style Facebook posts.**

**By capturing this data, UCO Bank can:**

* **Monitor real-time customer sentiment for campaigns..**
* **Identify pain points or repeated complaints (e.g., *"UCO mPassbook not working"*)**
* **Gather direct customer praise and suggestions to inform product development**
* **Improve customer service response and engagement based on feedback trends**
* **Generate structured data for further analytics, NLP sentiment tagging, or chatbot training**

**This scraping tool transforms raw social media interactions into actionable insights, supporting UCO Bank’s digital transformation and customer-first strategy.**

**Features Identified:**

**1. Dual Layout Handling:  
   - Detects and handles both modal popups and fullscreen post layouts, including special cases for video posts.**

**2. Advanced Comment Scraper:  
   - Loads all comments via scrolling and expands replies.  
   - Extracts username, comment text, and relative timestamp, cleaning unnecessary UI noise.**

**3. Post Meta Extraction:  
   - Captures post descriptions and post dates using resilient multi-layered selectors.  
   - Fallback to JavaScript DOM parsing for stubborn layouts.**

**4. Dynamic Anti-Bot Evasion:  
   - Randomized screen sizes and user-agent rotation.  
   - CDP-based Selenium fingerprint masking.**

**5. Robust Session Handling:  
   - Periodically resets browser to avoid detection.  
   - Cookie-based login (for video/fullscreen posts).  
   - Auto-recovery from blocks, captchas, or redirects.**

**6. CSV Output Schema:  
   - url | post\_description | username | comment | date**

**7. Sentiment Analysis Integration:**

**- Automatically tags each comment as positive, neutral, or negative.**

**- Includes a sentiment confidence score for each tag.**

**- Helps assess customer emotion trends over time.**

**8. Topic Categorisation using Semantic Similarity:**

**- Each comment is matched to a predefined topic (e.g., "Mobile App Issues", "Loan Enquiry").**

**- Uses advanced transformer embeddings to ensure accurate theme mapping.**

**- Falls back to “Other / Unclassified” for unmatched or ambiguous content.**

| **url** | **post\_description** | **username** | **comment** | **date** | **translated\_**  **comment** | **topic\_**  **categorisation** | **topic\_**  **confidence** | **sentiment** | **sentiment\_score** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| [**https://www.facebook.com/996436069352835/posts/818918620437915**](https://www.facebook.com/996436069352835/posts/818918620437915) | **From checkout to success...** | **Saddam Raza** | **Not working UCO mpassbook** | **36w** | **Not working UCO mpassbook** | **mobile app issues** | **0.773** | **negative** | **0.748** |
| [**https://www.facebook.com/996436069352835/posts/805309695132141**](https://www.facebook.com/996436069352835/posts/805309695132141) | **With the UCO Arhatiyas Scheme...** | **Aman Dhiman** | **Good Scheme UCO Bank** | **39w** | **Good scheme UCO Bank** | **Other / Unclassified** | **0.488** | **positive** | **0.786** |

**Model Output:**

**Technical Overview:**

**Driver Initialization:**

* Uses full Selenium evasion stack with user-agent spoofing and randomized screen sizes.
* Adds support for cookie-based login and maximized video post layout handling.
* Fingerprint masking via Chrome DevTools Protocol (CDP) for anti-bot evasion.

**Comment & Post Scraping:**

* Multi-tier XPath + JS DOM logic for both modal and fullscreen layouts.
* Automatic fallback to JavaScript execution for post text if XPath fails.
* Expands all comments and replies, including dynamic loading via scrolling.

**Selectors Strategy:**

* Tries multiple XPath paths ranked by priority.
* Filters out UI noise such as “Like”, “Reply”, “Edited”, and similar Facebook artifacts.

**Data Cleaning & Deduplication:**

* Normalizes text using Unicode NFC.
* Cleans translation artifacts (e.g., excessive punctuation, replacement characters).
* Eliminates comments that duplicate post descriptions.

**Translation & Language Detection:**

* Uses NLLB-200 model to translate comments written in various Indian scripts to English.
* Applies script-aware language detection with mixed-script handling and fallbacks.
* Avoids translation for English comments and preserves originals for failed cases.

**Topic Categorisation:**

* Uses SentenceTransformer embeddings to match comments to a curated banking-related topic dictionary.
* Assigns each comment a best-matched topic along with a similarity confidence score.
* Defaults to "Other / Unclassified" for low-confidence or unmatched comments.

**Sentiment Analysis:**

* Analyzes sentiment polarity (positive, neutral, negative) using XLM-Roberta base model.
* Returns a sentiment label with associated probability score per comment.

**Resilience Features:**

* Automatically retries scraping 2–3 times on failures with exponential backoff.
* Periodic saving of intermediate results into CSV and JSON for recoverability.
* Continues processing even if some comments fail translation or classification, ensuring robustness.

**Conclusion**

FB Prism is a production-grade Facebook scraping system designed for maximum layout compatibility and real-world **comment mining**. It adapts to UI changes, mimics human behaviour to prevent blocks, and ensures you get clean, structured data for **analytics**, **feedback mining**, or **sentiment studies**.  
 With the addition of translation, topic categorisation, and sentiment analysis, FB Prism now enables **automated comment intelligence**. UCO Bank can not only extract feedback but also **understand customer tone and intent**, **group recurring issues**, and **generate insights for operational and marketing teams**—without manual tagging or analysis.