1. **Quy trình hoạt động chung:**

* Chọn ảnh từ gallery.
* Crop bằng thư viện uCrop.
* Dùng ML Kit nhận diện chữ.
* Gửi chữ đến Gemini API để viết lại.
* Hiển thị kết quả ra TextView.

(Nói về mục đích chung của 2 đứa về muốn tạo 1 app chuyên dụng về 1 chức năng chính: Lấy văn bản từ 1 đoạn văn trên mạng => Thay vì dùng tay gõ lại => Rewrite lại version mượt hơn để user có thể tham khảo)

* **Dẫn dắt vào ý chính**

Trong Android hiện có 2 nhóm công cụ AI. Nhóm đầu là **AI nhận diện** như ML Kit – giúp trích xuất thông tin sẵn có như văn bản, khuôn mặt. Nhóm thứ hai là **Generative AI** – như Gemini hay OpenAI – có thể tạo ra nội dung mới từ đầu. Với đề tài AI tạo sinh, nhóm mình sử dụng kết hợp cả hai: ML Kit để lấy dữ liệu đầu vào, Gemini để viết lại nội dung theo cách mới.

Mô hình **Transformer** là **kiến trúc mạng nơ-ron** cực mạnh, được thiết kế để **xử lý chuỗi dữ liệu** như văn bản, âm thanh, hoặc mã nguồn. Đây là nền tảng của các mô hình AI tạo sinh như **ChatGPT, Gemini, BERT, T5,...**

1. **ML Kit – Google Machine Learning Kit** [**ML Kit  |  Google for Developers**](https://developers.google.com/ml-kit/)

* ML Kit là một thư viện AI miễn phí của Google, giúp xử lý các tác vụ Machine Learning ngay trên thiết bị Android. Trong app đang dùng **ML Kit Text Recognition** để nhận diện văn bản từ ảnh.

**+ Vision APIs**: Video and image analysis APIs to label images and detect barcodes, text, faces, and objects.

**+ Natural language APIs**: Natural language processing APIs to identify and translate between 58 languages and provide reply suggestions.

1. **Lý do chọn ML Kit**:

* Dễ tích hợp vào Android Studio.
* Không cần kết nối Internet (chạy offline).
* Nhận diện văn bản nhanh, chính xác bằng camera hoặc ảnh có sẵn.
* Hỗ trợ tiếng Việt.
* ML Kit hoạt động cực nhẹ, phù hợp với cả máy cấu hình thấp – rất thích hợp cho các app phổ thông ở Việt Nam

1. **Cách dùng**:

* Người dùng chọn ảnh từ thư viện.
* Dùng thư viện **uCrop** để cắt vùng muốn nhận diện.
* Gửi ảnh đã crop cho ML Kit Text Recognition.
* Nhận lại nội dung văn bản dạng String.

1. **Gemini API – Google AI Gemini 2.0 Flash**

Gemini API là nền tảng AI mới của Google, kế thừa Bard, tích hợp khả năng tạo nội dung văn bản, hình ảnh, và đối thoại thông minh. Nhóm mình dùng phiên bản **Gemini 2.0 Flash** (Hỗ trợ dùng free cho dev trải nghiệm).

1. **Lý do chọn Gemini 2.0 Flash**:

* Xử lý văn bản nhanh hơn GPT-4 nhưng vẫn giữ chất lượng tốt.
* API dễ gọi, có tài liệu rõ ràng.
* Giá rẻ hơn các AI lớn khác như GPT.
* Hỗ trợ rewrite nội dung cực tốt.

1. **Cách nhóm mình dùng**:

* Sau khi có văn bản từ **ML Kit**/Tự nhập tay (Tùy vào user) , nhóm mình tạo prompt:

“**Rewrite the following paragraph with different wording but keep the same meaning**:

[nội dung văn bản]”

* Gửi request POST đến endpoint của Gemini (dạng JSON): **gửi yêu cầu lên server của Gemini** bằng giao thức **HTTP POST**,
* Dùng **OkHttpClient** để kết nối API: Đây là thư viện trong Android giúp **gửi yêu cầu HTTP**. Nó hoạt động như “**người đưa thư**”:
* Bạn viết nội dung (JSON) →
* OkHttp mang đến **địa chỉ của Gemini API** →
* Chờ nhận lại thư trả lời.
* Nhận lại kết quả và hiển thị ra TextView.

“Gemini 2.0 Flash không chỉ rewrite mượt, mà còn có thể mở rộng lên chatbot, phân tích nội dung, hoặc tóm tắt – rất tiềm năng cho các dự án Android có AI.”