



ĐỒ ÁN

XÂY DỰNG HỆ THỐNG TRÍCH XUẤT
THÔNG TIN DANH THIẾP BẰNG
PADDLEOCR KẾT HỢP VỚI NER VÀ
TINH CHỈNH KẾT QUẢ BẰNG LLM

Sinh viên: Đức Minh Hoàng

Mã sinh viên: 2021605732

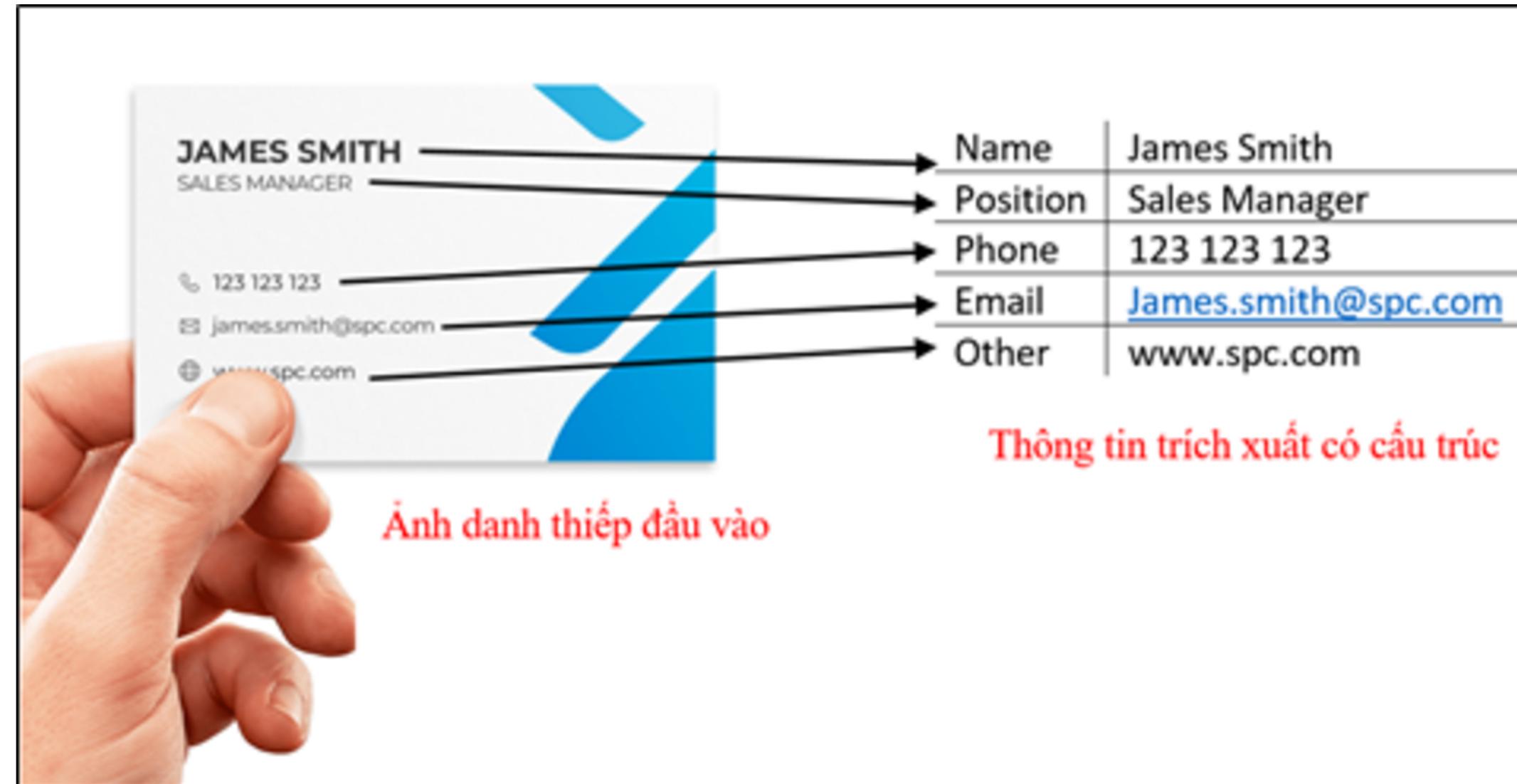
ĐỘNG LỰC & MỤC TIÊU



Làm sao tự động trích xuất thông tin danh thiếp nhanh, chính xác cho đội ngũ làm việc với khách hàng và đối tác?

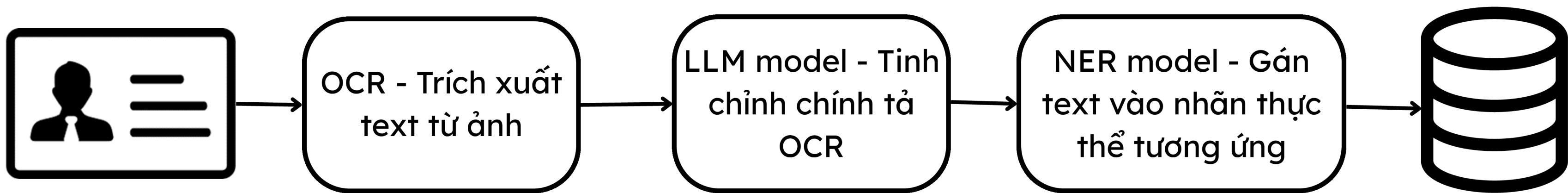
- Mục tiêu: Xây dựng hệ thống AI tự động trích xuất thông tin chính xác từ ảnh danh thiếp

PHÁT BIỂU BÀI TOÁN



Thách thức: Dữ liệu nhiễu, ảnh mờ,
lỗi nhận dạng từ OCR, đa nhãn.

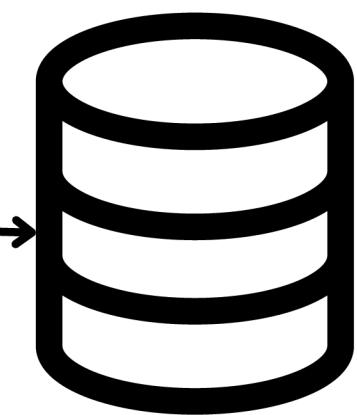
LUỒNG HỆ THỐNG



text ocr +
bounding
box

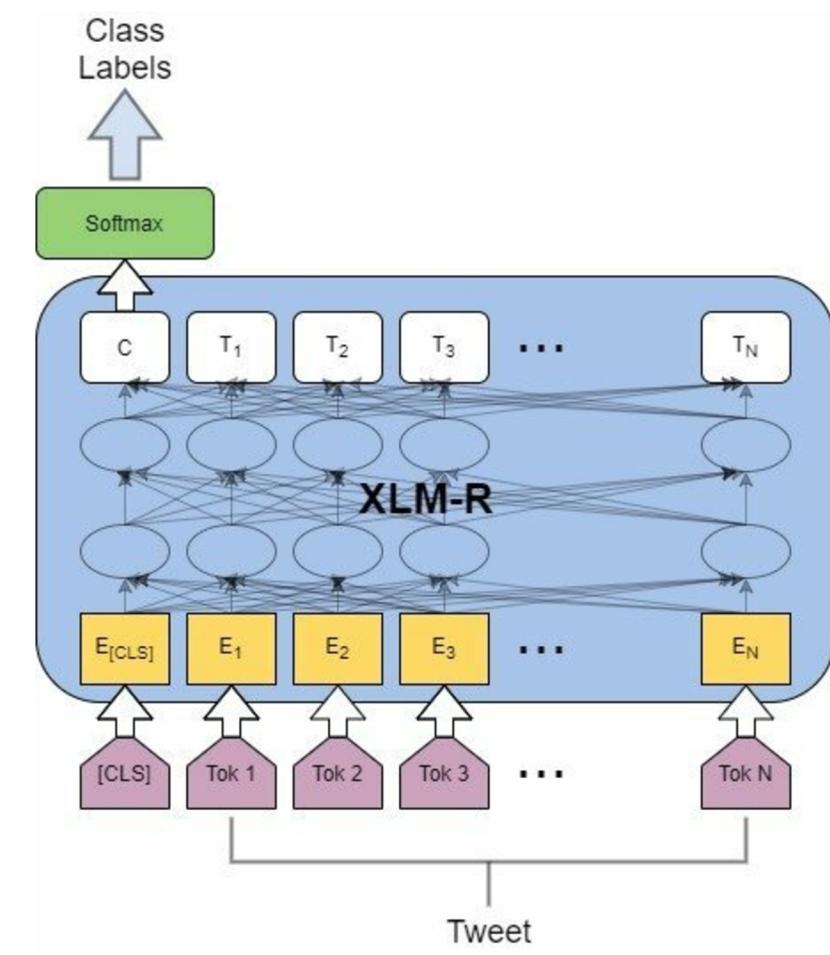
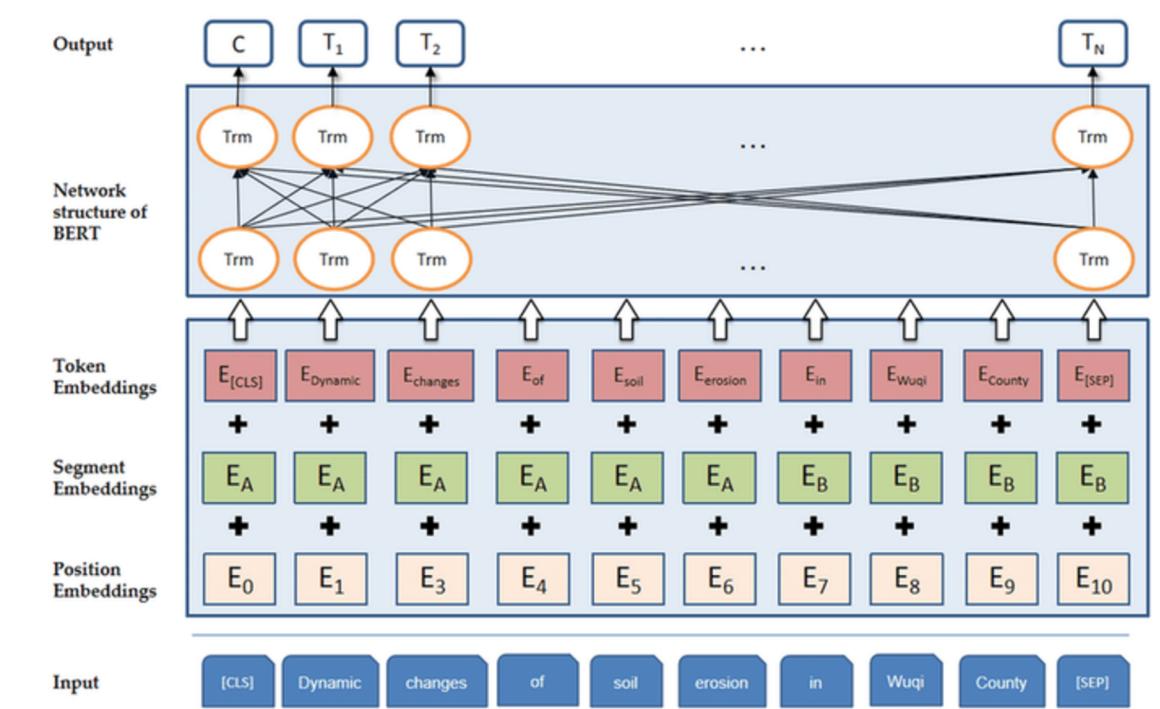
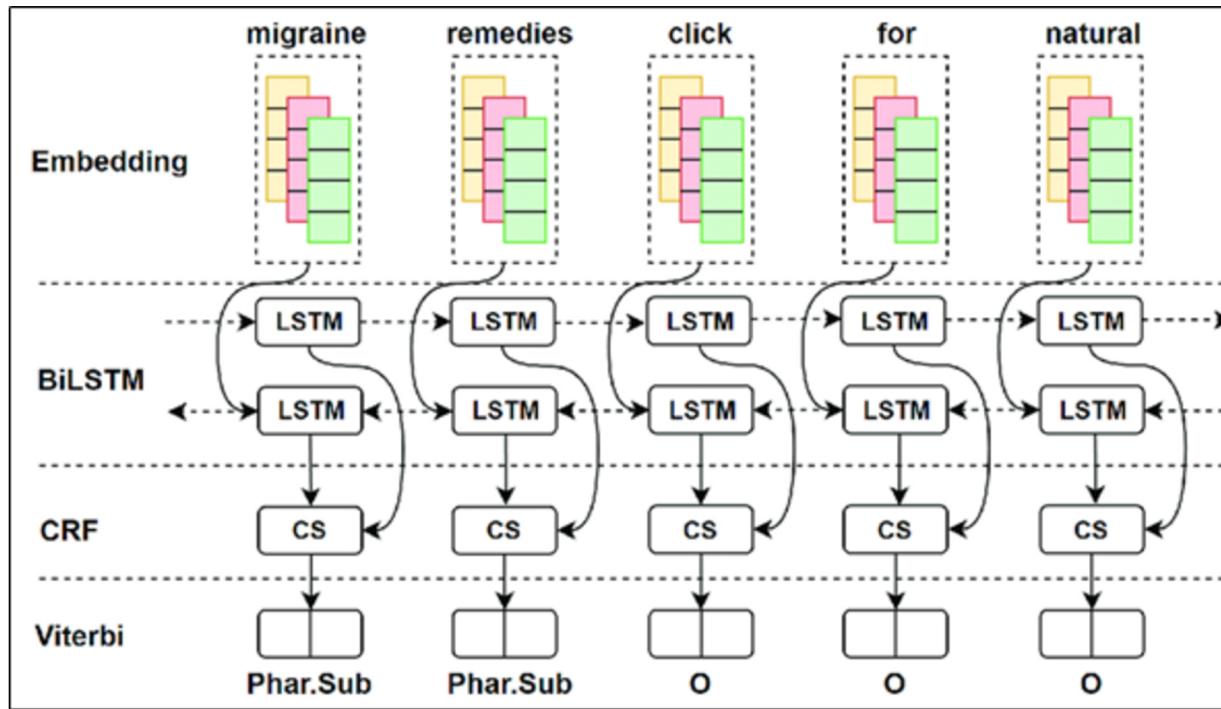


text sửa lỗi
chính tả +
bounding
box



Dữ liệu được
gán nhãn có
cấu trúc

NHẬN DẠNG THỰC THỂ CÓ TÊN



DỮ LIỆU THỰC NGHIỆM

Synthetic data format theo dạng Json theo chuẩn IOB (Inside-Outside-Beginning)

- Với 6000 cặp dữ liệu huấn luyện trên 100 trường hợp khác nhau
- Gồm 16 nhãn I-B của: Name, Position, Company, Address, Phone, Email, Department
- Và một nhãn O

```
{  
  "tokens": ["Nguyễn", "Văn", "An", "Manager", "ABC",  
  "Company", "123", "cầu", "Diễn", "an.nguyen@abc.com",  
  "0977123456"],  
  "labels": ["B-Name", "I-Name", "I-Name", "B-Position", "B-  
  Company", "I-Company", "B-Address", "I-Address", "I-  
  Address", "B-Email", "B-Phone"]  
}
```

FINETUNING MODEL

Sử dụng các pretrained models trên NER có sẵn trên HuggingFace

- PassbyGrocer/bert_bilstm_crf-ner-weibo
- dslim/bert-base-NER
- Davlan/xlm-roberta-base-ner-hrl

Các tham số huấn luyện

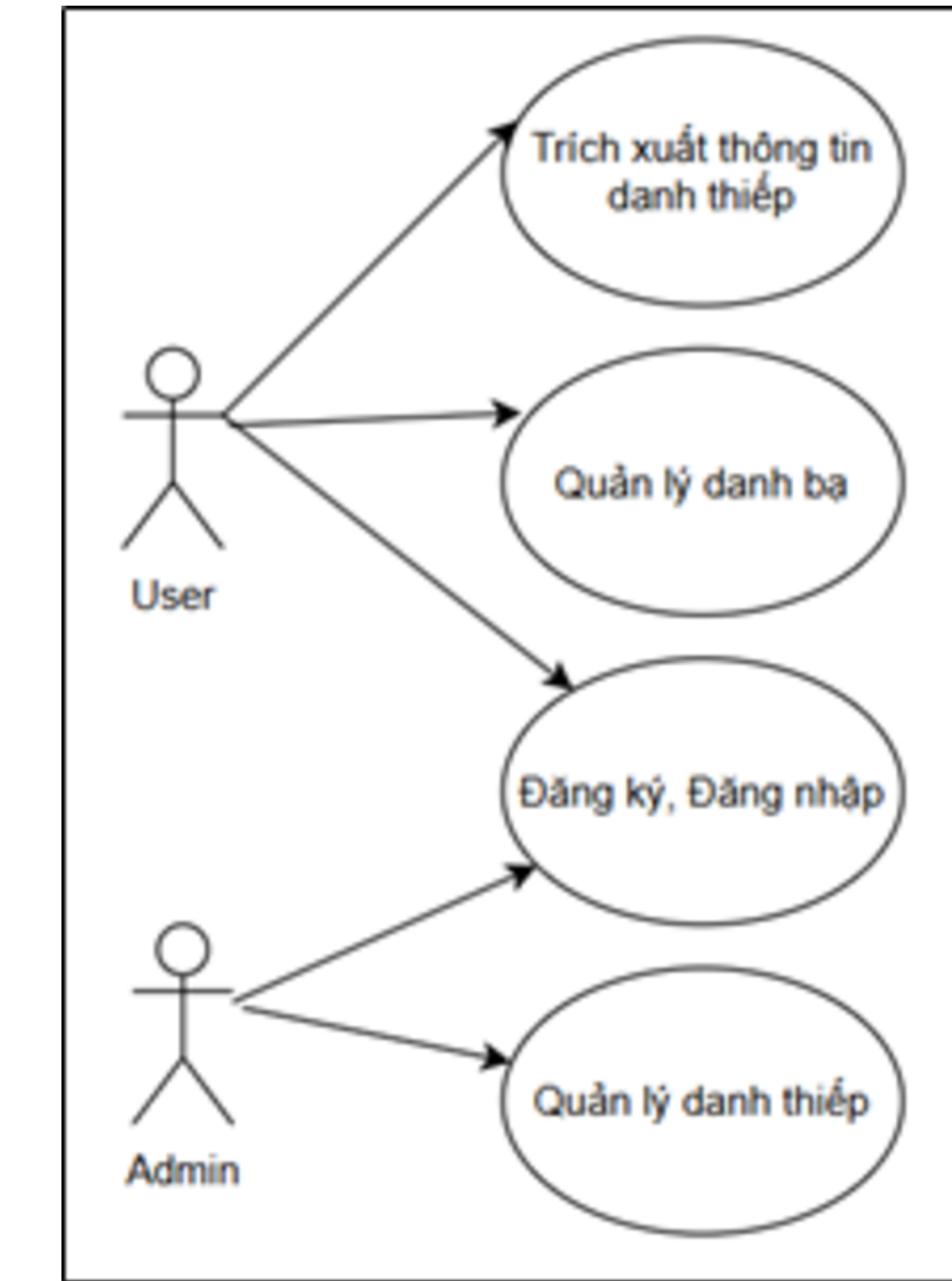
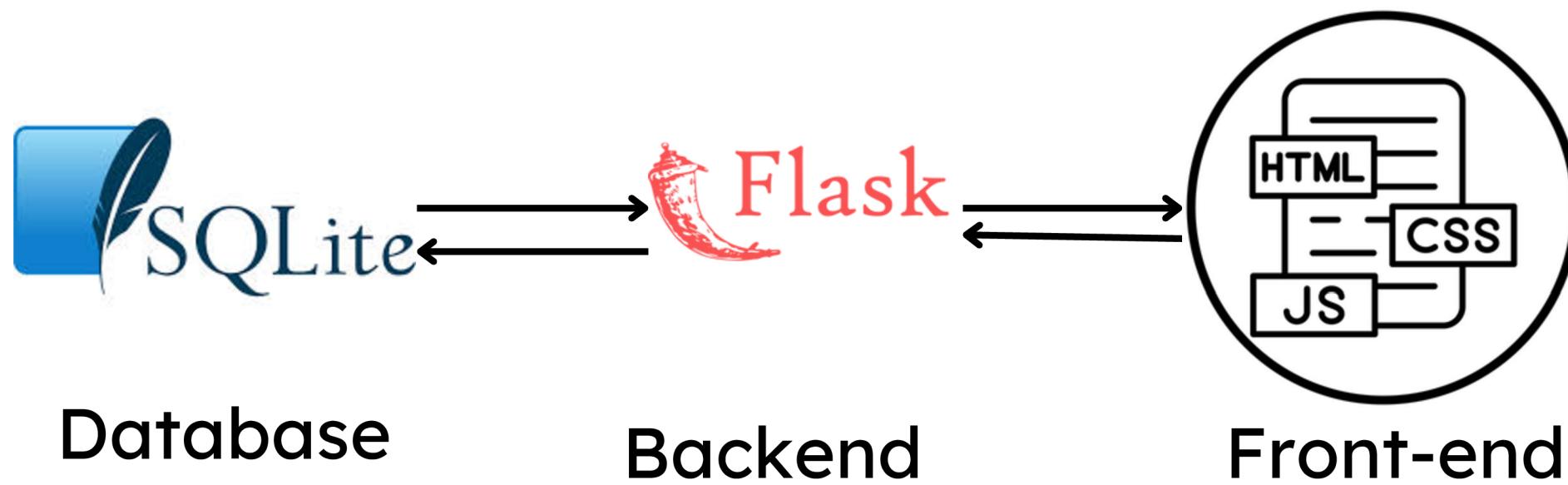
- Epoch: 100
- Batch size: 8
- Learning rate: 2e-5
- Weight decay: 0.01
- Tối ưu hóa: AdamW

KẾT QUẢ ĐÁNH GIÁ

Bảng so sánh kết quả trung bình từ $k = 10$ fold của ba mô hình BiLSTM-CRF, BERT và XLM-RoBERTa

Mô hình	Validation Loss	F1-scroce	Precision	Recall	Accuracy
BiLSTM-CRF	0.767s	0.593	0.553	0.553	0.872
BERT	0.771	0.612	0.557	0.683	0.869
XLM-RoBERTa	0.555	0.844	0.821	0.867	0.949

ỨNG DỤNG DEMO



KẾT LUẬN

- Đã xây dựng thành công hệ thống trích xuất thông tin danh thiếp tự động
- Kết hợp hiệu quả giữa thị giác máy tính (OCR), nhận xác thực thẻ có tên (NER) và LLM
- Ứng dụng có tiềm năng triển khai thực tế

Hạn chế:

- LLM còn nặng chưa tối ưu hóa chạy trên local (CPU)
- Chưa có hiệu chỉnh xoay với ảnh nghiêng

Định hướng:

- Xây dựng mô hình nhỏ gọn chính xác hóa chính tả từ OCR
- Bổ sung mô hình hiệu chỉnh xoay

Cảm ơn quý thầy cô đã
lắng nghe!