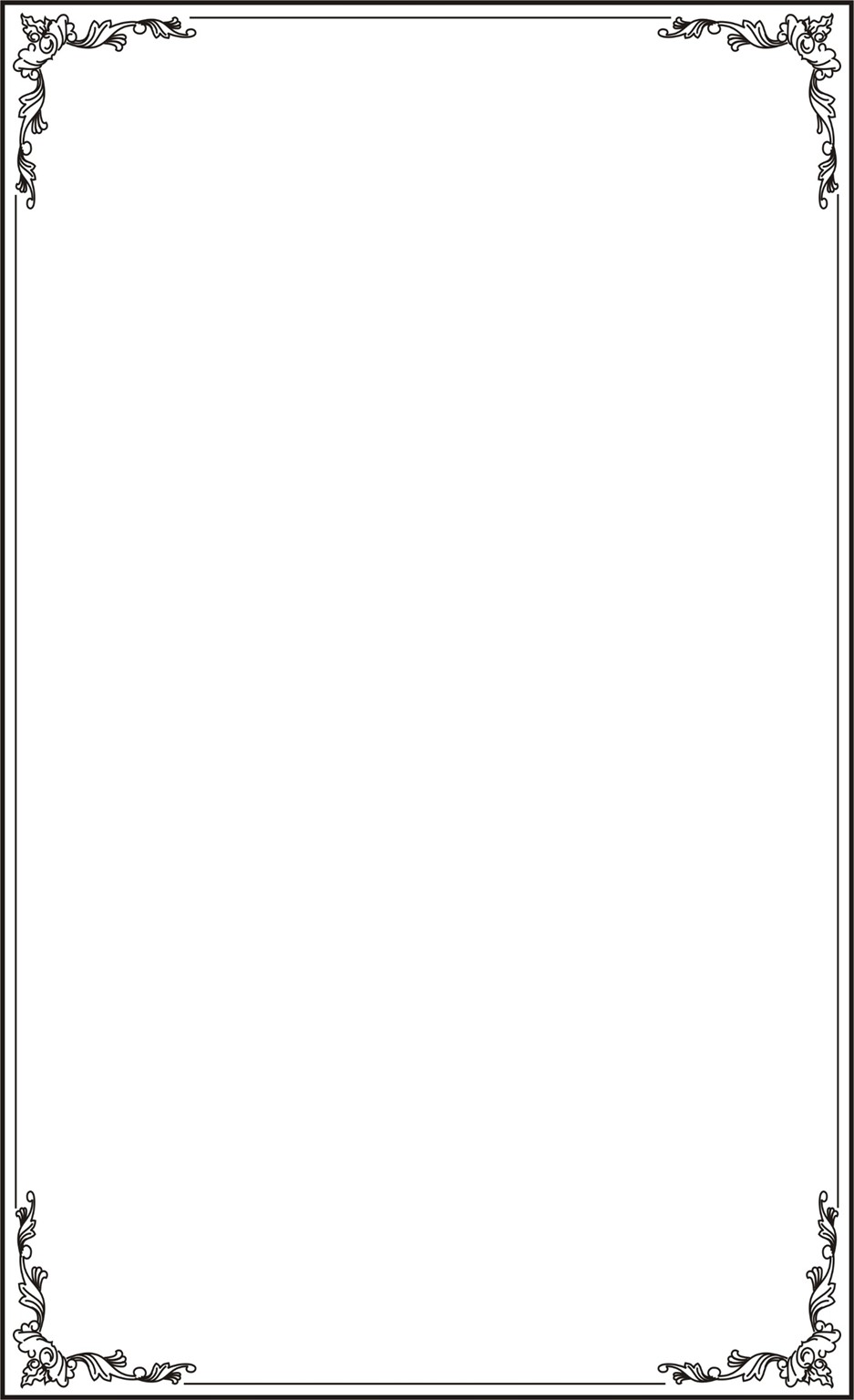
**HỌC VIỆN KỸ THUẬT QUÂN SỰ**



****

**ĐỒ ÁN TỐT NGHIỆP**

**CHUYÊN NGÀNH : CÔNG NGHỆ DỮ LIỆU**

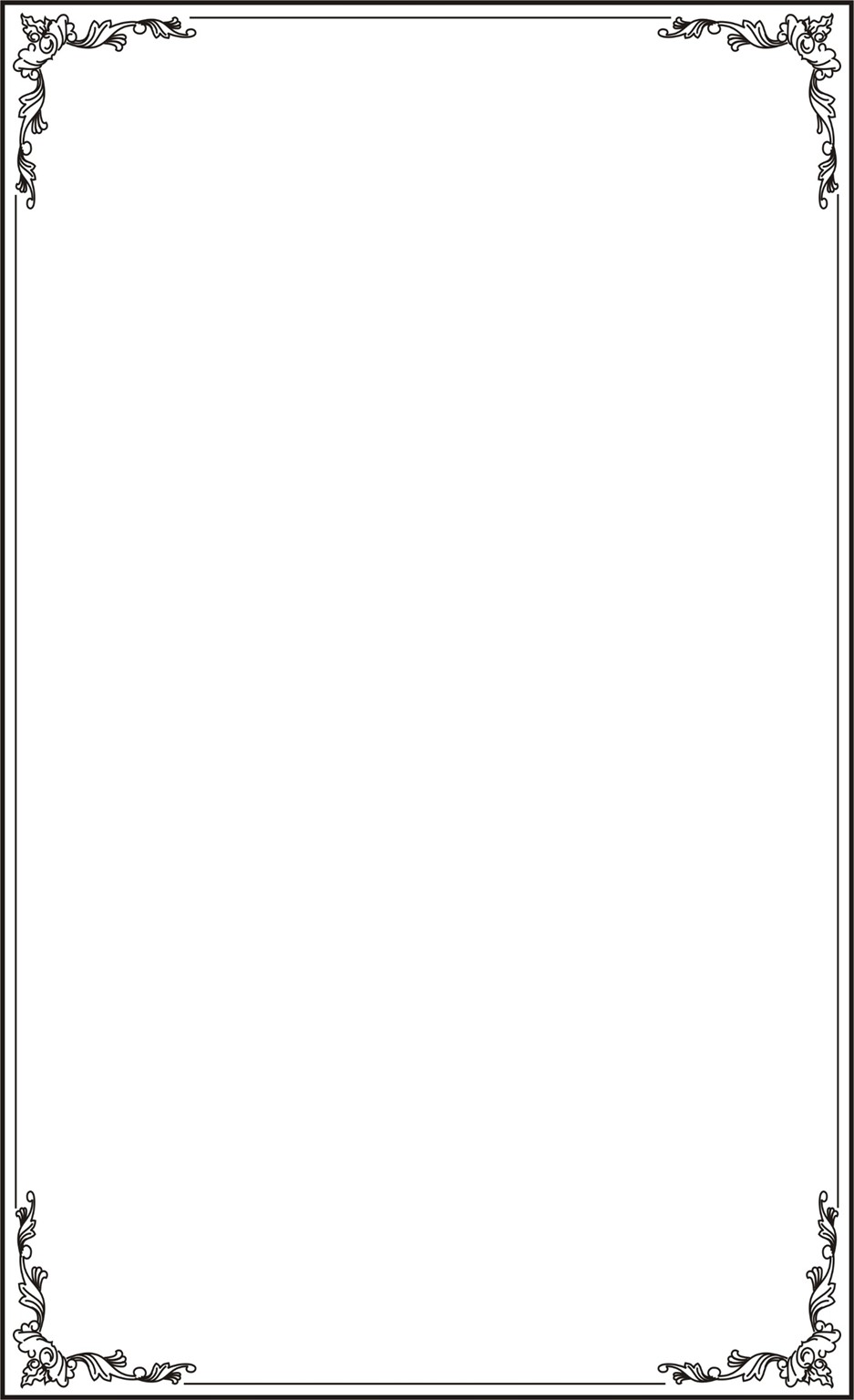
**Sinh viên thực hiện : Nguyễn Quốc Hiếu**

**Lớp : CNDL16B**

**Đề tài : Nghiên cứu xây dựng mô hình dịch song ngữ Việt – Lào**

**Hà Nội**

**HỌC VIỆN KỸ THUẬT QUÂN SỰ**



****

**ĐỒ ÁN TỐT NGHIỆP**

**CHUYÊN NGÀNH : CÔNG NGHỆ DỮ LIỆU**

**Sinh viên thực hiện : Nguyễn Quốc Hiếu**

**Lớp : CNDL16B**

**Đề tài : Nghiên cứu xây dựng mô hình dịch song ngữ Việt – Lào**

**GVHD: TS. Phan Việt Anh**

**Hà Nội**

HỌC VIỆN KỸ THUẬT QUÂN SỰ **CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM**

**KHOA CÔNG NGHỆ THÔNG TIN Độc lập - Tự do - Hạnh phúc**

**BỘ MÔN AN TOÀN THÔNG TIN**

**NHIỆM VỤ ĐỒ ÁN TỐT NGHIỆP**

Họ và tên: NGUYỄN QUỐC HIẾU Lớp: CNDL16B Khóa: 14

Ngành: Công nghệ thông tin

Chuyên ngành: Công nghệ dữ liệu

MỤC LỤC

MỞ ĐẦU

1. Tính cấp thiết, tình hình nghiên cứu liên quan đến đề tài
2. Mục tiêu, nhiệm vụ của đề tài
3. Phương pháp nghiên cứu
4. Đối tượng và phạm vi nghiên cứu
5. Bố cục của đồ án

CHƯƠNG 1 : GIỚI THIỆU BÀI TOÁN DỊCH MÁY

* 1. Bài toán dịch máy và ứng dụng

Trình bày bài toán dịch máy là gì, đầu ra, đầu vào

Ví dụ một số công cụ dịch

Các ứng dụng của nó trong thực tế

* 1. Các phương pháp

Trình bày các hướng tiếp cận đã được sử dụng để giải quyết bài toán này

-1.2.1 Dịch máy thống kê

-1.2.2 Dịch máy mạng nơron

CHƯƠNG 2 : Kiến thức cơ sở

2.1 Mạng nơ ron

2.2 Mạng nơ ron hồi quy và các biến thể

2.3 Kiến trúc encoder-decoder cho dịch máy

2.2 Nền tảng kiến thức cho hệ thống dịch máy

2.2.1 Kiến trúc Encoder-Decoder

2.2.2 Kỹ thuật Attention

CHƯƠNG 3 : XÂY DỰNG MÔ HÌNH

Mô tả cụ thể các thành phần của mô hình mà mình sử dụng. Ví dụ em chọn mô hình transformer

5-7 dòng mô tả khái quát mô hình, các thành phần

Các mục con mô tả chi tiết các thành phần. Tham khảo https://towardsdatascience.com/illustrated-guide-to-transformers-step-by-step-explanation-f74876522bc0

Chương 4 : HUẤN LUYỆN MÔ HÌNH VÀ XÂY DỰNG ỨNG DỤNG

4.1. Chuẩn bị dữ liệu

- Nêu các bước, giải pháp xử lý dữ liệu để có bộ dữ liệu

- Nếu các thống kê trên dữ liệu (số lượng mẫu, chiều dài của các câu, …), phân tích số liệu

4.2 Các độ đo

- Trình bày nội dung các độ đo dùng để đánh giá chất lượng mô hình

4.3 Huấn luyện

- Mô tả cách chia dữ liệu, các tham số khởi tạo: số lớp nơ ron, learning rate, số vòng lặp ….

- Phân tích kết quả huấn luyện

4.4. Xây dựng ứng dụng

- Trình bày chức năng, giao diện ứng dụng xây dựng được.

- Phân tích một số mẫu dữ liệu: hiển thị kết quả dịch một số câu, đánh giá mô hình tốt, không tốt trong những trường hợp nào (nếu có thể)

KẾT LUẬN

TÀI LIỆU THAM KHẢO