

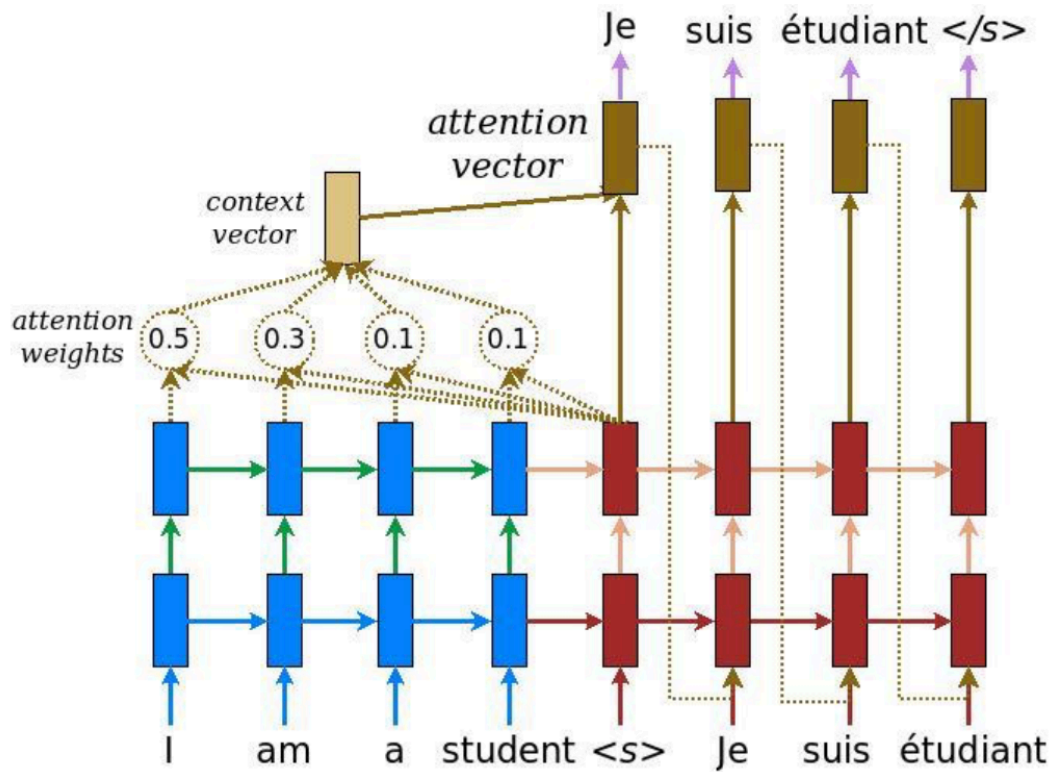
## CHƯƠNG 3: GIẢI PHÁP ĐỀ TÀI

### 3.1 TỔNG QUAN GIẢI PHÁP KIẾN TRÚC MÔ HÌNH

Nhóm sinh viên sử dụng mô hình kiến trúc đầu cuối (end-to-end hay còn được gọi với Sequence to Sequence Model – seq2seq) có ý tưởng từ bài báo *Sequence to Sequence Learning with Neural Networks* do nhóm tác giả đến từ google được ông bố vào năm 2014 tại Silicon Valley AI Lab đã trình bày ý tưởng cụ thể để xây dựng một mô hình mạng nơ-ron hồi quy tối ưu với hướng đi mới so với các hệ thống dịch máy truyền thống kết hợp cùng với cơ chế chú ý (Attention mechanism) từ bài báo *Effective Approaches to Attention-based Neural Machine* được thực hiện bởi nhóm tác giả đến từ đại học Stanford vào năm 2015. Từ đó nhóm sinh viên xây dựng một mô hình mạng nơ-ron hồi quy (Recurrent neural network) dùng để xây dựng một hệ thống dịch máy tiên tiến.

Mô hình seq2seq hoạt động dựa 2 mạng RNN kết hợp lại với một mạng RNN nhận nhiệm vụ mã hoá (encoder) câu đầu vào tiếng Anh thành một vector biểu diễn câu đầu vào và một mạng RNN giải mã (decoder ) có nhiệm vụ giải mã vector biểu diễn câu đầu vào và kết hợp với cơ chế chú ý (Attention mechanism) để giải mã câu đầu vào và cho ra kết quả câu đầu ra tiếng Việt tương ứng.

Kiến trúc tổng thể cho việc kết hợp 2 kiến trúc trên để xây dựng một mô hình dịch máy được minh họa cụ thể ở hình 3.1.



Hình 3.1: Tổng quan kiến trúc mô hình dịch máy

## 3.2 GIẢI PHÁP BIỂU DIỄN TỪ

### 3.2.1 Tổng quan về giải pháp

Để có thể sử dụng các mô hình Deep Learning (học sâu) phục vụ cho việc dịch máy, chúng ta cần biểu diễn các từ thành các số vì các mô hình chỉ làm việc với dữ liệu số. Vì thế dựa trên các kết quả tìm kiếm và thực nghiệm [], nhóm sinh viên đề xuất sử dụng Word Embedding (nhúng từ) dùng để biểu diễn các từ thành các vector số thực. Mô hình mà nhóm chọn là Word2vec với mục đích biểu diễn các từ tiếng Anh và tiếng Việt thành các vector số thực  $n$  chiều bằng nhau (mỗi chiều là một giá trị số thực) để phục vụ cho quá trình huấn luyện.

Word2vec là một mô hình học không giám sát (model unsupervised learning) nó dùng để thể hiện mối quan hệ giữa các từ, nó được kết hợp từ hai thuật toán Skip-gram và Continuous bag of words (CBOW). Ở đây nhóm sinh viên đề xuất sử dụng mô hình skip-gram cho biểu diễn từ. Với skip-gram, kích thước biểu diễn từ giảm từ kích thước bằng số từ trong bộ từ vựng xuống bằng chiều dài lớp ẩn. Hơn nữa các vector có ý nghĩa nhiều hơn về mặt mô tả mối quan hệ giữa các từ. Chi tiết mô hình đã được mô tả ở Chương 2.

### **3.2.2 Chi tiết giải pháp**

Nhóm sinh viên sử dụng đầu vào là tập dữ liệu được chia làm 2 tập tin chính chia làm 2 ngôn ngữ tiếng Anh và tiếng Việt.

### 3.3 GIẢI PHÁP XÂY DỰNG MÔ HÌNH DỊCH MÁY

#### 3.3.1 Tổng quan về giải pháp

Dựa trên các đánh giá thực tế và điều kiện phần cứng lẫn lượng dữ liệu (data) cho phép, nhóm sinh viên lựa chọn phương pháp học sâu (deep learning) để xây dựng mô hình mạng nơ-ron hồi quy (Recurrent neural network) trong mô hình dịch máy (machine neural translation). Mô hình được đào tạo từ đầu đến cuối từ những câu đã được biểu diễn dưới các nhúng từ (word embedding) để tạo ra các chuỗi đầu vào bộ mã hoá (encoder) và bộ giải mã (decoder). Do đó với lượng dữ liệu đủ lớn và khả năng tính toán, mô hình có thể tự học một cách chính xác để thực hiện việc dịch một câu từ tiếng Anh sang tiếng Việt.

#### 3.3.2 Mô hình mạng nơ-ron hồi quy (RNN) và khung huấn luyện

Cốt lõi của quá trình đào tạo một mô hình RNN là để nhận vào một văn bản tiếng Anh và tạo ra một văn bản tiếng Việt tương ứng. Để dễ hình dung ta có ví dụ một tập huấn luyện  $X = \{(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \dots\}$  với  $x$  là một vector các nhúng từ (word embedding) tương ứng với câu tiếng Anh đầu vào và  $y$  là một nhãn tức là một vector các nhúng từ (word embedding) tương ứng với câu tiếng Việt ở đầu ra. Mỗi câu tiếng Anh  $x^{(i)}$  là một chuỗi thời gian có độ dài  $T^{(i)}$ , trong đó mỗi đoạn thời gian nhất định là một vector nhúng từ (word embedding)  $x_t^{(i)}$ ,  $t = 1, 2, \dots, T^{(i)}$ . Mục tiêu của RNN là chuyển đổi đầu vào  $x$  thành một chuỗi xác suất ký tự cho nhãn  $y$ , với  $\hat{y}_t = P(w_t | x)$ , trong đó  $w_t$  thuộc các từ trong từ điển tiếng Việt và một vài ký tự đặc biệt khác.

Mô hình RNN được nhóm sinh viên chọn sử dụng là mô hình hồi quy với 3 thành phần chính. Thành phần đầu tiên là lớp nhúng từ (embedding), lớp (layer) này có nhiệm vụ chuyển các đầu vào của bộ mã hoá (encoder) và bộ giải mã (decoder) từ dạng int sang dạng nhúng từ (word embedding) để phục vụ cho công việc tính toán phía sau.

Thành phần thứ hai là bộ mã hoá (encoder), với bộ mã hoá chúng ta sử dụng Multi layer Bi-directional LSTM với số lượng layer và số lượng hidden units của LSTM cell được thiết lập trong param. Ngoài ra nhóm sinh viên còn sử dụng DropoutWrapper để thiết lập giá trị Drop Out cho các LSTM cell để tránh hiện tượng quá khớp (over-fitting) với dữ liệu huấn luyện.

### **3.3.3 Kết hợp mô hình với tìm kiếm chùm tia (Beam-search)**

### **3.5 GIẢI PHÁP XÂY DỰNG MÁY CHỦ**

Máy chủ (server) được nhóm sinh viên chọn Amazon EC2 làm máy chủ với mục đích tạo ra một cầu nối giữa mô hình đã được huấn luyện (model) và phía ứng dụng sản phẩm (client) – được xây dựng với React Native. Vì vậy trong giới hạn của khoá luận, máy chủ chỉ cung cấp duy nhất một giao diện lập trình (API) với chức năng chuyển đổi từ một văn bản (text) tiếng Anh thành một văn bản (text) tương ứng.

### **3.6 GIẢI PHÁP XÂY DỰNG ỨNG DỤNG**

Để ứng dụng hoá hệ thống dịch máy từ tiếng Anh sang tiếng Việt, nhóm sinh viên quyết định xây dựng web để ứng dụng kết quả của hệ thống vào một tình huống cụ thể có thể ứng dụng và thương mại hoá tốt.

Ứng dụng web do nhóm sinh viên xây dựng có chức năng chính là chuyển đổi văn bản tiếng Anh do người dùng nhập vào và đưa ra văn bản tiếng Việt tương ứng .

#### **3.6.1 Thiết kế giao diện ứng dụng**

### **3.6.2 Thiết kế kiến trúc ứng dụng**

## **3.7 TỔNG KẾT**

Thông qua chương 3, sinh viên đã làm rõ được các giải pháp cụ thể cho từng phần trong hệ thống dịch máy từ tiếng Anh sang tiếng Việt, hướng xây dựng máy chủ và cả ứng dụng trên nền tảng web.

Nhóm sinh viên đã trình bày một hệ thống dịch máy từ tiếng Anh sang tiếng Việt dựa trên việc học sâu (deep learning) từ đầu đến cuối có khả năng vượt trội và hiện đại trong hiện đại. Nhóm sinh viên tin rằng phương pháp này sẽ tiếp tục được cải thiện với các mô hình mới hơn, đơn giản hoặc phức tạp hơn khi tận dụng được sức mạnh tính toán phần cứng và kích thước dữ liệu được tăng thêm trong tương lai.

Chương kế tiếp nhóm sinh viên sẽ trình bày về các thư viện, công cụ và những khó khăn cụ thể nếu có cho các giải pháp đã trình bày ở chương này.