



# **MÔ HÌNH TRANSFORMER WITHOUT TEARS: CẢI THIẾN CHUẨN HÓA CỦA KỸ THUẬT SELF-ATTENTION**

**Trần Vĩ Hào - 19521482**

**Trương Quốc Bình - 19521270**

**Lê Thành Đạt - 17520332**

# Tóm tắt

- Lớp: CS519.M11.KHCL
- Link Github của nhóm: <https://github.com/datthanhle/CS519.M11.KHCL>
- Link YouTube video: <https://youtu.be/ky-eyqc82Bk>
- Ảnh + Họ và Tên của các thành viên



Trần Vĩ Hào

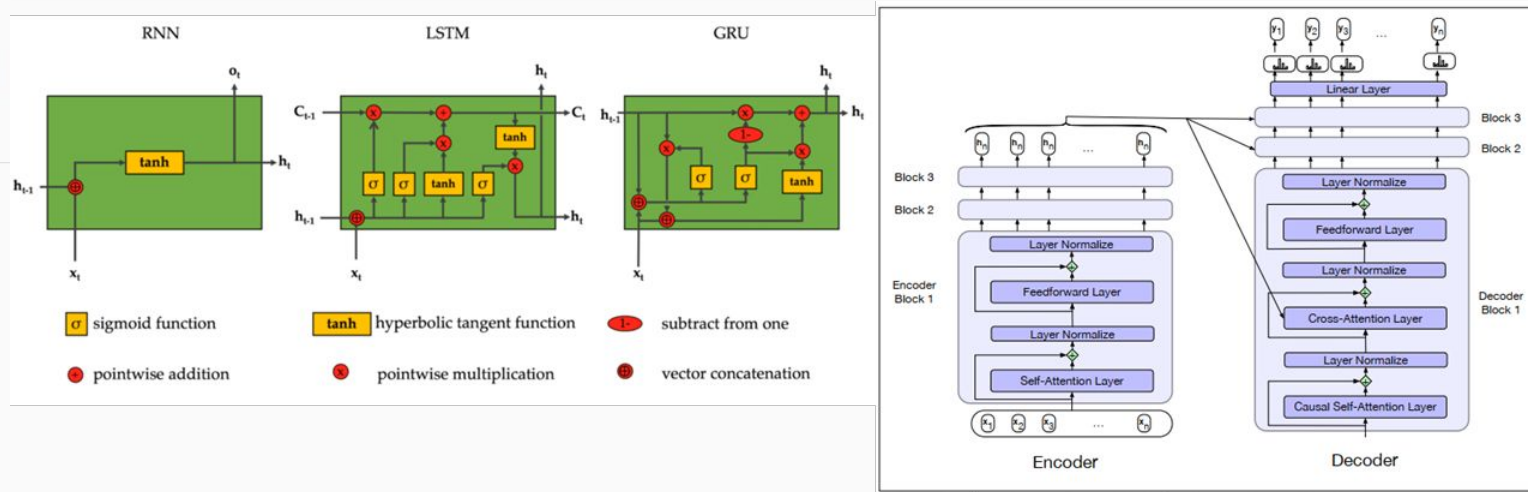


Trương Quốc Bình



Lê Thành Đạt

# Giới thiệu



INPUT

OUTPUT

INPUT

OUTPUT

Tiếng Anh

How are you?

Tiếng Việt

Bạn khỏe không?

OR

Tiếng Việt

Tôi là một sinh viên

Tiếng Anh

I'm a student.

# Mục tiêu

- Nghiên cứu thuật toán Transformer hiện có và cải tiến phương pháp chuẩn hóa của nó nhằm cải thiện quá trình huấn luyện và tạo ra phiên bản ***Transformer without tears***.
- Phát triển chương trình ứng dụng minh họa từ phiên bản ***Transformer without tears*** nhưng bước đầu chỉ với hai loại ngôn ngữ chính là tiếng Anh và tiếng Việt.
- Tạo ra một bộ dữ liệu song ngữ Anh-Việt mới.

# Nội dung và Phương pháp

- Nghiên cứu thuật toán Transformer truyền thống, huấn luyện mô hình để xác định độ đo BLEU trên tập dữ liệu IWSLT'15 English-Vietnamese.
- Nghiên cứu và áp dụng các kỹ thuật Postnorm, Prenorm, Scalenorm và Fixnorm vào thuật toán Transformer, so sánh và đánh giá các mô hình.
- Tự xây dựng bộ dữ liệu mới BHD-EnVi gồm khoảng một triệu câu tiếng Anh lẫn tiếng Việt.
- Nghiên cứu các kỹ thuật tăng cường dữ liệu (Data Augmentation) trong Xử lý ngôn ngữ tự nhiên để hỗ trợ cho việc xây dựng bộ dữ liệu mới.
- Huấn luyện mô hình **Transformer without tears** (phiên bản Transformer mà chúng tôi tự phát triển) sử dụng hai bộ dữ liệu IWSLT'15 English-Vietnamese và BHD-EnVi để so sánh và đánh giá các kỹ thuật đã sử dụng.
- Xây dựng chương trình ứng dụng minh họa.

# Nội dung và Phương pháp

- Tìm hiểu bản đồ nhận diện cho Transformer dựa theo một số kết quả đã được thực nghiệm trước đó, việc sử dụng Prenorm cho cả hai module encoder và decoder ảnh hưởng gì đến tỷ lệ chia của đầu ra? Tìm câu trả lời cho câu hỏi: “Liệu việc thay thế Postnorm bằng Prenorm sẽ hiệu quả hơn trong quá trình huấn luyện so với phương pháp gốc hay không?”
- Tìm hiểu cách khởi tạo trọng số của Postnorm dựa trên ý tưởng từ Glorot and Bengio.
- Tìm hiểu về hai kỹ thuật Scalenorm và Fixnorm, áp dụng hai kỹ thuật này vào việc thay thế các tham số scale và shift của Layernorm.
- Tìm hiểu cách đánh giá một mô hình dịch máy, cụ thể ở đây là Transformer và Transformer without tears bằng độ đo Bilingual Evaluation Understudy (BLEU).

# Nội dung và Phương pháp

- Chúng tôi tạo ra một bộ dữ liệu mới tên là BHD-EnVi bằng cách tổng hợp lại từ các bộ dữ liệu đã được công khai trên mạng, ngoài ra chúng tôi cũng sẽ thu thập thêm dữ liệu mới bằng cách crawl dữ liệu từ các trang web phim song ngữ, sách song ngữ, v.v... kết hợp với các phương pháp tăng cường dữ liệu.
- Huấn luyện mô hình Transformer without tears với từng trường hợp cụ thể (PostLayer, PreLayer, PreFixLayer, PreFixScale) chạy trên bộ dữ liệu IWSLT'15 English-Vietnamese và BHD-EnVi, so sánh và đánh giá kết quả dựa trên độ đo BLEU.
- Xây dựng chương trình ứng dụng trên nền Web cho phép người dùng nhập đầu vào một câu tiếng Việt và xem đầu ra câu tiếng Anh được dịch tương ứng với đầu vào, hoặc ngược lại.

# Kết quả dự kiến

- Báo cáo các phương pháp và kỹ thuật của mô hình Transformer without tears mà chúng tôi phát triển được sử dụng trong bài toán dịch máy. Kết quả thực nghiệm, đánh giá, so sánh các phương pháp với nhau và với mô hình Transformer truyền thống.
- Tập dữ liệu BHD-EnVi gồm một triệu câu tiếng Anh lẫn tiếng Việt sử dụng cho bài toán.
- Chương trình minh họa dịch máy, tương tự như Google Translate.



# Tài liệu tham khảo

- [1] Jingdong Wang, Ke Sun, Tianheng Cheng, Borui Jiang, Chaorui Deng, Yang Zhao, Dong Liu, Yadong Mu, Mingkui Tan, Xinggang Wang, Wenyu Liu, Bin Xiao: Deep High-Resolution Representation Learning for Visual Recognition. IEEE Trans. Pattern Anal. Mach. Intell. 43(10): 3349-3364 (2021).
- [2] Konstantin Sofiiuk, Ilia A. Petrov, Anton Konushin: Reviving Iterative Training with Mask Guidance for Interactive Segmentation. CoRR abs/2102.06583 (2021).
- [3] saic-vul, "ritm\_interactive\_segmentation," 2021. [Online]. Available: [https://github.com/saic-vul/ritm\\_interactive\\_segmentation.git](https://github.com/saic-vul/ritm_interactive_segmentation.git)
- [4] Xavier Glorot, Yoshua Bengio: Understanding the difficulty of training deep feedforward neural networks. AISTATS 2010: 249-256.