**TF-IDF** is a widely used method in Natural Language Processing (NLP) and text mining [[17], [18]](#re17). It helps assess how significant a word is within a document, based on both the frequency of that word in the document and the frequency of the word in the entire set of documents. In other words, TF-IDF allows us to determine which words are more prominent in a document compared to other documents in the same set.

In this study, we use TF-IDF as a **method to supplement information about the frequency of word occurrences, which helps to enhance the ability to distinguish between important and unimportant words. This contributes to improving the accuracy of fake news classification without interrupting the context learning process of Transformer models.**

**Specifically, TF-IDF is integrated as an additional layer into the features of the [CLS] token generated by Transformer models. This [CLS] token is the result of the hidden state from the final layer and represents the entire sentence, containing context information synthesized from all the words in the sentence. This approach will keep the process of learning context and the relationships between words in the sentence completely handled by the Transformer without being affected by TF-IDF.**

**TF-IDF** là một phương pháp được sử dụng rộng rãi trong xử lý ngôn ngữ tự nhiên (NLP) và khai thác văn bản [17], [18]. Nó giúp đánh giá mức độ quan trọng của một từ trong một tài liệu, dựa trên tần suất của từ đó trong tài liệu và tần suất của từ trong toàn bộ tập hợp tài liệu. Nói cách khác, TF-IDF cho phép chúng ta xác định từ nào nổi bật hơn trong một tài liệu so với các tài liệu khác trong cùng tập hợp

Trong nghiên cứu này, chúng tôi sử dụng TF-IDF như một phương pháp để **bổ sung các thông tin về tần suất xuất hiện của các từ, giúp tăng cường khả năng phân biệt giữa các từ quan trọng và không quan trọng. Việc này góp phần tăng độ chính xác trong việc phân loại tin giả mà không làm gián đoạn quá trình học ngữ cảnh của các mô hình Transformer.**

**Cụ thể, TF-IDF được tích hợp như một tầng bổ sung vào các đặc trưng của token [CLS] đã được các mô hình Transformer tạo ra. Token [CLS] này là kết quả từ hidden state của layer cuối cùng và đại diện cho toàn bộ câu, chứa thông tin ngữ cảnh tổng hợp từ tất cả các từ trong câu. Cách tiếp cận này sẽ giữ cho quá trình học ngữ cảnh và quan hệ giữa các từ trong câu được xử lý hoàn toàn bởi Transformer mà không bị ảnh hưởng bởi TF-IDF**

Sau khi hoàn thành các bước xử lý dữ liệu, bao gồm việc làm sạch và cân bằng dữ liệu như đã thảo luận trong các phần 3.2 và 3.3, chúng tôi đã thu được một bộ dữ liệu bao gồm các bài đăng trên mạng xã hội và tin tức từ các nguồn của Việt Nam. Bộ dữ liệu này chứa 1,406 mẫu, bao gồm cả tin tức thật và tin tức giả từ nhiều lĩnh vực khác nhau. **Trong đó gồm có 180 mẫu tin giả chúng tôi lấy từ bộ dữ liệu “VNFD - Vietnamese Fake News Datasets”. Phần tin tức còn lại gồm 1226 mẫu chúng tôi tự thu tập gồm các tin tức thật giả lẫn lộn. Trong đó có 845 tin tức được gán nhãn dựa trên các thông tin, kiến thức hoặc sự kiện thực tế đã được chứng minh là đúng hoặc sai, 381 tin còn lại chúng tôi đã gán nhãn thủ công bằng cách suy luận từ các kiến thức, sự kiện thực tế, hoặc dựa theo quan điểm của cộng đồng tại thời điểm thông tin được lan truyền.**

Chúng tôi sau đó đã chia ngẫu nhiên bộ dữ liệu thành hai phần, một phần 80% cho huấn luyện và 20% cho kiểm tra, với 1,124 mẫu dành cho huấn luyện và 282 mẫu dành cho kiểm tra. Phương pháp này giúp mô hình học được các mẫu hiệu quả, tăng khả năng đạt hiệu suất tốt hơn trên dữ liệu mới chưa thấy. Bộ dữ liệu có thể truy cập tại liên kết sau: [https://github.com/huynhtuan0106/Vietnamese-News-Dataset](https://github.com/huynhtuan0106/Vietnamese-News-Dataset" \t "_new).

**Prolem 3:**

Tôi đã thay đổi và chỉnh sửa các câu và cách diễn đạt sai được bạn nhắc đến trong bản thảo.

Tôi cũng đã chỉnh sửa lại số lượng mẫu trong bộ dữ liệu và cung cấp URL để truy cập đến bộ dữ liệu của chúng tôi.