



Welcome!!

Machine Learning



DETECTION FAKE NEWS

NT005





Giảng Viên:

Phạm Nguyễn Trường An – Lê Đình Duy

Thành viên:

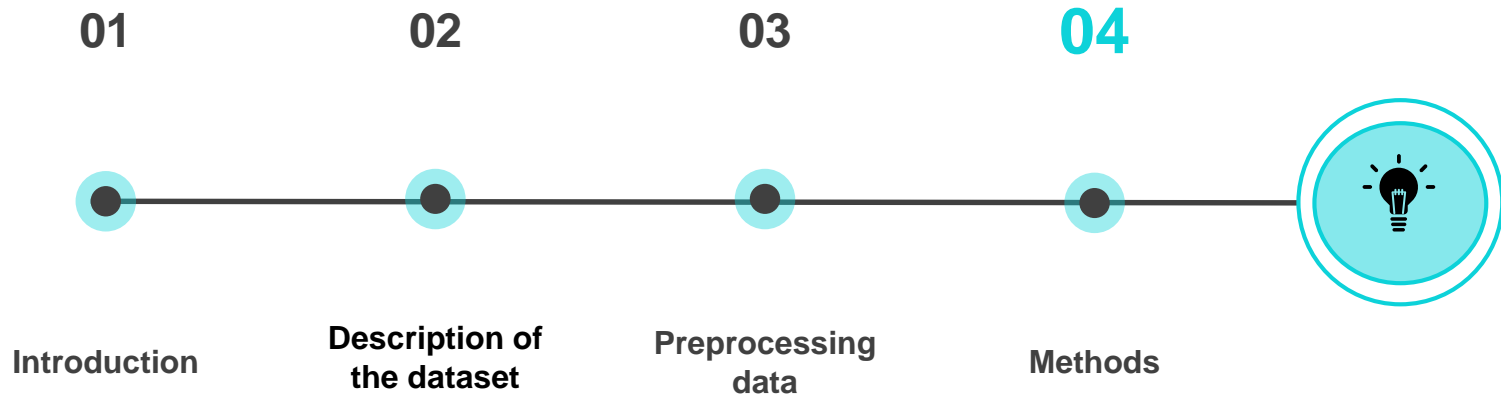
Nguyễn Văn Thái An - 17500216

Lê Hoàng Khánh – 17520623

Ngô Anh Vũ – 17522172

Nguyễn Hữu Nghĩa - 18521144

Detection Fake News



Introduction Fake News & Real News



Fake news là tình trạng thông tin bị sai lệch hoặc chỉ nêu ra một nửa sự thật và tốc độ lan truyền tin tức rất cao dẫn đến hiểu lầm trong dư luận, thậm chí kích động bạo lực.

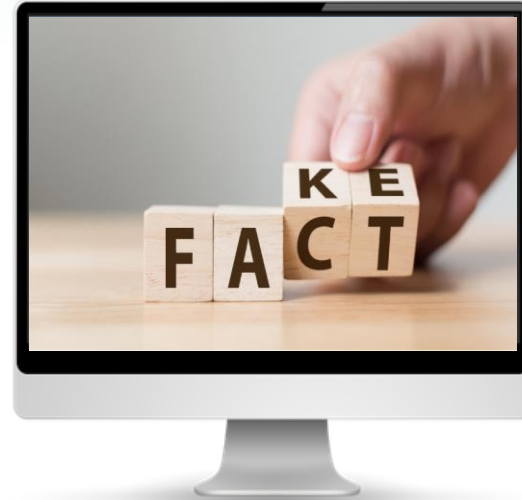
Hướng giải quyết từ đó nhóm đặt ra bài toán để nhận biết tin tức là thật hay giả dựa vào các mẫu tin tức đã được huấn luyện trước đó.

Input - Output



Input

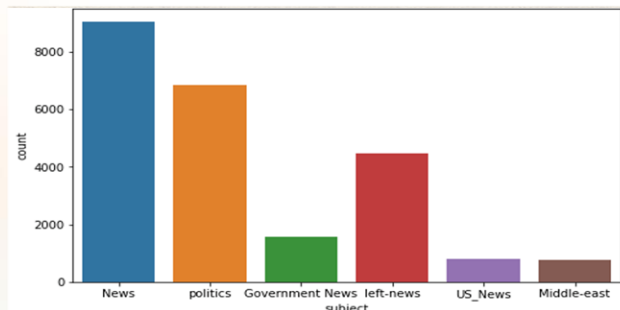
Một đoạn văn bản bất kì



Output

Kết quả thông báo đoạn tin đó là fake hay real

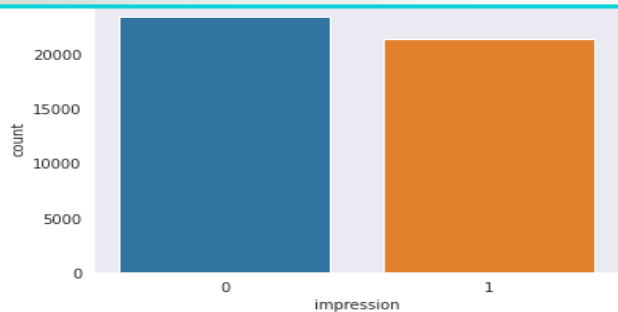
Introduction data



Điểm dữ liệu bao gồm 4 thuộc tính:

- Tiêu đề (title)
- Nội dung (text)
- Thể loại (subject)
- Ngày (Date)

Nội dung trong tập dữ liệu đa số tập trung vào tin tức thời sự tại Hoa Kỳ.



Các mẫu tin tức thu thập được từ **Kaggle**.

Có khoảng 40000 mẫu tin tức, bao gồm **Fake** và **Real** News, đã được gán nhãn sẵn.

Số lượng nhãn khá **cân bằng** nhau, vì vậy tốc độ huấn luyện mô hình tốt

Preprocessing data



Loại bỏ stop-words.



Stemming và Lemmatization (đưa các từ ngữ về chữ nguyên gốc của nó).



Loại bỏ những kí tự đặc biệt.

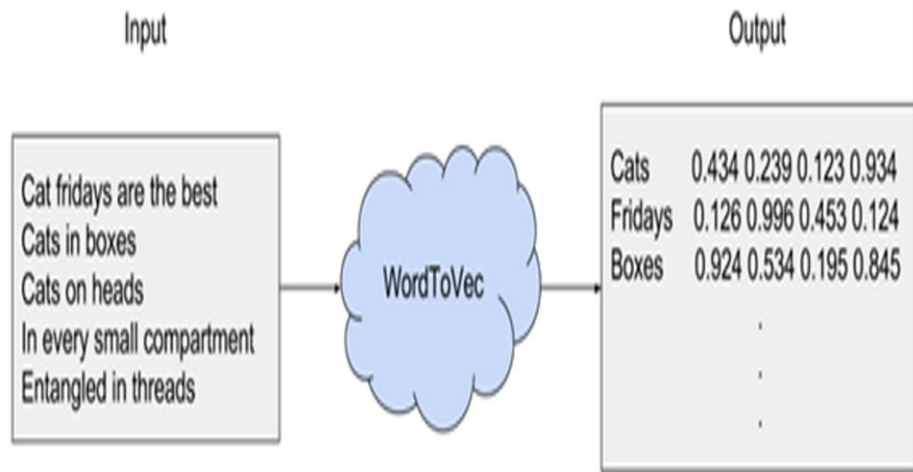


Chuyển đổi toàn bộ văn bản thành chữ thường.

Methods



Word2vec



- Là sự kết hợp giữa hai mô hình Skip-gram và Cbow model.
- Ánh xạ các từ thành các vector số.
- Nhược điểm của phương pháp là không thể nhận dạng được những từ không xuất hiện trong training set.

TF-IDF Vectorizer

$$\text{idf}(t, D) = \log \frac{|D|}{|\{d \in D : t \in d\}|}$$

$$f_{t,d} / \sum_{t' \in d} f_{t',d}$$

$$\text{tfidf}(t, d, D) = \text{tf}(t, d) \times \text{idf}(t, D)$$

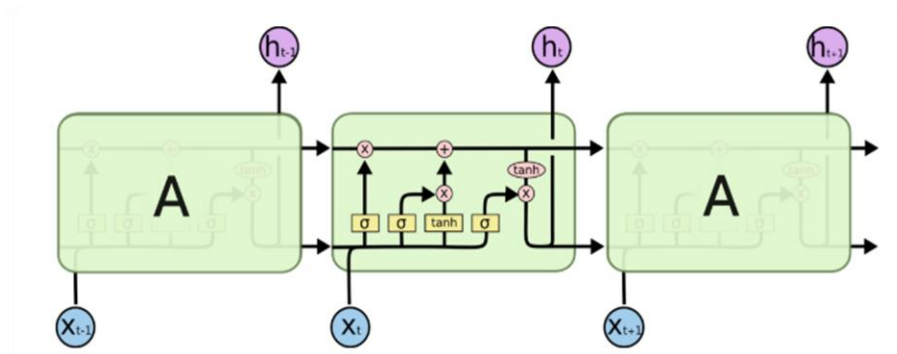
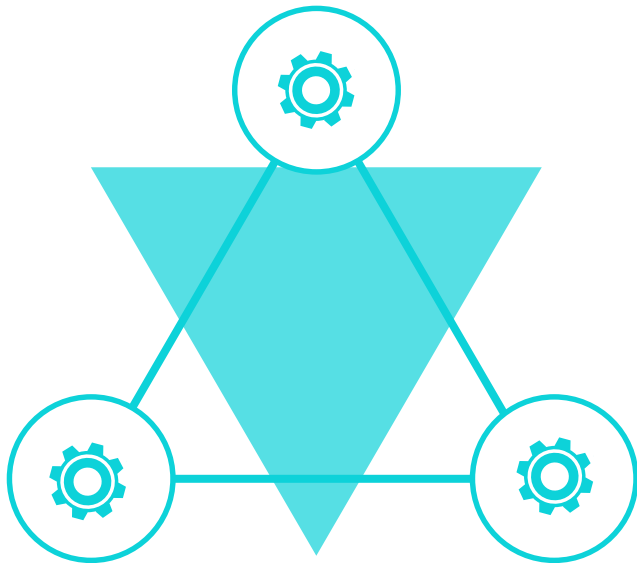
- Đầu tiên tính TF, tức tần số xuất hiện của 1 từ trong văn bản d.
- Tiếp theo, tính giá trị IDF, giúp tìm ra các từ xuất hiện nhiều không mang nhiều ý nghĩa.
- Giá trị TF-IDF cao giúp lọc ra những từ phổ biến đồng thời giữ lại những từ có ý nghĩa cao

LSTM

Là một phiên bản mở rộng của mô hình RNN

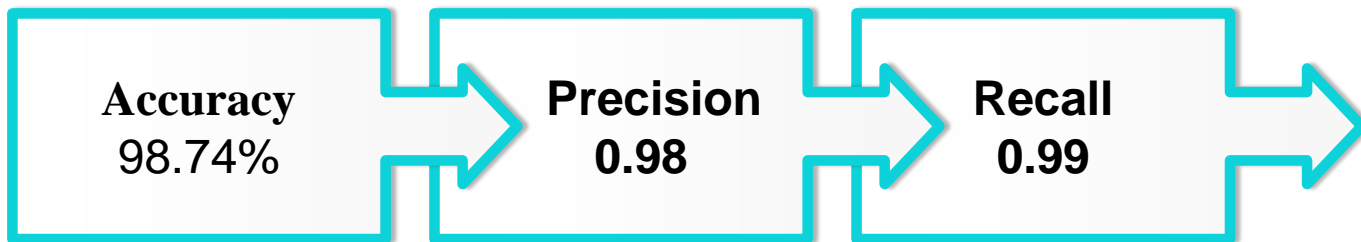
FORGET GATE

Forget gate sẽ lấy đầu vào là h_{t-1} và x_t và dùng hàm sigmoid để tính toán trọng số đầu ra. Nếu đầu ra tiệm cận 0, toàn bộ thông tin đó sẽ bị loại bỏ.



LTMS

Kết quả sau khi chạy



Phân tích lỗi và hướng phát triển

Điểm yếu:

- Chưa phát hiện tin giả khi thử nghiệm thay đổi tên nhân vật đề cập.
- Mô hình còn bỏ qua những từ đáng lẽ có trọng số quan trọng.
- Nhãn dữ liệu có thể bị “thiên vị” (bias) do phụ thuộc vào quan điểm của người gán nhãn.

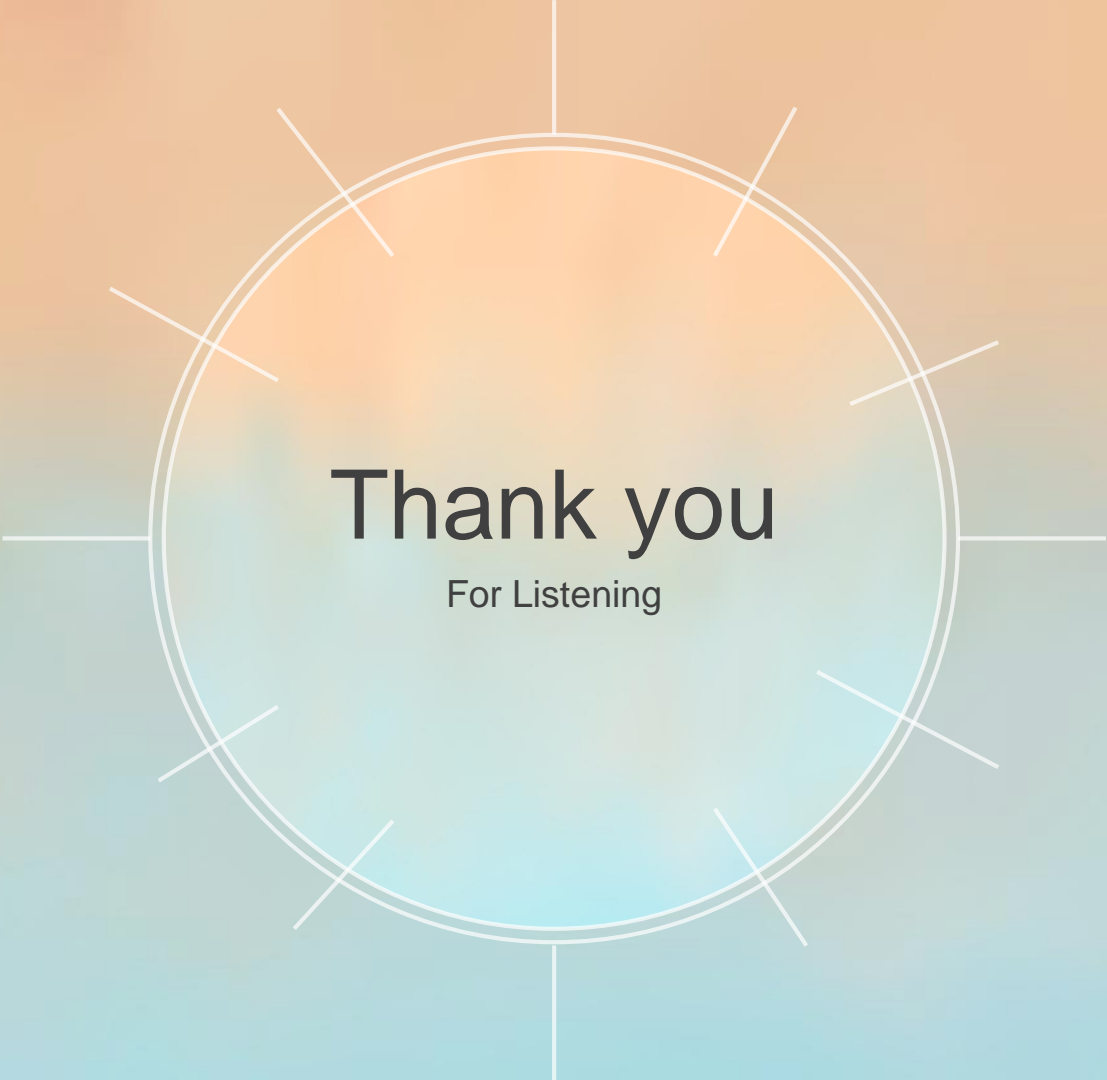
Hướng phát triển:

- Thu thập thêm mẫu tin của nhiều lĩnh vực hơn để làm giàu tập dữ liệu.
- Ưu tiên cao cho tên riêng của nhân vật trong mẫu tin tức.



DEMO





Thank you
For Listening