

PHÁT HIỆN BÌNH LUẬN THÔ TỤC, XÚC PHẠM

Trương Đức Vũ - 18520194 - CS114.K21.KHTN

Link Github:

<https://github.com/ducvuit/CS114.K21.KHTN>

Tóm tắt

Bài toán toán xác định một
bình luận có mang tính xúc
phạm, thô tục hay không

*Mấy loại súc vật ntn chỉ
đáng bỏ đi :))))))*



Thành viên nhóm



Trương Đức Vũ

Bài toán

**Việt Nam đứng
trong top các nước
có mức độ thiếu
văn hóa cao nhất
thế giới**



5 quốc gia có chỉ số DCI thấp nhất và cao nhất

Xếp hạng theo chỉ số DCI quốc gia trên 25 quốc gia tham gia khảo sát



Vương Quốc Anh
52%



Hà Lan
56%



Đức
58%



Malaysia
59%



Hoa Kỳ
60%



Việt Nam
78%



Nga
79%



Colombia
80%



Peru
81%



Nam Phi
83%

Bài toán

Bài viết của **TRẦN MINH NHỰT**

Fan cứng
[Profile Picture] 2
2 tuần Thích Trả lời

[Profile Picture]
Đút cái đĩa vào mồm nó bắt ăn
Gáy lên đi em? Sao không gáy nữa?
2 ngày Thích Trả lời

[Profile Picture]
Đừng trách vì sao tao hay chửi
2 tuần Thích Trả lời 1

Fan cứng
[Profile Picture]
2 tuần Thích Trả lời

[Profile Picture]
2 tuần Thích Trả lời 1

[Profile Picture]
Ừm, tụi nó làm như có mỗi mình làm dc và
vênh vào với cả người lớn
2 tuần Thích Trả lời

Trần Minh Nhựt
Viết bình luận... GIF 😊

📷 📺 👤 🗨️ 🔔 5 🔔 1 ☰

[Profile Picture] 36 phút · 🌟

Dm, thi Anh Văn có cái đề như shit

Thích · Bình luận · Chia sẻ

👍 **Nguyễn Thanh Fx** thích nội dung này.

[Profile Picture] **Nguyễn Thanh Fx** Dm, bên t cái đề cũng như shit đấy! Rót cmnr 😊
Thích · Trả lời · 34 phút

[Profile Picture] **Lê Minh Nhật** đề tào lao ^^
Thích · Trả lời · 33 phút

[Profile Picture] **Trần Minh Nhựt** Kìq nhưng t cũng đéo làm dc đề AV. Xác cmn định là out bà nó
rồi = =
Thích · Trả lời · 31 phút

[Profile Picture] **Trần Minh Nhựt** T cũng ếu hiểu cái đề nói nói con mẹ gì. What a fuckin test=))
Thích · Trả lời · 26 phút · Đã chỉnh sửa

[Profile Picture] **Trần Minh Nhựt** Vua long anh chưa? 🤔
Thích · Trả lời · 3 phút

Viết bình luận... 📷 😊

← Bình luận 🔍

[Profile Picture] **Nguyễn Duy**
Sunny thấy gớm ,gớm bằng mặt
[Profile Picture] mây ko??
8 phút Thích Trả lời

[Profile Picture] **Quang Nhật Nguyễn**
Nguyễn Duy bằng chứ 😊
6 phút Thích Trả lời

[Profile Picture] **Quang Nhật Nguyễn**
Tao mặt [Profile Picture] thiên hạ nên
sunny same same 😊 nhẹ
6 phút Thích Trả lời

[Profile Picture] **Nguyễn Duy**
Thắng mặt [Profile Picture] ,kêu hương
[Profile Picture] mở cmt đi ,reaction
chảnh xá lắm ,sao cụp đuôi
như [Profile Picture] vậy
5 phút Thích Trả lời 1

[Profile Picture] **Quang Nhật Nguyễn**
Ơ tao chảnh xá kem dâu nè
sao đéo rep tao 😊
2 phút Thích Trả lời

[Profile Picture] **Nguyễn Duy**
Chưa tìm thấy mặt [Profile Picture] mây
thôi ,nên quan tâm chuy
Hương trước, ai ngờ Hương
nhục quá khóa cmt rồi 🤡
Vừa xong Thích Trả lời

[Profile Picture] **Quang Nhật Nguyễn**
2273b/11/3 phạm thể hiện
xin mời tới
Vừa xong Thích Trả lời

Viết trả lời... GIF 😊

Bài toán

Và đó chính là lý do em chọn đề tài này để làm báo cáo cho môn học với mong muốn một Việt Nam văn minh hơn, thiện cảm hơn.

Cách giải quyết

Sử dụng machine learning nhằm phát hiện các bình luận thiếu văn hóa

Input: Một câu bình luận bất kì

output : 1 nếu câu bình luận thiếu văn minh, 0 nếu là bình thường

Mô tả dữ liệu

Bộ dữ liệu được crawl từ nhiều bài viết thuộc nhiều chủ đề khác nhau (chủ yếu là những bài phốt gây sốt cộng đồng mạng)

Kích thước dữ liệu: 4764 bình luận

- 2121 bình luận mang tính thô tục, xúc phạm
- Và những bình luận còn lại là bình thường

Kết quả

Accuracy hầu hết đều cho
kết quả trên 80%

Hoạt động tốt

Có vài trường hợp ngoại lệ
vd: Mẹ [1]

```
[INFO] Used TfidfVectorizer ...
[INFO] evaluating SVM...
      precision    recall  f1-score   support

     0       0.83     0.91     0.87     1293
     1       0.88     0.78     0.82     1089

 accuracy
macro avg       0.85     0.84     0.85     2382
weighted avg     0.85     0.85     0.85     2382

[INFO] evaluating Naive Bayes...
      precision    recall  f1-score   support

     0       0.76     0.93     0.84     1293
     1       0.89     0.66     0.76     1089

 accuracy
macro avg       0.83     0.79     0.80     2382
weighted avg     0.82     0.81     0.80     2382

[INFO] evaluating Logistic Regression...
      precision    recall  f1-score   support

     0       0.78     0.95     0.85     1293
     1       0.91     0.67     0.78     1089

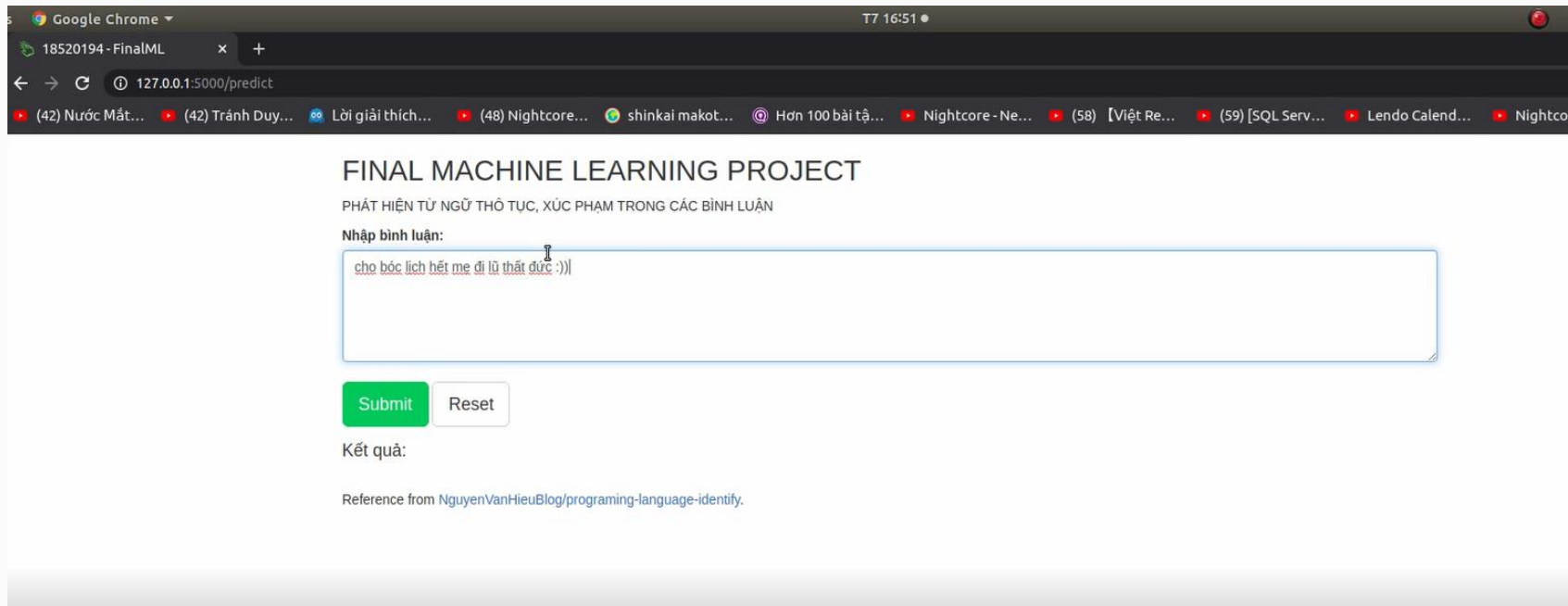
 accuracy
macro avg       0.84     0.81     0.81     2382
weighted avg     0.84     0.82     0.82     2382

[INFO] evaluating Decision Tree Classifier...
      precision    recall  f1-score   support

     0       0.84     0.86     0.85     1293
     1       0.82     0.80     0.81     1089

 accuracy
macro avg       0.83     0.83     0.83     2382
weighted avg     0.83     0.83     0.83     2382
```


Demo



Video chi tiết ở link trong báo cáo

demo

18520194 - FinalML x +

127.0.0.1:5000/predict

(42) Nước Mắt... (42) Tránh Duy... Lời giải thích... (48) Nightcore... shinkai makot... Hơn 100 bài tậ... Nightcore - Ne... (58) [Việt Re... (59) [SQL Serv... Lendo Calend... Nightcore Mo...

FINAL MACHINE LEARNING PROJECT

PHÁT HIỆN TỪ NGỮ THÔ TỤC, XÚC PHẠM TRONG CÁC BÌNH LUẬN

Nhập bình luận:

Submit Reset

Kết quả:

Nhận xét bình luận: **Bình luận mang nghĩa thô tục, xúc phạm!**

Reference from [NguyenVanHieuBlog/programing-language-identify](#).

demo

DETECTING OFENSIVE COMMNETS

Final project Machine learning

Truong Duc Vu - 18520194

Type any comment in Vietnamese here:

xin lỗi cl. Thanh minh xong cho cổ xong đ được lại quay ra xl làm trò cười à?

Submit

Reset

Predict result:

Title topic: Bình luận bình thường!

>.

Contact: Trương Đức Vũ.

Reference from [NguyenVanHieuBlog/programing-language-identify](#).

DETECTING OFENSIVE COMMNETS

Final project Machine learning

Truong Duc Vu - 18520194

Type any comment in Vietnamese here:

Submit

Reset

Predict result:

Title topic: Bình luận mang nghĩa thô tục, xúc phạm!

>.

Contact: Trương Đức Vũ.

Reference from [NguyenVanHieuBlog/programing-language-identify](#).

demo

DETECTING OFENSIVE COMMNETS

Final project Machine learning

Truong Duc Vu - 18520194

Type any comment in Vietnamese here:

đánh ghen thời 4.0

Submit

Reset

Predict result:

Title topic: **Bình luận mang nghĩa thô tục, xúc phạm!**

>.

Contact: [Trương Đức Vũ](#).

Reference from [NguyenVanHieuBlog/programing-language-identify](#).

DETECTING OFENSIVE COMMNETS

Final project Machine learning

Truong Duc Vu - 18520194

Type any comment in Vietnamese here:

I

Submit

Reset

Predict result:

Title topic: **Bình luận bình thường!**

>.

Contact: [Trương Đức Vũ](#).

Reference from [NguyenVanHieuBlog/programing-language-identify](#).

Kết luận và dự định tương lai

Mô hình có thể áp dụng vào để lọc các bình luận thô tục trong các trang báo mạng, mạng xã hội.

Tương lai

- Training thêm 1 model sửa chính tả tiếng việt
- Cài RNN cho mô hình

Tài liệu tham khảo

<https://towardsdatascience.com/building-a-better-profanity-detection-library-with-scikit-learn-3638b2f2c4c2>

<https://kipalog.com/posts/Machine-Learning---NLP--Text-Classification-su-dung-scikit-learn----python>

<https://viblo.asia/p/phan-loai-van-ban-tieng-viet-tu-dong-phan-1-yMnKM3baI7P>

https://github.com/langmaninternet/VietnameseTextNormalizer?fbclid=IwAR3EN_3JNG16ZhBRYw2x4HHUqNTybyFBZ9xpkm4ABVCDUBzRj0eILm5YYqo

<https://github.com/stopwords/vietnamese-stopwords>