Aspect Based Sentiment Analysis

Trần Minh Quang
Bùi Đăng Phúc
Lê Hồng Quân

University of Information Technology

Content

1. Introduction
2. Dataset Building Process
3. methods
5. Experiment

Motivation

- **Understand detailed evaluation**: ABSA helps you classify emotions not only as positive, negative, neutral but also clarify which aspect of the match the emotion applies to. This helps you have a more detailed and complete view of the source of emotions from the fan community.
- **Specific Feedback**: You can collect information about what fans like or dislike about players, teams, or comments from commentators. This can assist team managers, player managers, and content producers in improving the quality and appeal of matches.

Motivation

- **Track trends and changes**: You can track changes in opinions and sentiment over time, helping you better understand trends and fluctuations in the mood of your fan base.
- **Enhance the fan experience**: Understanding fan aspirations and desires for specific aspects of the match helps enhance the fan experience and create relevant and interesting broadcast or event content.

Input-Output

Input

- One comment of a football match video on youtube
- Maximize 200 word
- Can include emoji, teencode, unaccented

Onput

- Sentiment-aspect pairs (ex: Positiveplayer,Negativecommentator)
- Aspect can be one in {player, club, commentator, other}
- Aspect can be one in {Positive, Negative, Neutral}

Why is it necessary to collect data?

- Currently, comment datasets serving the problem of aspect based sentiment analysis (ABSA) on social media in general or comments on soccer videos on YouTube in particular are even rarer and there seems to be no such dataset about ABSA in this domain is public.
- Collecting and building a set of data to solve the problem in this domain is a necessary and urgent issue because the number of comments for videos on social media in general or football videos in particular is very large and the need to understand the emotions and opinions from these comments not only helps the operations of football teams for fans in Vietnam, or for the station to upgrade its services.

Why is it necessary to collect data?

- Building a Vietnamese dataset for the ABSA problem for comments on soccer videos on YouTube also contributes to solving the need for a model that can understand text data on social media, because in other domains, the text above, especially comments for videos, posts, etc., contains a lot of teen-code, spelling errors, emoji, etc. This leads to previous NLP models being unable to handle it.

Dataset Building Process

Where does the source of data come?

- Comments are taken from videos about Premier League, C1, C2 football matches,.. from official channels such as FPT, VTV,..
- Comments are kept intact without deleting emojis

Dataset Building Process

Characteristics of the dataset

- Contains many emojis: Psg năm nay như 💩
- Contains many abbreviations and teencodes: Đã rât lâu r mơi thay 1 Mu lỳ lợm trỡ lại .. Mu đag vao fom .. hi vog năm s se lm nên ch
- Spelling mistake: Có ai **cong** muốn mua camavinga nữa ko?
- The word is both unaccented and accented: Couto co noi ma nghien cuu bon live đá ntn va cung san sang len đá pen luon that la qua hay ko he giả chân chut nao het couto oi
- Talking in reverse: Mân Đàn

How do we label data?
Data were annotated by three people. We divided the dataset into two subsets. First, two annotators were asked to identify aspects and sentiments in two subsets (each annotator for one subset). If there is any concern check the concern column to 1. Then, the third annotator checked labeled data. If annotators disagreed on an assignment, three people were asked to examine and make the final decision for disagreed assignments and concerned assignments.

How do we label data?

- Definition of aspects
  - **"player"**: Comments are tagged with the "player" when they mention specific players
  - **"club"**: Comments are tagged with the "club" when they talk about the team or general statements about "defense", "attack" or the team's playing style.
  - **"commentator"**: Comments are tagged with the "commentator" when they refers to things related to match commentators.
  - **"other"**: Comments are tagged with the "other" when they do not mention to any aspects above or they are just general comments.

How do we label data?

- Definition of sentiments
  - **"position"**: Comments are tagged with the "positions" when they expresses satisfaction, praise or encouragement.
  - **"negative"**: Comments are tagged with the "negative" when they expresses dissatisfaction, criticism or complaints.
  - **"neutral"**: Comments are tagged with the "neutral" when they are unclear or incomplete in meaning. Or don't show any feelings

How do we label data?

- Example
  - "Thay thằng Onana ra giùm tao!!!" → *Negative#player*
  - "Nhà vua Real" → *Positive#club*
  - "BLV nhạt z" → *Negative#commentator*
  - "(@^^@)" → *Neutral#other*

Label statistics

	Club	Player	Commentator	Other
Positive	418	316	7	157
Negative	324	270	84	206
Neutral	241	145	14	344

Bảng 1: Training Set

	Club	Player	Commentator	Other
Positive	103	81	2	48
Negative	91	57	24	60
Neutral	55	31	10	69

Bảng 2: Test Set

Dataset Building Process

Difficulties when labeling

- When the real meaning is different from the written meaning.
  Example: Thủ môn chơi chân đẳng cấp Onana.

- When the emoji is different from the written meaning.
  Example: Psg năm nay như 💩

- A comment may have multiple labels, so that it's sometimes complicated to identify which aspects person talk about.

Evaluation

$$\text{Accuracy} = \frac{num\_correct\_predicts}{num\_samples}$$

where: a correct prediction which correct all aspects and sentiments

Evaluation

Let A the set of predicted aspects-sentiments(aspect-sentiment pairs), and B the set of annotated aspects-sentiments, precision, recall, and the F1 score are computed as follows:

$$\text{Precision} = \frac{|A \cap B|}{|A|}$$

$$\text{Recall} = \frac{|A \cap B|}{|B|}$$

$$\text{F1} = \frac{2xPrecisionxRecall}{Precision + Recall}$$

Vietnamese Corrector

- PhoBERT is trained on the ViWiki and ViNews datasets.
- So, it cannot handle Spelling mistake, unaccented cases.
- Vietnamese Corrector (bmd1905/vietnamese-correction) is a fine-tuned version of vinai/bartpho-syllable. The original dataset is avaiable at @duyvuleo/VNTC, customized for error correction task.

PhoBert

- PhoBERT is a Vietnamese natural language model based on the architecture of the BERT model (Bidirectional Encoder Representations from Transformers). Developed by Vingroup, this model aims to understand and represent the Vietnamese language effectively.
- PhoBERT is trained on the ViWiki and ViNews datasets.

PhoBERT: https://aclanthology.org/2020.findings-emnlp.92/

ViSoBERT

- ViSoBERT is an advanced language model for tasks on Vietnamese social networks. This is the first monolingual model specifically built for Vietnamese social network text.
- ViSoBERT is trained on the Vietnamese social media dataset

ViSoBERT: https://aclanthology.org/2023.emnlp-main.315/

Model

Comments	ViSoBERT	PhoBERT
concặc cáilồn gìđây🙂🙂🙂 *English*: Wut is dis fuckingd1ck 🙂🙂🙂	<s>, "conc", "ặc", "cái", "l", "ồn", "gì", "đây", "🙂🙂🙂", </s>	<s>, "c o n @ @", "c @ @", "ặ c", "c á @ @", "i l @ @", "ồ n", "g @ @","ì @ @", "đ â y", <unk>, <unk>, <unk>, </s>
e cảmơn anh😎😎 *English*: Thankyou 😎😎	<s>, "e", "cảm", "ơn", "anh", "😎", "😎", </s>	<s>, "e", "c ả @ @", "m @ @", "ơ n", "a n h", <unk>, <unk>, </s>
d4y l4 vj du cko mot cau teencode *English*: Th1s 1s 4 teencode s3nt3nc3	<s>, "d", "4", "y", "l", "4", "vj", "du", "cko", "mot", "cau", "teen", "code", </s>	<s>, "d @ @", "4 @ @", "y", "l @ @", "4", "v @ @", "j", "d u", "c k @ @", "o", "m o @ @", "t", "c a u"; "t e @ @", "e n @ @", "c o d e", </s>

Table 1: Actual social comments and their tokenizations with the tokenizers of the two pre-trained language models, ViSoBERT and PhoBERT.

Hình 1: Differences from tokenizations of ViSoBERT and PhoBERT

Model

Model	Emotion Regconition			Hate Speech Detection			Sentiment Analysis			Spam Reviews Detection			Hate Speech Spans Detection		
	Acc	WF1	MF1	Acc	WF1	MF1	Acc	WF1	MF1	Acc	WF1	MF1	Acc	WF1	MF1
Converting emojis to text															
$\text{PhoBERT}_{\text{Large}}$	66.08	66.15	63.35	87.43	87.22	65.32	76.73	76.48	76.48	90.35	90.11	77.02	92.16	91.98	86.72
Δ	↑ 1.37	↑ 1.49	↑ 0.80	↑ 0.11	↑ 0.24	↑ 0.18	↓ 0.21	↓ 0.12	↓ 0.12	↑ 0.23	↑ 0.08	↑ 0.14	↑ 0.72	↑ 0.52	↑ 0.16
$\text{TwHIN-BERT}_{\text{Large}}$	64.82	64.42	61.33	86.03	85.52	63.52	75.42	75.95	75.95	90.55	90.47	77.32	92.21	92.01	86.84
Δ	↑ 0.61	↑ 0.13	↑ 0.21	↓ 1.20	↓ 1.26	↓ 1.71	↓ 1.50	↓ 0.88	↓ 0.88	↑ 0.08	↑ 0.05	↑ 0.04	↑ 0.76	↑ 0.54	↑ 0.19
ViSoBERT [♣]	67.53	67.93	65.42	87.82	87.88	67.25	76.95	76.85	76.85	90.22	90.18	78.25	92.42	92.11	87.01
Δ	↓ 0.57	↓ 0.44	↓ 0.46	↓ 0.69	↓ 0.41	↓ 1.49	↓ 0.88	↓ 0.90	↓ 0.90	↓ 0.77	↓ 0.74	↓ 0.81	↑ 0.80	↑ 0.54	↑ 0.21
Removing emojis															
$\text{PhoBERT}_{\text{Large}}$	65.21	65.14	62.81	87.25	86.72	64.85	76.72	76.48	76.48	90.21	90.09	77.02	91.53	91.51	86.62
Δ	↑ 0.50	↑ 0.48	↑ 0.26	↓ 0.07	↓ 0.26	↓ 0.29	↑ 0.20	↑ 0.12	↑ 0.12	↑ 0.09	↑ 0.06	↑ 0.10	↑ 0.09	↑ 0.05	↑ 0.09
$\text{TwHIN-BERT}_{\text{Large}}$	62.03	62.14	59.25	86.98	86.32	64.22	75.00	75.11	75.11	89.83	89.75	76.85	91.32	91.33	86.42
Δ	↓ 2.18	↓ 1.15	↓ 1.87	↓ 0.25	↓ 0.46	↓ 1.01	↓ 1.92	↓ 1.72	↓ 1.72	↓ 0.64	↓ 0.67	↓ 0.43	↓ 0.13	↓ 0.14	↓ 0.23
ViSoBERT [♦]	66.52	67.02	64.55	87.32	87.12	66.98	76.25	75.98	75.98	89.72	89.69	77.95	91.58	91.53	86.72
Δ	↓ 1.58	↓ 1.35	↓ 1.33	↓ 1.19	↓ 1.19	↓ 1.79	↓ 1.58	↓ 1.77	↓ 1.77	↓ 1.27	↓ 1.23	↓ 1.11	↓ 0.04	↓ 0.04	↓ 0.08
ViSoBERT [♠]	**68.10**	**68.37**	**65.88**	**88.51**	**88.31**	**68.77**	**77.83**	**77.75**	**77.75**	**90.99**	**90.92**	**79.06**	**91.62**	**91.57**	**86.80**

Table 5: Performances of pre-trained models on downstream Vietnamese social media tasks by applying two emojis pre-processing techniques. [♣], [♦], and [♠] denoted our pre-trained language model ViSoBERT converting emoji to text, removing emojis and without applying any pre-processing techniques, respectively. Δ denoted the increase (↑) and the decrease (↓) in performances of the PLMs compared to their competitors without applying any pre-processing techniques.

Hình 2: Ability to handle raw text of ViSoBERT

Consideration

- We prefer ViSoBERT to PhoBERT:
  - PhoBERT is trained on the data sets ViWiki and ViNews, so it does not handle typical issues of comments on Social media such as emoji, teencode, abbreviations, etc.
  - Meanwhile, ViSoBERT is built and trained on a data set taken from social media, has a custom tokenizer to suit the processing of emoji, teencode,...

Experiment

We experiment with three varieties using PhoBERT (default, added emoji to tokenizer, and VietnameseCorrector+PhoBERT) and ViSoBERT
Finetune to suit the ABSA problem in youtube comments of football matches by adding a linear layer with output of 12 (4 aspect x 3 sentiment)

Experiment

- Environment: Google Colab with T4 GPU
- Batch size: 32
- Optimizer: AdamW with epsilon 1e-8 and learning rate 2e-5

Our Result

Model	Accuracy	F1
PhoBERT without emoji	0.468	0.52
PhoBERT with emoji	0.456	0.5
VietnameseCorrector +PhoBERT	0.478	0.54
ViSoBERT	0.548	0.59

THANKS FOR YOUR ATTENTION !