ASPECT SENTIMENT ANALYSIS USING BERT AND LSTM

(Phân tích khía cạnh tình cảm sử dụng mô hình Bert và LSTM – Một bản báo cáo)

Tóm tắt:

Thực hiện phân tích khía cạnh tình cảm sản phẩm. Bộ dữ liệu là các câu comment đánh giá sản phẩm của khách hành trên điện thoại. Bộ dữ liệu này là tiếng Việt.

Các phương pháp trước đây:

* Sử dụng Bi-LSTM: trên cùng bộ data
* Sử dụng Bert: bộ data tiếng Anh

Phương pháp mới đề xuất: kết hợp cả hai phương pháp sử dụng Bert và LSTM để tăng thêm độ chính xác.

Bộ dữ liệu:

* Có 10 aspect: 'SCREEN', 'CAMERA', 'FEATURES', 'BATTERY', 'PERFORMANCE', 'STORAGE', 'DESIGN', 'PRICE', 'GENERAL', 'SER&ACC'
* Mỗi aspect có 4 sentiment chính để thực hiện predict là: Positive, Negative, Neutral và None (tức là trong câu không chứa aspect đó).

Ví dụ:

Input = "Điện thoải ổn. Facelock cực nhanh, vân tay ôk , màn hình lớn, pin trâu ( liên quân , Zalo, YouTube ) một ngày mất khoảng 45 % ) tuy chỉ chip 439 nhưng rất mượt. Đa nhiệm khá ổn"

Kết quả dự đoán sẽ là:

Aspect Sentiment Prediction:

SCREEN: positive

CAMERA: none

FEATURES: positive

BATTERY: positive

PERFORMANCE: none

STORAGE: none

DESIGN: none

PRICE: none

GENERAL: positive

SER&ACC: none

Thống kê về bộ dữ liệu:

A table with numbers and a number

Description automatically generated

Huấn luyện mô hình:

* 10 epochs
* Thời gian: 35.5h
* Epoch 1/10, Train loss: 0.3192513349793041
* Epoch 2/10, Train loss: 0.21291347181527767
* Epoch 3/10, Train loss: 0.1819368628383417
* Epoch 4/10, Train loss: 0.15778480791458113
* Epoch 5/10, Train loss: 0.1388647394136237
* Epoch 6/10, Train loss: 0.12444866526665384
* Epoch 7/10, Train loss: 0.11172562608881653
* Epoch 8/10, Train loss: 0.10074596001443432
* Epoch 9/10, Train loss: 0.09115716066112019
* Epoch 10/10, Train loss: 0.08395827585330244

Đánh giá mô hình:

* Đánh giá hiệu suất tổng thể mô hình
* Đánh giá hiệu suất nhận diện khía cạnh tổng thể của mô hình
* Đánh giá hiệu suất nhận diện trên từng khía cạnh của mô hình
* Đánh giá hiệu suất nhận diện tình cảm tổng thể của mô hình
* Đánh giá hiệu suất nhận diện tình cảm trên từng khía cạnh của mô hình

1. Đánh giá hiệu suất tổng thể mô hình

Accuracy = 0.9140737410071943

Precision = 1.0

Recall = 0.7383983572895277

f1 = 0.8495157098984172

1. Đánh giá hiệu suất nhận diện khía cạnh tổng thể của mô hình

Accuracy = 0.9140737410071943

Precision = 1.0

Recall = 0.7383983572895277

F1-score = 0.8495157098984172

1. Đánh giá hiệu suất nhận diện trên từng khía cạnh của mô hình
2. Khía cạnh SCREEN

Accuracy = 0.9671762589928058

Precision = 1.0

Recall = 0.7582781456953642

f1 = 0.8625235404896422

1. Khía cạnh FEATURES

Accuracy = 0.9608812949640287

Precision = 1.0

Recall = 0.8927250308261405

f1 = 0.9433224755700326

1. Khía cạnh PERFORMANCE

(các dự đoán đều là none – tức là mô hình không thể nhận diện được khía cạnh này trên bộ dữ liệu kiểm thử)

Accuracy = 0.47302158273381295

Precision = 0.0

Recall = 0.0

f1 = 0.0

1. Khía cạnh GENERAL

Accuracy = 0.91681654676259

Precision = 1.0

Recall = 0.8818646232439336

f1 = 0.9372242958941296

1. Khía cạnh BATTERY

Accuracy = 0.9941546762589928

Precision = 1.0

Recall = 0.9878163074039362

f1 = 0.9938708156529938

1. Khía cạnh STORAGE

(các dự đoán đều là none – tức là mô hình không thể nhận diện được khía cạnh này trên bộ dữ liệu kiểm thử)

Accuracy = 0.9878597122302158

Precision = 0.0

Recall = 0.0

f1 = 0.0

1. Khía cạnh DESIGN

Accuracy = 0.9545863309352518

Precision = 1.0

Recall = 0.7651162790697674

f1 = 0.8669301712779973

1. Khía cạnh PRICE

Accuracy = 0.9676258992805755

Precision = 1.0

Recall = 0.8862559241706162

f1 = 0.9396984924623115

1. Khía cạnh SER&ACC

Accuracy = 0.9312050359712231

Precision = 1.0

Recall = 0.7695783132530121

f1 = 0.8697872340425532

1. Khía cạnh CAMERA

Accuracy = 0.987410071942446

Precision = 1.0

Recall = 0.9557661927330173

f1 = 0.9773828756058158

1. Đánh giá hiệu suất nhận diện tình cảm tổng thể của mô hình

Accuracy = 0.6537987679671458

Overall Precision, Recall, F1-Score:

Micro Precision: 0.9550

Micro Recall: 0.9550

Micro F1-Score: 0.9550

Macro Precision: 0.6855

Macro Recall: 0.6003

Macro F1-Score: 0.6332

Weighted Precision: 0.9463

Weighted Recall: 0.9550

Weighted F1-Score: 0.9481

1. Đánh giá hiệu suất nhận diện tình cảm trên từng khía cạnh của mô hình
2. Tình cảm trên khía cạnh SCREEN

precision recall f1-score support

negative 0.92 0.58 0.71 133

neutral 0.00 0.00 0.00 17

none 0.96 1.00 0.98 1922

positive 0.86 0.82 0.84 152

accuracy 0.96 2224

macro avg 0.69 0.60 0.63 2224

weighted avg 0.95 0.96 0.95 2224

1. Tình cảm trên khía cạnh FEATURES

precision recall f1-score support

negative 0.91 0.85 0.88 526

neutral 0.00 0.00 0.00 52

none 0.94 1.00 0.97 1413

positive 0.82 0.80 0.81 233

accuracy 0.92 2224

macro avg 0.67 0.66 0.66 2224

weighted avg 0.90 0.92 0.91 2224

1. Tình cảm trên khía cạnh PERFORMANCE

precision recall f1-score support

negative 0.00 0.00 0.00 454

neutral 0.00 0.00 0.00 116

none 0.47 1.00 0.64 1052

positive 0.00 0.00 0.00 602

accuracy 0.47 2224

macro avg 0.12 0.25 0.16 2224

weighted avg 0.22 0.47 0.30 2224

1. Tình cảm trên khía cạnh GENERAL

precision recall f1-score support

negative 0.94 0.72 0.82 361

neutral 0.00 0.00 0.00 83

none 0.78 1.00 0.88 658

positive 0.95 0.93 0.94 1122

accuracy 0.88 2224

macro avg 0.67 0.66 0.66 2224

weighted avg 0.86 0.88 0.87 2224

1. Tình cảm trên khía cạnh CAMERA

precision recall f1-score support

negative 0.82 0.82 0.82 197

neutral 0.50 0.01 0.03 71

none 0.98 1.00 0.99 1591

positive 0.85 0.95 0.90 365

accuracy 0.94 2224

macro avg 0.79 0.69 0.68 2224

weighted avg 0.93 0.94 0.93 2224

1. Tình cảm trên khía cạnh BATTERY

precision recall f1-score support

negative 0.85 0.89 0.87 397

neutral 0.00 0.00 0.00 92

none 0.99 1.00 0.99 1157

positive 0.85 0.95 0.90 578

accuracy 0.92 2224

macro avg 0.67 0.71 0.69 2224

weighted avg 0.89 0.92 0.91 2224

1. Tình cảm trên khía cạnh STORAGE

precision recall f1-score support

negative 0.00 0.00 0.00 6

neutral 0.00 0.00 0.00 3

none 0.99 1.00 0.99 2197

positive 0.00 0.00 0.00 18

accuracy 0.99 2224

macro avg 0.25 0.25 0.25 2224

weighted avg 0.98 0.99 0.98 2224

1. Tình cảm trên khía cạnh DESIGN

precision recall f1-score support

negative 0.89 0.24 0.38 105

neutral 0.00 0.00 0.00 28

none 0.95 1.00 0.97 1794

positive 0.87 0.89 0.88 297

accuracy 0.94 2224

macro avg 0.68 0.53 0.56 2224

weighted avg 0.92 0.94 0.92 2224

1. Tình cảm trên khía cạnh PRICE

precision recall f1-score support

negative 0.75 0.50 0.60 90

neutral 0.88 0.75 0.81 361

none 0.96 1.00 0.98 1591

positive 0.75 0.81 0.78 182

accuracy 0.92 2224

macro avg 0.84 0.76 0.79 2224

weighted avg 0.92 0.92 0.92 2224

1. Tình cảm trên khía cạnh SER&ACC

precision recall f1-score support

negative 0.89 0.48 0.63 201

neutral 0.00 0.00 0.00 27

none 0.91 1.00 0.95 1560

positive 0.94 0.87 0.90 436

accuracy 0.92 2224

macro avg 0.69 0.59 0.62 2224

weighted avg 0.90 0.92 0.90 2224

Tổng kết sơ bộ:

Các phương pháp mới này đã có hiệu suất nhận diện rất tốt và cải thiện kết quả đáng kể, tuy nhiên trên một vài khía cạnh mô hình còn tỏ ra chưa đạt kết quả mong đợi.