Emotion Recognition in Arabic Tweets

Mariam ElOraby, Rawan Reda, Salma El-Shafey

Outline

1. Introduction
2. Limitations
3. Methodology
4. Results
5. Other Approaches
6. Conclusion

Introduction

- **Emotion recognition** refers to the automatic detection and classification of emotions expressed in the text.

- We are focusing on the recognition of seven emotions: joy, sadness, anger, fear, love, sympathy, and surprise in **Arabic** text.

Dataset

emotone_ar on huggingface 🤗

8 categories: none, joy, sadness, anger, sympathy, fear, love and surprise.

Emotion Label	Count
none	1550
anger	1444
joy	1281
sadness	1256
love	1220
sympathy	1062
surprise	1045
fear	1207

Limitations

- Arabic is a **complicated language**, with many dialects
- Dataset also consists mostly of tweets in **Egyptian Arabic**.
- Dataset mostly consists of tweets about the **Olympics**
- Dataset we had found is considered very **small with 10K** tweets only.
- The dataset covered a **short time span** (July 31st, 2016 to August 20th, 2016)

Methodology

Data Preprocessing	Fine-tuning BERT	Model Evaluation	Select the the best performing model
Apply data cleaning and preprocessing	Tokenize the cleaned tweets using BERT’s tokenizer and fine-tune on BERT-based models.	Evaluate the performance of the classification models (with and without emojis)	Select the best performing model

Methodology

Data Preprocessing → **Fine-tuning BERT** → Model Evaluation → Select the the best performing model

We fine-tune seven BERT-based models on the aforementioned dataset and compare their results.

AraBERTv2-base
AraBERTv0.2-Twitter-base

CaMeLBERT-MSA (2 sizes)
CaMeLBERT-DA

MARBERTv2

mBERT

Methodology

Data Preprocessing → **Fine-tuning BERT** → Model Evaluation → Select the the best performing model

We fine-tune seven BERT-based models on the aforementioned dataset and compare their results.

AraBERTv2-base	CaMeLBERT-MSA (2 sizes)	MARBERTv2	mBERT
AraBERTv0.2-Twitter-base	CaMeLBERT-DA		

- ❏ Trained each model on data **with emojis** and **without emojis**
- ❏ Used **HuggingFace transformers** and **Google GPUs**
- ❏ Used **AdamW** optimizer (**default learning rate of 5e-5**, and **epsilon value of 1e-8**)
- ❏ Used a **batch size** of **32**
- ❏ Used **varied number of epochs (2/3)** to avoid over-fitting
- ❏ Used maximum **input length of 65**.

Methodology

Data Preprocessing → Fine-tuning BERT → **Model Evaluation** → Select the the best performing model

- Models were evaluated based on **accuracy**, precision and recall.

Model	Emojis Replaced	Emojis Removed
AraBERTv2-base	0.74 (3 epochs)	0.74 (3 epochs)
AraBERTv0.2-Twitter-base	0.78 (2 epochs)	0.78 (2 epochs)
MARBERTv2	**0.81** (2 epochs)	**0.81** (2 epochs)
CaMeLBERT-MSA	0.74 (3 epochs)	0.75 (3 epochs)
CaMeLBERT-MSA-16th	0.75 (3 epochs)	0.74 (3 epochs)
CaMeLBERT-DA	0.78 (2 epochs)	0.78 (2 epochs)
mBERT	0.70 (3 epochs)	0.70 (3 epochs)

Methodology

Data Preprocessing → Fine-tuning BERT → **Model Evaluation** → Select the the best performing model

- ❑ Models were evaluated based on accuracy, **precision and recall.**

	precision	recall	f1-score	support
none	0.78	0.83	0.80	159
anger	0.81	0.84	0.83	148
joy	0.65	0.67	0.66	122
sadness	0.75	0.68	0.71	105
love	0.84	0.76	0.80	122
sympathy	0.88	0.96	0.92	111
surprise	0.78	0.73	0.76	116
fear	0.95	0.94	0.94	124
accuracy			0.81	1007
macro avg	0.81	0.80	0.80	1007
weighted avg	0.81	0.81	0.80	1007

based on **MARBERT** results

Methodology

Data Preprocessing → Fine-tuning BERT → Model Evaluation → **Select the the best performing model**

Model	Emojis Replaced	Emojis Removed
AraBERTv2-base	0.74 (3 epochs)	0.74 (3 epochs)
AraBERTv0.2-Twitter-base	0.78 (2 epochs)	0.78 (2 epochs)
MARBERTv2	**0.81** (2 epochs)	**0.81** (2 epochs)
CaMeLBERT-MSA	0.74 (3 epochs)	0.75 (3 epochs)
CaMeLBERT-MSA-16th	0.75 (3 epochs)	0.74 (3 epochs)
CaMeLBERT-DA	0.78 (2 epochs)	0.78 (2 epochs)
mBERT	0.70 (3 epochs)	0.70 (3 epochs)

Results

As shown in the table below:

- mBERT gives the lowest accuracy.
- MARBERT gives the best results.
- AraBERTv0.2-Twitter-base gives better results than AraBERTv2-base.

Model	Emojis Replaced	Emojis Removed
AraBERTv2-base	0.74 (3 epochs)	0.74 (3 epochs)
AraBERTv0.2-Twitter-base	0.78 (2 epochs)	0.78 (2 epochs)
MARBERTv2	**0.81** (2 epochs)	**0.81** (2 epochs)
CaMeLBERT-MSA	0.74 (3 epochs)	0.75 (3 epochs)
CaMeLBERT-MSA-16th	0.75 (3 epochs)	0.74 (3 epochs)
CaMeLBERT-DA	0.78 (2 epochs)	0.78 (2 epochs)
mBERT	0.70 (3 epochs)	0.70 (3 epochs)

Results

As shown in the table below:

- CaMeLBERT-DA gives better results than both CaMeLBERT-MSA models.
- Existence of emojis doesn't make much difference.

Model	Emojis Replaced	Emojis Removed
AraBERTv2-base	0.74 (3 epochs)	0.74 (3 epochs)
AraBERTv0.2-Twitter-base	0.78 (2 epochs)	0.78 (2 epochs)
MARBERTv2	**0.81** (2 epochs)	**0.81** (2 epochs)
CaMeLBERT-MSA	0.74 (3 epochs)	0.75 (3 epochs)
CaMeLBERT-MSA-16th	0.75 (3 epochs)	0.74 (3 epochs)
CaMeLBERT-DA	0.78 (2 epochs)	0.78 (2 epochs)
mBERT	0.70 (3 epochs)	0.70 (3 epochs)

Examples

Tweet Text	Prediction	What We Expected
أنا مبسوط	joy	joy
احنا هانتخرج	sadness	none
ادا احنا هنتخرج!	surprise	surprise
احنا هنتخرج!	surprise	joy or surprise
مش عارف احل الاساينمنت	sadness	sadness or anger
الأكل حلو أوي	joy	joy
الأكل مش حلو أوي	sadness	sadness
الامتحان بكرا و انا مش عارف حاجة	fear	fear
الامتحان بكرا	fear	fear
تراني تاثرت	sympathy	sympathy
يعيني سهران طول اليل	sadness	sympathy
♡♡♡	love	love
!مش معقول اللي بيحصل ده	surprise	anger
وحشتيني جدا	joy	love
.بموت فيها بجد	joy	love or joy
فاكر لما كان عندك إهتمامات وشخصية قبل ما الرأسمالية تعرفك؟	surprise	sadness or surprise

LIVE INFERENCE :)

Other Approaches

- Rule-based Approach:

  SentiWordNet: sentiment scores for each sense of each word in WordNet.

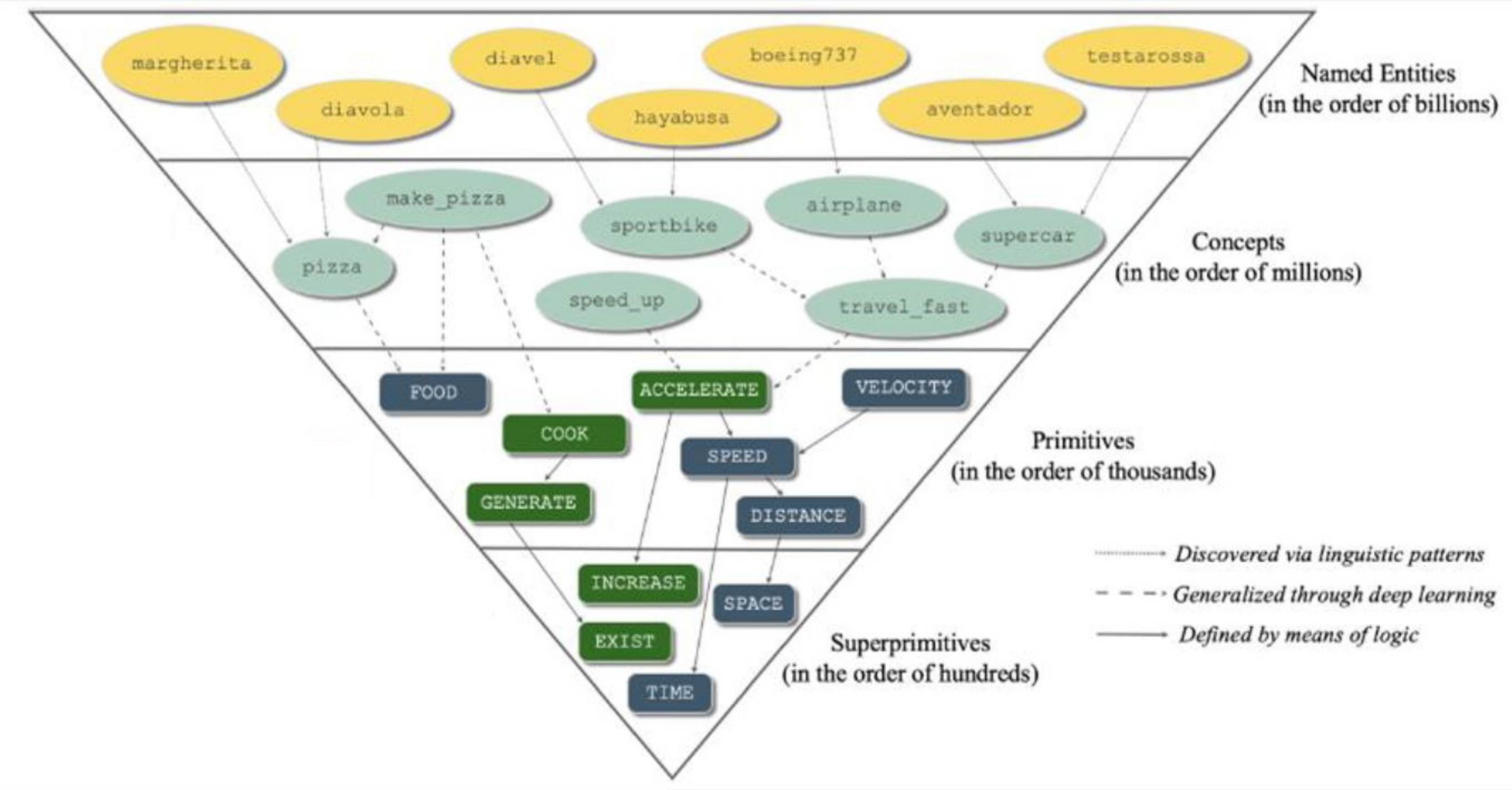
- Hybrid Approach: Machine Learning + Rule-based

Other Approaches - SenticNet 7

- Hybrid (Neurosymbolic) approach
- Based on common-sense knowledge

Other Approaches - SenticNet 7

-



Other Approaches - SenticNet 7

Lexicon	Year	CR	MR	Amazon	IMDb	Sanders	SST	STS	SE13	SE15	SE16
General Inquirer*	1966	56.56%	53.76%	59.63%	59.43%	46.81%	54.39%	54.59%	47.82%	51.12%	40.88%
LIWC	1993	52.42%	41.84%	57.33%	63.53%	52.45%	44.88%	67.68%	44.85%	43.96%	39.75%
ANEW	1999	51.55%	51.13%	50.66%	51.94%	50.08%	51.33%	47.35%	50.44%	48.26%	49.93%
WordNet-Affect*	2004	04.61%	05.05%	18.87%	28.99%	17.81%	04.82%	24.23%	15.92%	16.35%	10.54%
Opinion Lexicon	2004	72.98%	62.90%	65.76%	70.91%	67.89%	66.50%	74.09%	72.65%	64.01%	73.42%
Opinion Finder*	2005	62.05%	59.98%	59.48%	58.75%	51.22%	61.86%	60.72%	50.28%	53.57%	44.21%
Micro WNOp*	2007	20.39%	18.73%	44.48%	49.17%	22.95%	17.64%	28.13%	24.89%	26.58%	18.41%
Sentiment140	2009	65.50%	61.52%	66.64%	68.64%	70.92%	64.67%	76.88%	66.78%	60.94%	62.55%
SentiStrength*	2010	45.69%	41.72%	59.09%	60.18%	47.87%	41.57%	58.49%	42.32%	45.60%	35.46%
SentiWordNet	2010	64.60%	59.07%	62.36%	64.13%	61.68%	61.55%	63.23%	50.03%	60.53%	46.80%
AFINN	2011	70.59%	63.78%	66.63%	71.27%	**71.90%**	66.85%	78.27%	59.04%	**67.08%**	53.82%
SO-CAL	2011	65.58%	**64.58%**	**75.86%**	**78.67%**	52.78%	**67.33%**	63.51%	41.15%	37.63%	41.02%
EmoLex	2013	61.10%	56.03%	52.73%	51.94%	56.86%	59.06%	60.17%	66.21%	64.21%	66.40%
NOVAD*	2013	64.88%	56.91%	57.06%	56.81%	51.06%	58.88%	61.55%	61.10%	57.87%	58.16%
NRC HS Lexicon	2014	65.26%	58.53%	59.39%	63.49%	59.31%	61.58%	64.07%	70.45%	60.53%	72.72%
VADER	2014	**75.18%**	61.37%	67.03%	69.24%	71.81%	65.94%	**78.83%**	**74.88%**	**69.53%**	**74.05%**
MPQA	2015	68.20%	64.03%	62.43%	64.33%	61.03%	66.66%	71.03%	56.35%	58.28%	54.70%
SentiWords*	2016	62.71%	58.65%	58.11%	57.29%	53.59%	60.57%	60.44%	58.82%	57.46%	54.38%
HSSWE*	2017	71.33%	60.61%	67.08%	65.27%	**73.94%**	63.15%	78.27%	68.67%	64.83%	66.62%
Lingmotif-lex	2018	**76.08%**	**66.52%**	**73.34%**	**74.08%**	70.59%	**70.58%**	**79.11%**	**74.70%**	64.62%	**74.91%**
SenticNet 7	2022	83.60%	77.04%	81.53%	82.91%	80.54%	78.71%	90.08%	83.69%	81.67%	84.39%

Thank you! 😌