

A Rule-Based Approach to Implicit Emotion Detection in Text

CMCC

Vinicius Narciso da Silva - 11030214

A dark blue diagonal gradient bar that starts from the bottom left corner and extends towards the top right corner, covering the lower half of the slide.

Análise de Sentimentos

- Desafio em PLN;
- Recente;
- Subjetivo.

Abordagem por Matching Léxico

- Geralmente não tem um bom desempenho;
- Leva em conta um dicionário (ex: SentiWordNet) e calcula polaridade das palavras que compõe a sentença;
- Alto trabalho manual/computacional.

Naive Bayes (NB) supervisionado

- Utiliza-se de aprendizado de máquina;
- Altamente difundido;
- Utiliza um dicionário dinâmico*;
- Utiliza resultados previamente calculados para avaliar os próximos resultados;
- Atualmente alcança os melhores resultados em termos de F-measure (das técnicas mais difundidas).

Método baseado em regras

- Utiliza-se de regras previamente definidas por alguém;
- É semelhante ao matching léxico, porém utiliza-se de variáveis e define relações entre elas para estimar o sentimento (Direção, Tempo verbal, Polaridade Total, Polaridade de Evento);
- “Upgrade” do matching léxico.

Table 1. Rules for emotion detection.

Direction	Tense	<i>Input Variables</i>		<i>Output Emotion</i>
		Overall polarity	Event polarity	
Self	Future	Positive	Positive	Hope
Self	Future	Negative	Negative	Fear
Self	Present	Positive	Positive	Joy
Self	Present	Negative	Negative	Distress
Self	Past	Positive	Positive	Satisfaction
Self	Past	Negative	Negative	Fears-confirmed
Self	Past	Positive	Negative	Relief
Self	Past	Negative	Positive	Disappointment
Other	All	Positive	Positive	Happy-for
Other	All	Negative	Positive	Resentment
Other	All	Positive	Negative	Gloating
Other	All	Negative	Negative	Sorry-for

(a) Event-based.

<i>Input Variables</i>		<i>Output</i>
Direction	Polarity	Emotion
Self	Positive	Pride
Self	Negative	Shame
Other	Positive	Admiration
Other	Negative	Reproach

(b) Action-based.

<i>Input Variables</i>		<i>Output</i>
Event	Action	Emotion
Joy	Pride	Gratification
Distress	Shame	Remorse
Joy	Admiration	Gratitude
Distress	Reproach	Anger

(c) Compound emotions.

Implementação

Passo a passo

- Processar arquivo;
- Obter sentenças;
- Obter palavras;
- POS tagging;
- Dependency Parsing;
- Análise de tempo verbal;
- Cálculos de polaridade (total e parcial). <- Maior desafio

Resultados

Table 2. Statistics of the datasets. “Total” denotes the original number of sentences in each emotion category while “Implicit” denote the number of sentence which do no contain any emotion words according to WordNet-Affect.

Emotion Total Implicit			Emotion Total Implicit			Emotion Total Implicit		
Joy	1095	537	Joy	362	317	Happy	406	103
Fear	1095	366	Fear	160	130	Fearful	121	33
Anger	1096	483	Anger	66	60	Angry-Disgusted	174	84
Sadness	1096	488	Sadness	202	182	Sad	247	90
Disgust	1096	484	Disgust	26	24	Surprised	92	50
Shame	1096	581	Surprise	184	160	Total	1040	360
Guilt	1093	482	Total	1000	873	(c) Alm's		
Total	7667	3421	(b) SemEval					
(a) ISEAR								

Table 3. Performance comparison of F-measure results on the three datasets. Bold face values denote the best results obtained in each dataset.

Emotion	ISEAR			SemEval			Alm's		
	Lexicon	NB	Rule	Lexicon	NB	Rule	Lexicon	NB	Rule
Joy/Happy	33.4	61.2	69.6	39.7	71.7	59.9	58.8	63.5	81.8
Fear/Fearful	0	47.6	18.3	0	52.2	31.8	0	26.7	14.0
Anger/Angry-Disgusted	23.0	47.1	61.3	55.8	16.2	61.3	48.9	58.6	86.6
Sadness/Sad	25.6	55.4	68.0	47.8	56.0	71.5	61.0	56.0	79.6
Disgust	25.6	51.0	39.2	38.5	34.5	61.7	-	-	-
Average	21.5	52.5	51.3	36.4	58.2	57.3	42.2	56.0	65.5
Average (– Fear)	27.0	53.7	59.5	45.5	44.6	63.6	56.12	65.8	82.7

Resultados

JOY	2457
SADNESS	1799
FEAR	1450
ANGER	431
DISGUST	316
ITD	1213

Conclusões

Conclusões

- Melhoria entre 17-30% em relação ao Matching Léxico;
- Performance competitiva com o método supervisionado;
- Média de 82,7% de identificação para categorias “Happy”, “Anger-Disgusted”, “Sad.;
- Enviesado por base;
- Necessidade humana em toda iteração;
- Complexidade aumenta conforme a quantidade de variáveis aumenta;
- Limitado por uma tabela;
- Necessita um refinamento melhor.