

Análisis de Sentimiento

...

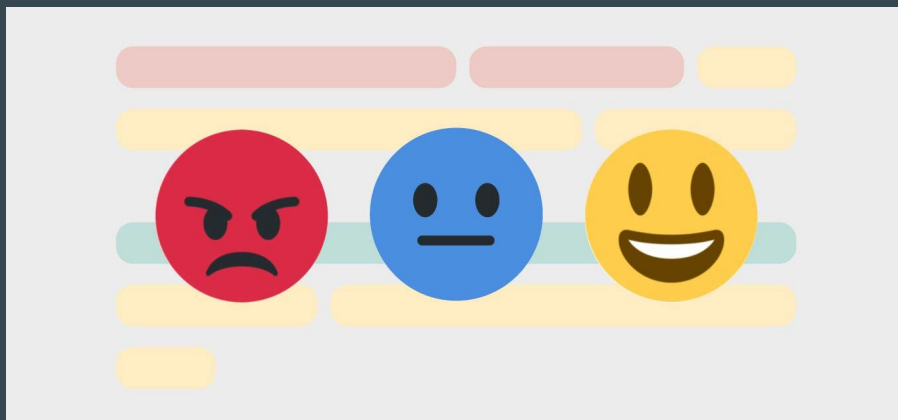
KeepCoding - Bootcamp de Big Data & Machine Learning

Índice

1. ¿Qué es?
2. Subjetividad, polaridad e intensidad
3. Aproximaciones
4. VADER
5. Machine Learning / Deep Learning
6. Retos

1. ¿Qué es?

- En NLP, el análisis de sentimiento es el proceso mediante el cual se busca detectar la **subjetividad**, **polaridad** e **intensidad** (a veces) en un fragmento de texto



2. Subjetividad, polaridad e intensidad

- Subjetividad

- Contenido objetivo (hechos) o subjetivo (opiniones)
 - Objetividad: *'hoy es jueves'*
 - Subjetividad: *'me gustan los perros'*

- Polaridad

- Orientación (media) de la emoción expresada.
- En general positivo vs negativo

- Intensidad

- Fuerza del sentimiento u opinión
 - *'me gustan los perros'*
 - *'me gustan mucho los perros'*

3. Aproximaciones

- **Herramientas de monitoreo**
 - En general, útiles para social media
 - Business Intelligence
 - De pago
- **Modelos basados en reglas**
 - En general, lexicons con palabras y pesos asociados a su sentimiento (valencia)
 - VADER (Valence Aware Dictionary and sEntiment Reasoner)
 - TextBlob
- **Machine Learning / Deep Learning**
 - Modelo pre-entrenado
 - Modelo pre-entrenado + fine tuning
 - Aprendizaje supervisado
 - Aprendizaje no supervisado

3. Aproximaciones

	Pros	Cons
Herramientas de monitoreo	Family friendly Visualizaciones Insights	Licencia Performance Muy limitado
Algoritmos basados en reglas	No necesidad corpus Ajustado a slang / emojis / etc Fácil implementación	Algo de programación Propósito concreto Modelo muy simple
Aprendizaje supervisado	Performance Flexibilidad Estado del arte	Algo + de programación Conocimiento específico Corpus

4. VADER

- Valence Aware Dictionary and sEntiment Reasoner
- Modelo de reglas con un **lexicon** *tuneado* con el propósito de detectar el sentimiento expresado en **redes sociales**
- Cada **palabra** del lexicon tiene asignado un **score**
- Estima **polaridad** e **intensidad**

```
#Baseline sentence
sentiment_analyzer_scores('The food here is good')

The food here is good----- {'neg': 0.0, 'neu': 0.58, 'pos': 0.42,
'compound': 0.4404}

#Punctuation
print(sentiment_analyzer_scores('The food here is good!'))
print(sentiment_analyzer_scores('The food here is good!!'))
print(sentiment_analyzer_scores('The food here is good!!!'))

The food here is good!----- {'neg': 0.0, 'neu': 0.556, 'pos': 0.44
4, 'compound': 0.4926}
None
The food here is good!!----- {'neg': 0.0, 'neu': 0.534, 'pos': 0.46
6, 'compound': 0.5399}
None
The food here is good!!!----- {'neg': 0.0, 'neu': 0.514, 'pos': 0.48
6, 'compound': 0.5826}
None
```

4. VADER

- Pros

- Buenos resultados en redes sociales y en textos con un lenguaje similar
- Slang / emojis / exclamaciones
- No requiere de corpus de entrenamiento
- Performance *acceptable*

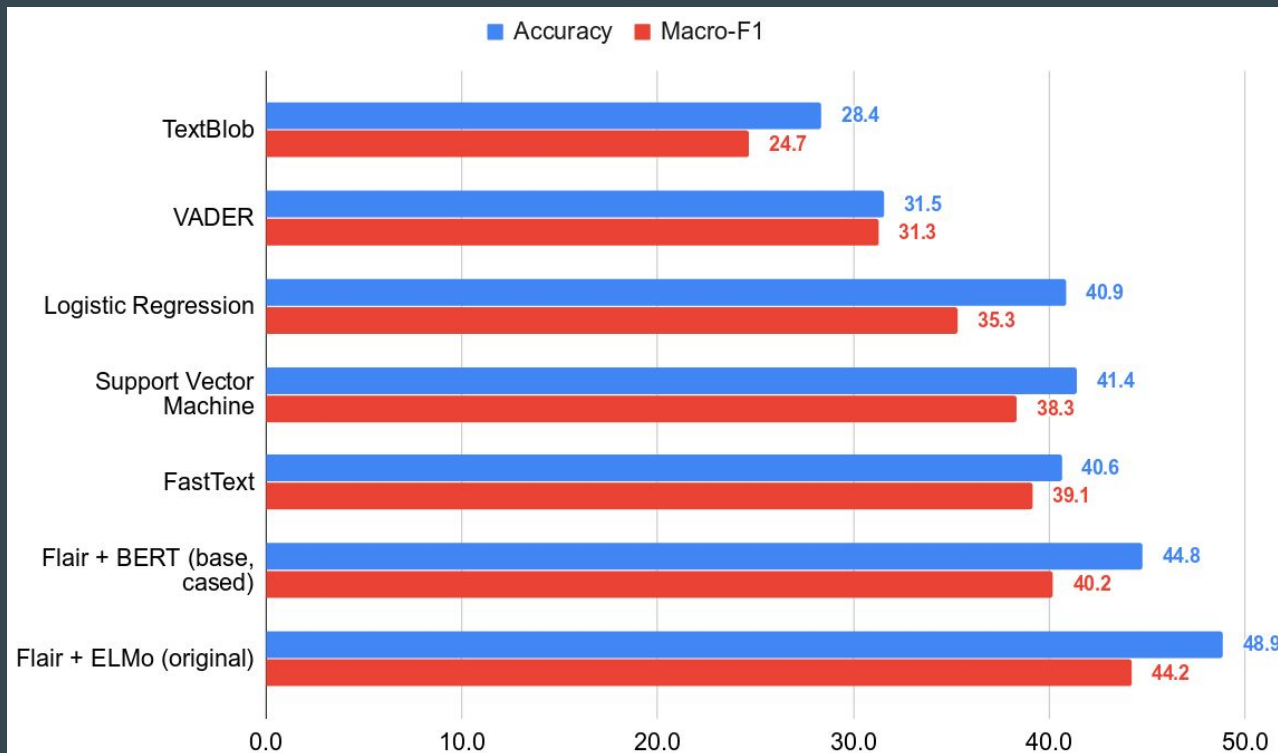
- Cons

- Modelo basado en reglas
- Muy específico
- Solo para Inglés

5. Machine Learning / Deep Learning

- Corpus con texto + etiquetas:
 - Aproximación igual a cualquier otro problema de aprendizaje supervisado que conozcamos
- Corpus sin etiquetas
 - Aproximación igual a otro problema de aprendizaje no supervisado que conozcamos
- Las features más utilizadas serán
 - **Bag-of-words** (term presence, term frequency, TF-IDF, ...)
 - **Word embeddings** (no contextuales y contextuales)
 - Otras features:
 - PoS tags
 - Entidades
 - Lexicon
 - ...

5. Machine Learning / Deep Learning



6. Algunos retos

- Negaciones

- *'Me gustan los perros'*
- *'No odio a los perros'*

- Ironía y sarcasmo

- *'Has sido de gran ayuda, gracias...'*
- *'No trabajes tan duro...'*

- Sentimiento negativo y positivo en el mismo texto

- *'A pesar de que la cena me gustó mucho, el postre estaba muy malo'*

- Neutralidad

- *'Me gustaría conocer más sobre el tema para opinar'*

- Otros retos comunes en NLP

- Errores gramaticales, OOV, desbalanceo de clases, etc.

7. Bonus - DeepMoji

DeepMoji has learned to understand emotions and sarcasm based on millions of emojis. Here's a [video](#) explaining a bit more. Type a sentence to see what our AI algorithm thinks.

You love hurting me, huh?

SUBMIT

Words are highlighted based on emotional impact. Click a word to turn it on/off.

you love hurting me , huh ?



high confidence

¡Vamos al lío!