

Procesamiento de Lenguaje Natural (NLP)

Clase 3: Análisis de Sentimientos

Prof. D.Sc. BARSEKH-ONJI Aboud

Faculty of Engineering
Universidad Anáhuac México Sur

17 de octubre de 2025

Agenda de la Clase

1. De los Números a la Acción
2. ¿Qué es el Análisis de Sentimientos?
3. El Algoritmo VADER
4. Ejemplo Práctico en MATLAB
5. Conclusiones y Próximos Pasos

Agenda

1. De los Números a la Acción
2. ¿Qué es el Análisis de Sentimientos?
3. El Algoritmo VADER
4. Ejemplo Práctico en MATLAB
5. Conclusiones y Próximos Pasos

¿Dónde nos quedamos?

Resumen de la Clase 2

¡Logramos el gran objetivo de la vectorización! Aprendimos a convertir texto en una matriz numérica usando **Bag-of-Words** y la ponderación **TF-IDF**.

Tokens Limpios

'bateri', 'dur', 'much', ... ↓ *Vectorización* ↓ **Matriz Numérica (TF-IDF)**

0.25	0.25	0	...
0	0	0.35	...

La Gran Pregunta

Ya tenemos nuestros datos en un formato que la máquina "entiende". **¿Ahora qué podemos hacer con ellos?**

Nuestra Primera Aplicación: Análisis de Sentimientos

Objetivo de Hoy

Usar nuestros conocimientos de NLP para realizar una de las tareas más comunes y útiles: determinar automáticamente la **emoción** o **polaridad** de un texto.

¿Este comentario es Positivo , Negativo , o Neutral ?

Vamos a explorar cómo una máquina puede responder a esta pregunta, una tarea fundamental conocida como **Análisis de Sentimientos** (o minería de opiniones).

Agenda

1. De los Números a la Acción
2. ¿Qué es el Análisis de Sentimientos?
3. El Algoritmo VADER
4. Ejemplo Práctico en MATLAB
5. Conclusiones y Próximos Pasos

Definición

Es el proceso de usar NLP para identificar y extraer información subjetiva de un texto. En términos simples, es la tarea de **clasificar la polaridad de un texto**.

¿Por qué es tan importante?

Para las Empresas:

- ¿Qué opinan los clientes de mi nuevo producto?
- Monitorear la reputación de la marca en redes sociales.
- Entender la satisfacción del cliente a partir de encuestas.

Para el Público:

- Analizar la opinión pública sobre un tema político.
- Resumir críticas de películas o libros.
- Detectar ciberacoso o discurso de odio.

Enfoques para el Análisis de Sentimientos

Existen varias formas de abordar este problema. Hoy nos centraremos en la más directa:

Enfoque Basado en Léxicos (Diccionarios)

- **La Idea:** Se utiliza un diccionario pre-construido (léxico) donde cada palabra tiene una puntuación de polaridad.
- **Ejemplo de Léxico:**
 - "excelente": +0.9
 - "bueno": +0.6
 - "malo": -0.7
 - "terrible": -0.9
- **¿Cómo funciona?** El algoritmo analiza un texto, busca las palabras en el léxico, y calcula una puntuación final sumando las polaridades.

Otro Enfoque: Basado en Machine Learning

En este caso, en lugar de un diccionario, se entrena un modelo con miles de ejemplos de textos ya etiquetados como positivos o negativos.

Agenda

1. De los Números a la Acción
2. ¿Qué es el Análisis de Sentimientos?
3. El Algoritmo VADER
4. Ejemplo Práctico en MATLAB
5. Conclusiones y Próximos Pasos

VADER: Una Herramienta Especializada

VADER: *Valence Aware Dictionary and sEntiment Reasoner*

Es un modelo basado en léxico que está específicamente **optimizado para analizar sentimientos en textos de redes sociales**.

¿Qué lo hace especial?

- **Considera la puntuación:** Reconoce que "!!!" intensifica un sentimiento.
- **Considera las mayúsculas:** Entiende que "ODIO" es más intenso que "odio".
- **Maneja negaciones:** Sabe que "no es bueno" invierte la polaridad de "bueno".
- **Reconoce modificadores:** Entiende que "muy bueno" es más positivo que solo "bueno".

En MATLAB..

¡Tenemos una función lista para usar que implementa VADER! Se llama `vaderSentimentScores`.

Interpretando las Puntuaciones de VADER

La función 'vaderSentimentScores' no devuelve una, sino **cuatro** puntuaciones para cada texto:

- **Positive**: La suma de las puntuaciones de las palabras positivas.
- **Negative**: La suma de las puntuaciones de las palabras negativas.
- **Neutral**: La proporción de palabras que no son ni positivas ni negativas.
- **Compound**: ¡La más importante! Es una puntuación **normalizada y agregada** de todas las demás, que va de -1 (muy negativo) a +1 (muy positivo).

Regla General para Clasificar con Compound

- Si $\text{Compound} \geq 0.05 \rightarrow$ **Positivo**
- Si $\text{Compound} \leq -0.05 \rightarrow$ **Negativo**
- Si no, es \rightarrow **Neutral**

Agenda

1. De los Números a la Acción
2. ¿Qué es el Análisis de Sentimientos?
3. El Algoritmo VADER
4. Ejemplo Práctico en MATLAB
5. Conclusiones y Próximos Pasos

Objetivo

Usar la función `vaderSentimentScores` para clasificar automáticamente un conjunto de comentarios.

Puntuando Comentarios con VADER

```
1 % 1. Datos de ejemplo
2 comentarios = [
3     "Amo este producto, es absolutamente fantastico.",
4     "Estoy seguro de que no me gusta.",
5     "El servicio fue horrible, nunca volvere.",
6     "El paquete llego a tiempo."
7 ];
8
9 % 2. Tokenizar los comentarios
10 documentos = tokenizedDocument(comentarios);
11
12 % 3. Obtener las puntuaciones de sentimiento
13 puntuaciones = vaderSentimentScores(documentos);
14
15 % 4. Mostrar la tabla de resultados
16 disp(puntuaciones)
```

Salida del Código: La Tabla de Puntuaciones

MATLAB nos devuelve una tabla muy clara con las 4 puntuaciones para cada comentario.

```
1 >> disp(puntuaciones)
2     0.1027
3    -0.2960
4    -0.5423
5     0
```

Salida del Código: La Tabla de Puntuaciones

MATLAB nos devuelve una tabla muy clara con las 4 puntuaciones para cada comentario.

```
1 >> disp(puntuaciones)
2     0.1027
3    -0.2960
4    -0.5423
5     0
```

La clave está en la columna 'Compound'

Ahora, podemos usar nuestra regla para interpretar estos números.

Interpretando el Resultado Final

Vamos a clasificar cada comentario basándonos en la puntuación Compound:

- **Comentario 1:** "Amo este producto..."
 - Compound = **0.1027** (≥ 0.05) → **POSITIVO**
- **Comentario 2:** "Estoy seguro..."
 - Compound = **-0.2969** (≤ -0.05) → **NEGATIVO**
- **Comentario 3:** "El servicio fue horrible..."
 - Compound = **-0.5423** (≤ -0.05) → **NEGATIVO**
- **Comentario 4:** "El paquete llegó a tiempo."
 - Compound = **0** (entre -0.05 y 0.05) → **NEUTRAL**

Resultados

Con solo unas pocas líneas de código, hemos construido un sistema funcional de análisis de sentimientos.

Agenda

1. De los Números a la Acción
2. ¿Qué es el Análisis de Sentimientos?
3. El Algoritmo VADER
4. Ejemplo Práctico en MATLAB
5. Conclusiones y Próximos Pasos

- **Aplicación Directa:** El Análisis de Sentimientos es una de las tareas más valiosas y directas del NLP.
- **Léxicos como Herramienta:** Los enfoques basados en léxicos (diccionarios) son una forma rápida y efectiva de analizar sentimientos sin necesidad de entrenar un modelo.
- **VADER:** Es una herramienta poderosa, especialmente para texto informal como el de redes sociales, y es muy fácil de usar en MATLAB.
- **Logro de Hoy:** Pasamos de tener datos numéricos a obtener **conclusiones accionables** sobre el contenido de un texto.

La Próxima Frontera: Construir un Modelo propio

Usar una herramienta pre-entrenada como VADER es fantástico, pero... ¿y si quisiéramos clasificar texto en categorías que VADER no conoce? Por ejemplo, clasificar noticias en "Deportes", "Tecnología" o "Negocios".

En la próxima y última clase, daremos el paso final: aprenderemos a **entrenar, probar y evaluar nuestro propio modelo de Machine Learning** para clasificar documentos desde cero.