

[†]Estos autores contribuyeron igualmente a este trabajo

En esta tarea desarrollamos y evaluamos clasificadores de sentimientos utilizando Naive Bayes (NB) y Regresión Logística (LR) en diferentes categorías del Multi-Domain Sentiment Dataset: "Books", "DVD", "Electronics" y "Kitchen". Los clasificadores fueron entrenados con representaciones de texto basadas en TF-IDF y características extraídas de lexicones. Se compararon los resultados de NB y LR en términos de precisión, recall, F1-score y accuracy.

Diferencias en Recall:

- Naive Bayes mostró un mejor recall en algunas categorías, como Electronics y Kitchen, lo que indica que el modelo fue capaz de identificar más ejemplos positivos correctamente, aunque con una tasa de falsos positivos más alta. Este comportamiento sugiere que NB es más conservador en términos de predicciones negativas, lo que puede ser útil en aplicaciones donde es crucial capturar la mayor cantidad de instancias positivas.

3. Sentiment Analysis: Clasificador Único para Múltiples Categorías

Mientras que la primera parte del proyecto involucró la construcción de clasificadores individuales para cada categoría, la segunda parte, que abordamos en este análisis, se enfoca en desarrollar un único clasificador que pueda generalizar a través de todas las categorías. Esto implica combinar los datos de todas las categorías y evaluar si un modelo consolidado puede igualar o superar el rendimiento de los clasificadores específicos de categoría.

3.1. Metodología

- Se recopilaron y combinaron todas las reseñas etiquetadas (positivas y negativas) de las cuatro categorías mencionadas. Esto resultó en un único conjunto de datos
- Se aplicó un preprocesamiento consistente al texto para normalizar y limpiar los datos
- El conjunto de datos combinado se dividió en conjuntos de entrenamiento y prueba, utilizando una proporción del 80 por ciento para entrenamiento y 20 por ciento para prueba.
- Se generaron dos tipos de representaciones TF y TF-IDF
- Se extrajeron características basadas únicamente en lexicones de sentimientos:
AFINN Lexicon: Proporciona puntajes de sentimiento para palabras individuales. SentiWordNet Lexicon: Asigna puntajes positivos y negativos a palabras basadas en su significado. WordStat Lexicon: Clasifica palabras en positivas, negativas y neutrales.

3.2. Resultados

3.2.1. Rendimiento de los Modelos con Características de Lexicones

Naive Bayes (Características de Lexicones):

Naive Bayes (Características de Lexicones):

- **Precisión:** 0.5364
- **Recall:** 0.9181
- **F1 Score:** 0.6772
- **Exactitud:** 0.5656

Regresión Logística (Características de Lexicones):

- **Precisión:** 0.7194
- **Recall:** 0.6877
- **F1 Score:** 0.7032
- **Exactitud:** 0.7119

La regresión logística superó al modelo de Naive Bayes en todas las métricas. Aunque Naive Bayes mostró un alto *recall*, su precisión y exactitud fueron significativamente menores, lo que indica un alto número de falsos positivos.

3.3. Análisis finales

- Las palabras con mayor peso positivo son adjetivos que típicamente se asocian con opiniones favorables, como great, excellent, best, easy, love.
- Las palabras con mayor peso negativo incluyen not, waste, worst, bad, poor, que son indicadores claros de reseñas negativas.
- La palabra not tiene un peso negativo significativo, lo que resalta su importancia en la inversión del significado de oraciones.