

Reporte de los datos suministrados

El siguiente reporte presenta los resultados utilizando el modelo LogisticRegression, el contenido de este se muestra a continuación:

- 1. Información de los datos
- 2. Steps con los parámetros del pipeline
- 3. Puntuación del modelo
- 4. Gráficas representativa del modelo
- 5. Gráfica de las reseñas

1. Información de los datos

Se encontraron 10 Filas y 1 Columnas, junto con 1 idiomas, los cuales fueron {'es': 10}.El conjunto de datos suministrado además contenía 0 nulos y 0 duplicados.

2. Steps pipeline

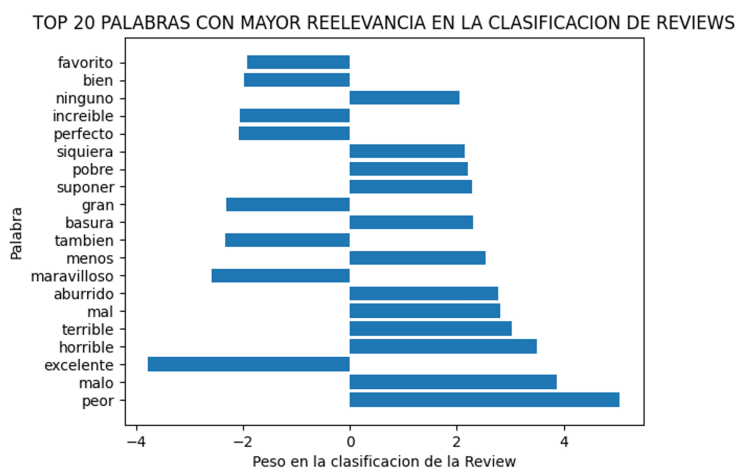
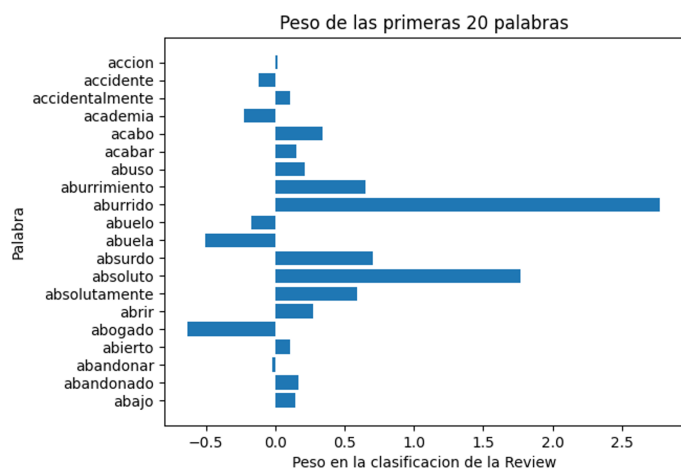
Parameters	Values
steps	[('vectorize', TfidfVectorizer(max_features=3000)), ('lr', LogisticRegression(solver='newton-cg'))]
verbose	False
vectorize	TfidfVectorizer(max_features=3000)
lr	LogisticRegression(solver='newton-cg')
vectorize__analyzer	word
vectorize__binary	False
vectorize__decode_error	strict
vectorize__dtype	<class 'numpy.float64'>
vectorize__encoding	utf-8
vectorize__input	content
vectorize__lowercase	True
vectorize__max_df	1.0
vectorize__max_features	3000
vectorize__min_df	1
vectorize__ngram_range	(1, 1)
vectorize__norm	l2
vectorize__preprocessor	None
vectorize__smooth_idf	True
vectorize__stop_words	None

Parameters	Values
vectorize__strip_accents	None
vectorize__sublinear_tf	False
vectorize__token_pattern	(?u)\b\w\w+\b
vectorize__tokenizer	None
vectorize__use_idf	True
vectorize__vocabulary	None
lr__C	1.0
lr__class_weight	None
lr__dual	False
lr__fit_intercept	True
lr__intercept_scaling	1
lr__l1_ratio	None
lr__max_iter	100
lr__multi_class	auto
lr__n_jobs	None
lr__penalty	l2
lr__random_state	None
lr__solver	newton-cg
lr__tol	0.0001
lr__verbose	0
lr__warm_start	False

3. Puntuación del modelo

Con los datos de entrenamiento se encontró un score de: 0.8487161693268563

4. Gráficas representativa del modelo



5. Gráfica de las reseñas

