

Proyecto Final Inteligencia Artificial

RECONOCIMIENTO DE GÉNERO POR VOZ

Maria Alejandra Acosta - Juliana Toro Camelo

Introducción

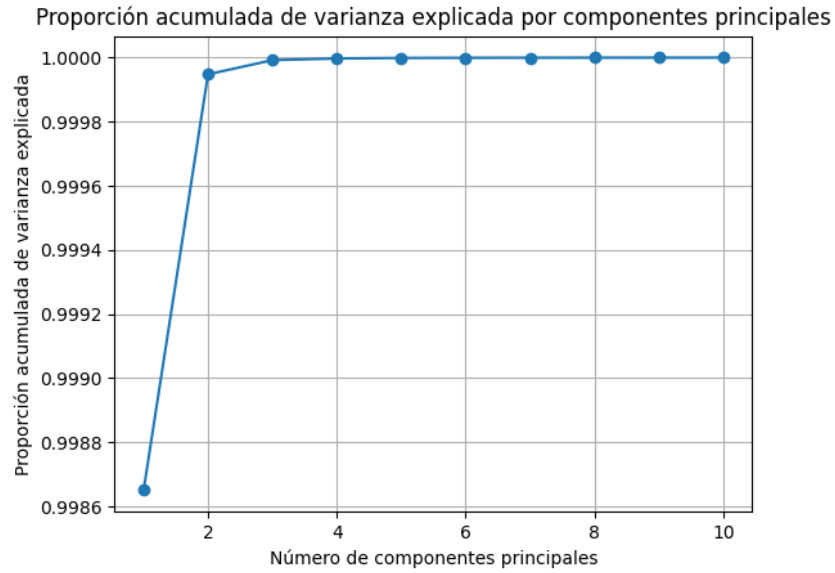
El conjunto de datos *Gender Recognition by Voice* contiene características acústicas de las voces de hombres y mujeres, extraídas de grabaciones de voz y se han utilizado para construir un modelo de aprendizaje automático que puede identificar si una voz pertenece a un hablante masculino o femenino. Incluye 3.168 grabaciones de voz etiquetadas con el género del hablante, así como 20 características acústicas diferentes, tal como la frecuencia fundamental, la energía, la entropía espectral y la tasa de cruces por cero.

Desarrollo

Primero fue necesario importar librerías y cargar los datos usando la biblioteca Pandas. Luego, se realizó un preprocesamiento de los datos, verificando y eliminando los valores nulos, también se utilizó el método LabelEncoder para convertir la columna de etiquetas categóricas a numéricas, se dividieron los datos en conjuntos de entrenamiento y validación utilizando train_test_split, se aplicó la normalización de características utilizando StandardScaler y se realizó una reducción de dimensionalidad utilizando PCA para extraer los componentes principales. Se implementaron varios modelos de clasificación; el SVM (Support Vector Machine), KNN (K-Nearest Neighbors) y Decision Tree. Se evaluaron los clasificadores utilizando métricas como F1-score y MCC (Matthews Correlation Coefficient) para medir su rendimiento. Además, se desarrolló un modelo de regresión lineal utilizando LinearRegression y se calculó el coeficiente de determinación (R^2) para evaluar su precisión. Se visualizó la regresión lineal mediante una gráfica, se calcularon los coeficientes de silueta para cada muestra utilizando silhouette_samples. Por último, se utilizó GridSearchCV para realizar una búsqueda en cuadrícula y encontrar los mejores hiperparámetros para el modelo ElasticNet. Se evaluó el modelo utilizando el coeficiente de determinación (R^2) y se imprimieron los mejores hiperparámetros encontrados.

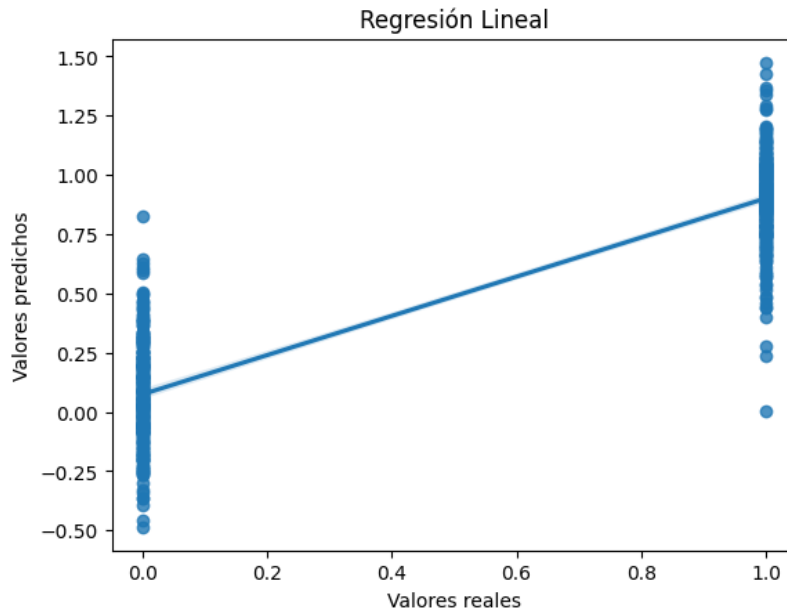
Resultados

- Tamaño del conjunto de entrenamiento: 2534
- Tamaño del conjunto de validación: 634



- Coeficiente de Determinación (R^2): 0.8227714136596993

	SVM	KNN	DT
F1-score	0.6539694103059606	0.7052525376249608	0.9716326117158115
MCC	0.30663916362122257	0.40896590991777787	0.9438057477137742



- Mejores Hiperparámetros: {'alpha': 1.0, 'l1_ratio': 0.5}

Conclusiones

1. Con un tamaño de conjunto de entrenamiento de 2534 muestras, se considera que es adecuado para realizar un entrenamiento y validación efectivos de los modelos. Si bien un conjunto de datos aún mayor puede brindar beneficios adicionales, el tamaño actual proporciona una cantidad razonable de datos para entrenar y evaluar los modelos de manera adecuada.
2. Los resultados, al evaluar métricas como F1-score y MCC en el conjunto de validación, mostraron que el modelo Decision Tree obtuvo el mejor rendimiento, seguido por KNN y SVM. Esto indica que el enfoque basado en árboles de decisión puede ser más adecuado para el problema de clasificación de voz en este caso.
3. El modelo de regresión lineal resulta con una capacidad razonable para predecir los valores de salida, explica aproximadamente el 82% de la variabilidad en los datos de salida.