

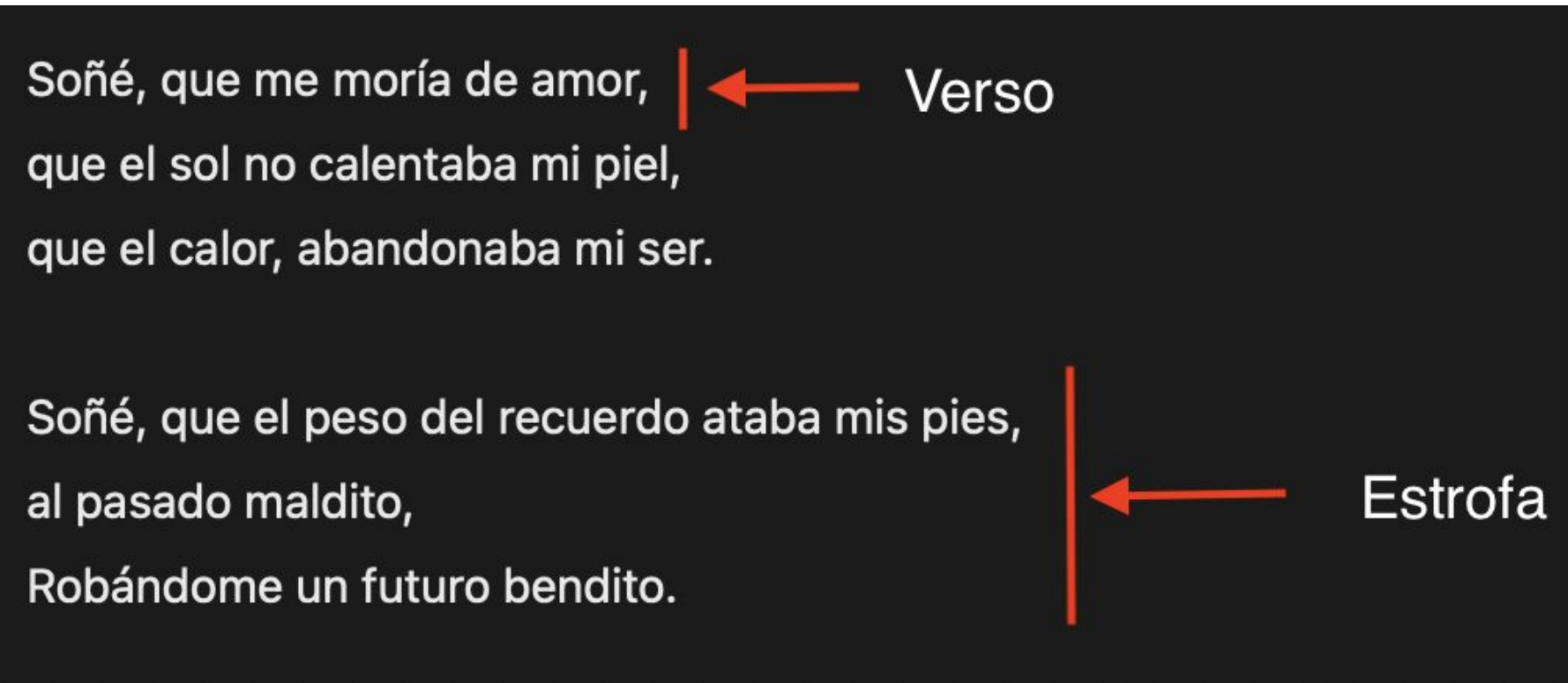
Generación de Poesía utilizando Redes Neuronales Recurrentes

Diego Alvarez



Introducción

El poema, es la composición poética básica. En general, se compone de múltiples versos y estrofas. Cada estrofa se compone de múltiples versos y un poema se compone de una o varias estrofas.



El objetivo principal de la siguiente implementación era generar un poema compuesto por 5 estrofas generado por un modelo de RNN. Para esto se buscó generar versos libres de un máximo de 10 palabras utilizando redes neuronales recurrentes. Para ello, se entrenaron varias RNN utilizando embeddings, LSTM y LSTM bidireccional.

Data Set

Como datos de entrenamiento, se utilizaron los textos de dos libros de poesía en español. El primero, Eres Arte, escrito por Diego Alvarez, es un libro escrito en Guatemala y publicado en Junio del 2020. El segundo, un clásico de la poesía, 20 poemas de amor y una canción desesperada, escrito por Pablo Neruda en 1924.

Arquitectura

Durante el entremiento de los modelos se realizaron diferentes pruebas utilizando LSTM y LSTM bidireccional. El meodelo con el mejor accuracy se logró utilizando una combinación de embedding, LSTM bidireccional y una capa de salida. Dicho modelo obtuvo un accuracy de 0.80 en comparación con 0.74 del mismo modelo utilizando LSTM sencillo.

Layer (type)	Output Shape	Param #
embedding (Embedding)	(None, 157, 100)	289500
bidirectional (Bidirectional)	(None, 300)	301200
dense (Dense)	(None, 2895)	871395
Total params: 1,462,095		
Trainable params: 1,462,095		
Non-trainable params: 0		

Cade destacar que el entrenamiento del modelo con LSTM bidireccional, a pesar de tener un mejor accuracy, se tomaba el doble de tiempo en cada iteración (38s aproximadamente),

Resultados

Después de generar 5 estrofas con distintos modelos, el resultado final se presenta en la imagen 03. Este poema fue el resultado del texto producido por el modelo más la edición del autor. La tabla 01 muestra el % de similitud y diferencia entre el resultado final y la salida de los modelos.

RESULTADOS FINALES	
SIMILITUD	DIFERENCIA
63.04%	36.96%

Quiero tu cuerpo

Quiero tu cuerpo,
He soñado que muero en ese momento,
Donde el arte funde en uno,
Nuestros corazones.

Busco en tí,
Que culminen las guerras,
Y platiquen las almas.

Dame luz de humo,
de estanque en calma,
de rosas de papel.

Entregate a mi,
Rayo de sol,
Como un túnel,
Del que huyen pájaros del cielo.

Quiero tu cuerpo alegre,
Tus luminosos ojos,
Y jamás volver sólo a ese café

AI/ALVOZA

Conclusión

Del entrenamiento se puede observar que los modelos pueden seguir mejorando su accuracy por lo que se recomienda que se entrene el modelo con más textos de poesía para evitar un overfitting de los datos. De esta manera el modelo podrá producir textos elocuentes de mayor longitud. Además cabe mencionar que entre las limitaciones de los modelos realizados se encuentra la carencia de conocimiento de métricas y rimas.

Predicción de Características del Café Utilizando Feed Forward Networks

Diego Alvarez



Introducción

El café es considerado el segundo producto en volumen físico más comercializado en el mundo. Su consumo, a nivel global, se ha duplicado en las últimas décadas, pasando de 92 millones de sacos en 1990 a 162 millones de sacos estimados en 2019. Además, se estima que se bebe alrededor de 4.000 millones de tazas y la demanda sigue en aumento.

Los dos tipos de granos de café son arábica y robusta. El primero es más apreciado en el mercado, mientras el segundo tiende a ser más amargo y menos apetecible, pero tiene una concentración del 50% más de cafeína que el arábica.

El objetivo de esta implementación fue predecir 10 características del café basado en factores cómo país de origen, altitud de cultivo, color, variedad y método de procesamiento.

- Acidez
 - Aroma
 - Sabor
 - Dulzura
 - Uniformidad
- Postgusto
 - Cuerpo
 - Balance
 - Limpieza
 - Humedad

Data Set

El set de datos utilizados fue una recopilación de catas de café realizadas por profesionales alrededor del mundo. El set de datos fue recopilado a lo largo de 5 años por el Coffee Quality Institute.

DATOS ANALIZADOS	
# Países	32
# Variedades de grano	28
# De métodos de procesamiento	5
# Total de muestras	1400

Arquitectura

Para el entrenamiento de este modelo se utilizó una Feed Forward network. La red constaba de tres capas. Una de entrada, una capa media y una capa de salida.

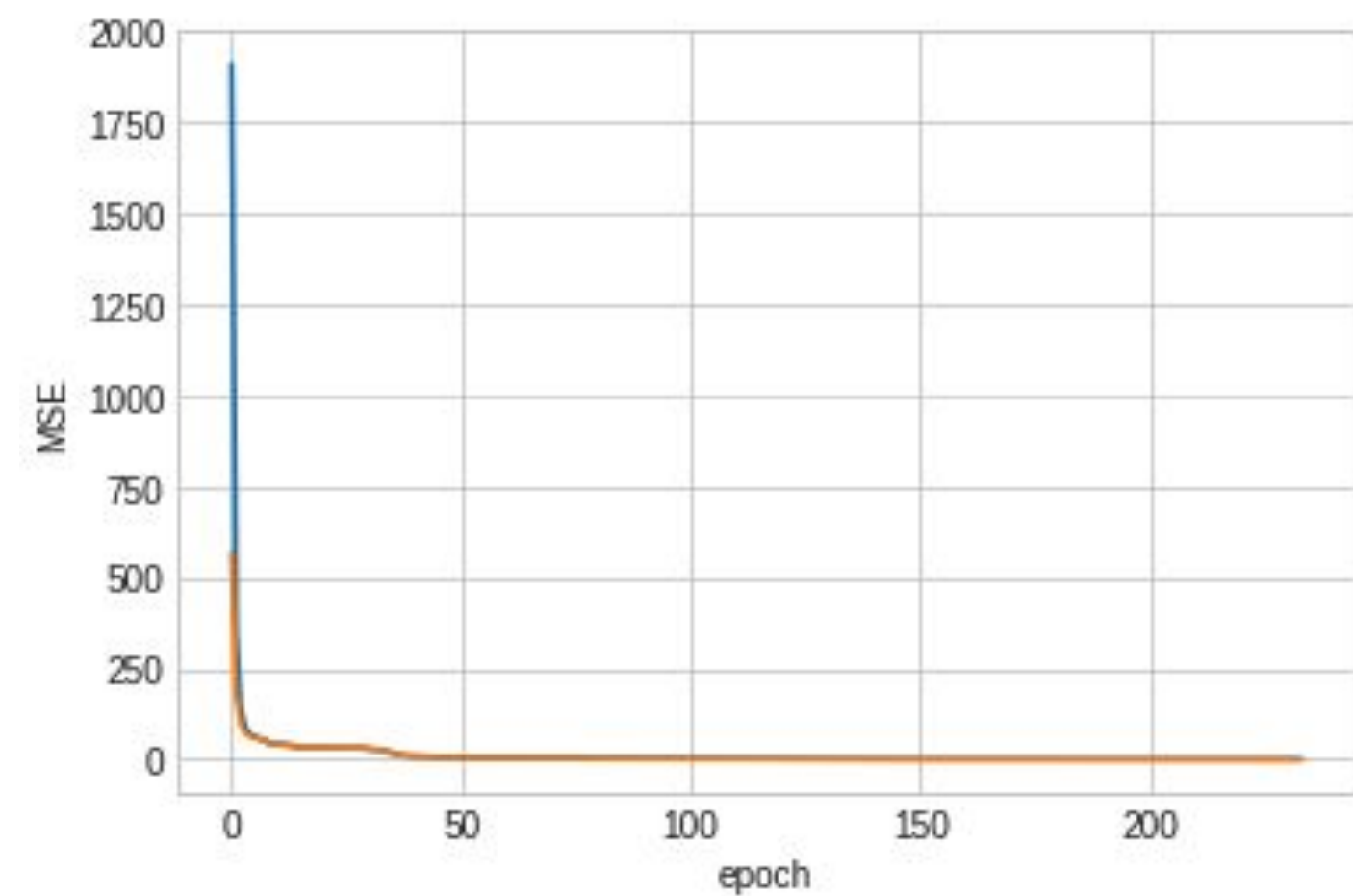
Layer (type)	Output Shape	Param #
=====		
dense (Dense)	(None, 11)	132
dense_1 (Dense)	(None, 6)	72
dense_2 (Dense)	(None, 10)	70
=====		
Total params: 274		
Trainable params: 274		
Non-trainable params: 0		

Además, antes de entrar el modelo se realizó el proceso de Feature engineering durante el cual se trataron valores nulos y outliers.

Resultados

Después de entrenar al modelo se obtuvieron los siguientes resultados:

RESULTADOS FINALES	
MSE TRAIN	MSE VAL
0.1179	0.2766



Conclusión

En conclusión, ya que el modelo posee un valor de error bajo, puede ser utilizado para comparar las predicciones de catadores inexpertos con las predicciones de catadores expertos recopiladas por Coffee Quality Institue. Esto puede ayudar a estudiantes de la catación del café a saber qué tan acertadas son sus predicciones en comparación con las del modelo.

Para expandir el modelo se podría recopilar información de los métodos de preparaciones y grado de tueste del café utilizados durante la catación para examinar los efectos que estos factores tienen en la tasa final del café.

Reconocimiento de Caracteres Utilizando Convolutional Neural Networks

Diego Alvarez



Introducción

El reconocimiento óptico de caracteres es el uso de tecnología para distinguir caracteres de texto impresos o escritos a mano dentro de imágenes digitales de documentos físicos, como un documento en papel escaneado. El proceso básico de OCR implica examinar el texto de un documento y traducir los caracteres a un código que se puede utilizar para el procesamiento de datos. A veces, el OCR también se denomina reconocimiento de texto.

El objetivo de esta implementación fue crear una red Convolutiva capaz de clasificar un caracter en una de 62 categorías. Para ello se entrenó un modelo capaz de clasificar una imagen en los caracteres A-Z, a-z y 0-9.

Data Set

Una de las dificultades con este dataset fue la cantidad limitada de imágenes disponibles para entrenar el modelo. Se utilizaron 55 muestras para cada caracter.

DATOS ANALIZADOS	
# 62 Categorías	A-Z a-z 0-9
# Muestras por caracter	55
# Total de muestras	3,410

Arquitectura

Para mejorar el accuracy del modelo se decidió utilizar la funcionalidad de image augmentation para generar diferentes versiones de una misma imagen. Esto ayudó a que el modelo tuviera más data disponible para su entrenamiento.

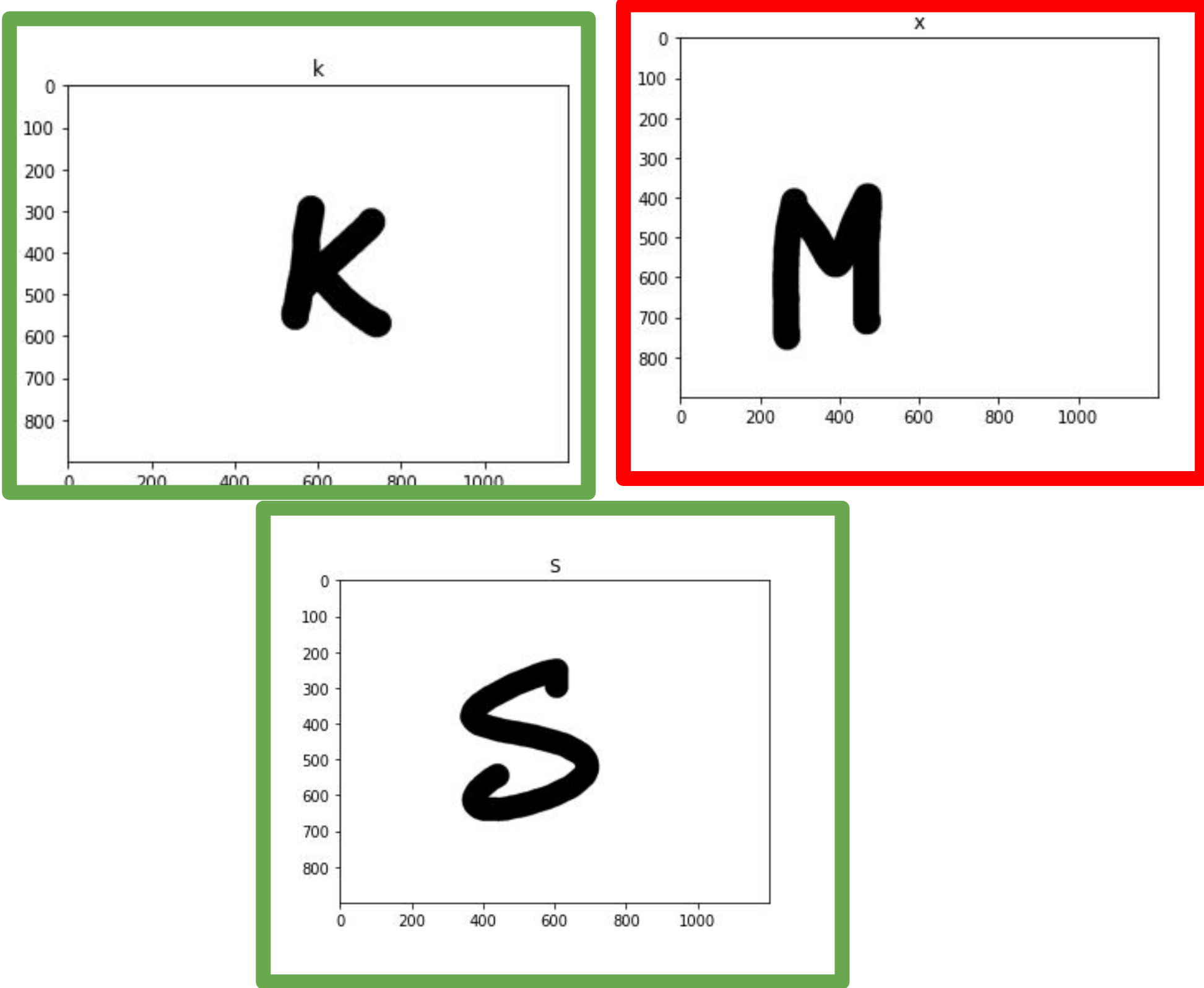
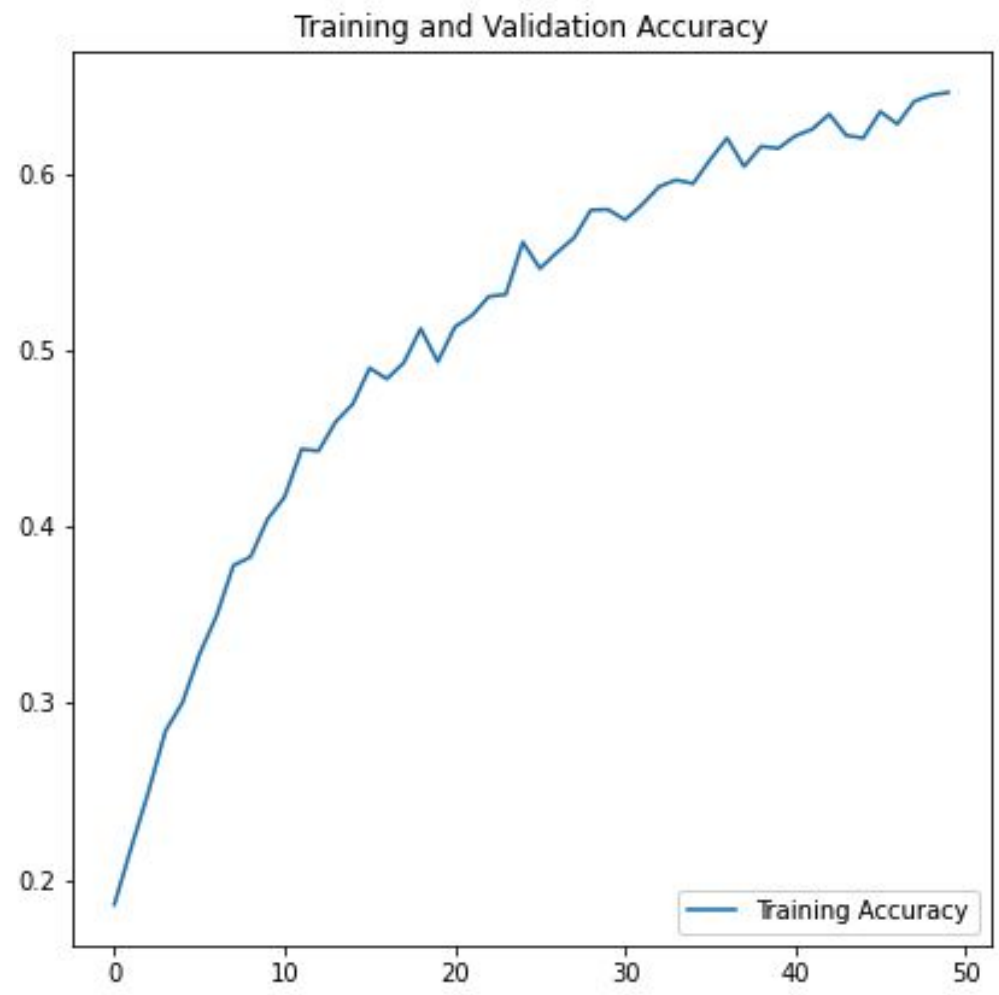
Layer (type)	Output Shape	Param #
=====		
batch_normalization_16 (Batch Normalization)	(None, None, None, 3)	12
conv2d_24 (Conv2D)	(None, None, None, 32)	896
max_pooling2d_20 (MaxPooling2D)	(None, None, None, 32)	0
conv2d_25 (Conv2D)	(None, None, None, 32)	9248
max_pooling2d_21 (MaxPooling2D)	(None, None, None, 32)	0
flatten_14 (Flatten)	(None, None)	0
batch_normalization_17 (Batch Normalization)	(None, 7200)	28800
flatten_15 (Flatten)	(None, 7200)	0
dense_20 (Dense)	(None, 512)	3686912
dropout_10 (Dropout)	(None, 512)	0
dense_21 (Dense)	(None, 62)	31806
=====		
Total params: 3,757,674		
Trainable params: 3,743,268		
Non-trainable params: 14,406		

Para el modelo final, se utilizaron como capas intermedias capas convolucionales, max pooling, batch normalization y flatten. Para la salida se utilizó una capa densa con activación Softmax para obtener la mejor predicción del modelo.

Cabe mencionar que el modelo tiempo significativamente alto en comparación con los otros modelos (1hr). Esto se debe a que el modelo tenía que leer los archivos de las imágenes mientras estaba entrenando.

Resultados

Después de entrenar el modelo, se llegó a un resultado final de Accuracy = 0.67. Al realizar pruebas con el modelo se puede observar que algunas predicciones aún no incorrectas, como se muestra a continuación.



Conclusión

En conclusión a pesar de que el proceso de data augmentation ayudó a mejorar el modelo, este no se encuentra aún en un nivel de accuracy aceptable para ser utilizado en un sistema. Se recomienda entrenar el modelo con más imágenes de caracteres.