

Entrega 1

Equipo de trabajo

CRISTHIAN JAVIER GONZALEZ RODRIGUEZ

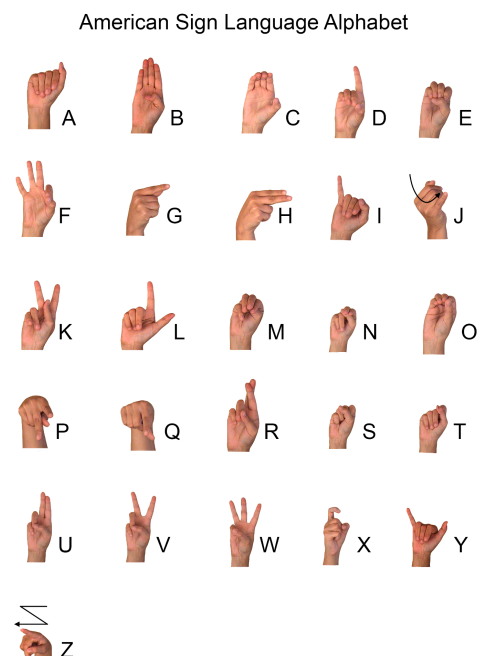
LAURA DANIELA MONSALVE GÓMEZ

MATEO RIVERA MONSALVE

Contexto de aplicación

Nuestra aplicación o desarrollo consistirá en un sistema de clasificación multiclase de imágenes del Alfabeto del lenguaje de Señas inglés, es decir, contamos con **27 clases** diferentes (uno por cada letra) numeradas (etiquetadas) del **0** al **26**, donde el **0** es una etiqueta vacía, el **1** sería el índice correspondiente a la letra **A** mientras que el **26** sería el correspondiente a la letra **Z**.

Es decir, el sistema será capaz de predecir la seña correspondiente a cada letra, individualmente, más no el lenguaje como tal, la analogía correcta sería por ejemplo que hay varios idiomas que comparten el mismo alfabeto o gran parte del mismo, pero las palabras son distintas, en este caso, buscamos clasificar cada elemento del alfabeto de señas, es decir, cada letra o dicho en otras palabras “la representación de cada letra en el alfabeto del lenguaje de señas” más no las palabras de dicho lenguaje.

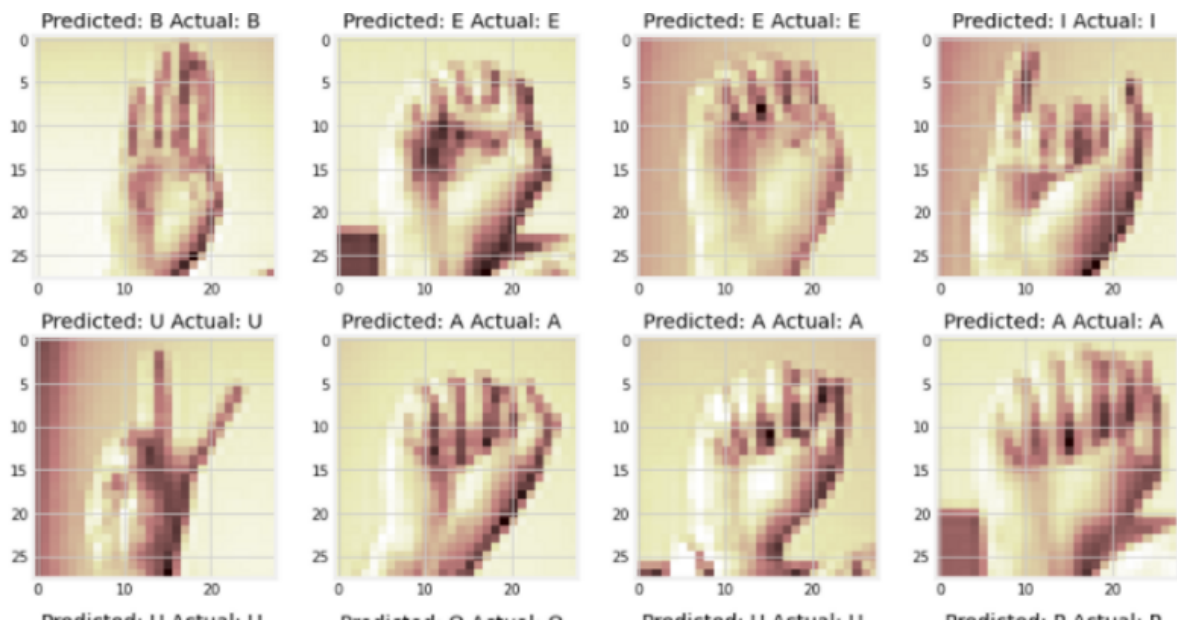


Las aplicaciones o usos que puede llegar a tener este sistema de clasificación son bastante amplios, puede permitir a las personas aprender o educarse en lo que respecta a este lenguaje como también facilitar la comunicación entre las personas, en términos pueden haber buenos beneficios.

Objetivo de machine learning

Como se ha venido comentando, contamos con un conjunto de fotos de manos de **28 × 28 píxeles (784 características)**, donde en cada imagen hay una mano realizando una posición del alfabeto del lenguaje de señas inglés (Difiere en el español prácticamente en la ausencia de la letra "ñ"), cada imagen está etiquetada con un número que varía entre 0 y 26 (Que es el índice de la letra a la que corresponde la posición excepto 0).

Dada una imagen de una mano haciendo una letra del alfabeto del lenguaje de señas inglés, queremos clasificarla en alguna de las letras del alfabeto inglés (Proyecto que podría extenderse para conseguir imágenes de la letra ñ para países hispanohablantes).

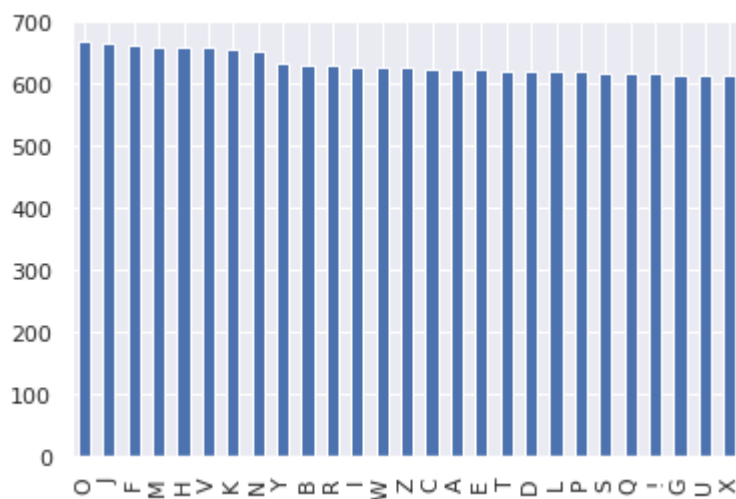


Dataset

Cada característica (Columna) es de tipo int64 con el siguiente rango: $[0, 255]$.

Este dataset se encuentra dividido por defecto en dos subconjuntos: Entrenamiento y Test, los subconjuntos tienen **12844** y **4267** muestras respectivamente.

El dataset completo tiene un tamaño en disco de **49.5 MB**.



La cantidad de muestras por clase es bastante balanceada, puesto que todas están en el rango $(600, 700)$, además se puede evidenciar en el histograma.

Métricas de desempeño

Se usará la matriz de confusión normalizada para hacer un análisis sobre qué tan bien está clasificando un modelo, y ver con qué clases puede estar confundiéndose a la hora de clasificar.

Como medida global se usará la exactitud (Accuracy) sobre cada una de las clases, sin embargo se estudiará más a profundidad cada clase con la sensibilidad (Recall) y la precisión (Positive Predictive Value).

Referencias y resultados previos

- Curso de **Fundamentos de Deep Learning** (UdeA) del profesor Raúl Ramos Pollán: <https://rramosp.github.io/2021.deeplearning/content/M02.html>
- C. Bishop, [**Pattern Recognition and Machine Learning**](#)
- Haykin, S. S., Haykin, S. S., Haykin, S. S., & Haykin, S. S. (2009). [**Neural networks and learning machines**](#) (Vol. 3). Upper Saddle River, NJ, USA:: Pearson.