Problem Statement Worksheet (Hypothesis Formation)

A charity project has contracted me to craft a model which can classify handwritten names at NLT 70% accuracy. A neural net model must be created for implementation by Q1 of next year.

H

1 Context

A charity project has over 400,000 names in their analog database that needs to be transferred to a newer digital database. The names of each person have been handwritten by each individual person, so there will be variation in spellings, inks, and styles. The project needs the names to be accurately classified at 70% accuracy or greater.

2 Criteria for success

Recognizing names and their spelling at a minimum of 70% accuracy.

3 Scope of solution space

This dataset consists of more than four hundred thousand handwritten names that must be input into a digital database.

4 Constraints within solution space

Character Recognition utilizes image processing technologies to convert characters on scanned documents into digital forms. It typically performs well in machine-printed fonts. However, it still poses difficult challenges for machines to recognize handwritten characters, because of the huge variation in individual writing styles. Other constraints could be spellings of similar names, ink used, etc.

5 Stakeholders to provide key insight

Dataset Owner: Landlord
Contributors: Jebastin Nadar, Aman Kumar, Aryan Tiwari

6 Key data sources

- https://www.kaggle.com/datasets/landlord/handwriting-recognition/discussion/180739?datasetId=818027
- License - CC0: Public Domain

H D E I P