# *Which machine and deep learning model is the best at recognizing handwriting?*

**Research Project - Turing Machine and Deep Learning**

Paulo de Souza Miranda

Daris Fadhilah

Aristotelis Giannakopoulos

Max Viertel Serrano

Abstract

In this report we study the accuracy and speed of two machine and deep learning models in recognizing handwriting. We answer the following research question *‘Which machine and deep learning model is the best at recognizing handwriting?’.* The models considered are a decision tree and a neural network. To run the models the data was first preprocessed to fit the necessary requirements. Our results showed that both decision trees and neural networks perform well in recognizing handwriting, however the performance of neural networks is significantly stronger. Therefore, the neural network model has the highest accuracy. In other settings, time required to train and validate is also an important criterion to choose a model, however for this context the time differences between the two models were not significant.

## Introduction

Machine and deep learning models are becoming increasingly relevant for solving a multitude of societal problems. In our report we will investigate the efficiency of a machine learning model in interpreting handwritten text. More specifically, we aim to answer the research question *'Which machine and deep learning model is the best at recognizing handwriting?'*. The type of task is classification, and we will use a neural network and decision tree models. In answering the research question, we will thus investigate the hypothesis of whether neural networks are the more accurate model. We furthermore investigate the effect that removing one dense layer from the neural network has on the model's accuracy and flexibility.

The dataset used for building the model is from Kaggle which is considered a reliable source for finding datasets. This dataset consists of more than four hundred thousand handwritten names collected through charity projects. It can be found under the following link: https://www.kaggle.com/datasets/landlord/handwriting-recognition

Previous research on the topic has been conducted and the prospective classification task is relevant for document digitization, data entry automation and forensic analysis.

## Methods

The central issue at hand is recognizing human handwriting. This is a complex process, as it involves complex structures (letters) which vary significantly across individuals. Furthermore, the underlying patterns and structures do not follow a linear relationship, complicating the issue at hand.

That means that any linear methods are likely to not perform well in capturing patterns of letters that vary across different "authors". For this reason, we believe neural networks and decision trees are well suited for the task at hand, given their flexibility and ability to capture non-linear relationships.

We therefore first split the data of the set into a training and test set. Each model, thus a neural network and a decision tree, will be trained on the same set to ensure equal conditions for success.

Seeking to answer the hypothesis, we compare the accuracies of each model on the test set. For the other question we seek to answer, our 'original' neural network consists of three dense layers whereas we compare it to one with two dense layers.

In terms of the preprocessing that was necessary in order to train our models, every word was split into individual letters. This was done through means of the "findContours()" function by OpenCV after having identified 'usable' instances of words with the help of "Fuzzy Distance" and "OCR".

We needed to do this because some instances of data we used were significantly different from the "typical" occurrence, consider for instance a printed word in front of the handwritten word that was necessary to train the data or evaluate a prediction. Dealing with these outliers while preserving their information in a generalized manner using code was not realistic in the scope of these papers, so these instances were removed from the employed data.

Results

To answer the first hypothesis, we compare the accuracy of the two models on the testing data. As shown in the accuracy measures in Figure 1 and Figure 2, we see that neural networks have the highest validation accuracy compared to decision trees, minimizing our expected risk.

Furthermore, we ran the neural network model with varying numbers of dense layers, namely one with three dense layers and one with two. The effect on the model's flexibility can be clearly seen from the figures (figure 2 and 3) presented below.

Figure 1: Accuracy, decision tree

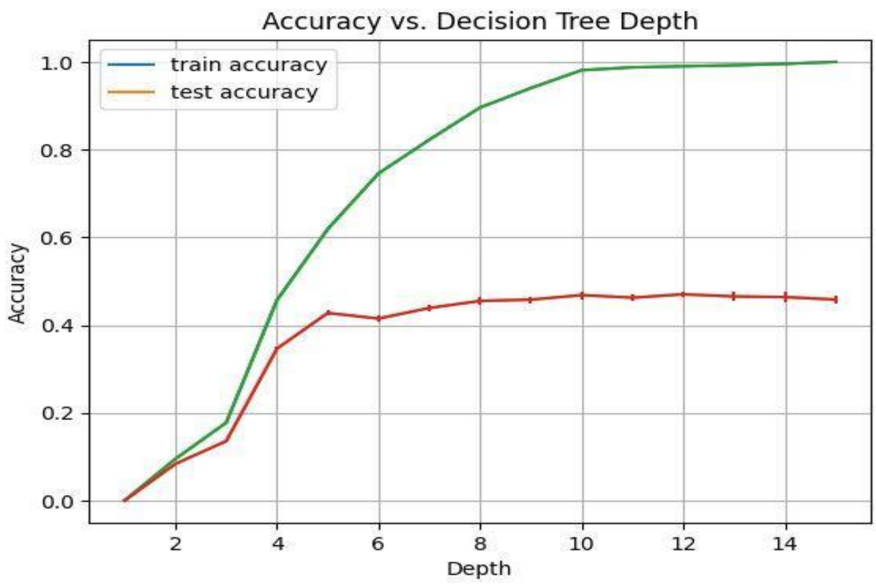



Figure 2: Accuracy, neural network

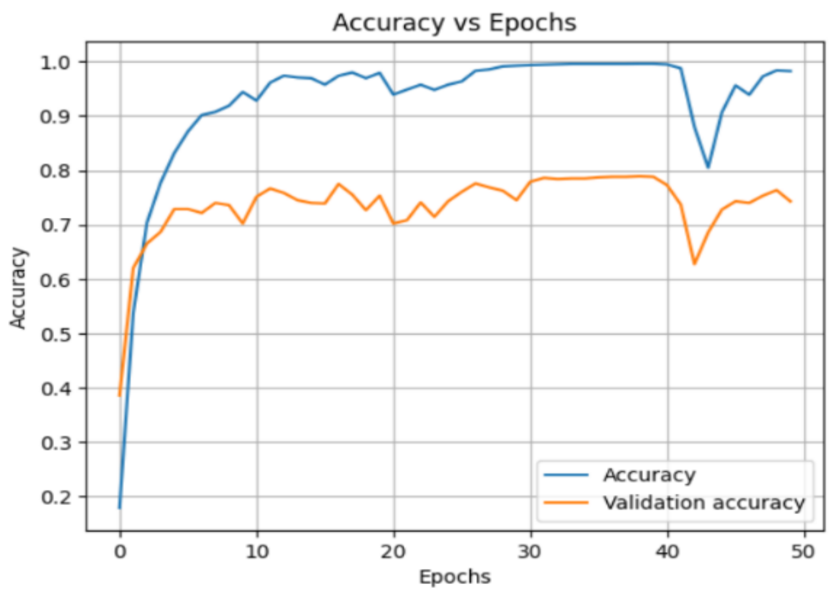
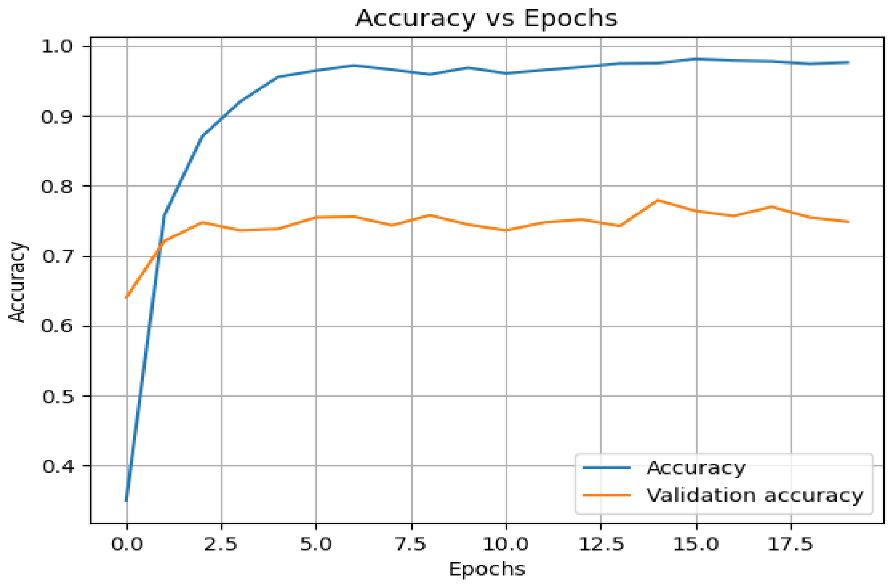


Figure 3: Accuracy neural network after removing one dense layer



## Discussion

As stated, we see that neural networks have a very strong validation accuracy measure. In this context and its magnitude, neural networks are thus always the best choice. Nonetheless, one must be aware that with increasing data and parameter size the computation time for training and validating a neural network goes up very significantly, potentially rendering the model less optimal than others.

Furthermore, the increased variation across epochs of the neural network with more dense layers can be attributed to the fact that more parameters are trained and updated across every epoch, allowing for more flexibility across these epochs and thus increasing the variability of the accuracy and loss measures. Naturally, the neural network with more dense layers performs better in terms of validation accuracy than the 'reduced one'.

Coming back to our research question, we do deem neural networks the best choice of model for recognizing handwriting, as in the magnitude and context of our research problem there was no significant time difference between decision trees and neural networks. It is also important to note that this ranking still involves some degree of subjectivity, as the question is how to weigh these performances. This is a consideration that changes across data considered and goals of the analysis.

An issue we encountered was heterogeneity across our data set, which caused the need for heavy data preprocessing. For example, sometimes there were other (printed) words and labels in front of the actual words we were trying to "read".

The homogeneity the preprocessing caused can be seen as one of the weaknesses of our data set, as it makes replicability on other data more difficult. Furthermore, in some cases we also had to remove specific instances of handwritten words, which reduced the data available to train our model, potentially causing loss of data points that would prove to be informative for the individual models, thereby improving accuracy and other performance measures.

Another potential issue with the comparison of our models is the choice of parameters, such as the number of dense layers to use and the tree depth. There is some difficulty associated with this because this is not fully comparable across models. The way we accounted for this was to choose the optimal measure of parameters that prevented overfitting and minimized expected risk for each individual model.

## Conclusion

The main findings of our paper were that neural networks are indeed the most accurate model in recognizing human handwriting compared to the other models used. This accuracy goes up with an increasing number of layers and neurons per layer.

The issue of this model, however, is its computation time, one disadvantage it has over its peers. Nonetheless, for the scope of this data set this notion is not so relevant, as computation does not take hours, for example. We therefore conclude that neural networks are the best machine and deep learning model for recognizing handwriting, well-aware that the accuracy – computation time tradeoff is not the same in other settings, which might affect the choice of model.

We have thus been able to determine the strength of neural networks for the referred to dataset, nonetheless it is to be noted that a heavy degree of preprocessing was involved and that these modifications resulted in a quite homogenous dataset, also with all capital letters. In this way, we optimized the conditions for our models to work in.

For further research, it would be interesting to expand the possibilities of inputs to machine and deep learning models to recognize handwriting, especially when considering activities such as document digitization and data entry automation, which involve a variety of different writing, ranging from printed to handwritten and employing different fonts in one document, for instance. Considering the performance of our models on this dataset, we are confident that the mentioned instances are issues that can reliably be tackled by machine learning models. Additionally, it would be interesting to compare the performance of inherently different machine learning models compared to neural networks for similar tasks.