

MACHINE LEARNING ENGINEER NANODEGREE

ML approach to bank_note_authentication

Capstone proposal

SandulaLavanya

January 30 , 2019

Proposal

Domain Background

Bank note authentication now a days gone popular since, Manual testing of all notes in transactions is very time consuming and untidy process and also there is a chance of tearing while handing notes. Banknotes are one of the most important assets of a country. Some miscreants introduce fake notes which bear a resemblance to original note to create discrepancies of the money in the financial market. It is difficult for humans to tell true and fake banknotes apart especially because they have a lot of similar features. Fake notes are created with precision, hence there is need for an efficient algorithm which accurately predicts whether a banknote is genuine or not. This paper proposes machine learning techniques to evaluate authentication of banknotes. Every year RBI (Reserve bank of India) face the counterfeit currency notes or destroyed notes. Handling of large volume of counterfeit notes imposes additional problems. Therefore, involving machines (independently or as assistance to the human Experts) makes notes recognition process simpler and efficient.

Automatic method for detection of fake currency note is very important in every country.

In this project i have made fake currency note detection technique using Machine learning with tensor flow and feature extraction from WTI(wavelet transformed images). It seems like Machine Learning can yield good results in bank note authentication. This is why I would like to try my own methods in this project.

The main motivation behind development of this project was to make a system for easy and quick detection of genuine and fake Indian currency notes.

Reference:

https://subscription.packtpub.com/book/big_data_and_business_intelligence/9781787124479/3/03lvl1sec46/using-neural-networks-for-classification

Problem Statement

The projects aims to create a machine learning model that can predict a note is genuine or forged by using the data of images(Skewness, curtosis , variance, entropy, class) using WTI. The task is to built an effective algorithm that takes the data taken from UCI Repository as input and outputs a binary form(1 or 0) for given data. The inputs will contain image variance(cont.),skewness(cont.),curtosis(cont.) and entropy(cont.) ,class(data of note as forged or genuine)of the bank note image.my algorithm need to find the genuine and forged notes returns 1 and 0.

Datasets and Inputs

In this project, I will use bank_note_authentication dataset from UCI Repository. This contains Img.Var, Img.Skew, Image.Curt, Entropy, Class. The dataset is copied and made it as a .csv (comma separated values). It can be easily read using pandas from scikit-learn. I will use these values as inputs and try to predict a correct binary value i.e 0 or 1. The high accuracy made in predicting (using tensorflow) it shows the efficiency of the model.

1. **variance_of_wavelet_transformed**, used as input.
2. **skewness_of_wavelet_transformed**, used as input.
3. **curtosis_of_wavelet_transformed**, used as input.
4. **entropy_of_image**, used as input.
5. **class**, used as target.

Solution Statement

In this project, I would like to use tensorflow to solve this problem. Tensorflow is an open source library built by Google, widely used in the field of machine learning and deep learning. The library is popular for its use of data-flow graphs to carry out numeric computations. I will use TensorFlow to build an artificial neural network that detects fake banknotes.

Building neural network using tensorflow

Choose a neural network architecture to represent the classification function. Depending on the number of inputs, neurons in the hidden layers, and outputs, the architecture and the result of the neural network will be different.

Benchmark model

The data I took from UCI Repository are imbalanced and the benchmark model for this project would be Random Forest Classifier. With given n_estimators that gives us maximum accuracy score. This model is proved to be successful in predicting the images (bank_notes) as genuine or forged about 99% accurately.

Reference:

This benchmark model uses the same input as the tensorflow and provide a benchmark performance for tensorflow.

Evaluation metrics

Since this is a classification problem, so the metrics of choice would be Accuracy score and f1 score. Accuracy measures how often the classifier makes the correct prediction. It's the ratio of the number of correct predictions to the total number of predictions. f1-score is weighted average (harmonic mean) of the precision and recall scores. This score can range from 0 to 1, with 1 the best possible F1 score (we take the harmonic mean as we are dealing with ratios).

Project Design

=>Imports

->importing libraries

=>Reading data using pandas

->.csv file saved from UCI Repository

=>Exploring data analysis(data set)

->relationship between forged and genuine notes using seaborn

=>Data preparation

->Standardising the data

=>Splitting the data into testing and training sets.

=>Defining parameters for neural network.

=>Defining weights and bias for neural network.

=>Calculating loss and optimizing it

=>Constructing a neural network

=>Training a neural network

=>Model evaluation

=>Comparing with the benchmark model

Reference:

Data set from the UCI Repository.

Guiding reference from https://www.neuraldesigner.com/learning/examples/banknote_authentication