FORGED CURRENCY PREDICTOR USING NEURAL NETS

Abstract:

This document is a model and instruction for prediction of the currency weather it is authentic or forged . For implementation various Machine learning algorithms are used. This project gives the accuracy of various models predicting the suspect. We used Random forest, Multi-layer perceptron for prediction.

Keywords- Python, Machine learning, Random Forest, Multilayer perceptron, neural nets, libraries used: Numpy, Pandas, Sklearn, matplotlib.pyplot.

CHAPTER 1:

INTRODUCTION:

Counterfeit money is imitation currency produced without the legal sanction of the state or government usually in a deliberate attempt to imitate that currency and so as to deceive its recipient.

Forged Currency notes possesses a threat to economy on a global scale in ways like:

- Devaluation of currency and inflation.
- Black marketing
- Dumping of goods at cheaper rates
- Non-reimbursement by banks.

Loss of public confidence

The main problem arises because of these counterfeit money are

- To detect forged currency notes.
- The fake bills can be hard to spot with an untrained eye, and sophisticated digital printers have made it easier for criminals to create higher-quality bad bills, filter. And the problem appears to go growing.
- In India, after demonetization the instances for fake currency usage doubled, which increased the need to find better ways to detect fake currency.

The project will find utility with small businesses which can include - grocery stores, petrol pump stations, and other businesses with similar transaction limits. The model's implementation also helps the general economy by increasing the confidence in the currency and reducing inflation.

The main objective of the project is to help the local businesses against currency fraud which may lead to huge personal as well as economic loss.

CHAPTER 2:
METHODOLOGY

PREPARATION OF MODEL:

- For preparing a model the following steps we followed are:
- Imported libraries: pandas,numpy,matplotlib.pyplot,sklearn
- Import the dataset
- Data preprocessing: Standardization of data using Normalization.
- Visualization
- Defining dependent and independent variables
- Splitting the data into training and test data.
- Fitting the model
- Calculate cost
- Applied optimizer to minimize cost.
- Prediction with the test data
  We used Multilayer perceptron as initial model then considered random forest for different level of accuracy.

  A. ABOUT THE LIBRARIES:

  • Scikit-learn: Scikit-learn is one the most popular ML libraries. It supports many supervised and unsupervised learning algorithms. Examples include linear and logistic regressions, decision trees, clustering, k-means and so on.

  • Pandas: Pandas is a very popular library that provides high-level data structures which are simple to use as well as intuitive.It has many inbuilt methods for grouping, combining data and filtering as well as performing time series analysis.Pandas can easily fetch data from different sources like SQL databases, CSV, Excel, JSON files and manipulate the data to perform operations on it. There are two main structures in the library.

• Matplotlib: It is a standard Python library used by every data scientist for creating 2D plots and graphs. With enough commands, you can make just about any kind of graph you want with Matplotlib. You can build diverse charts, from histograms and scatterplots to non-Cartesian coordinates graphs.

B. Determining dependent and independent variables:

In here we define dependent and independent variables and plot them as bar graph to show the difference in distribution of Forged and Authentic currency.

C. Dataset:

Our dataset is a CSV file that contains information extracted from (wavelet transformed) images of banknotes. There are 1,372 banknotes, each with the following attributes:

- **Image.Var** (Variance of Wavelet Transformed image (WTI))
- **Image.Skew** (Skewness of WTI)
- **Image.Curt** (Curtosis of WTI)
- **Entropy** (Entropy of image)
- **Class** (Whether or not the banknote was authentic)

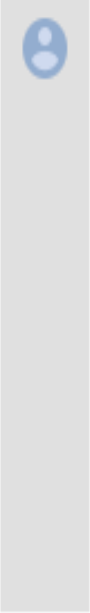


	Image.Var	Image.Skew	Image.Curt	Entropy	Class
0	3.62160	8.6661	-2.8073	-0.44699	0
1	4.54590	8.1674	-2.4586	-1.46210	0
2	3.86600	-2.6383	1.9242	0.10645	0
3	3.45660	9.5228	-4.0112	-3.59440	0
4	0.32924	-4.4552	4.5718	-0.98880	0

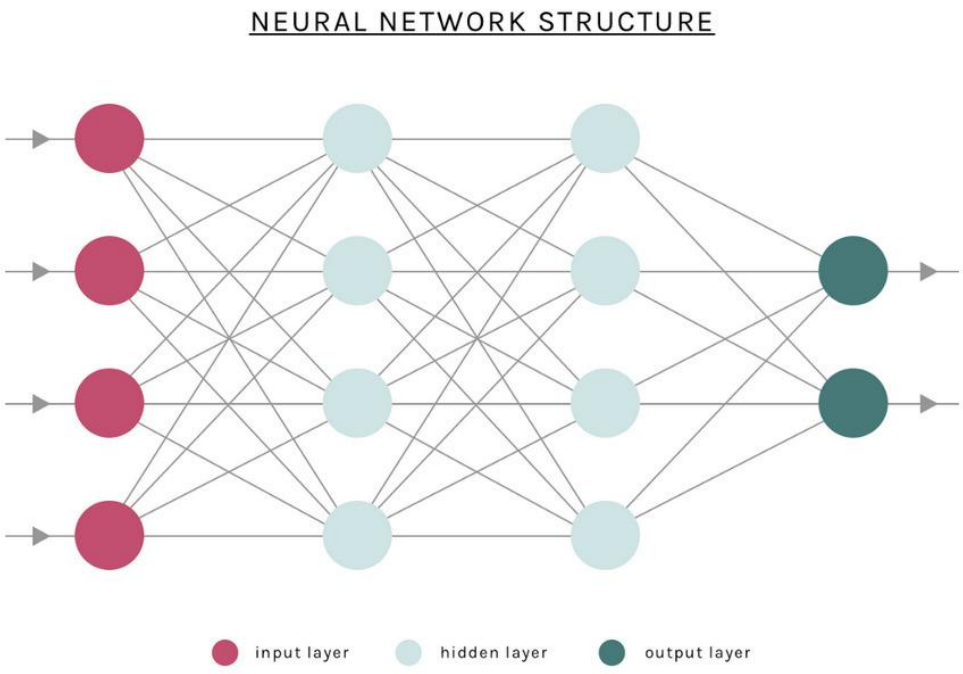
PROCEDDINGS:

Calculating projections along the x and y axes of an image. Standard deviation, skewness, kurtosis. These steps are repeated for all the images of currency notes and the dataset is prepared. We got the pre-processed currency notes dataset from UCI repository. Then using the matplotlib library the dataset target value is visualized and it is found that the number of authentic notes are more then the forged notes but the frequency is close. In addition to it, we find the relationship between the attributes by using pairplot() from Seaborn . Then the data given to us is normalized using standardization technquies. Later using the scikit library the data is split into training and test set in the ratio of 70:30.

MODELS:

For the modelling purpose neural nets were used and multi-perceptron model was used to wih parameters such as learning rate=0.01, training epocs=100, batch size=100 for the best performance. For the model 4 hidden layers were used in the neural net , given our dataet thetotal number of inputs given are 4 and the output classes were 2 namely 0(authentic) and 1 for (forged). The weights and bias for each layer in our network has to be defined, for this particular step a dictionary is created which jeeps tracks of weights and biases using the parameters which are already defined.

Hence the model has 3 layers(2 hidden layers and 1 output layer, excluding the input layer)



Cost and optimization:

For the optimization we use the softmax-cross entropy to calculate cost as
A **softmax layer**, allows the neural network to run a multi-class function. In short, the neural network will now be able to determine the probability that the target value is in the image, as well as the probability that additional objects are included as well.

For the minimization of cost obtained from the softmax-optmizer we used adam optimizer.For adam optimizer the learning rate adam converges very fast and the learning rate can be handcrafted. For the dictionary consisting of weights and bias a function is created that accepts the input x. For the activation function ReLu function is used. **ReLu** is less computationally expensive than tanh and **sigmoid** because it involves simpler mathematical operations. That is a **good** point to consider when we are designing deep neural nets.

TRAINING THE DATA:

- We have two loops in training:
    - The outer loop runs the epochs. , in our processing data the number of epochs used are 100 and for each epoch the learning rate is 0.01.
    - The inner loop runs the batches for each epoch. The batch size to run each epoch is 100.

After each epoch, we can print on the cost and append it to a list of costs. The way we can plot a line graph after training to visualise how our cost has been minimized.

CHAPTER 4:

RESULT AND CONCLUSION

MODEL EVALUTION:

After training the model using multilayer perceptron with 100 epochs, learning rate of 0.01 and the batch size of 100., the accuracy attained is of 95%. The reasons behind is that the dataset provided to us publically is already unclustered and balanced. Hence the layers of our neural net learn a lot faster as compared to learning for a clustered data.

For prediction other models otherthan Multi-layer perceptron was also used Random Forest and the model prediction was created using a confusion matrix. The accuracyof Random Forest was less than Multi-layer perceptron.

CONCLUSION:

Applying these two models we have demonstrated the feasibility of using multi layer perceptron to detect counterfeit notes. The process of detection of fake notes easy and quick under the trained data. By this we can ensure that under real and large data set the model can be trained and provides accurate results, which can help the people in recognising the fake currency and help in strengthening the economy.