**Progress Report: Animal Image Recognition**

The report mainly discusses the progress of the animal image recognition project. The report will discuss the methodologies, challenges and the approaches for addressing the challenges. Finally, the author will list out the future steps that needs to be completed.

**Quick Overview**

As the name suggests, the project aims to identify animal images that are categorized in ten categories, which is an improved version of digital recognition. Different from digit image, animal images contain more noise and colors, thus makes it harder to perform. The dataset contains 26179 images with ten different animal categories. The distribution of units in different categories is listed below:

Chart

Description automatically generated

It is clear to discover that the images of dogs and spiders dramatically outweigh the other categories, and it is possible that the two categories need to be truncated to similar data size.

**Data Preprocessing**

As multiple machine algorithms are included and the author decided to build a deep neural network, the parameters of deep neural network need to be different from the other machine learning algorithms’ parameters.

In deep neural network, TensorFlow library provides a build-in image data generator that could read and store the data in its certain form. For the function image data generator, it contains multiple parameters that could be tuned:

* rotation\_range: the maximum range the image will be rotated randomly as sample, range from 0 to 1
* shear\_range: the maximum range the image will be sheared randomly as sample, range from 0 to 1
* zoom\_range: the maximum range the image will be zoomed randomly as sample, range from 0 to 1
* horizontal\_flip: whether the image could be flipped horizontally for sampling, true/false
* vertical\_flip: whether the image could be flipped vertically for sampling, true/false
* rescale: the value that the whole matrix will multiply.
* validation\_split: the percentage of cross validation.

Except rescale and validation, all other will modify the original image. The purpose of allowing the function to randomly modify the image is to let the computer capture the feature of the image even if the image have been modified to some state. However, those values needs proper tuning since it the random range got too far, the image will be unrecognizable even for human. For example, if the image got sheared too much, the remaining part will be rather confusing.

As for rescale value, it will be 1/255 to make the rgb value between 0 to 1.

As for other machine learning algorithms, they will apply gray scale to compress the data into 1/3.

**Challenges & Approaches**

The author mainly encountered two challenges while preprocessing the data.

When training the Naïve Bayes Classifier, the model failed to accept matrix as input, and the output error indicate that the input was a 3D array and expect a 2D layer. The original solution was to create a huge pandas dataframe that records all the pixels by column using for loop. After the help from professor, it’s more time economic to flatten the matrix and consider it as input.

Also, the author previously mentioned that the image size will be resized to 224\*224. However, now the image size is now 128\*128 due to the limitation of computational power.

**Future Steps**

The previous week turned out to be a data cleaning process, and the following week will start training and tune the model and compare the performances of each algorithm. The author will apply the following algorithms below with comments:

Convoluted Neural Network: The most popular yet difficult algorithm that requires assembling and constant tuning. There will be a list of parameter and activation function that needs to be taken into consideration. Convoluted Neural Network requires a fixed set of parameters and ignores the sequence of each parameter, which fits well with the image dataset.

Naïve Bayes Classifier: The algorithm that author have high expectation besides convoluted neural network. Naïve Bayes Classifier performs good in image recognition since it performs better when each variable has less correlation.

Random Forest: As one of the ensemble learning, random forest by no means is useless as it could extract features and display what are the key pixels that might play a significant role.

Support Vector Machine: The algorithm that author will perform if time is sufficient. Due to the nature of support vector machine, it will be not suitable for this dataset since there will be ten categories, which makes the training more time consuming. What’s worse, SVM usually requires high computational power, with the fact that convoluted neural network is already time-consuming.