Term End Milestone-3 (Project -1)

IDENTIFICATION OF RICE VARIETIES

List Of expected Audience Questions and corresponding Answers

DSC, Bellevue University

Prashant Raghuwanshi

DSC680-T301 Applied Data Science (2225-1)

 Professor Catie Williams

18/03/2022

List Of expected Audience Questions and corresponding Answers

1. In the Real world how the data science team is going to capture the required rice grain data.

Answer: We have two ways to get rice grain preprocessed data

* Directly procure the required preprocessed grain data from agriculture university (or any available third party vendors) and have a contract with them for futures updates in rice varieties data
* Do the data processing in-house and procure the required equipment & Software for capturing & process the grain image data.
  + The images of rice samples were obtained first and images were processed using various image processing techniques. The resulting images are first converted to a grayscale image, then converted to a binary image and removed from the noise on the image. In the next phase, various morphological feature inference processes were applied to the obtained images.

1. For procuring data from vendors or universities what ethical considerations do we need to take care of.

Answers:

* we are going to request data validity, original work certification, and copyright approval from vendors. Also need to make sure the vendor should be responsible to arrange the required government clearances.

1. What would be our revenue model?

Answer:

* we are going to provide installable software packages and their license to robotics scanner companies. And also having a plan to get some revenue from after-sales support from scanner companies.

1. The presented solution or model is meant for identifying five varieties of rice

How you are going to make it fit for other varieties?

Answer:

* our developed solution can be used for any additional addon varieties of rice. However, before making it ready for any new variety of rice we need to retrain our model with updated or new datasets.

1. How do you conclude the LR model is the best fit?

Answer :

* We have built the model by using multiple classification algorithms and after computing the respective models’ accuracy and other factors, we concluded the LR is going to best fit for our use case.

1. How accurate is your Model?

Answer:

* In this project, I have used clean and preprocessed data that contains equal no of samples counts for each variety. I believe for this project I have used only 5 varieties of rice grains and have got the model max accuracy of 99.8%.

1. With the increase in counts of varieties of grains, are you expecting any degradation in, accuracy?

Answer :

* Theoretically, we might get degradation inaccuracy, however, we are going to build and evaluate the models by using evolving multiple algorithms, which may result in a steady accuracy percentage (similar to current accuracy)

1. How the Model is designed to behave with unknown grain data.

Answer:

* At present, the model is not enough mature to identify the unknown gain variety data, but definitely, we are going to add the additional unknown classification variety in our training datasets.

1. please elaborate more on provided Recommendations:

Answer :

* This model is suitable for implementation in any Automatic gain scanning system and it can be designed for many gains identification processes such as calibration of rice types and the separation of species from unwanted substances that may be present.

1. Do you have a recommendation for making the model prediction more consistent:

Answer:

* To increase the success rate in classification, more images can be obtained from species and it is thought that success rates can be increased by using morphological features as well as color and shape features.