UDACITY

Logout

Return to "Data Scientist Nanodegree" in the classroom

DISCUSS ON STUDENT HUB

Data Scientist Capstone

REVIEW | CODE REVIEW | HISTORY

Meets Specifications

Great Work!! 👏

Congratulations on completing the project!!

- Great analysis and you have chosen a nice topic to explore too! Your project proves that you have a good understanding of Data Science
- I certainly enjoyed walking through your code. It's very clean and I can clearly see the effort that has been put into this.
- Hopefully I will see some more informative projects from you in future

Good Luck in your Data Science journey, keep learning. 😊

Project Definition

Student provides a high-level overview of the project. Background information such as the problem domain, the project origin, and related data sets or input data is provided.

- Great Job in providing with an Overview at the beginning of the project!! 👏

The problem which needs to be solved is clearly defined. A strategy for solving the problem, including discussion of the expected solution, has been made.

- I can see some introduction information here but it is highly suggested that you highlight each of the section in markdowns cells and highlight them such :-

Problem Statement

- Here you need to mention what problems you trying to solve and what will be your objective and strategy

Metrics used to measure performance of a model or result are clearly defined. Metrics are justified based on the characteristics of the problem.

- Be sure to clearly mention about the metric!!
- Be sure to check out this Blog on evaluation metrics

Analysis

Features and calculated statistics relevant to the problem have been reported and discussed related to the dataset, and a thorough description of the input space or input data has been made. Abnormalities or characteristics about the data or input that need to be addressed have been identified.

- Be sure to explore the distribution of the target class (for example number of each dog breed). Do we have an unbalanced dataset or Balanced?
- Make sure you analyze the distribution of the classes. What are the image sizes. Are they of the same size or varying sizes? etc. Some descriptive stats would also be nice to see here.

Build data visualizations to further convey the information associated with your data exploration journey. Ensure that visualizations are appropriate for the data values you are plotting.

- Exploratory Visualizations have been done wonderfully well!! 😊

Methodology

All preprocessing steps have been clearly documented. Abnormalities or characteristics about the data or input that needed to be addressed have been corrected. If no data preprocessing is necessary, it has been clearly justified.

- It would be great if you add in some more information on how your preprocessed the data for computer vision task after the line Data Preprocessing of the images consist of transforming images to tensors. This includes resizing and normalizing pixel values.

FOR EXAMPLE

```
The preprocessing done in the "Prepare data" notebook consists of the following steps:
1. The list of images is randomized
2. The images are divided into a training set and a validation set
3. The images are split into square-shaped segments; random noise is used for padding
4. Each of the segments gets a label, which is "text" if the overlap between the segment and one
of the annotations is greater than a threshold, and "no-text" otherwise
```

- You can also add in the preprocess steps you used before loading the image data like converting to grayscale and normalizing the pixel values!!

The process for which metrics, algorithms, and techniques were implemented with the given datasets or input data has been thoroughly documented. Complications that occurred during the coding process are discussed.

The process of improving upon the algorithms and techniques used is clearly documented. Both the initial and final solutions are reported, along with intermediate solutions, if necessary.

- Here you need to mention all the hyperparameter tunning steps, or other steps that led you to your final model.

For example as given in your sample review of computer vision application

Refinement

```
As mentioned in the Benchmark section, the LeNet architecture trained with Caffe achieved a
good accuracy, around 80%. (This is considered good because around 20% of the positive
training examples contain only a single character, or just some portion of a character, and
as such are essentially false positives.)

To get the initial result, this architecture was ported to TensorFlow; the result was an
accuracy around 70%. This was proved upon by using the following techniques:
* Dynamic learning rate: whenever the loss function stopped decreasing, a learning rate drop
was added
* Weight decay: when overfitting was detected (the training and validation losses diverged too
much), the weight decay rate was increased
* Adding dropout to a layer: dropout randomly drops weights in the layer it's applied to
during training and scales the weights so that the network keeps working during inference.
This was later undone because the Android build of TensorFlow doesn't support it.
```

Results

If a model is used, the following should hold: The final model's qualities — such as parameters — are evaluated in detail. Some type of analysis is used to validate the robustness of the model's solution.

Alternatively a student may choose to answer questions with data visualizations or other means that don't involve machine learning if a different approach best helps them address their question(s) of interest.

The final results are discussed in detail.
Exploration as to why some techniques worked better than others, or how improvements were made are documented.

- You could improve the readme by providing more details of the results and details using visuals.

Conclusion

Student adequately summarizes the end-to-end problem solution and discusses one or two particular aspects of the project they found interesting or difficult.

- Great work with the conclusion in the blog. You can add more information on how you started and the steps followed to end up with the final solution.
- Be sure to mention the aspects that you found interesting.

Discussion is made as to how at least one aspect of the implementation could be improved. Potential solutions resulting from these improvements are considered and compared/contrasted to the current solution.

- I highly suggest you have a dedicated section on improvements in the readme. Do some more research on your problem and what all could be implemented, the idea here is just that you are aware of many possible improvements that are available for your set of problem.

For example as given in your sample review of computer vision application

Improvement

```
To achieve the optimal user experience, using more capable hardware and moving the text extraction process from the cloud to the device would be essential. This would reduce the processing time and give access to the outputs of all of the modules of the text extraction pipeline, which would, in turn, enable the foll
owing features:

* User-guided reading (e.g. read big text first, or read the text the user is pointing at)
* Better support for languages other than English
* Output filtering (e.g. ignore text smaller than some adjustable threshold)
* Passive text detection (auditory cue on text detection, perhaps with additional information encoded in the tone and volume)

The user experience could also be improved significantly by using MXNet, which is a deep learning library that is better optimized for mobile devices than TensorFlow. The speedup wouldn't be enough for running text extraction on the device, but it would reduce the classification delay significantly.
```

Deliverables

If the student chooses to provide a blog post the following must hold: Project report follows a well-organized structure and would be readily understood by a technical audience. Each section is written in a clear, concise and specific manner. Few grammatical and spelling mistakes are present. All resources used to complete the project are cited and referenced.

If the student chooses to submit a web-application, the following holds: There is a web application that utilizes data to inform how the web application works. The application does not need to be hosted, but directions for how to run the application on a local machine should be documented.

- Awesome work here, your Blog post was very attractive and informative. Your post clearly communicates the findings.!! 😊
- Checkout these tips for writing blogs
- Nice job with the title and image. I definitely would click! 👏

Student must have a Github repository of their project. The repository must have a README.md file that communicates the libraries used, the motivation for the project, the files in the repository with a small description of each, a summary of the results of the analysis, and necessary acknowledgements. If the student submits a web app rather than a blog post, then the Project Definition, Analysis, and Conclusion should be included in the README file, or in their Jupyter Notebook. Students should not use another student's code to complete the project, but they may use other references on the web including StackOverflow and Kaggle to complete the project.

- Your README looks great! Make sure you communicate the installation, project motivation, file descriptions and results.
- You could improve the readme by providing more details about each file and the technical pieces on the README. You might also show more of the results and discuss the motivation in more detail and using visuals.

Code is formatted neatly with comments and uses DRY principles. A README file is provided that provides. PEP8 is used as a guideline for best coding practices.

Best practices from software engineering and communication lessons are used to create a phenomenal end product that students can be proud to showcase!

- You have used markdown cells nicely where required 👏
- Check out this link for more info on DRY principals
- Here is the link for reference of the pep 8 guidelines: https://www.python.org/dev/peps/pep-0008/, https://www.youtube.com/watch?v=Sm0wwmEwqpI

DOWNLOAD PROJECT

RETURN TO PATH