

Digified Junior Machine Learning Task

Overview

The task is divided into multiple stages (Data generation - Core model - Interface - containerization).
The core stage of the task pipeline is name verification model with confidence threshold of how much the name is real.

The task pipeline goes as follow:

1- Data generation:

- Use the link below to generate real and fake names or you can use any other dataset.
[Names Dataset Link](#)
- The real data should contain 3 names separated with space, It also should consider the basic structure of the full name combination.
- The fake data may be any garbage data or real names with some of the wrong characters.

2- Core model:

- Using tensorflow v2 in python, develop a model that takes a text of three words and verify if it is correct name with a confidence threshold of your prediction.
- You may build on an open-source model or create your own, but the delivery should be in TF 2.
- Save the trained model or weights.
- For example:
"باسم وحيد السيد" is a real name with high confidence.
"باسمم وحتد السد" is a real name with low confidence.
How would you choose a threshold to classify as a "correct" or "incorrect" name?

3- Interface

- Define a simple HTTP API request for using your model.
- Create a web server to receive the request, and execute it using the saved model. This also needs to be in python.

4- Deployment:

- Containerize your final app into a docker image that can be tested. The build must be done using docker-compose.

Notes :

- Don't hesitate to skip and get back to any of the task stages.
- Your solution will be reviewed even if it didn't run correctly.
- You are encouraged to use any available open-source resources

Deliverables :

1. Buildable docker image which should include:
 - a. Model architecture code
 - b. Server app code
 - c. docker-compose.yml and any Dockerfiles
 - d. Dependencies and a way to install them within image
2. Test script: This should be using the external server, so it sends HTTP requests and returns results, including execution time.
3. Datasets and source links

4. README that clarifies:
 - a. How to build the docker image
 - b. How to test your solution
 - c. How to verify the result you claim