

Settyl Data Science And Machine Learning Engineer Task

Task -

Dataset:

<https://gist.github.com/farhaan-settyl/ecf9c1e7ab7374f18e4400b7a3d2a161>

Data Preprocessing: Preprocess the provided dataset to clean and format the external status descriptions and internal status labels for training the machine learning model.

Model Development: Develop a machine learning model using TensorFlow or similar frameworks to predict the internal status based on the external status descriptions. Experiment with different architectures and algorithms to achieve optimal performance.

Model Training and Evaluation: Train the developed model on the preprocessed dataset and evaluate its performance using appropriate metrics such as accuracy, precision, and recall.

API Development: Implement an API using FastAPI framework to expose the trained machine learning model. The API should accept external status descriptions as input and return the predicted internal status labels.

Testing and Validation: Test the developed API thoroughly to ensure its functionality and accuracy. Validate the predictions against a validation dataset to measure the model's generalization ability.

Documentation: Document the entire development process, including data preprocessing steps, model architecture, training procedure, API implementation, and testing results.

Submission:

- Public Github Repo

Bonus:

- Deployed API URL