

Skill Task - Machine Learning Engineer

Anda diberikan 2 tasks yang perlu dikerjakan sebagai bagian proses ML engineer.

1. Diberikan notebook berikut

<https://colab.research.google.com/drive/1XSnlhSlsw3LMc6mkXlw66sMc7ZwFAJKT>,

Anda diminta untuk mendeploy model yang **ANN** ke production.

- a. Buat model tersebut menjadi sesuatu yang mudah dideploy. Beberapa contoh: diserialisasi sebagai pickle atau joblib, dideploy menjadi model MLflow, dll.
- b. Deploy model tersebut sebagai scheduled job yang jalan setiap hari pada jam 10.00 WIB. Job ini membaca data dari table dan mengeluarkan hasil klasifikasinya ke sebuah table lain.
 - i. Format table input dan output bebas. Salah satu contoh:

```
table_input
```

```
+-----+-----+-----+-----+
| id | sepal_length | sepal_width | petal_length | petal_width |
+-----+-----+-----+-----+
| asdf | 6.0 | 3.0 | 4.0 | 1.5 |
| 1234 | 6.1 | 3.2 | 4.3 | 1.1 |
+-----+-----+-----+-----+
```

```
table_output
```

```
+-----+-----+-----+
|          executed_at          | id | class |
+-----+-----+-----+
| 2022-08-01 10:00:00 +07:00 | asdf | 0 |
| 2022-08-01 10:00:00 +07:00 | 1234 | 2 |
| 2022-08-02 10:00:00 +07:00 | asdf | 0 |
| 2022-08-02 10:00:00 +07:00 | 1234 | 1 |
+-----+-----+-----+
```

- ii. Database yang digunakan bebas. Beberapa contoh: mysql, postgres, sqlite, bigquery.
 - iii. Boleh diasumsikan table input tidak pernah bertambah/berubah datanya.
 - iv. Job scheduler yang digunakan bebas. Beberapa contoh: airflow, luigi, cron job, custom python.
- c. Deploy model tersebut sebagai API yang bisa digunakan untuk klasifikasi. API ini menghasilkan class hasil klasifikasi dari data input.

I. Format payload request/response bebas. Salah satu contoh:

Request:

```
[
  {
    "sepal_length": 6.0,
    "sepal_width": 3.0,
    "petal_length": 4.0,
    "petal_width": 1.5,
  },
  {
    "sepal_length": 6.1,
    "sepal_width": 3.2,
    "petal_length": 4.3,
    "petal_width": 1.1,
  },
  {
    "sepal_length": 5.7,
    "sepal_width": 2.9,
    "petal_length": 4.1,
    "petal_width": 1.6,
  }
]
```

Response:

```
[0, 2, 1]
```

- ii. Implementasi API bebas. Beberapa contoh: flask, FastAPI, MLflow.
- d. Buat ketiga task tersebut dalam satu repo, berbeda folder, dan kodenya dipush ke Github.
2. Propose sebuah arsitektur sistem ML engineering dan gambarkan dalam sebuah diagram. Dalam arsitektur ini harus lengkap mengandung tools/platform/infra di sisi development, testing, deployment/versioning, dan monitoring output prediksi terhadap aktual, serta opsional untuk mekanisme retraining. Jika diagram saja tidak cukup menjelaskan, disarankan untuk menjelaskan dalam bentuk tulisan juga.

- a. Buat dalam bentuk dokumen docs, pdf, ppt, atau sejenisnya.

Waktu pengerjaan 4 hari dari email dikirimkan dan dikonfirmasi balik oleh Anda dengan mengirimkan ke email naflah.adela@efishery.com, dimas.gs@efishery.com, agung.fajar@efishery.com. Saat pengerjaan sudah selesai, balas kembali email ini dengan menyertakan link ke Github untuk soal 1 dan dokumen untuk soal 2. Bila ada pertanyaan atau diskusi silakan tanyakan dengan membalas email ini.