

2023

# ML-Layer Desgin

AI EV-BATTERIJEN  
TADRAŁA, PIOTR P.P.

O-PP-CMK

## INHOUDSOPGAVE

1.0 Inleiding .....	2
2.0 Systeem.....	2
2.1 Context.....	2
2.2 mappen Structuur.....	3
3.0 Architectuur .....	3
3.1 Flask API .....	3
3.2 Model.....	4
4.0 Functionele Details .....	4
4.1 Predict.....	4
4.2 Retrain Endpoint.....	5
4.3 Anomaly Endpoint .....	5
5.0 Dataopslag .....	6
6.0 CI/CD & Deployment .....	6

## 1.0 INLEIDING

Dit document beschrijft het ontwerp van een Machine Learning API voor anomaly detection voor INNER, gericht op het herkennen van defecte EV-batterijen. Het systeem maakt het mogelijk om afbeeldingen in te posten en deze te classificeren als een anomaly of niet. Daarnaast kunnen gelabelde afbeeldingen worden gepost om het model te verbeteren met normale data, of om de reconstruction threshold opnieuw te evalueren. Het model maakt gebruik van een Convolutional Neural Network (CNN) voor anomaly detection op basis van reconstructiefout.

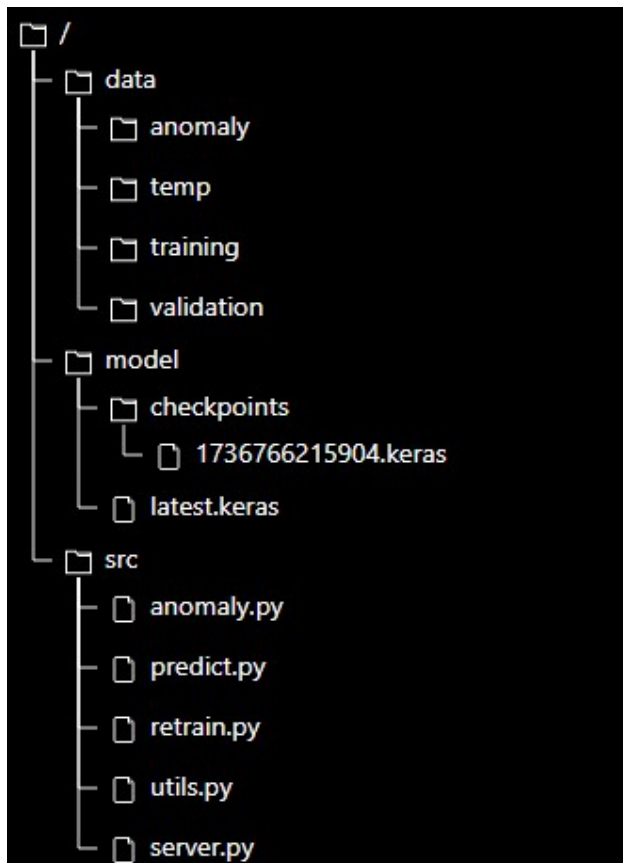
## 2.0 SYSTEEM

### 2.1 CONTEXT

Het systeem bestaat uit een Flask based API met de volgende functionaliteiten

- **Predict:** Berekent de reconstruction error en vergelijkt deze met eerder geëvalueerde threshold om te bepalen of de afbeelding een anomaly is.
- **Retrain:** Traint het model met de geleverde afbeelding om de nauwkeurigheid van het model te verbeteren.
- **Anomaly evaluation:** Evalueert de nieuwe threshold op basis van de bestaande data en de geposte anomaly afbeelding

## 2.2 MAPPEN STRUCTUUR



## 3.0 ARCHITECTUUR

### 3.1 FLASK API

De API is gebouwd met Flask en heeft drie endpoints:

Endpoint	Omschrijving
/predict	Accepteert afbeelding url en predict of de afbeelding een anomaly is.
/retrain	Accepteer afbeelding url en traint het model opnieuw.
/anomaly	Accepteert afbeelding url en evalueert de reconstruction error opnieuw.

## 3.2 MODEL

Het model is een TensorFlow CNN-model, gefinetuned op basis van bevindingen uit een eerder onderzoek "ML Model Onderzoek". Het model maakt gebruik van de reconstruction error (het verschil tussen de oorspronkelijke en de down/upscaled afbeelding) om anomaly te herkennen. Het model wordt geëvalueerd met behulp van validatie- en anomalydatasets.

## 4.0 FUNCTIONELE DETAILS

### 4.1 PREDICT

#### INPUT

Parameter	Type	Omschrijving
Image_url	String	URL van de afbeelding die je wilt predicten

#### PROCES

1. De afbeelding wordt opgehaald van de URL en gepreprocessed.
2. De afbeelding wordt door het model geanalyseerd om de reconstruction error te berekenen
3. Reconstruction error wordt vergeleken met een eerder gedefinieerde threshold.

#### OUTPUT

Parameter	Type	Omschrijving
Is_anomaly	Boolean	Geeft aan of de afbeelding een anomaly is.
Reconstruction_error	Number	Berekende reconstruction_error

## 4.2 RETRAIN ENDPOINT

### INPUT

Parameter	Type	Omschrijving
Image_url	String	URL van de afbeelding voor de retraining

### PROCES

1. De afbeelding wordt opgehaald en opgeslagen in zowel de trainings en tijdelijke mappen
2. Het model wordt opnieuw getraind met de nieuwe afbeeldingsdata en de prestaties worden geëvalueerd op de validatie en anomalydatasets

### OUTPUT

Parameter	Type	Omschrijving
new_validation_error	Number	Nieuwe reconstruction error van normal data
new_anomaly_error	Number	Nieuwe reconstruction error van anomaly data

## 4.3 ANOMALY ENDPOINT

### INPUT

Parameter	Type	Omschrijving
Image_url	String	URL van de afbeelding voor de reconstruction error evaluatie

### PROCES

1. De afbeelding wordt opgehaald en opgeslagen in de anomaly map
2. De prestaties van het model worden geëvalueerd op zowel de validatie als anomaly datasets

### OUTPUT

Parameter	Type	Omschrijving
new_validation_error	Number	Nieuwe reconstruction error van normal data
new_anomaly_error	Number	Nieuwe reconstruction error van anomaly data

## 5.0 DATAOPSLAG

Afbeeldingen worden lokaal opgeslagen in specifieke mappen

- **Training:** Afbeeldingen waar op het huidige model getraind is.
- **Validatie:** Afbeeldingen voor evaluatie van de normal reconstruction error.
- **Anomaly:** Afbeeldingen voor de evaluatie van de anomaly reconstruction error.
- **Temp:** Tijdelijke opslag voor afbeeldingen voor de retraining.

## 6.0 CI/CD & DEPLOYMENT

De ML Layer maakt deel uit van de INNER-monorepo, samen met de backend en de frontend. Deze is in GitHub geconfigureerd om bij elke update de Docker builds te runnen, indien alles in orde is, de Docker images naar de Azure Image Registry te pushen. Daar wordt de image vervolgens gebruikt door een container-webapp. Voor meer informatie, zie "Azure Deployment".