COREALIS
THE PORT OF THE FUTURE

COREALIS Piraeus Demo/Training Webinar

COREALIS Asset Management Predictor

Anil Goyal-NEC

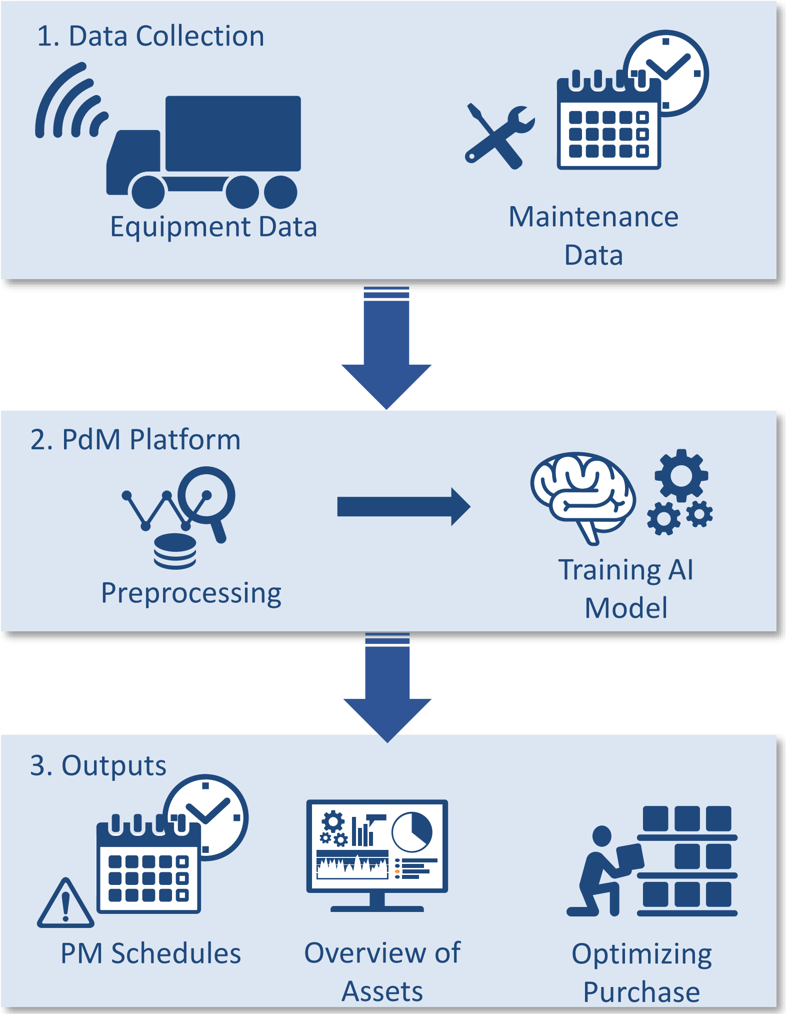
June 4th, 2020

Predictive Maintenance

- **Preventive vs Predictive Maintenance**:
  - **Preventive Maintenance**: Engineers inspect vehicles every couple of months. This method is costly and does not capture asset-specific conditions (for e.g. rides per day, occupancy level, weather conditions, etc.).
  - **Predictive Maintenance (PdM)** solutions takes into account the operating conditions of vehicles by leveraging data collected from individual assets to predict failures in future.

- **Overview:**
  - **Data Collection**: Equipment Data and Maintenance Data
  - **PdM Platform**: Preprocessing and Training AI Model
  - **Outputs**: Schedules, Overview of Assets and Purchase Optimization

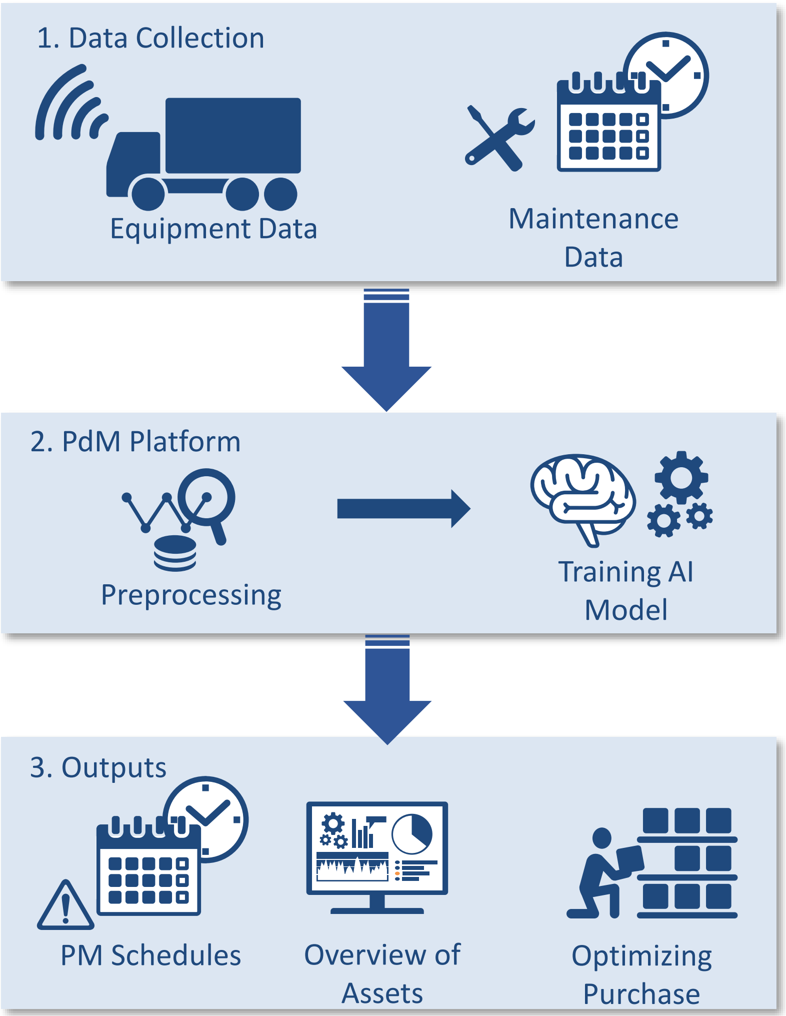
- **Customer Value:**
  - Reduced Asset Downtime,
  - Reduced maintenance and repair costs,
  - Optimize spare parts inventory



Market Analysis

- **Market Size (General)**:
  - Analysts consensus on growth rate at 25-30% **CAGR** in 2020-2025 period (Allied Market Research, Market Expert 24, ReportBuyer), but have diverse market size estimates ($15-30B)

- **PdM Solutions:**
  - Predictive Maintenance is an imbalanced classification problem where number of part failures is much smaller than predominant normal working parts.
  - To the best of our knowledge, PdM solutions (from Bosch, SAP, Huwaei, IBM and SAP) relies on statistical analysis.
  - Their solutions does not relies on advanced machine learning algorithms tailored for imbalanced classification tasks.



Technology Description (BRIGHT DAMVI)

IJCNN 2020

- **BRIGHT** is NEC's Ensemble Learning Technology:
  - Combine many weak learners to build a more efficient learner
  - Many specialized weak learners can beat a strong one
  - That's because "strong" learners (singletons) model the variable in a wide set of conditions, sacrificing memory of rare events to work in general

- **Meta-Learning**
  - learning how learners perform in which situation
  - A Meta-Learner learns from the behavior of other learners (= base learners)

- **BRIGHT DAMVI***: Meta-learning for imbalanced data classification
  - Learns a diversity-aware weighted majority vote classifier over the base classifiers
  - Delivers benefits in terms of F1-score and Average Precision especially in four main applications: predictive maintenance tasks, credit fraud detection, medical diagnosis and webpage classification.
  - Performs well even for highly imbalanced classification tasks (< 4% positive class examples).

*Diversity-Aware Weighted Majority Vote Classifier for Imbalanced Data

*Anil Goyal and Jihed Khiari, "Diversity-Aware Weighted Majority Vote Classifier for Imbalanced Data", The International Joint Conference on Neural Networks (IJCNN), 2020

Technology Description (BRIGHT DAMVI)

IJCNN 2020

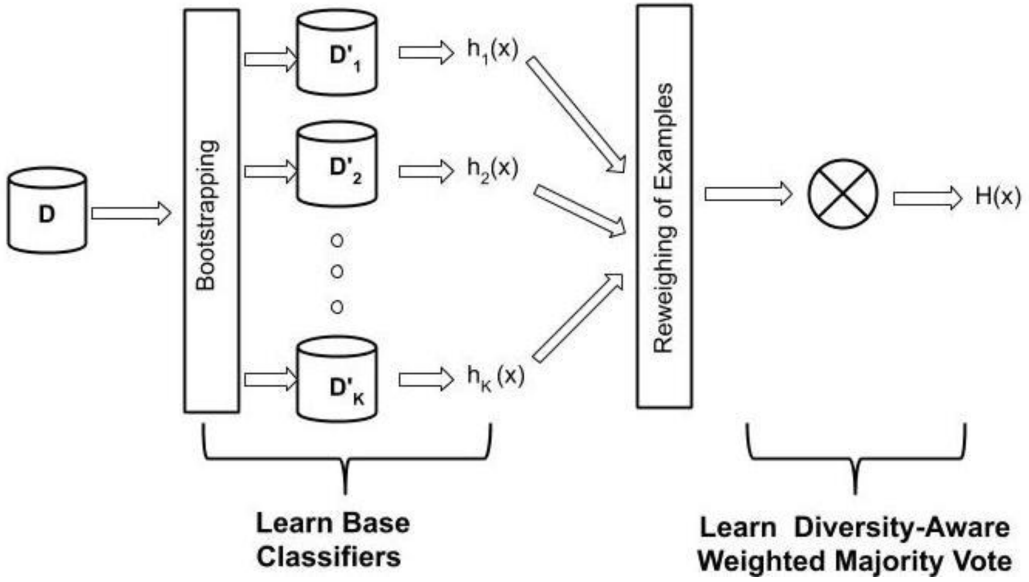
- **Base Classifiers:** Learn a set of base classifiers corresponding to K bootstrapped samples

- **Reweighing of Examples:** DAMVI increases the weights of positive examples (minority class) which are "hard" to classify with uniformly weighing of classifiers.

- **Learning Diversity-Aware Weighted Majority Vote:** Learned base classifiers are combined using diversity-aware weighted majority vote:

$$H(x) = \sum_{k=1}^{K} Q(k).h_k(x)$$

where, $h_K(x)$ is a base learner corresponding to subsampled data D_K and $Q(k)$ is the weight over base learner.



Experimental Results

COREALIS
THE PORT OF THE FUTURE

IJCNN 2020

Datasets

	#Attributes	#Examples	IR
Webpage	300	34780	3.03
Mammography	6	11183	2.32
Scania	170	60000	1.67
Protein Homo	74	145751	0.9
Credit Fraud	30	284807	0.17
PCT Data	17	816099	0.02

F1-score

	Webpage	Mammography	Scania	Protein Homo	Credit Fraud	PCT Data
S-DT	$.5062\pm.028^{\downarrow}$	$.5198\pm.009^{\downarrow}$	$.5848\pm.011^{\downarrow}$	$.5278\pm.005^{\downarrow}$	$.5610\pm.016^{\downarrow}$	$.8793\pm.011^{\downarrow}$
R-DT	$.4705\pm.021^{\downarrow}$	$.6053\pm.043^{\downarrow}$	$.6256\pm.019^{\downarrow}$	$.7290\pm.017^{\downarrow}$	$.7556\pm.013^{\downarrow}$	$\mathit{.9715}\pm.002^{\downarrow}$
A-DT	$.4693\pm.019^{\downarrow}$	$.4978\pm.034^{\downarrow}$	$.5807\pm.020^{\downarrow}$	$.5259\pm.019^{\downarrow}$	$.5653\pm.027^{\downarrow}$	$.8830\pm.009^{\downarrow}$
R-BG	$.4620\pm.016^{\downarrow}$	$\mathit{.6145}\pm.026^{\downarrow}$	$\mathit{.6845}\pm.014^{\downarrow}$	$\mathit{.7849}\pm.021^{\downarrow}$	$\mathit{.7703}\pm.020^{\downarrow}$	$.9691\pm.001^{\downarrow}$
S-BG	$\mathit{.6134}\pm.017^{\downarrow}$	$.5391\pm.017^{\downarrow}$	$.6493\pm.009^{\downarrow}$	$.6771\pm.009^{\downarrow}$	$.6839\pm.024^{\downarrow}$	$.9430\pm.006^{\downarrow}$
A-BG	$.4804\pm.021^{\downarrow}$	$.5169\pm.011^{\downarrow}$	$.6269\pm.007^{\downarrow}$	$.6346\pm.013^{\downarrow}$	$.6819\pm.030^{\downarrow}$	$.9312\pm.004^{\downarrow}$
BB	$.3445\pm.001^{\downarrow}$	$.4465\pm.030^{\downarrow}$	$.4317\pm.005^{\downarrow}$	$.4275\pm.008^{\downarrow}$	$.1376\pm.006^{\downarrow}$	$.8014\pm.006^{\downarrow}$
BRF	$.4098\pm.010^{\downarrow}$	$.3659\pm.014^{\downarrow}$	$.3822\pm.004^{\downarrow}$	$.4027\pm.009^{\downarrow}$	$.1255\pm.016^{\downarrow}$	$.2943\pm.007^{\downarrow}$
EE	$.4678\pm.011^{\downarrow}$	$.2534\pm.002^{\downarrow}$	$.4096\pm.006^{\downarrow}$	$.3350\pm.003^{\downarrow}$	$.0922\pm.007^{\downarrow}$	$.0881\pm.001^{\downarrow}$
DAMVI	$\mathbf{.7996}\pm.011$	$\mathbf{.6661}\pm.023$	$\mathbf{.7289}\pm.011$	$\mathbf{.8067}\pm.009$	$\mathbf{.8495}\pm.019$	$\mathbf{.9816}\pm.001$

Average Precision

	Webpage	Mammography	Scania	Protein Homo	Credit Fraud	PCT Data
S-DT	$.2794\pm.023^{\downarrow}$	$.2919\pm.010^{\downarrow}$	$.3526\pm.014^{\downarrow}$	$.3153\pm.005^{\downarrow}$	$.3482\pm.016^{\downarrow}$	$.7785\pm.001^{\downarrow}$
R-DT	$.3008\pm.016^{\downarrow}$	$.3811\pm.054^{\downarrow}$	$.3994\pm.024^{\downarrow}$	$.5347\pm.025^{\downarrow}$	$.5728\pm.020^{\downarrow}$	$.9447\pm.005^{\downarrow}$
A-DT	$.2481\pm.014^{\downarrow}$	$.2740\pm.034^{\downarrow}$	$.3483\pm.023^{\downarrow}$	$.3112\pm.019^{\downarrow}$	$.3516\pm.030^{\downarrow}$	$.7851\pm.016^{\downarrow}$
R-BG	$.4944\pm.010^{\downarrow}$	$\mathit{.7011}\pm.021$	$\mathit{.8097}\pm.016^{\downarrow}$	$.8495\pm.016$	$.8120\pm.030^{\downarrow}$	$\mathit{.9875}\pm.001$
S-BG	$.6219\pm.028^{\downarrow}$	$.6971\pm.025^{\downarrow}$	$.7275\pm.019^{\downarrow}$	$.8424\pm.013$	$.8135\pm.027^{\downarrow}$	$.9863\pm.001^{\downarrow}$
A-BG	$.4400\pm.024^{\downarrow}$	$.6261\pm.036^{\downarrow}$	$.6712\pm.018^{\downarrow}$	$.8276\pm.016$	$\mathit{.8137}\pm.035^{\downarrow}$	$.9847\pm.005^{\downarrow}$
BB	$.6302\pm.034^{\downarrow}$	$.6644\pm.037^{\downarrow}$	$.6745\pm.024^{\downarrow}$	$.8359\pm.018$	$.7516\pm.048^{\downarrow}$	$.9849\pm.001^{\downarrow}$
BRF	$.6930\pm.022^{\downarrow}$	$.6782\pm.023$	$.6877\pm.016^{\downarrow}$	$\mathit{.8549}\pm.014$	$.7615\pm.047^{\downarrow}$	$.6976\pm.010^{\downarrow}$
EE	$\mathit{.6969}\pm.038^{\downarrow}$	$.5967\pm.043^{\downarrow}$	$.7558\pm.014^{\downarrow}$	$\mathbf{.8561}\pm.012$	$.7672\pm.025^{\downarrow}$	$.0790\pm.001^{\downarrow}$
DAMVI	$\mathbf{.8331}\pm.013$	$\mathbf{.7142}\pm.039$	$\mathbf{.8335}\pm.007$	$.8267\pm.013$	$\mathbf{.8373}\pm.027$	$\mathbf{.9976}\pm.001$

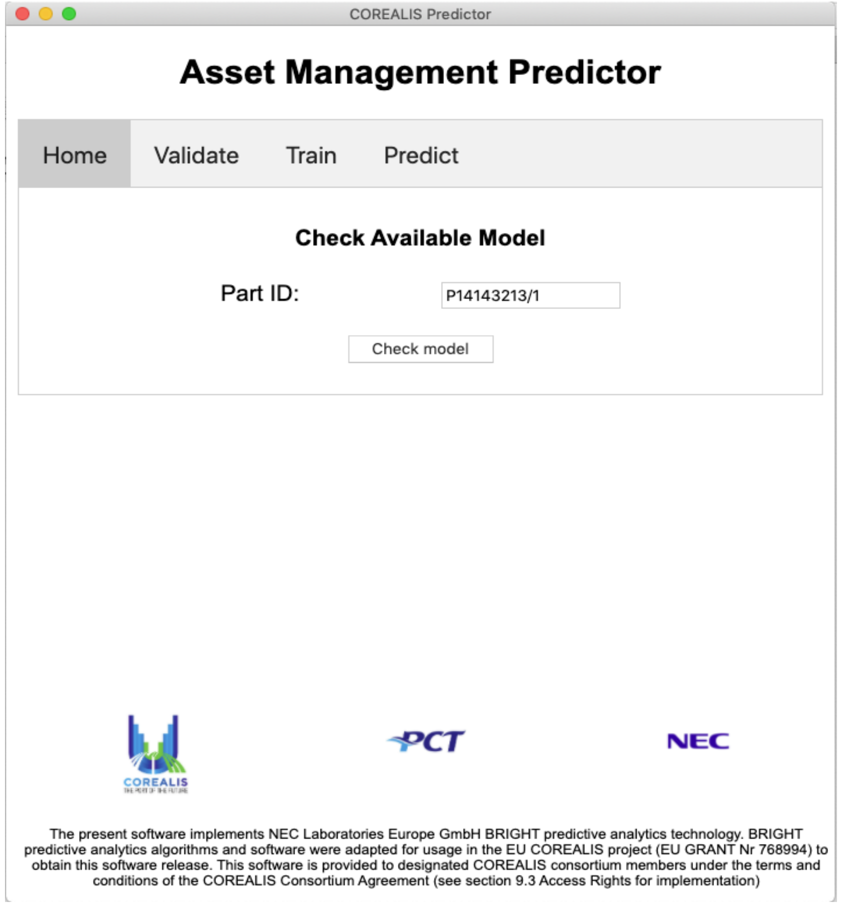
- DAMVI performs best on all datasets in terms of F1-score and for 5 out of 6 datasets in terms of Average Precision.
- DAMVI performs significantly better on PCT dataset with lowest imbalance ratio of 0.02.

Software Development

Web Interface

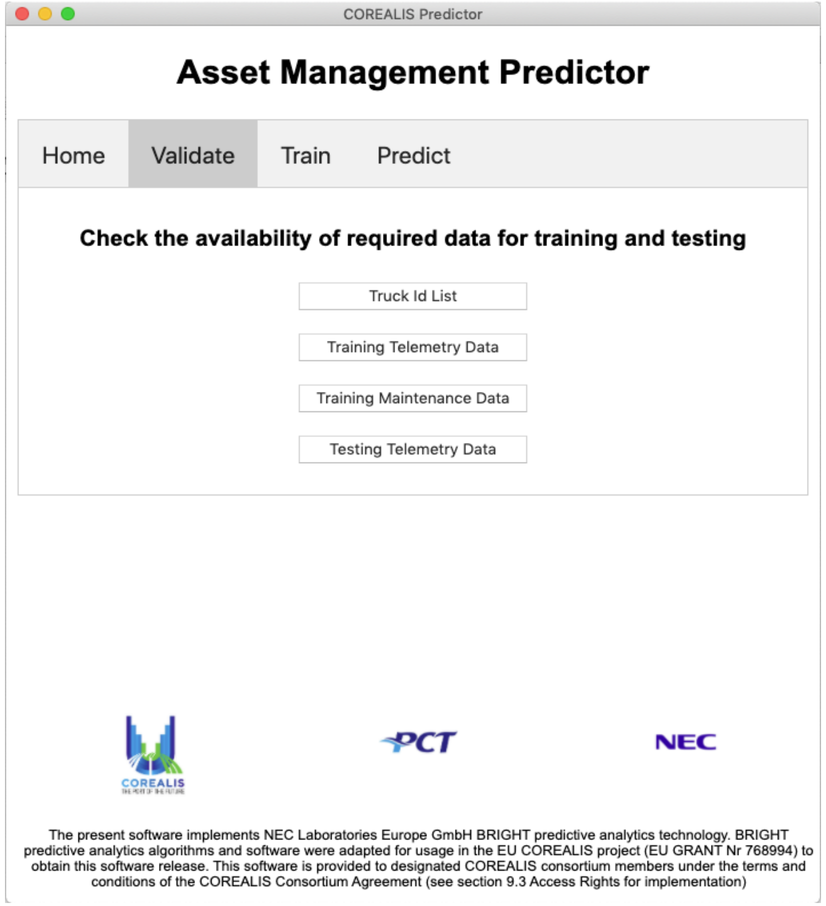
Step 1:

- User can check if there exists any learned model for a particular part (from past).



Step 2:

- User can validate if the required data for training and testing are stored at desired location or not.



Software Development

COREALIS
THE PORT OF THE FUTURE

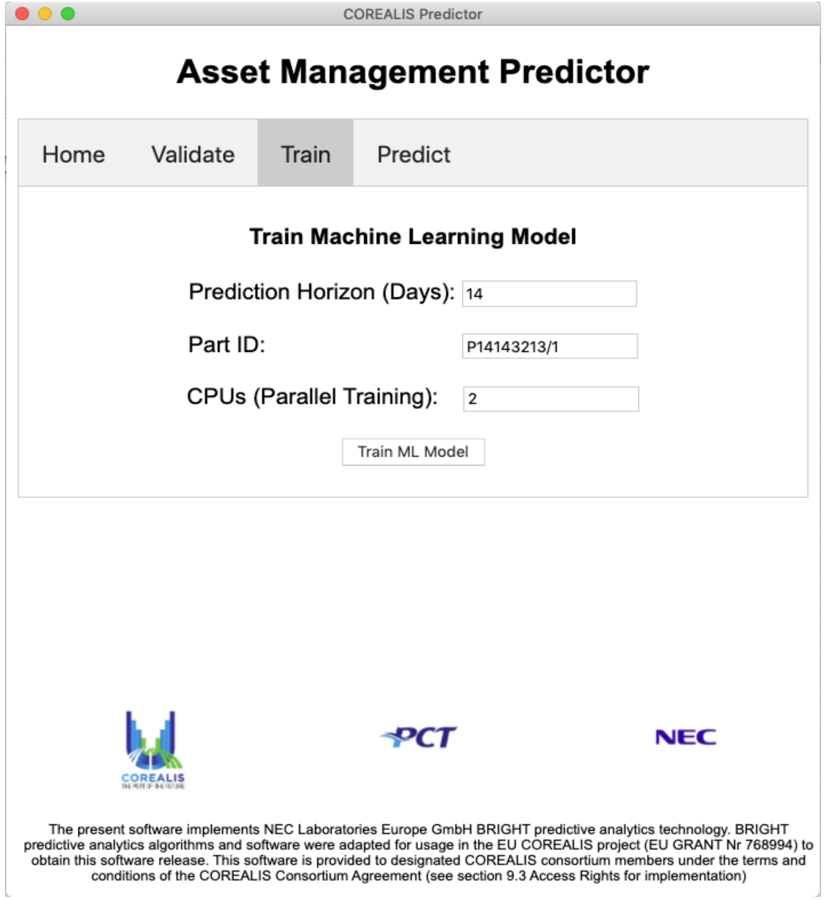
Web Interface

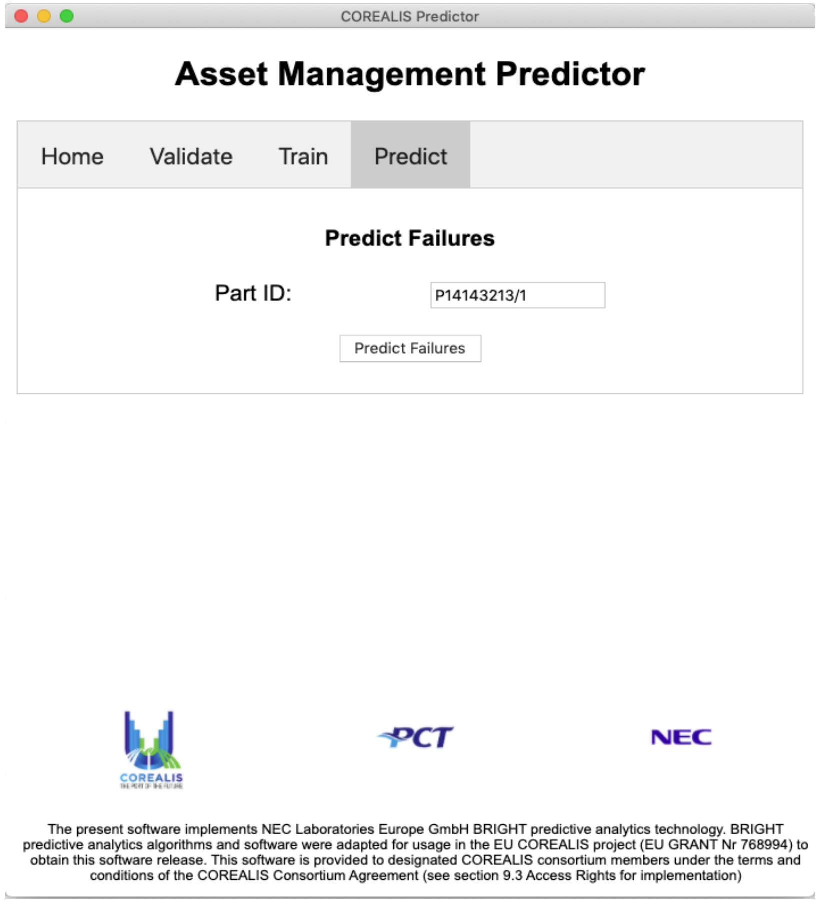
Step 3:

- Train the machine learning models for a particular part using prediction horizon and number of CPUs as input tomorrow.

Step 4:

- Failures predictions (along with probability) are saved in respective folder in csv file.





COREALIS
THE PORT OF THE FUTURE

www.corealis.eu
corealis_eu
COREALIS EU Project
Corealis_eu
info@corealis.eu

THANK YOU FOR YOUR ATTENTION

NEC

Anil Goyal

Anil.Goyal@neclab.eu

This project has received funding from the European Union's horizon 2020 research and innovation programme under grant agreement No. 768994