COREALIS
THE PORT OF THE FUTURE

COREALIS Final Conference

COREALIS Asset Management Predictor

Sebastien Nicolas - NEC

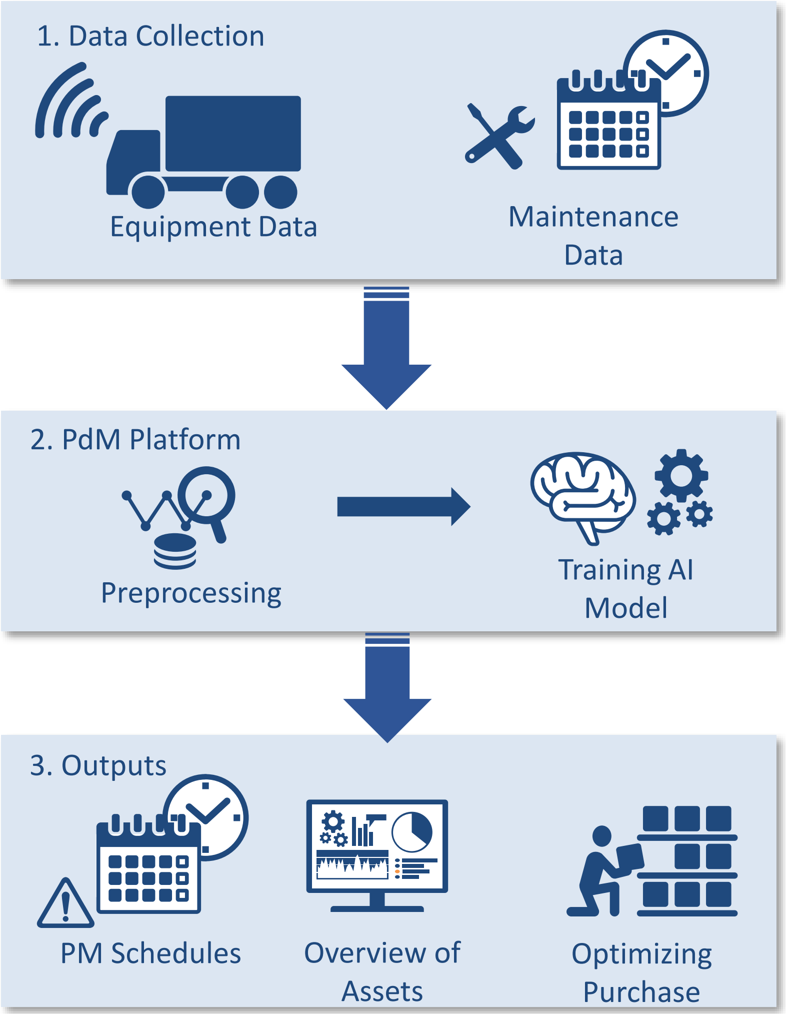
April 23rd, 2021

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

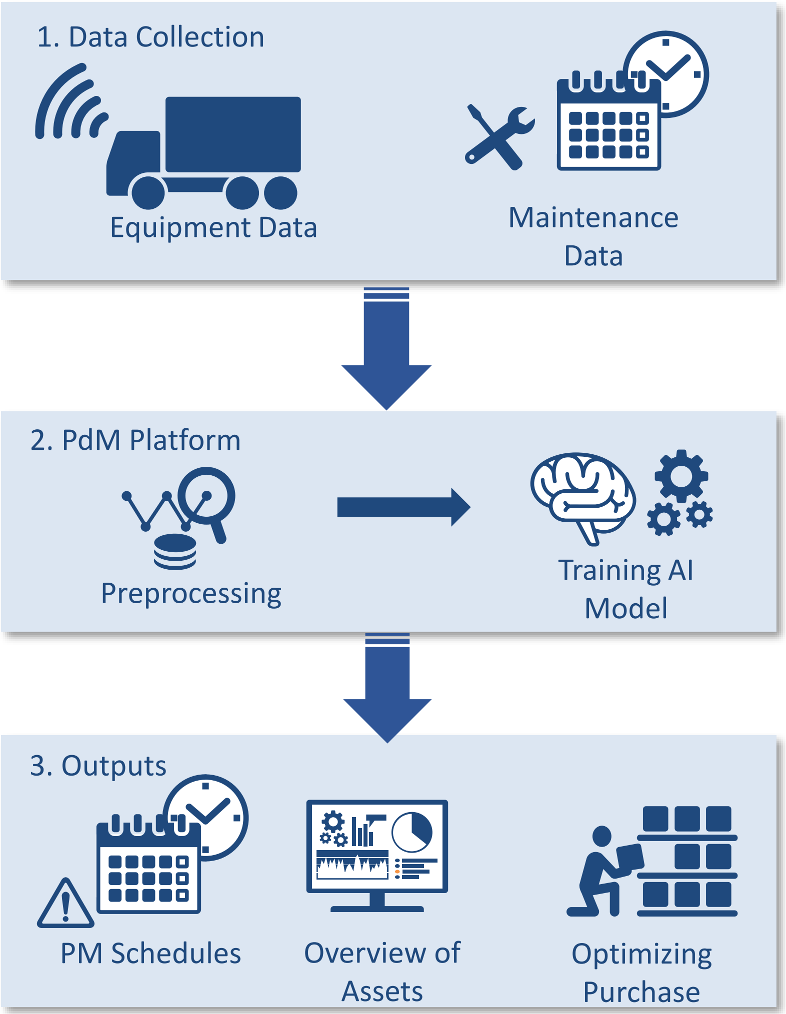
- **Customer Value:**
  - Reduced Asset Downtime,
  - Reduced maintenance and repair costs,
  - Optimize spare parts inventory



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

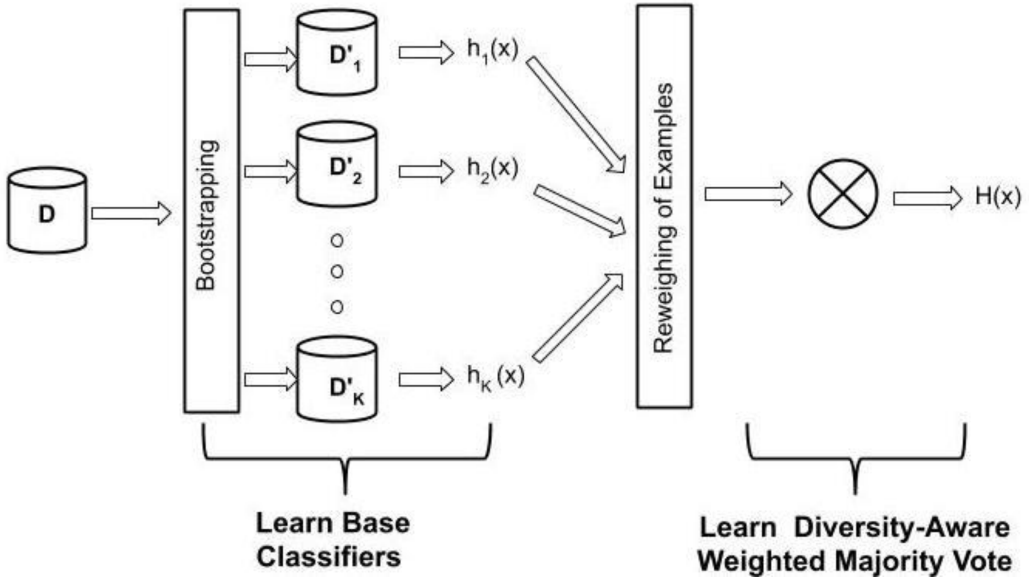
- **Base Classifiers:** Learn a set of base classifiers corresponding to K bootstrapped samples

- **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

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

- ❑ 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.

F1-score

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

Average Precision

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

Software Development

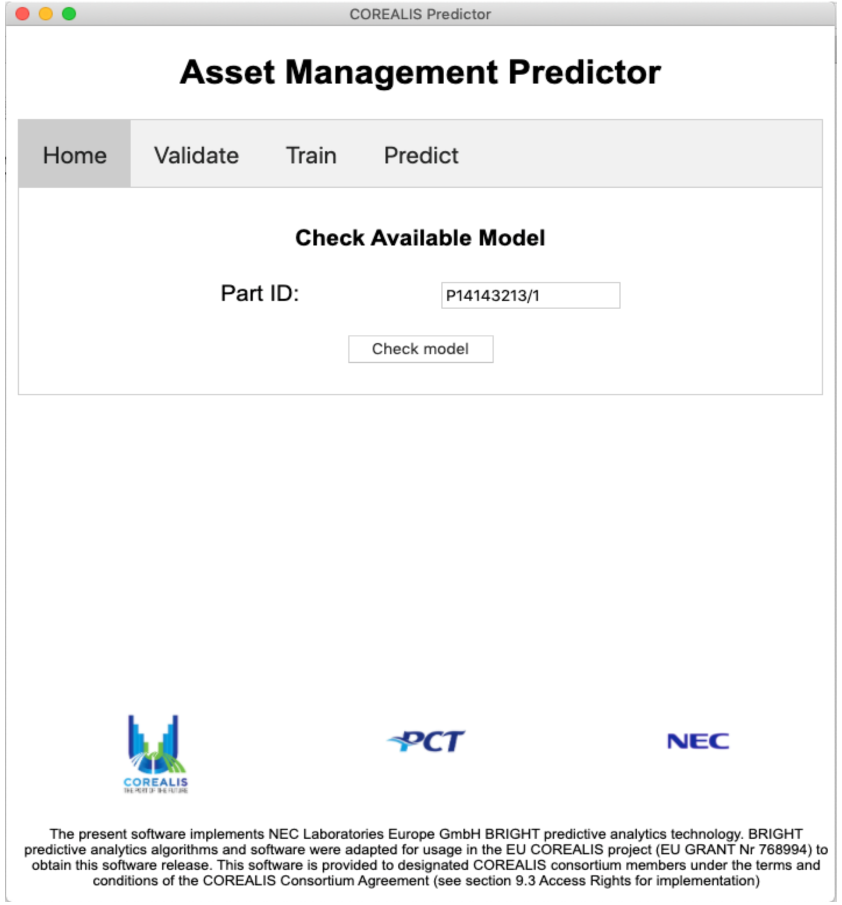
Web Interface

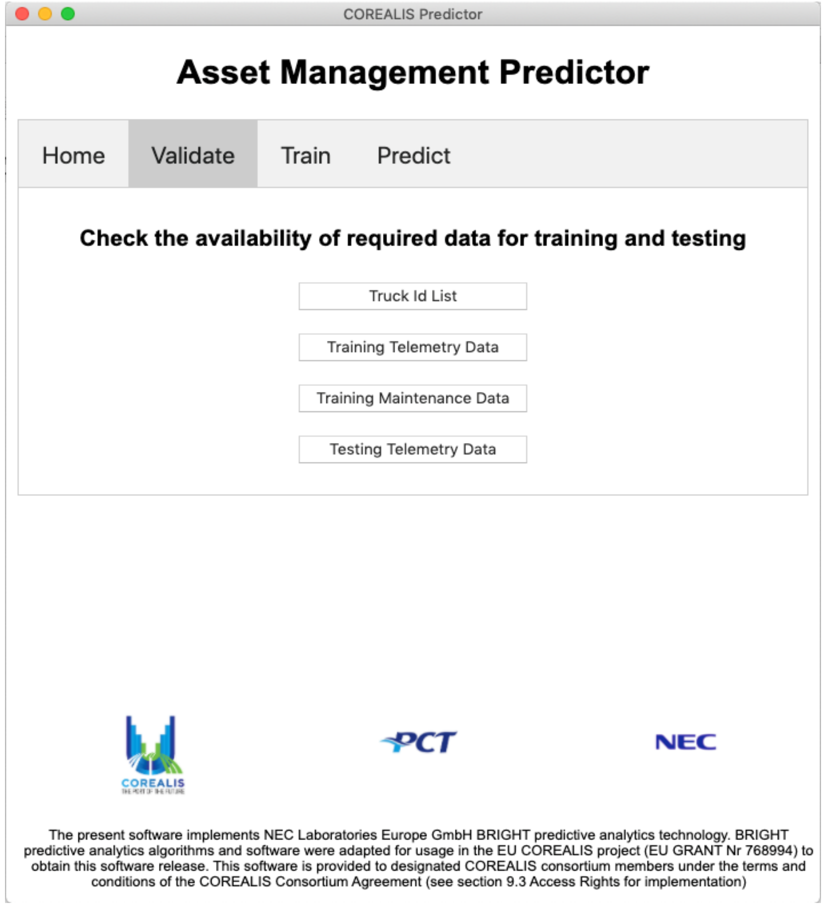
Step 1:

- 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

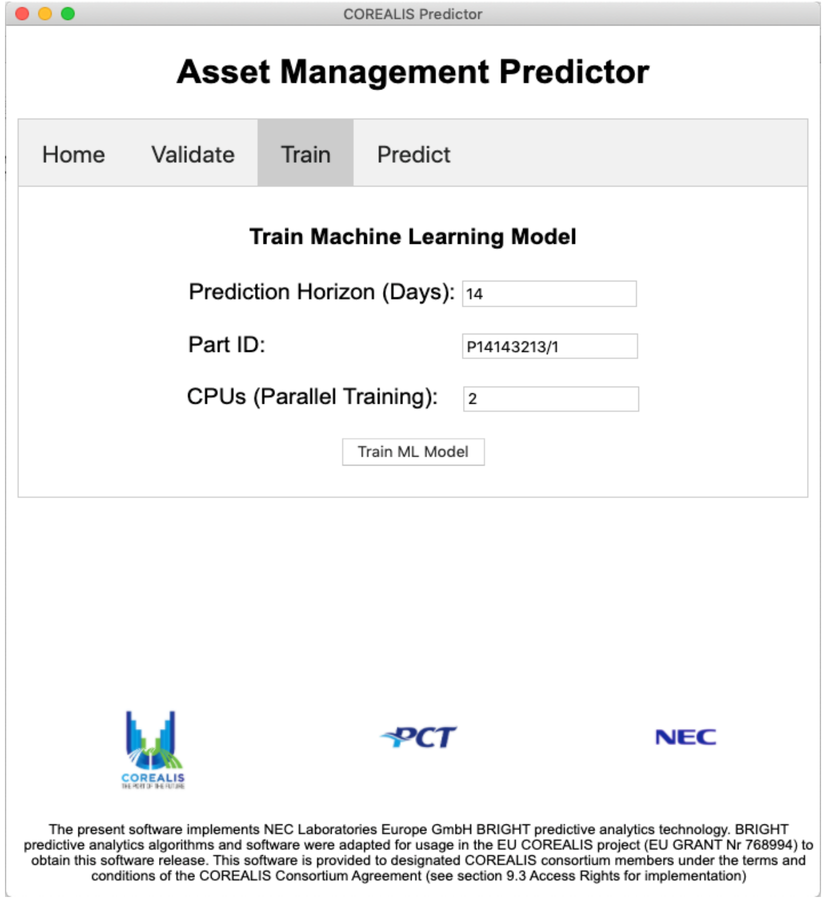
Web Interface

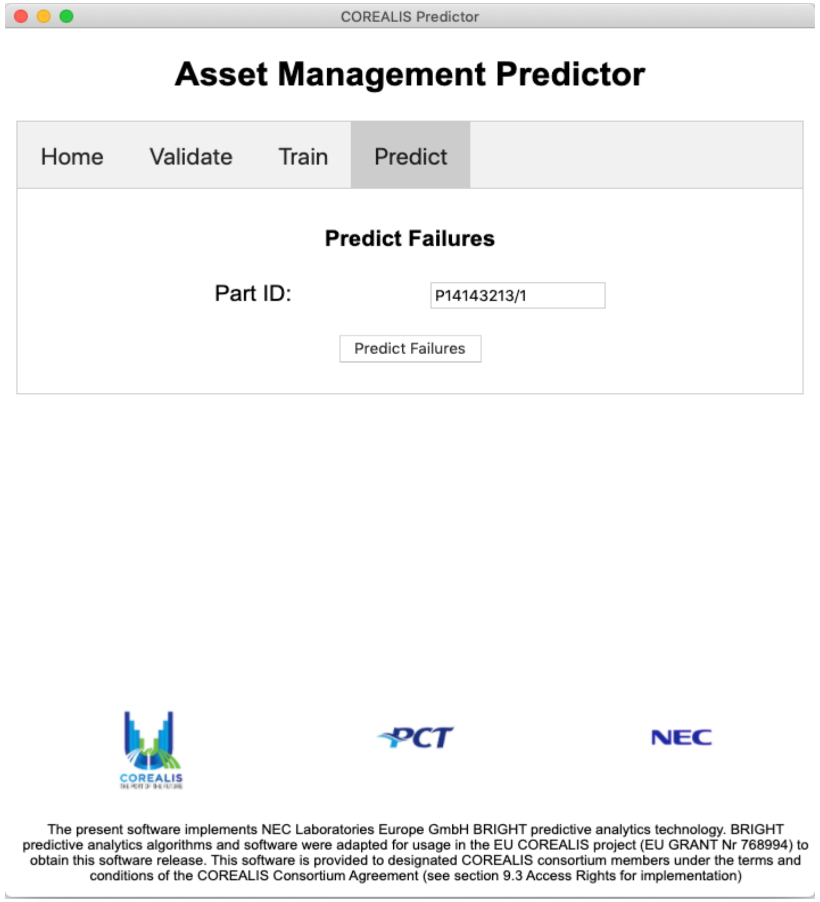
Step 3:

- 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

Sebastien Nicolas

sebastien.nicolas@neclab.eu

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