UNIVERSITY OF PATRAS

SCHOOL OF ENGINEEIRING

DEPT. OF MECHANICAL ENGINEERING & AERONAUTICS

FACULTY OF INDUSTRIAL MANAGEMENT

DIPLOMA THESIS

Development of industrial application using AI technologies

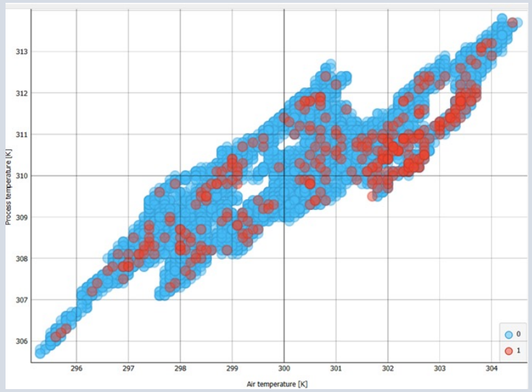
STAMATIOS PASPALIARIS, ID: 1048418

Supervisor: NIKOS LARAKAPILIDIS,Professor

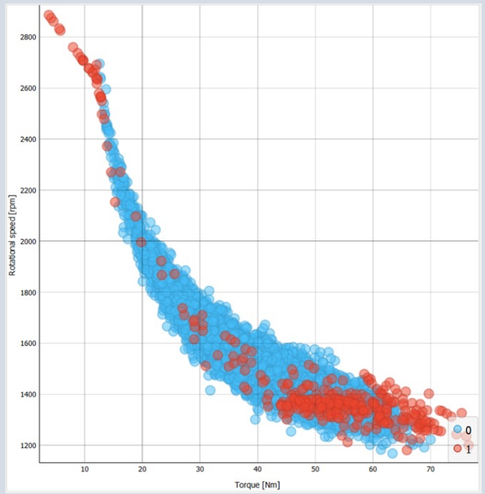
July 2023

Abstract: The aim of this thesis is to develop an AI system to support predictive maintenance decisions in industry. At the core of the implementation of these systems are Machine Learning algorithms. Based on this assumption, at a first level the paper focuses on presenting the implementation techniques, critical parameters and scopes of three main supervised Machine Learning algorithms. In particular, the algorithms chosen are decision trees, the Naïve Bayes classifier and the nearest neighbours (k-NN) algorithm. In addition, emphasis is also placed on the pre-processing stage of the data, which is considered particularly important for the efficient implementation of the algorithms, and the metrics for evaluating the effectiveness of the algorithms are listed. At a second level, the set of techniques is applied to real industry data. The results obtained from the implementation of different combinations, showed that the k - NN algorithm, is the most efficient over the others, achieving efficiency values ranging from 87% to 90%, depending on the choice of implementation parameters. Satisfactory results were also obtained by the Naïve Bayes classifier, while decision trees had significantly lower efficiency in the set of combinations considered.

Variables reduction using regression techniques



$$\Theta\Lambda^{new} = 115{,}04 + 0{,}65 * \Theta A$$



$$Ροπή^{new} = 114{,}86 - 0.05 * TΠ$$

Evaluation results according to specific metrics

Evaluation results according to Target variable

	AUC	CA	F1	Precision	Recall
70-30	Random sampling 50 repeats				
k-NN	89,0%	96,6%	94,9%	93,3%	96,6%
Decision Tree	58,4%	95,4%	95,4%	95,5%	95,4%
Naïve Bayes	85,0%	96,6%	94,9%	93,3%	96,6%
80-20	Random sampling 50 repeats				
k-NN	88,9%	96,6%	94,9%	93,3%	96,6%
Decision Tree	57,9%	95,4%	95,4%	95,5%	95,4%
Naïve Bayes	84,8%	96,6%	94,9%	93,3%	96,6%
90-10	Random sampling 50 repeats				
k-NN	89,0%	96,6%	94,9%	93,3%	96,6%
Decision Tree	59,5%	95,4%	95,5%	95,6%	95,4%
Naïve Bayes	84,8%	96,6%	94,9%	93,3%	96,6%

Evaluation results according to Failure Type varia

	AUC	CA	F1	Precision	Recall
70-30	Random sampling 50 repeats				
k-NN	87,7%	96,5%	94,8%	93,2%	96,5%
Decision Tree	61,5%	95,0%	95,1%	95,2%	95,0%
Naïve Bayes	86,0%	96,5%	94,8%	93,2%	96,5%
80-20	Random sampling 50 repeats				
k-NN	87,9%	96,5%	94,8%	93,1%	96,5%
Decision Tree	60,9%	95,1%	95,1%	95,1%	95,1%
Naïve Bayes	85,5%	96,5%	94,8%	93,1%	96,5%
90-10	Random sampling 50 repeats				
k-NN	88,1%	96,5%	94,8%	93,1%	96,5%
Decision Tree	60,1%	94,8%	94,9%	95,0%	94,8%
Naïve Bayes	85,7%	96,5%	94,8%	93,1%	96,5%