PO.066 Study of Feature-Selection-Algorithms with powerful 3D Insights for Wind Turbine Failure Prediction using SCADA Data

Alejandro Blanco[1], Jordi Solé-Casals[2], Pere Marti-Puig[2] , Juan José Cárdenas[1], Isaac Justicia[1], Jordi Cusidó[1]

[1]ITESTIT S.L., Terrassa, Spain - [2]Universitat de Vic, Vic, Spain

SMARTIVE

Abstract

One of the key steps in failure prediction using machine learning classifiers is to choose an optimal or near optimal set of inputs from tens to hundreds of variables. This task can be achieved with the implementation of unsupervised-supervised algorithms that aim to find out the most relevant and shortest set of variables related with the failure. Therefore, with the aim of study and select the best algorithm or algorithms of feature selection, we present a **thorough study of the state of the art of available techniques when applied to the specific area of wind turbine Operation & Maintenance**. In order to visualize the behavior of the selected variables we have choose sets of three variables for fault in order to make **3D plot** animations. Those provide **intuitive and powerful insights** about the behavior of the WTG until 21 days before failure. This helps us to confirm and improve the models used for failure prediction in Smartive platform.

Objectives

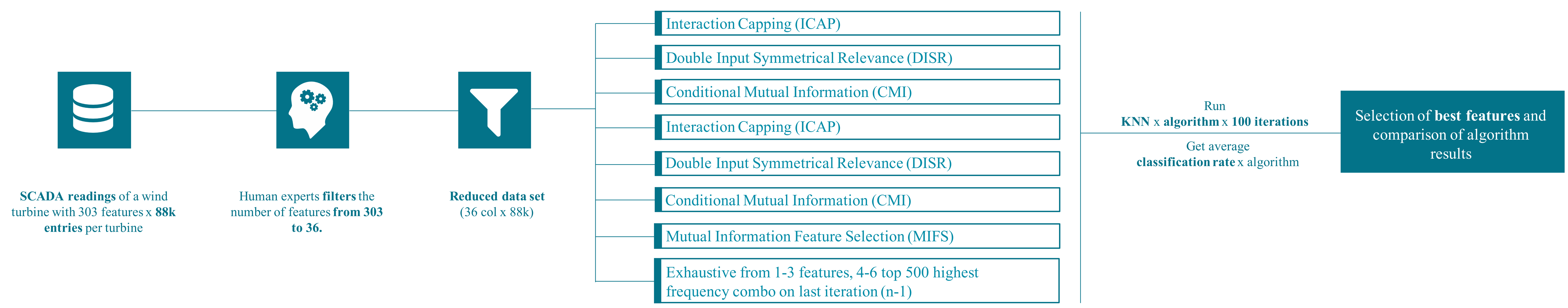
Improve efficiency of O&M tasks by mean of prognosis analysis of SCADA data in order to extract and test the found failure patterns and therefore:

- Reduce WTG down-time.
- Detect critical faults in earlier stages.
- Selection and visualization of the key indicators of failure over the time.
- Improve the prediction models accuracy by the selection.
- Reduce computation time and problem dimensionality.

Methods

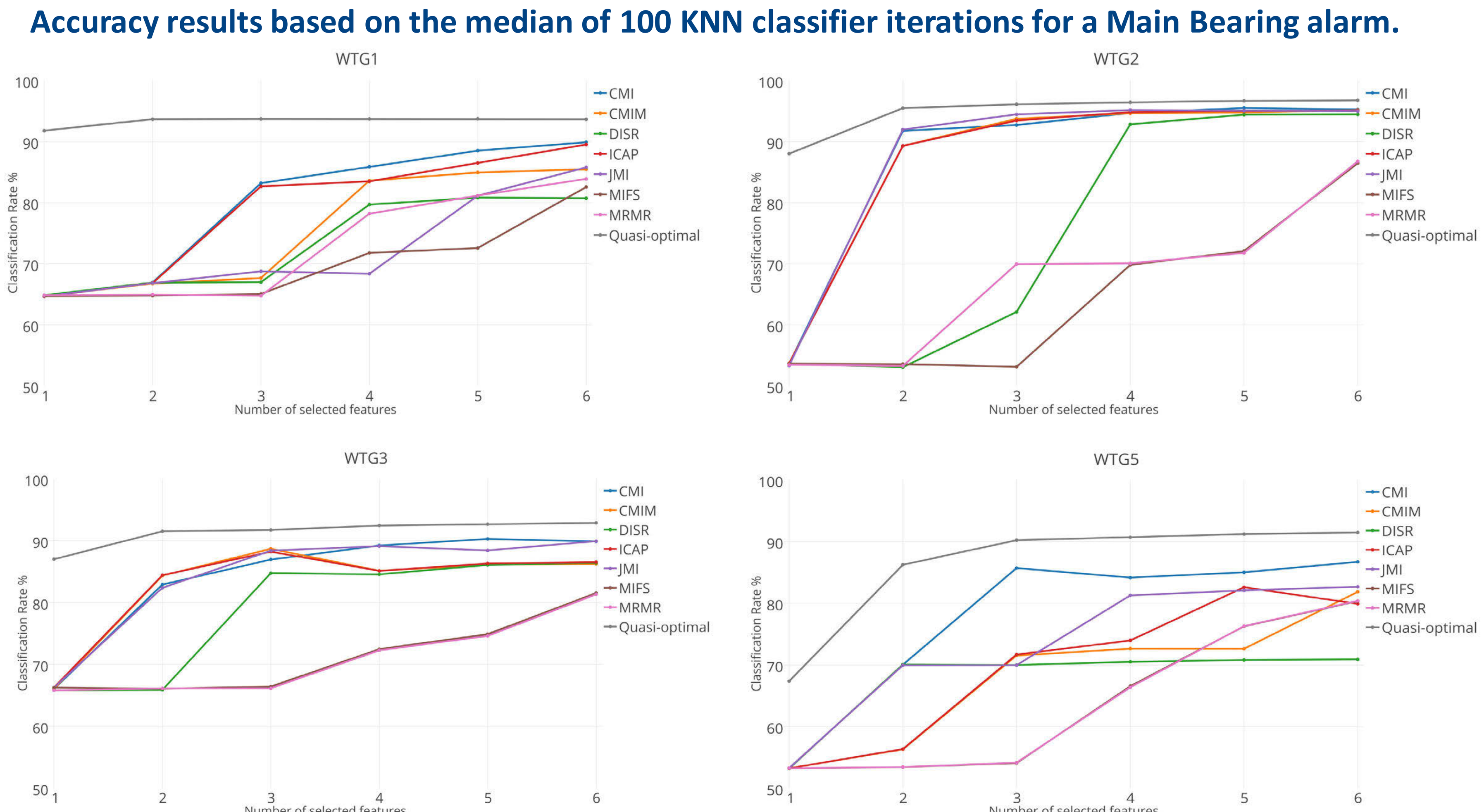
The process starts with the application of the feature selection algorithm over the SCADA data. A total of 8 algorithms were tested, among them:

Conditional mutual information[1], Double input symmetrical relevance[2], Min-redundancy Max-relevance[3], Conditional mutual information maximization[4], Joint mutual information[5], Interaction Capping[6], Mutual information feature selection[7] and quasi-optimal exhaustive method.

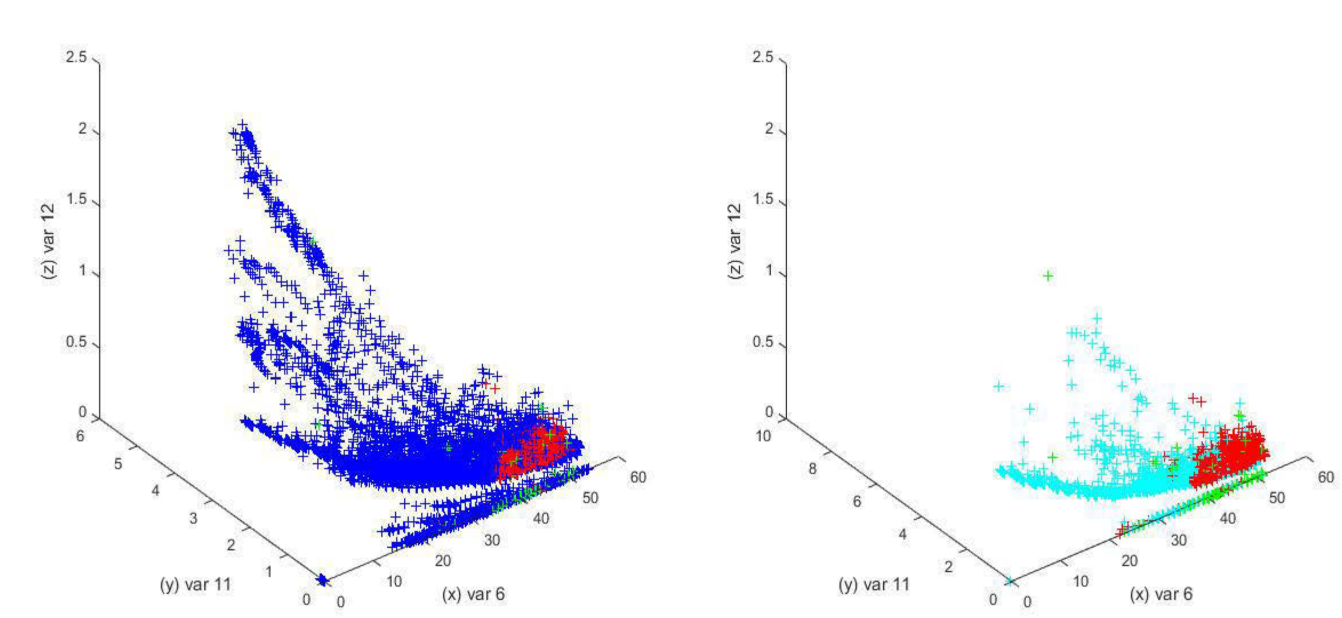


Results

Accuracy results based on the median of 100 KNN classifier iterations for a Main Bearing alarm.



3D plot that displays the variable converge before the alarm event (red zone).



The average computation time on reduced subset of 36 features x 88k entries:

- Exhaustive feature selection algorithms: 1-3 features exhaustive ~1h x WTG, 3-6 features top 500 histogram ~2h x WTG.
- Non-exhaustive feature selection algorithms: 1-3 ~1min x WTG, 3-6 ~5min x WTG.

Conclusions

- Feature selection algorithms improve the model accuracy over the use of all features as input, reducing the computation time and the input dimension space.
- The 3D selected input space shows “Healthy” and “alarms-failure” zones. This was seen for the 5 different WTG that are the same model. Even though the selected inputs were not exactly the same among the WTGs, they always bellowed to the same WTG’s system.
- The use of exhaustive methods are better with reduced set of inputs, but when the set of inputs grows the time increase exponentially and it becomes a non practical solution.
- The obtained results show the viability of automating one of the most critic steps when building a classifier for failure prediction in wind turbines.

References

1. H. Cheng, Z. Qin, C. Feng, Y. Wang, F. Li, Conditional mutual information-based feature selection analyzing for synergy and redundancy, *ETRI Journal 33 (2011) 210–218.*
2. P. E. Meyer, G. Bontempi, On the use of variable complementarity for feature selection in cancer classication, *Applications of Evolutionary Computing, Springer, 2006, pp. 91-102*
3. H. Peng, F. Long, C. Ding, Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy, Pattern Analysis and Machine Intelligence, *IEEE Transactions on 27 (2005) 1226–1238.*
4. F. Fleuret, Fast binary feature selection with conditional mutual information, *The Journal of Machine Learning Research 5 (2004) 1531-1555.*
5. H. H. Yang, J. E. Moody, Data visualization and feature selection: New algorithms for nongaussian data., *NIPS, volume 99, Citeseer, pp.687–693.*
6. A. Jakulin, Machine learning based on attribute interactions, *Ph.D. thesis, Univerza v Ljubljani, 2005*
7. R. Battiti, Using mutual information for selecting features in supervised neural net learning, Neural Networks, *IEEE Transactions on 5 (1994) 537-550.*