Università degli studi di Milano-Bicocca

Industry Lab

Final Project

Bosch GP5: Data Analysis and Leakage Coefficient Prediction

Authors:

Mauro Samarelli - 834196 - m.samarelli@campus.unimib.com

Ivan Mera - 783086 - i.merafranco@campus.unimib.it

June 9, 2020

Abstract

La prima parte di questo lavoro è stata effettuare un'analisi relativa alla pompa GP5 prodotta da Bosch. Le analisi effettuate sono state di tipo esplorativo includendo dunque le correlazioni tra le variabili nonché una stima del Leakage Coefficient(di seguito LC). Per la stima del Leakage Coefficient sono stati applicati modelli di regressione lineare, decision tree e random forest. Il modello più performante è risultato essere il modello di regressione lineare ottenendo un RMSE di val. La seconda parte di questo lavoro riguarda la costruzione di una dashboard interattiva che permettesse di monitorare il LC per ogni programma(pompa) all'interno del dataset. Inoltre la dashboard permette di vedere come varia il flusso in uscita al variare della pressione dato un certo range di valori per il LC. La dashboard è stata creata utilizzando il software Tableau.

1 Introduction

L'obiettivo di questo lavoro è poter garantire un continuo controllo sul LC in modo da capire quale tipologia di pompa sia più sensibile ad una variazione della pressione in uscita nonché analizzare eventuali correlazioni tra le variabili in oggetto o tra le due fasi del test della linea di assemblaggio. La GP5 è una pompa a bassa pressione che utilizza la benzina come lubrificante. Il suo funzionamento consiste nel ruotare degli ingranaggi in modo da generare una bassa pressione. Nel crearsi questa pressione il flusso di benzina viene portato verso la zona di alta pressione. The introduction should provide a clear statement of the problem posed by the project, and why the problem is of interest. It should reflect the scenario, if available. If needed, the introduction also needs to present background information so that the reader can understand the significance of the problem. A brief summary of the hypotheses and the approach your group used to solve the problem should be given, possibly also including a concise introduction to theory or concepts used later to analyze and to discuss the results.

2 Datasets

In this section the available data sets must be presented. The term dataset refers to any type of information source, for example web services for geolo-

cation fall into this category. In addition, all necessary data manipulation processes, such as cleaning and enrichment with external sources, must be presented and discussed.

3 The Methodological Approach

This is the central and most important section of the report. Its objective must be to show, with linearity and clarity, the steps that have led to the definition of a decision model. The description of the working hypotheses, confirmed or denied, can be found in this section together with the description of the subsequent refining processes of the models. Comparisons between different models (e.g. heuristics vs. optimal models) in terms of quality of solutions, their explainability and execution times are welcome.

Do not attempt to describe all the code in the system, and do not include large pieces of code in this section, use pseudo-code where necessary. Complete source code should be provided separately (in Appendixes, as separated material or as a link to an on-line repo). Instead pick out and describe just the pieces of code which, for example:

- are especially critical to the operation of the system;

- you feel might be of particular interest to the reader for some reason;

- illustrate a non-standard or innovative way of implementing an algorithm, data structure, etc..

You should also mention any unforeseen problems you encountered when implementing the system and how and to what extent you overcame them. Common problems are: difficulties involving existing software.

4 Results and Evaluation

The Results section is dedicated to presenting the actual results (i.e. measured and calculated quantities), not to discussing their meaning or interpretation. The results should be summarized using appropriate Tables and Figures (graphs or schematics). Every Figure and Table should have a legend that describes concisely what is contained or shown. Figure legends go below

the figure, table legends above the table. Throughout the report, but especially in this section, pay attention to reporting numbers with an appropriate number of significant figures.

5   Discussion

The discussion section aims at interpreting the results in light of the project's objectives. The most important goal of this section is to interpret the results so that the reader is informed of the insight or answers that the results provide. This section should also present an evaluation of the particular approach taken by the group. For example: Based on the results, how could the experimental procedure be improved? What additional, future work may be warranted? What recommendations can be drawn?

6   Conclusions

Conclusions should summarize the central points made in the Discussion section, reinforcing for the reader the value and implications of the work. If the results were not definitive, specific future work that may be needed can be (briefly) described. The conclusions should never contain "surprises". Therefore, any conclusions should be based on observations and data already discussed. It is considered extremely bad form to introduce new data in the conclusions.

References

The references section should contain complete citations following standard form. The references should be numbered and listed in the order they were cited in the body of the report. In the text of the report, a particular reference can be cited by using a numerical number in brackets as [?] that corresponds to its number in the reference list. LATEXprovides several styles to format the references