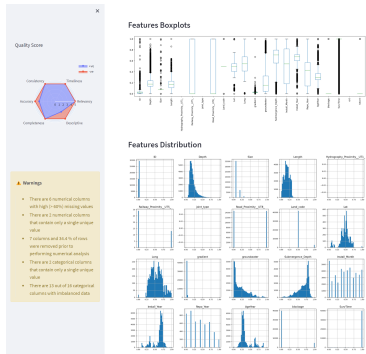


IDEA Tool



2022-07-06

Evaluate the appropriateness of data for the purposes of maintenance predictive analytics

Theme 2 Theme 5

Integrated Data Evaluation And Annotation

The aim of this tool is to:

1. evaluate the appropriateness of the data for the purposes of maintenance predictive analytics.
2. examine the information the data provides and the knowledge of the user about the data and the considered asset that allow them to interpret the data and build a reliable model appropriately.

The objective is not to build a model. Rather, the aim is to test whether there are enough data and knowledge of the appropriate fidelity to build a model.

The tool consists of three levels of assessment:

1. The first level assesses the metadata and user's knowledge to determine whether there is enough and relevant information of the intended analysis.
2. The second level conducts standard statistical and probabilistic measures to evaluate the accuracy and completeness of the associated numerical data.
3. The third level operationalises advanced nonlinear and machine learning measures to further assess the quality of the numerical data in relevant to the intended task.

The quality and appropriateness of the data are assessed based on the following criteria:

1. **Relevancy** - The data must be relevant to the requirements for the using purpose, that is, specific information must be recorded for specific models or analysis.
2. **Timeliness** - The data records must be labelled by accurate time and date; time frequency would affect the quality of the data; time and date of any incident (event) must be recorded; downtime and its causes must be recorded.
3. **Consistency** - The data format, recording scheme and preprocessing of raw data must be consistent.
4. **Accuracy** - The data needs to be accurate - outliers and noise reduce the quality level of the data. Data accuracy is balanced by data quantity.
5. **Completeness** - The data must not have too many missing records, in other words, the missing data ratio affects the quality of the data.
6. **Descriptive** - Data must be accompanied by sufficient explanation of resolution, fidelity, sampling, units of measurement and sources of noise.

Try the demo