Hochschule Hof
University of Applied Sciences

Master AIR

Predictive Maintenance and Condition Monitoring

Test exam summer term 2023

ALLOWED AIDS

Handwritten notes, lecture notes, calculator, books, PC with Jupyter-Hub

INDICATIONS FOR THE EXAMINEES

- This exam includes 3 pages. *Please check the exam for completeness before starting.*
- Only the written answers will be graded.
- Please return this exam.
- Write your matriculation number on the exam and on every paper that you return.

General Questions

Question 1 (/ 10)

Define the term 'Regression'! Define the term 'Classification'!
Briefly describe an application scenario for each term!

Case Study

Question 2

Your customer supplied a dataset with production data:

1. `f1` ... `f4`: Data measured during production and aggregated into feature-values for each product
2. `target`: discrete variable describing the quality for each product / set of feature values:
   - -1: product is too small
   - 0: product is inside tolerance band
   - 1: product is too big

The data is stored in the CSV-File `Data01.CSV`.

Question 3 **(/ 10)**

Upload the CSV-File to your Jupyter-Hub account, load the CSV-File and count the number of each value of the `target`-column!
What is the yield of the production? (Ratio bad parts to total parts produced)

Subquestion 3.a **(/ 5)**

Give the minimum-, average- and maximum-values for each of the feature columns!
Is it necessary to scale the feature values before proceeding in the analysis? Justify your answer!

Subquestion 3.b **(/ 5)**

Scale the features with an appropriate scaler and give the minimum-, average- and maximum-values for each of the scaled feature columns!

Subquestion 3.c **(/ 10)**

Perform a PCA-analysis on the scaled dataset and answer the following questions:

1. What is the smallest possible number of PCA components for this dataset? Justify your answer!
2. What is the biggest possible number of PCA components for this dataset? Justify your answer!
3. Draw a curve of the cumulated variance explained by the PCA-components over the number of PCA-components!!

Subquestion 3.d **(/ 10)**

Perform a PCA to reduce the number of features to two features and plot a dot for each data point (one of the PCA-features on the X-axis, on on the Y-axis)! Mark each data point in a color corresponding to it's target value! Give a sketch the resulting plot!

Subquestion 3.e **(/ 10)**

Choose a suitable scoring metric to measure the performance of a predictor for the target value from the feature columns!
Assuming a simple predictor that always predicts the most frequent target value, give a baseline for a minimum viable predictor!

Total points (/ 60)