

Tutorial Task 3: Analytics Challenge

Group No. 06

Names

- Nils Seibach

- Harshith Srinivas

- Thorsten Beumer

- Athary Paranjape

- Edgard Hols

Task Description (Intro)

This task revisits the Case Study "Kruise GmbH". It covers the phases of an Industrial Use Case:

a) Use Case Understanding and Planning
b) Data Understanding: Exploration, Preprocessing and Features
c) Modelling

The accompanying dataset is provided in PANDA and can be downloaded there. It also includes the use case and dataset description (see below).

The tasks are detailed in the boxes below.

Kruise GmbH: Detect Product Faults with a Smart Quality System

Use Case Description

Kruise GmbH is struggling with increasing quality problems and inefficiencies in maintenance in production of their highest-selling product, the washing machine. The front panels are particularly affected by the quality problems. Embossing with subsequent printing by a stamp results in a reject rate of 5-10%. Another problem is the frequent stoppage of the conveyor belts. This has resulted in a very expensive regular replacement of the bearings.

After consultation with an engineer, it is suspected that the vibrations of the ball bearings in the conveyor belt are related to the quality condition of the product being manufactured. Vibration data with corresponding time stamps are recorded by a Machine Database, historical data is available in the MES System. In addition, there are about 1500 production logs in the ERP System that show which product was manufactured at which time and the quality state (ok, nok = not ok) of the printing on the products is known from the quality logs in the QMS.

Print system

Conveyor belt

Ball Bearing

1

© Heinz Nixdorf Institut / Fraunhofer IPA

Kruise GmbH: Detect Product Faults with a Smart Quality System

Task and Dataset Description

Task: Help the Kruise GmbH and develop a monitoring system, that identifies the quality of the printing on the product using the historical vibration data and the quality and production log-data.

Description of the data set: There is data for one bearing with 2 acceleration sensors. The data set consists of individual files that are 1-second vibration signal snapshots recorded at specific intervals. Each file consists of a time series with 20 480 points at a sampling rate of ~20 kHz. The file name indicates when the data was collected. Each record (file) represents a production process and a corresponding product (row in production or quality log respectively).

Drive belt

Motor

Bearing

Radial Load

Accelerometers

Vibration measurements	production log	quality log
2023-11-08 16:01:45 P1.0.0001	2023-11-08 16:01:45 P1.0.0001	2023-11-08 16:01:45 P1.0.0001
2023-11-08 16:02:30 P1.0.0001	2023-11-08 16:02:30 P1.0.0001	2023-11-08 16:02:30 P1.0.0001
2023-11-08 16:03:15 P1.0.0001	2023-11-08 16:03:15 P1.0.0001	2023-11-08 16:03:15 P1.0.0001
2023-11-08 16:04:00 P1.0.0001	2023-11-08 16:04:00 P1.0.0001	2023-11-08 16:04:00 P1.0.0001
2023-11-08 16:04:45 P1.0.0001	2023-11-08 16:04:45 P1.0.0001	2023-11-08 16:04:45 P1.0.0001
2023-11-08 16:05:30 P1.0.0001	2023-11-08 16:05:30 P1.0.0001	2023-11-08 16:05:30 P1.0.0001
2023-11-08 16:06:15 P1.0.0001	2023-11-08 16:06:15 P1.0.0001	2023-11-08 16:06:15 P1.0.0001
2023-11-08 16:07:00 P1.0.0001	2023-11-08 16:07:00 P1.0.0001	2023-11-08 16:07:00 P1.0.0001
2023-11-08 16:07:45 P1.0.0001	2023-11-08 16:07:45 P1.0.0001	2023-11-08 16:07:45 P1.0.0001
2023-11-08 16:08:30 P1.0.0001	2023-11-08 16:08:30 P1.0.0001	2023-11-08 16:08:30 P1.0.0001
2023-11-08 16:09:15 P1.0.0001	2023-11-08 16:09:15 P1.0.0001	2023-11-08 16:09:15 P1.0.0001
2023-11-08 16:10:00 P1.0.0001	2023-11-08 16:10:00 P1.0.0001	2023-11-08 16:10:00 P1.0.0001
2023-11-08 16:10:45 P1.0.0001	2023-11-08 16:10:45 P1.0.0001	2023-11-08 16:10:45 P1.0.0001
2023-11-08 16:11:30 P1.0.0001	2023-11-08 16:11:30 P1.0.0001	2023-11-08 16:11:30 P1.0.0001
2023-11-08 16:12:15 P1.0.0001	2023-11-08 16:12:15 P1.0.0001	2023-11-08 16:12:15 P1.0.0001
2023-11-08 16:13:00 P1.0.0001	2023-11-08 16:13:00 P1.0.0001	2023-11-08 16:13:00 P1.0.0001
2023-11-08 16:13:45 P1.0.0001	2023-11-08 16:13:45 P1.0.0001	2023-11-08 16:13:45 P1.0.0001
2023-11-08 16:14:30 P1.0.0001	2023-11-08 16:14:30 P1.0.0001	2023-11-08 16:14:30 P1.0.0001
2023-11-08 16:15:15 P1.0.0001	2023-11-08 16:15:15 P1.0.0001	2023-11-08 16:15:15 P1.0.0001
2023-11-08 16:16:00 P1.0.0001	2023-11-08 16:16:00 P1.0.0001	2023-11-08 16:16:00 P1.0.0001
2023-11-08 16:16:45 P1.0.0001	2023-11-08 16:16:45 P1.0.0001	2023-11-08 16:16:45 P1.0.0001
2023-11-08 16:17:30 P1.0.0001	2023-11-08 16:17:30 P1.0.0001	2023-11-08 16:17:30 P1.0.0001
2023-11-08 16:18:15 P1.0.0001	2023-11-08 16:18:15 P1.0.0001	2023-11-08 16:18:15 P1.0.0001
2023-11-08 16:19:00 P1.0.0001	2023-11-08 16:19:00 P1.0.0001	2023-11-08 16:19:00 P1.0.0001
2023-11-08 16:19:45 P1.0.0001	2023-11-08 16:19:45 P1.0.0001	2023-11-08 16:19:45 P1.0.0001
2023-11-08 16:20:30 P1.0.0001	2023-11-08 16:20:30 P1.0.0001	2023-11-08 16:20:30 P1.0.0001
2023-11-08 16:21:15 P1.0.0001	2023-11-08 16:21:15 P1.0.0001	2023-11-08 16:21:15 P1.0.0001
2023-11-08 16:22:00 P1.0.0001	2023-11-08 16:22:00 P1.0.0001	2023-11-08 16:22:00 P1.0.0001
2023-11-08 16:22:45 P1.0.0001	2023-11-08 16:22:45 P1.0.0001	2023-11-08 16:22:45 P1.0.0001
2023-11-08 16:23:30 P1.0.0001	2023-11-08 16:23:30 P1.0.0001	2023-11-08 16:23:30 P1.0.0001
2023-11-08 16:24:15 P1.0.0001	2023-11-08 16:24:15 P1.0.0001	2023-11-08 16:24:15 P1.0.0001
2023-11-08 16:25:00 P1.0.0001	2023-11-08 16:25:00 P1.0.0001	2023-11-08 16:25:00 P1.0.0001
2023-11-08 16:25:45 P1.0.0001	2023-11-08 16:25:45 P1.0.0001	2023-11-08 16:25:45 P1.0.0001
2023-11-08 16:26:30 P1.0.0001	2023-11-08 16:26:30 P1.0.0001	2023-11-08 16:26:30 P1.0.0001
2023-11-08 16:27:15 P1.0.0001	2023-11-08 16:27:15 P1.0.0001	2023-11-08 16:27:15 P1.0.0001
2023-11-08 16:28:00 P1.0.0001	2023-11-08 16:28:00 P1.0.0001	2023-11-08 16:28:00 P1.0.0001
2023-11-08 16:28:45 P1.0.0001	2023-11-08 16:28:45 P1.0.0001	2023-11-08 16:28:45 P1.0.0001
2023-11-08 16:29:30 P1.0.0001	2023-11-08 16:29:30 P1.0.0001	2023-11-08 16:29:30 P1.0.0001
2023-11-08 16:30:15 P1.0.0001	2023-11-08 16:30:15 P1.0.0001	2023-11-08 16:30:15 P1.0.0001
2023-11-08 16:31:00 P1.0.0001	2023-11-08 16:31:00 P1.0.0001	2023-11-08 16:31:00 P1.0.0001
2023-11-08 16:31:45 P1.0.0001	2023-11-08 16:31:45 P1.0.0001	2023-11-08 16:31:45 P1.0.0001
2023-11-08 16:32:30 P1.0.0001	2023-11-08 16:32:30 P1.0.0001	2023-11-08 16:32:30 P1.0.0001
2023-11-08 16:33:15 P1.0.0001	2023-11-08 16:33:15 P1.0.0001	2023-11-08 16:33:15 P1.0.0001
2023-11-08 16:34:00 P1.0.0001	2023-11-08 16:34:00 P1.0.0001	2023-11-08 16:34:00 P1.0.0001
2023-11-08 16:34:45 P1.0.0001	2023-11-08 16:34:45 P1.0.0001	2023-11-08 16:34:45 P1.0.0001
2023-11-08 16:35:30 P1.0.0001	2023-11-08 16:35:30 P1.0.0001	2023-11-08 16:35:30 P1.0.0001
2023-11-08 16:36:15 P1.0.0001	2023-11-08 16:36:15 P1.0.0001	2023-11-08 16:36:15 P1.0.0001
2023-11-08 16:37:00 P1.0.0001	2023-11-08 16:37:00 P1.0.0001	2023-11-08 16:37:00 P1.0.0001
2023-11-08 16:37:45 P1.0.0001	2023-11-08 16:37:45 P1.0.0001	2023-11-08 16:37:45 P1.0.0001
2023-11-08 16:38:30 P1.0.0001	2023-11-08 16:38:30 P1.0.0001	2023-11-08 16:38:30 P1.0.0001
2023-11-08 16:39:15 P1.0.0001	2023-11-08 16:39:15 P1.0.0001	2023-11-08 16:39:15 P1.0.0001
2023-11-08 16:40:00 P1.0.0001	2023-11-08 16:40:00 P1.0.0001	2023-11-08 16:40:00 P1.0.0001
2023-11-08 16:40:45 P1.0.0001	2023-11-08 16:40:45 P1.0.0001	2023-11-08 16:40:45 P1.0.0001
2023-11-08 16:41:30 P1.0.0001	2023-11-08 16:41:30 P1.0.0001	2023-11-08 16:41:30 P1.0.0001
2023-11-08 16:42:15 P1.0.0001	2023-11-08 16:42:15 P1.0.0001	2023-11-08 16:42:15 P1.0.0001
2023-11-08 16:43:00 P1.0.0001	2023-11-08 16:43:00 P1.0.0001	2023-11-08 16:43:00 P1.0.0001
2023-11-08 16:43:45 P1.0.0001	2023-11-08 16:43:45 P1.0.0001	2023-11-08 16:43:45 P1.0.0001
2023-11-08 16:44:30 P1.0.0001	2023-11-08 16:44:30 P1.0.0001	2023-11-08 16:44:30 P1.0.0001
2023-11-08 16:45:15 P1.0.0001	2023-11-08 16:45:15 P1.0.0001	2023-11-08 16:45:15 P1.0.0001
2023-11-08 16:46:00 P1.0.0001	2023-11-08 16:46:00 P1.0.0001	2023-11-08 16:46:00 P1.0.0001
2023-11-08 16:46:45 P1.0.0001	2023-11-08 16:46:45 P1.0.0001	2023-11-08 16:46:45 P1.0.0001
2023-11-08 16:47:30 P1.0.0001	2023-11-08 16:47:30 P1.0.0001	2023-11-08 16:47:30 P1.0.0001
2023-11-08 16:48:15 P1.0.0001	2023-11-08 16:48:15 P1.0.0001	2023-11-08 16:48:15 P1.0.0001
2023-11-08 16:49:00 P1.0.0001	2023-11-08 16:49:00 P1.0.0001	2023-11-08 16:49:00 P1.0.0001
2023-11-08 16:49:45 P1.0.0001	2023-11-08 16:49:45 P1.0.0001	2023-11-08 16:49:45 P1.0.0001
2023-11-08 16:50:30 P1.0.0001	2023-11-08 16:50:30 P1.0.0001	2023-11-08 16:50:30 P1.0.0001
2023-11-08 16:51:15 P1.0.0001	2023-11-08 16:51:15 P1.0.0001	2023-11-08 16:51:15 P1.0.0001
2023-11-08 16:52:00 P1.0.0001	2023-11-08 16:52:00 P1.0.0001	2023-11-08 16:52:00 P1.0.0001
2023-11-08 16:52:45 P1.0.0001	2023-11-08 16:52:45 P1.0.0001	2023-11-08 16:52:45 P1.0.0001
2023-11-08 16:53:30 P1.0.0001	2023-11-08 16:53:30 P1.0.0001	2023-11-08 16:53:30 P1.0.0001
2023-11-08 16:54:15 P1.0.0001	2023-11-08 16:54:15 P1.0.0001	2023-11-08 16:54:15 P1.0.0001
2023-11-08 16:55:00 P1.0.0001	2023-11-08 16:55:00 P1.0.0001	2023-11-08 16:55:00 P1.0.0001
2023-11-08 16:55:45 P1.0.0001	2023-11-08 16:55:45 P1.0.0001	2023-11-08 16:55:45 P1.0.0001
2023-11-08 16:56:30 P1.0.0001	2023-11-08 16:56:30 P1.0.0001	2023-11-08 16:56:30 P1.0.0001
2023-11-08 16:57:15 P1.0.0001	2023-11-08 16:57:15 P1.0.0001	2023-11-08 16:57:15 P1.0.0001
2023-11-08 16:58:00 P1.0.0001	2023-11-08 16:58:00 P1.0.0001	2023-11-08 16:58:00 P1.0.0001
2023-11-08 16:58:45 P1.0.0001	2023-11-08 16:58:45 P1.0.0001	2023-11-08 16:58:45 P1.0.0001
2023-11-08 16:59:30 P1.0.0001	2023-11-08 16:59:30 P1.0.0001	2023-11-08 16:59:30 P1.0.0001
2023-11-08 16:59:45 P1.0.0001	2023-11-08 16:59:45 P1.0.0001	2023-11-08 16:59:45 P1.0.0001

2

© Heinz Nixdorf Institut / Fraunhofer IPA

a) Use Case Understanding and Planning

Business to Analytics Objectives

TASK DESCRIPTION

Describe the business objective for your use case and translate it into an analytics objective and define the expected output of the analytics model. Use the template to the right.

Business objective

Reducing big expenses for rejects via detecting current rejects and reducing time conveyor belt needs to be stopped

Analytics objective

- Correlate vibration data with the QA verdict

- Build classification model that can predict OK/NOK based on vibration data

Output of the model

Correlation between resources
Supervised model for classification

Analytics Canvas

TASK DESCRIPTION

Fill in the Analytics Canvas according to the Use Description. Use the icons for "available" in the template.

Make suggestions of how to further improve the Use Case by adding "missing" icons.

Analytics Use Case	Descriptive	Diagnostic	Predictive	Prescriptive
Data Analysis				
Data Pools	MES	QMS	ERP	
Data Description	Vibration Data (per frequency)	Production OK / NOK	Product quality with time stamp	
Automat.	MES		ERP	
Data Source	Manual			
Resource	Machine DB	QA Eval Verdict	Historical Data	

Constructs of the Analytics Canvas

Connection

Available

Missing

b) Data Understanding: Exploration, Preprocessing, Features

Data Visualization and Exploration

TASK DESCRIPTION

Visualize the data and paste the important visualizations into this field. Highlight and explain your insights and expected challenges with the dataset.

We can see that not only the amplitude of the vibrations, but also the frequencies of the vibrations might shift. So it would make sense to analyze the data from the fft to train a classifier.

Data Preprocessing and Feature Extraction

Which preprocessing is necessary for the data and why?	
Method of preprocessing	Explanation
Average Filter	To Reduce Noise in the data
Fast fourier transformation	To extract meaning full features from vibration data
Resampling the data	To train and evaluate the classifier over a balanced set of values of our target variable

TASK DESCRIPTION

Plan your Preprocessing and Feature Extraction with the tables provided above. Perform Preprocessing and Feature Extraction on the data accordingly and add visualizations below. You can put them next to your insights of the Exploration phase, if fitting.

Which features are useful and why?	
Feature	Explanation
We divide the frequency of the measured vibration into intervals and average the amplitude in those intervals	We can learn what vibrations occurred and in what intensity. Averaging over intervals reduces noise and reduces the number of features.

c) Modelling

Modelling

TASK DESCRIPTION

Which methods are suitable to model your analytics case? List possible models to the right and explain your choice. Train your model and past visualisation of results below.

Define Suitable Models and Explain your reasoning.

Model	Explanation
Decision Tree	We need to predict a categorical variable, so we should use a classification algorithm. De

Model Evaluation and Improvement

TASK DESCRIPTION

Evaluate your model and fill in the table below. How could you improve your model? Add the improvements in the analytics canvas using the red symbols.

Define indicators on which you can validate the performance of your model. Explain the results.

Indicator	Explanation
Confusion Matrix	To evaluate a classifier it is important to evaluate the performance over both error types. The confusion matrix does this.

