

Open Thesis / Project

Testing Robustness of Sensor-based Models

Motivation

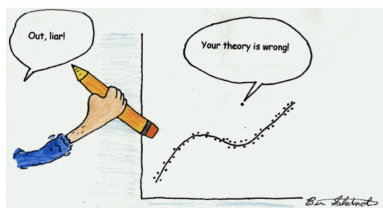
Robustness is a measure how sensitive a model is to smaller or larger deviations of the input from the assumptions the model relies on, or the distribution observed in the train data. These deviations are called outliers and may considerably degrade the performance of the predictive models. Classical robust statistics introduces a range of measures of robustness: breakdown value, sensitivity curve, influence function, maxbias curve, etc.. We will build on these measures and introduce an iterative procedure to both (1) measure model robustness and (2) conclude on sensor susceptibility to outliers under a given model. Moreover, you will test a novel method for improving model robustness and outlier filtering. In this thesis, you will be working with a data set which comprises activities of dairy cows and the corresponding lameness predictive models. Other data sets may also be used to justify the effectiveness of the proposed method. The results will be integrated into an existing cloud-based framework. **Interested? Contact us for more details!**

Target Group

Students in ICE and Computer Science.

Thesis Type

Master Project / Master Thesis.



An outlier is an observation that deviates from the fit suggested by the majority of the observations. Outliers, common in sensor data, may have a major impact on the the performance of predictive models. See: [Robust Statistics](#).

Goals and Tasks

In this project, you will focus on applying existing methods of robust statistics (1) in multivariate settings, and (2) as part of a novel iterative procedure, which will allow improving both input sensor data and the performance of predictive models that use this data. The approach will be tested on the provided data (cow activity data are of specific interest). The project includes the following tasks:

- Literature review on robust statistics and its application in the context of outlier detection and analysis of predictive models.
- Robustness analysis of several models operating on various data sets.
- Extend robustness measures to improve robustness of predictive models and to detect outliers in existing data.
- Integration of the solution into an existing cloud-based service.
- Summary of the results in a written report, oral presentation.

Requirements / Skills:

- A good knowledge of data analysis methods and statistics, enjoy working with real data.
- Interest to learn robust statistics (no prior knowledge is required), creative thinking.
- Programming skills in Python.

Contact Person

- Franz Papst (papst@tugraz.at)
- Dr. Olga Saukh (saukh@tugraz.at)



Institute for Technical Informatics
Networked Embedded Systems Group

