



TU Clausthal

Technische Universität Clausthal • Postfach 12 53 • 38670 Clausthal-Zellerfeld

Institut für Angewandte
Stochastik und Operations
Research

Arbeitsgruppe
Stochastische Optimierung
Professor Dr. Michael Kolonko

Bearbeiter/in:
Bärbel Heise-Kretzer

Telefon: (0 53 23) 72-2403
Telefax: (0 53 23) 72-2304
baerbel.heise@tu-clausthal.de

Ihr Zeichen/Ihr Schreiben vom

Mein Zeichen/Mein Schreiben vom

Clausthal-Zellerfeld, den
30. Januar 2018

To whom it concerns

Mr. Aditya Raj, born 26.12.1993 in India, is a student of the international master program 'Internet Technologies and Information Systems' at TU Clausthal, a joint program of the four universities at Clausthal, Braunschweig, Hannover, and Göttingen.

In March 2017, Mr. Raj joined my group at the Institute for Applied Stochastics and Operations Research for his Master Thesis. He became part of an ongoing research on machine learning for online quality control. This is a project in cooperation with Salzgitter Flachstahl GmbH, a major German steel company, focused on high quality flat rolled steel. In this project, we have developed a program system 'ROCA' to detect severe defects on the steel coil based on the data provided by the existing, though not sufficiently reliable, surface inspection system. ROCA is continuously extended by adding new tools for the detection of severe error events. One of its central components is the classification of errors based on error patterns, in particular the spatial clustering of the events.

The task of Mr. Raj was to develop a flexible algorithm to detect outliers in existing clusters. Here 'outlier' means an event not belonging to the cluster though spatially located within. This happens e.g. when a cluster of harmless events covers a singular severe defect.

Mr. Raj started with a survey on the relevant statistical literature on outliers. He implemented and tested many of these algorithms until he was able to formulate a new approach. Many of the existing algorithms cannot handle data with high dimensionality (our data points have up to 200 components of different characteristics), or showed generally poor results in our application. Mr. Raj's algorithm iteratively examines subsets of data points within a given cluster. In each iteration, the data points of the cluster are evaluated whether they are outliers with respect to the present subset. The decisions are considered as 'votes' on the status of the data point. These votes are collected and data points with many votes

Besuchsanschrift:

Gebäude B7
Erzstraße 1
38678 Clausthal-Zellerfeld

Telefon: (0 53 23) 72-2567
Telefax: (0 53 23) 72-2304
office@iasor.tu-clausthal.de
<http://www.iasor.tu-clausthal.de>

Briefanschrift:
Postfach 12 53
38670 Clausthal-Zellerfeld

Bankverbindung:
Sparkasse Hildesheim-Goslar-Peine
IBAN: DE71 259501300000022111
Swift/BIC Code: NOLADE21HIK

USt-Ident-Nr. DE811282802

are marked as outliers. In second stage, it is checked whether these outliers are ordinary one-dimensional outliers in at least one coordinate. In this way, Mr. Raj was able to reduce the chance of a false classification as outlier (false positive).

He tested his algorithm with realistic data into which artificial errors had been injected. It turned out that with respect to false negative and false positive results his approach was definitely superior to other algorithms from literature. This presents a real progress on the algorithmically side and will improve the quality control at the steel plant in practice.

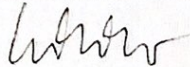
Mr. Raj worked at Salzgitter Flachstahl for some time to fully understand the details of the problem and later to evaluate his solution with real data and with the experience of the human steel experts.

He is a very engaged and highly motivated student, working very independently. In our regular meetings, he always presented his new findings in diligently prepared reports and was open for new ideas. He was able to understand the relevant technical details of the application in a very short time and to get acquainted with the theoretical background of the available statistical methods. He implemented his algorithm, first in the statistical language 'R' for testing and comparison, later in JAVA to incorporate it into the existing productive system.

Mr. Raj is one of the rare students that can work equally well on the mathematical theory and on its application to practical problems. This includes the ability to decide where trade-offs have to be made to arrive at a useful practical tool. Though the Master thesis is not yet quite finished, I am sure that he will receive an excellent grade for it.

Mr. Raj is a very agreeable and reliable person which I can recommend for challenging tasks, in particular at the borderline between theory and practice.

Mit freundlichen Grüßen



Prof. Dr. Michael Kolonko