Revision of doctoral (PhD) thesis

Name of the candidate: Gergely Horváth

Title of PhD thesis: Model–Based Analysis and Optimization of Industrial MDI Manufacturing

Reviewer: Gyula Ádám Dörgő, Sr Specialist Data and Process Mining, Borouge International

The outstanding work of Gergely Horváth provides a significant novelty in the increasingly actual topic of the application of machine learning and data analysis in the chemical industry. The work presents three case studies to demonstrate the applicability and performance of data-based modelling techniques in the production of methylenediphenyl diisocyanate. Each of these case studies was presented in a separate section: „Exploration and model-based analysis of MDA formation reactions", „Machine learning model development for MDA formation based on laboratory data", and „Soft-sensor development for product quality estimation with time delay and feature selection in industrial MDI production". The diverse nature of these case studies perfectly demonstrates how the toolboxes of data science and machine learning are universally applicable across different fields of engineering practice, moreover, highlighting the widespread knowledge of the candidate in the fields of chemical engineering and mathematical modeling.

It was revealing to read how the candidate analyzed and connected the deeply chemical and technical questions to mathematical modeling, formulated hypotheses, and proved them with data-based approaches. Special attention was paid to comparing the results with measurement data, which also had to be analyzed and preprocessed to meet the requirements of the research goals. The selected methods are justified, the case studies and results are evaluated, and the conclusions drawn are accurate and reflect the candidate's professional competence. It was a pleasure to read this work.

Formal and structural evaluation

From a structural perspective, after the introduction and literature review of model-based analysis and isocyanate production, the three case studies are presented in subsequent sections, providing the main findings of the thesis. As all of these sections describe independent and standalone analysis results, in my view, the selection of this structure is favorable and advantageous. The sections build on the already accepted journal articles of the candidate,

underpinning the scientific foundations of the results. The main part of the thesis consists of 152 pages and contains 142 references. Every part of the dissertation, from the introduction to the bibliography, is well-edited and constructed in a transparent and easy-to-follow structure. The figures are high-quality and self-explanatory, meeting the requirements of international publication standards.

Theses

Thesis #1: "I expanded the reaction system of existing models reported in the literature for the synthesis of Methylenedianiline by incorporating additional components and reaction pathways, I achieved enhanced model accuracy and reduced information loss."

(3 sub-theses)

The results of the thesis were published in a journal article and a conference abstract, for which the candidate has received a total of 4 citations according to the MTMT.

Thesis #2: "I justified that machine learning models efficiently support the understanding of Methylenedianiline synthesis variables on laboratory-scale, aiming to examine the impact of variables on Methylenedianiline quality and to assist large-scale industrial production."

(2 sub-theses)

The results of the thesis were published in a journal article and two conference abstracts.

Thesis #3: "Through the proposed intervention strategies, I justified that soft-sensor models can be well applied for Methylenediphenyl diisocyanate color prediction based on real industrial data in an explainable way, illustrated the non-linear behavior of the system and identified the best operational parameters for an industrial Methylenediphenyl diisocyanate production plant."

(4 sub-theses)

The results of the thesis were published in a journal article and two conference abstracts.

Based on the thesis work and the designated publication activity, it can be unequivocally stated that the work described in all three thesis groups is authentic and represents the candidate's own scientific results.

Question:

Following the internal presentation and discussion, and receiving the answers from the candidate to my comments, I have no formal comments left. I had no serious critical comments or concerns even during the previous evaluation round. My general question:

All of the findings are focusing on highly practical and industrially relevant findings. What are the next steps towards the industrial application of the findings (especially, the published soft-sensor solution)? What infrastructural, workforce, training, business process, etc. requirements are needed for a successful implementation, especially taking into account the CRISP-ML(Q) methodology discussed in the thesis?

Summary and Evaluation:

Considering the importance of the topic, its scientific significance, its novelty, the high standard of professional achievement, the detailedness demonstrated in the discussion of the analyses, and the publication record covering the topic (3 journal articles and 5 conference abstracts), the dissertation is an outstanding scientific achievement. The candidate has met the requirements for a PhD dissertation. In my opinion, the PhD dissertation of Gergely Horváth satisfies the conditions to present the findings at a public debate. Following the successful defense of Gergely Horváth's dissertation, I highly recommend that the candidate be awarded the PhD degree.

Vienna, 02.04.2026

Gyula Ádám Dörgő
Sr Specialist Data and Process Mining
Borouge International