

# Explainable deep learning for improved real-time monitoring of a chromatographic protein A capture step

Theresa Scharl<sup>1</sup>, Matthias Medl<sup>1</sup>, Friedrich Leisch<sup>1</sup>, and Astrid Dürauer<sup>2</sup>

<sup>1</sup>Universität für Bodenkultur Wien Institut für Statistik

<sup>2</sup>University of Natural Resources and Life Sciences

May 13, 2023

## Abstract

Model-based real-time monitoring of biopharmaceutical production is a major step towards Quality-by-Design in this field and the fundament for model predictive control. Data-driven models have been proven a viable option to model bioprocesses. In the high stakes setting of biopharmaceutical manufacturing it is essential to ensure high model accuracy, robustness and explainability. That is only possible when (i) the data used for modeling is of high quality, (ii) state-of-the-art modeling algorithms are employed and (iii) the input-output mapping of the model has been characterized. In this study we evaluate the accuracy of multiple data-driven models in predicting the monoclonal antibody concentration, dsDNA concentration, host cell protein concentration, and high molecular weight impurity content during elution from a protein A chromatography capture step. We demonstrate how permutation/occlusion-based methods can be used to gain understanding on dependencies learned by one of the most complex data-driven models; convolutional neural network ensembles. Finally, we present a workflow to test the model behavior in case of simulated sensor fouling and failure. This study represents a major step towards improved viability of data-driven models in biopharmaceutical manufacturing.

## Hosted file

Deep\_learning\_draft\_v7.docx available at <https://authorea.com/users/596739/articles/643180-explainable-deep-learning-for-improved-real-time-monitoring-of-a-chromatographic-protein-a-capture-step>