Inflection Point Detection for a Titrator

Graduate

Laurin Heitzer

Marco Stauber

Introduction: A titrator is an analytical instrument used to determine the concentration of a dissolved substance by incrementally adding a reagent of known concentration until a chemical reaction reaches completion. This reaction can be monitored through various methods, including measuring the solution's potential, using pH sensors, or observing color changes in an indicator solution. In a titration curve, measurements are plotted against the volume of the added titrant. The primary objective is to identify the inflection point, where the amount of titrant added neutralizes the analyte. While simple curve-fitting algorithms perform well under ideal conditions, real-world signals often exhibit significant noise, posing challenges for existing methods.

Definition of Task: The goal of this bachelor's thesis is to develop a machine learning (ML) algorithm to accurately identify inflection points in measured titration curves. The ML model will be trained exclusively on simulated datasets, where the ground truth of inflection points is known. A key component of this thesis is the development of a simulation tool to generate realistic titration curves. The final ML-based algorithm will be evaluated using both real-world and simulated data, with a comparative analysis against existing algorithms regarding accuracy and run-time.

Result: A software tool was developed to generate titration curves. Additionally, noise characteristics from real-world data were analyzed to create realistic noise models for these curves. These datasets serve as a foundation for training the ML model. Two approaches were evaluated: a digital signal processing (DigSig) model and a convolutional neural network (CNN). The DigSig model uses conventional signal processing techniques as a benchmark, while the CNN is trained on the generated curves.

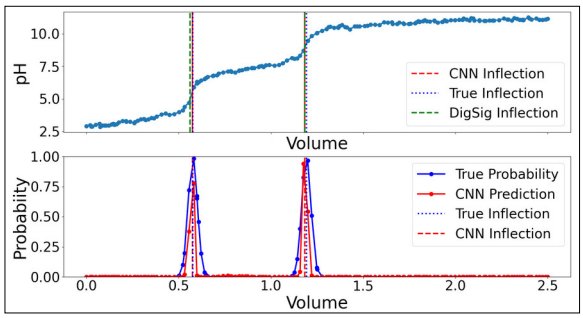
Results on real data:

- DigSig model: precision: 86.7% recall: 75%
- CNN model: precision: 92.7% recall: 73%

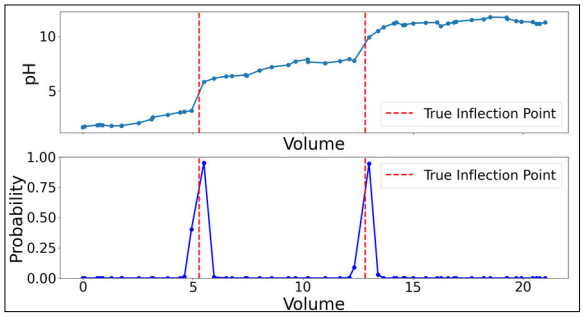
Results on generated data:

- DigSig model: precision: 93.6% recall: 93.5%
- CNN model: precision: 94.6% recall: 85.3%

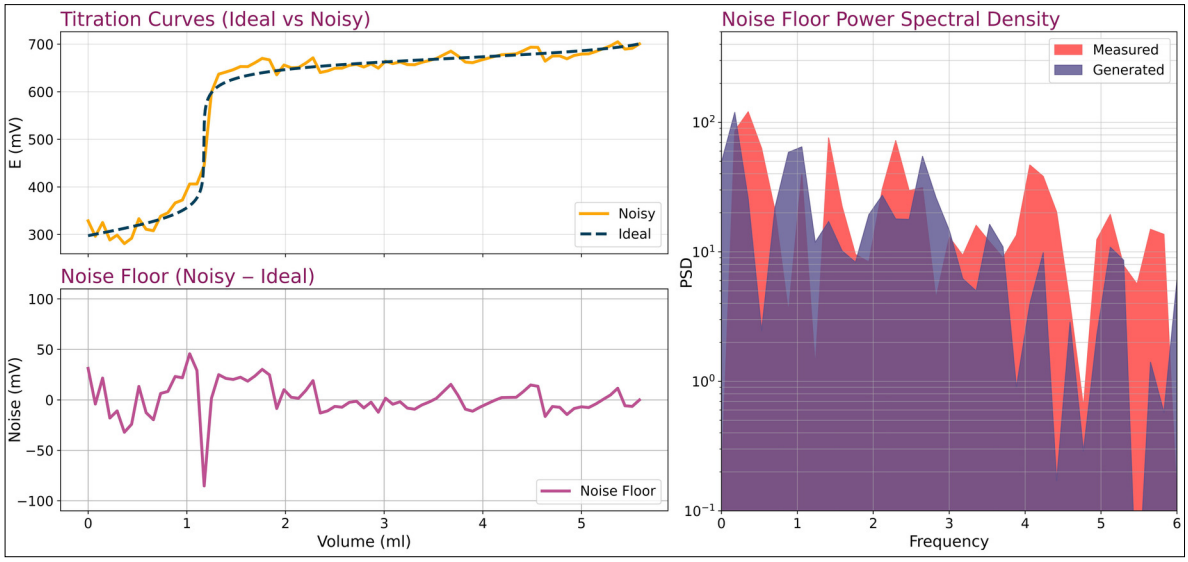
Prediction of two inflection points with the DigSig algorithm and the CNN model
Own presentment



Generating probability using sampled gaussian curves at inflection points.
Own presentment



Titration curve fitting and noise analysis
Curve data provided by industry partner



Advisors
Hannes Badertscher, Patrik Müller

Co-Examiner
Gabriel Sidler, Teamup Solutions AG, Uster, ZH

Subject Area
Artificial Intelligence