1158-68-360   **Yu-Hang Tang*** (`tang.maxin@gmail.com`), 1 Cyclotron Rd, Berkeley, CA 94720. *Outlier Detection on Experimental Data Using Gaussian Process Regression with A Non-Stationary Noise Model.* Preliminary report.

Supervised machine learning performs optimally using accurately labeled ground truth as training data. However, many next-generation high-throughput chemical experiment techniques produce intrinsically noisy data. As such, it is very challenging to train machine learning models using this type of noisy data. We propose a non-stationary noise model on top of the Gaussian process regression framework. Together with a graph kernel that operates directly on graph representations of molecules, our method can infer the reliability of chromatography data on a sample-by-sample basis. This information can improve the trained machine learning model and prevents overfitting. More importantly, our method can identify outliers at a per-sample granularity and can thus serve as a feedback for experiment automation and manual checkup. (Received March 03, 2020)