Contribution ID: **64**   Type: **not specified**

Hunting for signals using Gaussian Process regression

Friday 4 November 2022 15:00 (20 minutes)

We present a novel computational approach for extracting weak signals, whose exact location and width may be unknown, from complex background distributions with an arbitrary functional form. We focus on datasets that can be naturally presented as binned integer counts, demonstrating our approach on the datasets from the Large Hadron Collider. Our approach is based on Gaussian Process (GP) regression - a powerful and flexible machine learning technique that allowed us to model the background without specifying its functional form explicitly, and to separate the background and signal contributions in a robust and reproducible manner. Unlike functional fits, our GP-regression-based approach does not need to be constantly updated as more data becomes available. We discuss how to select the GP kernel type, considering trade-offs between kernel complexity and its ability to capture the features of the background distribution. We show that our GP framework can be used to detect the Higgs boson resonance in the data with more statistical significance than a polynomial fit specifically tailored to the dataset. Finally, we use Markov Chain Monte Carlo (MCMC) sampling to confirm the statistical significance of the extracted Higgs signature.

Primary author: GANDRAKOTA, Abhijith (Fermi National Accelerator Lab. (US))

Co-authors: MOROZOV, Alexandre; LATH, Amitabh (Rutgers, The State University of New Jersey); MURTHY, Sindhu (Carnegie-Mellon University (US))

Presenter: GANDRAKOTA, Abhijith (Fermi National Accelerator Lab. (US))

Session Classification: Anomaly Detection