Canadian Association of Physicists
Association canadienne des physiciens et physiciennes

Contribution ID: **4237** Type: **Oral Competition (Graduate Student) / Compétition orale (Étudiant(e) du 2e ou 3e cycle)**

(G*) Extracting Raman-like signals from CARS spectra with gradient boosting decision trees

Tuesday 28 May 2024 11:15 (15 minutes)

Coherent anti-Stokes Raman scattering (CARS) is a nonlinear optical process that is used for spectroscopy and imaging. The stimulated CARS signal is orders of magnitude stronger than in spontaneous Raman scattering, enabling CARS to achieve substantially faster acquisition speeds. This has positioned CARS as a desirable alternative to spontaneous Raman scattering as a contrast mechanism for chemical imaging. However, CARS suffers from the presence of a so-called non-resonant background (NRB) that distorts peak shapes and intensities, thus hindering the broader adoption of this powerful technique. The NRB makes quantitative analysis of CARS spectra nontrivial and reduces image contrast. NRB removal techniques that retrieve Raman-like signals from CARS spectra have thus become a central focus of the CARS literature. We present an original and accessible approach to NRB removal based on gradient boosting decision trees.

Gradient boosting decision trees are increasingly being used to win machine learning competitions, demonstrating their potential to compete with neural networks. Here, we apply the open-source gradient boosting framework XGBoost to NRB removal. A dataset of 100,000 stochastically generated CARS (input) and Raman-like (label) spectra was used for the training of the decision trees with a train-validation split of 80/20, while a dataset of 1000 independently generated pairs of spectra was used for testing. After hyperparameter tuning, the best decision tree yielded a Pearson correlation coefficient of r=.97 (p<.001) between retrieved and ground-truth Raman-like spectra, corresponding to a mean squared error (MSE) of 0.00047. When the trained model is applied to experimental CARS spectra obtained from samples with well-known Raman peaks, the model reproduces all of the expected Raman peaks for each of the samples that were tested. Our results establish gradient boosting decision trees as an effective tool for CARS NRB removal in lieu of neural networks.

Keyword-1

CARS

Keyword-2

NRB removal

Keyword-3

XGBoost

Author: SHAFE-PURCELL, John (Trent University)

Co-author: SLEPKOV, Aaron (Trent University)

Presenter: SHAFE-PURCELL, John (Trent University)

Session Classification: (DAMOPC) T1-8 Fundamental Physics with Atomic Systems | Physique fondamentale avec les systèmes atomiques (DPAMPC)

Track Classification: Technical Sessions / Sessions techniques: Atomic, Molecular and Optical Physics, Canada / Physique atomique, moléculaire et photonique, Canada (DAMOPC-DPAMPC)