****

臺灣大學應用力學研究所

演 講 公 告

****

主 講 人：徐振哲教授
臺灣大學化學工程學系

講 題：機器學習在電漿光譜進行有機物分析的應用

主 持 人： 舒貽忠教授

時 間： 111年3月7日（星期一）下午2時20分開始

地 點： 臺灣大學應用力學研究所國際會議廳

☆☆ 歡 迎 聽 講 ，敬 請 張 貼 ☆☆

Machine Learning for Volatile Organic Compounds Identification Using Plasma Spectroscopy

Cheng-che Hsu

Department of Chemical Engineering, National Taiwan University, Taiwan

chsu@ntu.edu.tw

In this talk, I will start with my own story how a problem I have been struggling for years can be solved rather easily by a researcher in the different field, which taught me the value of interdisciplinary collaboration.

The main theme of this presentation is sharing our recent progress using machine learning (ML) for analysis of optical emission spectroscopy (OES) of plasmas. The major effort is the assessment of multiple ML algorithms for plasma spectroscopy analysis. A dielectric-barrier- discharge- (DBD-) type microplasma generation device (MGD) is utilized as the testing system. OES emanating from plasmas ignited in ambient containing Ar and volatile organic compound (VOC) mixtures are utilized. The goal is to use plasma emission to predict the type of VOCs. Four VOCs, namely ethanol, acetone, isopropyl alcohol, and methanol, with various concentrations are tested. Algorithms assessed include PCA, linear discriminant analysis (LDA), supporting vector machine (SVM), convolution neural network (CNN), random forest, and decision tree. PCA is an unsupervised ML algorithm while the rest are supervised ones. Results show the great potential using ML with proper algorithm for spectroscopy analysis. Due to rather large variance of the emission spectra for VOCs of different concentrations, PCA is unable to capture the variance for different VOCs. LDA, SVM, and CNN are able to precisely classify VOC-type with an accuracy well above 95%. Random forest and decision tree, however, perform well only in the training stage but not in the testing data. Key spectral features for VOC classification are extracted using above-mentioned algorithms. It is shown that C_2 emission, as expected, is one of the most important feature for classification. We also observed that H emission is an important feature while CH emission is relatively unimportant. This is in contrast with those reported in the literature. Such observations demonstrate that ML and human knowledge complement each other.

To test transfer learning (TL) capability and potential, we take the CNN parameters obtained for a very different plasma system, plasmas in solution, and utilize as the initial condition to fine tune the CNN for VOC prediction. It is promising that the TL scheme is able to transfer the trained parameters across different plasma systems with very different feature and feature dimensions. Such a TL scheme provide a promising strategy reducing the required data amount, which can be used in cases where obtaining a large amount of data is difficult. Finally, I will discuss the challenges, opportunities, and limitation using ML for plasma spectroscopy analysis based on our current observation.