SIMS-24
24th International Conference on Secondary Ion Mass Spectrometry
8-13 September 2024
La Rochelle, France
www.sims-24.com

Data mining from rich SIMS data using machine learning

S. Aoyagi
Seikei University - Tokyo (Japon)
aoyagi@st.seikei.ac.jp

Secondary ion mass spectrometry (SIMS) provides so rich information that most of the information from SIMS data can be left unused. Multivariate analysis helps understanding complex SIMS data and recently a variety of machine learning methods have been applied to complicated SIMS data. Machine learning is generally classified three types, unsupervised, supervised and reinforcement learning. In this presentation, examples of the application of unsupervised and supervised learning methods are introduced. Unsupervised learning methods, such as principal component analysis (PCA), non-negativity matrix factorization (NMF) and autoencoder, are generally useful for extracting important features of a sample. Autoencoder, which is based on artificial neural networks (ANNs), is effective for non-linear data containing matrix effects. In terms of quantitative analysis, correction of matrix effects is essential. ANN-based methods are suitable for managing non-linear SIMS data. Feature extraction from organic and biological samples using autoencoder are compared with PCA and NMF in this presentation [1]. In addition, quantitative and qualitative analyses of organic SIMS data using an ANN-based supervised learning method was developed through Versailles Project on Advanced Materials and Standards (VAMAS) TWA2 A31 [2]. The identification of the peaks in complicated SIMS spectra is essential for understanding samples. A material prediction system using Random Forest (RF), which is a supervised learning method based on decision trees, has been developed for organic SIMS spectra by modifying the peptide prediction system developed through VAMAS TWA2 A26 [3].

References

[1] S. Aoyagi, K. Matsuda, Rapid Commun. Mass Spectrom., 37(4), e9445 (2023).
[2] S. Aoyagi et al., Anal. Chem., 95(40), 15078-15085 (2023).
[3] S. Aoyagi et al., Analytical Chemistry, 93, 9, 4191-4197 (2021).