Artificial Intelligence in UV-Visible Spectroscopy

M.Prathibha Bharathi*,
G.Manisha Reddy, M.Mythili, K. Shashikiran Reddy

Gokaraju Rangaraju College of Pharmacy
*Corresponding email id : prathibha8086@grcp.ac.in

ABSTRACT:

Since 2010, there has been a remarkable surge in the adoption of artificial intelligence (AI) methods for data analysis and modeling, particularly within the field of analytical chemistry. This growth has facilitated more efficient, accurate, and insightful solutions to complex problems, enabling better interpretation of chemical data, pattern recognition, and prediction. AI approaches, including machine learning algorithms and deep learning networks, have significantly advanced analytical techniques, leading to enhanced capabilities in areas such as spectral analysis, chemical sensing, and process optimization. The integration of AI in analytical chemistry continues to reshape the landscape, offering innovative ways to analyze and model chemical data with increased precision and speed. AI, when properly applied and evaluated using domain knowledge, is a useful tool for routine processing of large quantities of data and for assisting humans in performing routine or repetitive tasks with high repeatability and accuracy. AI also has significant potential for automating well-understood complex tasks, and for advancing discovery of new molecules and materials by integrating and combining known physical science principles with new innovative concepts and approaches. Neural networks are one of the many techniques used in AI, but they are not the only ones. AI encompasses a wide range of techniques and approaches, including rule-based systems, decision trees, genetic algorithms, and others.

Keywords: Artificial Intelligence, Neural networks, UV-Visible spectroscopy, Algorithms

INTRODUCTION:

UV Visible spectroscopy plays a vital role in various scientific and industrial applications, facilitating the analysis of materials based on their spectral characteristics. In recent years, the integration of image processing techniques and machine learning algorithms has emerged as a promising approach to enhance the accuracy and efficiency of spectroscopy. In fact, Chemometrics as an AI-based method, along with machine learning and deep learning (DL), a subfield of machine learning (ML). AI methods lead to an improved understanding of data. AI, when used to its full extent, it allows improved insight into the various processes, interactions, and characteristics of sample analysis data. AI-based techniques have been applied to chemical data since the 1970s. The paper specifically describes inverse modeling, preprocessing methods, and data modeling techniques as applied to spectral and image analysis for several analytical techniques.

Artificial Intelligence in Spectroscopy

1. AI in Vibrational Spectroscopy

Vibrational spectroscopy techniques, such as Fourier transform infrared (FT-IR) and Raman spectroscopy, have been demonstrated to be highly successful analytical methods and have proven to be low cost, require minimum sample preparation, be non-destructive, and yield highly valuable chemical information (1,2). There is a requirement for improvement and automation in data processing for classification and visualization of spectrochemical data measured from biological and medical samples. Multivariate classification is useful for assigning unknown samples to predefined groups and for information extraction from data. For this purpose, a tutorial article, representing the prerequisite algorithm requirements for machine learning approaches, has been published to demonstrate multivariate classification analysis using near-infrared (NIR), FT-IR, and Raman spectroscopy data (3).

Data preprocessing, data selection, feature extraction, classification, and model validation. These mathematical processes are essential for computation of practical spectrochemical classification models intended for analysis of biological samples. A paper demonstrated that convolutional neural networks (CNNs), a basic AI and ML tool, can be used to effectively classify vibrational spectroscopy data as well as to identify important spectral regions (4). CNNs are capable of image classification, and can be used to learn or define various visual interpretations of spectroscopic data. CNNs are able to reduce the requirements for rigorous data preprocessing, and are able to identify the most important spectral regions for analyzing the desired spectral features of interest. Spectroscopic data is most often preprocessed as part of the chemometric modeling work. Such preprocessing steps include various types of baseline correction, scatter correction, noise reduction or removal, feature selection, derivative processing, and other steps, which are applied to the spectra prior to model building. As a standard practice, preprocessing has traditionally been considered essential for optimizing standard chemometric models. A simple CNN architecture, using a single convolutional layer (shallow CNN), has demonstrated improved performance as compared to standard classification algorithms such as partial least squares (PLS) regression

2. **General Spectroscopic Applications of AI :**

Early work in implementing AI for spectroscopic applications described the identification of polyatomic molecules from their molecular spectra (5). This research concluded that AI systems provided one of the most promising potential developments for the future of analytical molecular spectroscopy, noting that AI systems presented a more spectroscopic methods. This research aims to leverage such advancements by developing a methodology for visible spectroscopy using polynomial regression and machine learning algorithms. Artificial intelligence (AI) comprises a broad spectrum of technologies and techniques aimed at enabling machines to mimic human cognitive functions, learn from data, and make autonomous decisions. Within the realm of AI, machine learning (4) is a pivotal subfield that empowers systems to improve their performance on tasks through experience.

AI is integrated with quantum chemistry and statistical mechanics for spectroscopic studies. The investigation of medium size to large size chromophores in their condensed phase by using UV-vis spectroscopy, a prototypical nitroxide radical (TEMPO) was evaluated by using different solvents. The paper explores the reliability, effectiveness, and robustness of the new

AI platform for application to complex UV-vis spectroscopic studies. One of the notable technique within machine learning is polynomial regression, which is one of type of regression analysis where the relationship between the independent variable (or variables) and the dependent variable

AI usage is increasing rapidly for analytical chemistry applications. are a useful for exploring in the growing field. This review discusses about the application of advancing computer processing power and algorithms of AI in chemistry for creating new compounds and to provide novel approaches in organic synthesis, drug discovery, and material design (5) . The use of AI to support analytical purposes has been mostly limited to data analysis methods like image recognition, vibrational spectroscopy, and mass spectrometry. AI technology to spectroscopy and separation science as well as sensor output and design.

It was reported that AI systems are able to compare measured spectra to computer-generated spectra, rather than just compare measured spectra to an archived spectral database. Such an AI approach offers more potential than manual methods to extract additional information from spectral analysis experiments. AI expert system for spectral interpretation and for the elucidation of chemical structures from combined spectroscopic data (EXSPEC) (6–9). The rationale and programming involved in creating such a system for interpretation of infrared, mass, and nuclear magnetic resonance (NMR) spectra for an earlier generation of computer systems. In research to better connect Raman spectroscopy measurements of biomedical and clinical samples to medical diagnosis and decision-making, an AI system was designed as a way to understand the chemical pathways and the progression of disease (10). This approach has potential use for the future of personalized medicine. The initial stages of this AI system were designed to optimize spectral preprocessing in order to reduce noise, fluorescence and baseline

Specifically, this research analyzed breast cancer tissue samples using Raman spectroscopy combined with principal component (PCA) and linear discriminate analysis (LDA). For this work, tissue microarray (TMA) breast biopsies were classified into multiple groups, such as luminal A, luminal B, HER2, and triple negative subtypes. The classification accuracy results of applying the AI system to the Raman spectra, which contains chemical information on tissue lipid, collagen, and nucleic acid content, were 70%, 100%, 90%, and 96.7% for luminal A, luminal B, HER2, and triple negative subtypes, respectively. This result is quite encouraging for further disease tissue classification research.

Because vibrational spectroscopic techniques have shown significant progress in assessing biomedical samples, they are being further investigated using automated and AI data processing. When properly implemented, AI is capable of identifying and connecting meaningful chemical and diagnostic relationships for clinical samples measured using infrared or Raman spectroscopy. A special article assesses the current use of AI in biomedical vibrational spectroscopy applications (11).

3. AI in Raman Spectroscopy :

ML, as a subfield of AI, is being used to extract, connect, and summarize information from large and complex analytical datasets for separation science, mass spectrometry, NMR, and atomic and molecular spectroscopy. The advancing of applications of AI technology is certainly the case in Raman and surface-enhanced Raman spectroscopy (SERS) techniques,

which involve large databases of complex vibrational spectra. Applying chemometric and analysis methods using manual techniques is no longer satisfactory for biomedical or diagnostic work. To implement accurate and powerful data analysis for Raman and SERS experiments, AI systems have recently been explored and implemented. A 2020 review article delves into the current uses of AI or machine (deep) learning techniques employed in Raman spectroscopy, including SERS (12).

An AI approach was applied to classify biomedical samples using Raman spectroscopy for the purposes of medical diagnosis and decision-making (13). The functional aspects of the AI system included automation of noise filtering (via a fuzzy controller), fluorescence background identification and correction, baseline optimization (using genetic algorithms), spectral normalization and scattering compensation (using the standard normal variate [SNV] algorithm), multivariate statistical analysis, sample data clustering, and projected decision-making. The AI program useful for classification of biomedical samples includes integration of fuzzy control, genetic algorithms, and principal component analysis (PCA), as well as system identification. The main goal for implementing the AI system is to incorporate a Raman probe on the end effectors of medical robots, to provide real-time information during robotic surgery (14), the application of Raman spectroscopy for noninvasive identification of mixture composition, including spectral preprocessing, and a detailed discussion of the application of AI analysis in Raman methods.

The use of ML in FT-Raman spectroscopy for trademark fingerprint assessment of fruit spirits (15). The combination of ML and Raman spectra was used in an attempt to compose reliable models for the classification of fruit spirits for trademark, geographical, and botanical origin. Raman spectroscopy is an excellent measurement technique for samples with high water content, such as alcoholic beverages, due to the relatively weak Raman water bending mode in the fingerprint region. The discriminant analysis for the optimized classification model. Geographical origin prediction was also achieved, but classification of botanical origin recognition was not successful. It was noted that spectral data used for the discriminant models involved both Stokes and anti-Stokes spectra, as well as individual spectral regions (windows). Several ML algorithm approaches were attempted. The paper offers an excellent discussion of the aspects of ML and Raman.

Raman spectroscopy provides an analytical method for non-destructively measuring the molecular structure for biomedical samples. Because of these benefits, Raman provides a suitable analytical technique for the assessment of inflammatory skin diseases and skin inflammation. The combination of AI with Raman spectroscopy to predict skin inflammation with high sensitivity and specificity was tested. For this feasibility work, Raman spectra, using excitation at 785 nm, were collected on mouse ear tissue where inflammation had been chemically induced. Principal component analysis (PCA) and AI were evaluated using receiver operating characteristic (ROC) curves to assess inflammation. The accuracy rate without AI implementation was 80.0% with an area under the curve (AUC) of 0.864. With AI implementation added, the accuracy rate was improved to 93.1% with an AUC of 0.972. The current findings demonstrate AI and Raman are able to provide highly accurate information on the presence and pathology of skin inflammation.

Raman spectroscopy has become known for its speed of analysis, limited requirements for sample preparation, ability to measure samples with high water content, and cost

effectiveness. With these features, Raman is quite suitable for analysis of foods and beverages, and has been studied relative to classification and authenticity of such products. Vegetable cold-pressed oils, which are of interest for nutritional health benefits, were tested for authentication analysis. In this study, oils from sesame, hemp, walnut, linseed, pumpkin, and sea buckthorn were tested for authentication using Raman spectra combined with ML. By using Raman spectra and ML, both adulteration and the degree of adulteration were successfully detected.

Improvements in chemometrics and data analysis techniques for Raman spectroscopy have allowed Raman to be used in more applications than ever before. However, there are still important issues to overcome when applying Raman spectroscopy that still limit its ability to fulfill its full analytical potential. These limitations include standardization of large datasets, intensity and baseline correction, fluorescence interference, measurement noise, changes in spectra relative to sample presentation variation, and so forth. In order to improve the quality and repeatability of Raman spectra, deep learning has been implemented as a strategy for enhanced correction of Raman spectral measurements.

4. AI in UV-vis Spectroscopy:

Ultraviolet-visible (UV-vis) spectroscopy is widely recognized as a workhorse for general analytical laboratory, process, and field measurements. Water quality monitoring is just one of the many applications where UV-vis exhibits its beneficial features, such as speed, environmental friendliness, non-destructive sample handling, and high analytical sensitivity. Water quality measurements using UV-vis for process or field applications have been quite common over the years. The application of ML to UV-vis spectroscopy provides an automated and “smart” platform for routine water analysis and screening.

Field water samples were collected, and spectra measured from 200 to 800 nm in a laboratory setting. Optimized CNN and PLS models were compared for water parameters resulting in an $R2$ for total organic carbon (TOC) between predicted values and reference values to be 0.927 for PLS and 0.953 for the CNN model. Also, $R2$ between predicted values and true values for total suspended solids (TSS) concentrations was 0.827 with the PLS model and 0.915 with the CNN model. It was concluded that CNN would be the preferred approach as compared to PLS for online water quality monitoring using UV–vis spectroscopy.

AI is being integrated with quantum chemistry and statistical mechanics for spectroscopic studies. One paper discusses the investigation of medium-to-large size chromophores in condensed phases using UV-visible spectroscopy, where a prototypical nitroxide radical (TEMPO) was evaluated in different solvents. The paper explores the reliability, effectiveness, and robustness of the new AI platform for application to complex UV-visible spectroscopic studies.

Summary:

Application of AI to spectroscopic studies, we are able to discern that this advancing field will continue to change the development and capabilities of spectroscopic data analysis and implementation well in the future. AI has great positive potential when used properly with expert domain knowledge, and also great potential for misuse when blindly applied or when input data is inadequate or not well understood.

References

1. Workman, J. J. M., Howard (2023). Artificial Intelligence in Analytical Spectroscopy, Part II: Examples in Spectroscopy. Spectroscopy, 38**,** 10-15.
2. Houhou, R. & Bocklitz, T. (2021). Trends in artificial intelligence, machine learning, and chemometrics applied to chemical data. Analytical Science .37,45-49.
3. Advances,2**,**128141.https://doi.org/10.1002/ansa.202000162.Available:https://chemistry-europe.onlinelibrary.wiley.com/doi/abs/10.1002/ansa.202000162.
4. Flores, Y. V., Polak, A., Jambet, J., Stothard, D. & Haertelt, M. (2023). Point of Interest Mid-Infrared Spectroscopy for Inline Pharmaceutical Packaging Quality Control. IEEE Sensors Journal, 23**,** 16115-16122.
5. Kong, X., Zhou, L., Li, Z., Yang, Z., Qiu, B., Wu, X., Shi, F. & Du, J. (2020). Artificial intelligence enhanced two-dimensional nanoscale nuclear magnetic resonance spectroscopy. npj Quantum Information, 6**,** 79-81.
6. Chen, H., Xu, L., Ai, W., Lin, B., Feng, Q. & Cai, K. (2020). Kernel functions embedded in support vector machine learning models for rapid water pollution assessment via near-infrared spectroscopy. Science of The Total Environment,24, 714-758.
7. Galán-Freyle, N. J., Ospina-Castro, M. L., Medina-González, A. R., Villarreal-González, R., Hernández-Rivera, S. P. & Pacheco-Londoño, L. C. (2020). Artificial Intelligence Assisted Mid-Infrared Laser Spectroscopy In Situ Detection of Petroleum in Soils. Applied Sciences, 10**,** 1319.
8. Contreras, J. & Bocklitz, T. (2024). Explainable artificial intelligence for spectroscopy data: a review. Pflügers Archiv - European Journal of Physiology. 41,56-59.
9. Lussier, F, Thibault, V, Charron, B, Wallace, G. Q, Masson, J. F. (2020) Deep Learning and Artificial Intelligence Methods for Raman and Surface-Enhanced Raman Scattering. TrAC, Trends Anal. Chem. 124,1036-1039.
10. Ye, Z. (2005) Artificial-Intelligence Approach for Biomedical Sample Characterization Using Raman Spectroscopy. IEEE Trans. Autom. Sci. Eng.
11. Pan, L, Zhang, P, Daengngam, C. Peng, S, Chongcheawchamnan, M. A (2022) Review of Artificial Intelligence Methods Combined with Raman Spectroscopy to Identify the Composition of Substances. J. Raman Spectrosc. 53 (1), 6–19.
12. Magdas, D. A, David, M, Berghian-Grosan, C. (2022). Fruit Spirits Fingerprint Pointed Out Through Artificial Intelligence and FT-Raman Spectroscopy. Food Control. 133,
13. Kanemura, Y, Kanazawa, M, Hashimoto, S. (2022) Assessment of Skin Inflammation Using Near-Infrared Raman Spectroscopy Combined with Artificial Intelligence Analysis in an Animal Model. Analyst 2022, 12. DOI: 10.1039/D2AN00193D
14. Berghian-Grosan, C, Magdas, D. A. (2020) Raman Spectroscopy and Machine-Learning for Edible Oils Evaluation. Talanta 218,
15. Jinadasa, M. W. N, Kahawalage, A. C, Halstensen, M, Skeie, N. O, Jens, K. J. (2022)Deep Learning Approach for Raman Spectroscopy. Recent Developments in Atomic Force Microscopy and Raman Spectroscopy for Materials Characterization . 5, 77-81.