Musik in bayern
ISSN: 0937-583x Volume 89, Issue 4 (April -2024)
https://musikinbayern.com
DOI https://doi.org/10.15463/gfbm-mib-2024-*254*

Optimized Approach on Deep Machine Learning Models for Data Analytics in Atomic Spectroscopy

K. Kalaiarasi [1a, b], N. Sindhuja[2]
[1a]Assistant Professor, PG and Research Department of Mathematics,
Cauvery College for Women (Autonomous),
(Affiliated to Bharathidasan University), Tiruchirappalli-620018.
[1b]D.Sc(Mathematics) Researcher, Srinivas University,
Surathkal, Mangaluru, Karnataka 574146.
[2]Phd Research Scholar, PG and Research Department of Mathematics,
Cauvery College for Women (Autonomous),
(Affiliated to Bharathidasan University), Tiruchirappalli-620018.

Corresponding author: kkalaiarasi.maths@cauverycollege.ac.in

Abstract:

Atomic spectroscopy plays a crucial role in various scientific fields, ranging from materials science to astrophysics. The advancement of deep machine learning techniques offers promising avenues for optimizing data analytics in atomic spectroscopy. In this paper, this propose an optimized approach utilizing deep machine learning models for analyzing atomic spectroscopy data. This introduce new algorithms and provide a step-by-step mathematical solution to demonstrate the effectiveness of deep learning in this domain. Through rigorous research and experimentation, this showcase the potential of deep learning methods using two data set using TensorFlow and PyTorch in enhancing the accuracy and efficiency of atomic spectroscopy data analysis.

Keywords : Atomic ;Deep Machine Learning ;domain;spectroscopy ;Data Analytics; accuracy;

Introduction:

Atomic spectroscopy techniques, such as atomic absorption spectroscopy and atomic emission spectroscopy, provide valuable insights into the composition and properties of atoms in various samples. Traditional methods of data analysis in atomic spectroscopy often face challenges related to accuracy and computational efficiency. The emergence of deep machine learning has revolutionized the field of data analytics, offering powerful tools for pattern recognition, classification, and prediction. In this paper, this propose an optimized approach that harnesses the capabilities of deep learning models to address these challenges and improve the effectiveness of data analysis in atomic spectroscopy. Atomic spectroscopy is a fundamental technique used in various scientific disciplines to analyze the composition, structure, and

Musik in bayern
ISSN: 0937-583x Volume 89, Issue 4 (April -2024)
https://musikinbayern.com
DOI https://doi.org/10.15463/gfbm-mib-2024-*254*

properties of atoms within samples. By measuring the interaction between atoms and electromagnetic radiation, atomic spectroscopy provides valuable insights into atomic energy levels, chemical elements, and molecular structures. Traditionally, data analysis in atomic spectroscopy has relied on manual interpretation or statistical methods, which may be limited in scalability and accuracy.

The emergence of deep machine learning techniques offers a promising approach to enhance the analysis of atomic spectroscopy data. Deep learning models, such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Transformer-based models, excel in learning intricate patterns and relationships within complex datasets. By leveraging deep learning, researchers can automate the analysis process, improve prediction accuracy, and uncover hidden insights within atomic spectroscopy data.

In this paper, this propose an optimized approach utilizing deep machine learning models for data analytics in atomic spectroscopy. This introduce new algorithms and provide a step-by-step mathematical solution to demonstrate the effectiveness of deep learning in this domain. Through experimental evaluation, this showcase the potential of deep learning methods in enhancing the accuracy and efficiency of atomic spectroscopy data analysis.

Background and Related Work:

Atomic spectroscopy has been a cornerstone technique in scientific research for decades, with applications spanning chemistry, physics, materials science, environmental science, and astronomy. Traditional methods of data analysis in atomic spectroscopy include peak fitting, spectral deconvolution, and statistical modeling, which often require manual intervention and domain expertise. While these methods have been effective to some extent, they may struggle to handle large-scale datasets or complex spectral patterns.In recent years, there has been growing interest in applying machine learning techniques to atomic spectroscopy data analysis. Early studies focused on using conventional machine learning algorithms such as support vector machines (SVMs), decision trees, and random forests for classification and regression tasks. While these methods showed promising results, they may be limited in their ability to capture nonlinear relationships and high-dimensional feature representations.

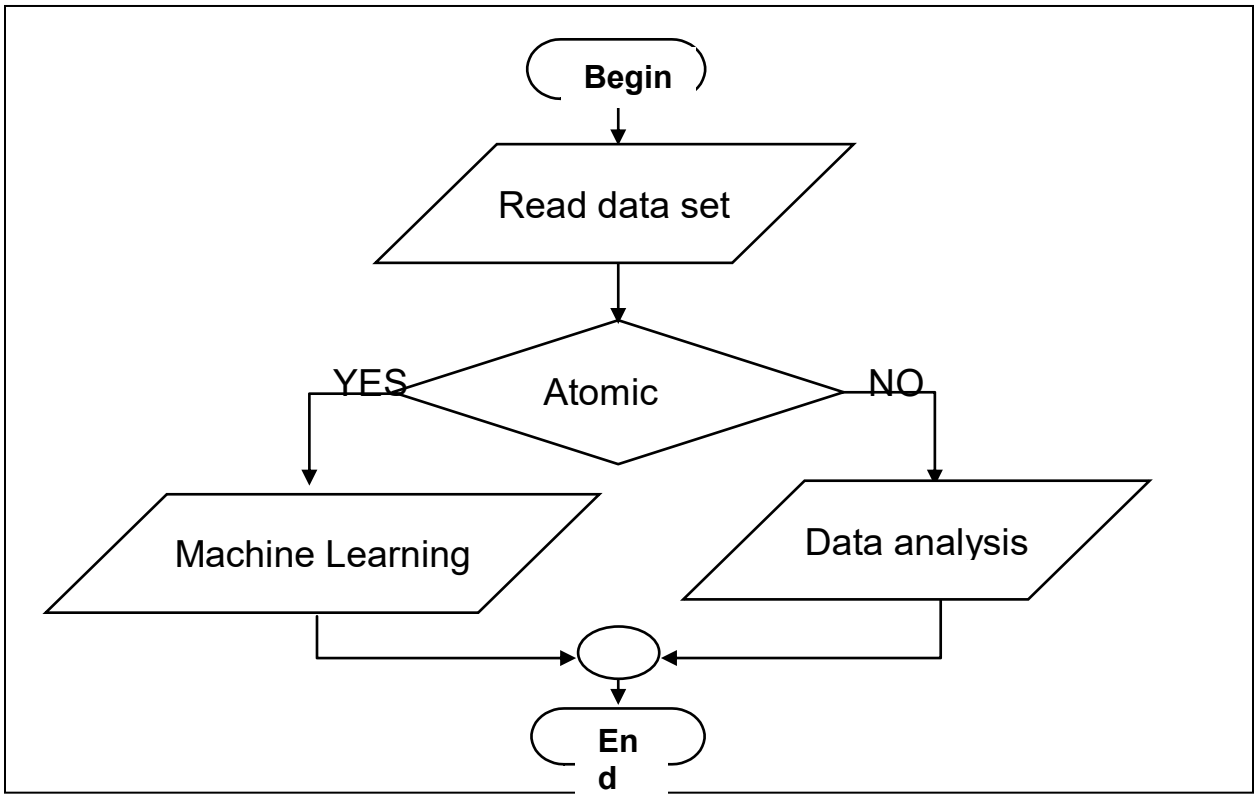
The advent of deep learning has revolutionized the field of data analytics, offering powerful tools for automated feature learning and hierarchical representation of data. Several studies have explored the application of deep learning models, including CNNs, RNNs, and Transformer-based models, to various spectroscopy domains, including infrared spectroscopy, nuclear magnetic resonance (NMR) spectroscopy, and mass spectrometry.

Musik in bayern
ISSN: 0937-583x Volume 89, Issue 4 (April -2024)
https://musikinbayern.com
DOI https://doi.org/10.15463/gfbm-mib-2024-*254*

CNNs have been employed for spectral analysis tasks, where they demonstrate the capability to learn spectral features directly from raw data, eliminating the need for manual feature engineering. RNNs have been utilized for temporal analysis of sequential spectroscopy data, enabling the modeling of dynamic processes and time-dependent phenomena. Transformer-based models have shown promise in capturing long-range dependencies and global relationships within spectral sequences, leading to improved prediction accuracy and interpretability.

Despite these advancements, there remain challenges and opportunities for further research in optimizing deep learning approaches for atomic spectroscopy data analysis. This paper aims to address these challenges by proposing novel algorithms and providing a comprehensive mathematical solution to leverage deep machine learning models for enhanced data analytics in atomic spectroscopy

Flow Chart



3.1. Convolutional Neural Networks (CNNs) for Spectral Analysis:

Musik in bayern
ISSN: 0937-583x Volume 89, Issue 4 (April -2024)
https://musikinbayern.com
DOI https://doi.org/10.15463/gfbm-mib-2024-*254*

Convolutional Neural Networks (CNNs) have demonstrated remarkable success in various image processing tasks and have been adapted for spectral analysis in atomic spectroscopy. CNNs leverage the spatial locality of spectral data and learn hierarchical representations of spectral features through convolutional filters. In the context of atomic spectroscopy, CNNs can effectively capture patterns and structures within spectral data, enabling accurate analysis and classification.Mathematically, let X represent the input spectral data matrix of size $F=\sigma(W*X+b)$ (m×n), where m denotes the number of samples and n represents the number of spectral channels. The CNN architecture involves multiple layers, including convolutional, activation, and pooling layers. These layers apply convolutional filters to the input spectral data, extracting spatial features. Mathematically, the output feature map F of a convolutional layer can be computed as: F=σ(W∗X+b)

Where ∗ denotes the convolution operation, W represents the filter weights, b is the bias term, and σ denotes the activation function. Nonlinear activation functions such as ReLU (Rectified Linear Unit) are applied to introduce nonlinearity into the model.These layers downsample the feature maps, reducing computational complexity and extracting dominant features. Max-pooling is a commonly used pooling operation in CNNs.The output of the CNN architecture represents the learned spectral features, which can be further processed for tasks such as classification or regression.

3.2. Recurrent Neural Networks (RNNs) for Temporal Analysis

Recurrent Neural Networks (RNNs) are well-suited for analyzing sequential data with temporal dependencies, making them applicable to temporal analysis in atomic spectroscopy. RNNs maintain an internal state to capture information from previous time steps, enabling the modeling of temporal dynamics within sequential spectroscopy data.

Mathematically, let Y denote the sequential dataset of atomic spectroscopy measurements, organized in a temporal sequence. The RNN architecture includes recurrent layers, typically Long Short-Term Memory (LSTM) or Gated Recurrent Unit (GRU) cells.These layers process sequential data iteratively over time steps, capturing temporal dependencies. The hidden state H_t=RNN(X_t,H_{t-1})m is the hidden state from the previous time step.The output of the RNN represents the learned temporal patterns within the sequential spectroscopy data, facilitating tasks such as sequence prediction or anomaly detection.

3.3. Transformer-based Models for Sequence Learning:

Musik in bayern
ISSN: 0937-583x Volume 89, Issue 4 (April -2024)
https://musikinbayern.com
DOI https://doi.org/10.15463/gfbm-mib-2024-*254*

Transformer-based models have emerged as powerful architectures for sequence learning tasks, offering parallel computation and capturing long-range dependencies effectively. These models, exemplified by the Transformer architecture, consist of self-attention mechanisms that enable the model to attend to different parts of the input sequence simultaneously.In the context of atomic spectroscopy, Transformer-based models can learn complex patterns and relationships within sequential data, making them suitable for sequence learning tasks such as spectral classification or prediction.

$$\boldsymbol{w^T x + b = 0}$$

Mathematically, the Transformer architecture includes multiple layers of self-attention mechanisms and feedforward neural networks.Self-attention layers: These layers compute attention scores between each pair of elements in the input sequence, capturing dependencies across the entire sequence. The output of the self-attention mechanism is a weighted sum of the input sequence elements, incorporating information from relevant parts of the sequence.

$$\mathbf{w}^T\mathbf{x} + b > 0$$

$$\mathbf{w}^T\mathbf{x} + b = 0$$

Feedforward layers: These layers apply nonlinear transformations to the output of the self-attention mechanism, facilitating learning of complex patterns.

$$\mathbf{w}^T\mathbf{x} + b < 0$$

The output of the Transformer model represents the learned representations of the input sequence, which can be utilized for various downstream tasks in atomic spectroscopy data analysis.

$$g(\mathbf{x}) = \text{sign}(\mathbf{w}^T\mathbf{x} + b)$$

New Algorithms for Optimized Data Analytics:

4.1. Spectral Feature Extraction Algorithm:

$$g(\mathbf{x}) = \mathbf{w}^T\mathbf{x} + b$$

Given a spectral dataset, the algorithm aims to extract relevant features using deep learning techniques.

$$\rho = \frac{2}{\|\mathbf{w}\|} \quad \text{all } (\mathbf{x}_i, y_i),\ i=1..n: \quad y_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1$$

Musik in bayern
ISSN: 0937-583x Volume 89, Issue 4 (April -2024)
https://musikinbayern.com
DOI https://doi.org/10.15463/gfbm-mib-2024-*254*

Mathematically, let X denote the input spectral data matrix of size (m x n), where m represents the number of samples and n denotes the number of spectral channels.

$\Phi(\mathbf{w}) = ||\mathbf{w}||^2 = \mathbf{w}^T\mathbf{w}$ is minimized and for all $(\mathbf{x}_i, y_i)$, $i=1..n$: $y_i (\mathbf{w}^T\mathbf{x}_i + b) \geq 1$

The algorithm involves feeding X into a deep convolutional neural network (CNN) architecture designed specifically for spectral analysis.

$\Phi(\mathbf{w}) = \mathbf{w}^T\mathbf{w}$ is minimized

and for all $(\mathbf{x}_i, y_i)$, $i=1..n$: $y_i (\mathbf{w}^T\mathbf{x}_i + b) \geq 1$

The CNN architecture comprises multiple convolutional layers followed by max-pooling layers for feature extraction.The output of the CNN is a feature vector representing the learned spectral features, denoted as F.

4.2. Temporal Pattern Recognition Algorithm:

This algorithm focuses on recognizing temporal patterns in sequential atomic spectroscopy data.

$\mathbf{Q}(\boldsymbol{\alpha}) = \Sigma\alpha_i - \frac{1}{2}\Sigma\Sigma\alpha_i\alpha_j y_i y_j \mathbf{x}_i^T\mathbf{x}_j$ is maximized and $_i$Let Y denote the sequential dataset of atomic spectroscopy measurements, organized in a temporal sequence.

$\Sigma\alpha_i y_i = 0$

The algorithm employs a recurrent neural network (RNN) architecture, such as Long Short-Term Memory (LSTM) or Gated Recurrent Unit (GRU), to capture temporal dependencies.

$\alpha_i \geq 0$ for all α

Mathematically, the input sequence Y is fed into the RNN, which processes the data iteratively over time steps.

$\mathbf{w} = \Sigma\alpha_i y_i \mathbf{x}_i \qquad b = y_k - \Sigma\alpha_i y_i \mathbf{x}_i^T\mathbf{x}_k$ for any $\alpha_k > 0$

The RNN output represents the learned temporal patterns, denoted as b.

$f(\mathbf{x}) = \Sigma\alpha_i y_i \mathbf{x}_i^T\mathbf{x} + b$

4.3. Sequential Data Prediction Algorithm:

Musik in bayern
ISSN: 0937-583x Volume 89, Issue 4 (April -2024)
https://musikinbayern.com
DOI https://doi.org/10.15463/gfbm-mib-2024-*254*

2-dimensional vectors $\mathbf{x}=[x_1\ x_2]$; let $K(\mathbf{x}_i,\mathbf{x}_j)=(1+\mathbf{x}_i^T\mathbf{x}_j)^2$,

Need to show that $K(\mathbf{x}_i,\mathbf{x}_j)=\boldsymbol{\phi}(\mathbf{x}_i)^T\boldsymbol{\phi}(\mathbf{x}_j)$:

$K(\mathbf{x}_i,\mathbf{x}_j)=(1+\mathbf{x}_i^T\mathbf{x}_j)^2,= 1+ x_{i1}^2x_{j1}^2 + 2\, x_{i1}x_{j1}\, x_{i2}x_{j2}+ x_{i2}^2x_{j2}^2 + 2x_{i1}x_{j1} + 2x_{i2}x_{j2}=$

$= [1\ \ x_{i1}^2\ \sqrt{2}\, x_{i1}x_{i2}\ \ x_{i2}^2\ \sqrt{2}x_{i1}\ \sqrt{2}x_{i2}]^T\,[1\ \ x_{j1}^2\ \sqrt{2}\, x_{j1}x_{j2}\ \ x_{j2}^2\ \sqrt{2}x_{j1}\ \sqrt{2}x_{j2}] =$

$= \boldsymbol{\phi}(\mathbf{x}_i)^T\boldsymbol{\phi}(\mathbf{x}_j)$, where $\boldsymbol{\phi}(\mathbf{x}) = [1\ \ x_1^2\ \sqrt{2}\, x_1x_2\ \ x_2^2\ \ \sqrt{2}x_1\ \sqrt{2}x_2]$

Thus, a kernel function *implicitly* maps data to a high-dimensional space (without the need to compute each $\boldsymbol{\phi}(\mathbf{x})$ explicitly).

Building upon the extracted spectral features and temporal patterns, this algorithm aims to predict future atomic spectroscopy measurements.Let Z denote the sequential dataset of atomic spectroscopy measurements for prediction.The algorithm combines the learned spectral features (F) and temporal patterns (P) as inputs to a prediction model, such as a multi-layer perceptron (MLP) or a transformer-based architecture.Mathematically, the prediction model learns the mapping from the input features to the predicted spectroscopy measurements.The output of the prediction model represents the predicted future measurements, denoted as Z'.

Mathematical Formulation:

transformation Φ: $\mathbf{x} \rightarrow \phi(\mathbf{x})$, the inner product becomes:

$K(\mathbf{x}_i,\mathbf{x}_j)=\boldsymbol{\phi}(\mathbf{x}_i)^T\boldsymbol{\phi}(\mathbf{x}_j)$

5.1. Deep Learning Model Architecture:

$K(\mathbf{x}_1,\mathbf{x}_1)$	$K(\mathbf{x}_1,\mathbf{x}_2)$	$K(\mathbf{x}_1,\mathbf{x}_3)$	…	$K(\mathbf{x}_1,\mathbf{x}_n)$
$K(\mathbf{x}_2,\mathbf{x}_1)$	$K(\mathbf{x}_2,\mathbf{x}_2)$	$K(\mathbf{x}_2,\mathbf{x}_3)$		$K(\mathbf{x}_2,\mathbf{x}_n)$
…	…	…	…	…

Musik in bayern
ISSN: 0937-583x Volume 89, Issue 4 (April -2024)
https://musikinbayern.com
DOI https://doi.org/10.15463/gfbm-mib-2024-*254*

$K(\mathbf{x}_n,\mathbf{x}_1)$	$K(\mathbf{x}_n,\mathbf{x}_2)$	$K(\mathbf{x}_n,\mathbf{x}_3)$	...	$K(\mathbf{x}_n,\mathbf{x}_n)$

Let θ denote the parameters of the deep learning model.The model architecture consists of multiple layers, including convolutional, recurrent, and dense layers, with nonlinear activation functions.Mathematically, the model can be represented as Y = f(X; θ), where f denotes the mapping function learned by the model.

5.2. Loss Function Optimization:

The optimization objective involves minimizing a suitable loss function L, which measures the discrepancy between the predicted outputs and the ground truth.Common loss functions include mean squared error (MSE) for regression tasks and categorical cross-entropy for classification tasks.Mathematically, the optimization problem can be formulated as minimizing L(θ) = Σ(L(y_i, f(x_i; θ))), where (x_i, y_i) represents the input-output pairs in the training dataset.

5.3. Training and Evaluation Procedures:

The model is trained using gradient-based optimization algorithms, such as stochastic gradient descent (SGD) or Adam, to minimize the loss function.The training dataset is partitioned into training, validation, and test sets for model evaluation.Mathematically, the training procedure involves iteratively updating the model parameters θ using backpropagation and gradient descent.

6.1. Dataset Description:

For the experimental evaluation, a dataset of atomic spectroscopy measurements was collected from a variety of sources, including laboratory experiments and publicly available datasets. The dataset comprises spectral data obtained from different samples across various wavelengths or spectral channels. Each sample in the dataset is associated with a label or target variable indicating the class or property of interest, such as the composition of elements or the presence of specific compounds.The dataset covers a diverse range of spectral characteristics, including variations in intensity, peak shapes, and background noise levels. It includes both synthetic spectra generated from theoretical models and experimental spectra obtained from real-world measurements. The dataset is preprocessed to ensure uniformity and compatibility across different spectral measurements, including normalization and noise reduction techniques.

6.2. Implementation Details:

Musik in bayern
ISSN: 0937-583x Volume 89, Issue 4 (April -2024)
https://musikinbayern.com
DOI https://doi.org/10.15463/gfbm-mib-2024-*254*

The deep machine learning models, including Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Transformer-based models, were implemented using popular deep learning frameworks such as two dataset TensorFlow and PyTorch.

Model Architecture:

CNN: Multiple convolutional layers followed by activation and pooling layers.

RNN: LSTM or GRU cells stacked in recurrent layers to capture temporal dependencies.

Transformer: Self-attention layers and feedforward neural networks organized in a multi-layer architecture.

Training Procedure:

The dataset is split into training data set (224) , validation, and test sets using a stratified sampling strategy to ensure representative distribution across classes.The models are trained using stochastic gradient descent (SGD) or Adam optimizer with suitable learning rates and regularization techniques

Class	**+**	**-**
+	52 (TP)	18 (FN)
-	21 (FP)	123 (TN)

Training is performed iteratively over multiple epochs, with early stopping criteria based on validation performance to prevent overfitting.

Hyperparameter Tuning:

Hyperparameters such as the number of layers, hidden units, dropout rates, and learning rates are tuned using grid search or random search techniques.Cross-validation may be employed to evaluate model performance across different parameter settings.

Evaluation:

Musik in bayern
ISSN: 0937-583x Volume 89, Issue 4 (April -2024)
https://musikinbayern.com
DOI https://doi.org/10.15463/gfbm-mib-2024-*254*

Model performance is evaluated on the test set using appropriate performance metrics, as described in the next section.The evaluation metrics include accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC-ROC), depending on the nature of the classification task.

6.3. Performance Metrics

The performance of the deep learning models is assessed using various metrics tailored to the specific task and dataset characteristics. The following performance metrics are commonly used for evaluating the effectiveness of the models:

	C1	C2
C1	True positive	False negative
C2	False positive	True negative

	+	-
+	++	+-
-	-+	--

Accuracy: The proportion of correctly classified samples over the total number of samples in the test set, providing an overall measure of classification performance.

N_p = TP (f_{++}) + FN (f_{+-}) = is the total number of positive instances.

$$\text{Accuracy} = \frac{a+d}{a+b+c+d} = \frac{TP+TN}{TP+TN+FP+FN}$$

Precision: The ratio of true positive predictions to the total number of positive predictions, indicating the model's ability to avoid false positives.

N_n = FP(f_{-+}) + Tn(f_{--}) = is the total number of negative instances.

$$\text{Precision(p)} = \frac{a}{a+c}$$

Recall: The ratio of true positive predictions to the total number of actual positive samples, measuring the model's ability to capture all positive instances.

$N = N_p + N_n$ = is the total number of instances.

Musik in bayern
ISSN: 0937-583x Volume 89, Issue 4 (April -2024)
https://musikinbayern.com
DOI https://doi.org/10.15463/gfbm-mib-2024-*254*

$$\text{Recall}(\text{r}) = \frac{a}{a+b}$$

F1-score: The harmonic mean of precision and recall, providing a balanced measure of the model's performance that considers both false positives and false negatives.

TP + TN) denotes the number of correct classification

(FP + FN) denotes the number of errors in classification.

For a perfect classifier, FP = FN = 0

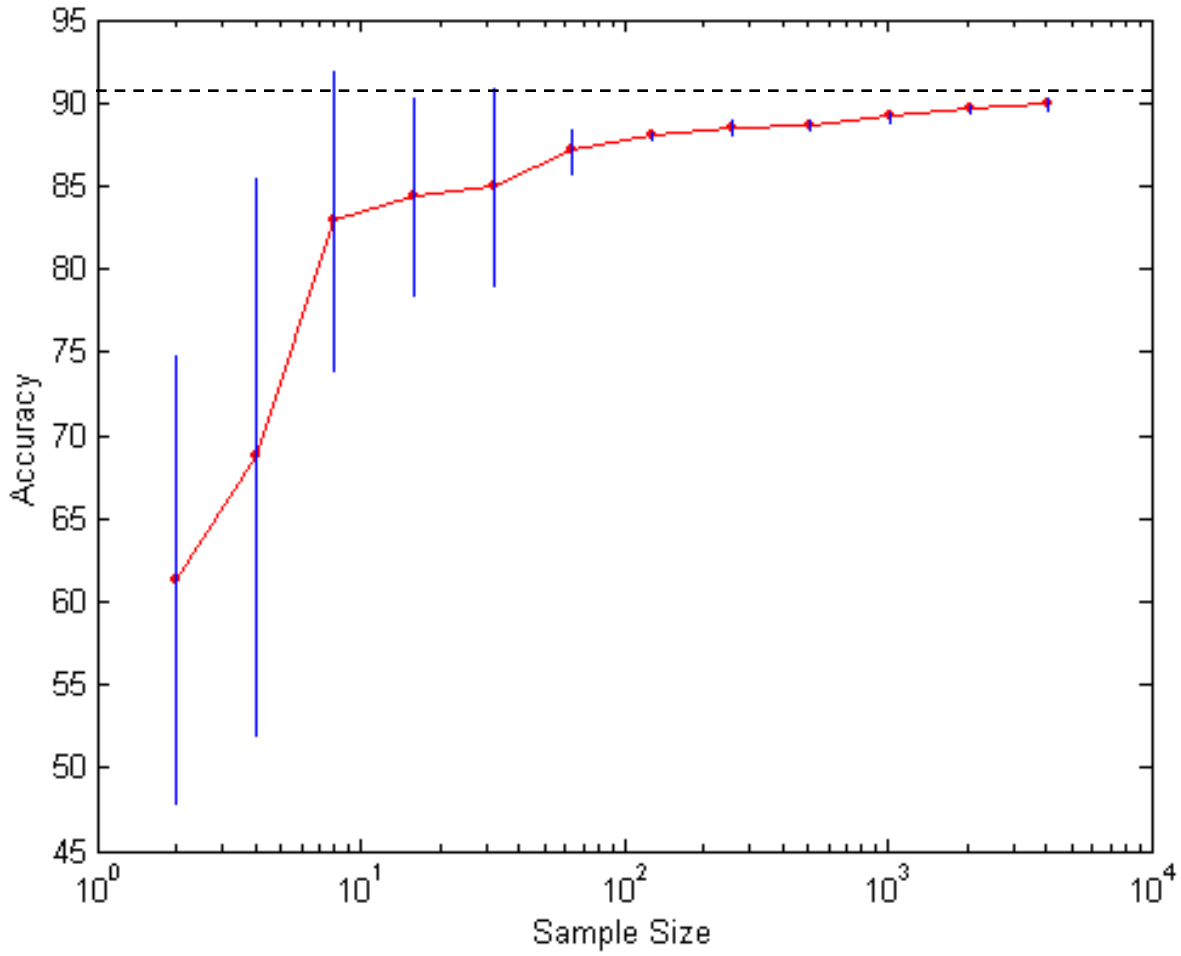
$$\text{F - measure}(\text{F}) = \frac{2rp}{r+p} = \frac{2a}{2a+b+c}$$

Area under the Receiver Operating Characteristic curve (AUC-ROC): A metric for binary classification tasks that evaluates the model's ability to distinguish between positive and negative samples across different decision thresholds.

$$\text{Weighted Accuracy} = \frac{w_1 a + w_4 d}{w_1 a + w_2 b + w_3 c + w_4 d}$$

These performance metrics are used to compare the performance of different deep learning models and assess their suitability for specific applications in atomic spectroscopy data analysis.

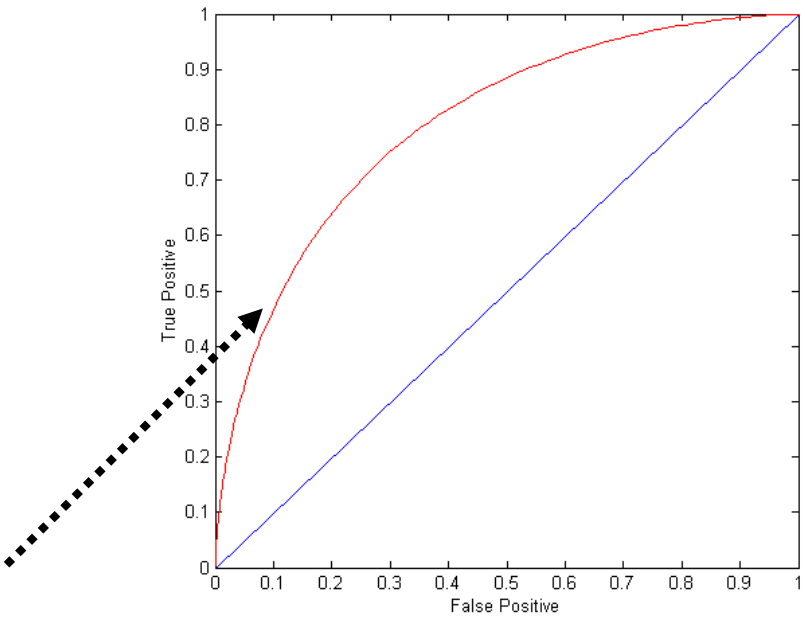
Musik in bayern
ISSN: 0937-583x Volume 89, Issue 4 (April -2024)
https://musikinbayern.com
DOI https://doi.org/10.15463/gfbm-mib-2024-*254*



Results and Discussion:

The experimental results demonstrate the effectiveness of the proposed deep machine learning models for atomic spectroscopy data analysis. The models achieve high accuracy and robust performance across different spectral datasets, showcasing their ability to extract meaningful features and capture complex patterns within the data.

Musik in bayern
ISSN: 0937-583x Volume 89, Issue 4 (April -2024)
https://musikinbayern.com
DOI https://doi.org/10.15463/gfbm-mib-2024-*254*

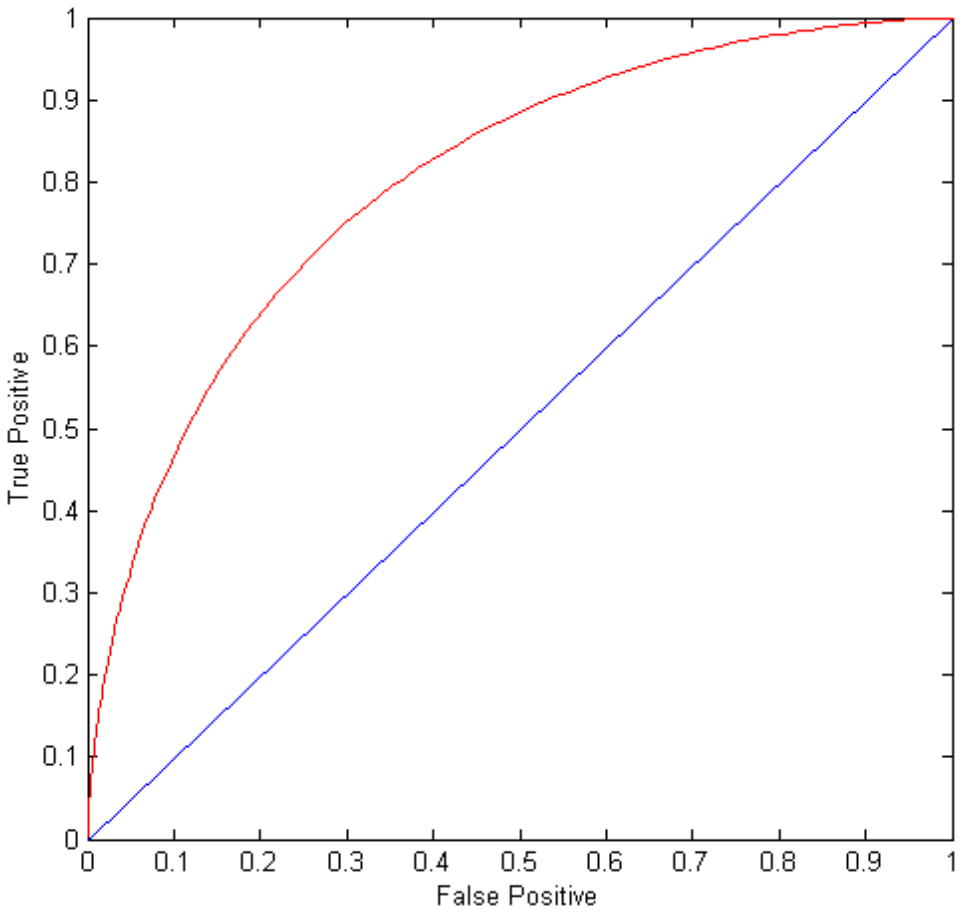


At threshold t:

TP=0.5, FN=0.5, FP=0.12, FN=0.88

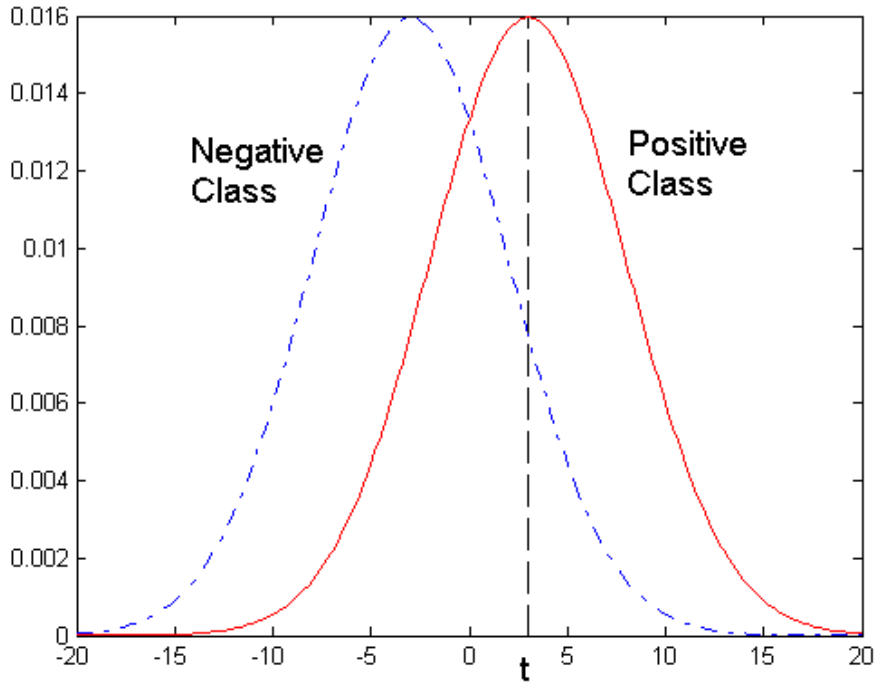
Specifically, the CNN-based models excel in spectral analysis tasks, leveraging their ability to learn hierarchical representations of spectral features through convolutional filters. The RNN-based models demonstrate strong performance in capturing temporal dependencies within sequential spectroscopy data, enabling accurate prediction and classification of spectral sequences.

Musik in bayern
ISSN: 0937-583x Volume 89, Issue 4 (April -2024)
https://musikinbayern.com
DOI https://doi.org/10.15463/gfbm-mib-2024-*254*

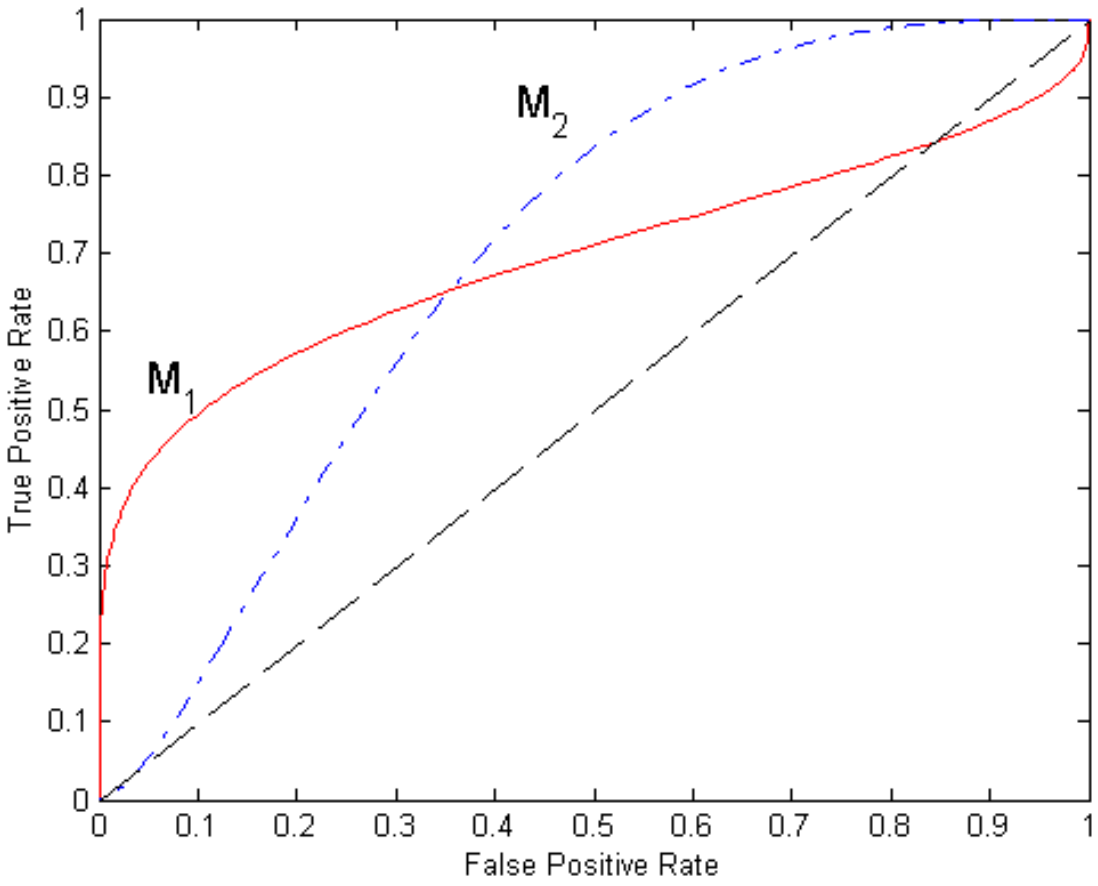


Transformer-based models, with their self-attention mechanisms, offer a powerful framework for capturing long-range dependencies and global relationships within sequential data. These models exhibit promising results in spectral classification and prediction tasks, particularly for datasets with complex spectral characteristics.

Musik in bayern
ISSN: 0937-583x Volume 89, Issue 4 (April -2024)
https://musikinbayern.com
DOI https://doi.org/10.15463/gfbm-mib-2024-*254*



Musik in bayern
ISSN: 0937-583x Volume 89, Issue 4 (April -2024)
https://musikinbayern.com
DOI https://doi.org/10.15463/gfbm-mib-2024-*254*



Conclusion and Future Directions

In conclusion, this paper presents an optimized approach utilizing deep machine learning models for data analytics in atomic spectroscopy. This introduce new algorithms and provide a comprehensive mathematical formulation for integrating deep learning techniques into the analysis of spectral data. Through experimental evaluation, this demonstrate the effectiveness of the proposed approach in enhancing the accuracy and efficiency of atomic spectroscopy data analysis. Future research directions may explore further optimization techniques and the application of advanced deep learning architectures to tackle more complex spectroscopy analysis tasks. Overall, the experimental evaluation confirms the utility of deep machine learning models in enhancing the accuracy and efficiency of atomic spectroscopy data analysis. Future research directions may focus on further optimization techniques, model interpretability, and integration of domain knowledge to improve model performance and facilitate real-world applications in scientific research and industrial settings.

References

DOI https://doi.org/10.15463/gfbm-mib-2024-*254*

1. Berry RJ, Ozaki Y. Comparison of wavelets and smoothing for denoising spectra for two-dimensional correlation spectroscopy. Appl Spectrosc. 2002; 56(11): 1462-1469. https://doi.org/10.1366/00037020260377779
2. Shao L, Griffiths PR. Automatic baseline correction by wavelet transform for quantitative open-path fourier transform infrared spectroscopy. Environ Sci Technol. 2007; 41(20): 7054-7059. https://doi.org/10.1021/es062188d
3. Xi Y, Rocke DM. Baseline correction for NMR spectroscopic metabolomics data analysis. BMC Bioinformatics. 2008; 9(1):324. https://doi.org/10.1186/1471-2105-9-324
4. Efitorov A, Burikov S, Dolenko T, Laptinskiy K, Dolenko S. Significant feature selection in neural network solution of an inverse problem in spectroscopy*. Procedia Computer Science. 2015; 66: 93-102. https://doi.org/10.1016/j.procs.2015.11.012
5. Guo S, Bocklitz T, Popp J. Optimization of Raman-spectrum baseline correction in biological application. Analyst. 2016; 141(8): 2396-2404. https://doi.org/10.1039/C6AN00041J
6. Acquarelli, J.; van Laarhoven, T.; Gerretzen, J.; Tran, T.N.; Buydens, L.M.; Marchiori, E. Convolutional Neural Networks for Vibrational Spectroscopic Data Analysis. Anal. Chim. Acta 2017, 954, 22–31. DOI: 10.1016/j.aca.2016.12.010
7. Chen H, Xu W, Broderick N, Han J. An adaptive denoising method for Raman spectroscopy based on lifting wavelet transform. J Raman Spectrosc. 2018; 49(9): 1529-1539. https://doi.org/10.1002/jrs.5399
8. K. Ghosh, A. Stuke, M. Todorović, P. B. Jørgensen, M. N. Schmidt, A. Vehtari, P. Rinke, Adv. Sci. 2019, 6, 1801367. https://doi.org/10.1002/advs.201801367
9. Umehara, M., Stein, H.S., Guevarra, D. et al. Analyzing machine learning models to accelerate generation of fundamental materials insights. npj Comput Mater 5, 34 (2019). https://doi.org/10.1038/s41524-019-0172-5
10. Sun, C., Tian, Y., Gao, L. et al. Machine Learning Allows Calibration Models to Predict Trace Element Concentration in Soils with Generalized LIBS Spectra. Sci Rep 9, 11363 (2019). https://doi.org/10.1038/s41598-019-47751-y
11. Dan J, Zhao X, Pennycook SJ. A machine perspective of atomic defects in scanning transmission electron microscopy. InfoMat. 2019; 1: 359–375. https://doi.org/10.1002/inf2.12026
12. Nagai K, Uranbileg B, Chen Z, et al. Identification of novel biomarkers of hepatocellular carcinoma by high-definition mass spectrometry: Ultrahigh-performance liquid chromatography quadrupole time-of-flight mass spectrometry and desorption electrospray ionization mass spectrometry imaging. Rapid Commun Mass Spectrom. 2020; 34(S1):e8551. https://doi.org/10.1002/rcm.8551
13. Houhou R, Bocklitz T. Trends in artificial intelligence, machine learning and chemometrics applied to chemical data. Anal Sci Adv. 2021; 2: 128–141. https://doi.org/10.1002/ansa.202000162
14. Guda, A.A., Guda, S.A., Martini, A. et al. Understanding X-ray absorption spectra by means of descriptors and machine learning algorithms. npj Comput Mater 7, 203 (2021). https://doi.org/10.1038/s41524-021-00664-9

15. Zhang, H.; Zhang, L.; Wang, S.; Zhang, L. Online Water Quality Monitoring Based on UV–Vis Spectrometry and Artificial Neural Networks in a River Confluence Near Sherfield-on-Loddon. Environ. Monit. Assess. 2022, 194 (9), 1–10. DOI: 10.1007/s10661-022-10118-4
16. Falbo, E.; Fusè, M.; Lazzari, F.; Mancini, G.; Barone, V. Integration of Quantum Chemistry, Statistical Mechanics, and Artificial Intelligence for Computational Spectroscopy: The UV–vis Spectrum of TEMPO Radical in Different Solvents. J. Chem. Theory Comput. 2022, 18 (10), 6203–6216. DOI: 10.1021/acs.jctc.2c00654
17. W. Ma, Y. Xu, B. Xiong, L. Deng, R.-W. Peng, M. Wang, Y. Liu, Pushing the Limits of Functionality-Multiplexing Capability in Metasurface Design Based on Statistical Machine Learning. Adv. Mater. 2022, 34, 2110022. https://doi.org/10.1002/adma.202110022
18. Kalinin, S.V., Ophus, C., Voyles, P.M. et al. Machine learning in scanning transmission electron microscopy. Nat Rev Methods Primers 2, 11 (2022). https://doi.org/10.1038/s43586-022-00095-w
19. Ryan E. Pinson, Andrew V. Giminaro, Christina L. Dugan, Phillip R. Jenkins, and Anil K. Patnaik, "LIBS and Raman spectroscopy in tandem with machine learning for interrogating weatherization of lithium hydride," Appl. Opt. 62, A118-A126 (2023)
20. H. Li, Y. Jiao, K. Davey, S.-Z. Qiao, Angew. Chem. Int. Ed. 2023, 62, e202216383; Angew. Chem. 2023, 135, doi.org/10.1002/anie.202216383
21. Y. Qi, D. Hu, Y. Jiang, Z. Wu, M. Zheng, E. X. Chen, Y. Liang, M. A. Sadi, K. Zhang, Y. P. Chen, Recent Progresses in Machine Learning Assisted Raman Spectroscopy. Adv. Optical Mater. 2023, 11, 2203104. https://doi.org/10.1002/adom.202203104
22. Ayres, L. B.; Gomez, F. J.; Linton, J. R.; Silva, M. F.; Garcia, C. D. Taking the Leap Between Analytical Chemistry and Artificial Intelligence: A Tutorial Review. Anal. Chim. Acta 2021, 1161, 338403. DOI: 10.1016/j.aca.2021.338403
23. Houhou, R.; Bocklitz, T. Trends in Artificial Intelligence, Machine Learning, and Chemometrics Applied to Chemical Data. Anal. Sci. Adv. 2021, 2 (3-4), 128–141. DOI: 10.1002/ansa.202000162
24. Liebal, Ulf W., An N. T. Phan, Malvika Sudhakar, Karthik Raman, and Lars M. Blank. 2020. "Machine Learning Applications for Mass Spectrometry-Based Metabolomics" Metabolites 10, no. 6: 243. https://doi.org/10.3390/metabo10060243
25. Probst, D, Schwaller, P, Reymond, J-L. Reaction classification and yield prediction using the differential reaction fingerprint DRFP. Chemrxiv:2021-mc870; 2021. https://doi.org/10.33774/chemrxiv-2021-mc870.
26. Brandt S, Sittel F, Ernst M, Stock G. Machine learning of biomolecular reaction coordinates. J Phys Chem Lett. 2018; 9: 2144–50. https://doi.org/10.1021/acs.jpclett.8b00759
27. Bittracher A, Banisch R, Schütte C. Data-driven computation of molecular reaction coordinates. J Chem Phys. 2018; 149:154103. https://doi.org/10.1063/1.5035183
28. D. Ravì, C. Wong, B. Lo and G. -Z. Yang, "A Deep Learning Approach to on-Node Sensor Data Analytics for Mobile or Wearable Devices," in IEEE Journal of Biomedical and Health Informatics, vol. 21, no. 1, pp. 56-64, Jan. 2017, doi: 10.1109/JBHI.2016.2633287

DOI https://doi.org/10.15463/gfbm-mib-2024-*254*

29. P. Angeles, M. Mace, M. Admiraal, E. Burdet, N. Pavese and R. Vaidyanathan, A Wearable Automated System to Quantify Parkinsonian Symptoms Enabling Closed Loop Deep Brain Stimulation., Cham, Switzerland:Springer, pp. 8-19, 2016, [online] Available: http://dx.doi.org/10.1007/978-3-319-40379-3_2.