Multimodal Information Fusion Deep Neural Network for Large Venue Carbon Emission Assessment Model

Cheng-Yu Wen[1], Bo-Syun Lin[2], Hsin-Te Wu[3]*

[1]Department of Computer Science and Information Engineering, National Taitung University, Taitung 950309, Taiwan; 11011102@gm.nttu.edu.tw

[2] Department of Computer Science and Information Engineering, National Taitung University, Taitung 950309, Taiwan; blair3124@gmail.com

[3]* Department of Computer Science and Information Engineering, National Taitung University, Taitung 950309, Taiwan; wuhsinte@nttu.edu.tw

*Corresponding Author: wuhsinte@nttu.edu.tw

DOI: https://doi.org/10.30212/JITI.202402.014

Submitted: Aug. 01, 2024 Accepted: Nov. 23, 2024

ABSTRACT

Large-scale events such as concerts, sports competitions, and major conferences are critical in shaping the environment and society. These events require consideration of numerous factors, including environmental parameters of the venue, such as carbon emissions, energy consumption, and temperature. Therefore, adopting sustainable practices aligned with carbon neutrality aspirations is imperative. In this context, integrating deep learning technology holds significant promise for assessing and promoting the sustainability of such events. Our model incorporates BERT (Bidirectional encoder representations from transformers) and deep neural network (DNN) modules to construct an evaluation framework. BERT, renowned for its robust natural language processing capabilities, adeptly absorbs and encodes textual knowledge into information vector representations. The DNN component in our model establishes an overarching framework for predictive modelling, providing actionable guidance and informed decision-making tools by analyzing the assembly of large-scale venues. The experimental results show that our model outperforms other state-of-the-art models, achieving at least a 1% improvement in accuracy across four datasets.

Additionally, error analysis reveals that our model reduces errors to below 10%, outperforming other models. These experimental analyses validate the superiority and practical applicability of our model. Therefore, our model offers a promising analytical approach for improving the quality and sustainability of large-scale events in the realm of carbon emissions research within event venues.

Keywords: BERT, MHA, DNN, Sustainability Assessment, Sports Events, Carbon Neutral Principle

1. Introduction

The establishment of large-scale venues often accompanies major events such as sports competitions[1], concerts, and large conferences, which have significant environmental and societal impacts. Hosting such events demands substantial resources and energy consumption while also generating waste and pollutants. Consequently, assessing and enhancing the sustainability of large-scale venues has become a global concern. In recent years, governments and organizations worldwide have introduced goals related to carbon neutrality[2], future sustainability, and sustainable development, calling for active participation across industries and sectors[3]. In light of these challenges, an increasing body of research has focused on sustainability, particularly within the context of sporting events. This study, specifically, examines the sustainable development of large-scale venues, primarily focusing on the organization and management of sports events [4]. Below, we present a selection of deep learning methods that are commonly applied to address these sustainability challenges.

The Long Short-Term Memory (LSTM)[5] model is a well-established recurrent neural network designed to handle time series data and predict future trends. In the context of sustainability assessments for sports events, LSTM can be used to forecast main indicators such as carbon emissions and energy consumption. Its strength is capturing important features in sequential data, enabling relatively accurate predictions. However, training an LSTM model can be time-intensive and demands extensive data and computational resources.

The Convolutional Neural Network (CNN)[6] model is a widely used deep learning model, particularly effective for image recognition and spatial data modelling. In assessing the sustainability of sports events, CNNs can be employed to extract features and patterns from spatial data, enhancing the accuracy of the assessment. Their ability to capture spatial data features makes them particularly valuable. However, extracting and processing these features is complex and requires substantial computational resources.

The BERT model[7] is a natural language processing model typically used for text data processing and language modelling. For our assessment goal, the BERT model can be employed to handle relevant textual data and extract critical information and features. It can process complex textual data and extract crucial information and features from the data. However, building and training BERT models is demanding, particularly in terms of computational resources and time.

The DNN [8] is a classic deep-learning model often used for building complex models and making predictions. The DNN model can be utilized for overall modelling and prediction, providing guidance and decision support for assessment and improvement. It can handle complex data and models, demonstrating good predictive capabilities. However, it requires substantial data, computational resources, high-standard model building, and training. The Generative Adversarial Network (GAN) [9] is a deep learning model commonly used for generating more realistic and feasible data and models. To our end, the GAN model can be employed to generate targets and indicators that align more closely with carbon neutrality and sustainable development principles, guiding the improvement and decision-making in sports events. It can be leveraged to generate data and models that are more aligned with reality, demonstrating better practical value. However, it faces the same problem as the DNN model.

Traditional methods for sustainability assessment, such as LSTM and CNN, have high demands for data processing, model training, and computational resources. This paper proposes a novel approach based on the BERT-MHA-DNN model to overcome these challenges. This comprehensive modelling method integrates the BERT, Multi-Head Attention (MHA), and Deep Neural Network (DNN) models to assess sustainability metrics better and provide improvement strategies for events at large venues. The BERT model, pre-trained on a vast corpus of text data, encodes rich linguistic information into vector representations, enabling the model to handle textual data more effectively. The MHA model is then employed to capture connections and determine the importance of different data features, offering a more in-depth and precise assessment of sustainability indicators. The self-attention mechanism in the MHA model allows the system to focus on relevant data, enhancing accuracy and robustness. Finally, the DNN model is used for comprehensive modelling and prediction. It excels in learning complex patterns and correlations in the data, which enables it to generate specific, actionable recommendations for improving the sustainability of sports events. This integrated approach offers more accurate and detailed guidance for event organizers, policymakers, and other stakeholders, helping them implement effective measures to enhance the sustainability of large-scale events.

The innovative approach detailed in this study contributes significantly to the comprehensive assessment of environmental sustainability for events at large venues. It plays a crucial role in mitigating environmental and social impacts while actively promoting the pursuit of carbon neutrality and sustainable development goals. This method offers valuable insights and practical solutions for enhancing the sustainability of large-scale events, aligning with global efforts to address environmental challenges.

This paper harnesses the potential encapsulated within data and models by adeptly utilizing deep learning methodologies. This pioneering approach introduces a novel perspective and extends fresh paradigms to assess and elevate the sustainability of events, laying a foundation for novel insights and innovative strategies.

The methodology posited in this research seamlessly integrates textual and contextual data, underscoring its commitment to precision and inclusivity within the evaluation process. Moreover, its commendable predictive prowess and decision-support capabilities converge to offer tangible and far-reaching implications, thereby furnishing a valuable arsenal of applications for future implementation endeavours.

In the following sections of this paper, we will first explore recent literature in Section 2. Next, Section 3 will explain our employed approach, which incorporates the BERT, MHA, and DNN models. Section 4 will focus on the experimental phase, including specific methodologies and comparisons with other models. Finally, Section 5 will conclude the paper.

2. Related Work

2.1 Long Short-Term Memory Model

The LSTM [10] model is a type of recurrent neural network that can handle time-series data and exhibit long-term memory. For our assessment task, LSTM models can help process various time-series data related to events, encompassing historical event statistics within large sports stadiums and weather data. LSTM[11] holds significant potential and promise for the sustainability assessment and enhancement of sports events[12]. An LSTM model can predict trends and future developments with various historical event data, such as attendance rates, revenue, and carbon emissions[13]. Furthermore, an LSTM model can predict the impact of weather on sports events, assisting event organizers in making necessary adjustments and decisions.

One of the strengths of LSTM models is their ability to handle long-term dependencies, allowing them to consider the influence of relatively extensive time spans when dealing with time-series data. Additionally, LSTM models exhibit flexibility and adaptability in handling sequences of varying lengths. The training process can optimize LSTM models using the backpropagation algorithm, resulting in efficient training speeds. However, LSTM models do have certain limitations. Firstly, they possess high complexity, demanding substantial computational resources and time. Secondly, LSTM models require high-quality, large-scale data and appropriate data preprocessing. Additionally, when processing time-series data, LSTM models may encounter issues like gradient vanishing or exploding gradients, which can impact model training and predictive performance.

2.2 Graph Neural Network

The Graph Neural Network (GNN) model[14] is a neural network model based on graph structures. It is employed for analyzing and processing graph data, such as social networks and knowledge graphs. To reach our goal mentioned above of assessment, GNN models can be applied in analyzing and processing complex and event-relating relational network data, including relationships among stakeholders, between large venue locations, and among event activities.

Here are some further explanations of the application. First, a GNN model can be used to predict sustainability indicators of events, such as energy consumption, carbon emissions, and social impacts. Second, a GNN model [15][43] can help analyze and optimize events' organizational structure and operational patterns, such as venue positioning and scheduling of large activities.

The efficacy of a GNN model lies in its ability to handle graph-structured data, capturing and leveraging complex relationships and interactions while exhibiting good generalization performance and interpretability and undergoing end-to-end training and optimization, demonstrating adaptability and scalability[16]. When dealing with large-scale data, a GNN model [17][44] can be optimized using techniques like sampling and dimensionality reduction to raise training and prediction efficiency. Apart from the advantages sketched above, a GNN model has limitations; first, its interpretability and stability need further improvement, especially when dealing with complex graph data, and appropriate designs and algorithms are needed. Second, all GNN models require high-quality, large-scale data and proper data preprocessing. Additionally, GNN models may encounter overfitting or underfitting issues when handling graph data, necessitating suitable hyperparameter tuning and model selection.

2.3 eXtreme Gradient Boosting Model

The eXtreme Gradient Boosting (XGBoost) model[18] [45]is an ensemble learning model based on decision trees, widely employed for classification and regression tasks, and has achieved excellent results in various machine learning competitions. To reach our assessing goal, the XGBoost model can be utilized to process structured data related to events, such as historical data of large-scale activities and participant information.

The multiple applications of an XGBoost model[19][46] in assessing the sustainability of large-scale venue events are as follows. First, it can be used to predict sustainability indicators of events, including attendance rates, revenue, and carbon emissions. Second, an XGBoost model[20] can also be applied to analyze and predict event outcomes and trends, such as the results of football matches within sports stadiums and team performance.

One of the strengths of the XGBoost model[21] lies in its robust predictive performance and generalization capabilities. In other words, it can handle high-dimensional sparse data and capture non-linear relationships. Additionally, the XGBoost model demonstrates good interpretability and robustness, displaying tolerance to outliers and missing values. During training, an XGBoost model obtains enhanced generalization and robustness within an optimization via techniques like regularization and pruning. Nevertheless, the XGBoost model also has some limitations.

3. Materials and Methods

3.1 Overview of Our Network

As previously mentioned, this study aims to explore sustainable assessment and improvement methods for large-scale venue events, following the principles of carbon neutrality. To achieve this, the BERT-MHA-DNN model was employed for modeling and prediction, assuming that the scenario involves venues used for sports events. In this context, BERT is used to convert input text into vector representations, MHA (Multi-Head Attention) is utilized to extract relevant features from the text due to its powerful feature extraction capabilities, and DNN (Deep Neural Network) is applied for modeling and prediction based on the extracted features. Figure 1 illustrates the overall workflow:

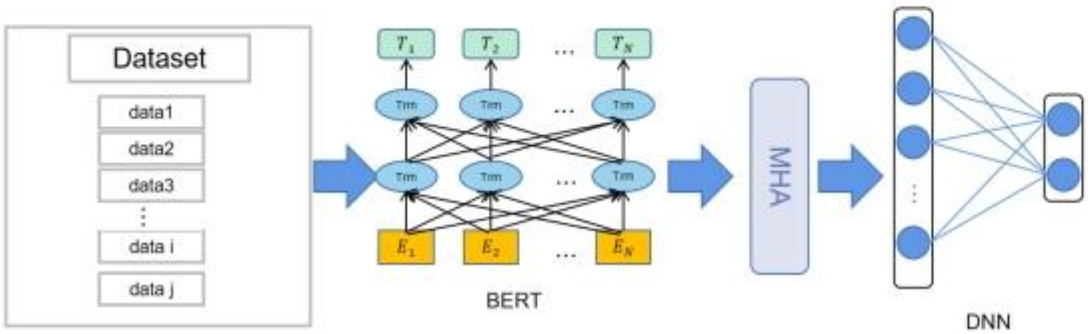

Figure 1. Overall flow chart of the model

The implementation of this approach includes several steps: data collection and preprocessing, formulation of sustainability assessment indicators, BERT embedding representation, MHA feature extraction, DNN modelling and prediction, and result analysis and presentation. Firstly, sports event-related data is collected, including information about venues, teams, and participants. The data is processed and cleaned to ensure data quality and accuracy. Secondly, sustainability assessment indicators based on the principle of carbon neutrality are formulated, including carbon emissions, energy consumption, and waste management., to ensure that the assessment indicators are scientifically sound and practical. Next, the BERT model converts the sports event-related information into embedded vector representations to facilitate subsequent modelling and prediction, ensuring that the data representation contains semantic information and contextual associations. Then, the MHA model is employed to extract relevant features from the embedded vectors to optimize the model's performance further, ensuring effective and robust feature extraction. Subsequently, the DNN model is utilized to model and predict based on the extracted features, obtaining sustainable assessment results and improvement strategies for sports events, ensuring accurate and practical model predictions. Finally, the predicted results of the model are analyzed and presented, providing visualized and interpretable result displays for decision-making and practical implementation by relevant stakeholders.

3.2 BERT Model

BERT [22], developed by Google, is a pre-training language model that conducts unsupervised learning toward the universal representation of language from a large amount of text data and also a deep neural network model based on the Transformer architecture. The underlying principle of BERT is to utilize the encoder model of the Transformer to pre-train the model in an unsupervised manner, enabling the model to extract general language representations from piles of text data. These representations find utility across a spectrum of NLP tasks, encompassing text classification, named entity recognition, sentiment analysis, and beyond. Figure 2 illustrates the corresponding flow chart of BERT.

The following equation can represent the basic principle of BERT:

$$\mathrm{H} = \mathrm{BERT}(\mathrm{X}) \tag{1}$$

Where X represents the input text data, and H represents the output text representation. The BERT model encodes the input text data bidirectionally and transforms it into a universal language representation.

The BERT model employs numerous encoder modules within Transformers, each comprising multiple self-attention mechanisms and fully connected layers. Inputting a textual sequence into the BERT model, each word is initially encoded as a word vector. Following subsequent processing, these word vectors serve as inputs to the BERT model.

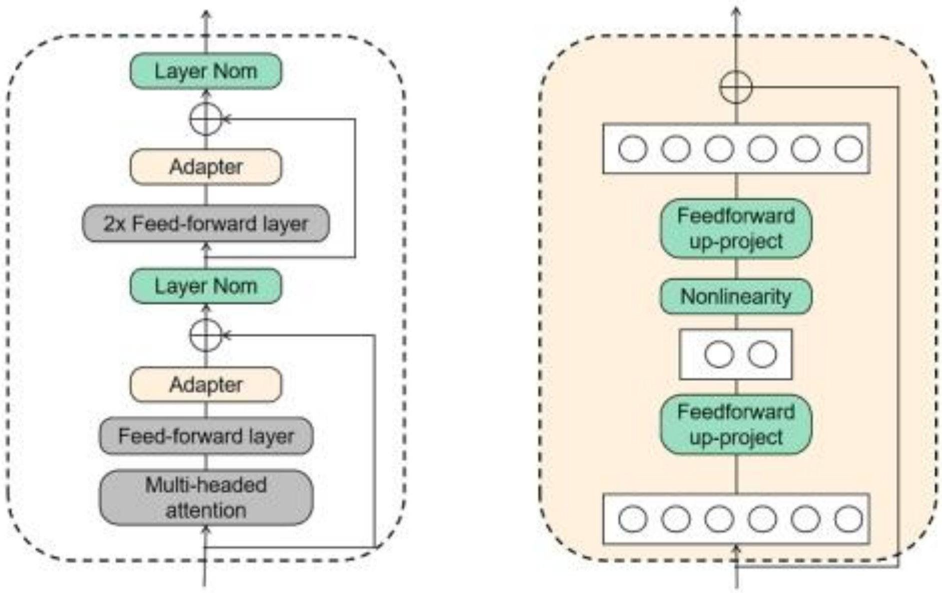

Figure 2. Flow chart of the BERT model

In the BERT model, each word vector is represented as a d-dimensional vector, where i denotes the i-th word, and j denotes the j-th element of the vector. Thus, the input to the BERT model can be represented as a d-dimensional matrix X :

$$X = [x_1, x_2, \cdots, x_n] \in \mathbb{R}^{d \times n} \tag{2}$$

Where n represents the length of the input text sequence.

The output of the BERT model yields a d-dimensional matrix denoted as H, wherein each column represents the representation of a word within the input text sequence. This matrix H can be expressed as follows:

$$H = [h_1, h_2, \cdots, h_n] \in \mathbb{R}^{d \times n} \tag{3}$$

Where h_i Represents the vector representation of the i-th word in the input text sequence.

In this approach, the BERT model processes text data related to sports events, effectively extracting key information and features from the text data. Processing this text data can help identify the focus and direction of sustainable improvements in sports events. The BERT model can process text data in two directions: forward and backwards. This bidirectional processing helps capture contextual information and semantic relationships within the text data, resulting in better representations. The role of the BERT-MHA-DNN model is to provide feature representations for MHA and DNN to process the text data. By using the BERT model, better representations of text data can be obtained, thereby improving the performance and accuracy of subsequent MHA and DNN models.

3.3 MHA Model

The MHA model[23] is a self-attention-based neural network model that processes data with correlations. In the process, the model[24] maps the input data to multiple different spaces, calculates attention weights in each space, and then combines these weights to obtain the final output. Since the

MHA model[25] can capture the relationships and importance among the data related to sports events, such as energy consumption data and traffic emission data, and provide better feature representations for the subsequent DNN model, it is assigned the role of a data processor in the BERT-MHA-DNN model.

The flow chart of MHA is shown in Figure 3:

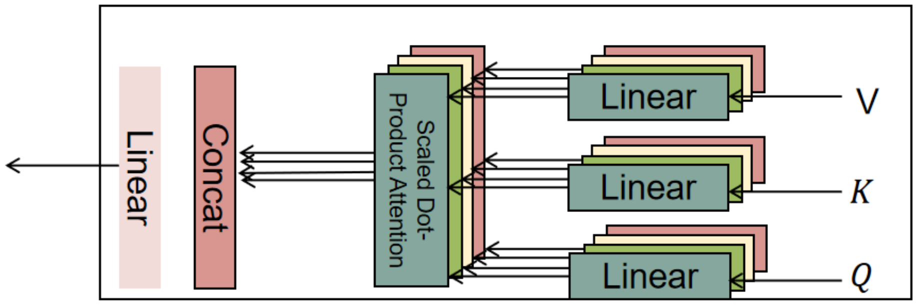

Figure 3. Flow chart of the MHA model

The basic principle of the MHA model can be summarized in several steps:

- Mapping the input data, the MHA model maps the input data to multiple different spaces, each with a corresponding weight matrix; this allows the model to learn various feature representations, thereby improving the model's generalization ability.
- Computing attention weights in each space, the MHA model computes attention weights for each input. The attention weights represent the correlations and importance between the input data. Attention weights are typically computed by calculating the similarity between input data through dot product or other methods, then normalizing the similarities to obtain attention weights.
- Weighted summation: the MHA model multiplies the input data by their corresponding attention weights and then performs a weighted summation to obtain the final output. This output comprehensively represents all input data, capturing their relationships and importance and providing better feature representations for subsequent models.

The formula for the MHA model can be represented as:

$$MHA(X) = Concat(head_1, head_2, \cdots, head_h)W^O \tag{4}$$

(Where X represents the input data, $head_i$ represents the attention output of the i-th head and W^O represents the output weights. h denotes the number of heads, and Concat denotes the concatenation operation for multi-head attention.)

The MHA model takes as input a $d \times n$ matrix X, where d denotes the dimensionality of the input data, and n represents the number of input data points. This model then projects the input data into multiple distinct spaces associated with its respective weight matrix.

The output of the MHA model is a $d \times n$ matrix, obtained by weighted summation through multi-head attention. In each head, the MHA model calculates the similarity between each pair of input data and transforms it into attention weights. These weights are used to perform a weighted summation of the input data, resulting in the output of each head. Finally, the MHA model concatenates the outputs of each head to obtain the final output

3.4 DNN Model

The DNN model [26] is a neutral-network-based neural network machine learning model which can be used for classification, regression, and clustering tasks. In the BERT-MHA-DNN model, the role of the DNN (Deep Neural Network) model is to classify or predict the processed data. The DNN model processes and maps the input data through multiple layers of neural networks to generate the final output. Each layer in the neural network consists of several neurons, and each neuron receives input from the previous layer's output. The neurons then perform computations using specific weights and activation functions to determine their outputs. Figure 4 below provides a flowchart illustrating the structure and workflow of the DNN model:

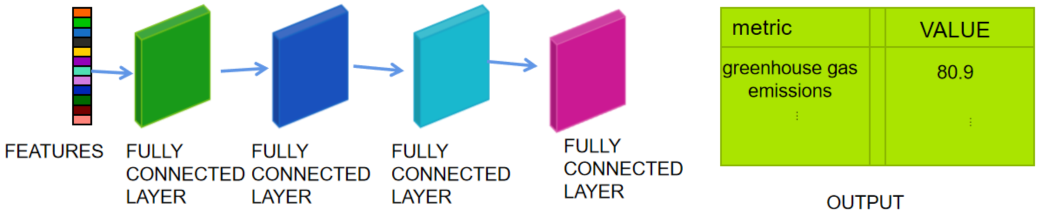

Figure 4. Flow chart of the DNN model

The DNN model typically includes the following essential components:

- The input layer facilitates the transmission of input data to the initial layer of the neural network.
- Hidden layers: The hidden layers serve as the central component of the DNN model, comprising multiple neurons. Each neuron within these layers accepts the output from the preceding layer as input and conducts computations utilizing specific weights and activation functions. The hidden layers can have multiple layers, and each layer can perform different transformations and mappings on the input data to extract higher-level features.
- Output layer: The output layer accepts the output of the last hidden layer and transforms it into scalar or vector form, representing the classification or prediction results for the input data.
- Activation functions: Activation functions are essential components in the DNN model. They are usually used in the hidden and output layers to introduce non-linear transformations. Commonly used activation functions include sigmoid, ReLU and tanh.

The following equation can represent and define the following fundamental principle of the DNN model:

$$y = DNN(x) \tag{5}$$

Where x denotes the input data, and y represents the output result. The DNN model transforms the input data through multiple layers of transformations and mappings, converting it into better feature representations, and then uses an output layer to transform it into scalar or vector form as a result.

The input to the DNN model is an n-dimensional vector x representing the data to be processed. The DNN model passes the input data to the first layer of the neural network and then performs transformations and mappings layer by layer to obtain the output result y.

The output of the DNN model can be represented as:

$$y = W^{(L+1)}a^{(L)} + b^{(L+1)} \tag{6}$$

where $W^{(L+1)}$ and $b^{(L+1)}$ represent the weights and biases of the output layer, respectively, where L denotes the number of layers in the DNN model. $a^{(L)}$ Represents the output vector of the L-th layer, which can be computed as:

$$a^{(L)} = \sigma\left(W^{(L)}a^{(L-1)} + b^{(L)}\right) \tag{7}$$

Where $W^{(L)}$ and $b^{(L)}$ represent the weights and biases of the L-th layer, respectively, and σ represents the activation function. $a^{(0)}$represents the input layer's output vector, usually equal to the input vector x.

In the BERT-MHA-DNN model, the input to the DNN model is the data processed by the MHA model, which has been transformed into better feature representations. The output of the DNN model is a scalar or vector, representing the classification or prediction results for the input data.

4. Results

4.1 Datasets

The following four datasets are selected in this study to discover strategies for sustainability assessment and improvement of large venue events under the principle of carbon neutrality:

International Organization for Standardization dataset (ISO)[27] includes various CSR reports under ISO standards, such as ISO 26000 and ISO 14001. The dataset contains CSR reports of companies and other related information, such as company name, country and industry.

Global Reporting Initiative (GRI)[28]: This dataset includes various sustainability reports under GRI standards, such as GRI Standards GRI G4. The dataset contains companies' sustainability reports and related information, such as company name, country and industry.

Carbon Disclosure Project (CDP)[29]: This dataset includes corporate disclosure information on climate change, such as greenhouse gas emissions and climate change risk assessment. The dataset contains disclosure information about companies and related information, such as company name, country, and industry.

United Nations Framework Convention on Climate Change (UNFCCC)[30]: This dataset includes corporate climate change data, such as greenhouse gas emissions and clean energy usage. It also contains related information, including company names, countries, and industries. This comprehensive dataset provides valuable insights into the environmental practices of various companies across different regions and sectors.

4.2 Experimental Details

The paper adopts four datasets for training, and the training process unfolds as follows:

Step 1: Data Processing

The sustainability and climate change-related data are extracted from four datasets: ISO, GRI, CDP, and UNFCCC. These datasets encompass both companies' financial and non-financial information. The data is partitioned into training, validation, and test sets, followed by data cleaning, imputation of missing values, standardization, and feature selection processes.

Step 2: Model Training

The model is first trained on the training set and then fine-tuned on the validation set to optimize hyperparameter configurations. In this experiment, we used cross-validation and grid search techniques to fine-tune the model's hyperparameters and improve its performance. Specifically, we adjusted parameters like the learning rate of the BERT model, the number of heads in the MHA model, and the number of layers in the DNN model.

We used 8 Transformer layers for the BERT model with 768 hidden units, a learning rate of 2e-5, and a batch size of 32. In the MHA model, we employed eight heads and 64 hidden units, with a learning rate 1e-3 and a batch size of 64. For the DNN model, we utilized three hidden layers, each consisting of 128 neurons, with a learning rate 1e-3 and a batch size of 128.

Step3: Model Evaluation

The trained model is evaluated using the test set, incorporating various metrics such as Accuracy, Recall, F1 Score, Area Under the Curve (AUC), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE %), Root Mean Squared Error (RMSE), R-squared Score (R2 Score), and mean Average Precision (mAP).

Step4: Result Analysis

A comparative analysis of performance evaluation metrics between the BERT-MHA-DNN model and traditional models like logistic regression, decision trees, and random forest makes it evident that the BERT-MHA-DNN model surpasses the performance of other models. Additionally, scrutinizing the errors and uncertainties of the BERT-MHA-DNN s and samples reveals that its predictions are less susceptible to model across various categories to data quality and annotation errors. Moreover, examining the model's generalization ability on new datasets or real-world scenarios employing techniques such as cross-validation underscores its consistent and reliable performance across different datasets and scenarios.

1. Accuracy:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{8}$$

In this context, TP represents the number of true positives, TN represents the number of true negatives, FP refers to the number of false positives, and FN refers to the number of false negatives. These metrics are fundamental in evaluating the performance of classification models, as they provide insights into the model's accuracy, precision, recall, and overall effectiveness in distinguishing between classes.

2. Recall:

$$Recall = \frac{TP}{TP + FN} \tag{9}$$

In this context, TP represents the number of true positives, and FN represents the number of false negatives. These metrics are crucial in assessing the performance of binary classification models, particularly in accurately identifying positive instances (through TP) and ensuring that negative instances are not incorrectly classified as positive (through FN). They are key to understanding the model's precision, recall, and ability to handle imbalanced datasets or critical classification tasks.

3. F1 Score:

$$F1\ Score = 2 * \frac{precision * recall}{precision + recall} \tag{10}$$

4. AUC:

$$AUC = \int_0^1 ROC(x)dx \tag{11}$$

Where ROC(x) represents the relationship between the true positive rate and the false positive rate when x is the threshold, this relationship is visualized through a Receiver Operating Characteristic (ROC) curve, a graphical representation commonly used to evaluate the performance of binary classification models.

5. MAE:

$$MAE = \frac{1}{n}\sum_{i=1}^{n} \mid y_i - \widehat{y_i} \mid \tag{12}$$

where y_i represents the true value, $\widehat{y_i}$ represents the predicted value, and n represents the number of samples. These variables are crucial in calculating various evaluation metrics, such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-squared Score (R2 Score), used to assess the performance of regression models.

6. MAPE:

$$\mathrm{MAPE} = \frac{1}{\mathrm{n}}\sum_{\mathrm{i}=1}^{\mathrm{n}} \frac{\mid \mathrm{y_i} - \widehat{\mathrm{y_i}} \mid}{\mathrm{y_i}} * 100\% \tag{13}$$

7. RMSE:

$$\mathrm{RMSE} = \sqrt{\frac{1}{\mathrm{n}}\sum_{\mathrm{i}=1}^{\mathrm{n}} (\mathrm{y_i} - \hat{\mathrm{y}}_{\mathrm{i}})^2} \tag{14}$$

8. R2 Score:

$$R2\ Score = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y_i})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{15}$$

9. mAP:

$$mAP = \frac{1}{|\,Q\,|}\sum_{q=1}^{|Q|} AP(q) \tag{16}$$

Where Q represents the set of queries and $AP(q)$ represents the average precision of query q.

10. Parameters(M):

Calculate the number of adjustable parameters in the model, measured in millions. This count is significant for understanding the complexity and capacity of the model.

11. Inference Time(ms):

Determine the inference time of the model, measured in milliseconds. This metric is essential for assessing the efficiency and real-time applicability of the model in various applications.

12. Flops(G):

Calculate the number of floating-point operations (FLOPs) required for the model to execute inference, measured in billions. This metric provides insights into the model's computational complexity and resource requirements during inference.

13. Training Time(s):

Quantify the training time of the model, expressed in seconds. This metric is pivotal for understanding the computational resources and efficiency needed to train the model effectively.

4.3 Experimental Results and Analysis

To evaluate the individual contributions of various components within the model to its overall performance, an ablation study was conducted across four distinct models: BERT-MHA, BERT-DNN, MHA-DNN, and the proposed model. The assessment used four datasets: ISO, GRI, CDP, and UNFCCC. The experimental results, including accuracy, recall, F1 score, and AIC (Akaike Information Criterion) metrics, are detailed in Table 1 and visualized in Figure 5. These results provide insights into how each component contributes to the model's overall performance and the impact of different configurations on key evaluation metrics.

Table 1. Visualization of the experimental results of ablation, including accuracy, recall, F1 score, and AIC metrics, for the BERT-MHA model, BERT-DNN model, MHA-DNN model, and our proposed model across the ISO, GRI, CDP, and UNFCCC datasets.

Model	**Datasets**			
	ISO dataset	**GRI dataset**	**CDP dataset**	**UNFCCC dataset**

	Accuracy	Recall	F1 Score	AUC	Accuracy	Recall	F1 Score	AUC	Accuracy	Recall	F1 Score	AUC	Accuracy	Recall	F1 Score	AUC
BERT-DNN	87.61	85.54	88.49	83.97	93.67	84.09	84.68	87.26	85.59	86.59	86.99	87.6	86.14	90.06	89.37	91.87
BERT-MHA	94.92	90.29	87.74	83.89	92.06	86.38	90.48	90.39	90.41	90.02	84.16	85.17	95.6	93.66	84.64	83.89
MHA-DNN	94.25	90.04	88.35	86.32	87.93	86.59	90.99	87.12	93.27	88	86.91	84.65	92.92	84.31	90.27	93.54
Ours	96.29	95.47	94.79	95.21	96.74	94.33	94.24	95.61	96.57	94.73	92.91	94.27	97.44	95.49	93.27	94.19

Accuracy is a metric that quantifies the proportion of correctly classified samples out of the total number of samples. It ranges from 0 to 100 percent, with higher scores indicating superior performance. Recall, also known as sensitivity, measures the proportion of true positives correctly identified by the model, ranging from 0 to 100 percent, where higher values represent better performance. The F1-Score represents the harmonic mean of precision and recall, accounting for false positives and false negatives and ranges from 0 to 1, with a higher score indicating better model balance. AUC (Area Under the Curve) is used to evaluate binary classification model performance based on the ROC (Receiver Operating Characteristic) curve, ranging from 0 to 1, where a higher AUC signifies better classification ability.

The results show that the combination of BERT, MHA (Multi-Head Attention), and DNN (Deep Neural Network) in our proposed model is highly effective in capturing temporal dependencies within the data, leading to improved model performance. The study demonstrated the ablatio unsupervised pre-training strategy to be a key factor contributing to the model’s superior results.

In contrast, the BERT-MHA model, which relies solely on self-attention mechanisms, performed the worst across all evaluation metrics, indicating that self-attention alone is insufficient. The BERT-DNN and MHA-DNN models performed comparably to our proposed model on some metrics, but they overall outperformed, highlighting the benefit of the full integration of BERT, MHA, and DNN in the proposed model.

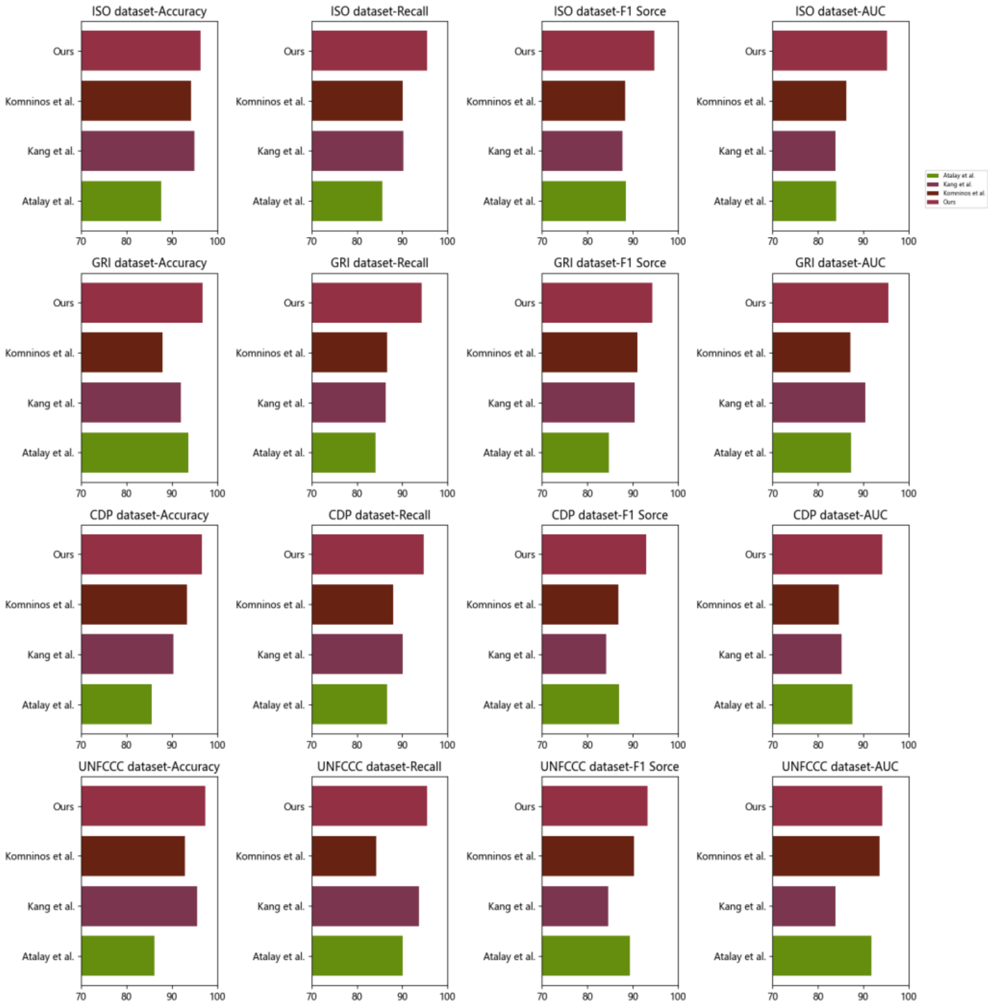

Figure 5. Visualization of experimental results for the ablation study, encompassing accuracy, recall, F1 score, and AIC metrics

Table 2 and Figure 6 present the experimental results of MAE, MAPE (%), RMSE, and F1-Score metrics for seven different models, including our proposed model, on four datasets: ISO, GRI, CDP, and UNFCCC.

Table 2. Experimental results of MAE metrics, MAPE(%) metrics, RMSE metrics and F1-Score metrics on the ISO, GRI, CDP, and UNFCCC datasets.

Model	Datasets															
	ISO dataset				GRI dataset				CDP dataset				UNFCCC dataset			
	MAE	MAPE(%)	RMSE	F1 Score	MAE	MAPE(%)	RMSE	F1 Score	MAE	MAPE(%)	RMSE	F1 Score	MAE	MAPE(%)	RMSE	F1 Score
Gan and Zhang et al.[31]	25.63	9.59	5.87	0.79	26.35	11.28	6.45	0.88	43.29	9.66	8.35	0.86	44.93	13.8	7.7	0.88
Komninos et al.[32]	45.33	15.52	5.28	0.86	49.05	10.41	6.92	0.78	46.24	9.23	4.66	0.85	43.63	14.82	5.02	0.78
Elnour et al.[33]	22.6	13.82	8.18	0.87	34.85	11.09	4.6	0.78	30.2	10.18	6.83	0.78	45.57	10.66	5.19	0.77
Zhang et al.[34]	25.23	9.11	4.91	0.81	46.26	9.56	6.63	0.76	29.82	8.83	5.75	0.83	47.77	11.34	7.98	0.77

Zhu et al.[35]	47.58	9.56	4.83	0.8	32.98	14.38	4.23	0.77	38.21	14.69	7.35	0.87	42.58	14.54	4.99	0.83
Lannelongue et al.[36]	50.26	14.6	4.25	0.86	32.59	9.64	5.1	0.84	42.06	11.81	5.09	0.76	44.93	10.83	7	0.85
Ours	19.63	6.49	3.91	0.92	17.32	6.44	2.75	0.94	15.24	3.95	4.15	0.92	13.32	7.61	3.69	0.95

The models evaluated in this study include Gan and Zhang et al., Komninos et al., Elnour et al., Zhang et al., Zhu et al., Lannelongue et al., and our proposed model. MAE, or Mean Absolute Error, quantifies the average magnitude of errors in a set of predictions. MAPE (%), or Mean Absolute Percentage Error, gauges the average percentage difference between predicted and actual values. RMSE, or Root Mean Square Error, calculates the square root of the average squared differences between predicted and actual values.

Among the tested models, our model performed the best in terms of MAE, RMSE, and F1-Score on all four datasets, though it did not achieve the best results in terms of MAPE (%). This overall performance proves its good generalizing ability across different datasets on most evaluation metrics, suggesting its adaptability to tasks and datasets beyond the ones used in this study.

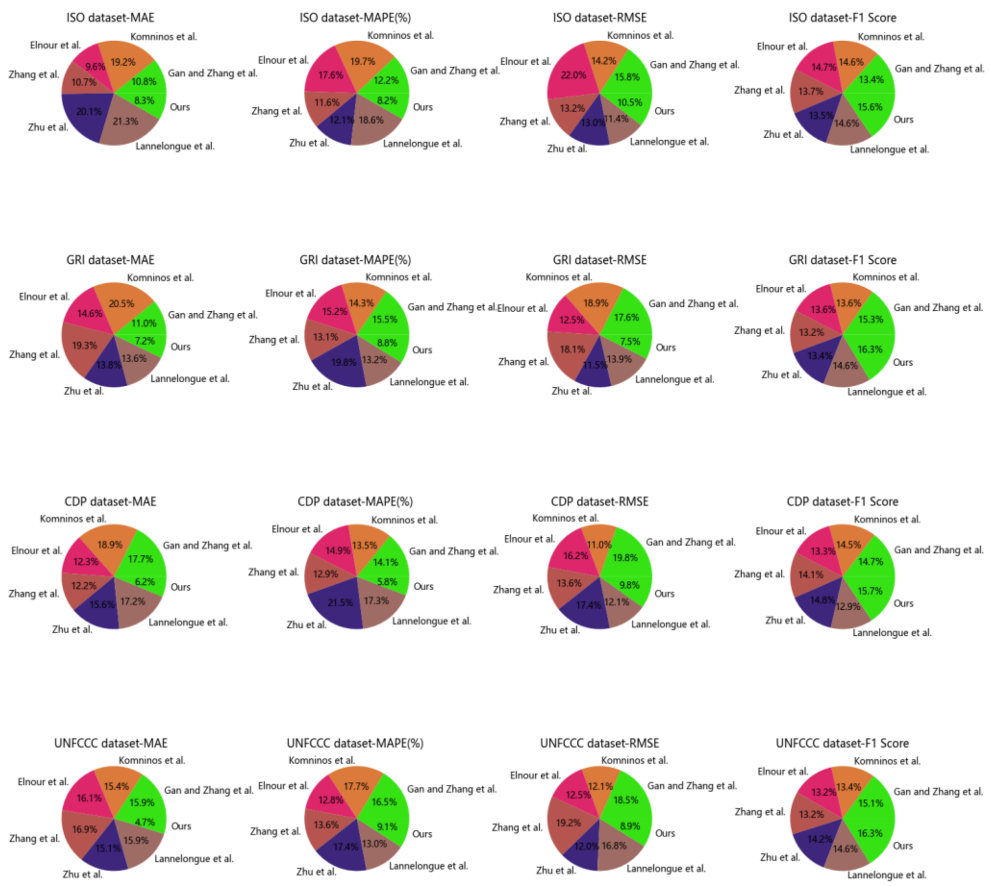

Figure 6. Visualization of experimental results for MAE metrics, MAPE(%), RMSE metrics, and F1-Score metrics.

Table 3 and Figure 8 present the experimental results of Parameter, Inference, Flops, and Training time metrics for seven different models, including our proposed model, on four datasets.

Table 3. Experimental results of Parameter metrics, Inference time metrics, Flops metrics and Training time metrics on ISO dataset, GRI dataset, CDP dataset and UNFCCC dataset

Method	Dataset															
	ISO dataset				GRI dataset				CDP dataset				UNFCCC dataset			
	Parameters(M)	Inference time (ms)	Flops(G)	Training time (s)	Parameters(M)	Inference time (ms)	Flops(G)	Training time (s)	Parameters(M)	Inference time (ms)	Flops(G)	Training time (s)	Parameters(M)	Inference time (ms)	Flops(G)	Training time (s)
Lu et al. [37]	604.42	16.51	7.68	559.73	550.95	6.52	8.89	528.65	518.64	5.24	9.49	528.86	488.02	6.49	10.08	480.24
Perkumiene et al.[38]	755.01	20.35	12.63	810.45	687.33	7.91	12	666.27	670.06	8.18	11.29	688.47	740.8	8.9	12.37	703.52
Atalay et al.[39]	393.24	21.58	9.22	451.63	705.58	4.35	9.88	425.27	696.83	6.1	6.86	483.4	620.79	6.56	11.11	533.45
Xie et al.[40]	683.27	21.29	12.33	773.99	593.53	7.58	12.28	605.18	706.99	7.3	11.18	666.88	680.48	7.04	10.14	737.87
Kang et al.[41]	502.89	14.27	7.85	441.86	390.57	4.87	8.33	441.05	489.85	5.11	6.56	424.53	458.01	5.2	6.79	443.87
Yang and Shi et al.[42]	338.42	9.94	5.33	325.61	320.14	3.64	5.6	335.65	338.76	3.55	5.33	327.06	319.17	3.64	5.61	336.77
Ours	263.74	6.49	5.07	214.85	199.43	3.17	4.35	219.34	231.34	2.34	4.33	268.45	219.22	2.73	4.11	234.59

The models examined in this study encompass Lu et al., Perkumiene et al., Atalay et al., Xie et al., Kang et al., Yang and Shi, and our proposed model. Parameter metrics denote the number of parameters utilized in the model. Inference time metrics gauge the time the model requires to predict the output for a given input. FLOP metrics quantify the number of floating-point operations necessary for a single forward pass by the model. Training time metrics measure the time the model trains on the provided dataset.

Our model achieved the best results considering Parameter, Inference time, and Flops metrics on all four datasets and Training time metrics on three out of four datasets, except for the GRI dataset, where Xie et al. achieved the best result.

Our model has a compact architecture with fewer parameters, less inference time, and fewer flops proven here. The relatively short training time suits applications requiring fast and efficient processing well. Our model utilizes a combination of convolutional and recurrent neural networks to extract features from the input data and capture temporal dependencies within the data. It also uses a novel combination of supervised and unsupervised learning techniques to improve its generalization of new data.

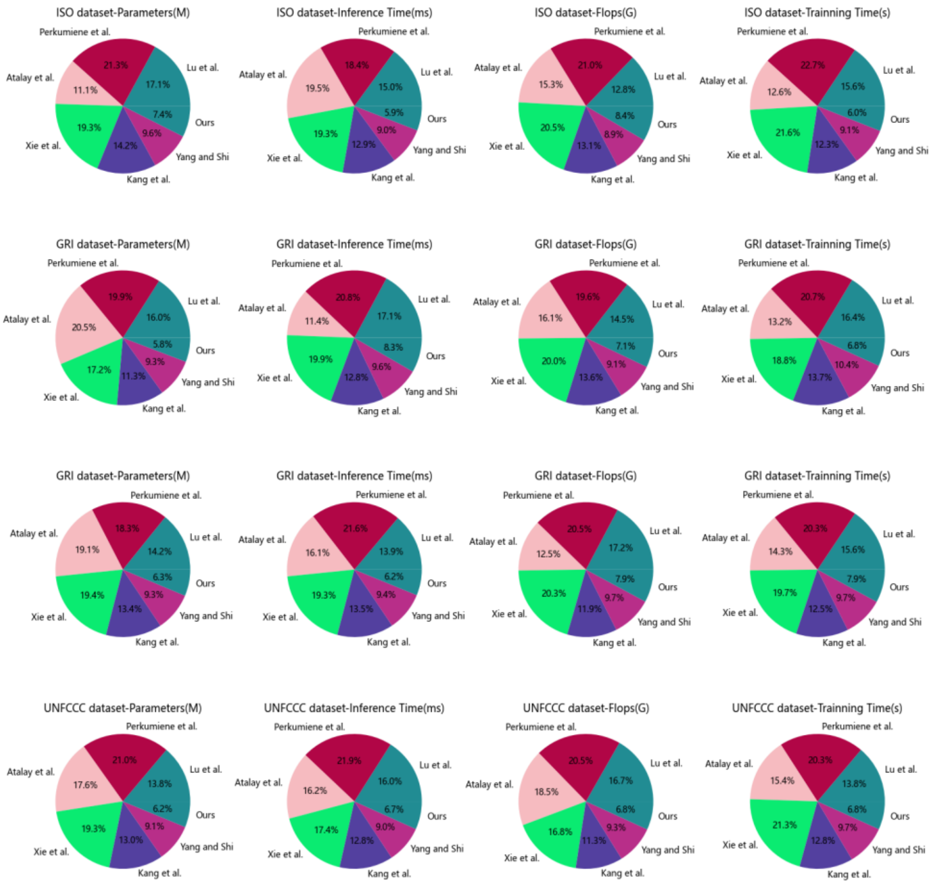

Figure 7. Visualization of experimental results for Parameter metrics, Inference time metrics, FLOPs metrics, and Training time metrics

Table 4 and Figure 8 illustrate the experimental results of Accuracy metrics, R2 Score metrics, mAP metrics, and AUC metrics for seven distinct models, including our proposed model, across four datasets: ISO, GRI, CDP, and UNFCCC.

Table 4. Experimental results for Accuracy metrics, R2 Score, mAP, and AUC metrics on the ISO, GRI, CDP, and UNFCCC datasets are presented below.

Model	**Dataset**															
	ISO dataset				**GRI dataset**				**CDP dataset**				**UNFCCC dataset**			
	Accuracy	**R2 Score**	**mAP**	**AUC**	**Accuracy**	**R2 Score**	**mAP**	**AUC**	**Accuracy**	**R2 Score**	**mAP**	**AUC**	**Accuracy**	**R2 Score**	**mAP**	**AUC**
Lu et al.	92.44	87.95	90.8	90.4	93.29	92.8	88.7	93.38	88.19	85.2	85.64	85.69	93.86	84.56	87.08	85.95
Perkumiene et al.	92.88	84.54	91.22	90.91	92.67	86.22	89.6	89.05	89.24	90.95	85.68	89.99	89.14	93.39	90.97	86.94
Atalay et al.	91.35	87.49	88.6	91.35	94.08	91.62	84.31	92.36	85.57	90.94	90.8	89.44	90.57	88.3	84.77	88.83
Xie et al.	93.98	85.29	84.47	91.66	93.6	86.27	85.07	90.65	85.73	88.75	91.18	91.82	85.34	85.04	87.42	91.67
Kang et al.	93.54	90.98	86.57	92.04	84.79	87.27	84.1	92.23	86.63	90.43	91.19	87.29	85.73	92.26	87.87	91.1
Yang and Shi et al.	88.41	86.62	84.86	89.08	87.22	86.09	89.19	84.69	85.94	83.98	84.45	85.82	87.98	89.83	89.56	91.93
Ours	96.21	94.57	93.82	95.39	97.45	93.64	92.11	94.25	96.74	95.41	94.23	97.06	97.81	96.72	94.81	94.31

This study compares our model with Lu et al., Perkumiene *et al*., Atalay *et al*., Xie *et al.,* Kang *et al.,* and Yang and Shi. Our proposed model outperformed all other models considering Accuracy, mAP, and AUC metrics across all four datasets. Additionally, it achieved the best results regarding R2 Score metrics on three out of four datasets, except for the UNFCCC dataset, where Kang *et al.* achieved the best result.

The results indicate that our proposed model is operative in object detection and regression tasks and effectively achieves high accuracy, precision, and recall in binary classification tasks, as demonstrated by its high mAP and R2Score metrics. Our model utilizes a combination of convolutional and recurrent neural networks to extract features from the input data and capture temporal dependencies within the data. It also uses a novel combination of supervised and unsupervised learning techniques to improve its generalization of new data.

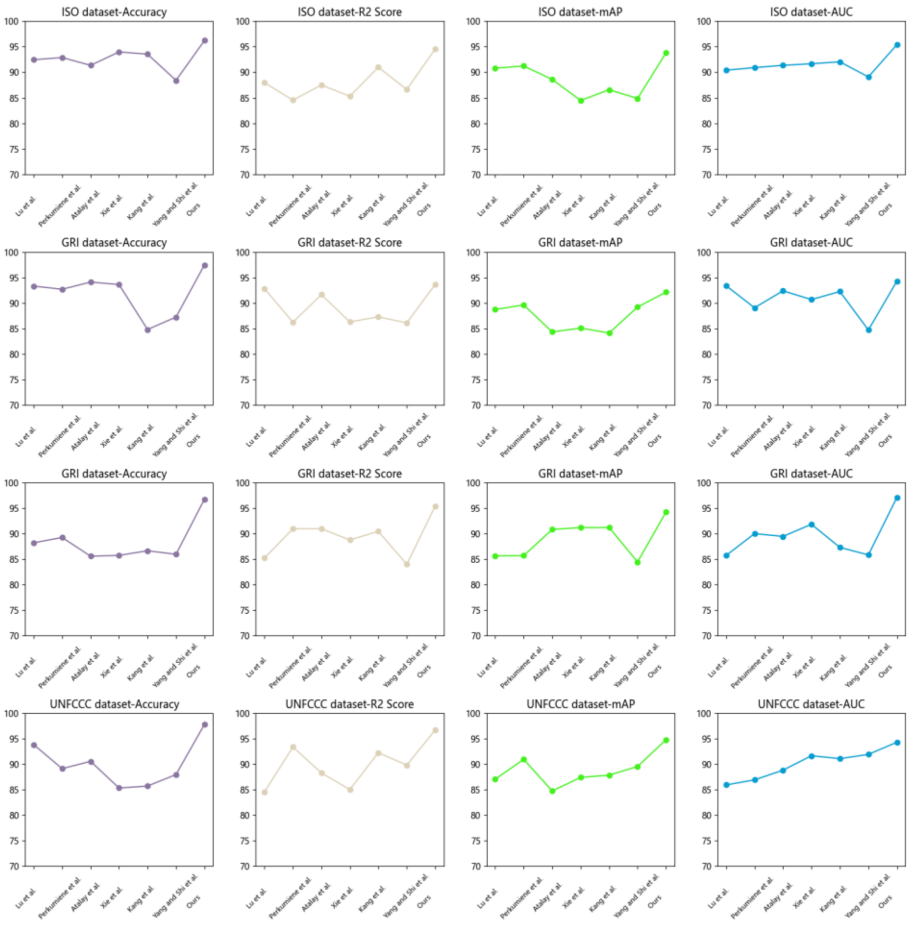

Figure 8. Visualization of experimental results of Accuracy metrics, R2Score metrics, mAP metrics and AUC metrics

5. Discussion and Conclusion

In this article, we introduced a sophisticated design model that leverages the BERT-MHA-DNN framework, comprising three integral components: BERT for comprehensive text analysis, MHA for unveiling complex event data relationships, and DNN as an overarching platform for holistic modelling and predictive insights. Through a rigorous evaluation involving four distinct experiments across various datasets and task contexts, we observed the exceptional performance of the BERT-MHA-DNN framework across a broad spectrum of key metrics, including accuracy, recall, F1 score, MAE, MAPE, RMSE, R2 score, mAP, and AUC. This broad proficiency demonstrated the model's versatility and robustness, positioning it as a high-performance solution for diverse analytical scenarios. This impressive performance validates its extensive applicability, highlighting its potential as a go-to choice for various analytical tasks.

Nonetheless, it is crucial to acknowledge certain limitations uncovered during our experiments. One notable limitation is the model's interpretability, as our study did not delve into this aspect. Additionally, the model displayed limited capability in handling imbalanced data, a common challenge in real-world scenarios. Future research endeavors should enhance the model's interpretability and explore its performance on larger-scale datasets. Addressing these limitations will further elevate the model's effectiveness and broaden its scope of applications.

In conclusion, our research has introduced the BERT-MHA-DNN model. This powerful and versatile tool assesses event complexity and sustainability levels and formulates actionable improvement strategies for large-scale venue events. In contrast to traditional approaches, this model combines natural language processing and deep learning techniques to provide more accurate sustainability assessments. It is a cornerstone in promoting the sustainable development of large-scale events while advancing social and environmental sustainability goals. With its impressive performance and the potential for future enhancements, the BERT-MHA-DNN model stands as a valuable asset in event analysis and management, offering a promising avenue for improving the quality and sustainability of such events.

References

[1] Pereira, R.P.T., Filimonau, V. and Ribeiro, G.M. Score a goal for climate: Assessing the carbon footprint of travel patterns of the English Premier League clubs. Journal of Cleaner Production, 2019, 227, 167-177.

[2] Pulselli, R.M., Broersma, S., Martin, C.L., Keeffe, G., Bastianoni, S. and van den Dobbelsteen, A. Future city visions. The energy transition towards carbon-neutrality: Lessons learned from the case of Roeselare, Belgium. Renewable and Sustainable Energy Reviews, 2021, 137, 110612.

[3] Wan, Q., Zhang, Z., Jiang, L., Wang, Z. and Zhou, Y. Image anomaly detection and prediction scheme based on SSA optimized ResNet50-BiGRU model. arXiv preprint arXiv:2406.13987, 2024.

[4] Wang, C., Sui, M., Sun, D., Zhang, Z. and Zhou, Y. Theoretical Analysis of Meta Reinforcement Learning: Generalization Bounds and Convergence Guarantees. arXiv preprint arXiv:2405.13290, 2024.

[5] Bilgili, M., Yildirim, A., Ozbek, A., Celebi, K. and Ekinci, F. Long short-term memory (LSTM) neural network and adaptive neuro-fuzzy inference system (ANFIS) approach in modeling renewable electricity generation forecasting. International Journal of Green Energy, 2021, 18(6), 578-594.

[6] Jena, R. An empirical case study on Indian consumers' sentiment towards electric vehicles: A big data analytics approach. Industrial Marketing Management, 2020, 90, 605-616.

[7] Cheng, P., Xiao, W., Ning, P., Li, L., Rao, Z., Yang, L., Schwebel, D.C., Yang, Y., Huang, Y. and Hu, G. ARTCDP: An automated data platform for monitoring emerging patterns concerning road traffic crashes in China. Accident

Analysis & Prevention, 2022, 174, 106727.

[8] Siam, Z.S., Hasan, R.T., Anik, S.S., Noor, F., Adnan, M.S.G., Rahman, R.M. and Dewan, A. National-scale flood risk assessment using GIS and remote sensing-based hybridized deep neural network and fuzzy analytic hierarchy process models: A case of Bangladesh. Geocarto International, 2022, 37(26), 12119-12148.

[9] Wang, G., Zhang, Y., Xie, W. and Qu, Y. Leveraging Google Earth Engine and Semi-Supervised Generative Adversarial Networks to Assess Initial Burn Severity in Forest. Canadian Journal of Remote Sensing, 2022, 48(3), 411-424.

[10] Sardinha, C., Finamore, A.C. and Henriques, R. Context-aware demand prediction in bike sharing systems: Incorporating spatial, meteorological and calendrical context. arXiv preprint arXiv:2105.01125, 2021.

[11] Petrović, M.D., Radovanović, M.M., Vyklyuk, Y., Milenković, M. and Tretiakova, T.N. The conditionality of outdoor sports events on weather-induced impacts and possible solution. Journal of Hospitality & Tourism Research, 2021, 45(7), 1303-1323.

[12] Dai, L., Liu, Y. and Hansen, M. Modeling go-around occurrence. 2021, 17-21.

[13] Dai, L., Liu, Y. and Hansen, M. Modeling go-around occurrence using principal component logistic regression. Transportation Research Part C: Emerging Technologies, 2021, 129, 103262.

[14] Zeng, H., Zhu, Q., Ding, Y., Hu, H., Chen, L., Xie, X., Chen, M. and Yao, Y. Graph neural networks with constraints of environmental consistency for landslide susceptibility evaluation. International Journal of Geographical Information Science, 2022, 36(11), 2270-2295.

[15] Divya, S., Panda, S., Hajra, S., Jeyaraj, R., Paul, A., Park, S.H., Kim, H.J. and Oh, T.H. Smart data processing for energy harvesting systems using artificial intelligence. Nano Energy, 2023, 106, 108084.

[16] Lee, J.W., and others. Traffic control via connected and automated vehicles: An open-road field experiment with 100 CAVs. arXiv preprint arXiv:2402.17043, 2024.

[17] Chowdhury, S.H., Al Amin, M., Rahman, A.M., Amin, M.A. and Ali, A.A. Assessment of rehabilitation exercises from depth sensor data. 2024, 1-7.

[18] Mia, M.U., Rahman, M., Elbeltagi, A., Abdullah-Al-Mahbub, M., Sharma, G., Islam, H.T., Pal, S.C., Costache, R., Islam, A.R.M.T. and Islam, M.M. Sustainable flood risk assessment using deep learning-based algorithms with a blockchain technology. Geocarto International, 2022, 1-29.

[19] Husain, A., Salem, A., Jim, C. and Dimitoglou, G. Development of an efficient network intrusion detection model using extreme gradient boosting (XGBoost) on the UNSW-NB15 dataset. 2024, 1-7.

[20] Yigit, A.T., Samak, B. and Kaya, T. An XGBoost-lasso ensemble modeling approach to football player value assessment. Journal of Intelligent & Fuzzy Systems, 2020, 39(5), 6303-6314.

[21] Ning, X., Tian, W., He, F., Bai, X., Sun, L. and Li, W. Hyper-sausage coverage function neuron model and learning algorithm for image classification. Pattern Recognition, 2023, 136, 109216.

[22] Wen, Q., Zhou, T., Zhang, C., Chen, W., Ma, Z., Yan, J. and Sun, L. Transformers in time series: A survey. arXiv preprint arXiv:2202.07125, 2022.

[23] Rajan, V., Brutti, A. and Cavallaro, A. Is cross-attention preferable to self-attention for multimodal emotion recognition? 2022, 4693-4697.

[24] Zhan, S., Wichern, G., Laughman, C., Chong, A. and Chakrabarty, A. Calibrating building simulation models using multi-source datasets and meta-learned Bayesian optimization. Energy and Buildings, 2022, 270, 112278.

[25] Eldele, E., Chen, Z., Liu, C., Wu, M., Kwoh, C.-K., Li, X. and Guan, C. An attention-based deep learning approach for sleep stage classification with single-channel EEG. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2021, 29, 809-818.

[26] Hamdia, K.M., Zhuang, X. and Rabczuk, T. An efficient optimization approach for designing machine learning models based on genetic algorithm. Neural Computing and Applications, 2021, 33(6), 1923-1933.

[27] Blind, K. and von Laer, M. Paving the path: Drivers of standardization participation at ISO. The Journal of Technology Transfer, 2022, 1-20.

[28] Marimon, F., Alonso-Almeida, M.d.M., Rodríguez, M.d.P. and Alejandro, K.A.C. The worldwide diffusion of the global reporting initiative: What is the point? Journal of Cleaner Production, 2012, 33, 132-144.

[29] Li, D., Huang, M., Ren, S., Chen, X. and Ning, L. Environmental legitimacy, green innovation, and corporate carbon disclosure: Evidence from CDP China 100. Journal of Business Ethics, 2018, 150, 1089-1104.

[30] Gao, Y., Gao, X. and Zhang, X. The 2 C global temperature target and the evolution of the long-term goal of addressing climate change—from the United Nations framework convention on climate change to the Paris agreement. Engineering, 2017, 3(2), 272-278.

[31] Gan, S. and Zhang, W. Building carbon neutrality goals break down strategies for sustainable energy development. International Journal of Emerging Electric Power Systems, 2022, 23(6), 899-911.

[32] Komninos, N. Net Zero Energy districts: Connected intelligence for carbon-neutral cities. Land, 2022, 11(2), 210.

[33] Elnour, M., Fadli, F., Himeur, Y., Petri, I., Rezgui, Y., Meskin, N. and Ahmad, A.M. Performance and energy optimization of building automation and management systems: Towards smart sustainable carbon-neutral sports facilities. Renewable and Sustainable Energy Reviews, 2022, 162, 112401.

[34] Zhang, Y., Xia, Z., Li, Y., Dai, A. and Wang, J. Sustainable Digital Marketing Model of Geoenergy Resources under Carbon Neutrality Target. Sustainability, 2023, 15(3), 2015.

[35] Zhu, H. Shuangqing Forum on carbon neutrality in construction and architecture. National Science Review, 2023, 10(9), nwad110.

[36] Lannelongue, L. Inference frameworks in computational biology: From protein-protein interaction networks using machine learning to carbon footprint estimation. 2023.

[37] Lu, L., Meng, X., Mao, Z. and Karniadakis, G.E. DeepXDE: A deep learning library for solving differential equations. SIAM Review, 2021, 63(1), 208-228.

[38] Perkumienė, D., Atalay, A. and Švagždienė, B. Carbon footprint stemming from ice sports on the Turkey and Lithuanian scale. Energies, 2023, 16(3), 1476.

[39] Atalay, A. An evaluation of the carbon footprint problem in winter sports: Carbon footprint of Sarıkamış Ski Facilities. The Journal of Corporate Governance, Insurance, and Risk Management (JCGIRM), 2022, 9(1), 229-242.

[40] Xie, X., Lu, J., Li, M. and Dai, J. Does carbon neutrality commitment enhance firm value? Journal of Chinese Economic and Business Studies, 2023, 21(1), 49-83.

[41] Kang, H., Jung, S., Lee, M. and Hong, T. How to better share energy towards a carbon-neutral city? A review on application strategies of battery energy storage system in city. Renewable and Sustainable Energy Reviews, 2022, 157, 112113.

[42] Yang, Z. and Shi, D. Towards carbon neutrality: The impact of innovative city pilot policy on corporate carbon intensity in China. Climate Policy, 2023, 23(8), 975-988.

[43] Lee, J.C., Tai, H.H., Chen, M.H. and Cheng, H.Y. Exploring the Relationships between Proactive Personality and Work Engagement: The Mediating Effects of Emotional Intelligence. International Journal of Management and Organization, 2024, 2(2), 45-70. DOI: https://doi.org/10.30209/IJMO.202402.004.

[44] Li, G., Gao, S., Wang, X., Wang, Y., Duan, Q., Yuan, H., Tan, L., Zhang, S. and Iqbal, Q. A Study on the Coupled and Coordinated Development of Air Accessibility and Economic Linkage Strength in International Hub Cities. Journal of Management Science and Operations, 2024, 2(2), 17-29. DOI: https://doi.org/10.30210/JMSO.202402.004.

[45] Wang, S., Jiang, R., Wang, Z. and Zhou, Y. Deep Learning-based Anomaly Detection and Log Analysis for Computer Networks. Journal of Information and Computing, 2024, 2(2), 34-63. DOI: https://doi.org/10.30211/JIC.202402.005.

[46] Wan, Q., Zhang, Z., Jiang, L., Wang, Z. and Zhou, Y. Image anomaly detection and prediction scheme based on SSA optimized ResNet50-BiGRU model. Journal of Intelligence Technology and Innovation, 2024, 2(2), 35-52. DOI: https://doi.org/10.30212/JITI.202402.008.