[1] Sneha Mishra
[2]Umesh Chandra Jaiswal

Ensemble Methods with Statistics and Machine Learning on the Class Imbalance Problems of EEG data

Journal of Electrical Systems

Abstract: -Class imbalance in EEG data sets is a significant issue that affects the quality of outputs. The uncertainty in the data sets, which can be small or large, can lead to class imbalance problems (CIP). This imbalance can lead to highly imbalanced predictive models. The selection of random samples for algorithms can result in high variation of classes. Data sets of EEG are generated as image data sets and are often random and never repeat, causing a high variation in classes. To address this issue, various sampling methods on the data and heuristics are being developed to develop predictive models which were in vogue based on the level of the uncertain state.

Keywords: Class imbalances, Deep Learning, EEG Image Data, SMOTE.

I. INTRODUCTION

***Class Imbalance Problem* [1]**

Machine learning models often face class imbalance in various tasks such as fraud detection, disease diagnosis, network intrusion, face recognition, fault detection, spam filtering, credit score, forecasting uncommon events, and object detection. These issues can lead to models performing worse on underrepresented minority classes and better on majority classes. Factors such as fraud detection, disease diagnosis, intrusion into the network, face recognition, fault detection, spam filtering, credit score, uncommon events, and object detection can be complex for models to accurately identify and classify individuals. To address these imbalances, machine learning models must focus on addressing the imbalance through specific procedures and addressing the bias in their predictions.

Several statistical strategies and metrics can effectively address class imbalance in machine learning. These include prior probability, sampling techniques like under sampling, oversampling, and SMOTE [2], cost-sensitive learning, precision, recall, and F1-score, ROC curves, precision-recall curves, lift charts, cumulative lift, Kolmogorov-Smirnov charts, and hypothesis testing. These methods help manage unbalance in Bayesian models, alter class distribution within training data, and highlight the significance of minority groups. Precision, recall, and F1-score are superior measures than accuracy when dealing with imbalanced classes. ROC curves [3] illustrate the balance between true positives and false positives, while Precision-Recall curves compromise accuracy and completeness. Lift charts show the enhancement in prediction value relative to a baseline random model, while Kolmogorov-Smirnov charts display empirical cumulative distribution functions.

To address class imbalance in datasets used for image classification, various strategies are employed. These include data-level solutions like oversampling or under sampling, algorithmic solutions like weighted cross-entropy, transfer learning, data augmentation, ensemble approaches, exclusive models, region-based object detectors, semi-supervised learning, metric learning, and generative models. Oversampling or under sampling techniques alter class distributions within the training data, while algorithmic solutions use loss functions like weighted cross-entropy to allocate greater weights to minority class instances. Transfer learning uses pre-trained models to provide valuable feature representations, even for classes with less data. Data augmentation involves generating more photos of the minority class through various transformations. Ensemble approaches train multiple models on balanced data, enhancing the ability to identify and classify minority groups. Generative models use Generative Adversarial

1, Research Scholar, Department of Information Technology and Computer Application, Madan Mohan Malaviya University of Technology, Gorakhpur-273010, UP, India.
2, Professor, Department of Information Technology and Computer Application, Madan Mohan Malaviya University of Technology, Gorakhpur-273010, UP, India
Corresponding author E-mail address: snehamishra216@gmail.com

Networks (GANs) to create artificial instances of underrepresented categories and equalize the training dataset. Additional metrics are necessary to accurately assess performance beyond accuracy.

Image classification in the domain of EEG data sets face overwhelming class imbalance problems in varied orders, despite all the significant progress about deep learning. Actually, within extremely small subgroups of the overall EEG data sets, the features of the images that are anticipated to be present are found. A great number of dynamic correlations are involved in the process of classification. These correlations are tied to a specific certainty and state that exists over the course of time. Patients may have different explanations for the non-uniformity of the graphs, and the patterns that are present in the EEG data sets all have their own unique characteristics. It is believed to analyze deep learning models, while the irregularities are brought about by the ambiguity that exists inside the classes. Classification of EEG data has symptomatic peculiarities with regard to pre-ictal and ictal categories, though these characteristics appear in minor amounts in the data sets and are not considered to be as anomalies, however the classifier shall have to be formulated with such criteria.

Therefore, the commonality and rarity of the traits in the positive classes can be ignored by a large sample with diverse features.

The incorporation of real-time data presents challenges for categorization, classification and annotation in deep learning. Classification in deep learning is the process of categorizing data sets based on a decision-making problem. Real-time data is composed of several types of information. The features in real-time data sets frequently indicate qualities that are typically not taken into account during deep learning computations. Decision-making imbalances arise from the uncertain presence of particular attributes, when dataset instances are classified differentially into one target class and another class [1][2]. As a result, handling unbalanced datasets in deep learning is linked to the issue of class imbalance. The target class with positive presence—that is, a positive class that might have less high intensity instances—is another characteristic of the analysis on the issue of uncertainty and the class imbalance. These are also known as positive occurrences or minor examples; they are smaller yet have high intensities in the positive class.

Improving prediction rates and incorporating ambiguous positive cases into the positive class are the goals of addressing class imbalances. Numerous techniques are put forth in the literature to achieve realistic and efficient prediction rate convergence. The most common approaches for addressing class imbalances involve sampling strategies to convert unbalanced datasets into dispersed datasets that are well-balanced and appropriate for deep learning classifiers [4]. Most sampling strategies [2] provide advantages since they can be easily applied to many situations, allowing researchers to use different classifier categories without limitations. The class imbalance problems can be solved with a rudimentary initiative of using sampling methods which contains the methods for duplicating positive examples in the positive classes until all the number of positive instances are equal to the number of negative instances of the source data sets which is the candidate for classification. The knowledge about each distinct instance is used in composing the classifier, together with all the common characteristics that are unclear and contribute to an uneven distribution; it must also exercise caution to avoid excessive fitting to the training data. In general, positive examples are grouped together with other positive instances around the perimeter by inheriting the features that are suitable for classifying the instance as positive [2].

Machine learning solutions for class imbalance include random oversampling, random undersampling, SMOTE, and data cleansing. Algorithmic solutions include weighted loss functions, threshold adjustment, and ensemble approaches. Kernel-based learning employs kernel functions to map data into a higher-dimensional space, while cost-sensitive learning takes into account the expenses linked to misclassifications or mistakes. Sophisticated sampling involves near-miss sampling [5] and selecting informative subsets of information. The optimal strategy depends on the dataset and algorithms used, but often combining data-level and algorithm-level solutions is the most effective approach. Ensemble approaches integrate multiple models to enhance minority group identification [6]. Kernel-based learning uses one-class SVM and KPCA for feature transformation [7]. Cost-sensitive learning considers the cost associated with misclassifications or errors during training

II. BACKGROUND

Class Imbalance Problem with Deep Learning: Deep learning strategies can address class imbalance by re-weighting the loss function, applying class balancing techniques, adjusting thresholds, focusing on complex examples, utilizing data augmentation, training multiple models, using binary classifiers for each class, and using metric learning for unbalanced distributions [8]. Class-specific batch sampling and early stopping [9] can also be employed. The effectiveness of these strategies depends on network architecture, loss function, and degree of imbalance, with a combination of strategies often being most effective.

EEG data presents class imbalance due to infrequent aberrant neurological events. Optimizing for accuracy may not identify minority groups, while scarcity of data can lead to overfitting. Variability in EEG signals contributes to uncertainty. Loss reweighting, oversampling, and ensemble techniques can address imbalance, while Bayesian models or Monte Carlo dropout can enhance dependability.

EEG data and picture classification face class imbalance and uncertainty due to the infrequent occurrence of aberrant neurological events and variability in signals. Techniques like loss reweighting, oversampling, and ensemble techniques can address imbalances. Bayesian models or Monte Carlo dropout can provide insights into uncertainty measurement. Image classification faces imbalanced distribution of object classes and model uncertainty, especially for minority classes with low training samples. Data augmentation and focal loss can help mitigate overfitting and improve model trustworthiness.

Wang et al.[10] in their work, analyzes different oversampling and undersampling approaches for imbalance with deep learning. Besides, it also proposes the methods that evaluates oversampling and undersampling techniques for convolutional neural networks on image datasets. It is found that replication oversampling improved accuracy on minority classes but not overall performance. SMOTE oversampling improved performance by generating synthetic minority class examples. Ensemble undersampling improved performance by training multiple models on balanced class subsets.

Cui, Y. et al., [11] proposes a class balancing loss function that considers effective sample size for each class. Further, proposes a new loss that weights classes based on their "effective number of samples", considering inter-sample distances within a class to estimate diversity. The loss is integrated into the deep network architecture for image classification and shows state-of-the-art performance on long-tailed datasets. It is compatible with existing remedies like data resampling and re-weighting.

Khan, S. et al.[12] in their article illustrates using uncertainty to dynamically weight loss and perform data undersampling. Khan et al. also proposes an uncertainty-based class imbalance learning method that uses predictive uncertainty to dynamically weight loss and perform data undersampling. The method is robust and adaptive, preventing overfitting to majority classes and reducing bias towards majority classes. Its key strengths include its fully automated balancing approach and adaptability.

Buda, M. et al.[13] evaluates multiple techniques including oversampling, thresholding and loss weighting. Buda et al.'s study on class imbalance in convolutional neural networks reveals that oversampling, focused undersampling, and cost-sensitive training are effective in handling class imbalance. Combining these techniques improves performance. The study also highlights that no single technique is universally ideal, and assembling classifiers trained with different methods is most robust.

In the studies by Johnson and Khoshgoftaar [14] in the Journal of Big Data (2019) provides an overview of deep learning techniques for class imbalance across image, text, and speech applications. It discusses data-level, algorithm-level, and hybrid approaches, including sampling, loss weighting, and transfer learning. The study highlights the importance of hyperparameter tuning and generative adversarial networks for effective solutions.

Wang et al. [15] proposes a self-supervised contrastive loss objective to address class imbalance in images. The loss involves augmented positive and negative samples, with minority classes contributing more. This embedding exposes a linear classifier to balanced class separability and can be combined with supervised cross-entropy loss. It offers greater generalization and flexibility in class representations. Skewness of the distribution, rarity, size of samples determines the nature of the classes that are derived while solving the class imbalance problem. Certain parameters related to class imbalance problem are summarized and tabulated below:

Table 1: A general summary of class imbalance problems

Ref.	Skewness in Class Distribution	Rarity (%)	Sample Sizes (No.s)	Class Nature
Wang et al. [15]	mean < median, skew is left	40% - 60%	10	minority
Cui, Y. et al. [11]	mean < median, skew is left	20% - 50%	15	minority
Khan, S. et al [12]	mean < median, skew is left	20% - 70%	8	minority
Buda, M. et al [13]	mean < median, skew is left	15% - 75%	6	minority

Table 2: A general summary of class imbalance problems – after application ML algorithms

Ref.	Skewness in Class Distribution	Rarity (%)	Sample Sizes (No.s)	Class Nature
Wang et al. [15]	mean ≈ median	55%	10	balanced classes
Cui, Y. et al. [11]	mean ≈ median	65%	12	balanced classes
Khan, S. et al [12]	mean ≈ median	60%	10	balanced classes
Buda, M. et al [13]	mean ≈ median	75%	15	balanced classes

Class Imbalance Problem with Statistics: A class imbalance occurs when there are substantially fewer instances of one or more classes in a dataset than there are of other classes. It is possible that as a consequence of this, machine learning models will become biased in favor of the majority class, rendering them less effective in dealing with the minority class. Statistical methods [16] include class distribution analysis, resampling, weighted classes, ensemble methods, evaluation metrics, cost-sensitive learning, stratified sampling, anomaly detection, cost-benefit analysis, and transfer learning can be applied to address this. It is crucial to take into account the particulars of your dataset and the issue you are attempting to resolve. You should also test out various strategies to see how they affect the model's functionality.

Wang, Minku, et al. [17] (2021) introduces two new ensemble approaches called Online Under-Over-Bagging (OUOBag) and Online Under-Over-Boosting (OUOBoost) for handling class imbalance in online learning scenarios. These methods combine undersampling of majority classes with oversampling of minority classes, and are designed to adapt to changing class imbalance ratios in non-stationary environments. They achieve better AUROC and G-mean compared to existing methods.

Class Imbalance Problem in EEG Data: To address dataset imbalance [18], Martins et al. [19] in their projects has employed a DA stage in an automatic PPR detection technique. A cross-validation based approach is applied on PPR and non-PPR windows, where non-PPR windows are found under sampled, while to attain the increased number of PPR windows. Prior to applying DA, the raw EEG data were windowed and labelled. ML models 2C-KNN and DL-NN were trained and evaluated using the final data. In spite of increased morphological variability brought on by various waveforms, DA enhanced detection outcomes. While both approaches worked well, DL-NN produced superior outcomes.

In the works of Aslam et al. [20] the dataset is cleaned of noise, for example, a sliding window of size 30 is applied with selected 50% which also overlaps. Short Time Fourier Transform (STFT) is typically used in improving of the signal-to-noise ratio (SNR). In order to solve the class imbalance that exists between the preictal and interictal phases, the overlapping window is utilized in preictal data. A non-overlapping window of size 30 is utilized in order to translate the interference data into the frequency domain.

III. RIVIEW OF CIP SOLUTIONS

The Class Imbalance Problem (CIP) is a significant challenge in machine learning and statistics, where one class outnumbers another, leading to biased models. Solutions involve strategies to balance distribution or improve learning processes for minority classes.

Tanimoto et al[21] in their paper introduces a novel problem setting, imbalanced classification with positivity, and proposes a versatile method to address it. The technique is more adaptable than current regression or rank-based methods and outperforms traditional cost-sensitive classification in severely unbalanced situations by taking use of near-miss positive cases to decrease class imbalance.

Hoyos et. al [22] in their paper presents the Relevant Information-based UnderSampling (RIUS) technique, which is intended to enhance classification performance in situations when the data is unbalanced. With fewer samples, RIUS extracts the underlying structure of the majority class by applying the information-preservation principle. To improve data representation, it is used with the Clustering-based Undersampling method (CBUS). According to experimental findings, RIUS and CRIUS lessen information loss and disclose the relevant structure of the data.

Thabtah et al [23] has acknowledged that the class imbalance problem and uncertainty significantly affect the raw data used in the models for detecting fraudster operations, intrusion detection in networks, and medical diagnosis. Because they concentrate on the majority class while neglecting the minority class, modern machine learning algorithms find it difficult to address this challenge.In this work, the purpose is to highlight the implications of class imbalance on classification models by analyzing the effects of various class imbalance ratios on the performance of classifiers. The study attempts to assist researchers in more efficiently addressing this issue.

As opined by Zheng et al., [24] uncertainty in data is a common issue in both synthetic frameworks of theoretical research and also in related real-world applications. Conventional classification algorithms have difficulties in effectively learning and making predictions from unbalanced data. To address this issue, they commonly use techniques such as oversampling, undersampling, or hybrid sampling approaches. Nevertheless, existing techniques often use arbitrary sample ratios, resulting in unreliable classification performance. In order to enhance the accuracy of classification, three genetic algorithms were suggested to autonomously ascertain the appropriate sampling ratios for oversampling, undersampling, and hybrid techniques.

In their study Bria et. al [25] presents a two-stage deep learning method for managing class imbalance in tiny lesion detector training. The process includes imparting a complicated set of decision trees by utilizing an algorithm that reduces the total amount of background samples. Further to this, the remaining data are transferred to a Convolutional Neural Network (CNN), which is able to reap the benefits of the Data Center's painstaking extraction and redistribution, respectively. Microcalcification detection and microaneurysm detection were the two examples of classification problems that were investigated in this work. Both of these problems had a considerable imbalance in the distribution of their data. On both occasions, the DC-CNN demonstrated superior performance in comparison to CNNs that were trained using techniques such as one-class classification, hard mining, cost-sensitive learning, and either over- or undersampling. According to the results of the test, the DC-CNN was nearly ten times faster than CNN.

Statistical Methods for Ensemble Classification: Ensemble classification methods include bagging, boosting, stacking, Bayesian Model Averaging, Negative Correlation Learning, Dempster-Shafer Theory, and Mean Regularized Committee Machines [26]. These techniques combine multiple classifiers into an improved model, aiming to promote diversity among base learners. Bagging reduces variance by resampling training data and voting for the final prediction. Boosting sequentially trains classifiers, focusing on correcting errors from the previous ensemble. Stacking trains meta-learners to combine predictions from multiple base learners.

Li et. al [27] in their research evaluates ensemble learning models for credit scoring, comparing their performance in accuracy. The AUC and KS statistic assess the discriminative power of a model, which is crucial in unbalanced classification tasks. The Brier score evaluates the accuracy of projected probability in terms of calibration. Operational duration quantifies the effectiveness of computing. Model operational time quantifies the pace at which inferences are made during deployment. It is crucial to monitor, particularly for applications that need immediate data updates. The time taken for each prediction is reported in units such as seconds or milliseconds. A lower value is preferable. The Brier score quantifies the degree of calibration of anticipated probability. Values vary from 0 for flawless calibration to 1 for bad calibration. A lower value is preferable. Assists in evaluating the dependability of probability generated by a model.AUC, which stands for "area under the ROC curve," is a method that evaluates the discriminative capacity of a classifier without regard to any threshold. It is helpful for situations with unbalanced classes. Values that are closer to one imply that the positive and negative groups are differentiated more clearly. The Kolmogorov-Smirnov (KS) statistic measures the disparity between the score distributions of the positive and negative classes. Greater values imply superior differentiation between the classes. Graphically represented using a Kolmogorov-Smirnov plot.Five baseline classifiers are used as benchmarks. Experimental findings of the article by the Li et. al [27] show ensemble learning outperforms individual learners, except for AdaBoost. Random forest has the best performance in five metrics, while XGBoost and LightGBM are close competitors

Dutta et. al [28] in their article presents an ensemble method for network anomaly detection, leveraging deep models like Deep Neural Network (DNN) and Long Short-Term Memory (LSTM) and a meta-classifier (logistic regression) using the principle of stacked generalization. This study presents an ensemble approach incorporating deep learning algorithms and stacked generalization for effective anomaly-based network intrusion detection. Using feature engineering methods and dimensionality reduction, the proposed stacked ensemble framework outperformed state-of-the-art individual classifiers and meta-classifiers. The framework can eliminate the challenge of providing recent network traffic datasets and provide acceptable accuracy in detecting anomaly behaviors. Traditional machine learning methods are inefficient due to the vast network traffic in critical Cyber-Physical Systems (CPSs).

Table 3: A Review of Solutions for Class Imbalance Problems

Ref.	**Method**	**Classification Accuracy**	**Recall**		**F-Measure**	**AUROC**	**Training Stability** *(min 100 epochs)*	**Prediction Probability** *(min 100 epochs)*
			Minority	**Majority**				
Tanimoto et al [21]	Near-Miss Positive Instances	84%	0.78	0.76	0.86	0.79	80%	60%
Hoyos et. al [22]	Relevant Information-based UnderSampling coupled with Clustering-based Undersampling algorithm	87%	0.64	0.62	0.92	0.83	85%	70%
Thabtah et al [23]	Study on Varying Classifier Accuracies	Varied Accuracies	min 0.53 (approx)	min 0.42 (approx)	0.66	0.89	90%	80%
Zheng et al [24]	Genetic Algorithms	92%	0.72	0.74	0.94	0.92	90%	90%
Bria et. al [25]	Two-Stage Deep Learning Method for Managing Class Imbalance	94%	0.76	0.72	0.92	0.84	90%	90%
Dutta et. al [28]	Concept of Meta-Classification using Deep Learning Techniques	Varied Accuracies	min 0.48 (approx)	min 0.44 (approx)	-	-	-	-

Therefore, ensemble methods, which combine multiple models, can address class imbalance problems using statistical techniques. In summary, these methods include bagging, boosting, stacking, Bayesian model averaging, negative correlation learning, Dempster-Shafer theory, and regularization methods. The techniques improve diversity and balance between classes, mitigate individual model biases, and use resampling, meta-learning, Bayesian statistics, and regularization to achieve better performance.

3.1 Data Sets

EEG data sets are used in machine learning tasks like brain-computer interfaces and seizure detection, but class imbalance is common due to abnormal EEG patterns [29]. Small sample sizes for minority classes and challenges in collecting abnormal data lead to skews. Data resampling techniques like SMOTE and ADASYN [30] are used to oversample minority classes, while ensemble and active learning methods improve minority class recognition. Careful dataset design, resampling, and algorithm selection are crucial for handling EEG data imbalance.

IV. COMPRATIVE STUDY

According to the scientific consensus, one of the most important aspects of comparisons is the algorithm level and data level approaches for machine learning. On the other hand, ensemble approaches using statistics are available as data level approaches. Ensemble approaches that apply core statistical properties are available as data-level approaches. Data-level approaches, including over-sampling, under-sampling, and hybrid data-level approaches [31], alter the data used for training rather than the learning method that is followed in order to achieve the goal of rebalancing the class distributions.

Algorithm-level methods improve classifiers by modifying algorithms, with CSL being the most natural modification [32]. It assigns varying costs to different classes, challenging to determine the appropriate misclassification cost.

To develop a cost-sensitive learning model of machine learning [33], a dataset shall be identified criticallyandestimate of the cost of the classes. Define class-specific costs for each class, representing the consequences of misclassification. Integrate the cost matrix into the classifier's training process. An appropriate classification algorithm shall be chosen that supports cost-sensitive learning. Use cross-validation techniques to evaluate the model's effectivenessandfine-tune the hyperparameters [34] of the chosen algorithm are considered in cost-sensitive learning. The performance of the model is evaluated iteratively using the metrics until desired balance is achieved amongst accuracy and cost of misclassification.

The table shows some of the comparative measures that are considered between statistics-based ensemble methods and machine learning methods in developing solutions for CIP.

Table 4: Table illustrating some of the salient comparative properties

S.No.	Machine Learning methods	Ensemble based Statistics methods
1	From the computed FP, TP, FN and TN in ML, Precision, Recall, F-Measure, Sensitivity, Specificity and Accuracy in ROCAUC are compared with the methods.	A Kappa statistic shall be calculated for the Expected Accuracy and Observed Accuracy. Mathews Correlation Coefficient (MCC) shall be statistically useful to determine the
2	Robustness of the Classes is computed based on the AUC and PR curve.	Gradient Boosting and Optimization methods shall be employed to determine the robustness of the classes.
3	Data Preprocessing and Data Augmentation is essential in order to rule out the outlier data for the determination of classes, which maintains stability.	Perturbations to the Datasets with random noise of different levels is introduced to test the stability, whether the classes remain same even after introducing the random noise
4	Transfer Learning helps in assessing the model's ability, to consider the data transfer on to the imbalanced dataset and solve the problem if the datasets have the nature "same but distinct".	An Experimental design with adaptive learning shall be employed to dynamically update the data distributed changes.

V. RECOMMENDATIONS

Ensemble methods in statistics enhance predictive performance and robustness of statistical models [35]. A systematic approach involves defining the problem, gathering data, selecting base models, and implementing ensemble methods. Performance is evaluated by comparing it to individual models, considering bias-variance trade-off, interpretability, and robustness. Real-world deployment and documentation of the design improve predictive power and reliability.

The performance of the ensemble methods for the statistical application shall be experiment and evaluated with domain-specific models [36]. There is a scope of analyzing important factors such as the bias-variance trade-off, interpretability and explainability, and further to assess the ensemble's robustness by introducing noise or perturbations to the data.

An issue that frequently arises in supervised learning methods, such as deep CNNs [37], is class imbalance. A class-balanced ensemble is presented in this study [37], which is derived from a single CNN model and is applied to large-scale unbalanced data. Experiments show the method's effectiveness on different imbalance levels and alleviate the vanishing gradient problem.

Ensemble methods in machine learning solve class imbalance problems by combining multiple base models to improve classification performance and mitigate biases [38]. Common techniques include random oversampling, undersampling, SMOTE, and more. Cost-sensitive learning penalizes minority misclassification, while stacking and blending combine multiple base classifiers. Hybrid methods combine strategies for specific imbalance problems. Experimentation is crucial for understanding model performance.

To address class imbalance issues in machine learning, experimental designs using statistical methods should consider the limited availability of data and fluctuations in minority classes. This involves employing stratified sampling to maintain proportional representation of classes, with most promising measures to assess the quality of the process such as F1-score, precision and recall, and statistical tests like McNemar's test to assess disparities in classifier performance across minority classes. Statistical bootstrapping methods can be used for resampling on unbalanced data, and hypothesis testing on metrics across classifiers and sampling approaches can be conducted using t-tests and ANOVA. Visualizing model scores' distributions using QQ plots and Kolmogorov-Smirnov charts can help assess the separation of model scores. Metrics like Brier score, reliability diagrams, and Expected Calibration Error can be used to assess the calibration of the classifier. Uplift modeling can be employed to measure incremental improvement over a baseline model, and synthetic sample generation methods like SMOTE can be compared based on produced metrics.

Hence, the statistical methods are ideal for solutions of class imbalance problem with domain-specific data sets, whereas, the machine learning models contribute to the unsupervised and supervised methods of building a model to escape with biased data sets.

VI. CONCLUSIONS

Ensemble methods and statistics are essential tools for EEG data analysis in various fields such as brain-computer interfaces, neurology, and cognitive science. These methods include Random Forests, Gradient Boosting [39], Neural Networks, and Stacking [40]. Descriptive statistics summarize EEG data, while time-frequency analysis provides statistical information about EEG signals. Statistical tests can reveal significant differences in brain activity. Correlation analysis helps identify patterns and changes in brain activity. Feature selection and multivariate analysis reduce the dimensionality of EEG data while preserving relevant information. Challenges in EEG data analysis include noise handling, class imbalance, temporal dynamics, and inter-subject variability. Interpretability and visualization techniques can help gain insights into the brain's response patterns. Cross-validation techniques can assess the performance and generalization of ensemble models and statistical analyses. The choice of methods and statistical tests should be tailored to the research goals and the nature of the EEG data collected. Collaboration with domain experts can be beneficial for informed EEG data analysis.

A general summary of the experiments in the research consensus by eminent authors has been annotated in Table 1 and the intervention of machine learning to the solutions of class imbalance problems are stated in Table 2. Table 3 illustrate some of the performance measures that are observed from the collected articles, which conclude with measures shall be taken for improvising the model to mitigate with the varying class accuracies while demonstrating the new classes. Machine learning algorithms with evolutionary methods shall be a good recommendation for solving CIP on biased datasets, statistics and ensemble based methods shall be recommended for domain-specific known large datasets.

References

[1] Japkowicz, Nathalie, and Shaju Stephen. "*The Class Imbalance Problem: A Systematic Study.*" Intelligent Data Analysis 6, No. 5 (2002): 429-449.

[2] Fernández, Alberto, Salvador Garcia, Francisco Herrera, and Nitesh V. Chawla. "*SMOTE for learning from imbalanced data: progress and challenges, marking the 15-year anniversary.*" Journal of Artificial Intelligence Research Vol. 61 (2018): 863-905.

[3] Fawcett, Tom. "*An introduction to ROC analysis.*" Pattern recognition letters 27, no. 8 (2006): 861-874.

[4] Johnson, Justin M., and Taghi M. Khoshgoftaar. "*Survey on deep learning with class imbalance.*" Journal of Big Data 6, No. 1 (2019): 1-54.

[5] Gnoni, Maria Grazia, and Joseph Homer Saleh. "*Near-miss management systems and observability-in-depth: Handling safety incidents and accident precursors in light of safety principles.*" Safety Science 91 (2017): 154-167.

[6] Zhou, Todd, and Hong Jiao. "*Exploration of the stacking ensemble machine learning algorithm for cheating detection in large-scale assessment.*" Educational and Psychological Measurement 83, no. 4 (2023): 831-854.

[7] Kallas, Maya, Clovis Francis, Lara Kanaan, Dalia Merheb, Paul Honeine, and Hassan Amoud. "*Multi-class SVM classification combined with kernel PCA feature extraction of ECG signals.*" In 2012 19th International Conference on Telecommunications (ICT), pp. 1-5. IEEE, 2012.

[8] Bria, Alessandro, Claudio Marrocco, and Francesco Tortorella. "*Addressing class imbalance in deep learning for small lesion detection on medical images.*" Computers in biology and medicine 120 (2020): 103735.

[9] Reiss, Tal, Niv Cohen, Liron Bergman, and YedidHoshen. "*Panda: Adapting pretrained features for anomaly detection and segmentation.*" In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 2806-2814. 2021.

[10] Wang, Shoujin, Wei Liu, Jia Wu, Longbing Cao, QinxueMeng, and Paul J. Kennedy. "*Training deep neural networks on imbalanced data sets.*" In 2016 international joint conference on neural networks (IJCNN), pp. 4368-4374. IEEE, 2016.

[11] Cui, Yin, MenglinJia, Tsung-Yi Lin, Yang Song, and Serge Belongie. "*Class-balanced loss based on effective number of samples.*" In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 9268-9277. 2019.

[12] Khan, Salman, Munawar Hayat, Syed WaqasZamir, JianbingShen, and Ling Shao. “*Striking the right balance with uncertainty*.” In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 103-112. 2019.
[13] Buda, Mateusz, Atsuto Maki, and Maciej A. Mazurowski. “*A systematic study of the class imbalance problem in convolutional neural networks.*” Neural networks 106 (2018): 249-259.
[14] Johnson, Justin M., and Taghi M. Khoshgoftaar. “*Survey on deep learning with class imbalance.*” Journal of Big Data 6, no. 1 (2019): 1-54.
[15] Zhu, Jianggang, Zheng Wang, Jingjing Chen, Yi-Ping Phoebe Chen, and Yu-Gang Jiang. “*Balanced contrastive learning for long-tailed visual recognition.*” In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 6908-6917. 2022.
[16] Menon, Aditya, HarikrishnaNarasimhan, ShivaniAgarwal, and Sanjay Chawla. “*On the statistical consistency of algorithms for binary classification under class imbalance.*” In International Conference on Machine Learning, pp. 603-611. PMLR, 2013.
[17] Wang, Shuo, Leandro L. Minku, and Xin Yao. “*Resampling-based ensemble methods for online class imbalance learning.*” IEEE Transactions on Knowledge and Data Engineering 27, no. 5 (2014): 1356-1368.
[18] Fan, Jiahao, Chenglu Sun, Chen Chen, Xinyu Jiang, Xiangyu Liu, Xian Zhao, Long Meng, Chenyun Dai, and Wei Chen. “*EEG data augmentation: towards class imbalance problem in sleep staging tasks.*” Journal of Neural Engineering 17, no. 5 (2020): 056017.
[19] Martins, Fernando Moncada, Víctor Manuel González Suárez, José Ramón VillarFlecha, and Beatriz GarcíaLópez. “*Data Augmentation Effects on Highly Imbalanced EEG Datasets for Automatic Detection of Photoparoxysmal Responses.*” Sensors 23, no. 4 (2023): 2312.
[20] Aslam, Muhammad Haseeb, Syed Muhammad Usman, Shehzad Khalid, Aamir Anwar, RoobaeaAlroobaea, Saddam Hussain, JasemAlmotiri, Syed SajidUllah, and AmanullahYasin. “*Classification of EEG signals for prediction of epileptic seizures.*” Applied Sciences 12, no. 14 (2022): 7251.
[21] Tanimoto, Akira, So Yamada, Takashi Takenouchi, Masashi Sugiyama, and Hisashi Kashima. “*Improving imbalanced classification using near-miss instances.*” Expert Systems with Applications 201 (2022): 117130.
[22] Hoyos-Osorio, J., A. Alvarez-Meza, GenaroDaza-Santacoloma, A. Orozco-Gutierrez, and GermánCastellanos-Dominguez. “*Relevant information undersampling to support imbalanced data classification.*” Neurocomputing 436 (2021): 136-146.
[23] Thabtah, Fadi, SuhelHammoud, FiruzKamalov, and Amanda Gonsalves. “*Data imbalance in classification: Experimental evaluation.*” Information Sciences 513 (2020): 429-441.
[24] Zheng, Ming, Tong Li, Liping Sun, Taochun Wang, Biao Jie, Weiyi Yang, Mingjing Tang, and Changlong Lv. “*An automatic sampling ratio detection method based on genetic algorithm for imbalanced data classification.*” Knowledge-Based Systems 216 (2021): 106800.
[25] Bria, Alessandro, Claudio Marrocco, and Francesco Tortorella. “*Addressing class imbalance in deep learning for small lesion detection on medical images.*” Computers in biology and medicine 120 (2020): 103735.
[26] Hamza, Mounir, and Denis Larocque. “*An empirical comparison of ensemble methods based on classification trees.*” Journal of Statistical Computation and Simulation 75, no. 8 (2005): 629-643.
[27] Li, Yiheng, and Weidong Chen. “*A comparative performance assessment of ensemble learning for credit scoring.*” Mathematics 8, No. 10 MDPI (2020): 1756.
[28] Dutta, Vibekananda, MichałChoraś, MarekPawlicki, and RafałKozik. “*A deep learning ensemble for network anomaly and cyber-attack detection.*” Sensors 20, No. 16 MDPI (2020): 4583.
[29] Daftari, Charmi, Jainish Shah, and Manan Shah. “*Detection of epileptic seizure disorder using EEG signals.*” In Artificial intelligence-based brain-computer interface, pp. 163-188. Academic Press, 2022.
[30] Halim, Anthony Mas, MahendraDwifebri, and FhiraNhita. “*Handling Imbalanced Data Sets Using SMOTE and ADASYN to Improve Classification Performance of Ecoli Data Sets.*” Building of Informatics, Technology and Science (BITS) 5, no. 1 (2023): 246-253.
[31] Lin, Cian, Chih-Fong Tsai, and Wei-Chao Lin. “*Towards hybrid over-and under-sampling combination methods for class imbalanced datasets: an experimental study.*” Artificial Intelligence Review 56, no. 2 (2023): 845-863.
[32] Mienye, IbomoiyeDomor, and Yanxia Sun. “*Performance analysis of cost-sensitive learning methods with application to imbalanced medical data.*” Informatics in Medicine Unlocked 25 (2021): 100690.
[33] Fernández, Alberto, Salvador García, MikelGalar, Ronaldo C. Prati, BartoszKrawczyk, Francisco Herrera, Alberto Fernández et al. “*Cost-sensitive learning*.” Learning from imbalanced data sets (2018): 63-78.
[34] Elgeldawi, Enas, AwnySayed, Ahmed R. Galal, and Alaa M. Zaki. “*Hyperparameter tuning for machine learning algorithms used for arabic sentiment analysis*.” In Informatics, vol. 8, no. 4, p. 79. MDPI, 2021.
[35] Seni, Giovanni, and John Elder. “*Ensemble methods in data mining: improving accuracy through combining predictions*”, Morgan & Claypool Publishers, 2010.
[36] Acar, Esra, Frank Hopfgartner, and SahinAlbayrak. “*A comprehensive study on mid-level representation and ensemble learning for emotional analysis of video material.*” Multimedia Tools and Applications 76 (2017): 11809-11837.
[37] Johnson, Justin M., and Taghi M. Khoshgoftaar. “*Survey on deep learning with class imbalance.*” Journal of Big Data 6, No. 1 Springer (2019): 1-54.

[38] Pirizadeh, Mohsen, NafisehAlemohammad, Mohammad Manthouri, and MeysamPirizadeh. "*A new machine learning ensemble model for class imbalance problem of screening enhanced oil recovery methods.*" Journal of Petroleum Science and Engineering 198 (2021): 108214.

[39] Feng, Ji, Yang Yu, and Zhi-Hua Zhou. "*Multi-layered gradient boosting decision trees*." Advances in neural information processing systems 31 (2018).

[40] Divina, Federico, Aude Gilson, Francisco Goméz-Vela, Miguel García Torres, and José F. Torres. "*Stacking ensemble learning for short-term electricity consumption forecasting.*" Energies 11, No. 4 MDPI (2018): 949.