International Journal of Computing and Digital Systems
2025, VOL.17, NO.1, 1-11

http://dx.doi.org/10.12785/ijcds/1571110606

Prediction of Drug Risks Consumption by Using Artificial Intelligence Techniques

Ruba Talal Ibrahim[1] and Hanan Anas Aldabagh[2]

[1]*Department of Computer Science,University of Mosul, Mosul,Iraq*
[2]*Department of Computer Science,The General Directorate of Education in Nineveh Governorate,Mosul,Iraq*

Received 19 September 2024, Revised 30 January 2025, Accepted 31 January 2025

Abstract: Drug abuse and addiction have reached unprecedented heights, destroying and weakening society. It is a dangerous and deadly weapon that has a major impact on individuals. Clinical evaluation by experts is the most common method for diagnosing addicted patients and isolating them, but this requires equipment, tools, and human effort. Therefore, in this paper, a new hybridization model (EXT- HBOS) between supervised algorithm (Extra tree) and unsupervised algorithm (histogram-based outlier scores) as well as many states of art machine learning techniques (Extremely Randomized Trees, Cat Boost and Light Gradient Boosting Machine) were used to predict drug-addicted patients based on survey online dataset from Kaggle. The dataset was analyzed, discussed, and rebalanced using random oversampling, also the Grey Wolf Optimization (GWO) algorithm was used for tuning important hyperparameters and get the best one. The results were analyzed and discussed using different performance and statistical methods and showed that the hybrid model (EXT- HBOS) did the best on all measures. It gained 90% accuracy score and 74% Cohen's kappa score. Also, The results illustrated that Neuroticism (Nscore) is the most important factor that tempts an individual to abuse drugs such as heroin.

Keywords: Artificial Intelligent, Drug, Heroin, Grey Wolf Optimization, Machine Learning, Prediction, Extra tree, histogram-based outlier scores

1. INTRODUCTION

Drug abuse is an epidemic that threatens the developed and advanced countries of the world, and its dangers do not stop at the borders of a particular country, and this is a fact that confirmed by social, psychological and health scientists. It has become a state that cannot be controlled, just as the use of electronic games and smartphones cannot be controlled, or the excessive consumption of alcohol [1].The National Centre for Drug Abuse stated in its latest statistics that 50% of people over the age of 12 have used drugs at least once, in addition to an increase in deaths of more than 700,000 people due to drug overdoses. Also, according to the National Centre, an alcohol drinking issue impacts 28.320 million individuals, as well as 20.4% of the population. Tobacco or nicotine products (vape) are used by 57.277 million people. A drug problem affects 25.4% of illegal drug users. Opioid problems impact 24.7% of individuals with drug addictions (this includes prescription pain pills or pain killers as well as heroin) [2]. Drug abuse (addiction) is a disease that affects a person's brain and behavior because it contains chemical substances that lead to the difficulty to manage the use of any legal or illegal drugs or medicine [3]. Drugs are a group of substances that cause addiction and poison the nervous system, and lead to drowsiness and sleep or lack of consciousness. Along with causing psychological and physical dependence and negative effects on both the individual and society, it can also lead to a change in personality, low functional and cognitive performance, or a sense of apathy, loss of correct judgment of things, family disintegration and divorce issues, or the spread of crimes to obtain money. There are two sources: natural (opium, morphine, cocaine, etc.), and synthetic (heroin, amphetamines, etc.).

Addiction to drugs can begin with the trial use of a pleasure drugs in certain social contexts, and it becomes more common for some people and more frequent than for others. Heroin is a commonly used opiate that causes mental and psychological illnesses, such as stress and depression, and can lead to death [4]. Clinical evaluation by experts is the most common way to diagnose and isolate patients [5], but this requires tools, equipment, and human effort, in addition to requiring a cash budget from the patients to reach the health centers. Therefore, many papers have turned to using artificial intelligence techniques in practical evaluation and diagnosis of dependencies. Digital healthcare is as accurate in conducting clinical examinations for drug abuse

E-mail address: rubatalal@uomosul.edu.iq, hanan_aldabagh1985@nan.epedu.gov.iq

as a machine learning and deep learning algorithm that processes big data under rigorous conditions [6].Machine learning algorithms have tremendous potential for exploring evolutionary horizons as they predict the results of future samples of large, unorganized, and variable data. Machine learning algorithms take important relevant features from the data after it has been cleaned, and analyzed, then train these features using the training process, and finally, testing their performance on a subset of the data. The goal of this paper is to apply state of art machine learning techniques to online survey data from Kaggle as a basic dataset to predict and treat drug-addicted patients because clinical treatment requires significant time and effort in addition to financial expenses, and the contributions of the paper are:

1) State-of-the-art machine learning techniques such as (LGBM, Extra Trees, Cat Boost, and Histogram-based outlier scores) were applied to predict drug-addicted patients.
2) A novel hybrid model called (EXT- HBOS) consisting of supervised machine learning (Extra tree) with unsupervised machine learning (histogram-based outlier scores) was applied.
3) The data were prepared by transforming and normalizing operations, and then was divided and re-balanced by using random oversampling.
4) The Gray Wolf Optimization Algorithm (GWO) was used for tuning the hyperparameters and getting the best one in all proposed models.
5) The results were discussed using performance and statistical measurements.

The rest of study organized as follows: Section 2 will discuss Related Works , Section 3 will present Background of Models, Section 4 will discuss Methodology, Section 5 will display Prediction, and Section 6 will present Results and Analysis. Finally, Section 7 will present the Conclusion.

2. Related Works

Recently, a number of researches have demonstrated the potential for employing machine learning techniques as a successful strategy to accurately and predictably estimate the risks of drug usage. In 2016 [7], researchers presented a study on the use of heroin and amphetamine using machine learning algorithms (Elastic Net). The goal of the study was to determine the number of participants who abused heroin alone, to identify participants who abused amphetamine alone, and to determine those who abused both, in addition to identifying non-participants. The study relied on many important characteristics (demographic, character, Psychiatric, problems, Neurocognitive, impulsivity), including psychological illness, which determined the number of abusers of heroin and amphetamine together. The results showed that the study achieved better results in AUC: 0.863 (heroin) and 0.712 (amphetamine).

In 2017 [8], researchers proposed a study on predicting treatment for substance use disorders using a number of machine learning algorithms (logistic regression, random forests, penalized regression, deep learning neural networks, and super learning). The paper was based on a national data set that included 99,013 respondents. The dataset have 10 characteristics have been used for the patient (age, sex, race, social condition, education, job condition, pregnancy at the time of acceptance, ancient warriors' condition, and living condition), 3 therapy features (i.e. severity, drug-backed ophthalmic treatment, residency duration), the primary source of referral, A brief description of the substance that is causing the problem (with other drugs only categories, alcohol only, or drugs and alcohol), as well as the problem of mental health. The algorithms were examined with a test sample. The AUC for the algorithms ranged from 0.793 and 0.820. They were outperformed by Super Learning, which was the first stage of targeted learning, a framework for analysis that produces double robust impact assessment and inference with a smaller set of assumptions than typical parametric approaches.

In 2019 [9], study predicted the abuse of two types of drugs and central stimulants (methamphetamine and amyl), which affect the general health of society. 12 personal attributes whole, comprising demographic data (age, gender, ethnicity, nationality and education level) and character traits, were included in the original dataset. This dataset was classified using a number of machine learning algorithms (Random Forest, XGBoost, and Light-GBM), which proved their effectiveness compared to the KNN algorithm Whereas XGBoost, and Light-GBM achieved an accuracy of 0.77.

In 2020 [10], researchers used artificial intelligence methods (Gradient Boosting and word2vec) for the early diagnosis of opioids. The paper analyzed the commercial claim dataset that contains details on both medical insurance claims and personal diagnosis from 2006 to 2018 into six groups (features) to obtain good results, as the sensitivity was 0.85 and the specificity was 0.88.

Also, in 2020 [11], researchers predict adults at risk for opioid use. Random forest and decision tree algorithms were used with down sampling to handle unbalanced classes. The prediction was made using the NSDUH dataset (National Survey on Drug Use and Health) that contain demographic data (gender, age, race), socioeconomic status, Physical Psychological data. The paper achieved sensitivity of 0.81 and specificity of 0.81.

In 2021 [12], researchers classified the use of heroin drug after identifying important features extracted from the NSDUH. The paper used three different methods from the Random Forest algorithm to explain how to extract features from an unbalanced medical dataset with multiple variables and achieved precision = 0.69 and an F1-score measure of 0.53.

In 2022 [13], researchers predicted the use of opioid drugs for People who suffer from attention deficit and hyperactivity using a number of supervised machine learning algorithms (Decision Tree, Decision Bayesian Classifier, Random Forest, and Improved Decision Tree). The paper relied on a dataset from NSDUH after collecting features from each observation in it, then re-cleaning, coding, and selecting the important features using an algorithm called

Chi-Square. The results showed that the improved decision tree obtained best accuracy of 99.21 because it was based on the Maclaurin method's approximation formula, which allows for the creation of a decision tree in a short period of time.

In 2023 [14], researchers classified drug usage into two groups using five machine learning algorithms: Gaussian Naive Bais, Random Forests, Decision Trees, Logistic Regression, and Nearest Neighbors. The open-source data "UCI repository" was used, and different results were obtained, ranging from (70-98%) to classify the consumption of different types of drugs: Alcohol, Amyl, Amphetamine, Benzos, Cannabis, Caff, Choc, Crack, Coke, Ecstasy, Ketamine, Heroin, LSD, Legalh, Meth, Nicotine Mushrooms, and VSA. However, the results for Heroin were between 60%-88%.

In 2024 [15], a paper used the Australian (ATOS) dataset to build a predictive model of clinical risk for heroin users, which included heroin usage, overdose recovery, and mortality over various time periods. The study took into account a variety of factors, including study duration, sexual trauma, prison experience, mental handicap, previous criminal convictions, and benzodiazepine use. The researchers used an ensemble learning strategy that integrated multiple machine learning algorithms (Randam Forest, Sport Vector Machine, and Elastic Net). The analytical and statistical results indicated good percentages ranging from 0.73% to 0.91%.

In 2024 [16], researchers used knowledge graphs, natural language processing (NLP), and artificial intelligence (AI) to predict drug interactions. When it comes to drug interaction prediction, deep neural networks outperformed conventional techniques. A detection accuracy of over 85% was achieved for complicated drug interactions. According to the study, ensemble models enable the prediction of adverse drug events with greater precision, which results in safer drug administration procedures.

In 2025 [17], the study employed traditional machine learning approaches as well as LSTM, BiLSTM, and Recursive LSTM algorithms for deep learning to predict drug consumption. The study used simple data from the Finnish National Drug Survey to describe the elements that influence drug use. The researchers combined numerous preprocessors, including Smote, in the oversampling and downsampling to create SmoteTomek, a hybrid approach that surpassed the previous two. The results showed that traditional methods of machine learning outperformed LSTM algorithms due to the simplicity of the data and lack of complexity, with accuracy rates ranging from 93% to 99%.

3. BACKGROUND OF MODELS

A. Light Gradient Boosting Machine (LGBM)

GBDT is a framework that relied on gradient-boosting and decision trees. It is one of the most widely used machine learning methods in several tasks in the field of artificial intelligence, such as prediction [18], classification [19], and learning rank [20]. However, due to the increase in data and the complexity of the features contained within it, GBDT needs to scan all instances of the data. Thus, it became somewhat unsatisfactory, so LGBM appeared, which accelerated the work of GBDT more than 20 times by relying on two technologies. The first is GOSS, which takes the most significant gradients in estimating data acquisition. To decrease the number of features. The second technique EFB is used to bundle mutually exclusive features [21].

B. Extremely Randomized Trees (Extra Trees)

Extra Trees is an ensemble learning technique that incorporates randomization into the tree-growing process [22]. Many decision trees are employed, and samples from each tree are taken at random to achieve originality in data set selection. Furthermore, the characteristics are picked at random, which is why it has that name [23]. The method selects a split value at random rather than computing a locally optimal value for splitting the data using Gini or entropy. As a result, the trees are dissimilar and diversified [24].

C. Cat Boost Algorithm

Cat Boost is a machine-learning ensemble approach that corresponds to the GBDT (gradient boosted decision tree) family. Since its debut in late 2018, researchers have successfully used Cat Boost for machine learning studies including Big Data. It is based on a decision tree algorithm and is characterized by containing implicit processing, which is converting categorical data into numerical data. It does not require pre-processing and is therefore fast to implement. On similar-sized ensembles, the Cat Boost library provides a GPU execution of the learning algorithm and a CPU representation of the scoring method that are significantly faster than existing gradient-boosting libraries [25].

D. Histogram based outlier scores (HBOS)

HBOS is a non-parametric statistical approach that uses feature-specific densities from univariate histograms. It allows calculating categorical and numerical unlabelled data with high performance and minimal execution time [26]. The algorithm is used to explain the distributions of the dataset's features in the form of histograms, which employ bins (densities) to express the frequency and probability of each feature. Initially, the HBOS algorithm used the information from every feature independently. Following that, it was improved to combine all feature histograms to calculate the amount of the algorithm's anomaly score [26]. The steps of algorithms are:

1) To measure every feature, create a histogram and divide the result by the highest number.
2) Normalize the features to lie between 0 and 1, such that a maximum of 1 may be reached by the data.
3) Using the heights of the bins in the histogram, determine the HBOS of each feature in the dataset.
4) Algorithm HBOS used Equation 1 to provide the anomaly score for assessing an instance(q)of a

dataset (x^d), where(d) is the number of features [27].

$$HBOS(q) = \sum_{i=1}^{d} log_2 \left[\frac{1}{histogram(q_i)} \right] \quad (1)$$

5) Data that exceeds the threshold is considered abnormal and vice versa.

4. RESEARCH METHODOLOGY AND APPROACH

A. Background of the Research Study

The PyCharm platform was used as a framework for implementing the research results. During the programming stage, Python libraries such as Scikit-Learn, known for their ML capabilities, and Niapy were used to apply the Grey Wolf Optimizer algorithm from the Swarm Intelligence (SI) technique. The dataset was analysed using four different ML techniques: LightGBM (LGBM), Cat Boost, Extra Trees models and hybrid (EXT- HBOS). Figure 1 shows Methodology Framework.

B. Dataset Description

The dataset contains 1,885 participants, and for each participant there are 12 quantitative features(inputs). The Table 1 contains demographic information about the participants (ID, Age, Gender, Education, Country Ethnicity). It also contains the five-character traits, in addition to two features: impulsivity and sensation-seeking [28]. Eighteen drugs, both legal and illegal, were used as a means of reflecting the outcomes that were asked of the participants, which are also listed in the Table I.

TABLE I. Dataset Description [28]

Feature	Description
ID	Identification
Age	Age range of participant
Gender	Male or Female
Education	Level of education
Country	Country of origin
Ethnicity	Ethnicity/Race of participant
Nscore, Escore, Oscore,Ascore, Cscore	NEO Five-Elements Inventory Neuroticism score, Extraversion score, Openness to experience score,Agreeableness scoreand Conscientiousness score
Impulsive	Quantified BIS-11 impulsiveness score
SS	Quantified Impulsive Sensation Seeking score
Drug	Various drugs like (alcohol,amphetamines, benzodiazepines, amyl nitrite, cannabis, cocaine, chocolate, caffeine,crack, heroin, ecstasy, ketamine, LSD, legal highs, methadone, nicotine,magic mushrooms, and Volatile Substance Abuse (VSA))

C. Correlation Analysis (CA)

A table that shows the correlation coefficients of numerous attributes is called a correlation matrix [29]. This matrix illustrates how the dataset's features relate to the result and to one another in Figure 2. The graph shows a strong correlation between impulsive features and SS (0.62), which indicates that impulsive people have a high amount of impulsive feeling, and the relation between Nscore and Escore shows a negative correlation (-0.43), which indicates that people with a high degree of neuroticism are less extroverted, etc. As for the correlation of these features with heroin addiction, Ascore and Escore have the highest correlation, meaning that the degree of Agreeableness and Conscientiousness have the highest effect on heroin addiction, followed by Escore which represents Extraversion score, and in contrast there is the lowest correlation of heroin addiction with SS and Impulsive.

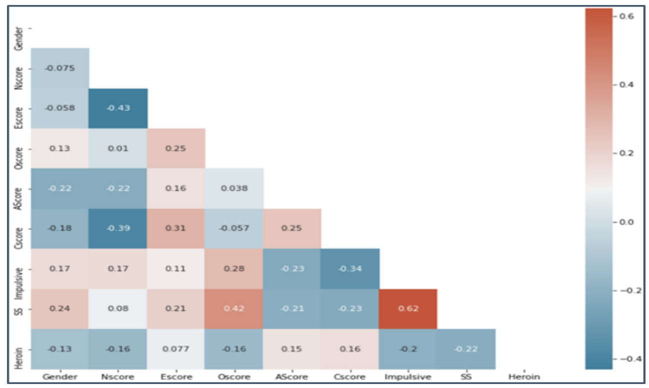

Figure 2. Correlation Matrix of Dataset

D. Data Preparation

Data preparation stage included a number of critical steps: data transformation, normalization, splitting, and resample training data.

- Data Transformation (discretization)
  The dataset has only four categorical fields: age, education, country, and ethnicity; the rest are numerical. Because most ML algorithms expect all data to be numeric, these fields were converted to numeric using one-hot encoding technique.

- Data Normalization
  Without altering the underlying distribution, the Min-Max Scaler is used to standardize the data to a specified range, frequently precisely between 0 and 1. When scaling values to a particular range, it guarantees that data's original shape is maintained. Equation 2 provides the feature's normalization on a scale from 0 to 1 [30]:

$$f_{scaled} = (f - f_{min})/(f_{max} - f_{min}) \quad (2)$$

  Where f_max stands for the feature's maximum value and f_min stands for its minimum value. The Min-Max Scaler function was employed to accomplish this.

- Data Splitting
  First, the data was divided into two groups, 80% of the data going toward training and the remaining 20% going toward the testing.

- Resampling Training Data

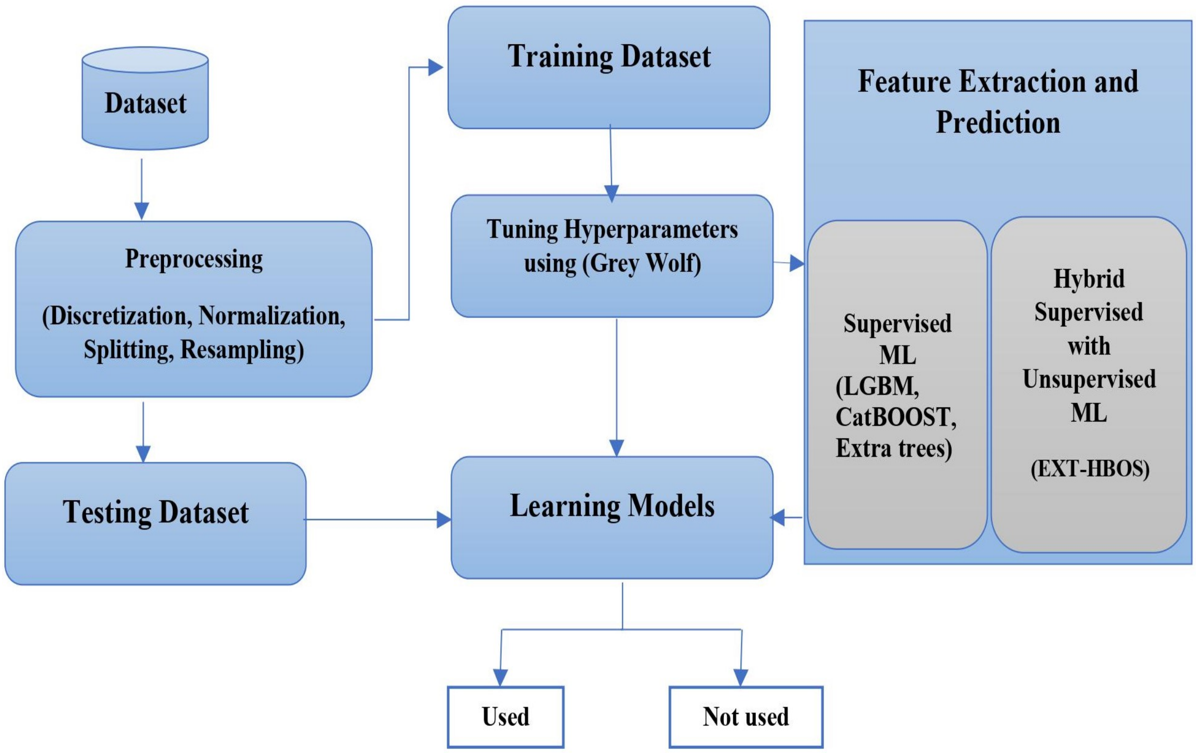

Figure 1. Methodology Framework

The augmented impact may cause classifiers to perform poorly on the minority class but well on the majority. In order to produce a more equal distribution of class instances from unbalanced datasets, resampling is commonly employed.
Techniques for resampling include random oversampling or undersampling. While examples from the minority class are repeated during random oversampling to balance the datasets, samples from the majority group are eliminated during undersampling to balance the set [31].
In order to obtain accuracy and high performance in the model due to the unbalanced data, we adjusted it to balance by undersampling or oversampling [32],[33]. In the dataset utilized for this study, there were 280 heroin users and 1604 non-users. Because of the data imbalance, this would bias the prediction results.
The classifiers will perform poorly on the minority class but well on the majority class. Oversampling was used to balance the data and produce 1604 heroin users and non-users.

- Tuning Hyperparameters
  It can be difficult to choose the best hyperparameters for classification algorithms because it has a significant impact on how effectively a prediction model works [34]. The gray wolf optimizer was used to do this process.

A recent pack intelligence optimization technique (GWO) is widely applied in numerous important fields.
It primarily mimics the hunting strategy and hierarchical structure of the grey wolf race pack in order to attain optimization through these behaviours.
The Grey Wolf Optimization (GWO) was proposed by Seyedali Mirjalili et al. in 2014. The GWO simulates the unique hunting and prey-seeking characteristics of the grey wolf [35]. The group of canines that are still alive includes grey wolves.
In the group, each wolf has a distinct role, and wolves work together to accomplish objectives. The grey wolf population was divided according to four social hierarchy stages by the GWO Figure 3.
The wolf, who holds the top rank, makes decisions about actions like hunting. The greatest choice for wolf is wolf, who holds the second rank, is subordinate to wolf, and together they assist make judgments. Wolf is the third position, below wolf and wolf, and is in charge of duties including scouting and hunting. The lowest position, wolf (the fourth), is in charge of looking after the wolf pack.
Grey wolves hunt by following, chasing, and attacking their prey [36]. The pseudocode of GWO illustrates in Figure 4 [37],[38].

Figure 3. The Grey Wolf Social Hierarchy's distribution and their individual tasks [39]

Algorithm 1: GWO

Initialize: x_i denoted the population of grey wolf that selected randomly, $i = 1 \dots n$
Initialize: parameter a, A and C
Initialize: the iteration number $t = 1$, max denoted max number of iterations
Calculate: fitness of each gray wolf
Select: x_α = best gray wolf
x_β = second best gray wolf
x_δ = third best gray wolf
While t < max do
For each wolf do
Initialize randomly r_1 and r_2
Use $x(t+1) = \frac{x_\alpha + x_\beta + x_\delta}{3}$ to update the position of current gray wolf
End for
Update: a, A and C
Calculate: fitness of all gray wolf
Update: x_α, x_β and x_δ
$t = t + 1$
End while
Return x_α

Figure 4. Algorithim of GWO

The GWO will be used to adjust the hyperparameters of LGBM Classifier, Cat Boost Classifier, Extra Trees Classifier and hybrid (EXT- HBOS). The best Hyperparameters are shown in the Table II.

TABLE II. Best Hyperparameters of the Models

Models	Hyperparameters
LGBM	boosting type='gbdt', numleaves=36, learning rate=0.19, n estimators=149
Cat Boost	iterations=91, learning rate=0.1387, depth=11
Extra Trees	n estimators=73, max depth=29, min samples split=2, random state=6
hybrid(EXT-HBOS)	n estimators=73, max depth=29, min samples split=2, random state=6

5. PREDICTION USING SUPERVISED AND UNSUPERVISED MACHINE LEARNING (HYBRID EXT- HBOS)

At this stage, the algorithms (LGBM, Cat Boost and Extra Trees and hybrid EXT- HBOS) were applied to predict heroin consumption. The hybrid (EXT- HBOS) consists of unsupervised machine learning (HBOS), which contains unlabeled features and supervised algorithm (Extra tree). The aim is to add more features that were discovered by (HBOS), called anomaly Score to the original data. Then this fusion data was entered into Extra tree algorithm to predicate heroin consumption.The following steps explain how the hybrid model (EXT-HBOS) works:

- Inputs: Dataset after preprocessing and resampling steps(original features).
- Step1: Applying HBOS algorithm to produce anomaly Score as new features.
- Step2: Fusion both original features and HBOS feature to obtain new features.
- Step3: Tuning Extra Trees on new features by applying GWO algorithm.
- Outputs: Make prediction using Extra Trees.Figure 5 display fusion features.

	Original Features												HBOS Feature
	Gender	Education	Country	Ethnicity	Nscore	Escore	Oscore	AScore	Cscore	Impulsive	SS	Heroin Consumption	HDOS Score
0	1	Doctorate degree	UK	White	-0.67825	1.93886	1.43533	0.76096	-0.14277	-0.71126	-0.21575	1	-150.863207
1	1	Professional certificate/ diploma	UK	White	-0.46725	0.80523	-0.84732	-1.62090	-1.01450	-1.37983	0.40148	1	-152.660407
2	0	Masters degree	UK	White	-0.14882	-0.80615	-0.01928	0.59042	0.58489	-1.37983	-1.18084	1	-155.729975
3	0	Doctorate degree	UK	White	0.73545	-1.63340	-0.45174	-0.30172	1.30612	-0.21712	-0.21575	1	-151.419626
4	0	Left school at 18 years	Canada	White	-0.67825	-0.30033	-1.55521	2.03972	1.63088	-1.37983	-1.54858	1	-141.968578

Figure 5. Fusion Features

The ExtraTrees and HBOS models were chosen for their ability to handle high-dimensional data efficiently. HBOS is also computationally efficient because it generates graphs for each feature without labeling them, making it easier to discover outliers or extreme values. Furthermore, ExtraTrees uses random tree partitioning or subtrees to reduce dataset overfitting. As a result, both models have been combined because they complement one another. The first model identifies multi-feature interactions, whereas the second model detects deviations in the distribution of individual features, which is not necessarily valid in real-time datasets. We found that our hybrid model(EXT-HBOS) included additional features that contributed to increasing prediction and interpretation accuracy.
The hybrid model's implementation in clinical settings decreases the impacts on drug users since it predicts patients at risk of drug addiction early on and demonstrates that neurotic individuals are more vulnerable to this risk due to emotional instability, sadness, anxiety, and stress.

6. ANALYSIS RESULTS AND COMPARISON WITH ANOTHER WORK

After constructing the models in practical implementation, the impacts of each model must be evaluated. Evaluation criteria are primarily concerned with the model's accuracy. However, it is better to use other performance measures besides statistical measures with data to show how well the models work. This section will present and discuss the outcomes of the preceding models.

A. Results and Analysis using Performance Test

In this work, the accuracy, precision, recall, F1-score and Cohen's Kappa metrics were used to evaluate the performance of AI models and predict heroin consumption, which are defined from Equations 3,4,5,6 [40],[41]:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

$$Recall = \frac{TP}{TP + FN} \tag{5}$$

$$F1score = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{6}$$

where the true positive, true negative, false positive, and false negative are denoted, respectively, by TP, TN, FP, and FN. Furthermore, Cohen's Kappa assesses the classifier's performance against its chance-only performance. Stated differently, a high variance between accuracy and null error rate indicates a high Kappa value for a given model. Furthermore, the kappa value reveals the degree of agreement between two raters. Table III displays the parameters for calculating Cohen's kappa score [42].

TABLE III. Calculating Cohen's kappa Score Parameters [41]

Parameters	Description
P_0	(TP+TN) / (TP+TN+FP+FN)
$P_{positive}$	(TP+FP) (TP+FN) / (TP+TN+FP+FN)^2
$P_{negative}$	(FN+TN) (FP+TN) / (TP+TN+FP+FN)^2
P_e	$P_{positive} + P_{negative}$
Cohen's kappa	$(P_0 - P_e)/(1 - P_e)$, P_0 refers to overall agreement, while P_e refers to chance agreement

The measure's values range from zero to one. If the kappa value is 0, no agreement exists between the classes; if it is 1, perfect agreement exist [43].

In this study, the Table IV shows that hybrid (EXT- HBOS) had the best evaluation metrics with an accuracy of 90% and Cohen's kappa of 74%, precision of 91%, recall of 98% and f1-score of 94% for the test dataset, because of merging the anomalous features with the original ones and getting multi-model learning and high-level prediction.In addition, Extra Trees classifier had the better evaluation metrics with an accuracy of 88% and Cohen's kappa of 62%, precision of 88%, recall of 98% and f1-score of 93% for the test dataset because the random choice of a splitting value for a feature is the most essential and distinguishing in Extra Trees.

TABLE IV. Performance Comparison Between Three Models

Models	Accuracy	Cohen's kappa	Precision	Recall	F1-score
LGBM	76%	51%	92%	79%	85%
Cat Boost	83%	60%	88%	93%	91%
Extra Trees	88%	62%	88%	98%	93%
EXT- HBOS	90%	74%	91%	98%	94%

The Figure 6, shows the confusion matrix of three state of art machine learning as well as hybrid (EXT- HBOS) model which achieved best results in accuracy and Cohen's kappa score with 90% and 74% respectively. These results show the ratio of correctly classified instances to the total number of cases. Therefore, hybrid (EXT- HBOS) model performed better at explaining the relationship between the dataset's properties and outcome parameter.

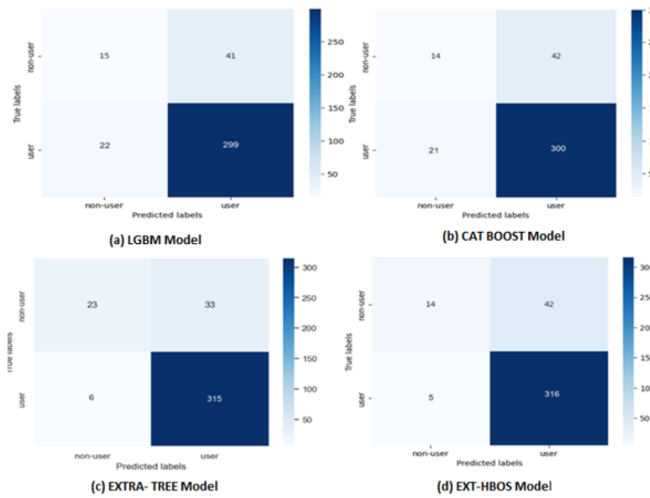

Figure 6. Confusion Matrix of Models

Furthermore, the results of the hybrid model (EXT-HBOS) were compared with another similar study that used the same dataset [14]. Table V shows that the hybrid model outperformed the previous study in terms of accuracy.

TABLE V. Comparison with Previous Study

Models	Dataset	Methods	Accuracy
Previous Study [14]	UCI repository [28]	Decision Tree, Random Forest, Gaussian Naive Bais, K Neighbors Classifier, Logistic Regression	60%-88%
EXT- HBOS	UCI repository [28]	hybridization between Supervised and Unsupervised ML	90%

The method of predicting heroin consumption is influenced by a variety of factors. The relevance level for each attribute in the prediction procedure is shown in Figure 7. It is noticed that Nscore is by the most important factor, followed by Ascore, Escore and Cscore that same influence, and then Ethnicity_Black and Ethnicity_Mixed-Black/Asian, which are the least important. This is because there is a connection between neuroticism and drug usage, as those with high neuroticism exhibit worry, tension, and emotional instability. As a result, people with this tendency may turn to drug abuse as a kind of self-treatment. Higher neuroticism scores [44] are significantly connected with an increased likelihood of lifetime drug use. The adjusted odds ratio (AOR) for drug abuse was 1.05 (95% CI: 1.01, 1.10; p = 0.013), indicating that higher levels of neuroticism enhanced the chance of drug misuse.

Also,the association between childhood adversity and the chance of drug addiction is stronger in those with high neuroticism, implying that neuroticism triggers the consequences of childhood adversity,leading to drug abuse behaviors [45].

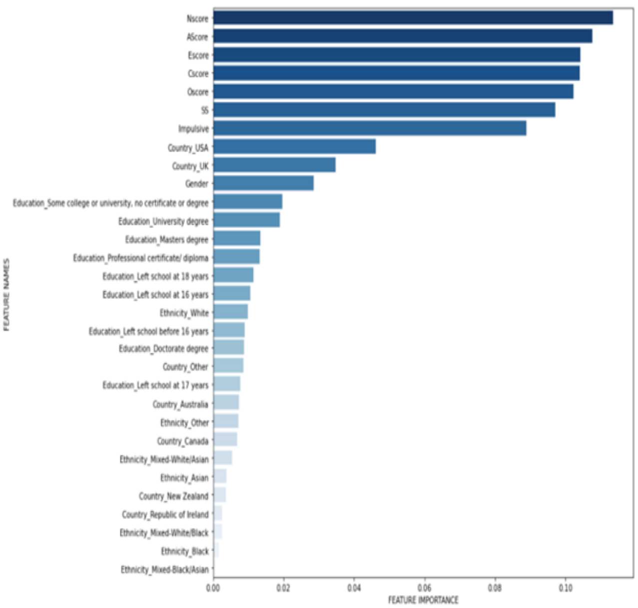

Figure 7. Matrix of Important Features

B. Results and Analysis Using Statistical Test

The Friedman test is used to measure how much the hybrid (EXT- HBOS) outperforms the other three models.

- Friedman Test
  According to the alternative hypothesis, the prediction errors of the proposed models are not the same and differ,must be accepted when the Friedman statistical test is in the region of rejection, and vice versa. In the first dimension of Table VI, a significant differences were observed in the average ranking of hybrid EXT-HBOS and LGBM because F=128(Friedman Test) is greater than 5.99 (state decision rule)[46],[47], also p-value=0.001 which is less than significance level (0.05), therefore the null hypothesis was rejected and the alternative hypothesis was accepted.
  The second, third and fourth dimensions followed in achieving the same results. This demonstrates the superiority of the hybrid (EXT- HBOS) model.
  As for the fifth dimension, there was no statistically significant difference between the models, so the null hypothesis was chosen over the alternative hypothesis[48].
  To achieve more reliability, the Wilcoxon test was applied in the next subsection.
- Wilcoxon Test
  The first dimension of Wilcoxon test in Table VII demonstrates that there is a statistically significant difference in the mean value of the hybrid (EXT- HBOS) and LGBM and the p-value=0.005 which is less than significance level (0.05).
  Also, according to the results, the negative ranking was based on due to the value of z=-2.803, which means that mean rank is 0.00, which is less than critical value 8 [49],[50]. As a result, accept the alternative hypothesis while rejecting the null hypothesis. In the dimensions second and third, there is a statistical difference in the mean value for the hybrid (EXT- HBOS) with CatBoost and Extra Trees respectively. Also, there is a statistical difference in the mean value in the four dimensions between Extra Trees and LGBM. The p-value was 0.005 and negative ranking was based on, so reject the null hypothesis and accept the alternative hypothesis.

TABLE VI. Friedman Test of Three Models

First Dimension	Friedman Test
EXT- HBOS – LGBM	H1: $e1 \neq e3$, e1=e5 , $e5 \neq e3$, F= 128 > 5.99, Chi_square = 20.000, P-Value =0.001, Reject null hypothesis and accept alternative hypothesis
Second Dimension	Friedman Test
EXT- HBOS – CatBoost	H1: e5=e1, $e5 \neq e3$, $e1 \neq e3$, Reject null hypothesis and accept alternative hypothesis
Third Dimension	Friedman Test
EXT- HBOS –Extra Trees	H1: $e3 \neq e1$,$e3 \neq e5$,e1=e5 Reject null hypothesis and accept alternative hypothesis
Fourth Dimension	Friedman Test
Extra Trees – LGBM	H1: $e4 \neq e3$, $e4 \neq e6$, e3=e6 Reject null hypothesis and accept alternative hypothesis
Fifth Dimension	Friedman Test
Extra Trees – CatBoost	H0: $e6 \neq e4$,e6=e3 Accept null hypothesis and reject alternative hypothesis

Based on the preceding statistical measurements, we conclude that the four models (hybrid EXT- HBOS, Extra Trees, LGBM, and Cat Boost) are normally distributed, with statistically significant differences, and that the hybrid EXT-HBOS model which combined supervised algorithm (Extra Trees) with unsupervised (HBOS) algorithm achieved the highest accuracy and performance.

7. CONCLUSIONS

The primary goal of this paper was to evaluate and compare how well the hybrid (EXT- HBOS), LGBM, Cat Boost, and Extra Trees Classifiers performed in predicting heroin consumption that has a substantial influence on those

TABLE VII. Wilcoxon Test of Three Models

Dimensions	Algorithm	Mean	P-Value	Null Hypothesis
First	EXT- HBOS LGBM	50.9000 41.2300	0.005	Mean rank = 0.00 <8, z=2.803 (negative rank) Reject Null hypothesis
Second	EXT- HBOS Cat-Boost	50.9000 46.5000	0.005	
Third	EXT- HBOS Extra Trees	50.9000 49.5000	0.005	
Fourth	Extra Trees LGBM	49.5000 41.2300	0.005	

who use it. Accuracy, precision, recall, and F1-score metrics are used to evaluate models. The outcomes show that hybrid EXT- HBOS Classifier outperforms Extra tree, Cat Boost and LGBM in terms of performance. The superiority of hybrid EXT- HBOS Classifier due to fused anomaly score features with original data to achieve multi feature extraction and get high level prediction of heroin consumption. This study helps to increase the accuracy and reliability of forecasting people's heroin intake by addressing the difficulties faced by machine learning in this area. Heroin eliminates the natural painkillers produced by the brain. These painkillers are called endorphins (and they also cause happiness), so the body's tolerance for any pain, no matter how slight, decreases. After treating addiction, the body begins to produce endorphins again, but the damage that occurs in the brain may take years to be treated, which has a number of negative repercussions, including liver and heart disease, high blood pressure, and skin issues brought on by repetitive injections, such as boils and bruises, among many others.

References

[1] B. Chhetri, L. M. Goyal, and M. Mittal, "How machine learning is used to study addiction in digital healthcare: A systematic review," *International Journal of Information Management Data Insights*, vol. 3, no. 2, p. 100175, 2023.

[2] N. C. for Drug Abuse Statistics, "Drug abuse statistics," Accessed: Aug. 9, 2024. [Online]. Available: https://drugabusestatistics.org.

[3] D. M. Walker and E. J. Nestler, "Neuroepigenetics and addiction," *Handbook of clinical neurology*, vol. 148, pp. 747–765, 2018.

[4] A. Ambekar, A. Agrawal, R. Rao, A. K. Mishra, S. K. Khandelwal, R. K. Chadda *et al.*, "Magnitude of substance use in india," *New Delhi: Ministry of Social Justice and Empowerment, Government of India*, vol. 23, 2019.

[5] S. Joseph, "Drug demand reduction programme in india: Recommendations for the future," *Shanlax International Journal of Arts, Science and Humanities*, vol. 7, no. 2, 2019.

[6] R. E. Baker, J.-M. Pena, J. Jayamohan, and A. Jérusalem, "Mechanistic models versus machine learning, a fight worth fighting for the biological community?" *Biology letters*, vol. 14, no. 5, p. 20170660, 2018.

[7] W.-Y. Ahn and J. Vassileva, "Machine-learning identifies substance-specific behavioral markers for opiate and stimulant dependence," *Drug and alcohol dependence*, vol. 161, pp. 247–257, 2016.

[8] L. Acion, D. Kelmansky, M. van der Laan, E. Sahker, D. Jones, and S. Arndt, "Use of a machine learning framework to predict substance use disorder treatment success," *PloS one*, vol. 12, no. 4, p. e0175383, 2017.

[9] Z. Qiao, T. Chai, Q. Zhang, X. Zhou, and Z. Chu, "Predicting potential drug abusers using machine learning techniques," in *2019 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS)*. IEEE, 2019, pp. 283–286.

[10] Z. Segal, K. Radinsky, G. Elad, G. Marom, M. Beladev, M. Lewis, B. Ehrenberg, P. Gillis, L. Korn, and G. Koren, "Development of a machine learning algorithm for early detection of opioid use disorder," *Pharmacology Research & Perspectives*, vol. 8, no. 6, p. e00669, 2020.

[11] A. S. Wadekar, "Understanding opioid use disorder (oud) using tree-based classifiers," *Drug and alcohol dependence*, vol. 208, p. 107839, 2020.

[12] M. Beattie and C. Nicholson, "Feature extraction for heroin-use classification using imbalanced random forest methods," *Substance Use & Misuse*, vol. 56, no. 1, pp. 123–130, 2021.

[13] K. Jain, A. Singh, P. Singh, and S. Yadav, "An improved supervised classification algorithm in healthcare diagnostics for predicting opioid habit disorder," *International Journal of Reliable and Quality E-Healthcare (IJRQEH)*, vol. 11, no. 1, pp. 1–16, 2022.

[14] M. Almahmood, H. Najadat, D. Alzu'bi, L. Abualigah, R. A. Zitar, S. Abualigah, and F. Al-Saqqar, "Predictive model of psychoactive drugs consumption using classification machine learning algorithms," *Applied and Computational Engineering*, vol. 8, no. 1, pp. 738–743, 2023.

[15] C. Marel, M. H. Afzali, M. Sunderland, M. Teesson, and K. L. Mills, "Predicting risk of heroin overdose, remission, use, and mortality using ensemble learning methods in a cohort of people with heroin dependence," *International Journal of Mental Health and Addiction*, pp. 1–19, 2024.

[16] P. Sonaji, L. Subramanian, and M. Rajesh, "Artificial intelligence-driven drug interaction prediction," *World Journal of Biology Pharmacy and Health Sciences*, vol. 17, no. 2, pp. 297–305, 2024.

[17] A. Unlu and A. Subasi, "Substance use prediction using artificial intelligence techniques," *Journal of Computational Social Science*, vol. 8, no. 1, p. 21, 2025.

[18] M. Richardson, E. Dominowska, and R. Ragno, "Predicting clicks: estimating the click-through rate for new ads," in *Proceedings of the 16th international conference on World Wide Web*, 2007, pp. 521–530.

[19] P. Li, "Robust logitboost and adaptive base class (abc) logitboost," *arXiv preprint arXiv:1203.3491*, 2012.

[20] C. J. Burges, "From ranknet to lambdarank to lambdamart: An overview," *Learning*, vol. 11, no. 23-581, p. 81, 2010.

[21] M. R. Machado, S. Karray, and I. T. De Sousa, "Lightgbm: An effective decision tree gradient boosting method to predict customer loyalty in the finance industry," in *2019 14th International Conference on Computer Science & Education (ICCSE)*. IEEE, 2019, pp. 1111–1116.

[22] P. Zuccolotto, M. Sandri, and M. Manisera, "Spatial performance analysis in basketball with cart, random forest and extremely randomized trees," *Annals of Operations Research*, vol. 325, no. 1, pp. 495–519, 2023.

[23] P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees," *Machine learning*, vol. 63, pp. 3–42, 2006.

[24] J. T. Hancock and T. M. Khoshgoftaar, "Catboost for big data: an interdisciplinary review," *Journal of big data*, vol. 7, no. 1, p. 94, 2020.

[25] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, "Catboost: unbiased boosting with categorical features," *Advances in neural information processing systems*, vol. 31, 2018.

[26] U. Binzat and E. Yıldıztepe, "The adjusted histogram-based outlier score-ahbos," *Mugla Journal of Science and Technology*, vol. 9, no. 1, pp. 92–100, 2023.

[27] I. Aguilera-Martos, M. García-Barzana, D. García-Gil, J. Carrasco, D. López, J. Luengo, and F. Herrera, "Multi-step histogram based outlier scores for unsupervised anomaly detection: Arcelormittal engineering dataset case of study," *Neurocomputing*, vol. 544, p. 126228, 2023.

[28] KHADIJA, "Kaggle dataset," Accessed: May. 5, 2024. [Online]. Available: https://www.kaggle.com/code/obeykhadija/drug-consumption-prediction/notebookGetting-Started.

[29] C. J. Burges, "From ranknet to lambdarank to lambdamart: An overview," *Learning*, vol. 11, no. 23-581, p. 81, 2010.

[30] A. Gökhan, C. O. Güzeller, and M. T. Eser, "The effect of the normalization method used in different sample sizes on the success of artificial neural network model," *International journal of assessment tools in education*, vol. 6, no. 2, pp. 170–192, 2019.

[31] M. S. Kraiem, F. Sánchez-Hernández, and M. N. Moreno-García, "Selecting the suitable resampling strategy for imbalanced data classification regarding dataset properties. an approach based on association models," *Applied sciences*, vol. 11, no. 18, p. 8546, 2021.

[32] R. Mohammed, J. Rawashdeh, and M. Abdullah, "Machine learning with oversampling and undersampling techniques: overview study and experimental results," in *2020 11th international conference on information and communication systems (ICICS)*. IEEE, 2020, pp. 243–248.

[33] H. A. Aldabagh and G. A. Altalib, "A survey for using ai techniques for predicting covid-19," in *2022 International Conference on Computer Science and Software Engineering (CSASE)*. IEEE, 2022, pp. 1–63.

[34] R. K. Tan and Ş. Bora, "Parameter tuning in modeling and simulations by using swarm intelligence optimization algorithms," in *2017 9th International Conference on Computational Intelligence and Communication Networks (CICN)*. IEEE, 2017, pp. 148–152.

[35] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Advances in engineering software*, vol. 69, pp. 46–61, 2014.

[36] Y. Hou, H. Gao, Z. Wang, and C. Du, "Improved grey wolf optimization algorithm and application," *Sensors*, vol. 22, no. 10, p. 3810, 2022.

[37] K. Li, S. Li, Z. Huang, M. Zhang, and Z. Xu, "Grey wolf optimization algorithm based on cauchy-gaussian mutation and improved search strategy," *Scientific Reports*, vol. 12, no. 1, p. 18961, 2022.

[38] M. H. Hashem, H. S. Abdullah, and K. I. Ghathwan, "Grey wolf optimization algorithm: A survey," *Iraqi Journal of Science*, pp. 5964–5984, 2023.

[39] E. Dada, S. Joseph, D. Oyewola, A. A. Fadele, H. Chiroma *et al.*, "Application of grey wolf optimization algorithm: recent trends, issues, and possible horizons," *Gazi University Journal of Science*, vol. 35, no. 2, pp. 485–504, 2022.

[40] H. A. Aldabagh and G. A. Altalib, "Predicting the effect of covid-19 on physical activity of survivors using gso and hybrid intelligent model," in *2022 2nd International Conference on Advances in Engineering Science and Technology (AEST)*. IEEE, 2022, pp. 739–745.

[41] F. Uysal, F. Hardalaç, O. Peker, T. Tolunay, and N. Tokgöz, "Classification of shoulder x-ray images with deep learning ensemble models," *Applied Sciences*, vol. 11, no. 6, p. 2723, 2021.

[42] R. T. Ibrahim and F. M. Ramo, "Classification of personality traits by using pretrained deep learning models," in *2021 7th International Conference on Contemporary Information Technology and Mathematics (ICCITM)*. IEEE, 2021, pp. 42–47.

[43] J. Landis, "The measurement of observer agreement for categorical data," *Biometrics*, 1977.

[44] L. G. Ongeri, "Prevalence of personality disorders among substance abusers in drug rehabilitation centres in kenya," Ph.D. dissertation, 2011.

[45] J. N. Davies, S. C. Harty, and J. M. Boden, "The roles of extraversion and neuroticism in the relationship between childhood adversity and adolescent substance misuse," *Personality and mental health*, 2024.

[46] R. T. Ibrahim and F. M. Ramo, "Hybrid intelligent technique with deep learning to classify personality traits," *International Journal of Computing and Digital Systems*, vol. 13, no. 1, pp. 231–244, 2023.

[47] H. A. Aldabagh and R. T. Ibrahim, "Hybrid intelligent technique between supervised and unsupervised machine learning to predict water quality," *International Journal of Computing*, vol. 17, no. 1, pp. 1–14, 2025.

[48] A. N. Younis and F. M. Ramo, "Developing viola jones' algorithm for detecting and tracking a human face in video file," *Int J Artif Intell*, vol. 12, no. 4, pp. 1603–1610, 2023.

[49] A.-S. Muna and F. M. Ramo, “Performance evaluation for face mask detection based on mult modification of yolov8 architecture,” *Informatyka, Automatyka, Pomiary w Gospodarce i Ochronie Środowiska*, vol. 14, no. 2, pp. 89–95, 2024.

[50] H. A. Younis and Y. F. Mohammad, “Hybridization of self supervised learning models for enhancing automatic arabic speech recognition,” in *2024 4th International Conference of Science and Information Technology in Smart Administration (ICSINTESA)*. IEEE, 2024, pp. 650–655.