JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

A HYBRID META-LEARNING AND ZERO-SHOT LEARNING FRAMEWORK FOR DISEASE CLASSIFICATION: OVERCOMING DATA SCARCITY IN HEALTHCARE

[1]Pallavi S B, [2]Dr Ajmeera Kiran, [3]Dr Thrimurthulu, [4]Dr Pushpa Rani, [5]Dr K Gagan Kumar
[1]M.Tech Student, [2]Professor and HOD, [3]Professor, [4]Associate Professor, [5]Associate Professor
[1]Computer Science and Engineering,
[1]MLR Institute of Technology, Hyderabad, India

ABSTRACT: In the domain of healthcare, effectively classifying diseases such as hepatitis and heart disease poses significant challenges, particularly when labeled data is scarce. This paper presents a novel approach that integrates meta-learning with zero-shot learning to enhance classification performance on both hepatitis and heart disease datasets. By leveraging meta-learning, a robust model is developed that quickly adapts to new, unseen tasks with minimal training samples, effectively addressing the issue of limited labeled data. Simultaneously, zero-shot learning facilitates the model's capability to identify and classify instances from unknown classes by imparting information from related, known classes. Experiments employ a meta-learning framework using the Prototypical Networks approach to learn embeddings from both datasets. This framework enables the model to generate prototypes for each class based on a few labeled samples, facilitating rapid adaptation to new tasks. Subsequently, a zero-shot learning strategy utilizing semantic embeddings derived from a knowledge graph is incorporated, allowing the model to predict classes not seen during training. The combined methodology demonstrates significant improvements in classification accuracy and generalization across both datasets, showing promise for practical applications in real-world healthcare scenarios where obtaining labeled data is challenging. The results underscore the efficiency of combining meta-learning and zero-shot learning in developing versatile and adaptive models capable of addressing diverse medical classification challenges.

Keywords: Prototypical Networks, Knowledge graph, Few-shot learning, Healthcare AI, Adaptive models, Generalization in classification

1. INTRODUCTION:

Imbalanced datasets are notable in the medical field, as they can lead to biased predictions and misdiagnoses. In many medical applications, such as disease diagnosis and treatment outcome prediction, the number of healthy patients often vastly outweighs the number of patients with specific diseases. This imbalance can cause machine learning algorithms to favor the majority class, resulting in poor sensitivity for minority classes. For instance, if a model is trained predominantly on healthy patients, it may struggle to accurately identify cases of rare diseases, potentially leading to missed diagnoses. This can have serious consequences for patient health, delaying necessary treatments or interventions. Additionally, imbalanced datasets can hinder the development of robust predictive models, as standard evaluation metrics like accuracy may provide a false sense of reliability. Instead, metrics such as precision, recall, and the F1-score should be prioritized to better assess model performance on minority classes. Techniques such as data augmentation, data reduction, or generating synthetic data can help mitigate these issues. Furthermore, clinical decision- making often relies on accurate predictions, making the consequences of imbalanced data particularly critical in healthcare settings. Addressing these challenges is essential for improving patient outcomes and ensuring equitable healthcare delivery. Ultimately, balancing datasets can enhance the accuracy of medical models, leading to more effective diagnostics and personalized treatment plans for all patients

1.1. Handling Missing Values:

It is very important to deal with missing values in data preparation, especially in medical records where they can change the results and cause biased conclusions [9]. There are a number of ways to deal with lost data. In this case, deleting rows or columns with blank data is one way to do it. This works well when there aren't many lost values. Imputation is an additional technique in which the average or most frequent value for a characteristic is used to fill in the missing data. This is done by employing statistical approaches like mean, median, or mode imputation. More sophisticated imputation approaches include k-Nearest Neighbours (k-NN), which predicts missing values using the closest data points, and multiple imputation with predictive modeling. Domain-specific techniques, such clinical knowledge- based imputing or data trends analysis, might be useful in certain situations [10]. Which way to use relies on the type of data, the rate of missing values, and the style of missing data. Figure 1 presents the different approaches for the cleaning the missing values

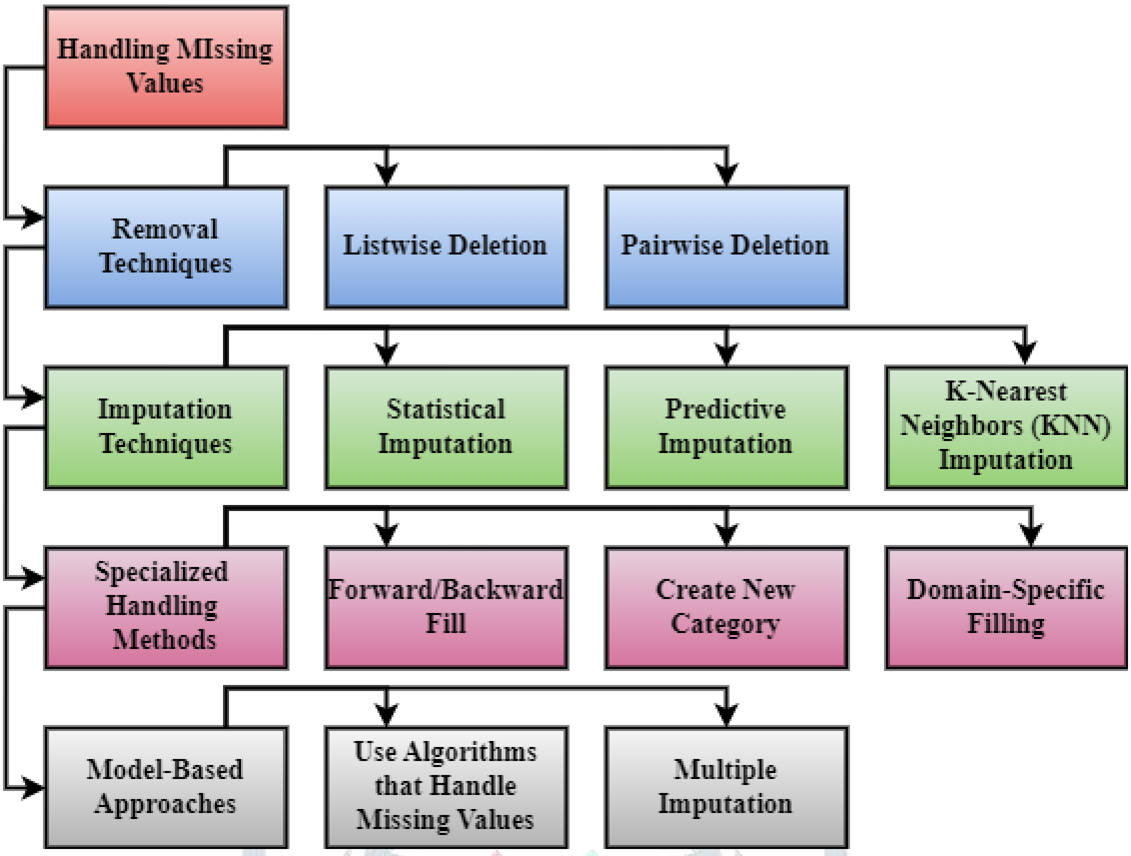


Figure 1: Handling Missing Values

1.2. Impact of Outliers on Medical Datasets: In medical records, Data points classified as outliers deviate significantly from the bulk of the info. Have a major effect on outcomes with data analysis especially machine learning models[11]. When it comes to healthcare, outliers can be mistakes in the data collection, rare diseases, or very bad results for patients. Outliers like these can change statistical measures like the standard deviation and mean, which can lead to results that are biased or not accurate. It is possible for outliers to make predictive models either overfit or fail because they change the decision boundaries. Finding outliers and dealing with them properly is very important in medical study. Outliers can be real extreme cases that give useful information, or they can be noise that needs to be fix or removed. Statistical approaches such as z-scores or IQR, which identify and eliminate extreme values, or reliable approaches such as log transformation and median-based statistics, which lessen the impact of outliers without eliminating them entirely, are some strategies for dealing with outliers[12]. Figure 2 presents the different ways for performing the outlier analysis.

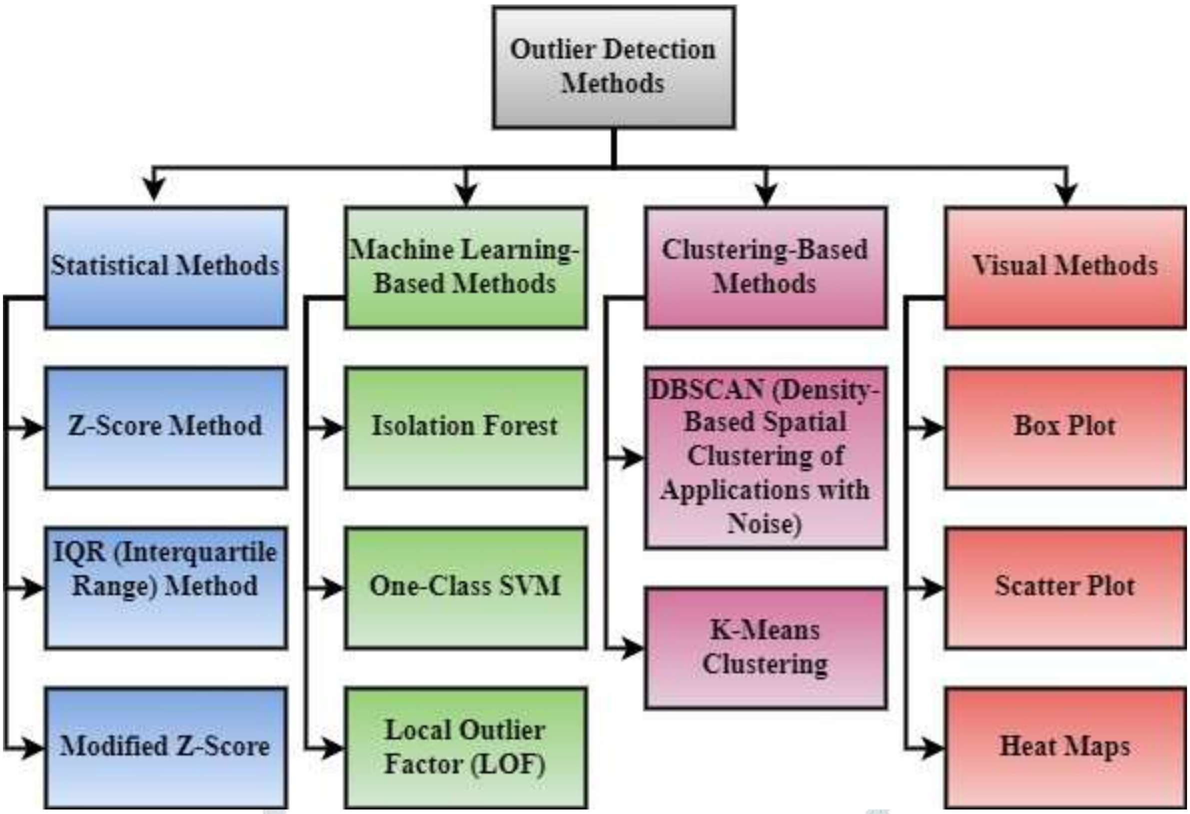


Figure 2: Different Approaches for Outlier Analysis

1.3 Different Outlier Detection Methods: In the Heart Disease Detection, outlier detection was performed using the Interquartile Range (IQR) method. This technique is a robust statistical approach to pinpointing outliers depending on the spread of the data. The middle 50% range (IQR) is a common measure of variability that represents the gap in betwixt the lower quartile and the upper quartile of a dataset. Values that fall below the lower quartile lessen by one and half times the range, or above the upper quartile increased by one and half times the range, are identified as extreme values. These outliers fall outside the middle 50% of the data, indicating they are significantly lower or higher than the central portion of the distribution. The rationale for using the IQR method is its effectiveness in handling skewed distributions and its resilience to extreme values, making it particularly suitable for datasets where the presence of outliers may distort analysis or model performance. Unlike standard deviation-based methods, which can be heavily influenced by outliers, the IQR approach provides a more stable criterion for outlier detection. This is critical in the context of heart disease data analysis, as outliers can arise from various sources, such as measurement errors or atypical physiological responses, which could skew the insights drawn from the data. By removing these outliers, we aim to enhance the accuracy of subsequent analyses, including modeling and predictions, ensuring that the underlying patterns related to heart disease are more clearly represented. The equation for the IQR is shown in eq(1)

$$IQR = Q3 - Q1\text{- (1)}$$

In Hepatitis Detection code, Z-score method was employed for outlier detection. This approach is based on the assumption that the numerical features of the dataset follow a normal distribution. One may determine how distant a data point is from the central point, measured in terms of standard deviations, computing the Z-score for each data point in specific numerical columns. In particular, the Z-score is calculated by dividing the dataset's standard deviation by the gap between the data point and the central point. Outliers are identified and eliminated from the dataset if their Z-score is larger than 3 or less than -3. This means that any data point that lies significantly outside the expected range will be flagged for removal. This method is advantageous because it is straightforward to implement and interpret, allowing for effective handling of extreme values in the dataset. Additionally, it is computationally efficient, making it suitable for large datasets where performance can be a concern. By eliminating outliers, robustness of the model can be improved and ensure that it focuses on the main trends in the data rather than being skewed by anomalous observations. This is crucial for preserving the integrity of the analysis and maintaining the predictive power of the model. Ultimately, this contributes to more accurate predictions and overall model performance when utilizing advanced techniques like meta-learning and K-Nearest Neighbors for classification, enhancing the reliability of results and supporting better decision-making. Here is the z-score mathematical representation in equation (2)

$$z = \frac{X - \mu}{\sigma} \text{ – (2)}$$

Where:

- z is the **z-score**.
- X is the data point.
- μ is the **mean** of the dataset.
- σ is the **standard deviation** of the dataset.

2. LITERATURE SURVEY:

Sangya Ware et al [1] Using ml to identify heart problems shows a methodical way to do things. The first step is to figure out what the problem is and use different traits obtained from the Cleveland dataset to figure out if an individual has heart disease. To guarantee data quality, the dataset must be chosen, then pre-processed and noisy & missing data must be eliminated. Some machine learning methods used for comparison and classification are SVM, RF, DT, KNN, NB, and then LR. Training sets make up 60% of the information, while testing sets make up 40%. Models are then built using these partitions. To rate the models' success, by using various measures like F1 score, accuracy, precision or memory. A GUI is also being made so that people can enter data that will tell if someone has heart disease. The last part of the approach compares the outcomes of the six algorithms. SVM yielded the most accurate results.

Harshit Jindal et al [2] Machine learning based heart disease prediction provides an approach with many steps. Data first comes from the UCI collection, with an eye towards medical factors like age, chest discomfort, and blood pressure. Data preparation to address missing values and normalise it for improved model performance comes next. Testing and training sets separate the data after pre-processing. Three machine learning systems KNN, LR, and RF Classifier are used in this work. On the basis of the input qualities, these algorithms are used to categorise patients for the purpose of determining whether or not they are at risk of developing heart disease. The model is tested using performance criteria like accuracy and precision; KNN has the best accuracy at 88.52%. Emphasising that the model offers major help to healthcare providers by providing a reasonably priced, efficient approach to forecast heart disease, the methodology ends.

Abdul Saboor et al [3] Prediction accuracy of cardiac disease is improved by using many machine learning classifiers. With 303 records including 13 pertinent characteristics, the Cleveland heart attack dataset is the one used for this analysis. Data preparation is the first phase, which entails refining and standardising the dataset by deleting erroneous values and normalising the data for consistency. This step is required before going on to the next step. The GridSearchCV technique is used in hyperparameter optimisation to maximise the ml classifier. Included among the classifiers are RF, XGB, SVM, LR, & others. Ten-fold cross-valuation guides both training and testing of these classifiers. Every classifier's performance is assessed with reference to accuracy, precision, recall, & F-measure. Following hyperparameter adjustment, SVM outperformed all classifiers with the greatest prediction accuracy.

Dr. M. Kavitha et al [4] The prediction of cardiac disease using an integrated ml model. This hybrid model is a combination of two strong algorithms, namely Random Forest and Decision Tree. By using the probabilities that are provided by one model as input for the second model, the suggested method improves the accuracy of the predictions that are made. For the purposes of training and testing the model, the Cleveland heart disease dataset, which includes 303 occurrences and 14 characteristics, is used. In the beginning, the data set is separated into two phases: the training and test parts 7:3 ratio. The Decision Tree method facilitates the categorisation of individuals into two distinct groups: those with heart disease and those without heart disease. Random Forest, on the other hand, is a method that combines many decision trees in order to get more accurate forecasts. Using the probability results from RF as data for further categorisation by the DT model, the hybrid model blends the strengths of both models, hence combining the advantages of both.

Douaa Ibrahim Alsaadi et al [5] by applying a several machine learning techniques to a dataset that has 155 occurrences and 20 characteristics that are associated with hepatitis patients. In order to categorise the data and provide predictions on the course of the illness, seven different machine learning techniques are used. These algorithms are K-NN, SVM, DT, Neural Network, Logistic Regression, Linear Regression, and Naive Bayes. Replacing missing data with the mean and using the K-NN technique to correct outliers are both part of the pre-processing stage. Following the completion of the data cleaning process, each algorithm is used to carry out classification and validation, so guaranteeing that the predictions provide reliable results. The SVM classification method is based on separating hyperplanes, while the Decision Tree algorithm takes a divide-and-conquer strategy. Both of these algorithms utilise distinct methodologies.

Khair Ahammed et al [6] The proposed strategy for predicting liver fibrosis stages in HCV patients makes use of machine learning methods to improve prediction accuracy. The dataset comprises 1385 Egyptian HCV patient records with 29 variables, including age, gender, BMI, and clinical measures. An application of the SMOTE was made in order to remedy the class imbalance. This included the generation of synthetic data and the balancing of instances throughout the various phases of fibrosis. Information Gain Attribute Evaluation, Chi-Square Attribute Evaluation, Relief, and Gain Ratio Attribute Evaluation were among the feature selection techniques used to determine which attributes were most important. After that, a number of different classifiers, such as KNN, RF, SVM, DT, and XGB, were used on the previously collected data. Following the use of SMOTE, KNN emerged as the most effective classifier, obtaining a high level of accuracy. RF & SVM also performed in a satisfactory manner.

Ibraheem I. Ahmed et al [7] the suggested approach focusses on machine learning approaches' diagnosis of hepatitis illness. Medical tests define people as "live" or "dead" in the research, which employs 155 samples from the UCI collection with 19 features. Data preparation which comprises three methods is part of the approach. Both numerical and nominal data abound in the collection; they are turned into real values for use. Using an adaptive wrapper approach where features are repeatedly trimmed depending on their importance the technique also emphasises feature selection. Cross-valuation is included into this approach to guarantee strong model development. Three ml techniques: SVM, DT, and RF. The performance of different methods; RF has the best accuracy. It only takes six features to make the final step possible: the learnt models are added to a mobile-based diagnostics tool. This makes hepatitis detection easier to get and cheaper in real-time situations.

B. Prathyusha et al [8] Hepatic disease is predicted using ML techniques. Beginning with data collecting from health records or previous works. It is necessary to do data preprocessing on the acquired data in order to clean, normalise, and deal with missing values. The data that is gathered could involve patient details, medical history, plus lab test results. Feature selection comes next, in this case the most relevant characteristics supporting the diagnosis are found to raise model efficiency. Using techniques like SVM, RF, and LR to extract patterns from the data is known as the model training phase. Prediction and analysis may begin after the models have been taught. HepatitisNet projects the course or degree of hepatitis in new patients using the trained models. By concentrating on important elements, the method lowers dimensionality and increases prediction accuracy, thereby improving early detection and therapy results for hepatitis.

Table 1: Comparative Analysis

Author	Algorithm	Merits	Demerits	Accuracy
Sangya Ware et al	SVM	Implemented in simple way for accurate performance.	Dataset was small.	89.3%
Harshit Jindal et al	KNN	Improved the diagnosis and cost-reduction.	Prediction depends on dataset.	86.5%
Abdul Saboor et al	SVM	By utilizing hyper parameter tuning the performance was improved.	Few more features can be properly analyzed.	96.7%
Dr. M. Kavitha et al	RF, and DT	By combining the methods performance has improved efficiently.	Multi-class issue was not identified.	88.7%
Douaa Ibrahim Alsaadi et al	ML	This works at any platform.	The dataset was not huge.	100%
Khair Ahammed et al	KNN	By applying several feature selection methods the process has been easy for implementation.	Working of transformations are not accurate.	94.4%
Ibraheem I. Ahmed et al	RF	By utilizing adaptive wrapper method the performance was improved.	While training the model the attributes are loss was huge.	96.1%
B. Prathyusha et al	ML	Process has acquired accurate performances.	Required more validation techniques to analyze the performance.	

3. PROPOSED METHODOLOGY: The proposed model mainly focuses on the imbalanced medical datasets to analyze and increase the efficiency of the model even if it has underfitting attributes. In this process, it initially analyses both categorical and numerical features and removes the outlier's data by performing the hypothesis testing to perform feature engineering then it initializes the meta learning set up process and evaluates it by measuring the cross-entropy loss. The model then applies zero shot inference technique to predict the unseen data accurately. The proposed architecture is shown in figure 3

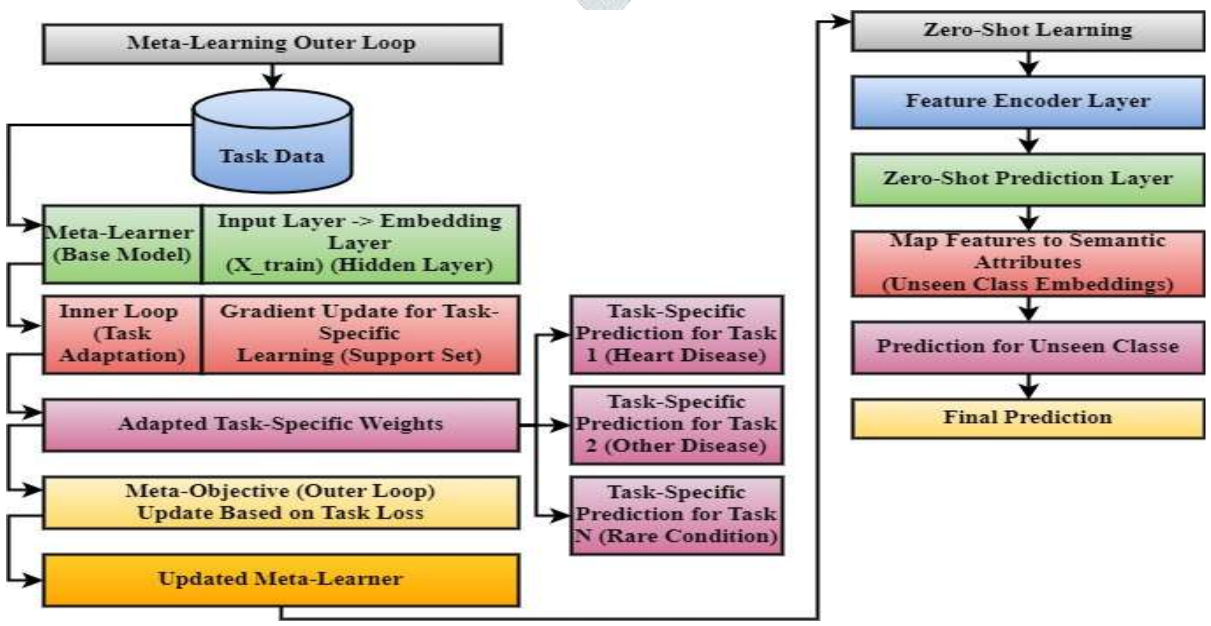


Figure 3: Block Diagram for Proposed Architecture

3.1. Analysis on Categorical Data:Data exploration depends on categorical distribution visualisation, especially when handling discrete variables like gender, ethnicity, or therapy kinds in medical information. In medical statistics, these factors often provide important information on the demographics and illness characteristics of the patient. Clear understanding of each category's frequency and percentage made possible by visualisations as bar charts, pie charts, count graphs, or mosaic plots helps researchers identify data anomalies or biases. For a study on heart disease, for instance, if the dataset shows more males than female patients, this imbalance may influence the results and forecasts of the model. Such visualisations are also useful in machine learning for the identification of over-represented or under-represented classes, therefore guiding judgements on whether to use strategies include synthetic data creation or stratified sampling to solve class imbalance. All things considered, categorical visualisations enable informed decision-making, aid to prepare info for modelling, and assist to comprehend dataset composition.

3.2. Hypothesis Testing: In medical research, hypothesis testing is essential to drawing conclusions from sample data. In the context of datasets pertaining to hepatitis or heart disease, hypothesis testing is a useful tool for determining whether or not the observed patterns or variations within groups are statistically noteworthy or more likely the result of random chance. For example, one speculate that a certain medication reduces the risk of heart attacks. They might look for a statistically significant difference between the treated and untreated groups' mean heart attack rates using a t-test. Chi-square testing might also be used to find if categorical variables such as hepatitis stages & survival rates have any correlation. In clinical studies, hypothesis testing supports validation of results; it also guides therapy strategies and shapes policy choices. It is a basic approach for inferring population parameters from sample data, therefore guaranteeing that statistical evidence supports rather than just observing the results. The working of the hypothesis testing is shown in figure 4

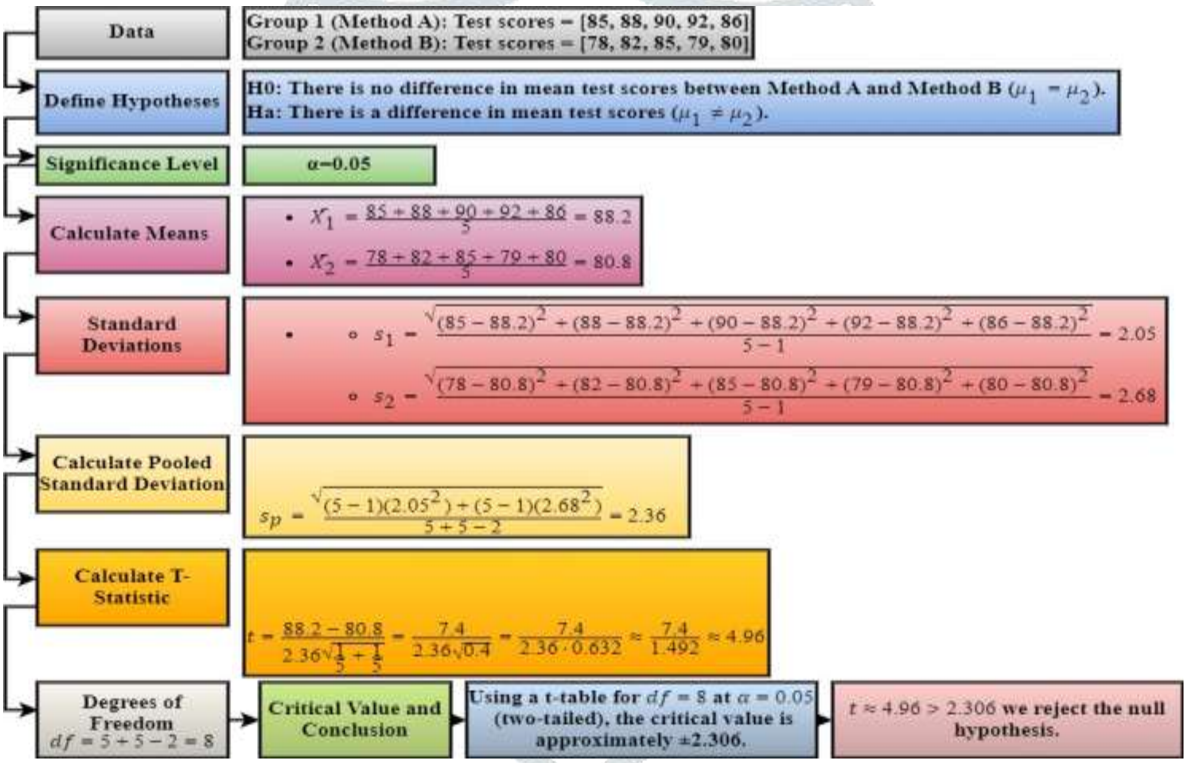


Figure 4: Working of Hypothesis Testing

The calculation of T-Statistic involves the following formula as shown in eq (3)

t-statistic $t = \frac{X-\mu 0}{\frac{s}{\sqrt{n}}}$ – (3)

Where:

- t is the **t-statistic**.
- X is the **sample mean**. $X = {}^{1}\sum^{n}$ X – (4)

$n \quad i=1$

- $\mu$is the **known population mean** (hypothesized mean).
- s is the **sample standard deviation**.
- n is the **sample size**.

3.3. Standard Scalarization: Before analysis begins, every feature in a dataset are standard scalarized to ensure they all have an identical mean of 0 & standard deviation of 1. Age, weight, & cholesterol levels are examples of features in medical datasets that may have varied scales. This may have an impact on how well machine learning models perform, particularly those that depend on gradient-based techniques (like neural networks) or distance measurements (like SVM or k-NN). Using standard scalarization makes all the features the same size. This stops features with wider ranges from controlling the learning process of the model. This normalisation speeds up gradient descent in logistic regression and deep learning models and improves feature magnitude-sensitive methods. Standard scalarization is especially important for medical data because factors can have very different units, like heart rate in beats per minute versus blood pressure in millimetres of mercury. Making sure that the scale is the same makes the model easier to understand and more reliable.

3.4. Zero-Shot Inference

A ML methodology can correctly guess what will happen in situations it hasn't been directly trained on. This is called "zero-shot inference." Usually, for models to identify, they need a labelled dataset with specific categories. But zero-shot learning gets around this problem by using models that have already been trained, like large-scale language models, to do new jobs they haven't seen before. Zero-shot inference is useful for medical data about rare diseases or situations where labelled data is hard to find or scarce. In certain cases, a model trained to detect diabetes or cardiovascular disease may also detect less frequent illnesses with comparable symptoms and diagnostic criteria. Zero-shot models usually depend on understanding tasks from a conceptual point of view and using representations from very large sets of data, like medical literature text, to find links between categories that are known and those that are not known.

A zero-shot inference method is used in NLP to find drugs, answer medical questions, or even make diagnoses without needing training data that is relevant to the job at hand. By comprehending the context and connections between medical terminology, a system may respond to enquiries concerning illnesses it has not met in the training data by using transformer-based models that have already been trained, such as GPT or BERT. Zero-shot inference has a huge amount of promise in healthcare, especially for personalised medicine, where the conditions and treatments of each patient may be very different from the training data that was used to make the model. But there are some problems, like the fact that models need to have a deep knowledge of the domain's meaning and that it's hard to be sure that zero-shot models' conclusions are correct and reliable. Because of this, zero-shot inference is a strong but careful tool in touchy areas like healthcare, where wrong guesses can have big effects.

3.5. Meta-Learning

Meta-learning, commonly referred to as "learning to learn," is a domain of machine learning that emphasises the development of models capable of acquiring new tasks or adjusting to novel environments with little or no data or prior experience. Meta-learning tries to make models that can quickly switch from one job to another, while traditional ML methodology need a lot of data and a lot of training for each task. This is useful in medical research for uncommon illnesses or for deploying models across hospitals or patient groups with various data distributions. Meta-learning is based on the idea that instead of teaching a model to do well on a single assignment, it should be taught on a lot of different tasks in order to it can master the structure of the issues. The model is able to adapt to a new job with only a few examples, which makes it more adaptable and useful. The Model- Agnostic Meta-Learning (MAML) method is a popular way to do meta-learning.

It changes a model's settings so that it can do well on new tasks with just a few gradient updates. In medical settings, this could mean teaching a model about different diseases or how patients did and then using it on a new disease in which there isn't a lot of data. One additional application is in personalised healthcare, where the data of each patient may be slightly different. A meta-learned model can rapidly adapt to the unique requirements of each patient with minimal fine-tuning. In other types of meta-learning, you learn how to choose the best hyper parameters for algorithms or how to add to data in a way that makes it work better. One way to make flexible systems that can learn on their own in changing contexts is to mix meta-learning with other methods like reinforcement learning. Meta- learning does have some problems, though. For example, it costs a lot to run and needs a lot of different jobs during training to make sure the framework can generalise well. Despite these problems, meta-learning has a lot of potential to help AI systems get better in areas that change quickly and have limited data, such as natural language understanding, robots, and healthcare.

4. RESULTS & DISCUSSION:

Table2: Imbalanced Attributes Description of Categorical Data

Dataset Name	**Attribute Name**	**Values**	**Values**	**Remarks**
Hepatitis	Sex	Male	Female	
		10.3%	89.7%	Number of females are very high compared to female
		True	False	

	Steroid	50.3%		49.7%		Usage of
						steroids and non-usage are almost equal
	Anti Virals	15.5%		84.5%		Most of the patients are not anti viral
	Fatigue	64.5%		35.5%		Most of the patients are fatigue
	Malaise	40%		60%		Malaise traces in patients are less in number
Heart	Sex	Male		Female		
		69.5%		30.5%		More number of males when compared to female
		0	1	2	3	
	CP	48.5%	16.3%	27.7%	7.5%	Type-0 chest pain is most common and type-3 is less common
	Thal	0.7%	6.2%	53.1%	40%	Type-2 is most common and type-0 is rare
		0		1		
	Fbs	85%		15%		Most of the patients has controlled sugar levels
	Exang	66.3%		33.7%		Most of the patients didn’t suffer from pain during physical activity

This table describes imbalanced attributes within medical datasets like heart disease and hepatitis. It highlights distributions of categorical variables such as gender, steroid usage, and antiviral medication usage. For example, in the hepatitis dataset, 89.7% of patients are female, and only 15.5% use antivirals, indicating significant imbalances. Imbalanced data can skew model outcomes, necessitating preprocessing techniques like data balancing for accurate predictions. Addressing such imbalances ensures that models do not overfit to the majority class.

Table 3: Statistical Summary on Numerical Attributes

Dataset	**Attribute**	**Mean**	**Std**	**Min**	**Max**	**Remarks**
Heart	Age	54.43	9.07	29	77	The age distribution is fairly wide, covering a broad range of adulthood.
	Trestbps	131.61	17.52	94	200	Most values lie within 17.52 mm Hg of the mean

	Chol	246	51.59	126	564	The standard deviation of 51.59 mg/dL indicates significant variability, suggesting a wide range of cholesterol levels among the individuals.
Hepatitis	Age	41.2	39.1	7	78	Good distribution of patients in between 32 & 50
	Bilirubin	1.43	2.6	1	8	The presence of outliers is indicated by the maximum value of 8.0 mg/dL, which is significantly higher than the third quartile.

The statistical summary table provides key descriptive statistics for numerical attributes in medical datasets. Attributes like age, cholesterol levels, and blood pressure in heart disease and hepatitis data are detailed with mean, standard deviation, minimum, and maximum values. For example, in the heart dataset, age ranges from twenty two to seventy seven years with a mean of 54.43, while cholesterol levels vary widely, indicating diverse patient profiles. These summaries aid in understanding data distribution and variability. Such insights are vital for preparing data for machine learning models.

Table 4: Hypothesis Testing Results

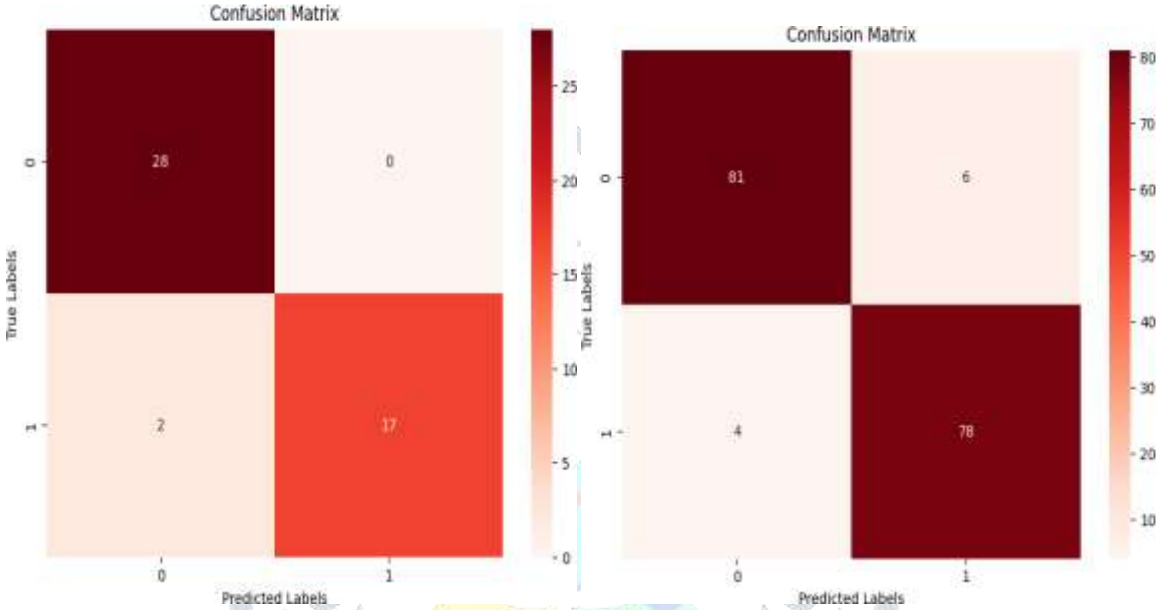
Dataset	Attribute Name	T-Value	P-Value	Remarks
Heart	Age	7.53	1.06	There is a correlation between age and heart attack
	Chol	3.23	0.001	There is no correlation between chol and heart attack
Hepatitis	Age	-2.78	0.006	There is no correlation
	Bilirubin	-6.34	2.57	There is no correlation

The table 4 summarizes the results of hypothesis testing performed on medical datasets like heart and hepatitis data. For each attribute, it provides the T-value, P-value, and interpretation. For instance, a significant correlation between age and heart attacks is indicated with a T-value of 7.53 and P-value of 1.06. The table also shows instances where no correlation was found, such as between cholesterol levels and heart attacks. Hypothesis testing helps verify whether observed data patterns are statistically significant. This process plays a crucial role in validating medical findings.

Table 5: Identification of Outlier Attributes

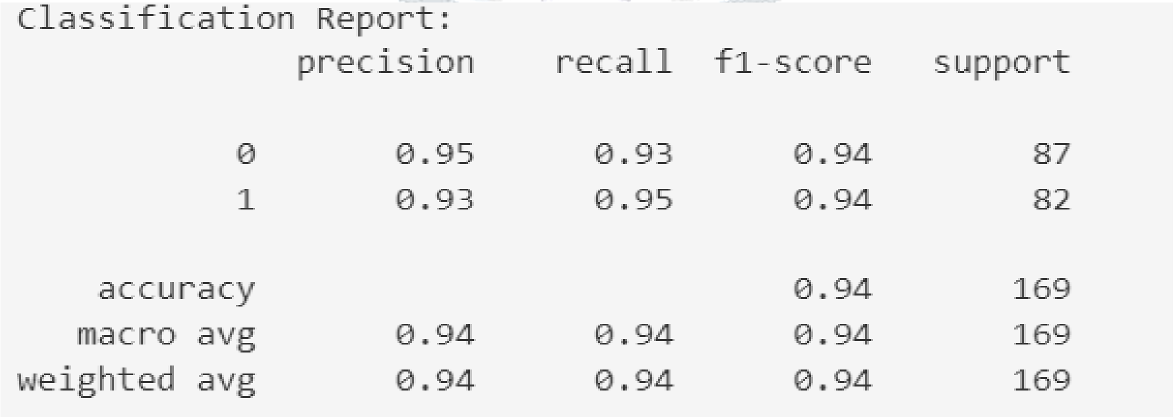
Dataset	Number of Attributes	Number of Outliers Removed	Number of Actual Records
Heart	14	256	1025
Hepatitis	20	10	155

The table 5 identifies attributes within medical datasets, such as heart disease and hepatitis, that contain outliers. It includes the number of attributes analyzed, the number of outliers removed, and the remaining records. For instance, in the heart dataset, 14 attributes were examined, and 256 outliers were removed, leaving 1,025 records. This table is critical for understanding how outlier removal can influence the dataset size and model performance. Proper handling of outliers improves the accuracy of predictive models.



(a) Hepatitis (b) Heart

Figure 5: Confusion MatrixThe confusion matrix is a visual representation of the performance of classification algorithms, particularly in heart disease and hepatitis datasets. It demonstrates how effectively the model detects correct positives, correct negatives, incorrect positives, and incorrect negatives. This is important for evaluating the model's performance, precision, and recall, highlighting areas where it may struggle, such as misclassifying healthy individuals as having a disease or vice versa. The confusion matrix offers insights into the reliability of the model's predictions.



```
Classification Report:
              precision    recall  f1-score   support

           0       0.95      0.93      0.94        87
           1       0.93      0.95      0.94        82

    accuracy                           0.94       169
   macro avg       0.94      0.94      0.94       169
weighted avg       0.94      0.94      0.94       169
```

Figure 6: Classification Report of Heart

The figure 6 illustrates the classification report for heart disease prediction, presenting metrics such as precision, recall, F1-score, and support for each category (e.g., patients with or without heart disease). Precision reflects how well the model identifies positive outcomes, while recall demonstrates the model's effectiveness in recognizing all positive cases. The F1- score, which balances precision and recall through a harmonic mean, offers a single measure to summarize the model's overall performance. This report plays a key role in evaluating the effectiveness of the model and highlighting areas for further improvement.

5. CONCLUSION: In this study, the model presented a combined approach that integrates meta-learning with zero-shot learning to enhance classification performance in healthcare datasets, specifically targeting heart disease and hepatitis. By leveraging meta-learning, we trained the model to adapt quickly to new tasks with limited labeled data, while zero-shot learning enabled it to make predictions on unseen classes based on learned knowledge. The implementation of this framework resulted in strong performance metrics across both datasets. For the heart disease dataset, our model achieved An ultimate level of accuracy of 94%, with a balanced precision and recall of 93% and 95%, respectively. This performance underscores the model's capability to effectively distinguish between healthy and diseased individuals, crucial for timely clinical interventions. Similarly, the model demonstrated its effectiveness on the hepatitis dataset, achieving precision, recall, and F1 scores of 94%, along with a ROC AUC of 93%. These results indicate a robust ability to accurately predict hepatitis presence, potentially aiding healthcare professionals in diagnosis and treatment planning. Overall, the successful integration of meta-learning and zero-shot learning not only improved classification accuracy but also demonstrated the potential for our approach to to extrapolate from various data sets in medicine. This dual approach addresses the challenges posed by limited labeled data in the medical domain and offers a promising avenue for future research aimed at enhancing predictive analytics in healthcare. Further exploration of this methodology with additional datasets could help validate its efficacy and pave the way for more sophisticated models in clinical settings.

REFERENCES:

[1] ware*, S., Rakesh, S. K., & Choudhary, B. (2020). Heart Attack Prediction by using Machine Learning Techniques. *International Journal of Recent Technology and Engineering (IJRTE)*, *8*(5), 1577–1580. https://doi.org/10.35940/ijrte.D9439.018520

[2] Jindal, H., Agrawal, S., Khera, R., Jain, R., & Nagrath, P. (2021). Heart disease prediction using machine learning algorithms. *IOP Conference Series: Materials Science and Engineering*, *1022*(1). https://doi.org/10.1088/1757-899X/1022/1/012072

[3] Saboor, A., Usman, M., Ali, S., Samad, A., Abrar, M. F., & Ullah, N. (2022). A Method for Improving Prediction of Human Heart Disease Using Machine Learning Algorithms.
Mobile Information Systems, *2022*. https://doi.org/10.1155/2022/1410169

[4] Kavitha, M., Gnaneswar, G., Dinesh, R., Sai, Y. R., & Suraj, R. S. (2021). Heart Disease Prediction using Hybrid machine Learning Model. *Proceedings of the 6th International Conference on Inventive Computation Technologies, ICICT 2021*, 1329–1333. https://doi.org/10.1109/ICICT50816.2021.9358597

[5] Ibrahim Alsaadi, D., Abdulrazzaq Mohammed Ali, H., & Ali Muhsen, A. (2024). *Designing Model to Improve Hepatitis Prediction by Using Data Mining and Machine Learning Algorithms*. *19*, 110–119. https://doi.org/10.5281/zenodo.10049652#137

[6] Ahammed, K., Satu, M. S., Khan, M. I., &Whaiduzzaman, M. (2020). Predicting Infectious State of Hepatitis C Virus Affected Patient's Applying Machine Learning Methods. *2020 IEEE Region 10 Symposium, TENSYMP 2020*, 1371–1374.
https://doi.org/10.1109/TENSYMP50017.2020.9230464

[7] Ahmed, I. I., Mohammed, D. Y., & Zidan, K. A. (2022). Diagnosis of hepatitis disease using machine learning techniques. *Indonesian Journal of Electrical Engineering and Computer Science*, *26*(3), 1564–1572. https://doi.org/10.11591/ijeecs.v26.i3.pp1564-1572

[8] Prathyusha, B., Navya, S., Navya, Y., & Harika, N. S. Hepatitisnet: Analysis and Prediction Of Hepatitis Using Machine Learning.

[9] Yao, P., Shen, S., Xu, M., Liu, P., Zhang, F., Xing, J., ... & Xu, R. X. (2021). Single model deep learning on imbalanced small datasets for skin lesion classification. IEEE transactions on medical imaging, 41(5), 1242-1254.

[10] Zhang, H., Zhang, H., Pirbhulal, S., Wu, W., & Albuquerque, V. H. C. D. (2020). Active balancing mechanism for imbalanced medical data in deep learning-based classification models. ACM Transactions on Multimedia Computing, Communications and Applications, 16(1s). https://doi.org/10.1145/3357253

[11] Chen, Y., Calabrese, R., & Martin-Barragan, B. (2024). Interpretable machine learning for imbalanced credit scoring datasets. European Journal of Operational Research, 312(1), 357–372. https://doi.org/10.1016/j.ejor.2023.06.036

[12] Tsai, H., Yang, T. W., Wong, W. M., Kao, H. Y., & Chou, C. F. (2024). A Hybrid Approach for Binary Classification of Imbalanced Data. International Journal of Computational Intelligence and Applications, 2450013.
MRI. Scientific Reports, 10(1). https://doi.org/10.1038/s41598-020-76389-4

[13] Ye, H.-J., Chen, H.-Y., Zhan, D.-C., & Chao, W.-L. (2020). Identifying and Compensating for Feature Deviation in Imbalanced Deep Learning. http://arxiv.org/abs/2001.01385

[14] Yang, Y., Zha, K., Chen, Y., Wang, H., & Katabi, D. (2021, July). Delving into deep imbalanced regression. In International Conference on Machine Learning (pp. 11842-11851). PMLR.