Table of Contents

Chapter: 1 INTRODUCTION

Monkey pox is a zoonotic viral infection due to the monkey pox virus, which was initialy found in 1959 during the investigating pox like illness laboratory research monkeys in Denmark. These monkeys were part of controlled breeding colony, maintained for scientific and medical experiments. Although the monkeys themselves were not the natural reservoir of the virus, they acted as the initial carriers through which the virus was isolated and studied. This led to the virus being named “Monkey pox”.

This incident was the beginning of human surveillance and research on the disease.

Since then scenario, the virus was classified as a zoonotic pathogen, primarily transmitted from animals (likely rodents) to humans. For decades, monkey pox remained confined to central and West African rainforests.

However, the unprecedented global outbreak in 2022 marked a turning point, with over 104 countries reporting confirmed cases across six continents.

Traditional diagnosis of monkey pox involves clinical evaluation of monkey pox involves clinical evaluation followed by confirming testing initially through PCR (Polymerase Chain Reaction).

However, these methods are not feasible for mass deployment. Early and efficient diagnosis is considered as crucial, not only for patient care but also for containing outbreaks by promoting timely isolation and treatment.

The symptoms of monkey pox closely resemble to small pox and chicken pox, though generally milder. The initial symptoms including fever, fatigue, headache, swollen lymph nodes and skin rashes. Due to overlap of symptoms with other viral infection early diagnosis is essential to prevent further transmission.

The rise in global monkey pox cases and limitation of diagnosis motivate to explore the application of Artificial Intelligence along with Machin Learning.

In this study, Machine Learning (ML) Model are train to analyze symptomatic data from patient and to detect weather the individual is monkey pox positive or negative. In this, the objective to identify most effective ML model in terms of performance and accuracy to explore how such model can detect early decision making.

This research to determine the most appropriate model for early and accurate detection. There for Machine Learning to the contributing to the broader field AI based medical diagnostics.

For machine learning models- Decision Tree, Logistic Regression, Random forest, Naïve Bayes- were implemented and evaluated based on performance metrics such as accuracy precision, recall, F1-Score. The study concludes the implication of ML for disease surveillance for future direction.

Chapter:2 Why This Topic

- Monkey Pox is a communicable disease for which no human physical contact is required as the detection is done through machine learning models.

- The process of detection is carried out through training and testing models of machine learning.

- Using different models for the detection of Monkey Pox in order to analyze which model exhibits the best performance of detection.

The choice of monkey pox detection using machine learning is motivated by several factors:

- Emerging Public Health Concern: Monkey pox, though rare, poses a significant threat due to its communicable nature and global spread.

- Advanced Technology Application: Machine learning offers a novel approach to enhance diagnostic accuracy in medical science.

- Non-Invasive Detection: ML models can diagnose monkey pox based on symptomatic data, reducing the need for physical contact.

- Model Comparison: Evaluating multiple ML models allows identification of the most effective algorithm for monkey pox detection.

Research Questions

- Develop Predictive ML Models
- Identify Key Predictive Features
- Explore Multimodal Symptoms Analysis
- Evaluate Model
- Enable Rapid Clinical Decision Support

Chapter: 4 Literature Review

- “Alnaji (2024) emphasized on monkey pox contact tracing and its breakage management. Effective contacts tracing policies were used to reduce the spreading of disease which was detected through the study. The fundamental of the study is based on contact tracing.”
- “Kulkarni and Verma(2024) identified monkey pox disease classification using a histogram of SVM resulting in effective disease classification from images in clinical settings, the study demanded the investigation on larger and detailed datasets.”
- “Gupta et al. (2023): Worked on monkey pox early detection on healthcare monitoring using block chain technology it resulted in order to secure disease surveillance with the potential for early warning systems”
- “Kundu Rahman (2024) worked on monkey pox skin lesion dataset using image synthesis. The study diversified skin lesions datasets and improve the performance of lesion classification models.”
- “Ren & Wang (2023) used multiple model data fusion to investigate the outcome as per the diagnosis accuracy exploring multiple sources of data.”
- “Singh et al. (2024) worked on modeling monkey pox transmission on basis mobility data for which structural data was used and the reason for the transmission of disease was to be identified.”
- “Rampogu (2023) worked on Monkey pox transmission networks using graph-based analysis, the study detected the key transformation to guide targeted intervention and it is dependent for availability of contact tracing data

- “Maqsood & Forkert (2024) employed multiple modal data for Monkey pox risk assessment and the outcome was an intensive evaluation of individual and community risk factors. However, the study requires access to diverse data sources which are challenges in real-world deployment”
- “Rallapalli et al. (2023) developed a study on monkey pox vaccine effectiveness using Bayesian statistical models and casuals. The study quantifies the impact of various factors on effectiveness on vaccine. Using Baysian statistical models when casual inference.”
- “Chunhapran Maliyeam (2024) operated on deep learning methods using the symptoms monkey pox rash and lesions for classification in order to research precisely the process of disease to be detected.”
- “Pikulkaew et al. (2024) elaborated a research on the basis of monkeypox lesion classification by analyzing through image processing which will further improve the accuracy of skin lesion classification giving idea to detect the possibility of this disease.”
- “Gairola and Kumar (2022) researched on Monkey pox diagnosis on the basis of symptom prediction, lesion detectionin order to detect the risk using multiple tasking deep neural networks and to study for the evaluation of monkeypox status .”
- “Eliwa et. al (2023) applied CNN for skin lesion classification of monkey pox for automatic skin lesion to investigate the skin classification of monkey pox in order to study rapid diagnosis .”
- “ Shah (2022) used CNN for segmentation of monkey pox skin lesion based on images of dermatology. The study detected aacurate marking of monkey pox. The research granted contents to public to generate awareness about the spread of this monkeypox disease.”
- “ Mohbey et al. (2024) used CNN for monkeypox analysis on the basis of information available on social media .This reseach offers materials to generate awareness of monkeypox amongst in the society.”

- “Yolcu Oztel (2024)Used CNN along with Vision transformation for monkeypox skin lesion classification .This research was conducted to generalize other skin diseases as well but it stopped visual analysis.”
- . “Kakulapati (2023)UNETS and VGG16 were applied for the diagnosis of monkeypox skin lesion detection having multiple datasets but not necessarily containing every skin tone.’’
- “Towhidul Islam et. al (2022) conducted the research with the help of web scrapping in order to make a database which are effected by the measles, cow pox, chicken pox and small pox. The same kind of images are detected for each class enabling to improve in the best version to other datasets. Special vadidation is required from the doctors for the confirmation of the disease. This will enhance the better opportunity to model for the classification of different skin lesion.’’
- “Ahsan et. al. (2022) this was the research conducted for the diagnosis of image which was carried out for the detection of covid-19 patient for machine learning was proved to perform with high potention for the detection of diseases.”
- “Kolluri et al. (2022) carried out study through the preparation of open soarce which was based on the data of official worls health organization through the verification of factual soarces.”
- “ Patel et al.(2022) In the study basically focus on artificial intelligence enabling machine learning as the helpful tool to detect the pros and cons of the disease for future scope.”
- “Eid et al. (2022) presented a novel approach for the accurate prediction of monkeypox confirm cases for which different techniques were adopted which were based on tissue in optimized long memory.
- This technique was useful for the detection of monkeypox in order to assist the diagnosis .”

Chapter: 5 Research Gaps in Monkey Pox Studies

- **Limited Data on Long-Term Effects**
- **Vaccine Effectiveness Uncertainty**
- **Insufficient AI-Based Diagnostic Models**
- **Understudied Transmission Patterns**
- **Zoonotic Reservoirs Not Fully Understood**

Chapter: 6 Methodology

Use Monkey Pox dataset focusing on common symptoms.

1. Collection of data

2. Data Preprocessing

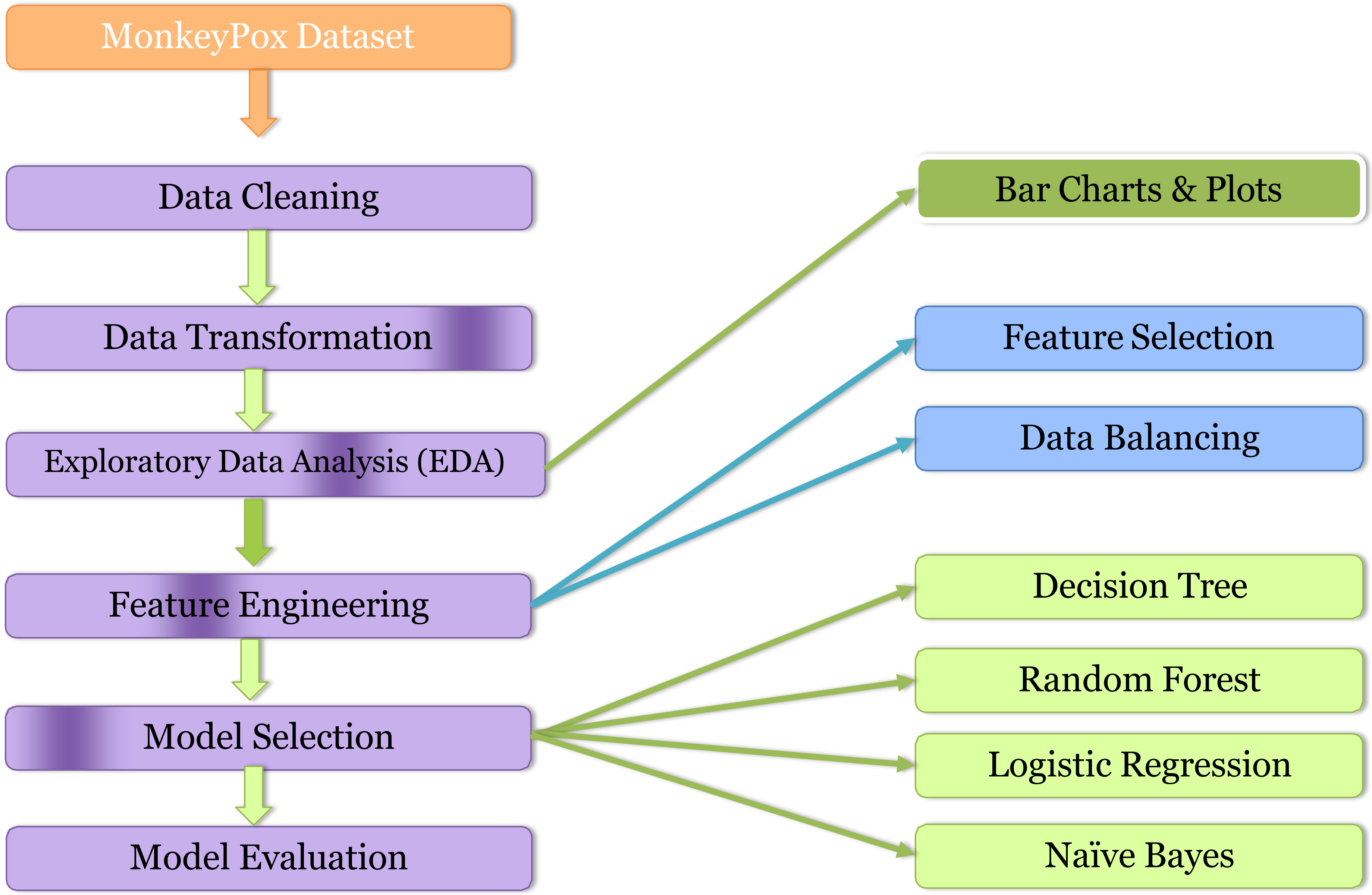
3. Exploratory Data Analysis (EDA)

4. Feature Selection

5. Model Selection & Training

- **[i] Applied Decision Tree Classifier**
- **[ii] Applied Logistic Regression**
- **[iii] Applied Naïve Bayes**
- **[iv] Random Forest**

6. Model Evaluation



MonkeyPox Dataset
Data Cleaning
Data Transformation
Exploratory Data Analysis (EDA)
Feature Engineering
Model Selection
Model Evaluation
Bar Charts & Plots
Feature Selection
Data Balancing
Decision Tree
Random Forest
Logistic Regression
Naïve Bayes

Chapter: 7 Data Overview and Visualization

Fig:1

	Patient_ID	Systemic Illness	Rectal Pain	Sore Throat	Penile Oedema	Oral Lesions	Solitary Lesion	Swollen Tonsils	HIV Infection	Sexually Transmitted Infection	MonkeyPox
0	P0	NaN	False	True	True	True	False	True	False	False	Negative
1	P1	Fever	True	False	True	True	False	False	True	False	Positive
2	P2	Fever	False	True	True	False	False	False	True	False	Positive
3	P3	NaN	True	False	False	False	True	True	True	False	Positive
4	P4	Swollen Lymph Nodes	True	True	True	False	False	True	True	False	Positive

Data Collection resource:

In this study, data is collected from " GitHub" source known as the "Monkey pox symptom based Dataset"

Number of Records: 25,000 patients

Number of Features:

11 attributes (Systemic Illness, Rectal Pain, HIV, Oral Lesions, etc.)

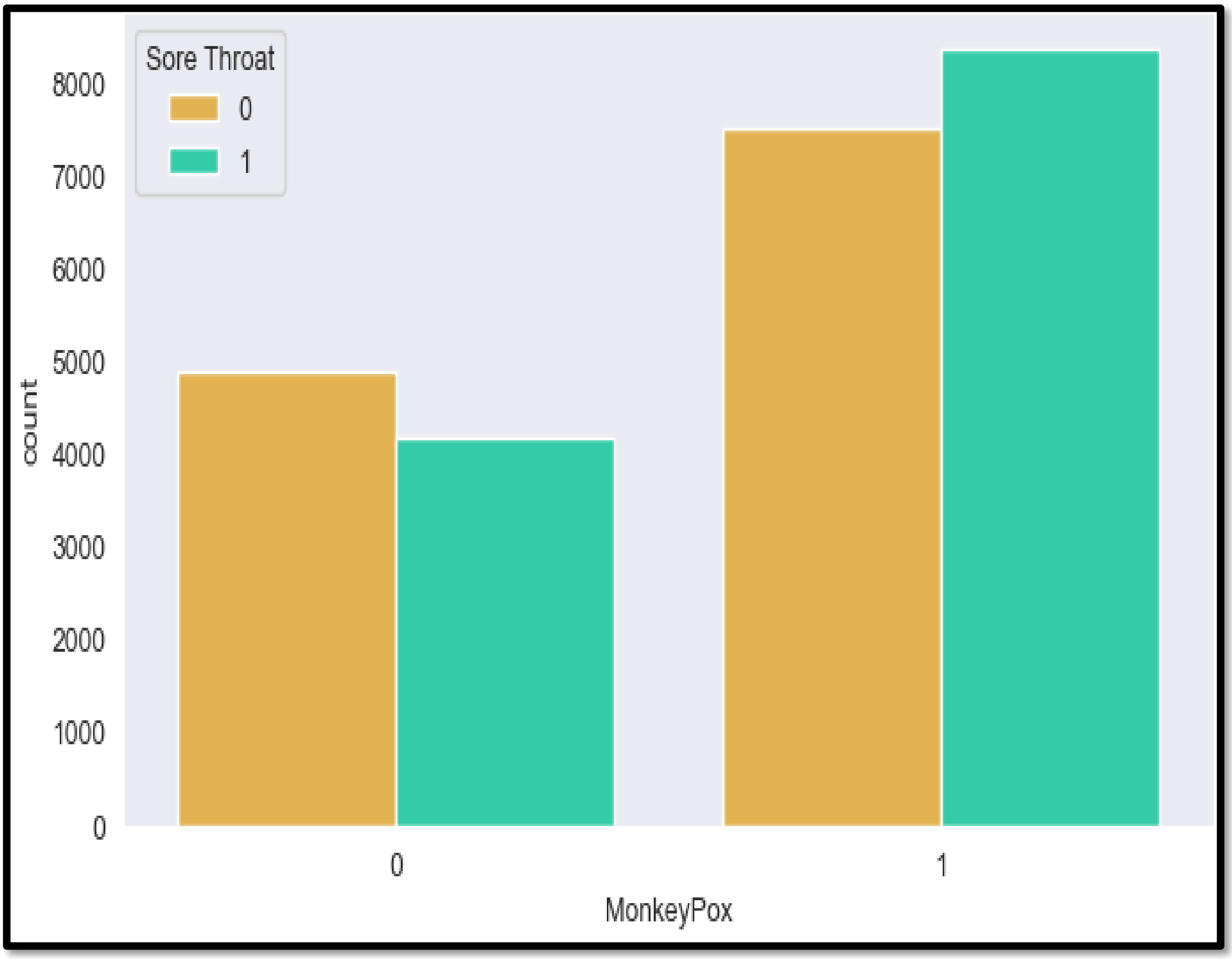
Monkey Pox (Positive) patients counted: 15909

Monkey Pox (Negative) Patient counted: 9091

Target Variable: Monkey Pox (Positive/Negative)

Understanding Some Feature

Fig:2




```
Total Observation of Dataset :  25000
Total MonkeyPox Positive :  15909
Total MOnkeyPox Negative :  9091
Sore Throat Positive :  12554
Sore Throat Negative :  12446
MonkeyPox Positive & Sore Throat Positive :  8370
MonkeyPox Positive & Sore Throat Negative:  7539
MonkeyPox Negative & Sore Throat Positive:  4184
MonkeyPox Negative & Sore Throat Negative:  4907
```

The bar graph shows the frequency distribution of the symptom sore throat in patients, categorized by their monkey pox test result-positive (1) or negative (0)

Color Representation:

- Green Bar: Sore throat present (positive)
- Yellow Bar: Sore throat absent(negative)

Monkey pox Classification:

“1” on x-axis =Monkey pox positive

“0” on x-axis=Monkey Pox negative

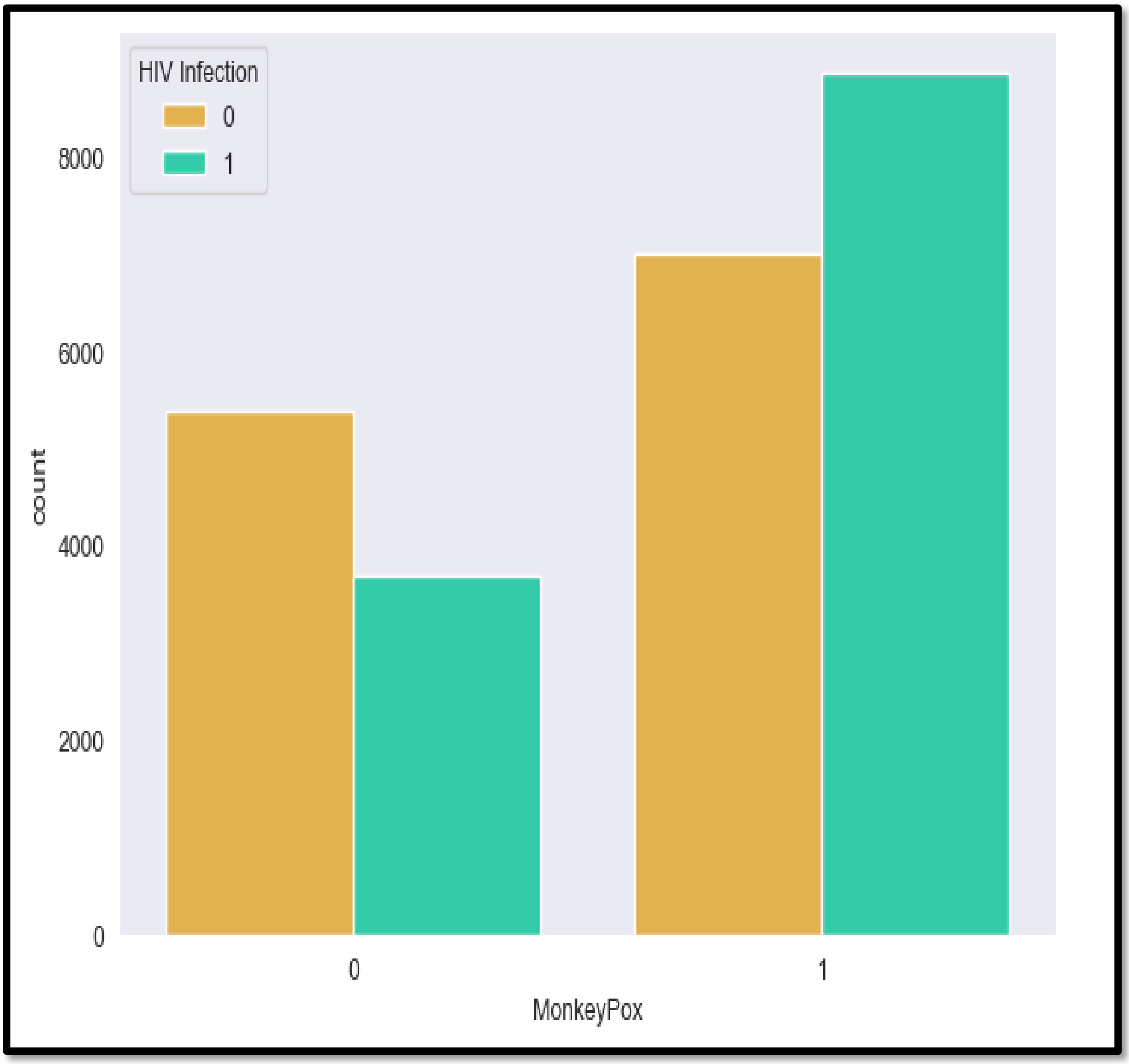
Implication for Model Training:

The frequency of sore throat in positive cases makes it a relevant symptom, contributing to the predictive model for identifying monkey pox

Explanation of Bar Graph and Output

Understanding Some Feature

Fig:3




```
Total Observation of Dataset :  25000
Total MonkeyPox Positive :  15909
Total MonkeyPox Negative :  9091
Total HIV Infection Positive :  12584
Total HIV Infection Negative :  12416
MonkeyPox Positive & HIV Infection Positive :  8887
MonkeyPox Positive & HIV Infection Negative:  7022
MonkeyPox Negative & HIV Infection Positive:  3697
MonkeyPox Negative & HIV Infection Negative:  5394
```

Explanation of Bar Graph and Output

The bar graph shows the frequency distribution of the symptom HIV Infection in patients, categorized by their monkey pox test result-positive (1) or negative (0)

Color Representation:

- Green Bar: HIV Infection present (positive)
- Yellow Bar: HIV Infection absent(negative)

Monkey pox Classification:

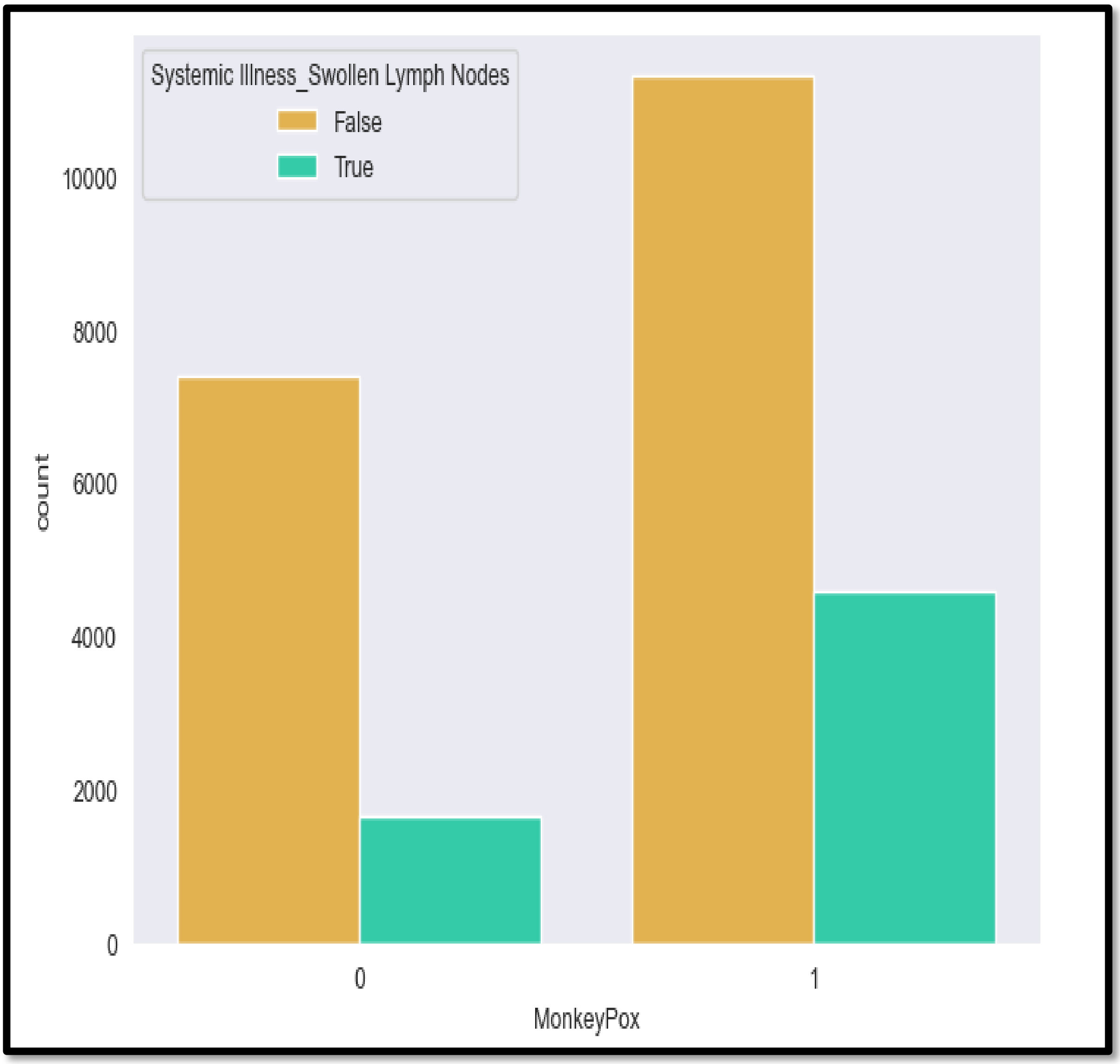
- “1” on x-axis =Monkey pox positive
- “0” on x-axis=Monkey Pox negative

Implication for Model Training:

- The frequency of HIV Infection in positive cases makes it a relevant symptom, contributing to the predictive model for identifying monkey pox

Understanding Some Feature

Fig:4




```
Total Observation of Dataset :  25000
Total MonkeyPox Positive :  15909
Total MOnkeyPox Negative :  9091
Systemic Illness_Swollen Lymph Nodes Positive :  6252
Systemic Illness_Swollen Lymph Nodes Negative :  18748
MonkeyPox Positive & Systemic Illness_Swollen Lymph Nodes Positive :  4581
MonkeyPox Positive & Systemic Illness_Swollen Lymph Nodes Negative:  11328
MonkeyPox Negative & Systemic Illness_Swollen Lymph NodesPositive:  1671
MonkeyPox Negative & Systemic Illness_Swollen Lymph Nodes Negative:  7420
```

Explanation of Bar Graph and Output

The bar graph shows the frequency distribution of the symptom Systemic Illness in patients, categorized by their monkey pox test result-positive (1) or negative (0)

Color Representation:

- Green Bar: Systemic Illness present (positive)
- Yellow Bar: Systemic Illness absent(negative)

Monkey pox Classification:

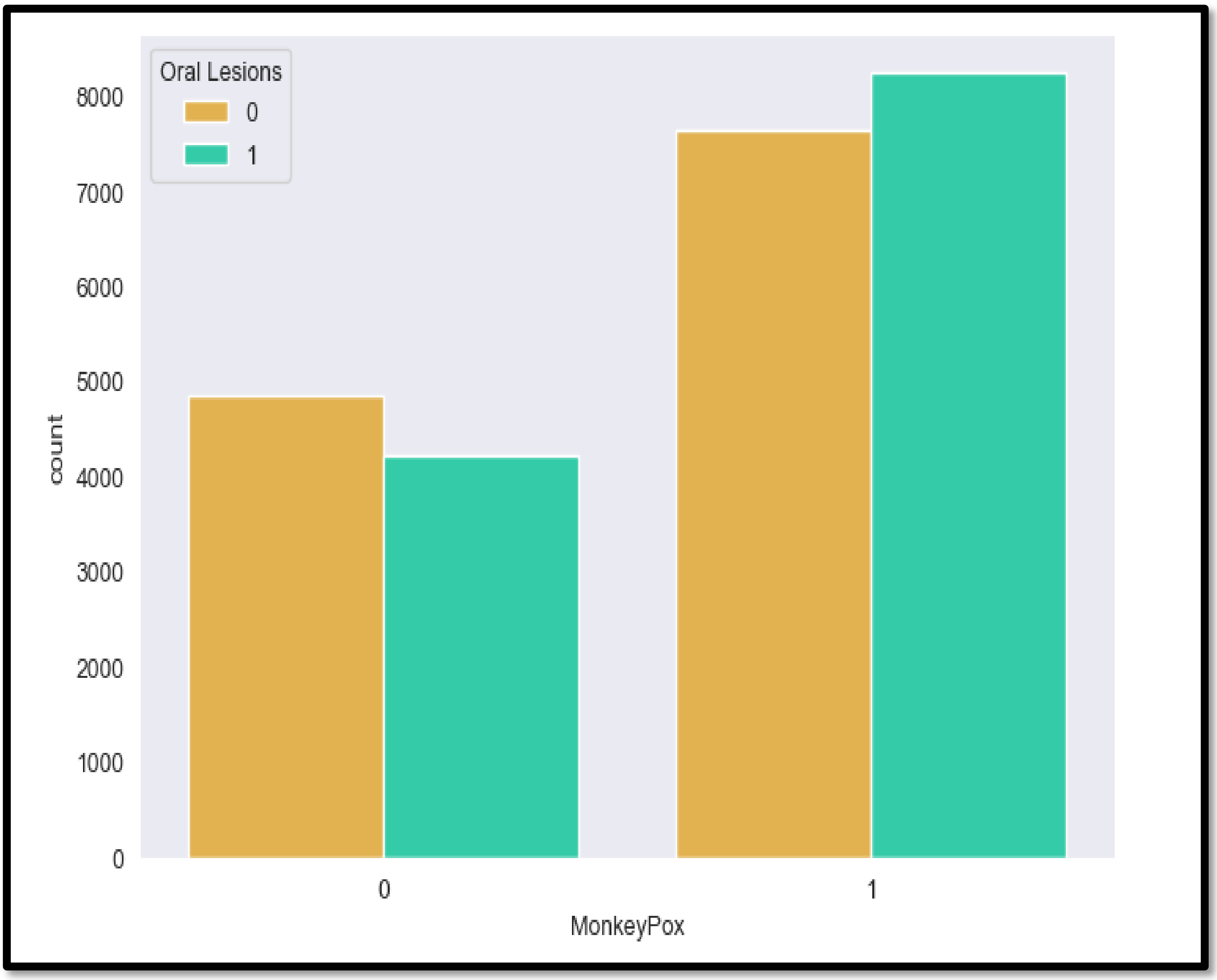
- “1” on x-axis =Monkey pox positive
- “0” on x-axis=Monkey Pox negative

Implication for Model Training:

- The frequency of Systemic Illness in positive cases makes it a relevant symptom, contributing to the predictive model for identifying monkey pox

Understanding Some Feature

Fig:5




```
Total Observation of Dataset :  25000
Total MonkeyPox Positive :  15909
Total MonkeyPox Negative :  9091
Oral Lesions Positive :  12486
Oral Lesions Negative :  12514
MonkeyPox Positive & Oral Lesions Positive :  8258
MonkeyPox Positive & Oral Lesions Negative:  7651
MonkeyPox Negative & Oral Lesions Positive:  4228
MonkeyPox Negative & Oral Lesions Negative:  4863
```

Explanation of Bar Graph and Output

The bar graph shows the frequency distribution of the symptom Oral Lesions in patients, categorized by their monkey pox test result-positive (1) or negative (0)

Color Representation:

- Green Bar: Oral Lesions present (positive)
- Yellow Bar: Oral Lesions absent(negative)

Monkey pox Classification:

- “1” on x-axis =Monkey pox positive
- “0” on x-axis=Monkey Pox negative

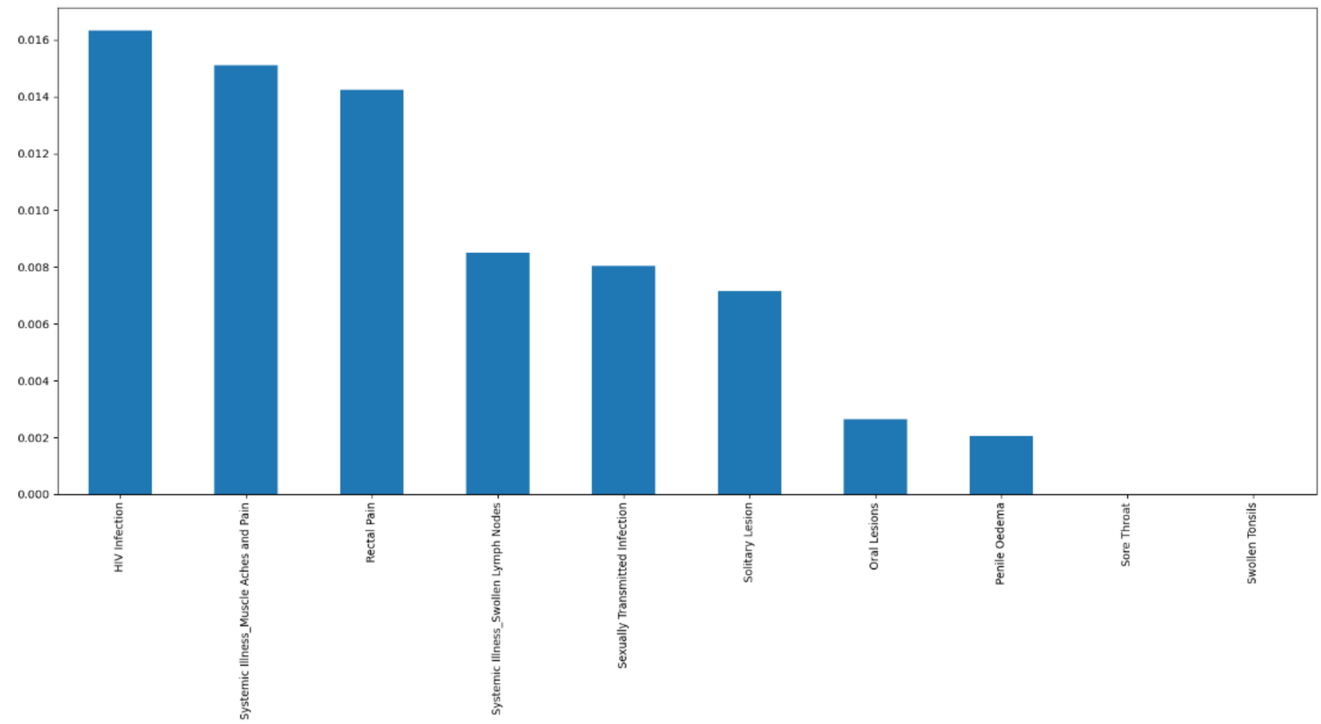
Implication for Model Training:

- The frequency of Oral Lesions in positive cases makes it a relevant symptom, contributing to the predictive model for identifying monkey pox

Correlation Coefficient

HIV Infection	0.016323
Systemic Illness_Muscle Aches and Pain	0.015099
Rectal Pain	0.014240
Systemic Illness_Swollen Lymph Nodes	0.008514
Sexually Transmitted Infection	0.008034
Solitary Lesion	0.007161
Oral Lesions	0.002648
Penile Oedema	0.002049
Sore Throat	0.000000
Swollen Tonsils	0.000000

Fig:6



If the correlation between these variables is observed then weak relationship is detected due to correlation value range between 0 to 0.25 .This will give rise to extremely weak correlation. Thus, disease variables has no correlations.

Chapter: 8 Data Preparation

- Data preparation is process of framing raw data into require data for analysis after processing.
- The initial methods are to collect to clean and label the initial data into require format suitable for machine learning algorithms followed data exploration and visualization.

- This method also called “Preprocessing”.

- Data undergoes transformation to ensure compatibility with analytical tool. The transformation process ensures normalization or numerical value transformation to common range encoding categorical variable into numerical form.

- Data preparation involves detection of outlier. These are carefully analyse for the errors to be corrected.

- Finally dataset divided into subset such as training and testing samples using different model. This helps us to evaluate between the outputs of the performance of different models. It also help us to detect the efficiency of different models with best performance of which model to be detected.

- Data preparation improve the quality of dataset ensuring analysis in a relevant way to study about the nature of the dataset on which we are working

Chapter: 9 Model Training & Evaluation

- **Data Preparation:**
  We prepared data for model training and evaluation before training any model data should neat and clean so

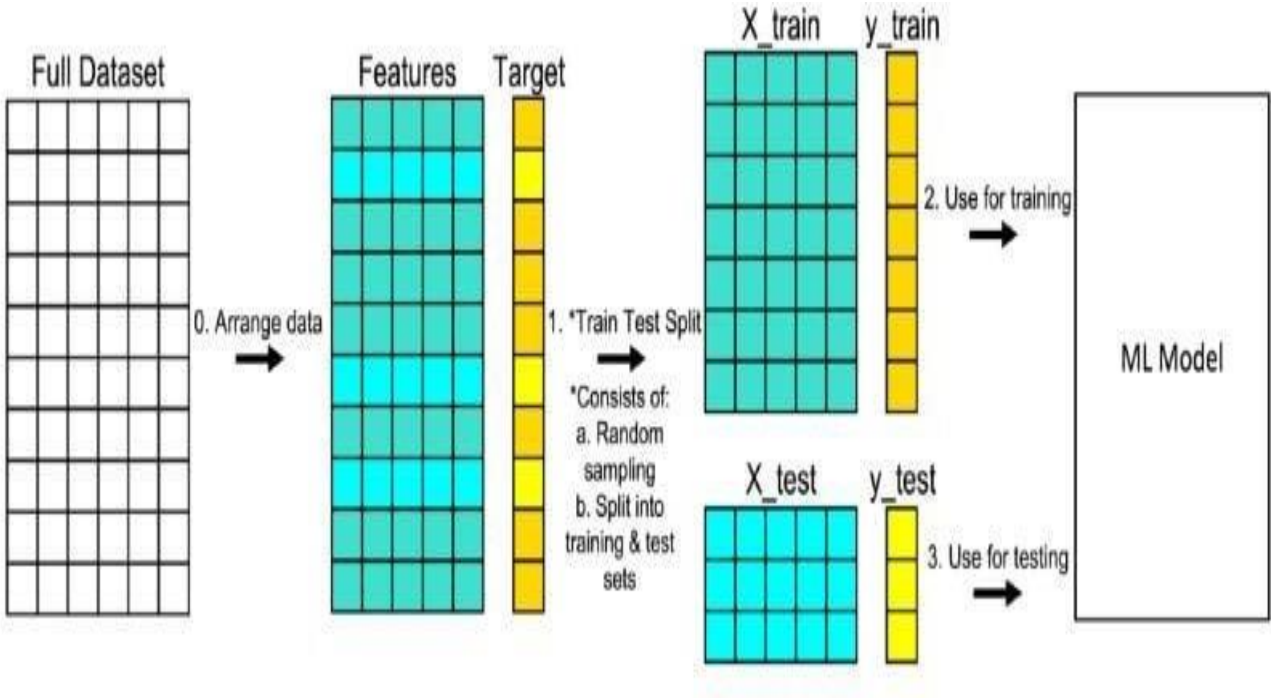
- **Train-Test Split:**
  The data is splitted into 80% fot the training process, and 20% for test data. Such data separation applies to all algorithms to be tested

- **Models Training**
  We are taking deferent -2 model training and testing we got best accuracy of both training and testing that model we select for further task.

- **Hyper parameter Tuning:**
  Hyper parameter tuning basically use for overcome the model overfitting and under fitting.

Chapter: 10 Train and Test split

Fig:7



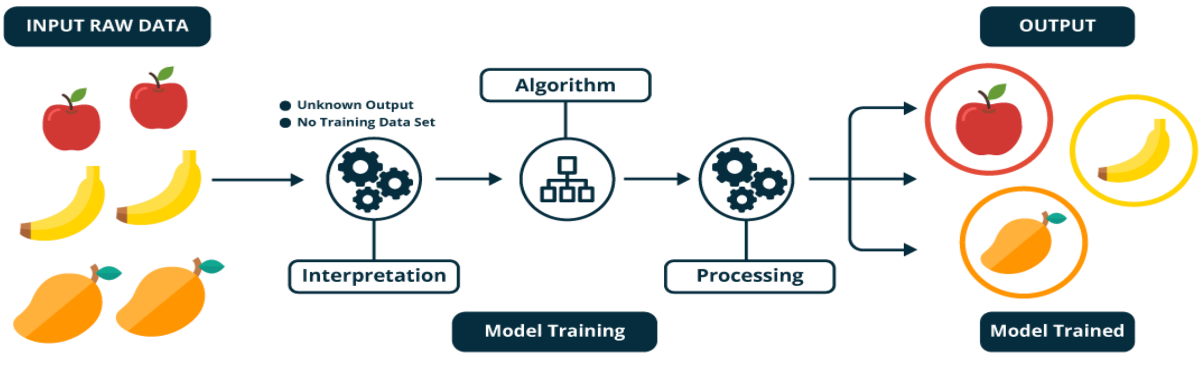
After feature selection we need to break down our data into two parts in this scenario we split our data 80% used for train our model and 20 % data is used for testing our data.

Where overall observation is 25000 patient which are affected to monkey pox disease. While twenty thousand data is used to train model and five thousands data is used to test the model.

Finally I found the accuracy of the model on training time and model accuracy on testing time which is best for training and testing both that model will be used for monkey pox case analysis after that select model and deployment any sever.

Model training and evaluation:

Fig:8



• Decision Tree: Achieved 68.69% training accuracy and 67.16% testing accuracy. Precision: 69% (positive), 59% (negative); Recall: 87% (positive), 32% (negative); F1-score: 77% (positive), 41% (negative).

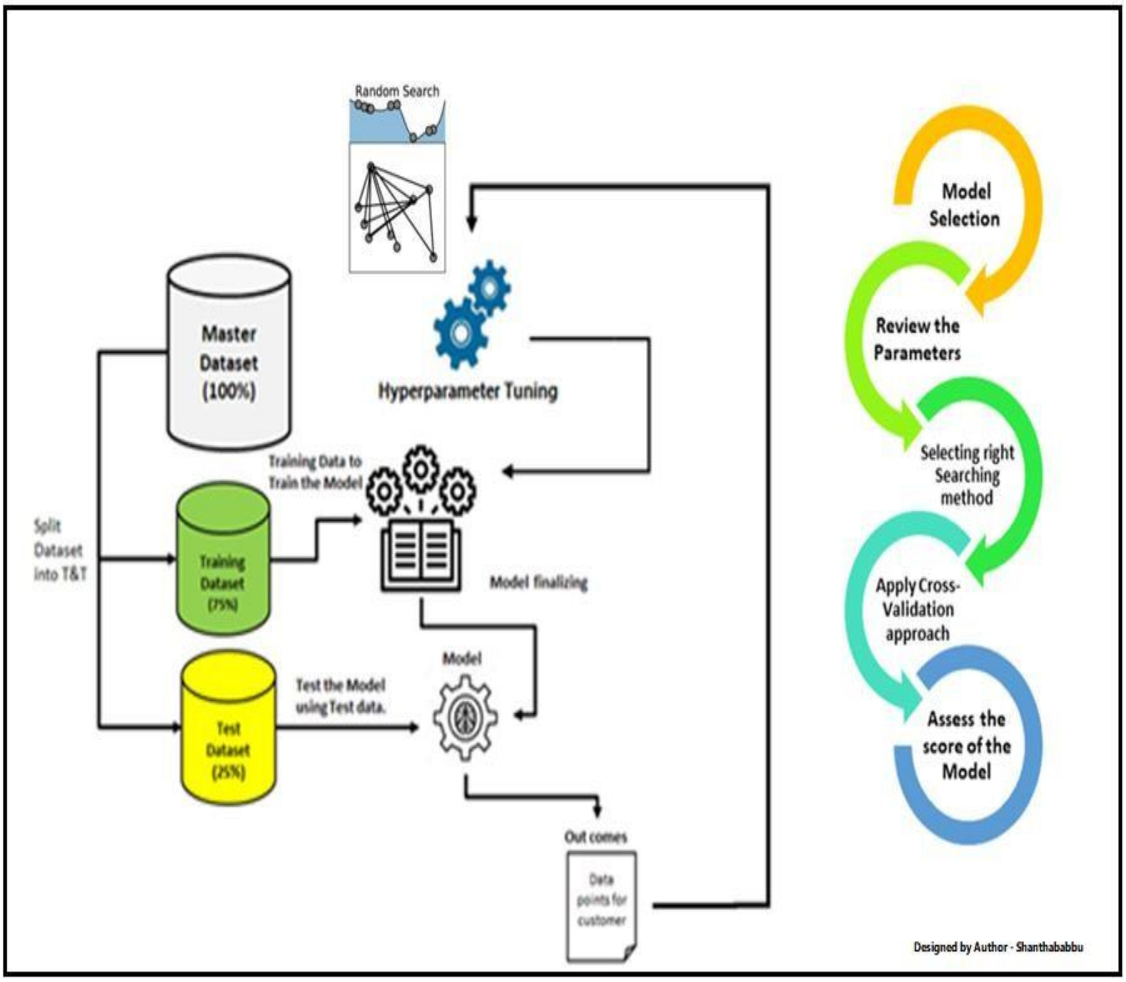
• Random Forest: Achieved 68.25% training accuracy and 67.76% testing accuracy. Precision: 68% (positive), 64% (negative); Recall: 92% (positive), 26% (negative); F1-score: 78% (positive), 37% (negative).

• Logistic Regression: Achieved 67.63% training accuracy and 67.56% testing accuracy. Precision: 70% (positive), 59% (negative); Recall: 87% (positive), 34% (negative); F1-score: 77% (positive), 43% (negative).

• Naïve Bayes: Achieved 67.7% training accuracy and 67.24% testing accuracy. Precision: 69% (positive), 60% (negative); Recall: 88% (positive), 31% (negative); F1-score: 77% (positive), 40% (negative).

Chapter: 11 Understanding Hyper parameter Tuning

Fig:9



Hyper parameter tuning is the process of selecting the best values for hyper parameters in order to avoid overfitting which is related to machine learning models.

The quality of excellent hyper parameter tuning requires good performance of machine learning models which are based on metrices for their specific tasks to be performed. Since, the tuning demands better version of performace , it is also called “Parameter optimization”..

Hyper parameter tuning is a process in which multiple experiments are carried out along with different trials containing different values .This process is carried out until best value is evaluated.

Data Splitting

We divide that information into 80% training data, and 20% test data. Such data division is used with all the algorithms to be tested Data splitting is a method for dividing a dataset into two or more subsets. When the data is split into two halves, the initial half is utilized to test or validate the data, and the remaining half is utilized to train the model. The reason behind constructing an ML model based on training sets is to learn pattern data in order to generalize the model on novel and unseen data. The second is the test set, where the model is provided with hypothetical data to predict the outcomes and verify the output to measure the model's performance

Classification Algorithm

We apply different machine learning models (ML) algorithm to carry out algorithms for meticulous research through optimization of performance by comaring outcomes. In our research ,The Decision Tree algorithm, Logistic Regression, Random Forest(RF) and Naïve Bayes were applied

Model Training and Evaluation

Decision Tree

	precision	recall	f1-score
0	0.59	0.32	0.41
1	0.69	0.87	0.77

Train Accuracy: 0.68695
Test Accuracy: 0.6716

Shows the classification results of the Decision Tree algorithm with a precision of 69% positive, and 59% negative. The recall value obtained is 87% which are positive and 32% were negative , . On F1-Score 77% positive, and 41% negative. The accuracy value obtained is 68.69% on training and 67.16% accuracy on testing.

Random Forest

	precision	recall	f1-score
0	0.64	0.26	0.37
1	0.68	0.92	0.78

Train Accuracy: 0.68255
Test Accuracy: 0.6776

Shows the Random Forest algorithm with precision of 68% positive and 64% negative. The recall values obtained is 92% which were positive and 26% values to be negative whereas F1-score to be &8% positive and 37% negative. The accuracy value to be 68.25% on training and 67.76% accuracy on testing.

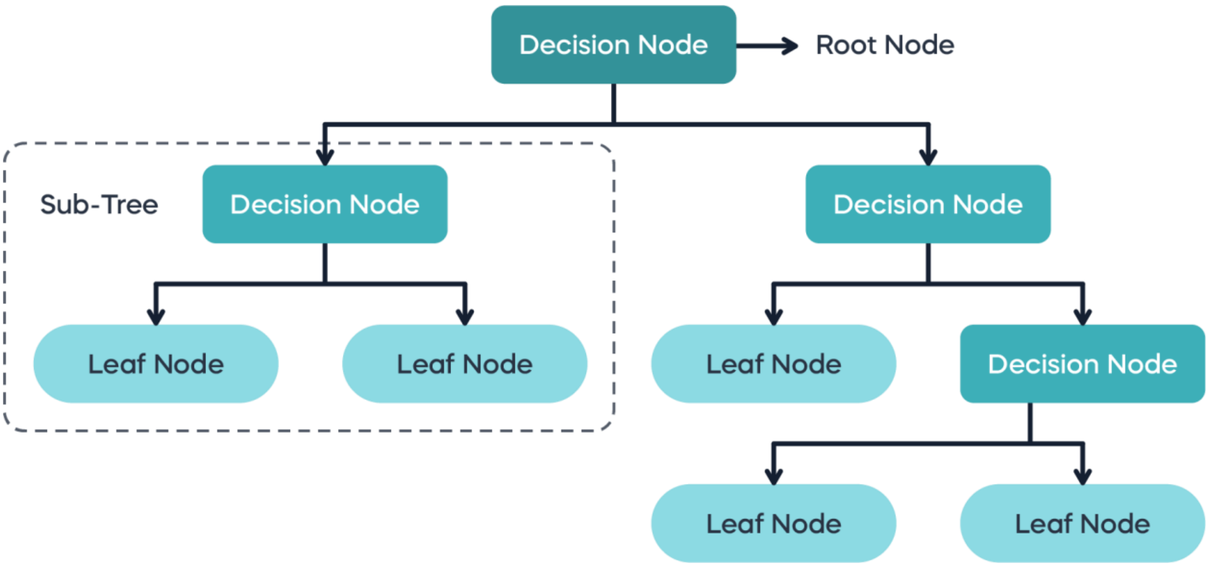
Logistic

	precision	recall	f1-score
0	0.59	0.34	0.43
1	0.70	0.87	0.77

Train Accuracy: 0.6763
Test Accuracy: 0.6756

Understanding Decision Tree

Fig:10



A Decision Tree is a supervised learning algorithm is applicable on classification and regression problems. It resembles to a tree like structure which makes its functioning easy to understand and analyse further outputs in detailed format.

A decision tree contains two elements – Nodes and Branches. The root node gives information of the splitting of data through initial feature whereas the internal nodes signify further decision on the basis of different features.

The leaf nodes are also called terminal node represent the final output.

One of the advantage of decision tree is it's capacity to function on model complex relationships due to long linear decision boundaries.They can evaluate both categorical and numerical values as a matter of fact this model acts as one of the efficient way because it works on both numerical and categorical variables without scaling or normalization.

The major drawback of decision tree is its tendency to overfitting when they are too deep and start accumulating noise in the training data. In order to cure this, pruning or setting constraints like maximum samples per leaf are commonly used.

In this study, the decision tree algorithm was applied as one of the models for classifying monkey pox cases. Its function was evaluated using standard metrics.

Application to Monkeypox Detection

In this research, the Decision tree model was used to classify the cases of monkeypox on the basis of symptoms based on demographic attributes. It’s structured made convenient for easy visualization of the feature based decision paths. Although , the Decision Tree offered quick and rapidly interpreted results but its predictive performance was found to be lower compared to random forest model. However, this model is useful in healthcare analysis due to it’s efficient tendency to interpret. In Monkeypox detection system , this model has foundational methods enabling to perform in more complex algorithms

Evaluate Decision Tree Model

	precision	recall	f1-score
0	0.59	0.32	0.41
1	0.69	0.87	0.77

Train Accuracy: 0.68695

Test Accuracy: 0.6716

Precision: precision emphasizes with proportion of correct positive predictions

Precision=True Positives | True Positives + False Positives

Recall (True Positive Rate): recall is related to the the ratio of correctly predicted positive instances out of all actual positive instances

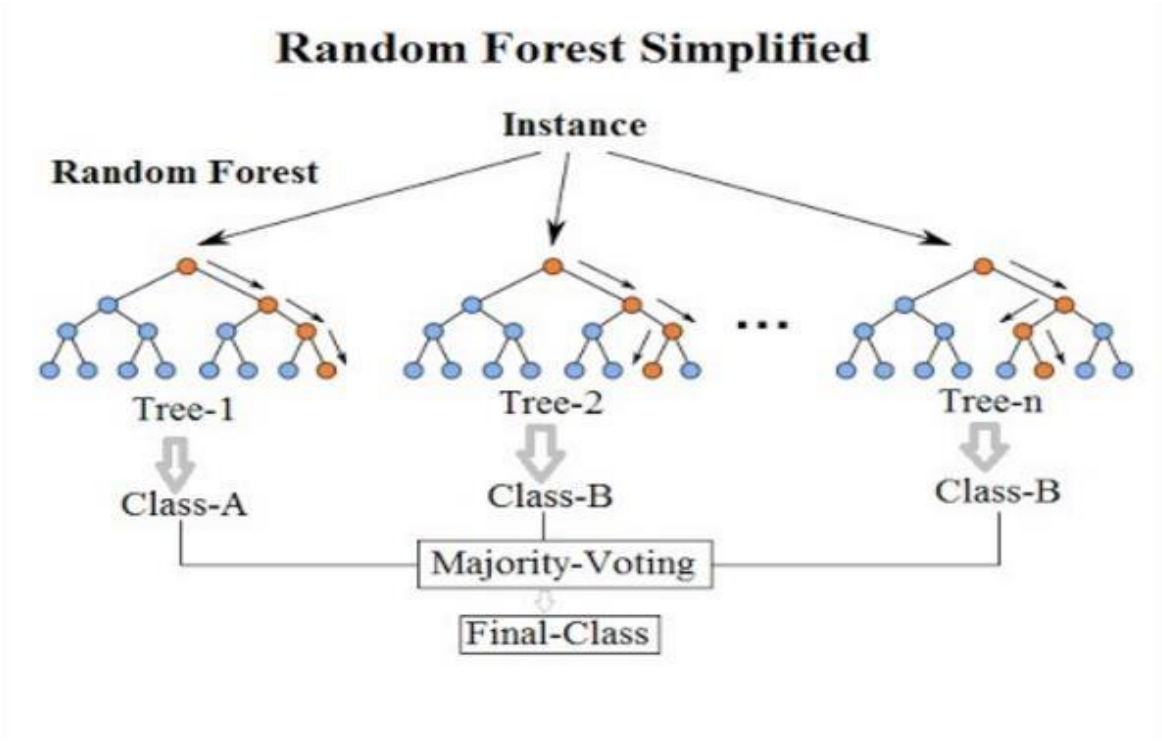
Recall=True Positives| True Positives + False Negatives

F1 Score: F1 score is the mean of Precision and Recall

F1 Score=2∗Precision∗Recall|Precision + Recall

Understanding Random Forest

Fig:11

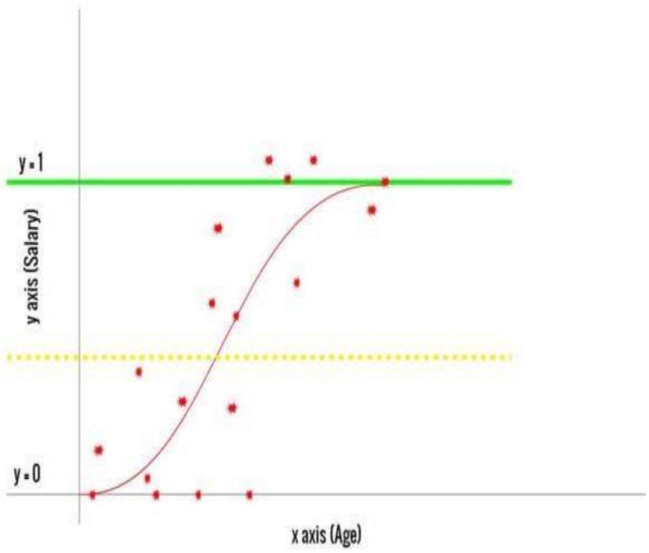


Decision tree is a supervised learning containing non-parametric algorithms which are helpful for classification and regression problems. It is basically tree based structure which resembles to be hierarchical sequence consisting of root nodes , branches, internal nodes and leaf nodes.Decision tree acts as a classifier enabling numerous decision trees. It trains each tree with different random components of the dataset and then averages of the results.

This is done to increase the accuracy in prediction. Random Forest is an ensemble-learning approach.

Logistic Regression

Fig:12



Logistic regression is a supervised machine learning algorithm used to classify various tasks where the target is about the prediction of the probability of a scenario belonging to a certain class.

. Logistic regression is a kind of an algorithm which analyses the relation between two data.

. Logistic regression determines the output of a categorical variable ensuring the outcome to be categorical.

. It gives the answer in the form of “yes or no”, “0 or 1”, “True or False” giving probabilistic values between 0 and 1

Evaluation of Logistic Regression model

	precision	recall	f1-score
0	0.59	0.34	0.43
1	0.70	0.87	0.77

Train Accuracy: 0.6763

Test Accuracy: 0.6756

Precision: deals with the positive proportion correct predictions

Precision=True Positives | True Positives + False Positives

Recall :it relates the ratio of correct predicted positive instances to the total number of actual positive instances

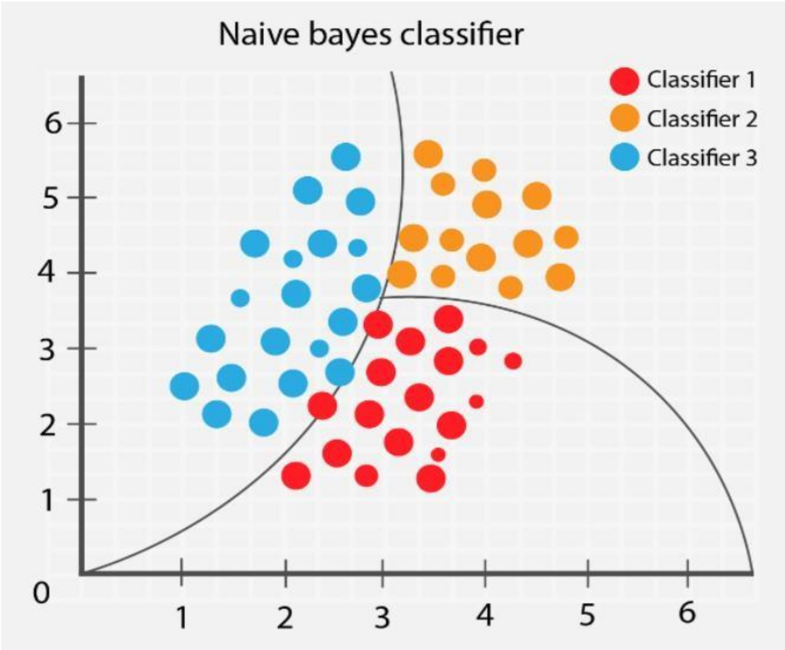
Recall=True Positives| True Positives + False Negatives

F1 Score: F1 score is the harmonic mean of precision and recall.

F1 Score=2∗Precision∗Recall|Precision + Recall

Naive Bayes

Fig:13



Naïve Bayes is a supervised machine learnin"text classification". They employ principles of probability to carry out classification tasks.

.Naive Bayes Classifier is a classifier which operates on probabilistic values to which the application becomes easy due to less parameters used to build the machine learning models which are compatible enough for making predictionsin less time as compared to other classifiers.

. It is a probabilistic classifier because it analyses about one attribute in the model is self dependent i.e independent in the presence of other models. Thus, each attribute is the summation of predictions without any connection being independent to itself.

Evaluation Naïve Bayes Model

	precision	recall	f1-score
0	0.60	0.31	0.40
1	0.69	0.88	0.77

```
Train Accuracy:  0.677
Test Accuracy:  0.6724
```

<u>Precision:</u> precision is concerned with the proportion of positive predictions that are correct.

<u>Precision</u>=True Positives | True Positives + False Positives

<u>Recall</u> (True Positive Rate): recall is concerned with the ratio of correctly predicted positive instances out of all actual positive instances.

Recall=True Positives| True Positives + False Negatives

<u>F1 Score:</u> F1 score is the harmonic mean of precision and recall.

F1 Score=2∗Precision∗Recall|Precision + Recall

Chapter: 12 Discussion

After the training and testing of each model above ,the performance is determined on the basis of interpretation using algorithms in terms of accuracy, precision, recall and F1-score. Random Forest Classifier was the best performer amongst the four models (Decision Tree,Random Forest, Logistic Regression,Naïve Bayes) as it achieved the highest score:67.76% accuracy.

The highest monkey pox positive precision result falls in the Random Forest Classifier with 67.76%. then comes the performance of Logistic Regression in the second place with value 67.56% then Naïve Bayes 67.24% in the third place. Hence, the monkey pox negative precision being the highest was evaluated to be 64% in case of Random Forest ,then comes naïve bayes with 60% and then Decision tree in third place with 59% along with logistic Regression in fourth place with the same value as 59%.The highest monkeypox positive monkeypox recall is located in random forest with value of 92% then Naïve Bayes in second place with value 88% and then third Decision tree with 87% then comes Logistic Regression in the fourth place with 87%. While the highest monkey pox recall of 34% was achieved in Logistic Regression ,32% in Decision Tree, Naïve Bayes as 40% and fourth Random Forestas 37%.

Thus, analysis of each model through training and testing was performed using machine learning algorithms when visualized from F1-score result, Random Forrest and Decision Tree algorithms ranked first

Chapter: 13 Conclusion

. In this research , our target was comparison numerous machine learning models (4 models used) in order to study their performances in studying monkey pox disease . Hence, the analysis of each model is carried out through testing these models to be utilized as a classifier.

We examine the performance of Decision Tree,Random Forest, Logistic Regression and Naïve Bayes algorithms using performance metrices that show accuracy , precision, recall, F1-score and time taken values.

Since the best accuracy was 67.76%, the research demonstrated that Random Forest performance was the best amongst all four models i.e Logistic Regression, Naïve Bayes and Decision Tree in monkey pox disease analysis.

The analysis methodology of the monkeypox disease can be a platform to assist nations and government policies particularly the health sector in order to determine the cause of monkeypox disease along with it’s prevention amongst the patients in order to detect this disease timely to avoid future inconveniences.

Thus, these models can be used for the best performance to analyze the detection of certain diseases such as monkeypox.

References

- World Health Organization (WHO). (2022). Monkey Pox Fact Sheet. Retrieved from https://www.who.int/health-topics/Monkey Pox
- Centers for Disease Control and Prevention (CDC). (2022). Monkey Pox Information. Retrieved from https://www.cdc.gov/poxvirus/Monkey Pox/index.html
- Adler, H., et al. (2022). Clinical features and management of human Monkey Pox: A retrospective observational study in the UK. The Lancet Infectious Diseases, 22(8), 1153-1162.
- European Centre for Disease Prevention and Control (ECDC). (2022). Risk assessment: Monkey Pox multi-country.
- John P. Thornhill: https://www.nejm.org/doi/full/10.1056/NEJMoa2207323
- Md Manjurul Ahsan, Muhammad Ramiz Uddin: https://arxiv.org/abs/2206.01862
- Sourav Saha, Trina Chakraborty, Rejwan Bin Sulaiman: https://arxiv.org/abs/2302.10277
- Xiangting Li, Sara Habibipour, Tom Chou, and Otto O: https://arxiv.org/abs/2308.03714
- Ajit patil & Kurle:https://www.journalofinfection.com/article/S0163-4453%2824%2900257-3/fulltext
- Alnaji https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0300216
- Kulkarni https://link.springer.com/chapter/10.1007/978-981-99-7862-5_13
- Gupta: Block chain-enabled healthcare monitoring system for early monkey pox detection.

- The Journal of Supercomputing : https://link.springer.com/article/10.1007/s11227-023-05288-y

- Kundu: https://doi.org/10.1109/ACCESS.2024.3370838
- Ren : https://doi.org/10.1016/j.bsheal.2023.05.006
- Singh: A Comparative Study of Time Series Regression and Machine Learning Models, https://doi.org/10.1007/s00284-023-03531-6:
- Islam: https://doi.org/10.1101/2022.08.08.503193
- Ahsan: Image Data collecting and using of deep learning-based model in detecting Monkeypox disease using VGG16 https://arxiv.org/abs/2206.01862

- Maqsood: https://doi.org/10.1016/j.eswa.2024.124584

  H S, M., Rallapalli https://doi.org/10.1109/EASCT59475.2023.10392701,https://doi.org/

- Chunhapran https://doi.org/10.1007/978-3-031-58561-6_14,

- Eliwa https://www.nature.com/articles/s41598-023-41545-z

- Kolluri, https://arxiv.org/abs/2209.09300
- Eid, https://www.mdpi.com/2227-7390/10/20/3845
- Figure1 is referred or taken from jupyter notebook local host.
- Figure2, Figure3, Figure4, Figure5 and Figure 6 all figures taken from local host jupyter notebook.

Presented by:
Vimal Gautam

THANK YOU