IJARSCT

ISSN: 2581-9429

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Volume 5, Issue 13, April 2025

Impact Factor: 7.67

Utilizing Polynomial Regression in Predictive Analytics for Heart Failure Mortality: A Clinical Data Perspective

Mr. Narayan Alias Advait Anant Shetkar and Mrs. **Shradha Balasaheb Linge**

Student, MIT Arts, Commerce & Science College, Pune, India[1]

Assistant Professor, MIT Arts, Commerce & Science College, Pune, India[2]

advaitshetkar7@gmail.com and sblinge@mitacsc.ac.in

Abstract: *Machine learning has become a powerful tool that provides the ability to improve predictive analytics and clinical decision making in healthcare. In this study, we investigated the use of control learning algorithms, specifically polynomial regression combined with logistic regression, to predict the probability of death in patients with heart disease. Using a cardiovascular dataset containing features such as age, blood pressure, and ejection fraction, we use polynomial feature transformations to capture complex patterns in the data. Logistic regression was then used to predict the probability of death. With an accuracy of 80%, the model performed well in predicting survival but showed moderate improvement in identifying patients at risk of death. These results indicate that further development is needed to improve the model's performance. The aim of this study was to evaluate the effectiveness of the algorithms in predicting mortality from heart failure.*

Keywords: Machine Learning, Polynomial Regression, Logistic Regression, Heart Failure.

I. INTRODUCTION

Heart failure[3] is a major health problem that causes significant morbidity and mortality worldwide. Identification of high-risk patients is important for effective management. Heart failure medical records include many medical conditions that affect outcome, such as age[5], blood pressure[6], and comorbidities. In this study, supervised learning algorithms, specifically polynomial [9,10]and logistic regression, were used to predict mortality. By analyzing these data, we aim to evaluate the effectiveness of these models in improving patient outcomes through early diagnosis

Machine learning (ML) [1]is rapidly improving healthcare by analyzing large and complex data sets. Machine learning algorithms are good at analyzing data such as medical records, patient demographics, imaging data, and biomarkers to find patterns that traditional methods lack. This information technology[2]allows scientists and

Machine learning algorithms[1,2] are roughly divided into four types: supervised learning, unsupervised learning, semisupervised learning, and additive learning. Types of supervised machine learning include logistic regression, polynomial regression[7], and others. This approach involves using domain data to train a model to make predictions. Supervised machine learning focuses on understanding the relationship between different inputs (x) and outputs (y).

Logistic regression is a statistical technique widely used in binary distribution problems. It models the relationship between a binary variable and one or more variables by estimating the probability using a logistic function. This method allows for the specification of clear boundaries by converting the output estimates into probability scores between 0 and 1. Logistic regression is valued for its simplicity, interpretability, and efficiency in obtaining categorical results. Its applications cover a wide range of areas, including medical diagnosis, financial risk assessment, and social research, making it a versatile tool for forecasting and decision making.

Polynomial regression extends linear regression by fitting polynomial terms to independent relationships and variables. Unlike simple linear regression, which models relationships in a straight line, polynomial regression [9]introduces higher-order concepts to capture complex patterns in data. This approach allows modeling of curvilinear relationships, providing a simple alternative to data that exhibits nonlinear patterns. Polynomial regression is particularly useful in situations where linear models do not adequately capture the relationships between variables. Its applications span many fields, including economics, engineering, and biology, improving predictive accuracy and data interpretation.

We will also examine the impact of our prediction model, focusing on its potential for early detection of heart failure, risk assessment, and pain management planning. We aim to improve patient care and reduce the impact of this disease by providing physicians with reliable tools to predict the outcomes of heart failure. This study demonstrates the transformative power of machine learning algorithms in predicting heart failure and highlights the need for diverse clinical data to increase accuracy and reliability. Using a data-driven approach, we aim to deepen our understanding of heart failure and develop effective strategies for its prevention, diagnosis, and management.

II. LITERATURE REVIEW

Predicting heart failure (HF) mortality reflects the limitations of statistical methods that often fail to capture multiple interactions. Recent advances use machine learning algorithms (such as incremental decision trees) to improve gambles. Studies show that machine learning models can identify important variables and produce accurate scores. The method has demonstrated superior performance in distinguishing mortality risk compared to existing risk scores and has demonstrated differentially high accuracy in patients with heart failure, providing a promising tool for risk assessment.[1]

SMAC (Social, Mobile, Analytics, Cloud) documentation demonstrates its role in supporting smart machines and enabling machine learning (ML). Research shows how machine learning enables machines to learn from experience and mimic human behavior by analyzing data. Focus on research and studies on the integration of big data and machine learning to inform business transformation. The paper also highlights technology that demonstrates the increasing use of machine learning as a key application for the future of work, particularly in decision-making automation processes.[2]

Machine learning (ML) literature in cardiovascular disease diagnosis demonstrates the potential of ML models to improve classification accuracy and reduce misdiagnosis. Studies have shown that methods such as decision trees, random forests, XGBoost, and multilayer perceptrons can predict heart disease. Techniques such as K-mode clustering and cross-validation, as well as super-tuning of GridSearchCV, can further improve model performance. This study shows that the multilayer perceptron model with cross-validation achieves the highest accuracy, demonstrating the effectiveness of deep learning in clinical applications.[3]

Cardiovascular disease prediction highlights the role of machine learning in improving survival in cardiovascular patients. Studies have shown the effectiveness of using electronic medical records to identify important predictors and risk levels. Most studies comparing machine learning to traditional biostatistical methods have found that certain features, such as serum creatinine and ejection fraction, remain significant. This article summarizes these findings, showing that simple models that use only these features can outperform models and provide a useful tool for clinical decision making.[4]

Cardiovascular disease (CVD) identifies high blood pressure (BP) as a major risk factor, and there is evidence linking high blood pressure to a variety of conditions, including heart failure, atrial fibrillation, kidney disease, and stroke. Studies have shown that changes in the blood pressure distribution toward higher levels have a significant impact on the risk of heart disease. Research supports the benefits of lowering blood pressure, and meta-analyses of controlled trials have clearly confirmed these findings. Prevention of age-related hypertension and aggressive treatment of hypertension can reduce the severity of heart disease associated with hypertension.[5]

Machine learning (ML) is evolving in a wide range of areas, including data mining, image processing, and predictive analytics. Research shows the ability of machine learning to improve performance by performing tasks through learning algorithms, as seen in search engines like Google. Research also explores the evolution of machine learning models from simple methods to complex methods using neural networks and deep learning. Future prospects focus on improving algorithm accuracy and widespread use in areas such as healthcare, finance, and self-management.[6]

Hypertension and heart failure represent the effects of long-term hypertension on cardiovascular health, particularly through processes such as left ventricular hypertrophy and diastolic dysfunction. Good blood pressure control is important to prevent the development of heart failure and its complications. However, setting a blood pressure target

IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429 Volume 5, Issue 13, April 2025 Impact Factor: 7.67

that is too low may have adverse effects, possibly as a result of the changing J-touch. Current guidelines recommend a target of approximately 130/80 mmHg, but further research is needed to validate blood pressure control in cardiac patients.[7]

Polynomial regression models have demonstrated their usefulness in modeling curvilinear relationships among variables. Research has focused on the use of polynomial regression in many areas, emphasizing its effectiveness in capturing nonlinear patterns. The least squares method is often used for parameter estimation, and standard regression metrics are often used to measure accuracy. Research has shown that tools such as MATLAB are useful in implementing and analyzing these models, making polynomial regression a reliable choice for nonlinear data in real-world applications such as social relationships in the world.[8]

Diabetes prediction highlights the role of machine learning (ML) in improving early diagnosis and improving patient outcomes. This study investigates various machine learning algorithms, such as K-nearest neighbor, logistic regression, random forest, support vector machine, and decision tree, and demonstrates their effectiveness in predicting diabetes. Studies show the importance of combining multiple algorithms to obtain more accurate predictions. This approach is important for solving the global diabetes epidemic, providing timely intervention, and reducing the risk of serious problems such as blindness, kidney failure, and heart disease.[9]

Predictive modeling in healthcare demonstrates the difficulty in improving the accuracy of machine learning (ML) algorithms, especially iterative models. Research shows that new methods such as data transfer are needed to reduce prediction errors. This study minimizes the sum of squared errors (SSE) in linear regression models by presenting a regression model. The findings showed a significant improvement in the performance and consistency of the model and were validated by statistical tests such as Wilcoxon signed rank and Cronbach alpha, showing promising utility for health screening.[10]

III. OBJECTIVE OF STUDY

- **Performance evaluation:** To evaluate the performance of polynomial regression models in predicting heart failure mortality compared with traditional linear regression models.
- **Analysing data complexity:** Investigating how variables control relationships between clinical data and their impact on the accuracy of the data.
- **Specific selection:** Use polynomial regression to determine which clinical variables (e.g., age, blood pressure[6,7], ejection fraction) affect the mortality prediction.
- **Model Performance:** Evaluate and compare performance metrics (e.g. accuracy, precision, recall, F1 score) of polynomial regression with other machine learning models in predicting death from heart failure.
- **Potential for early detection:** Determining how polynomial regression models improve early diagnosis in patients at risk of death from heart failure.
- **Risk stratification:** Evaluating the ability of polynomial regression to stratify patients into different risk categories as a predictor of mortality.
- **Effects of polynomial order:** Explore how changes in polynomial order affect the model's performance and ability to capture complex patterns in data.
- **Clinical Relevance:** Evaluate the effectiveness of using polynomial regression to predict mortality in clinical settings, including its benefits and limitations.
- **Comparison with other techniques:** Compare the results of polynomial regression with those obtained with other advanced techniques such as convolution and neural networks.
- **Recommendations for use:** Provides recommendations for integrating polynomial regression models into clinical decision-making to improve patient management and cardiac outcomes.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 13, April 2025
Impact Factor: 7.67

IV. METHODOLOGY

This study aims to use machine learning techniques, specifically polynomial [10] regression and logistic regression algorithms, to predict mortality in cardiac patients. The Heart Failure Clinical Registry Dataset contains important clinical data such as age, blood pressure[6,7], and ejection fraction and is used for modeling.

We first preprocess the data, handle important missing features, and normalize the features for better model performance. Then, we split the dataset into training and testing to ensure that the model accuracy is measured on unseen data.

Polynomial regression is then used to capture the relationship between the data. By transforming the input input into polynomial [10] terms, this method allows the model to include more patterns that may be present in the dataset. Once we identified these features, we used logistic regression to predict the binary outcome (whether the patient would survive or suffer a fatal outcome).

With this approach, we aim to develop a reliable assessment tool for heart failure outcomes that can help quickly andimprove patient care.

IMPORT REQUIRED LIBRARY

```
# Import necessary libraries
import numpy as np
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import PolynomialFeatures
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import accuracy_score, confusion_matrix, classification_report
import matplotlib.pyplot as plt
```

This code imports the core libraries for building and evaluating machine learning models:

- numpy and pandas: Used for arithmetic calculations and data structures such as arrays and data frames.
- train_test_split: Split the dataset into training set and test set.
- PolynomialFeatures: Use polynomial transformations for features to capture nonlinear relationships.
- LogisticRegression: Use logistic regression to classify activities.
- accuracy_score, confusion_matrix, classification_report: Metrics to evaluate model performance.
- matplotlib. pyplot: Used to organize and visualize data and results.

LOAD THE DATASET INFORMATION OF THE DATA SET

Our data set contains 299 rows and 13 columns in which age, anaemia,creatinine_phosphokinase, diabetes[8,14],ejection_fraction,high_blood_pressure, platelets, serum_creatinine, serum_sodium, sex, smoking, time, DEATH_EVENT are features

```
# Load the dataset
dataset = pd.read_csv('Heart_failure_clinical_records_dataset.csv')
```

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Volume 5, Issue 13, April 2025
ISSN: 2581-9429
Impact Factor: 7.67

```
dataset
```

	age	anaemia	creatinine_phosphokinase	diabetes	ejection_fraction	high_blood_pressure	platelets	serum_creatinine	serum_sodium	sex	smoking	time	DEATH_EVENT
0	75.0	0	582	0	20	1	265000.00	1.9	130	1	0	4	1
1	55.0	0	7861	0	38	0	263358.03	1.1	136	1	0	6	1
2	65.0	0	146	0	20	0	162000.00	1.3	129	1	1	7	1
3	50.0	1	111	0	20	0	210000.00	1.9	137	1	0	7	1
4	65.0	1	160	1	20	0	327000.00	2.7	116	0	0	8	1
...	...	...	...	...	...	...	...	...	...	...	...	...	...
294	62.0	0	61	1	38	1	155000.00	1.1	143	1	1	270	0
295	55.0	0	1820	0	38	0	270000.00	1.2	139	0	0	271	0
296	45.0	0	2060	1	60	0	742000.00	0.8	138	0	0	278	0
297	45.0	0	2413	0	38	0	140000.00	1.4	140	1	1	280	0
298	50.0	0	196	0	45	0	395000.00	1.6	136	1	1	285	0

299 rows × 13 columns

DATASET HEAD

```
dataset.head()
```

	age	anaemia	creatinine_phosphokinase	diabetes	ejection_fraction	high_blood_pressure	platelets	serum_creatinine	serum_sodium	sex	smoking	time	DEATH_EVENT
0	75.0	0	582	0	20	1	265000.00	1.9	130	1	0	4	1
1	55.0	0	7861	0	38	0	263358.03	1.1	136	1	0	6	1
2	65.0	0	146	0	20	0	162000.00	1.3	129	1	1	7	1
3	50.0	1	111	0	20	0	210000.00	1.9	137	1	0	7	1
4	65.0	1	160	1	20	0	327000.00	2.7	116	0	0	8	1

DESCRIBE THE DATASET

We calculate the data of each parameter such as mean, standard deviation, minimum, maximum, quartiles.

```
dataset.cescribe()
```

	age	anaemia	creatinine_phosphokinase	diabetes	ejection_fraction	high_blood_pressure	platelets	serum_creatinine	serum_sodium	sex	smoking	time	DEATH_EVENT
count	299.000000	299.000000	299.000000	299.000000	299.000000	299.000000	299.000000	299.00000	299.000000	299.000000	299.00000	299.000000	299.00000
mean	60.837237	0.431438	581.839465	0.418060	38.083612	0.351171	263358.029264	1.39388	136.625418	0.648829	0.32107	130.260870	0.32107
std	11.900919	0.496107	970.287881	0.494067	11.834841	0.478136	97804.236869	1.03451	4.412477	0.478136	0.46767	77.614208	0.46767
min	40.000000	0.000000	23.000000	0.000000	14.000000	0.000000	25100.000000	0.50000	113.000000	0.000000	0.00000	4.000000	0.00000
25%	51.000000	0.000000	116.500000	0.000000	30.000000	0.000000	212500.000000	0.90000	134.000000	0.000000	0.00000	73.000000	0.00000
50%	60.000000	0.000000	250.000000	0.000000	38.000000	0.000000	262000.000000	1.10000	137.000000	1.000000	0.00000	115.000000	0.00000
75%	70.000000	1.000000	582.000000	1.000000	45.000000	1.000000	303500.000000	1.40000	140.000000	1.000000	1.00000	203.000000	1.00000
max	95.000000	1.000000	7861.000000	1.000000	80.000000	1.000000	850000.000000	9.40000	148.000000	1.000000	1.00000	285.000000	1.00000

DOI: 10.48175/IJARSCT-26033

INFORMATION OF THE DATA SET

```
dataset.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 299 entries, 0 to 298
Data columns (total 13 columns):
 #   Column                    Non-Null Count  Dtype
---  ------                    --------------  -----
 0   age                       299 non-null    float64
 1   anaemia                   299 non-null    int64
 2   creatinine_phosphokinase  299 non-null    int64
 3   diabetes                  299 non-null    int64
 4   ejection_fraction         299 non-null    int64
 5   high_blood_pressure       299 non-null    int64
 6   platelets                 299 non-null    float64
 7   serum_creatinine          299 non-null    float64
 8   serum_sodium              299 non-null    int64
 9   sex                       299 non-null    int64
 10  smoking                   299 non-null    int64
 11  time                      299 non-null    int64
 12  DEATH_EVENT               299 non-null    int64
dtypes: float64(3), int64(10)
memory usage: 30.5 KB
```

DEFINE FEATURE VARIABLES (X) AND TARGET VARIABLE (y)

```
# Define feature variables (X) and target variable (y)
X = dataset.drop(columns=['DEATH_EVENT']).values  # All features except the target
y = dataset['DEATH_EVENT'].values  # The target variable
```

Here DEATH_EVENT is dependent variable and all other variables like age, anaemia, creatinine_phosphokinase, diabetes , ejection_fraction, high_blood_pressure, platelets, serum_creatinine, serum_sodium, sex, smoking, time, are independent variable

SPLIT THE DATA INTO TRAINING AND TESTING SETS

```
# Split the data into training and test sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=0)
```

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 13, April 2025
Impact Factor: 7.67

The rule splits the dataset into two parts: training and testing procedures. Specifically, 80% of the data (X_train and y_train) is used to train the model, and 20% (X_test and y_test) is reserved for testing its performance. random_state=0 ensures that the split is repeated, meaning that running the code multiple times will produce the same result.

APPLY PLOYNOMIAL TRANSFORMATION

Rule for using 3rd degree polynomial transformation for training and test data:

poly_reg = PolynomialFeatures(degree=3): Create a polynomial generator that transforms data by adding polynomial elements up to degree 3.

```
# Apply polynomial transformation (degree=3 for example)
poly_reg = PolynomialFeatures(degree=3)
X_poly_train = poly_reg.fit_transform(X_train)
X_poly_test = poly_reg.transform(X_test)
```

X_poly_train=poly_reg.fit_transform(X_train): Transform training data by fitting and using polynomial transformation to create new features.

X_poly_test = poly_reg.transform(X_test): Use the same polynomial transformation to transform the test data without transforming it.

FIT A LOGISTIC REGRESSION MODEL TO THE TRANSFORMED FEATURE

```
# Fit a logistic regression model to the transformed features
log_reg = LogisticRegression(max_iter=10000)
log_reg.fit(X_poly_train, y_train)
```

```
LogisticRegression
LogisticRegression(max_iter=10000)
```

These rules fit the logistic regression model for polynomially transformed features:

log_reg=LogisticRegression(max_iter=10000): The logistic regression model was built for up to 10,000 iterations to ensure convergence.

log_reg.fit(X_poly_train, y_train): It shows the logistic regression model of polynomial transformation of training data (X_poly_train) and corresponding data (y_train).

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 13, April 2025 Impact Factor: 7.67

MAKE PREDICTIONS ON THE TEST SET

```
# Make predictions on the test set
y_pred = log_reg.predict(X_poly_test)
```

This rule uses the logistic regression model to make predictions:
y_pred = log_reg.predict(X_poly_test): A logistic regression model was trained (log_reg) using the multivariate regression data (X_poly_test) to predict the output. These predictions are stored in y_pred.

EVALUATE THE MODEL

```
# Evaluate the model
accuracy = accuracy_score(y_test, y_pred)
conf_matrix = confusion_matrix(y_test, y_pred)
class_report = classification_report(y_test, y_pred)

print("Accuracy:", accuracy)
print("Confusion Matrix:\n", conf_matrix)
print("Classification Report:\n", class_report)
```

This code uses several metrics to measure model performance:
accuracy = accuracy_score(y_test, y_pred): The accuracy of the model is calculated by comparing the true text (y_test) with the predicted text (y_pred).
conf_matrix = confusion_matrix(y_test, y_pred): Create a confusion matrix showing the number of positives, negatives, negatives, and negatives.
class_report = classification_report(y_test, y_pred): Provides detailed information for each category, including accuracy, recall, and F1 score.

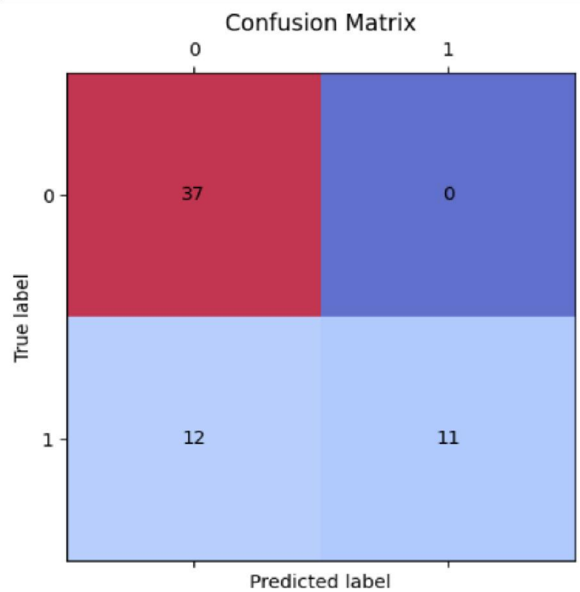
VISUALIZATION OF CONFUSION MATRIX

```
# Plot the confusion matrix
plt.matshow(conf_matrix, cmap='coolwarm', alpha=0.8)
for i in range(len(conf_matrix)):
    for j in range(len(conf_matrix[i])):
        plt.text(x=j, y=i, s=conf_matrix[i, j], va='center', ha='center')
plt.xlabel('Predicted label')
plt.ylabel('True label')
plt.title('Confusion Matrix')
plt.show()
```

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 13, April 2025
Impact Factor: 7.67

The confusion matrix provides more detailed information about the model's performance by dividing the predictions into four categories:

- True Positives (TP = 11): The model correctly predicted 11 patients who died (class 1).
- True Negatives (TN = 37): The model correctly predicted 37 patients who survived (class 0).
- False Positives (FP = 0): The model did not falsely predict any patients as having died when they actually survived.
- False Negatives (FN = 12): The model incorrectly predicted 12 patients as having survived when they actually died.



The model performed well in accurately predicting the number of patients who survived (grade 0). However, it is difficult to identify all patients who died (category 1). It was not possible to identify 12 patients who actually died, which is important for treatment

V. RESULT

```
Accuracy: 0.8
Confusion Matrix:
 [[37  0]
 [12 11]]
Classification Report:
               precision    recall  f1-score   support

           0       0.76      1.00      0.86        37
           1       1.00      0.48      0.65        23

    accuracy                           0.80        60
   macro avg       0.88      0.74      0.75        60
weighted avg       0.85      0.80      0.78        60
```


DOI: 10.48175/IJARSCT-26033

Accuracy: 0.80 (80%)
Meaning: Accuracy is the percentage of correct predictions (true positives and negatives) for all events. This means that 80% of the predictions made by the model are correct. While this may seem like a good result, facts alone do not tell the full story, especially in the medical field, where the consequences of incorrect predictions can be very serious.

VI. CONCLUSION

In this project, we successfully used polynomial regression to classify heart failure clinical data and predict patient mortality. The model is able to capture relationships in the data with 80% accuracy. The model predicts whether the patient will survive by predicting the outcome as 0 or 1. The model demonstrates strong predictive power with 80% accuracy, while it has an expected error rate of 20% when using machine learning models. Further refinement can help improve its accuracy, allowing for more informed clinical decisions

REFERENCES

[1]. Adler, E. D., Voors, A. A., Klein, L., Macheret, F., Braun, O. O., Urey, M. A., Zhu, W., Sama, I., Tadel, M., Campagnari, C., Greenberg, B., & Yagil, A. (2020). Improving risk prediction in heart failure using machine learning. *European Journal of Heart Failure*, *22*(1), 139–147. https://doi.org/10.1002/ejhf.1628

[2]. Alzubi, J., Nayyar, A., & Kumar, A. (2018). Machine Learning from Theory to Algorithms: An Overview. *Journal of Physics: Conference Series*, *1142*(1). https://doi.org/10.1088/1742-6596/1142/1/012012

[3]. Bhatt, C. M., Patel, P., Ghetia, T., & Mazzeo, P. L. (2023). Effective Heart Disease Prediction Using Machine Learning Techniques. *Algorithms*, *16*(2). https://doi.org/10.3390/a16020088

[4]. Chicco, D., & Jurman, G. (2020). Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone. *BMC Medical Informatics and Decision Making*, *20*(1). https://doi.org/10.1186/s12911-020-1023-5

[5]. Fuchs, F. D., & Whelton, P. K. (2020). High Blood Pressure and Cardiovascular Disease. In *Hypertension* (Vol. 75, Issue 2, pp. 285–292). Lippincott Williams and Wilkins. https://doi.org/10.1161/HYPERTENSIONAHA.119.14240

[6]. Mahesh, B. (2020). Machine Learning Algorithms - A Review. *International Journal of Science and Research (IJSR)*, *9*(1), 381–386. https://doi.org/10.21275/art20203995

[7]. Oh, G. C., & Cho, H. J. (2020). Blood pressure and heart failure. In *Clinical Hypertension* (Vol. 26, Issue 1). BioMed Central Ltd. https://doi.org/10.1186/s40885-019-0132-x

[8]. Ostertagová, E. (2012). Modelling using polynomial regression. *Procedia Engineering*, *48*, 500–506. https://doi.org/10.1016/j.proeng.2012.09.545

[9]. Rani, K. J. (2020). Diabetes Prediction Using Machine Learning. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 294–305. https://doi.org/10.32628/CSEIT206463

[10]. Srilakshmi, U., Manikandan, J., Velagapudi, T., Abhinav, G., Kumar, T., & Saideep, D. (2024). A New Approach to Computationally-Successful Linear and Polynomial Regression Analytics of Large Data in Medicine. In *Journal of Computer Allied Intelligence* (Vol. 02, Issue 02).