Title  
Author: Farjana Bristy (your email)  
Affiliation: Independent Researcher (or “Independent”)  
Date: [DD Month 2025]

**Abstract (100–150 words)**  
We apply three machine learning models — Logistic Regression, Random Forest, and XGBoost — to the UCI Heart Disease dataset to predict the presence of heart disease. We compare performance using accuracy, precision, recall, F1-score, and AUC. Our best model achieved an AUC of [AUC], indicating [brief interpretation]. The study highlights [one or two key findings: e.g., most important predictors] and discusses clinical relevance and limitations.

**1. Introduction (0.5 page)**

* Motivation (importance of early prediction)
* Brief literature mention (1–2 sentences referencing common works)
* Paper contributions (bullet: quick, reproducible comparison; clear pipeline; feature importance)

**2. Dataset (0.5 page)**

* Describe source (UCI / Kaggle Heart Disease dataset)
* Number of instances and features
* Preprocessing steps (missing values, scaling)

**3. Methods (1 page)**

* Models used (hyperparameters)
* Training/test split (80/20), cross-validation if used
* Metrics (accuracy, precision, recall, F1, AUC)
* Reproducibility note (random seed, software versions)

**4. Results (1 page)**

* Table of quantitative results (copy from notebook)
* Confusion matrix and ROC for best model (embed figures)
* Feature importance chart and short interpretation

**5. Discussion (0.5 page)**

* What the results mean clinically / academically
* Comparison to expected performance in literature (brief)
* Strengths: reproducibility, clarity
* Limitations: small dataset, possible overfitting, no external validation

**6. Conclusion (0.25 page)**

* One-paragraph summary and next steps (external validation, more features, deep learning on larger datasets)

**References (0.25 page)**

* UCI Heart Disease dataset citation
* One or two seminal ML-in-healthcare references

**Appendix (optional)**

* Short code snippet or link to GitHub/Colab