model_metrics.csv

```
Model,Type,Hyperparameters,MAE,RMSE,R2
Linear Regression,Regression,"-",845.6,1024.3,0.73
Random Forest,Regression,"n_estimators=200,max_depth=20,min_samples_split=2",
622.4,785.9,0.87
XGBoost,Regression,"n_estimators=200,learning_rate=0.1",641.2,811.5,0.85
```

README.md

🏥 Healthcare Billing Prediction Project

Objective:
Predict hospital billing amounts using patient demographics and admission data.

Project Structure

- `healthcare_modeling.ipynb` – main code
- `healthcare_preprocessed.csv` – cleaned dataset
- `model_metrics.csv` – final model results
- `Capstone_Final_Report.pdf` – summary report

Key Insights

- Random Forest achieved the highest accuracy (R^2 = 0.87)
- Admission type and medical condition are top drivers of cost
- Model supports automated billing estimation for hospitals

Tech Stack

Python | pandas | scikit-learn | XGBoost | Matplotlib | Jupyter

Dataset

Kaggle Healthcare Dataset by Prasad Patil

🏫 Model Summary

Model	MAE	RMSE	R^2
Linear Regression	845.6	1024.3	0.73

Model	MAE	RMSE	R^2
Random Forest	622.4	785.9	0.87
XGBoost	641.2	811.5	0.85
