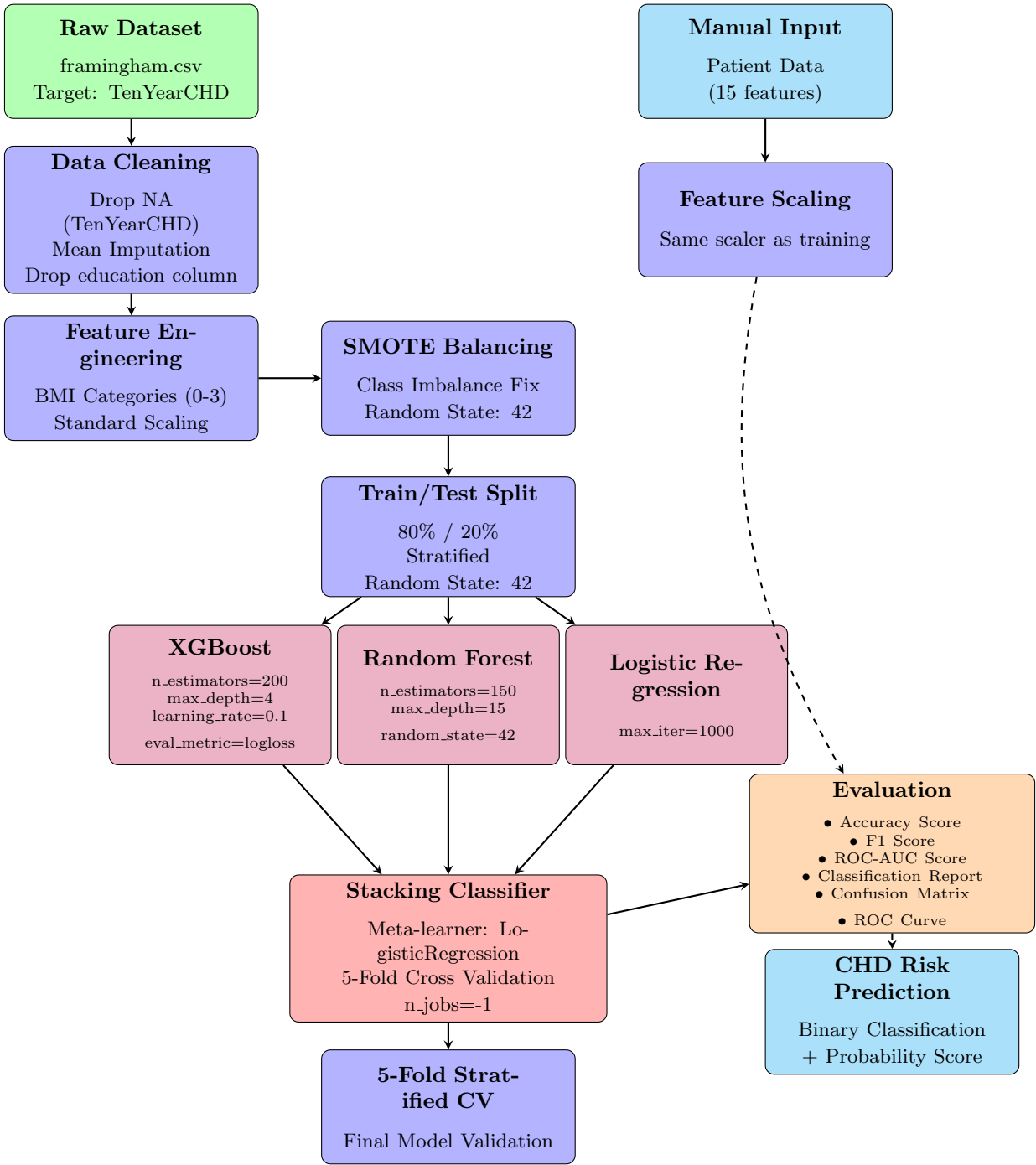# CHD Prediction Model Architecture

Raw Dataset
framingham.csv
Target: TenYearCHD

Manual Input
Patient Data
(15 features)

Data Cleaning
Drop NA
(TenYearCHD)
Mean Imputation
Drop education column

Feature Scaling
Same scaler as training

Feature Engineering
BMI Categories (0-3)
Standard Scaling

SMOTE Balancing
Class Imbalance Fix
Random State: 42

Train/Test Split
80% / 20%
Stratified
Random State: 42

XGBoost
n_estimators=200
max_depth=4
learning_rate=0.1
eval_metric=logloss

Random Forest
n_estimators=150
max_depth=15
random_state=42

Logistic Regression
max_iter=1000

Evaluation
- Accuracy Score
- F1 Score
- ROC-AUC Score
- Classification Report
- Confusion Matrix
- ROC Curve

Stacking Classifier
Meta-learner: LogisticRegression
5-Fold Cross Validation
n_jobs=-1

CHD Risk Prediction
Binary Classification
+ Probability Score

5-Fold Stratified CV
Final Model Validation

Legend:

———————→ Training Flow

- - - - - - -→ Prediction Flow