

## Model Development Phase Template

Date	29 June 2025
Team ID	SWTID1749634408
Project Title	Early Prediction for Chronic Kidney Disease Detection: A Progressive Approach to Health Management
Maximum Marks	6 Marks

### **Model Selection Report:**

In the forthcoming Model Selection Report, various models will be outlined, detailing their descriptions, hyperparameters, and performance metrics, including Accuracy or F1 Score. This comprehensive report will provide insights into the chosen models and their effectiveness.

Model	Description	Performance Metric (e.g., Accuracy)
Random Forest	Ensemble of decision trees; robust, handles complex relationships, reduces overfitting, and provides feature importance for loan approval prediction.	Accuracy score = 98.75%
Decision Tree	Simple tree structure; interpretable, captures non-linear relationships, suitable for initial insights into loan approval patterns.	Accuracy score = 98.75%
KNN	Classifies based on nearest neighbors; adapts well to data patterns, effective for local variations in loan approval criteria.	Accuracy score = 98.75%

Gradient Boosting	Gradient boosting with trees; optimizes predictive performance, handles complex relationships, and is suitable for accurate loan approval predictions.	Accuracy score = 98.75%
XGBoost Classifier	An optimized gradient boosting algorithm; highly efficient and scalable, suitable for medical prediction tasks like CKD risk detection.	Accuracy score = 98.75%
ADA Boost Classifier	Boosts weak learners to form a strong classifier; effective in improving accuracy for imbalanced datasets common in CKD detection.	Accuracy score = 98.75%
Stochastic Gradient Boosting	Incorporates randomness into gradient boosting to reduce overfitting and enhance generalization, suitable for CKD data complexity.	Accuracy score = 98.75%
CAT Boost Classifier	Efficient gradient boosting on categorical features; handles missing values well and is suited for CKD datasets with varied attributes.	Accuracy score = 98.75%
Extra Trees Classifier	An ensemble of randomized decision trees; improves variance reduction and handles high-dimensional medical data effectively.	Accuracy score = 100%