

Model Development Phase Template

Date	12 July 2024
Team ID	SWTID1720092248
Project Title	Revolutionizing Liver Care: Predicting Liver Cirrhosis Using Advanced Machine Learning Techniques
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 Selection Report:

Model	Description	Hyperparameters	Performance Metric (e.g., Accuracy, F1 Score)
logistic regression	A basic linear model that uses the logistic function to model the probability of the binary outcomes. It is simple, interpretable, and works well for linearly separable data.	C, solver	Accuracy : 0.996606 f1_score: 0.914286

logistic regression CV	An extension of logistic regression that performs cross-validation to find the best regularization parameter, which helps in avoiding overfitting and improving model performance.	Cs, cv, solver	Accuracy : 0.996606 f1_score: 0.914286
XGBoost	An advanced implementation of gradient boosting that provides parallel tree boosting which is fast, accurate, and widely used in machine learning competitions. It handles missing values and performs well with both structured and unstructured data.	n_estimators, learning_rate, max_depth	Accuracy : 0.997738 f1_score: 0.941176
Ridge classifier	A linear classifier that uses ridge regression for training, adding L2 regularization to the logistic regression, which helps in handling multicollinearity	alpha	Accuracy : 0.977376 f1_score: 0.642857

	and preventing overfitting.		
KNN	A non-parametric, instance-based learning algorithm that classifies a data point based on how its neighbors are classified. It is simple and effective but can be computationally expensive.	n_neighbors	Accuracy : 0.935520 f1_score: 0.387097
Random Forest	An ensemble learning method that constructs multiple decision trees during training and outputs the mode of the classes as the prediction. It reduces overfitting and improves accuracy.	n_estimators, max_depth	Accuracy : 1.000000 f1_score: 1.000000
Support Vector Classifier	A powerful classification method that finds the hyperplane that best separates the classes in the feature space. It	C, kernel	Accuracy : 0.997738 f1_score: 0.941176

	works well for high-dimensional data and can handle non-linear relationships using kernel trick.		
--	--	--	--