



INDIVIDUAL ASSIGNMENT

TECHNOLOGY PARK MALAYSIA

CT046-3-M-AML

APPLIED MACHINE LEARNING

APDMF2112DSBA(DE)(PR)

APRIL 2022

**TITLE: E-COMMERCE PACKAGE DELIVERY TIME
PREDICTION USING MACHINE LEARNING**

LEE KEAN LIM

TP065778

LECTURER: PROF. DR. MANDAVA RAJESWARI

ABSTRACT

Proportional to the rapid growth of e-commerce industry is the package delivery services. By providing accurate package delivery time, operation efficiency enhances and customer retention rate increases. The adoption of machine learning methods to predict the package delivery time has garnered popularity due to the wide applicability of machine learning and the accuracies of the predictions. Common issue with machine learning in this domain is feature selection as variety of data can be collected from each leg of the supply chain. This study aimed to develop prediction models using machine learning algorithms to predict the package delivery time. An e-commerce shipping dataset from Kaggle consisting of customer and package shipment information is utilized for this prediction task. The prediction task is a binary classifier to predict whether the package is delivered on time or not on time. The prediction models are developed using Logistic Regression, XGBoost, AdaBoost, and Random Forest algorithms. Furthermore, each algorithm is used to build two models namely the basic model with hyperparameters as default setting values and an optimized model with hyperparameter optimization using grid search over a wide range of hyperparameter combinations. Among the four developed models, the optimized AdaBoost model obtained the best performance with an accuracy of 69.82% and outperformed the literature that uses the same dataset.

TABLE OF CONTENTS

ABSTRACT	i
TABLE OF CONTENTS	ii
LIST OF TABLES	v
LIST OF FIGURES	vi
LIST OF ABBREVIATIONS	vii
SECTION 1: INTRODUCTION	1
1.1 INTRODUCTION	1
1.2 RESEARCH QUESTION	3
1.3 AIM & OBJECTIVES	3
1.3.1 Aim	3
1.3.2 Objectives	3
SECTION 2: LITERATURE REVIEW	4
2.1 LITERATURE REVIEW	4
2.2 DISCUSSION	9
SECTION 3: METHODS	11
3.1 DATASET	11
3.2 MACHINE LEARNING ALGORITHMS	12
3.2.1 Logistic Regression	12
3.2.2 Extreme Gradient Boosting (XGBoost)	12
3.2.3 Random Forest (RF)	13
3.2.4 Adaptive Boosting (AdaBoost)	13
3.3 EVALUATION METRIC	14
3.3.1 Confusion Matrix	14

3.3.2	Accuracy	15
3.3.3	Specificity	15
3.3.4	Sensitivity	15
3.3.5	Receiver Operating Characteristics (ROC)	16
3.3.6	Kappa	16
3.4	R PACKAGES	17
SECTION 4: DATASET PREPARATION		18
4.1	EXPLORATORY DATA ANALYSIS	18
4.1.1	Dataset Initial View	18
4.1.2	Data Distribution of Categorical Variables	20
4.1.3	Data Distribution of Continuous Variables	22
4.1.4	Correlation Between Features	24
4.1.5	Identifying Missing Value	25
4.2	DATA PRE-PROCESSING	26
4.2.1	Dropping and Recoding Features	26
4.2.2	Factoring Categorical Variables	27
4.2.3	Missing Value Imputation	27
4.2.4	Feature Encoding	28
4.2.5	Feature scaling	29
4.2.6	Train and Test Data Split	30
SECTION 5: MODEL IMPLEMENTATION		31
5.1	LOGISTIC REGRESSION	31
5.1.1	Logistic Regression with Cross Validation	31
5.1.2	Logistic Regression Model Performance	32
5.2	EXTREME GRADIENT BOOSTING (XGBoost)	33

5.2.1	Basic XGBoost model	33
5.2.2	Optimized XGBoost model	34
5.2.3	XGBoost Model Performance	35
5.3	ADAPTIVE BOOSTING (AdaBoost)	38
5.3.1	Basic AdaBoost Model	38
5.3.2	Optimized AdaBoost Model	39
5.3.3	AdaBoost Model Performance	40
5.4	RANDOM FOREST	43
5.4.1	Basic Random Forest Model	43
5.4.2	Optimized Random Forest Model	44
5.4.3	Random Forest Model Performance	44
SECTION 6: ANALYSIS & RECOMMENDATIONS		48
6.1	RESULT ANALYSIS	48
6.2	RESULTS COMPARISON	49
SECTION 7: CONCLUSION		51
REFERENCES		52
APPENDIX A: MODEL OPTIMIZATION RESULTS		53

LIST OF TABLES

Table 2.1: Literature review matrix	4
Table 3.1: Dataset variables description	11
Table 3.2: Dataset variables data type	12
Table 3.3: Cohen's kappa value interpretation	17
Table 3.4: R packages used	17
Table 4.1: Frequency table of warehouse block after recoding	26
Table 4.2: Recoding for product importance	26
Table 4.3: Class frequency of train and test set	30
Table 5.1: Confusion matrix for logistic regression model	32
Table 5.2: Evaluation metric results for logistic regression model	32
Table 5.3: Confusion matrix for XGBoost models	36
Table 5.4: Evaluation metric results for XGBoost models	36
Table 5.5: Confusion matrix for AdaBoost models	40
Table 5.6: Evaluation metric results for AdaBoost models	40
Table 5.7: Confusion matrix for RF models	45
Table 5.8: Evaluation metric results for Random Forest models	45
Table 6.1: Evaluation results compilation of predictive models	48
Table 6.2: Model performance of peers from the same dataset	50

LIST OF FIGURES

Figure 3.1: Example confusion matrix (Chauhan, 2020)	14
Figure 3.2: Example ROC curve (Chauhan, 2020)	16
Figure 4.1: Glimpse of the first few observations of the dataset	18
Figure 4.2: Internal structure of the dataset	19
Figure 4.3: Data distribution of every categorical variable in the dataset	20
Figure 4.4: Data distribution of every numerical variable in the dataset	22
Figure 4.5: Correlation between features	24
Figure 4.6: Missing values proportion	25
Figure 4.7: Internal structure of dataset after dropping	26
Figure 4.8: Internal structure of dataset after factor	27
Figure 4.9: Missing value imputation result	28
Figure 4.10: Internal structure after feature encoding	29
Figure 4.11: Summary statistics of features after feature scaling	30
Figure 5.1: Logistic regression model with cross validation snippet	31
Figure 5.2: ROC curve for basic logistic regression	33
Figure 5.3: Snippet of basic XGBoost model building and fitting	34
Figure 5.4: Snippet of optimized XGBoost model building and fitting	35
Figure 5.5: ROC curve for basic XGBoost	37
Figure 5.6: ROC curve for optimized XGBoost	38
Figure 5.7: Snippet of basic AdaBoost model building and fitting	39
Figure 5.8: Snippet of optimized AdaBoost model building and fitting	40
Figure 5.9: ROC curve for basic AdaBoost	42
Figure 5.10: ROC curve for optimized AdaBoost	43
Figure 5.11: Snippet of basic Random Forest model building and fitting	44
Figure 5.12: Snippet of optimized Random Forest model building and fitting	44
Figure 5.13: ROC curve for the basic Random Forest	46
Figure 5.14: ROC curve for the optimized Random Forest	47

LIST OF ABBREVIATIONS

AdaBoost.....	Adaptive Boosting
ANN.....	Artificial Neural Network
CatBoost.....	Categories Boosting
CV	Cross Validation
XGBoost	Extreme Gradient Boosting
GRU	Gated Recurrent Unit
KNN.....	K-Nearest Neighbor
LightGBM.....	Light Gradient Boosting Machine
LSTM.....	Long Short-Term Memory
MAE.....	Mean Absolute Error
RF.....	Random Forest
ROC	Receiver Operating Characteristic
RMSE.....	Root Mean Square Error
SVM.....	Support Vector Machine
SVR.....	Support Vector Regression

SECTION 1

INTRODUCTION

1.1 INTRODUCTION

The e-commerce has seen an exponential growth in recent years due to the movement restriction order caused by the COVID-19 pandemic which pushed people to commerce online more than ever. In which the increased in e-commerce activities led to the increased demand of package delivery services. The package delivery time plays an important role in maintaining supply chain efficiency and customer retention (Khiari & Olaverri-Monreal, 2020). Therefore, accurate delivery time allows the optimization of delivery routes and provides information to customers to allow time coordination to ensure high successful delivery rate. In addition, providing accurate delivery time increases the customer experience and confidence towards the delivery company and seller (Hildebrandt & Ulmer, 2021). Since e-commerce lacked the physical component of being able to examine the product upfront, customers associate brands and products with the quality of service the sellers provide which one of the indicator is on time product delivery (Magiya, 2020).

The prediction of package delivery time is a very challenging task due to the multitude of uncertainties which involve in every leg of the supply chain. As the world is getting more interconnected, products are being sold and bought from different parts of the world. In which this would involve the use of different transportation mode to ship packages, which each of them carries their own unique challenges. The uncertainties can include traffic conditions, weather conditions, etc. In addition, peak seasons due to holidays or weekends may contribute to increased uncertainties as delivery orders surge. Sundaram and Tallamraju (2021) mentioned that providing information to customers about the estimated delivery date is a double-edged sword as the actual delivery can be earlier or later even by just one day from the estimated date would lead to a bad customer experience and would likely result in a customer churn.

Due to the multitude of uncertainties involved, the utilization of machine learning and deep learning approaches are adopted in predicting the package delivery time as claimed by their capabilities of being able to solve complex and non-linear problems. In particular, the utilization

of boosting-based algorithms has garnered popularity in predicting the delivery time due to the fast computation and high accuracy (Magiya, 2020; Sundaram & Tallamraju, 2021). Examples of widely used boosting-based algorithms include Light Gradient Boosting Machine (LightGBM), XGBoost, Adaptive Boosting (AdaBoost), etc. Which in the industry, accurate delivery time information can provide benefits of optimizing operation flow and resource planning and management. However, to develop an accurate delivery time prediction model, it requires variety of data which covers the trip information, operational information, and customer information (Wu & Wu, 2019). Therefore, such volume and variety of data requires high computational power and time which delays the information segregation. Other types of algorithms are also used in delivery time prediction which include Artificial Neural Networks (ANN), Long-Short Term Memory (LSTM), Random Forest (RF), Support Vector Regression (SVR), Naïve Bayes, Logistic Regression, etc.

The common issue with producing an accurate delivery time prediction model is the feature selection phase (Servos *et al.*, 2019). Voluminous data is generated from each phase of the package delivery which the package traverse through multiple transition hubs before finally arriving to the last-mile hub for the final delivery. This would include multiple mode of transportation for the package such as using ships, planes, and road vehicles, where each of them tracks a different set of data and information. Therefore, large number of features are generated from a single trip of package delivery. Feature selection is utmost important to identify significant features that provide the most impact to the development of the model while eliminating the least significant features to optimize the model by reducing complexity and reducing the processing time.

This study investigates an e-commerce dataset consisting of customer and shipment data extracted from Kaggle to predict the last-mile package delivery time. The delivery time prediction is a classification problem with a binary output of either package has been delivered on time or not on time. The prediction algorithms utilized in this study includes Logistic Regression, XGBoost, AdaBoost, and Random Forest (RF).

1.2 RESEARCH QUESTION

For the purpose of this project, the following questions will be addressed:

1. What are the machine learning algorithms suitable for predicting the shipping product delivery time?
2. How are the model performance to be improved?
3. How are the prediction models to be evaluated?

1.3 AIM & OBJECTIVES

1.3.1 Aim

The aim of this study is to develop prediction models that predict the delivery time of shipping products as a classification problem whether products are delivered on time or not on time by using machine learning algorithms.

1.3.2 Objectives

The objectives of the study are as followed:

1. To identify suitable machine learning algorithms for prediction of shipping product delivery time.
2. To develop shipping product delivery time prediction models and apply hyperparameter optimization techniques to achieve better accuracy.
3. To evaluate performance of the prediction models against existing related works.

SECTION 2

LITERATURE REVIEW

2.1 LITERATURE REVIEW

A literature review on ten existing works related to the logistic domain of predicting the package delivery time is performed and tabulated in Table 2.1. Following the table, will be a discussion on the findings based on the literature review.

Table 2.1: Literature review matrix

Reference	Dataset & Size (Row x Col)	Methodology	Result		Conclusion	Recommendation	Comment
(Sundaram & Tallamraju, 2021)	Trip information data (1000000 x 8)	<ul style="list-style-type: none">- Predicts shipping time of fashion products.- Feature selection performed after one-hot encoding on categorical variables.- Hyperparameter optimization to solve overfitting issue.- Algorithms: LightGBM, Random Forest, XGBoost, LSTM- Evaluation: Accuracy	LightGBM	Accuracy: 68.10%	LightGBM gave best result.	N/A	<ul style="list-style-type: none">- No mention of what feature is selected.- No mention of normalizing or scaling of features.
			XGBoost	Accuracy: 58.13%			
			LSTM	Accuracy: 57.25%			
			RF	Accuracy: 55.45%			
(Ahmed <i>et al.</i> , 2021)		<ul style="list-style-type: none">- Predicts parcel delivery time.	XGBoost	MAE: 14.57 minutes	XGBoost and LightGBM	Neural network of different	<ul style="list-style-type: none">- Neural network did

	Trip information data (211392 x 10)	<ul style="list-style-type: none"> - Data merge from three datasets. - Preprocessing applied to remove duplicates and outliers. - Imputation for missing values by interpolation. - Data formatting applied on the different dataset. - Performed feature selection by SHAP. - Algorithms: XGBoost, LightGBM, LSTM, biLSTM, GRU, SVR - Evaluation: RMSE, MAE 		RMSE: 26.51 minutes	provides the best performance.	architecture and hybrid models with more complex architecture can be explored.	not outperform boosting-based algorithms. - No mention of normalizing and scaling of features.
			BiLSTM	MAE: 16.30 minutes			
				RMSE: 29.96 minutes			
			GRU	MAE: 16.56 minutes			
				RMSE: 29.97 minutes			
			LightGBM	MAE: 14.40 minutes			
				RMSE: 26.54 minutes			
			LSTM	MAE: 16.60 minutes			
				RMSE: 29.97 minutes			
			SVR	MAE: 25.77 minutes			
				RMSE: 49.48 minutes			
(Ogura <i>et al.</i> , 2021)	Vessel trip information data (414x 12)	<ul style="list-style-type: none"> - Studies effect of weather on vessels arrival time prediction. - Dataset was small thus chose the use of Naïve Bayes. - Algorithm: Naïve Bayes - Evaluation: Accuracy 	Naïve Bayes	Accuracy: 90%	Only one model is utilized, but the weather condition plays a significant role in	Data such as congestion status en route and at destination port can be considered to improve prediction accuracy.	Naïve bayes can have better accuracy than other methods such as SVM when dataset is small.

					prediction accuracy.		
(Chandra, 2021)	E-commerce dataset (10999,12)	<ul style="list-style-type: none"> - Predicts parcel delivery time. - Performed feature scaling using StandardScaler. - Algorithms: Random Forest, KNN - Evaluation: Accuracy 	RF	Accuracy: 65.27%	Random Forest performed better than KNN.	N/A	- No mention of feature selection.
			KNN	Accuracy: 63%			
(Bagga, 2021)	E-commerce dataset (10999,12)	<ul style="list-style-type: none"> - Predicts parcel delivery time. - Direct implementation of raw data to algorithms. - Algorithms: Logistic Regression, KNN - Evaluation: Accuracy 	Logistic Regression	Accuracy: 65%	Logistic regression and KNN obtained same result.	N/A	- No mention of normalizing and scaling of features.
			KNN	Accuracy: 65%			
(Kumar, 2021)	E-commerce dataset (10999,12)	<ul style="list-style-type: none"> - Predicts parcel delivery time. - Performed feature selection by chi square test, eliminating least significant features. - Performed feature scaling using MinMaxScaler. - Performed hyperparameter tuning using grid search only on RF. - Algorithms: SVM, RF, ANN - Evaluation: Accuracy 	ANN	Accuracy: 67%	ANN provides better performance than SVM and RF but not significantly lot.	N/A	- No hyperparameter tuning performed on ANN model.
			SVM	Accuracy: 66%			
			RF	Accuracy: 66%			
(Khiari & Olaverri-	Trip information	<ul style="list-style-type: none"> - Predicts delivery time in postal services. 	Gradient boosting	MAE: 61.67 minutes	- Boosting-based	To implement weather data and	- No mention of

Monreal, 2020)	data (193684 x 7)	<ul style="list-style-type: none"> - Performed feature engineering from raw dataset. - Algorithms: Gradient boosting, Adaboost, Extreme gradient boosting, Light gradient boosting, Histogram gradient boosting, Catboost - Evaluation: MAE, RMSE 		RMSE: 238.61 minutes	algorithms has great potential in predicting travel time. - LightGBM, Catboost, and Histogram boosting yielded a low computational time without compromising accuracy.	intermediate GPS data to improve delivery time prediction accuracy.	normalizing or scaling of features. - Evaluated the computational time of the prediction models.
			AdaBoost	MAE: 133.21 minutes			
				RMSE: 223.13 minutes			
			Histogram gradient boosting	MAE: 48.08 minutes			
				RMSE: 247.04 minutes			
			XGBoost	MAE: 34.41 minutes			
				RMSE: 402.81 minutes			
			CatBoost	MAE: 67.85 minutes			
				RMSE: 162.17 minutes			
			LightGBM	MAE: 48.10 minutes			
				RMSE: 246.77 minutes			
(Magiya, 2020)	Trip information data (28269 x 39)	<ul style="list-style-type: none"> - Predicts parcel delivery time by motorcycle. - Performed feature selection. 	XGBoost	Accuracy: 42.21%	Feature selection is important for model to	To study how different delivery vehicle types would affect delivery time.	- No mention of normalizing and scaling of features.

		<ul style="list-style-type: none"> - Performed cross validation and hold out validation. - Algorithm: XGBoost - Evaluation: Accuracy 			perform better.		
(Wu & Wu, 2019)	Trip information data (350000 x 10)	<ul style="list-style-type: none"> - Predicts parcel delivery time. - Performed feature normalization using padding method. - Algorithm: Stacked LSTM - Evaluation: RMSE, MAPE 	LSTM	RMSE: 63.58 minutes	- LSTM significantly outperforms the benchmarks on real related works.	Using data about sequence of route might improve prediction accuracy.	Combining two models output by a fully connected neural network for prediction.
(Servos <i>et al.</i> , 2019)	Trip information data (101136 x 36)	<ul style="list-style-type: none"> - Predicts freight delivery time. - Performed data cleansing to remove outliers and denoise. - Feature engineering and selection of significant features. - Used grid search and cross validation for parameter tuning. - Algorithms: Randomized Trees, AdaBoost, SVR - Evaluation: RMSE, MAE 	Randomized Trees	MAE: 108.39 minutes	Support vector regression performed best.	To be evaluated on a larger transport fleet and different transport mode.	- No mention of normalizing and scaling of features.
				RMSE: 119.94 minutes			
			AdaBoost	MAE: 103.67 minutes			
				RMSE: 113.83 minutes			
			SVR	MAE: 31 minutes			
				RMSE: 38.88 minutes			

2.2 DISCUSSION

The logistic chain is vast and involve multiple mode of shipment, typically using ships, planes, and road vehicles. Generally, trip information data for package delivery using ships and road vehicles are frequently used in predicting the package delivery time instead of trip information data of planes. This is due to the highly uncertain events associated with this type of shipping mode such as traffic conditions, weather conditions, travel path restrictions, etc. While for deliveries using planes, less uncertainties are present due to the highly regulated and planning involved in every shipment and does not conform to the traffic and weather conditions like other mode of transport.

The most frequently used prediction algorithm is the boosting-based algorithm such as AdaBoost, XGBoost, etc. The boosting-based algorithms have garnered popularity due to the high computational speed and highly accurate result. This is due to the requirement of real time output needed for the companies to perform decision making and planning. However, in the case of Servos *et al.* (2019), it was found that SVR outperforms AdaBoost. Following that, the second frequently used algorithms are neural networks. This is due to the multidimensional data collected from each leg of the supply chain, which neural networks are claimed to perform well on such large and complex dataset. The use of neural network algorithm can be seen in the works of Kumar (2021) which performed better as compared to SVM and RF.

Missing data are typical, as data is collected from different vehicles using sensors which might malfunction due to certain reasons. However, only Ahmed *et al.* (2021) has performed missing data imputation by using interpolation method while others did not mention on missing data and how to deal with missing data. Generally, data is abundant when it is obtained from a logistic company as current technologies allowed easy implementation and tracking of vehicle movements. Therefore, authors may opt to drop the missing data instead of imputing the missing data.

Furthermore, it was observed that feature scaling and feature selection is performed on majority of the works (Chandra, 2021; Kumar, 2021; Magiya, 2020; Servos *et al.*, 2019; Wu & Wu, 2019). Feature scaling is performed due to the data consisting of various information typically in different value ranges such as trip distance, package weight, number of stops, etc. By having features in the

same scale, this would improve model performance especially distance-based algorithms such as KNN. In addition, feature selection is performed to reduce model complexity as not all features provide significant impact to the model performance.

For the evaluation metric, two major groups of metrics are observed depending on the type of problems. For regression type problems, the metric MAE and RMSE is frequently used. This type of works predicts the exact time of delivery which involves more complex dataset containing higher feature numbers. It is not able to directly compare the performance between the authors due to the different dataset, subdomain, and transportation mode used which varies greatly dependent on the location of study. On the other hand, typical metric used for classification type problems is accuracy. Which mostly predicts whether package delivered on time or not on time. Comparing the works of Kumar (2021), Bagga (2021), and Chandra (2021) which originates from the same dataset. It is observed that ANN by Kumar (2021) provides the best model performance achieving an accuracy of 67%. However, the accuracies achieved by similar works average about 65% which the result achieved by ANN is not a significant improvement. Therefore, more work can be done by utilizing other methods of machine learning approaches to improve the prediction result.

SECTION 3

METHODS

3.1 DATASET

The dataset is an e-commerce shipment dataset which comprise of shipping trip information and delivery package information. The dataset comprises of 10999 observations and 12 variables. In addition, the **dependent variable** of this dataset is the feature “**Reached on time**” which is a binary class. The dataset will be used for predicting package delivery time as a classification problem of whether package is delivered on time or not on time. Table 3.1 tabulates the description for each variable in the dataset.

The dataset is retrieved from Kaggle which the link to the dataset is as followed:

<https://www.kaggle.com/datasets/prachi13/customer-analytics>

Table 3.1: Dataset variables description

No.	Variable	Description	Labels
1	ID	Customer ID number	1 – 10999
2	Warehouse block	Warehouses which are divided into blocks	A, B, C, D, E
3	Mode of Shipment	Package transportation mode	Ship, Flight, and Road
4	Customer care calls	Customer enquiry calls about the shipment	2 – 7
5	Customer rating	Customer rating with 1 as lowest (Worst) and 5 as highest (Best)	1 – 5
6	Cost of the product	Product price in US Dollars.	96 – 310
7	Prior purchases	Prior purchase count by customers	2 – 10
8	Product importance	Product importance rating with ranking	low, medium, high
9	Gender	Male and female	M, F
10	Weight in gms	Package weight in grams	1001 – 7846
11	Discount offered	Discount offered on specific product	1 – 65
12	Reached on time	Target variable of product arrival time. Where 1 indicates not on time, while 0 indicates reached on time	1, 0

Table 3.2 tabulates the data type of each variable indicating the numeric or categorical variable type. It is further divided where numeric variable would be divided into continuous or integer type variable. In addition, categorical variable would be divided into nominal and ordinal type variable type.

Table 3.2: Dataset variables data type

Numerical		Categorical	
Continuous	Integer	Nominal	Ordinal
Cost of the product	Customer care calls	Warehouse block	Customer rating
Weight in gms	Prior purchases	Mode of shipment	Product importance
Discount offered		Gender	

3.2 MACHINE LEARNING ALGORITHMS

The machine learning algorithms used in the prediction models for this study will be introduced below. Four machine learning algorithms are identified based on similar use case from the literature review. The implementation of the prediction models will be performed in RStudio.

3.2.1 Logistic Regression

Logistic regression is a supervised machine learning algorithm used to predict the relationship between a dependent variable and independent variable. It is typically used for classification type problems to predict a binary class output. Application of logistic regression can be seen in medical sectors predicting a tumor as benign or malignant, incoming email classifying as spam or not spam, etc. The logistic regression is a very basic prediction algorithm which is not expected to outperform other algorithms used in this study. It is served as a benchmark model in this study as a comparison to other prediction algorithms that are more powerful.

3.2.2 Extreme Gradient Boosting (XGBoost)

XGboost is a type of boosting ensemble learning algorithm which aimed to correct errors made by previous models. The base classifier of XGBoost uses decision trees. The error correction process is sequential and iterative until no further improvements can be made. It is very popular due to its computational speed and model performance which is due to the feature of XGBoost where parallelization is adopted in the computational process. The computational speed is much faster as

compared to its peers such as AdaBoost and RF. It is chosen to be used in this study due to the speed and performance that the algorithm provides.

3.2.3 Random Forest (RF)

Random Forest is a type of bagging ensemble learning algorithm where each base classifier is trained independently. Generally, the base classifier uses decision trees where each decision tree trains on a different subset of the training data. This provides diversity in each decision tree which typically results in better prediction performance as compared to a single model fitting the entire training dataset. The predictions made by the ensemble members are then aggregated using voting in the case of classification problem. Based on the literature review, boosting algorithms are a popular choice for package delivery time prediction thus is chosen for use in this study. In addition, related works that used the same dataset as this study have utilized random forest algorithm in prediction which can served as a comparison in performance after applying hyperparameter tuning (Bagga, 2021; Chandra, 2021; Kumar, 2021).

3.2.4 Adaptive Boosting (AdaBoost)

AdaBoost is a type of boosting ensemble learning algorithm where the base classifier training process is sequential and iterative. Generally, the base classifier uses decision trees where each decision tree trains on the entire training dataset with add-on of sample weights from previous model. The underlying of boosting ensembles is to correct the previous prediction errors of fitted model. The second model attempts to fix the errors of previous model by fitting to the dataset and placing a higher weightage on data samples with high error rate that the previous model encountered. Therefore, greater attention is placed on samples with higher error. This process is iterated until the sample error rate remains constant. The final prediction is made by summing up each weight multiply by each tree of the entire iterative process. Based on the literature review, boosting algorithms are a popular choice for package delivery time prediction thus is chosen for use in this study. In addition, related works that used the same dataset as this study have not utilized AdaBoost algorithm in prediction which can set as the benchmark on performance of AdaBoost algorithm for the specific dataset (Bagga, 2021; Chandra, 2021; Kumar, 2021).

3.3 EVALUATION METRIC

Evaluation metrics are indicators that measure performance of the predictive model. Different evaluation measures exist dependent on the type of model either as classification or regression model. The evaluation metrics are performed to identify the predictive capability of the model when unseen data is introduced thus providing information on how well the model generalize. Following would list few evaluation metrics utilized in this study which predicts a binary classification problem.

3.3.1 Confusion Matrix

A confusion matrix is a representation of prediction results displayed in a matrix format. Typically, a confusion matrix is used for binary classification to describe the performance of the model. However, extension of the confusion matrix exists to cater for higher number of classes. Figure 3.1 shows a typical confusion matrix with binary class labels.

		Predicted Values	
		Negative	Positive
Actual Values	Negative	TN True Negative	FP False positive
	Positive	FN False Negative	TP True Positive

Figure 3.1: Example confusion matrix (Chauhan, 2020)

Prediction from a classification model typically is categorized into one of the quadrants of the confusion matrix. Following show the description of each quadrant of a confusion matrix:

- True Positive (TP): Predicted true, and actual is true
- True Negative (TN): Predicted false, and actual is false

- False Positive (FP): Predicted true, but actual is false
- False Negative (FN): Predicted false, but actual is true

These outcomes from the confusion matrix are further used to derive other evaluation metrics to provide more information on the model performance.

3.3.2 Accuracy

One of the metrics derived from the confusion matrix is accuracy. Accuracy is the measure of correctly predicted outcomes compared to the total outcome. A high accuracy indicates the model can generalize well to unseen data.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

3.3.3 Specificity

Another one of the metrics derived from the confusion matrix is specificity. Specificity is the measure of how often the negative class is predicted correctly. It is the proportion of true negatives to the sum of classes that should be classified as negative. A high specificity would indicate fewer false positive results.

$$Specificity = \frac{TN}{TN + FP}$$

3.3.4 Sensitivity

Another one of the metrics derived from the confusion matrix is sensitivity. Sensitivity is the measure of how often the positive class is predicted correctly. It is the proportion of true positives to the sum of classes that should be classified as positive. A high sensitivity would indicate fewer false negative results.

$$Sensitivity = \frac{TP}{FN + FP}$$

3.3.5 Receiver Operating Characteristics (ROC)

The area under the ROC is another method to evaluate model performance. A plot of sensitivity against (1-Specificity) forms the ROC curve. The ROC can be used to visualize the trade off between true positive rate and false positive rate. Figure 3.2 shows a typical ROC curve plotted. Performance of a model is compared to the random guessing which is indicated as dashed line in the figure. Generally, the aim is to obtain a model that provides a curve that is closer to the top left corner of the chart which area under the curve is closer to one which indicates a better model. However, thresholds can be set depending on scenario which might result in other models being more favorable.

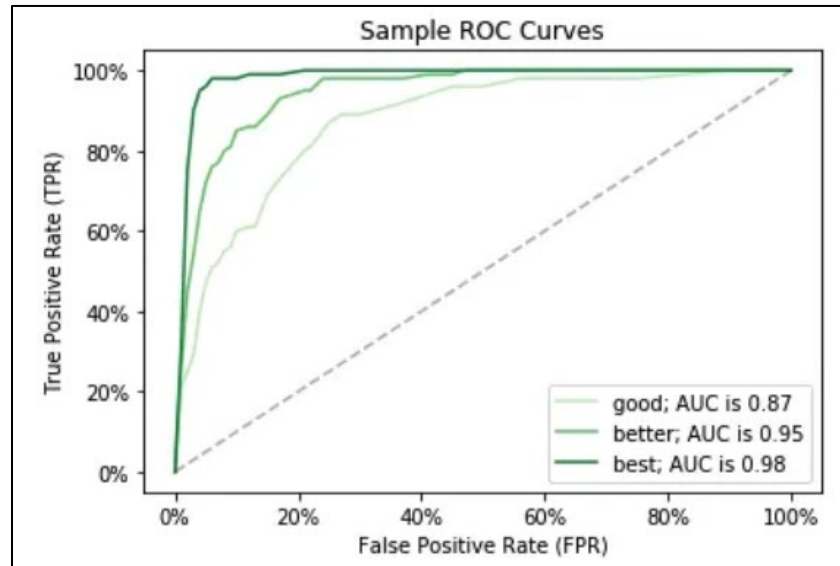


Figure 3.2: Example ROC curve (Chauhan, 2020)

3.3.6 Kappa

The Cohen's kappa is a model performance evaluation metric used for a classification model. It is typically used when the class distribution is significantly imbalanced and a multi-class output problem. It is calculated based on confusion matrix in addition considering the imbalance in class distribution. In contrast to the accuracy metric which does not consider the class distribution. The range of Cohen's kappa is from zero to one, which the value closer to one indicates a better classifier. Table 3.3 shows the interpretation of Cohen's kappa value ranges.

Table 3.3: Cohen's kappa value interpretation

Cohen's kappa value	Indicator
0.00 – 0.20	Poor agreement
0.21 – 0.40	Fair agreement
0.41 – 0.60	Moderate agreement
0.61 – 0.80	Good agreement
0.81 – 1.00	Perfect agreement

3.4 R PACKAGES

This section describe the R packages used in the study which includes for data exploratory analysis, data pre-processing, model development and validation. Table 3.4 lists the R packages used in this study along with their descriptions.

Table 3.4: R packages used

R Package Name	Description
DataExplorer	For data exploration and treatment
mice	For missing value imputation
dplyr	For data manipulation
caTools	For statistical analysis tool
caret	For model development
ROCR	For visualization of performance scoring

SECTION 4

DATASET PREPARATION

4.1 EXPLORATORY DATA ANALYSIS

The exploratory data analysis is the initial glimpse on the overall dataset to understand the characteristics of the dataset and to guide the following data pre-processing step.

4.1.1 Dataset Initial View

↑	↕ i.ID	↕ Warehouse_block	↕ Mode_of_Shipment	↕ Customer_care_calls	↕ Customer_rating	↕ Cost_of_the_Product
1	NA	D	Flight	4	2	177
2	2	F	Flight	4	5	216
3	NA	A	Flight	2	2	183
4	4	B	NA	3	3	176
5	NA	C	Flight	2	2	NA

↕ Prior_purchases	↕ Product_importance	↕ Gender	↕ Discount_offered	↕ Weight_in_gms	↕ Reached.on.Time_Y.N
3	low	F	NA	1233	1
2	low	M	59	NA	1
4	low	M	NA	3374	NA
4	medium	M	10	1177	1
3	NA	F	46	2484	1

Figure 4.1: Glimpse of the first few observations of the dataset

Figure 4.1 shows the first five observation of the dataset along with all the features. Few observations are identified and listed as followed:

- Missing values are present which is annotated by 'NA' thus **missing value imputation** is required.
- Categorical variables such as warehouse block, mode of shipment, product importance, and gender require **feature encoding**.
- The value range of the features differs greatly, such as weight in grams, discount offered, and cost of product thus require **feature scaling**.
- The ID column can be **removed** as it does not provide any information and significance to the model development.

- The second instance in the feature warehouse block is showing 'F'. Based on the metadata, warehouse block only has category 'A', 'B', 'C', 'D', and 'E'. Further investigation is required to identify the why category 'F' is present in the dataset.

```
'data.frame': 10999 obs. of 12 variables:
 $ i..ID      : int  NA 2 NA 4 NA 6 7 8 9 10 ...
 $ warehouse_block : chr  "D" "F" "A" "B" ...
 $ Mode_of_shipment : chr  "Flight" "Flight" "Flight" NA ...
 $ Customer_care_calls: int  4 4 2 3 2 3 NA 4 3 3 ...
 $ Customer_rating  : int  2 5 2 3 2 1 4 1 4 2 ...
 $ Cost_of_the_Product: int  177 216 183 176 NA 162 250 233 150 164 ...
 $ Prior_purchases  : int  3 2 4 4 3 3 3 2 3 3 ...
 $ Product_importance : chr  "low" "low" "low" "medium" ...
 $ Gender           : chr  "F" "M" "M" "M" ...
 $ Discount_offered  : int  NA 59 NA 10 46 12 3 48 NA 29 ...
 $ weight_in_gms     : int  1233 NA 3374 1177 2484 1417 2371 2804 1861 1187 ...
 $ Reached.on.Time_Y.N: int  1 1 NA 1 1 1 1 1 1 1 ...
```

Figure 4.2: Internal structure of the dataset

Figure 4.2 shows the internal structure of the dataset. Few observations are identified and listed as followed:

- The dataset contains **10999 observations** and **12 variables**.
- The features that require encoding into **factor** are warehouse block, mode of shipment, gender, and reached on time.
- The feature product importance is an ordinal variable, it is suggested to perform **label encoding** on such variable.

4.1.2 Data Distribution of Categorical Variables

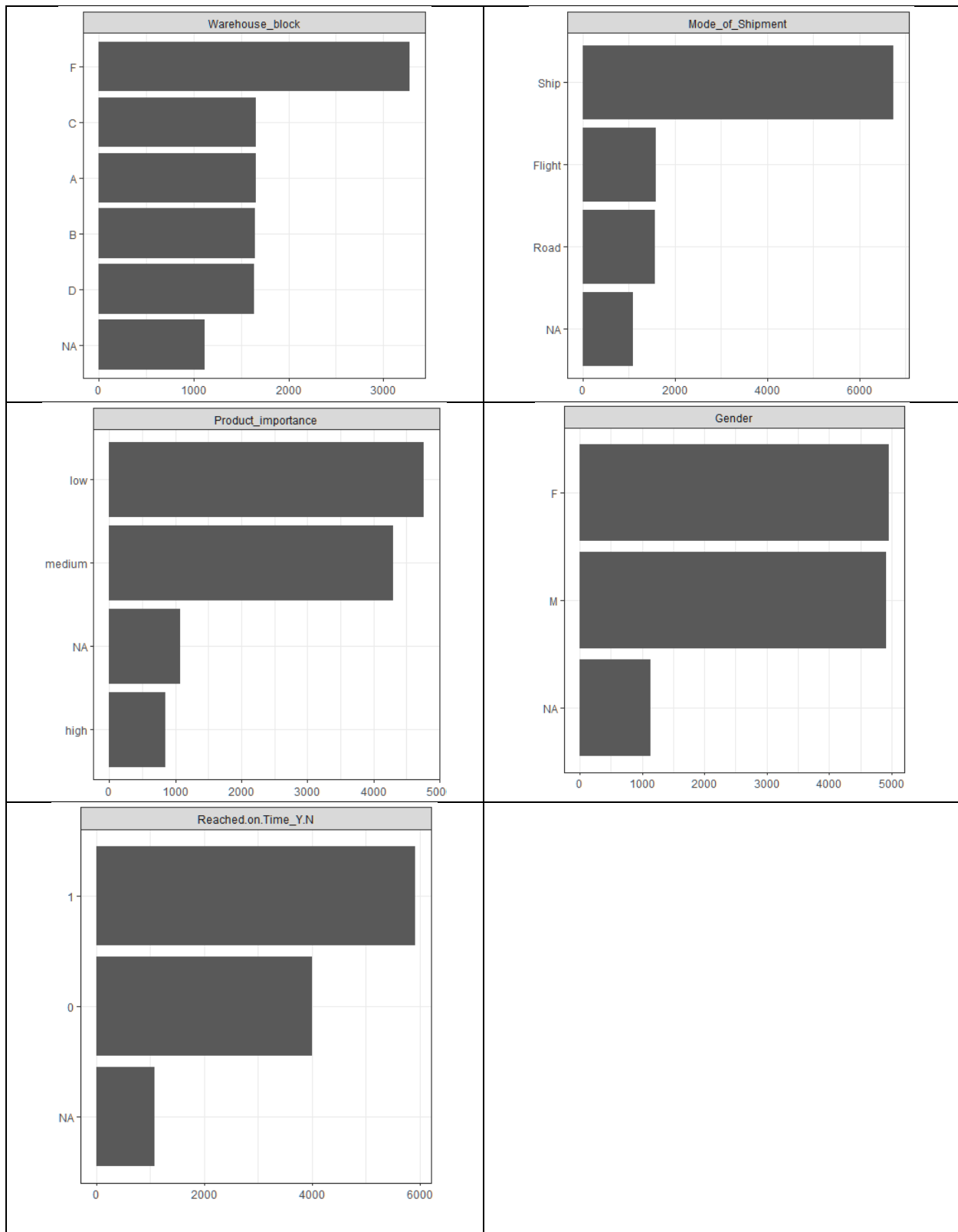


Figure 4.3: Data distribution of every categorical variable in the dataset

Figure 4.3 shows the data distribution of every categorical feature in dataset. The 'NA' observed in the bar chart of every feature indicates missing values which will not be discussed here. The vertical axis shows the categories available while the horizontal axis shows the frequency. This is applied to every bar chart in Figure 4.3. Observations of every feature will be listed as followed:

1. Warehouse block

- As previously identified the presence of warehouse block 'F' which does not align with the metadata. **Recode** of the warehouse block 'F' into warehouse block 'E' will be performed.
- Warehouse block 'F' is identified to be of the highest frequency with almost twice the amount as compared to other warehouse blocks.

2. Mode of shipment

- The shipment mode using ships has the highest frequency as compared to flight and road. It is significantly more than the other mode of shipment with approximately over four times more as compared to others.

3. Product importance

- Majority of the products fall under the low and medium importance which in total accounts for 90% of the total shipment.

4. Gender

- The frequency of male and female is distributed almost equally.

5. Reached on time (**Target Variable**)

- There is a higher number of products that did not reach on time as compared to products reaching on time.
- The proportion of products did not reach on time is about **60%**, while products that reached on time is about **40%**. Therefore, no balancing of the dataset is required.

4.1.3 Data Distribution of Continuous Variables

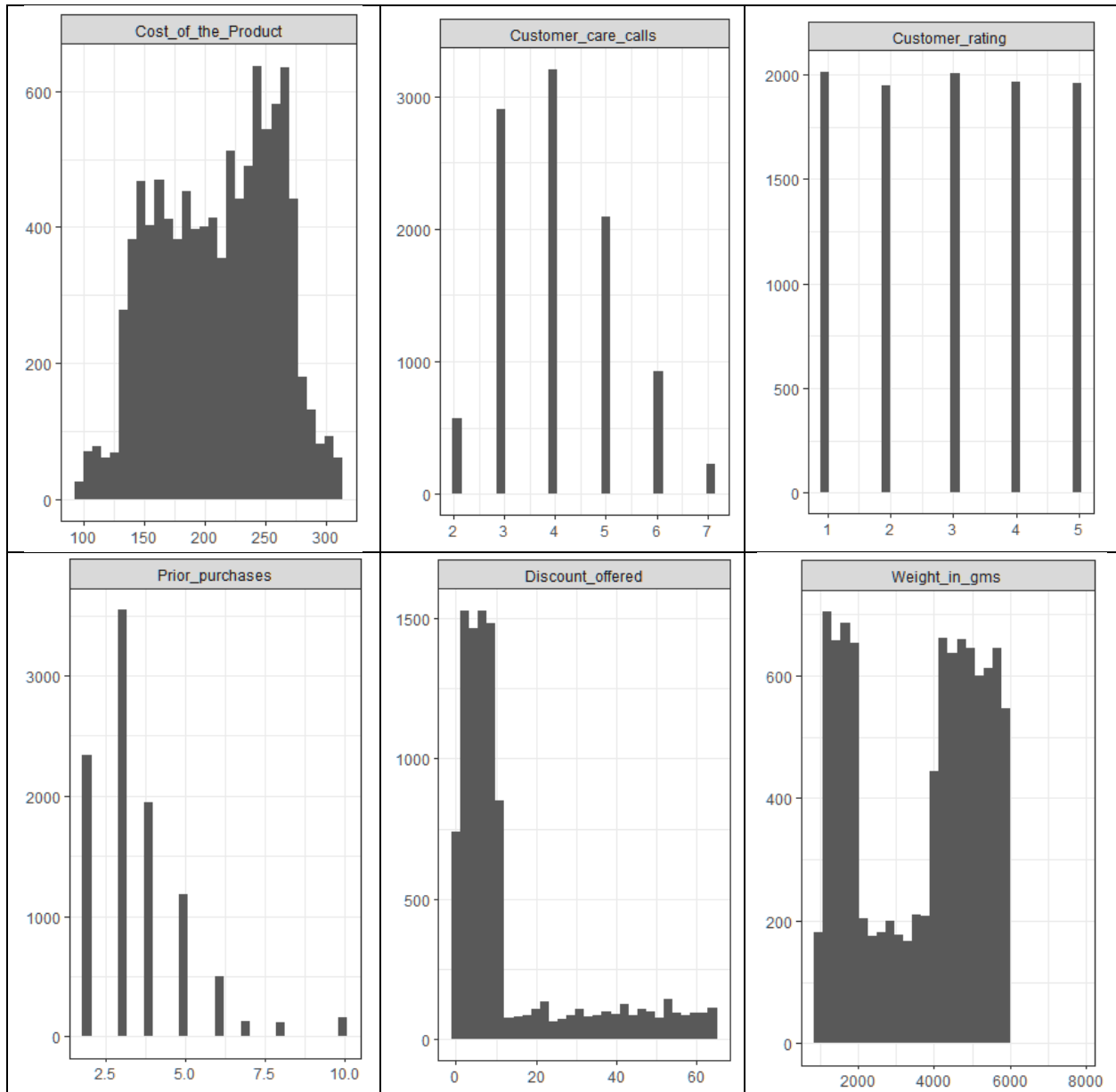


Figure 4.4: Data distribution of every numerical variable in the dataset

Figure 4.4 shows the data distribution of every numerical feature in dataset. The vertical axis shows the frequency while the horizontal axis shows the feature values. This is applied to every histogram in Figure 4.4. Observations of every feature will be listed as followed:

1. Cost of the product

- Two major groups of product cost interval are identified to have high frequency.

- Majority of the product cost is between about 225 to about 275.
- Second major group of product cost is identified between about 140 to about 224.

2. Customer care call

- Majority of the customer called between three or four times to the customer care.

3. Customer rating

- There is not much variation between the ratings from one to five as provided by the customer.

4. Prior purchases

- Majority of the customers have three prior purchases count.

5. Discount offered

- Majority of the discount offered is between 2% to 8%.

6. Weight in grams

- Two major groups of product weight interval are identified to have high frequency.
- Majority of the product weight is between about 1000 to about 2000 grams.
- Second major group of product weight is identified between about 4000 to about 6000 grams.

4.1.4 Correlation Between Features

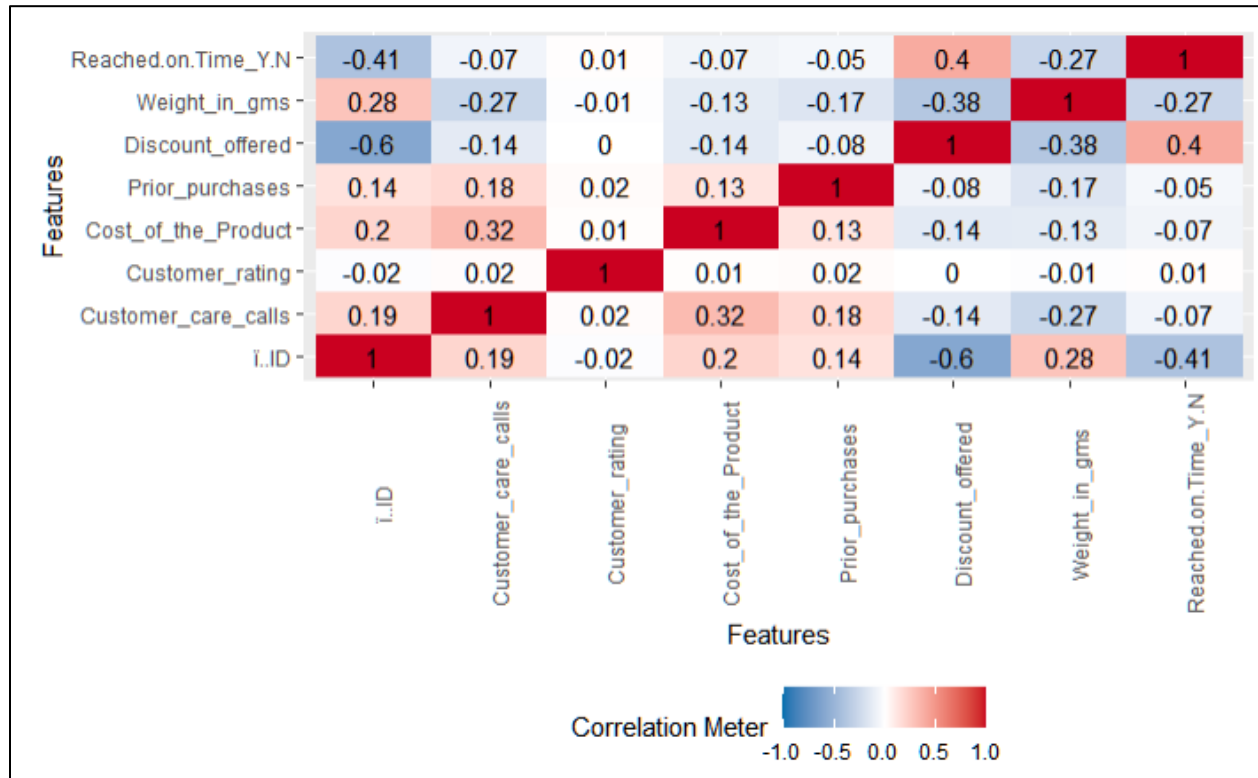


Figure 4.5: Correlation between features

Figure 4.5 shows the correlation matrix between numerical features. Observations from the correlation matrix are as followed:

- The discount offered has a medium positive correlation with the target variable reached on time with a correlation of 40%. A point increase in discount offered would imply a moderate increase in product reaching on time.
- Weight in grams has a weak negative correlation with the target variable reached on time with a correlation of -27%. A point increase in weight would imply a fair decrease in product reaching on time.
- Weight in grams has a weak negative correlation with customer care call frequency with a correlation of -27%. A point increase in weight would imply a fair decrease in customer care call frequency.
- Weight in grams has a medium negative correlation with discount offered with a correlation of -38%. A point increase in weight would imply a moderate decrease in discount offered.

- Cost of product has a medium positive correlation with customer care call frequency with a correlation of 32%. A point increase in cost of product would imply a moderate increase in customer care call frequency.

4.1.5 Identifying Missing Value

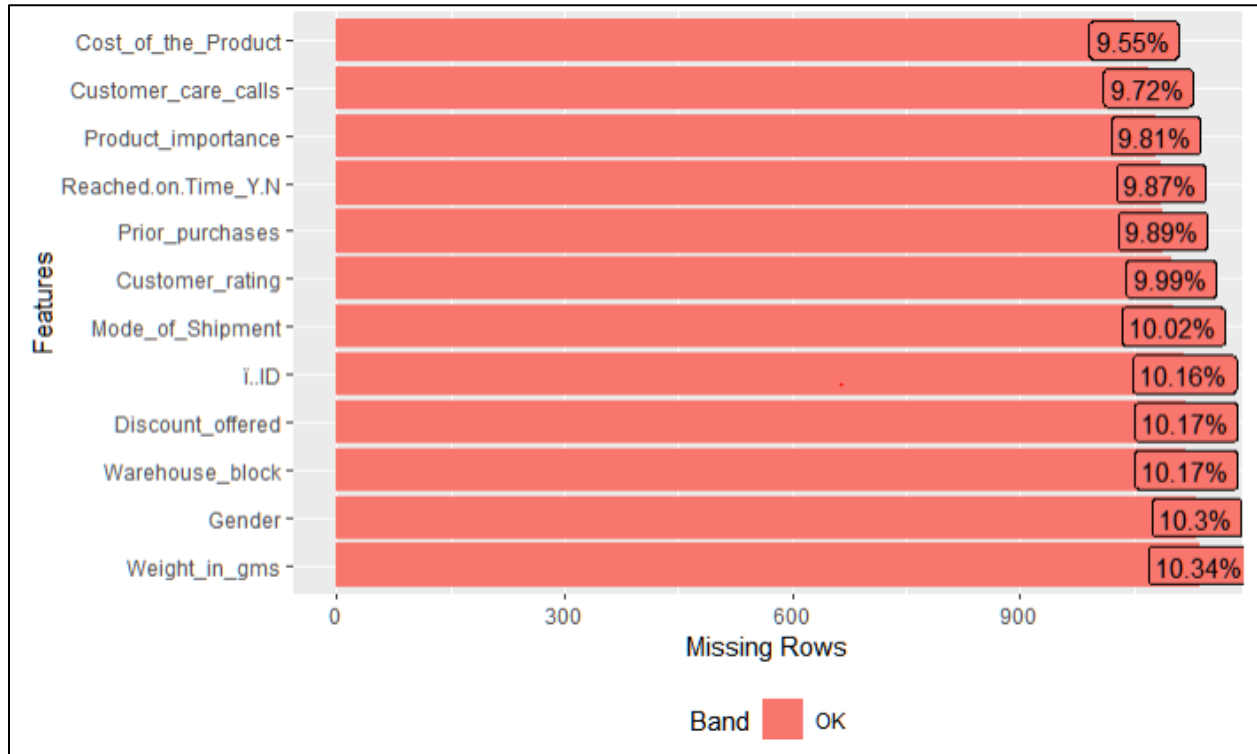


Figure 4.6: Missing values proportion

Figure 4.6 shows the proportion of missing values present in each feature. It is observed that all features have missing values and the proportion of missing values is about 10%. Since the proportion of missing value for every feature is low, **missing value imputation is suggested**. There is no need to drop any columns or rows.

4.2 DATA PRE-PROCESSING

4.2.1 Dropping and Recoding Features

The ID column will be dropped from the dataset as it is not informative in the model development stage. Figure 4.7 shows the internal structure of the dataset after dropping the ID column, leaving the dataset with 11 variables.

```
'data.frame': 10999 obs. of 11 variables:
 $ warehouse_block : chr "D" "F" "A" "B" ...
 $ Mode_of_shipment : chr "Flight" "Flight" "Flight" NA ...
 $ Customer_care_calls: int 4 4 2 3 2 3 NA 4 3 3 ...
 $ Customer_rating : int 2 5 2 3 2 1 4 1 4 2 ...
 $ Cost_of_the_Product: int 177 216 183 176 NA 162 250 233 150 164 ...
 $ Prior_purchases : int 3 2 4 4 3 3 3 2 3 3 ...
 $ Product_importance : chr "low" "low" "low" "medium" ...
 $ Gender : chr "F" "M" "M" "M" ...
 $ Discount_offered : int NA 59 NA 10 46 12 3 48 NA 29 ...
 $ weight_in_gms : int 1233 NA 3374 1177 2484 1417 2371 2804 1861 1187 ...
 $ Reached.on.Time_Y.N: int 1 1 NA 1 1 1 1 1 1 1 ...
```

Figure 4.7: Internal structure of dataset after dropping

As previously mentioned, the need to recode the feature warehouse block 'F' into 'E'. The recode function is used to recode the warehouse block 'F' values into 'E'. Table 4.1 shows the frequency of each warehouse block after recoding. In addition, the feature product importance which is an ordinal variable is recoded with the following encodings shown in Table 4.2.

Table 4.1: Frequency table of warehouse block after recoding

Warehouse block	A	B	C	D	E
Frequency	1659	1649	1661	1636	3275

Table 4.2: Recoding for product importance

Original value	Recoded value
Low	1
Medium	2
High	3

4.2.2 Factoring Categorical Variables

As previously mentioned, four categorical features will be converted to factor. The `as.factor` function is used to convert the categorical features into a factor. Figure 4.8 shows the internal structure of the dataset after converting categorical features into factor.

```
'data.frame': 10999 obs. of 11 variables:
 $ warehouse_block : Factor w/ 5 levels "A","B","C","D",...: 4 5 1 2 3 5 4 5 1 2 ...
 $ Mode_of_shipment : Factor w/ 3 levels "Flight","Road",...: 1 1 1 NA 1 1 1 NA 1 1 ...
 $ Customer_care_calls: int 4 4 2 3 2 3 NA 4 3 3 ...
 $ Customer_rating : int 2 5 2 3 2 1 4 1 4 2 ...
 $ Cost_of_the_Product: int 177 216 183 176 NA 162 250 233 150 164 ...
 $ Prior_purchases : int 3 2 4 4 3 3 3 2 3 3 ...
 $ Product_importance : int 1 1 1 2 NA 2 1 1 1 2 ...
 $ Gender : Factor w/ 2 levels "F","M": 1 2 2 2 1 1 1 NA NA 1 ...
 $ Discount_offered : int NA 59 NA 10 46 12 3 48 NA 29 ...
 $ Weight_in_gms : int 1233 NA 3374 1177 2484 1417 2371 2804 1861 1187 ...
 $ Reached.on.Time_Y.N: Factor w/ 2 levels "0","1": 2 2 NA 2 2 2 2 2 2 2 ...
```

Figure 4.8: Internal structure of dataset after factor

Based on Figure 4.8 the following observations are identified and listed as followed:

- The feature warehouse block after applying factor contains five levels.
- The feature mode of shipment after applying factor contains three levels.
- The feature gender after applying factor contains two levels.
- The feature reached on time after applying factor contains two levels.

4.2.3 Missing Value Imputation

The missing values present in the dataset is imputed using the `mice` package. Figure 4.9 shows the proportion of missing values after imputation. Based on the figure, all features now have zero missing values.

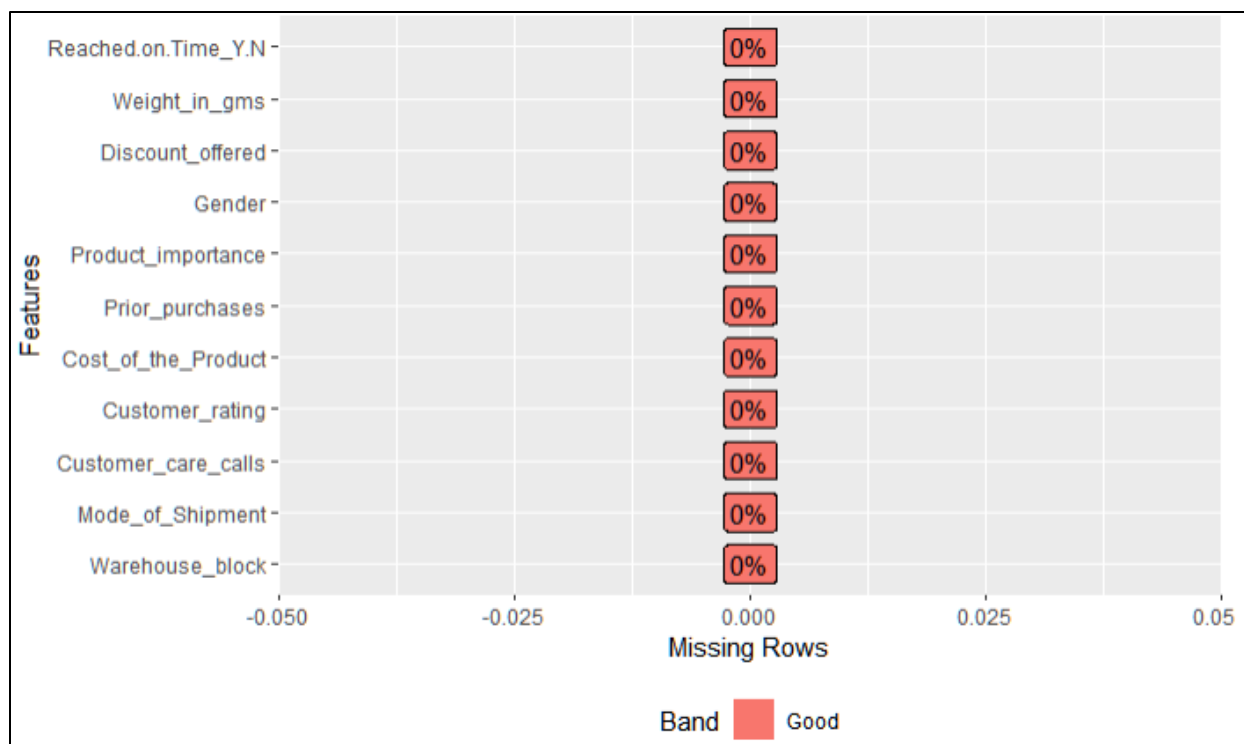


Figure 4.9: Missing value imputation result

4.2.4 Feature Encoding

Two types of feature encoding are performed namely label encoding and one-hot encoding. Label encoding is performed on the feature gender as the values are only binary which is redundant if applied one-hot encoding. One-hot encoding is applied using dummify function from the DataExplorer package. Figure 4.10 shows the internal structure of the dataset after feature encoding.

```

'data.frame': 10999 obs. of 17 variables:
 $ Customer_care_calls : int 4 4 2 3 2 3 5 4 3 3 ...
 $ Customer_rating : int 2 5 2 3 2 1 4 1 4 2 ...
 $ Cost_of_the_Product : int 177 216 183 176 201 162 250 233 150 164 ...
 $ Prior_purchases : int 3 2 4 4 3 3 3 2 3 3 ...
 $ Product_importance : num 1 1 1 2 1 2 1 1 1 2 ...
 $ Gender : num 1 0 0 0 1 1 1 1 1 1 ...
 $ Discount_offered : int 42 59 23 10 46 12 3 48 32 29 ...
 $ Weight_in_gms : int 1233 1645 3374 1177 2484 1417 2371 2804 1861 1187 ...
 $ Reached.on.Time_Y.N : Factor w/ 2 levels "On Time","Not on Time": 2 2 2 2 2 2 2 2 2 2 ...
 $ Warehouse_block_A : int 0 0 1 0 0 0 0 0 1 0 ...
 $ Warehouse_block_B : int 0 0 0 1 0 0 0 0 0 1 ...
 $ Warehouse_block_C : int 0 0 0 0 1 0 0 0 0 0 ...
 $ Warehouse_block_D : int 1 0 0 0 0 0 1 0 0 0 ...
 $ Warehouse_block_E : int 0 1 0 0 0 1 0 1 0 0 ...
 $ Mode_of_Shipment_Flight : int 1 1 1 0 1 1 1 1 1 1 ...
 $ Mode_of_Shipment_Road : int 0 0 0 0 0 0 0 0 0 0 ...
 $ Mode_of_Shipment_Ship : int 0 0 0 1 0 0 0 0 0 0 ...

```

Figure 4.10: Internal structure after feature encoding

Based on Figure 4.10 the following observations are identified and listed as followed:

- The feature warehouse block is one-hot encoded and divided into five features, each representing one warehouse block.
- The feature mode of shipment is one-hot encoded and divided into three features, each representing one mode of shipment.
- The total number of variables is now increased to 17 variables.

4.2.5 Feature scaling

Feature scaling is performed since the range of values of the dataset varies widely. Normalization is applied for scaling the features to confine the value range between zero and one. This is performed to achieve better model convergence. Figure 4.11 shows the summary statistics of the features after applying feature scaling. The `preProcess` function from the `caret` package is utilized to perform the feature scaling with the parameter 'method' set as 'range' to perform normalization on features.

Customer_care_calls	Customer_rating	Cost_of_the_Product	Prior_purchases	Product_importance	Gender	Discount_offered
Min. :0.0000	Min. :0.0000	Min. :0.0000	Min. :0.0000	Min. :0.0000	Min. :0.0000	Min. :0.00000
1st Qu.:0.2000	1st Qu.:0.2500	1st Qu.:0.3458	1st Qu.:0.1250	1st Qu.:0.0000	1st Qu.:0.0000	1st Qu.:0.04688
Median :0.4000	Median :0.5000	Median :0.5514	Median :0.1250	Median :0.5000	Median :1.0000	Median :0.09375
Mean :0.4121	Mean :0.4997	Mean :0.5339	Mean :0.1963	Mean :0.3013	Mean :0.5025	Mean :0.19393
3rd Qu.:0.6000	3rd Qu.:0.7500	3rd Qu.:0.7243	3rd Qu.:0.2500	3rd Qu.:0.5000	3rd Qu.:1.0000	3rd Qu.:0.14062
Max. :1.0000	Max. :1.0000	Max. :1.0000	Max. :1.0000	Max. :1.0000	Max. :1.0000	Max. :1.00000
Weight_in_gms	Reached.on.Time_Y.N	warehouse_block_A	warehouse_block_B	warehouse_block_C	warehouse_block_D	warehouse_block_E
Min. :0.0000	On Time :4454	Min. :0.0000	Min. :0.0000	Min. :0.0000	Min. :0.0000	Min. :0.0000
1st Qu.:0.1243	Not on Time:6545	1st Qu.:0.0000	1st Qu.:0.0000	1st Qu.:0.0000	1st Qu.:0.0000	1st Qu.:0.0000
Median :0.4603		Median :0.0000	Median :0.0000	Median :0.0000	Median :0.0000	Median :0.0000
Mean :0.3857		Mean :0.1701	Mean :0.1683	Mean :0.1663	Mean :0.1657	Mean :0.3297
3rd Qu.:0.5918		3rd Qu.:0.0000	3rd Qu.:0.0000	3rd Qu.:0.0000	3rd Qu.:0.0000	3rd Qu.:1.0000
Max. :1.0000		Max. :1.0000	Max. :1.0000	Max. :1.0000	Max. :1.0000	Max. :1.0000
Mode_of_Shipment_Flight	Mode_of_Shipment_Road	Mode_of_Shipment_Ship				
Min. :0.0000	Min. :0.0000	Min. :0.0000				
1st Qu.:0.0000	1st Qu.:0.0000	1st Qu.:0.0000				
Median :0.0000	Median :0.0000	Median :1.0000				
Mean :0.1611	Mean :0.1582	Mean :0.6807				
3rd Qu.:0.0000	3rd Qu.:0.0000	3rd Qu.:1.0000				
Max. :1.0000	Max. :1.0000	Max. :1.0000				

Figure 4.11: Summary statistics of features after feature scaling

Based on Figure 4.11 the following observations are identified and listed as followed:

- All features now have a value range between zero and one.
- The target variable reached on time remains as a factor.

4.2.6 Train and Test Data Split

The data is split into two subsets namely training set and testing set. The training set will be used for fitting the model while the testing set will be used for validation. The split ratio adopted for the dataset is 80% allocated to the training set and 20% allocated to the testing set. The data split is performed using sample.split function from the caTools package which preserve the relative ratio of the class labels. Table 4.3 shows the target variable class frequency after applying the data split.

Table 4.3: Class frequency of train and test set

	Class frequency		Total
	On Time	Not on Time	
Training Set	3563	5236	8799
Testing Set	891	1309	2200

Based on Table 4.3 the following observations are identified and listed as followed:

- The proportion of each class after splitting remained the same as the original dataset which is about 40% for the On Time class and 60% for the Not on Time class.
- 80% of the total observations falls into the training set with a total 8799 observations.
- 20% of the total observations falls into the testing set with a total 2200 observations.

SECTION 5

MODEL IMPLEMENTATION

Four machine learning algorithms are chosen and implemented to identify the best performer to be compared to related literature. The model development process will be discussed in this section which include model fitting to training data subset, validation by testing data subset, and evaluation metric computation for evaluating the performance of each model. All models will be built using the caret package.

Each of the algorithm used to develop the model would consist of a base model with hyperparameters settings as default and an optimized model where hyperparameter tuning is applied. The prediction task is a classifier that predicts the delivery time of the package with output either as delivery is on time or not on time.

5.1 LOGISTIC REGRESSION

Since logistic regression does not have any hyperparameters for tuning thus only one model is developed. The performance of the logistic regression model will serve as a benchmark to the other models in this study.

5.1.1 Logistic Regression with Cross Validation

The logistic regression model is built using the generalized linear model method from the caret package. A 10-fold repeated cross validation (CV) is applied to the fitting model. Figure 5.1 shows a snippet of the code used for developing the model.

```
# Train set
tc1_2 = trainControl(method='repeatedcv',number = 10, repeats = 3)
cls1_2 = train(Reached.on.Time_Y.N ~., data = trainSet, method = "glm",metric = 'Accuracy',family = binomial, trControl = tc1_2)
summary(cls1_2)

predTrain1_2 = predict(cls1_2, subset(trainSet, select = -(Reached.on.Time_Y.N)))

# Test set
predTest1_2 = predict(cls1_2, subset(testSet, select = -(Reached.on.Time_Y.N)))
```

Figure 5.1: Logistic regression model with cross validation snippet

5.1.2 Logistic Regression Model Performance

The logistic regression model performance will be discussed in this section. Table 5.1 shows the confusion matrix for the logistic regression model at training and testing stage. These results are further processed to calculate for other evaluation metrics. Table 5.2 shows the evaluation metrics computed based on the confusion matrix of the model.

Table 5.1: Confusion matrix for logistic regression model

Base Model	Train Set			Test Set		
	Reference Prediction	On Time	Not on Time	Reference Prediction	On Time	Not on Time
	On Time	2151	1606	On Time	519	398
	Not on Time	1412	3630	Not on Time	372	911

Table 5.2: Evaluation metric results for logistic regression model

	Base Model	
	Train Set	Test Set
Accuracy	0.6570	0.6500
Kappa	0.2944	0.2772
Sensitivity	0.6037	0.5825
Specificity	0.6933	0.6960

Based on Table 5.2, the following observations are identified and listed as followed:

- Accuracy on training set obtained 65.7% which is slightly higher than the testing set of 65%. The difference is small which indicates model is a good fit.
- Kappa on the testing set obtained 0.2772 which indicates a fair agreement in the predictive performance.
- Sensitivity on the testing set obtained 0.5825 which indicates model is moderately accurate in correctly classifying positive outcomes.
- Specificity on the testing set obtained 0.6960 which indicates model has good accuracy in correctly classifying negative outcomes.

Figure 5.2 shows the comparison of Receiver Operating Characteristic (ROC) curve for the logistic regression model for the training data and testing data. The area under the ROC curve is computed and labeled in each diagram. It is observed that the testing data area under ROC is 0.6390 which

is slightly lower than the training data of 0.6480. This indicates model has a good fit but model has weak performance in identifying the positive and negative class.

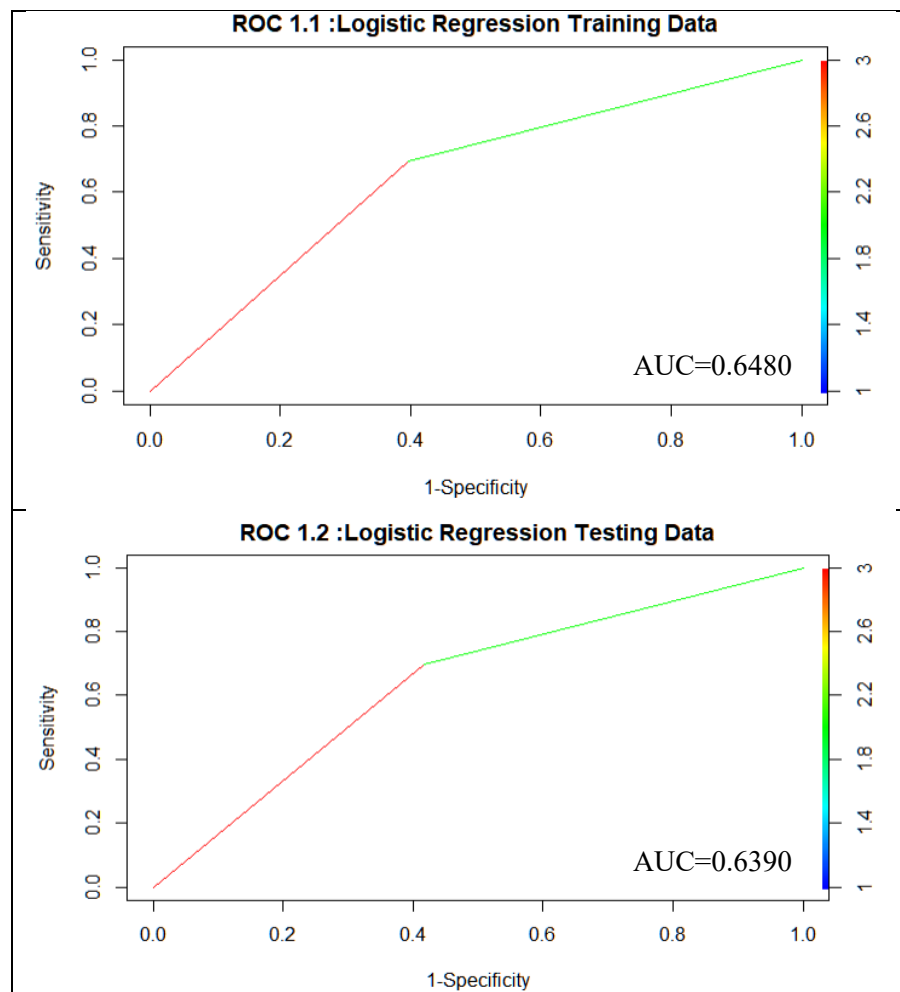


Figure 5.2: ROC curve for basic logistic regression

5.2 EXTREME GRADIENT BOOSTING (XGBOOST)

Two models are developed based on the XGBoost algorithm. First model is a basic model that adopts a fixed and default hyperparameter setting while the second model undergoes hyperparameter optimization to identify best performing model.

5.2.1 Basic XGBoost model

The basic XGBoost model is built using `xgbTree` method from `caret`. The following list the hyperparameters along with their description. These hyperparameter values are the default values

and will be fixed when fitting the model. Figure 5.3 shows a snippet of basic XGBoost model building and fitting to the training and testing set.

- ‘eta’: Learning rate = 0.3
- ‘gamma’: Minimum loss reduction = 0
- ‘max_depth’: Maximum depth of tree = 6
- ‘subsample’: Subsample ratio of training instances = 1
- ‘colsample_bytree’: Subsample ratio of columns when constructing each tree = 1
- ‘min_child_weight’: Minimum sum of instance weight needed in a child = 1
- ‘nrounds’: Number of trees in the final model = 100

```
grid = expand.grid(eta=0.3, nrounds=100, max_depth=6, gamma=0, colsample_bytree =1, min_child_weight=1, subsample=1)|
cls2_1 = train(Reached.on.Time_Y.N~., data=trainSet, method="xgbTree", tuneGrid = grid, metric='Accuracy')
# Training set
predTrain2_1 = predict(cls2_1, trainSet)
# Testing set
predTest2_1 = predict(cls2_1, testSet)
```

Figure 5.3: Snippet of basic XGBoost model building and fitting

5.2.2 Optimized XGBoost model

The optimized XGBoost model is built using xgbTree method from caret. The optimization of the XGBoost model is performed using grid search method on a range of hyperparameter values. In addition, 10-fold repeated CV is applied to the fitting model during the hyperparameter optimization process. The hyperparameter value ranges for the grid search are listed as followed:

- ‘eta’: 0.0.5
- ‘gamma’: from 0.1 to 1 with step 0.1
- ‘max_depth’: from 5 to 10 with step 1
- ‘subsample’: 0.5
- ‘colsample_bytree’: 0.5
- ‘min_child_weight’: from 1 to 20 with step 2
- ‘nrounds’: from 50 to 200 with step 50

Based on the range of hyperparameter values, 2640 combinations of hyperparameter value settings are generated which the model would iterate through to identify the best performing

hyperparameter value combination. Results from each iteration will be attached under Appendix A. Following that, a XGBoost model will be built upon the best performing hyperparameter value combination. Figure 5.4 shows a snippet of XGBoost model with grid search of the selected hyperparameter ranges and fitted to the training and testing set. The best performing hyperparameter combination identified are as followed:

- ‘eta’: 0.05
- ‘gamma’: 0.6
- ‘max_depth’: 5
- ‘subsample’: 0.5
- ‘colsample_bytree’: 0.5
- ‘min_child_weight’: 15
- ‘nrounds’: 50

```
tc2_2 = trainControl(method="repeatedcv", number=10, repeats=3, allowParallel = T, verboseIter = F)
grid=expand.grid(eta = 0.05,
                 max_depth = seq(5,10,1),
                 gamma = seq(0.1,0.1), |
                 subsample = 0.5,
                 colsample_bytree = 0.5,
                 min_child_weight= seq(1,20,2),
                 nrounds = seq(50,200,50))

cls2_2 = train(Reduced.on.Time.Y.N~., data=trainSet, method="xgbTree", trControl=tc2_2, tuneGrid =grid, metric='Accuracy', verbosity=0)
# Training set
predTrain2_2 = predict(cls2_2, trainSet)
# Testing set
predTest2_2 = predict(cls2_2, testSet)
```

Figure 5.4: Snippet of optimized XGBoost model building and fitting

5.2.3 XGBoost Model Performance

The model performance before and after applying hyperparameter optimization for the XGBoost models will be discussed in this section. Table 5.3 shows the confusion matrix for both the XGBoost models at training and testing stage. The results are used to calculate the evaluation metrics. Table 5.4 shows the evaluation metrics computed based on the confusion matrix for the XGBoost models.

Table 5.3: Confusion matrix for XGBoost models

	Train Set			Test Set		
Base Model	Reference Prediction	On Time	Not on Time	Reference Prediction	On Time	Not on Time
	On Time	3333	574	On Time	580	405
	Not on Time	230	4662	Not on Time	311	904
Optimized Model	Reference Prediction	On Time	Not on Time	Reference Prediction	On Time	Not on Time
	On Time	2954	1962	On Time	706	512
	Not on Time	609	3274	Not on Time	185	797

Table 5.4: Evaluation metric results for XGBoost models

	Base Model		Optimized Model	
	Train Set	Test Set	Train Set	Test Set
Accuracy	0.9086	0.6745	0.7078	0.6832
Kappa	0.8133	0.3359	0.4284	0.3790
Sensitivity	0.9354	0.6510	0.8291	0.7924
Specificity	0.8531	0.6906	0.6253	0.6089

Based on Table 5.4, the following observations are identified and listed as followed:

- The accuracy of the optimized model is higher as compared to the base model which obtained 68.32% and 67.45% respectively. This indicates the hyperparameter optimization has obtained a better performance model.
- The accuracy between the training and testing set for the base model has a wide difference which indicates model is overfitting. However, in the optimized model, the difference was small, which indicates model has a good fit.
- Kappa for both the models have fair agreement in the predictive performance.
- Sensitivity for the base model obtained 0.6510 which indicates model has moderate accuracy in correctly classifying positive outcomes. However, in the optimized model which obtained 0.7924 indicates model has good accuracy in correctly classifying positive outcomes.
- Specificity for the base model obtained 0.6906 which indicates moderate accuracy in correctly classifying negative outcomes. Similarly, in the optimized model which obtained 0.6089 indicates model has moderate accuracy in correctly classifying negative outcomes.

Figure 5.5 shows the comparison of ROC curve for the basic XGBoost model in the training and testing data. The area under the ROC curve is computed and labeled in each diagram. It is observed that the testing data area under ROC is 0.671 which is lower than the training data of 0.913. This indicates the model is slightly overfitting but provides acceptable performance in identifying the positive and negative class.

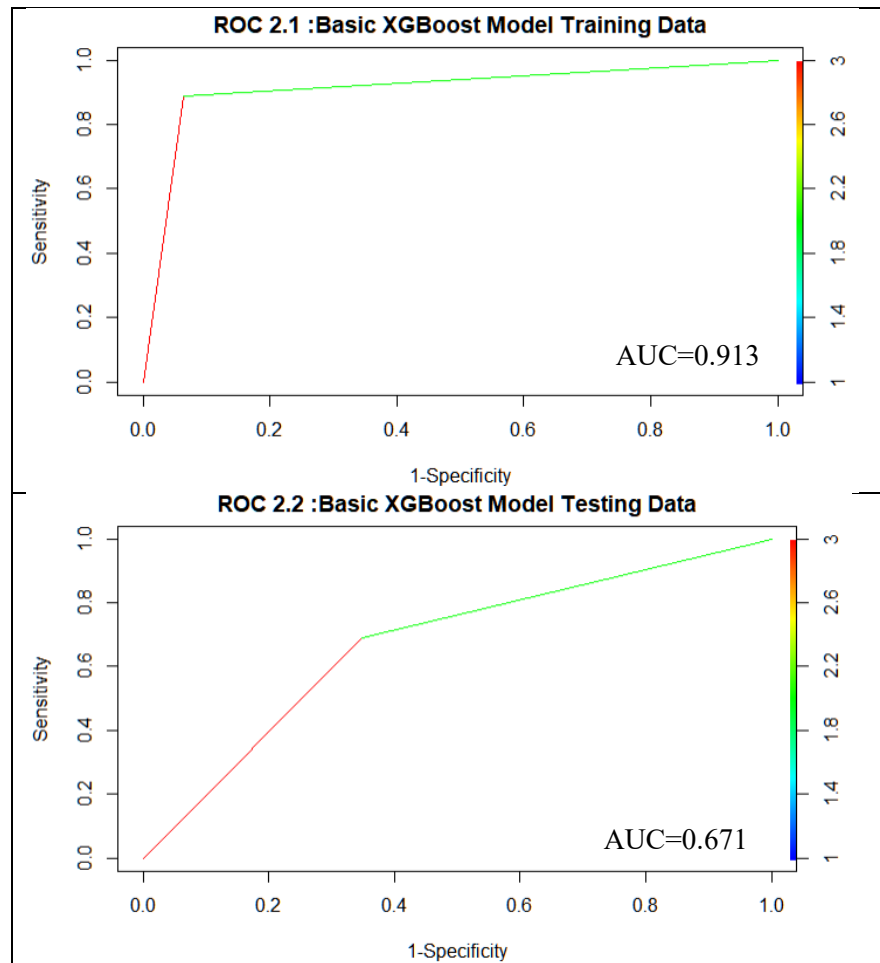


Figure 5.5: ROC curve for basic XGBoost

Figure 5.6 shows the comparison of ROC curve for the optimized XGBoost model in the training and testing data. The area under the ROC curve is computed and labeled in each diagram. It is observed that the testing data area under ROC is 0.701 which is slightly lower than the training data of 0.727. This indicates model is good fit and provides acceptable model performance in identifying the positive and negative class.

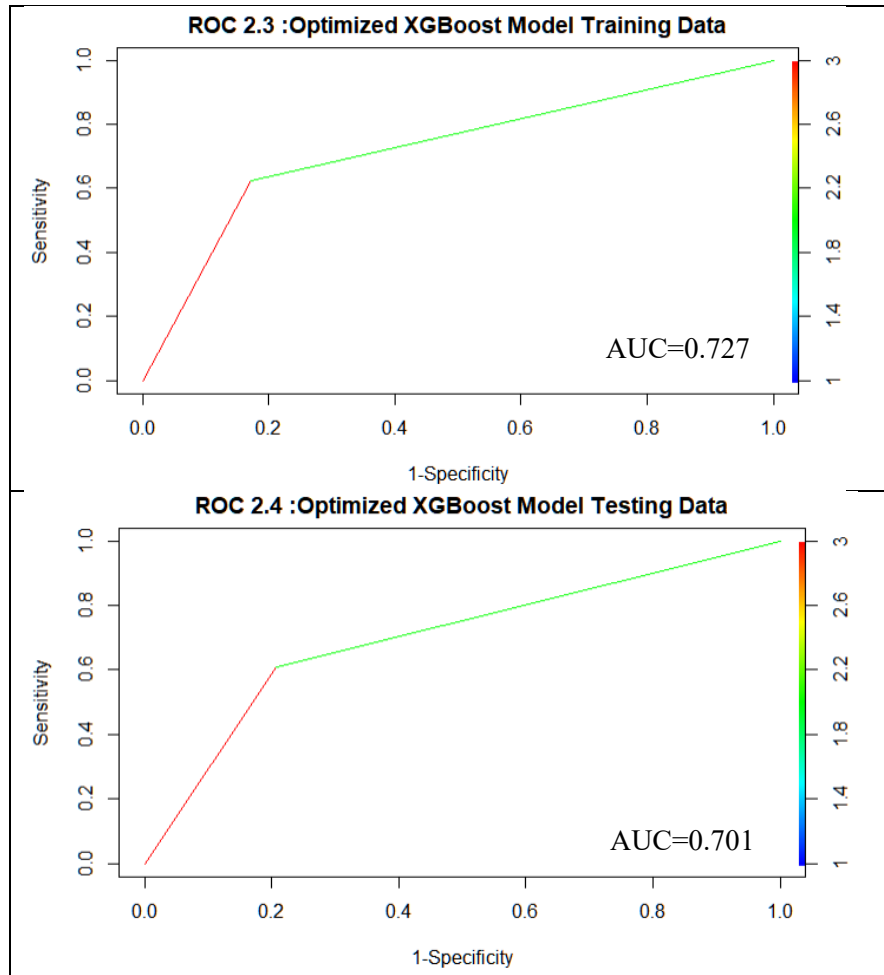


Figure 5.6: ROC curve for optimized XGBoost

5.3 ADAPTIVE BOOSTING (ADABOOST)

Two models are developed based on the AdaBoost algorithm. First model is a basic model that adopts a fixed and default hyperparameter setting while the second model undergoes hyperparameter optimization to identify best performing model.

5.3.1 Basic AdaBoost Model

The basic AdaBoost model is built using AdaBag method from caret. The following list the hyperparameters along with their description. These hyperparameter values are the default values and will be fixed when fitting the model. Figure 5.7 shows a snippet of basic AdaBoost model building and fitting to the training and testing set.

- 'mfinal': Number of trees = 50

- ‘maxdepth’: Maximum tree depth = 1

```
grid = expand.grid(mfinal=50, maxdepth=1)
cls4 = train(Reached.on.Time_Y.N~., data=trainSet, method="AdaBag", tuneGrid = grid, metric='Accuracy')
# Training data
predTrain4_1 = predict(cls4,trainSet)
# Testing data
predTest4_1 = predict(cls4,testSet)
```

Figure 5.7: Snippet of basic AdaBoost model building and fitting

5.3.2 Optimized AdaBoost Model

The optimized Adaboost model is built using AdaBag method from caret. The optimization of the AdaBoost model is performed using grid search method on a range of hyperparameter values. In addition, 10-fold repeated CV is applied to the fitting model during the hyperparameter optimization process. The hyperparameter value ranges for the grid search are listed as followed:

- ‘mfinal’: from 50 to 500 with step 50
- ‘maxdepth’: from 1 to 10 with step 1

Based on the range of hyperparameter values, 100 combinations of hyperparameter value settings are generated which the model would iterate through to identify the best performing hyperparameter value combination. Results from each iteration will be attached under Appendix A. Following that, an AdaBoost model will be built upon the best performing hyperparameter value combination. Figure 5.8 shows a snippet of AdaBoost model with grid search of the selected hyperparameter ranges and fitted to the training and testing set. The best performing hyperparameter combination identified are as followed:

- ‘mfinal’: 200
- ‘maxdepth’: 5

```

tc4_2 = trainControl(method="repeatedcv", number=10, repeats=3, allowParallel = T)
metric = "Accuracy"
set.seed(globalSeed)

grid=expand.grid(mfinal = seq(50,500,50), maxdepth=seq(1,10,1))
cls4_2 = train(Reached.on.Time_Y.N~., data=trainSet, method="AdaBag", trControl=tc4_2, tuneGrid =grid )
# Training set
predTrain4_2 = predict(cls4_2, trainSet)
# Testing set
predTest4_2 = predict(cls4_2, testSet)

```

Figure 5.8: Snippet of optimized AdaBoost model building and fitting

5.3.3 AdaBoost Model Performance

The model performance before and after applying hyperparameter optimization for the AdaBoost models will be discussed in this section. Table 5.5 shows the confusion matrix for both the AdaBoost models at training and testing stage. The results are used to calculate the evaluation metrics. Table 5.6 shows the evaluation metrics computed based on the confusion matrix for the AdaBoost models.

Table 5.5: Confusion matrix for AdaBoost models

Base Model	Train Set			Test Set		
	Reference Prediction	On Time	Not on Time	Reference Prediction	On Time	Not on Time
	On Time	3540	3148	On Time	889	770
	Not on Time	23	2088	Not on Time	2	539
Optimized Model	Reference Prediction	On Time	Not on Time	Reference Prediction	On Time	Not on Time
	On Time	3342	2480	On Time	833	606
	Not on Time	221	2756	Not on Time	58	703

Table 5.6: Evaluation metric results for AdaBoost models

	Base Model		Optimized Model	
	Train Set	Test Set	Train Set	Test Set
Accuracy	0.6396	0.6491	0.6930	0.6982
Kappa	0.3441	0.3600	0.4216	0.4298
Sensitivity	0.9935	0.9978	0.9380	0.9349
Specificity	0.3988	0.4118	0.5264	0.5371

Based on Table 5.6, the following observations are identified and listed as followed:

- The accuracy of the optimized model is higher as compared to the base model which obtained 69.82% and 64.91% respectively. This indicates the hyperparameter optimization has obtained a better performance model.
- The accuracy of the testing set is higher than the training set for both the models which indicates model is slightly underfitting.
- Kappa for both the models have fair agreement in the predictive performance.
- Sensitivity for both the models obtained near perfect score which indicates outstanding accuracy of the model in correctly classifying positive outcomes.
- Specificity for both the models obtained a moderate score which indicates moderate accuracy of the model in correctly classifying negative outcomes.

Figure 5.9 shows the comparison of ROC curve for the basic AdaBoost model in the training and testing data. The area under the ROC curve is computed and labeled in each diagram. It is observed that the testing data area under ROC is 0.705 which is slightly higher than the training data of 0.696. This indicates model is slightly underfitting but an acceptable performance of the model in identifying the positive and negative class.

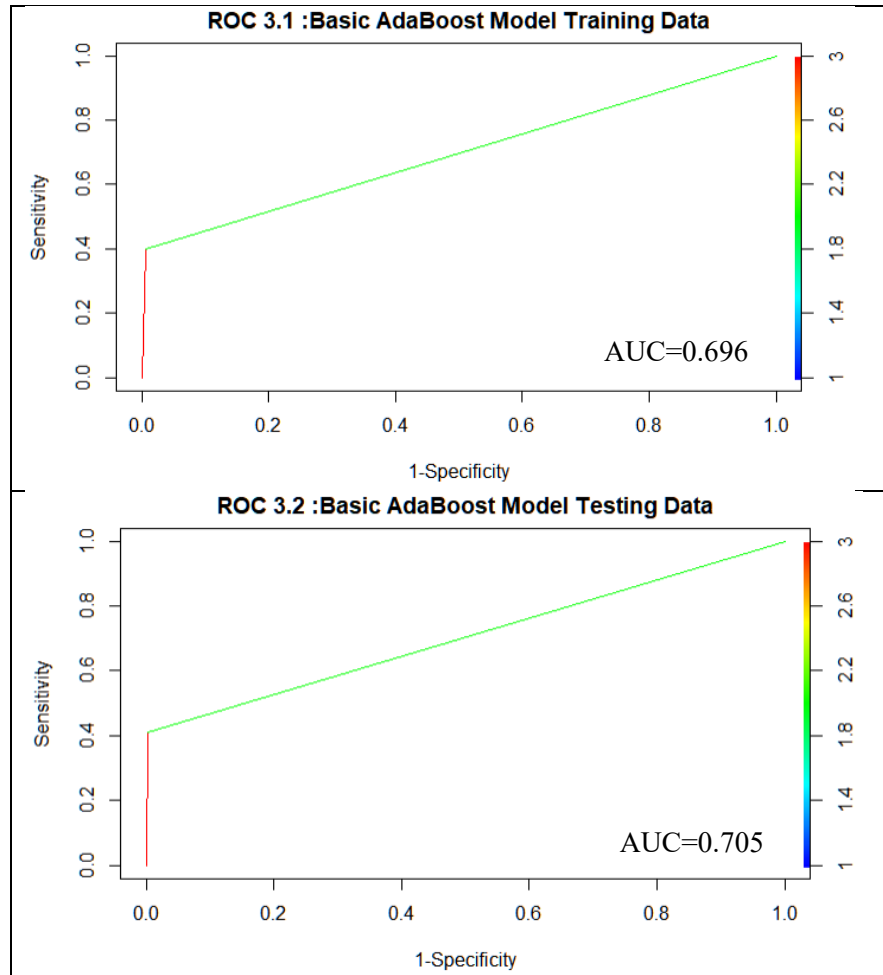


Figure 5.9: ROC curve for basic AdaBoost

Figure 5.10 shows the comparison of ROC curve for the optimized AdaBoost model in the training and testing data. The area under the ROC curve is computed and labeled in each diagram. It is observed that the testing data area under ROC is 0.736 which is just few points higher than the training data of 0.732. This indicates model is slightly underfitting but an acceptable performance of the model in identifying the positive and negative class.

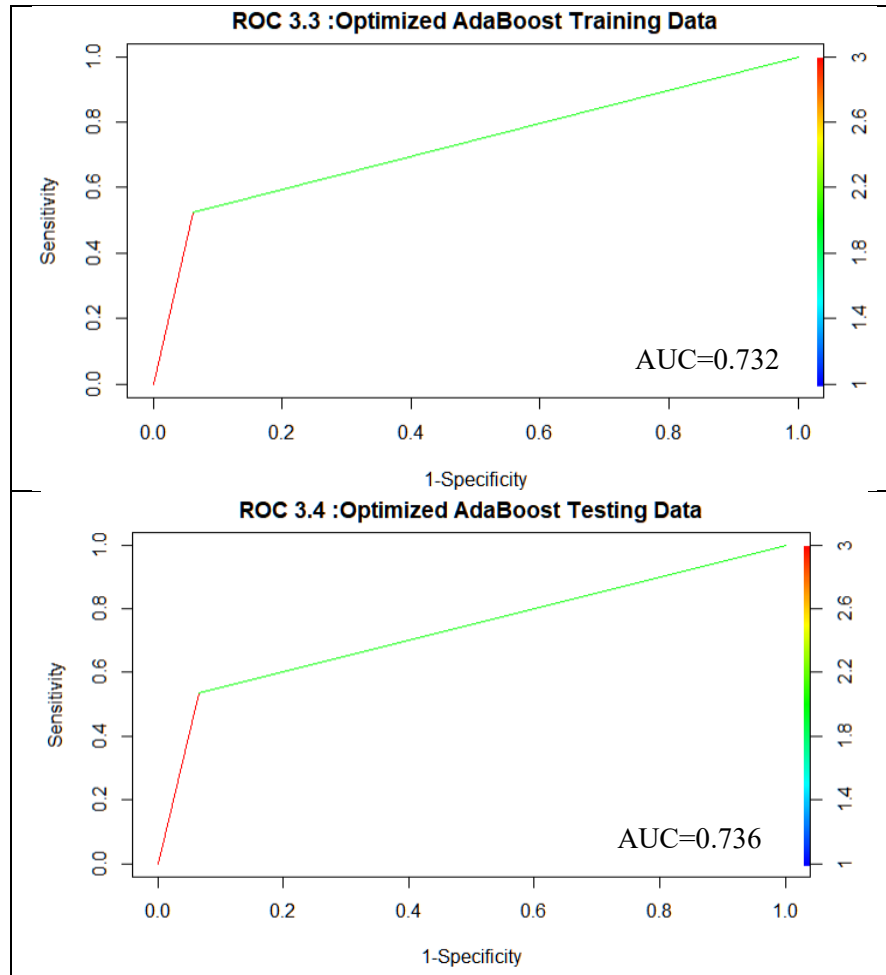


Figure 5.10: ROC curve for optimized AdaBoost

5.4 RANDOM FOREST

Two models are developed based on the Random Forest algorithm. First model is a basic model that adopts a fixed and default hyperparameter setting while the second model undergoes hyperparameter optimization to identify best performing model.

5.4.1 Basic Random Forest Model

The basic RF model is built using `rf` method from `caret`. The following list the hyperparameter along with the description. The hyperparameter value is the default value and will be fixed when fitting the model. Figure 5.11 shows a snippet of basic RF model building and fitting to the training and testing set.

- 'mtry': Random sampling of number of variables = 1

```

grid = expand.grid(mtry=1)
cls5 = train(Reached.on.Time_Y.N~., data=trainSet, method="rf", metric='Accuracy', tuneGrid=grid)
# Training set
predTrain5_1 = predict(cls5, trainSet)
# Testing set
predTest5_1 = predict(cls5, testSet)

```

Figure 5.11: Snippet of basic Random Forest model building and fitting

5.4.2 Optimized Random Forest Model

The optimized RF model is built using rf method from caret. The optimization of the RF model is performed using grid search method on a range of hyperparameter values. In addition, 10-fold repeated CV is applied to the fitting model during hyperparameter optimization process. The hyperparameter value range for the grid search is listed as followed:

- 'mtry': from 1 to 17 with step 1

Based on the range of hyperparameter values, 17 combinations of hyperparameter value settings are generated which the model would iterate through to identify the best performing hyperparameter value. Results from each iteration will be attached under Appendix A. Following that, a RF model will be built upon the best performing hyperparameter. Figure 5.12 shows a snippet of RF model with grid search of the selected hyperparameter ranges and fitted to the training and testing set. The best performing hyperparameter identified is as followed:

- 'mtry': 2

```

tc5_2 = trainControl(method="repeatedcv", number=10, repeats=3)
set.seed(globalSeed)
tuneGrid = expand.grid(mtry=seq(1:17))

cls5_2 = train(Reached.on.Time_Y.N~., data=trainSet, method="rf", metric='Accuracy', tuneGrid=tuneGrid, trControl=tc5_2)
# Training set
predTrain5_2 = predict(cls5_2, trainSet)
# Testing set
predTest5_2 = predict(cls5_2, testSet)

```

Figure 5.12: Snippet of optimized Random Forest model building and fitting

5.4.3 Random Forest Model Performance

The model performance before and after applying hyperparameter optimization for the RF models will be discussed in this section. Table 5.7 shows the confusion matrix for both the RF models at

training and testing stage. The results are used to calculate the evaluation metrics. Table 5.8 shows the evaluation metrics computed based on the confusion matrix for the RF models.

Table 5.7: Confusion matrix for RF models

	Train Set			Test Set		
Base Model	Reference Prediction	On Time	Not on Time	Reference Prediction	On Time	Not on Time
	On Time	595	86	On Time	81	52
	Not on Time	2968	5150	Not on Time	810	1257
Optimized Model	Reference Prediction	On Time	Not on Time	Reference Prediction	On Time	Not on Time
	On Time	3323	949	On Time	683	480
	Not on Time	240	4287	Not on Time	208	829

Table 5.8: Evaluation metric results for Random Forest models

	Base Model		Optimized Model	
	Train Set	Test Set	Train Set	Test Set
Accuracy	0.6529	0.6082	0.8649	0.6873
Kappa	0.1729	0.0592	0.7282	0.3813
Sensitivity	0.1670	0.0909	0.9326	0.7666
Specificity	0.9836	0.9603	0.8188	0.6333

Based on Table 5.8, the following observations are identified and listed as followed:

- The accuracy of the optimized model is higher as compared to the base model which obtained 68.73% and 60.82% respectively. This indicates the hyperparameter optimization has obtained a better performance model.
- The accuracy of the testing set is slightly lower than the training set for both the models which indicates model is slightly overfitting.
- Kappa for the base model obtained a poor agreement in the predictive performance. However, for the optimized model the kappa obtained a fair agreement in the predictive performance.
- Sensitivity for the base model is extremely poor which indicates poor accuracy for the model in correctly classifying positive outcomes. However, in the optimized model the sensitivity is high which indicates model is good at correctly classifying positive outcomes.

- Specificity for the base model is near perfect which indicates very high accuracy for the model in correctly classifying negative outcomes. However, in the optimized model the specificity is moderate which indicates moderate accuracy of the model in correctly classifying negative outcomes.

Figure 5.13 shows the comparison of ROC curve for the basic RF model in the training and testing data. The area under the ROC curve is computed and labeled in each diagram. It is observed that the training data area under ROC is 0.575 which is slightly higher than the testing data of 0.526. This indicates model is a good fit, but the model is showing weak performance in identifying the positive and negative class.

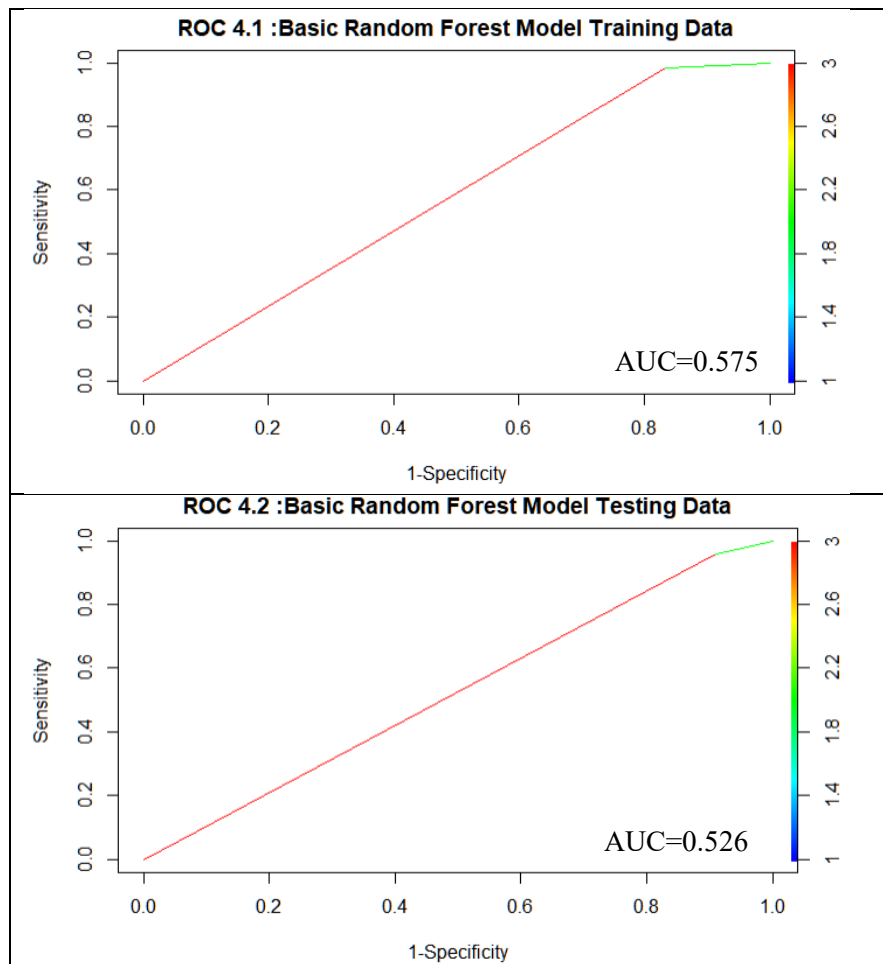


Figure 5.13: ROC curve for the basic Random Forest

Figure 5.14 shows the comparison of ROC curve for the optimized RF model in the training and testing data. The area under the ROC curve is computed and labeled in each diagram. It is observed that the training data area under ROC is 0.876 which is slightly higher than the testing data of 0.701. This indicates model has a good fit and provides acceptable performance in identifying the positive and negative class.

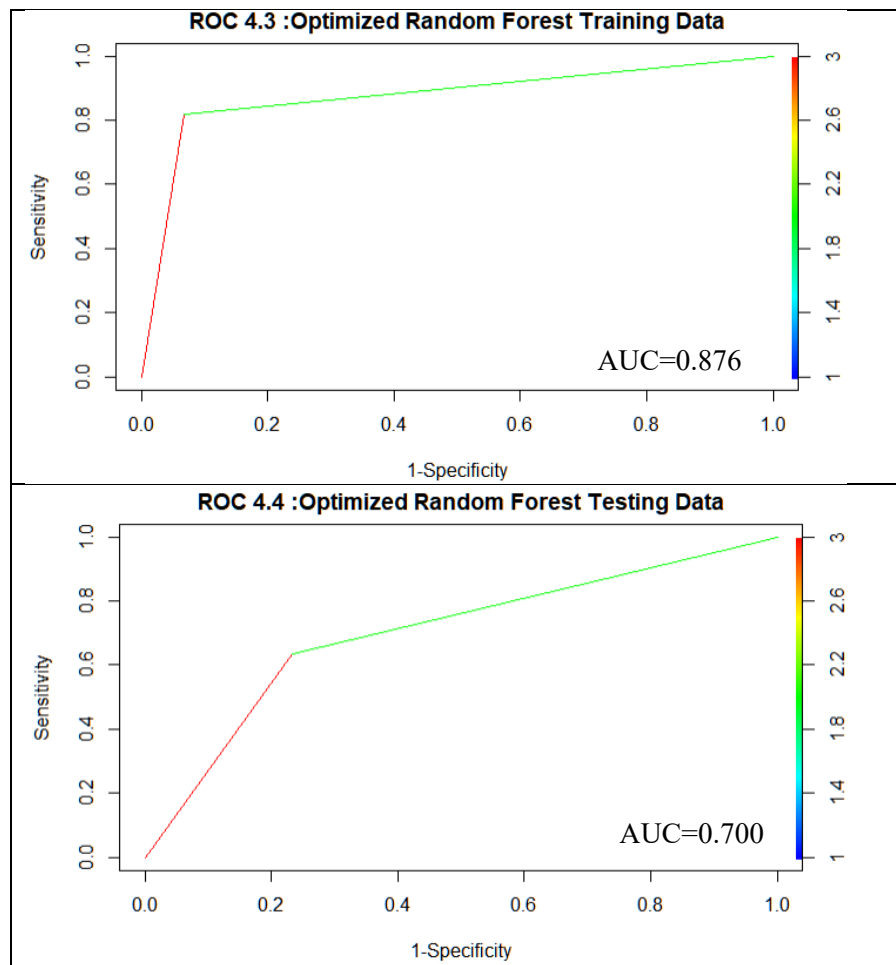


Figure 5.14: ROC curve for the optimized Random Forest

SECTION 6

ANALYSIS & RECOMMENDATIONS

6.1 RESULT ANALYSIS

Table 6.1: Evaluation results compilation of predictive models

Models	Hyperparameters & resampling		Accuracy	Kappa	Sensitivity	Specificity	AUROC
Logistic Regression	- Repeated 10-fold cv		0.6500	0.2772	0.5825	0.6960	0.6390
Basic XGBoost	- eta: 0.3	- colsample_bytree: 1	0.6745	0.3359	0.6510	0.6906	0.6710
	- gamma: 0	- min_child_weight: 1					
	- max_depth: 6	- nrounds: 100					
	- subsample: 1	- Repeated 10-fold cv					
Optimized XGBoost	- eta: 0.05	- colsample_bytree: 0.5	0.6832	0.3790	0.7924	0.6089	0.701
	- gamma: 0.6	- min_child_weight: 15					
	- max_depth: 5	- nrounds: 50					
	- subsample: 0.5	- Repeated 10-fold cv					
Basic AdaBoost	- mfinal: 50	- Repeated 10-fold cv	0.6491	0.3600	0.9978	0.4118	0.7050
	- maxdepth: 1						
Optimized AdaBoost	- mfinal: 200	- Repeated 10-fold cv	0.6982	0.4298	0.9349	0.5371	0.7360
	- maxdepth: 5						
Basic Random Forest	- mtry: 1		0.6082	0.0592	0.0909	0.9603	0.5260
Optimized Random Forest	- mtry: 2	- Repeated 10-fold CV	0.6873	0.3813	0.7666	0.6333	0.7000

Table 6.1 shows the compilation of all evaluation metric results for all the developed models in this study. In addition, the hyperparameters and resampling settings on each model is presented. Model with the best result in each evaluation metric is highlighted in bold as shown in the table. Based on Table 6.1, the following observations are identified and listed as followed:

- In terms of accuracy, optimized AdaBoost model obtained the highest accuracy of 69.82% which indicates the model provides good performance in ratio of correctly identifying the prediction as compared to the total observations.
- In terms of kappa, optimized AdaBoost model obtained the highest kappa of 0.4298 which indicates the model provides moderate agreement in the predictive performance.
- In terms of sensitivity, basic AdaBoost model obtained the highest sensitivity of 0.9978 which indicates the model provides a near perfect score at correctly classifying positive outcomes.
- In terms of specificity, optimized AdaBoost model obtained the highest specificity of 0.7360 which indicates the model provides a good score at correctly classifying negative outcomes.
- In terms of AUROC, optimized AdaBoost model obtained the highest area of 0.7360 which indicates the model is very likely to be able to distinguish between the positive and negative class.

In overall, the optimized AdaBoost model has obtained the highest score in four of the five evaluation metrics. Therefore, it can be concluded that the optimized AdaBoost model is the best performing model in this study.

6.2 RESULTS COMPARISON

This section discusses the comparison of results from this study to peers that have utilized the same dataset. Table 6.2 shows the compilation of results from literature using different models to predict the package delivery time. In addition, models from this study will be included in the table as a comparison in model performance.

Table 6.2: Model performance of peers from the same dataset

No.	Author	Prediction Model	Accuracy (%)
1	From this study	AdaBoost	69.82
2	From this study	RF	68.73
3	From this study	XGBoost	68.32
4	(Kumar, 2021)	ANN	67.00
5	(Kumar, 2021)	RF	66.00
6	(Kumar, 2021)	SVM	66.00
7	(Chandra, 2021)	RF	65.27
8	(Bagga, 2021)	KNN	65.00
9	(Bagga, 2021)	Logistic regression	65.00
10	(Chandra, 2021)	KNN	63.00

Based on Table 6.2, the best model performance from literature is by Kumar (2021) which uses ANN model and achieved an accuracy of 67%. However, all three prediction models from this study have outperformed the ANN model by at least 1.32% in terms of accuracy. This may be due to Kumar (2021) on not performing hyperparameter optimization on the ANN model. Although extensive hyperparameter optimization is performed on the models from this study, the model performance was just slightly better than the ANN model. Therefore, with proper hyperparameter tuning on the ANN model, better performance might be expected.

In specific comparison of the RF model result from this study to literature. It is observed that RF model from this study has outperformed the RF models from literature. In addition, it is observed that RF model was used by Kumar (2021) and Chandra (2021) which achieved an accuracy of 67% and 65.27% respectively. On further investigation into the works of Chandra (2021). The author did not perform any hyperparameter tuning on the model instead only relying on the basic RF model for the prediction. However, in the works of Kumar (2021), some hyperparameter tuning was performed on the number of trees parameter and the split quality criterion parameter. The parameter used in tuning the RF model in this study was the number of random sampling variables. This might indicate that the number of random sampling variables has a higher effect on the model performance.

SECTION 7

CONCLUSION

The increased package delivery frequency from the rapid expansion of e-commerce led to the generation of voluminous data which can be utilized to extract information such as predicting the delivery time of the packages. With the advancement of technologies, big data analytics has opened more doors to companies trying to gain value from all sorts of data available at hand. Which in this study, four machine learning algorithms are explored, and multiple prediction models are developed in trying to accurately predict the package delivery time. The algorithms involve are Logistic Regression, XGBoost, AdaBoost, and Random Forest. The best performing model is identified to be the optimized AdaBoost. In which, being able to accurately predict package delivery time enhances the operation efficiency and customer retention rate. The models in this study have outperformed the models from literature. However, the accuracy of the models can be further improved if larger dataset and features about weather condition can be incorporated in the prediction models. In addition, hyperparameter optimization can be further tuned to achieve even higher accuracy. Furthermore, ensembles of different algorithms can be explored to examine the effects of combining different models on the prediction accuracy.

REFERENCES

- Ahmed, I., Weerasingha, I., RESHADAT, V., & BROEDERS, S. (2021). *Travel Time Prediction and Explanation*. (Master), Eindhoven University of Technology, Retrieved from <http://dx.doi.org/10.3390/electronics11010106>
- Bagga, S. (2021). *Insights into shipping data*. 1. Retrieved from <https://www.kaggle.com/code/sakshibagga/insights-into-shipping-data>
- Chandra, M. G. (2021). *Commerce Analytics KNN, Random Forest*. 5. Retrieved from <https://www.kaggle.com/code/gaganmaahi224/commerce-analystics-knn-random-forest>
- Chauhan, N. S. (2020). *Model Evaluation Metrics in Machine Learning*. Retrieved from <https://www.kdnuggets.com/2020/05/model-evaluation-metrics-machine-learning.html>
- Hildebrandt, F. D., & Ulmer, M. W. (2021). Supervised learning for arrival time estimations in restaurant meal delivery. *Transportation Science*.
- Khiari, J., & Olaverri-Monreal, C. (2020). *Boosting algorithms for delivery time prediction in transportation logistics*. Paper presented at the 2020 International Conference on Data Mining Workshops (ICDMW).
- Kumar, T. S. (2021). *Ecommerce Shipping - EDA and Prediction*. 3. Retrieved from <https://www.kaggle.com/code/santhoshtsk/ecommerce-shipping-eda-prediction>
- Magiya, J. (2020). *Predicting Package Delivery Time For Motorcycles In Nairobi*. (Master), KCA University, Retrieved from <http://dx.doi.org/10.13140/RG.2.2.27105.94567>
- Ogura, T., Inoue, T., & Uchihira, N. (2021). Prediction of Arrival Time of Vessels Considering Future Weather Conditions. *Applied Sciences*, 11(10), 4410. doi:<http://dx.doi.org/10.3390/app11104410>
- Servos, N., Liu, X., Teucke, M., & Freitag, M. (2019). Travel time prediction in a multimodal freight transport relation using machine learning algorithms. *Logistics*, 4(1), 1. doi:<https://doi.org/10.3390/logistics4010001>
- Sundaram, N., & Tallamraju, R. B. (2021). Online Fashion Commerce: Modelling Customer Promise Date. *arXiv preprint*. doi:<https://doi.org/10.48550/arXiv.2105.00315>
- Wu, F., & Wu, L. (2019). *DeepETA: a spatial-temporal sequential neural network model for estimating time of arrival in package delivery system*. Paper presented at the Proceedings of the AAAI Conference on Artificial Intelligence.

APPENDIX A

MODEL OPTIMIZATION RESULTS

AdaBoost Optimization Iterative Results

maxdepth	mfinal	Accuracy	Kappa
1	50	0.639617	0.344154
1	100	0.639617	0.344154
1	150	0.639617	0.344154
1	200	0.639617	0.344154
1	250	0.639617	0.344154
1	300	0.639617	0.344154
1	350	0.639617	0.344154
1	400	0.639617	0.344154
1	450	0.639617	0.344154
1	500	0.639617	0.344154
2	50	0.680415	0.360929
2	100	0.680302	0.360802
2	150	0.680302	0.360729
2	200	0.680264	0.360662
2	250	0.68034	0.360795
2	300	0.68034	0.360795
2	350	0.68034	0.360795
2	400	0.680377	0.360881
2	450	0.68034	0.360795
2	500	0.68034	0.360795
3	50	0.680453	0.361072
3	100	0.680188	0.360546
3	150	0.680264	0.360699
3	200	0.680226	0.360632
3	250	0.680188	0.360563
3	300	0.680074	0.360362
3	350	0.68015	0.360478
3	400	0.680302	0.360745
3	450	0.680226	0.360612
3	500	0.68034	0.360814

maxdepth	mfinal	Accuracy	Kappa
4	50	0.684737	0.404116
4	100	0.685343	0.405683
4	150	0.685608	0.406093
4	200	0.685949	0.406588
4	250	0.685797	0.40627
4	300	0.685646	0.406085
4	350	0.685077	0.405001
4	400	0.685494	0.405847
4	450	0.685456	0.405737
4	500	0.685343	0.405564
5	50	0.686251	0.40831
5	100	0.686214	0.408535
5	150	0.686479	0.409194
5	200	0.686554	0.409335
5	250	0.686516	0.409241
5	300	0.686327	0.40886
5	350	0.686441	0.409067
5	400	0.686327	0.408842
5	450	0.686175	0.408523
5	500	0.686138	0.408459
6	50	0.684773	0.402406
6	100	0.685758	0.404768
6	150	0.684736	0.402984
6	200	0.684812	0.403352
6	250	0.68466	0.403128
6	300	0.684357	0.402473
6	350	0.684584	0.403066
6	400	0.684281	0.402437
6	450	0.684281	0.402408
6	500	0.684281	0.402428

maxdepth	mfinal	Accuracy	Kappa
7	50	0.682159	0.393473
7	100	0.682653	0.395362
7	150	0.681137	0.393148
7	200	0.681894	0.394608
7	250	0.68216	0.395172
7	300	0.682917	0.396735
7	350	0.682576	0.396299
7	400	0.68269	0.396525
7	450	0.682614	0.396452
7	500	0.682235	0.395879
8	50	0.681324	0.390442
8	100	0.679924	0.38942
8	150	0.680113	0.389827
8	200	0.680341	0.390469
8	250	0.680758	0.391598
8	300	0.680758	0.391564
8	350	0.680303	0.39071
8	400	0.680266	0.390857
8	450	0.680379	0.390884
8	500	0.680531	0.391218
9	50	0.679206	0.385486
9	100	0.679054	0.386239
9	150	0.678901	0.386045
9	200	0.679129	0.38673
9	250	0.679128	0.386774
9	300	0.678977	0.386593
9	350	0.67928	0.3872
9	400	0.679015	0.386726
9	450	0.67875	0.386204
9	500	0.678939	0.386428
10	50	0.678674	0.382522
10	100	0.679812	0.385955
10	150	0.679205	0.385348
10	200	0.679129	0.385245
10	250	0.678825	0.384898
10	300	0.678825	0.385025
10	350	0.67856	0.384534
10	400	0.678749	0.384806
10	450	0.678711	0.38488
10	500	0.678976	0.385279

Random Forest Optimization Iterative Result

mtry	Accuracy	Kappa
1	0.600031	0.025266
2	0.676173	0.361281
3	0.671702	0.351499
4	0.666097	0.335035
5	0.665603	0.33257
6	0.663632	0.327403
7	0.663746	0.327098
8	0.663823	0.326875
9	0.662913	0.325542
10	0.663746	0.326856
11	0.663406	0.325998
12	0.662383	0.323208
13	0.663065	0.324945
14	0.662381	0.322834
15	0.662534	0.323317
16	0.66136	0.320838
17	0.662343	0.322937

XGBoost Optimization Iterative Result

#	eta	max_depth	gamma	colsample_bytree	min_child_weight	subsample	rounds	Accuracy	Kappa
1	0.05	5	0	0	1	0.5	50	0.680766842	0.377800541
2	0.05	5	0	0	3	0.5	50	0.678818323	0.37722468
3	0.05	5	0	0	5	0.5	50	0.680391449	0.376219199
4	0.05	5	0	0	7	0.5	50	0.679070264	0.376265497
5	0.05	5	0	0	9	0.5	50	0.680418154	0.377799919
6	0.05	5	0	0	11	0.5	50	0.679622296	0.375389116
7	0.05	5	0	0	13	0.5	50	0.681516294	0.376528205
8	0.05	5	0	0	15	0.5	50	0.679188024	0.374655629
9	0.05	5	0	0	17	0.5	50	0.680302911	0.382168893
10	0.05	5	0	0	19	0.5	50	0.681592784	0.380822905
11	0.05	5	0.1	0	1	0.5	50	0.680342223	0.375618126
12	0.05	5	0.1	0	3	0.5	50	0.681251067	0.376382456
13	0.05	5	0.1	0	5	0.5	50	0.679323327	0.375911916
14	0.05	5	0.1	0	7	0.5	50	0.678031185	0.372823381
15	0.05	5	0.1	0	9	0.5	50	0.68047172	0.377515195
16	0.05	5	0.1	0	11	0.5	50	0.678446951	0.37502452
17	0.05	5	0.1	0	13	0.5	50	0.678977208	0.375073117
18	0.05	5	0.1	0	15	0.5	50	0.68071972	0.378448435
19	0.05	5	0.1	0	17	0.5	50	0.682249421	0.385112389
20	0.05	5	0.1	0	19	0.5	50	0.681819024	0.381325037
21	0.05	5	0.2	0	1	0.5	50	0.678017154	0.374367784
22	0.05	5	0.2	0	3	0.5	50	0.680849017	0.378342262
23	0.05	5	0.2	0	5	0.5	50	0.681023785	0.378786909
24	0.05	5	0.2	0	7	0.5	50	0.68117659	0.378842357
25	0.05	5	0.2	0	9	0.5	50	0.680381265	0.377483578
26	0.05	5	0.2	0	11	0.5	50	0.681286272	0.378438979
27	0.05	5	0.2	0	13	0.5	50	0.67924279	0.375530086
28	0.05	5	0.2	0	15	0.5	50	0.679602133	0.376898352
29	0.05	5	0.2	0	17	0.5	50	0.680458912	0.378812317
30	0.05	5	0.2	0	19	0.5	50	0.681404251	0.380350555
31	0.05	5	0.3	0	1	0.5	50	0.679434981	0.375442789
32	0.05	5	0.3	0	3	0.5	50	0.681144108	0.378124118
33	0.05	5	0.3	0	5	0.5	50	0.678089184	0.375801205
34	0.05	5	0.3	0	7	0.5	50	0.681554259	0.378680909
35	0.05	5	0.3	0	9	0.5	50	0.682359415	0.381232317
36	0.05	5	0.3	0	11	0.5	50	0.679700091	0.377048337
37	0.05	5	0.3	0	13	0.5	50	0.683070747	0.382895748
38	0.05	5	0.3	0	15	0.5	50	0.681365555	0.378795192
39	0.05	5	0.3	0	17	0.5	50	0.683016228	0.382386327
40	0.05	5	0.3	0	19	0.5	50	0.680458933	0.378583752
41	0.05	5	0.4	0	1	0.5	50	0.679632633	0.378881138
42	0.05	5	0.4	0	3	0.5	50	0.679548869	0.376307115
43	0.05	5	0.4	0	5	0.5	50	0.679743886	0.375284882
44	0.05	5	0.4	0	7	0.5	50	0.67944811	0.373024225
45	0.05	5	0.4	0	9	0.5	50	0.679478308	0.373385111
46	0.05	5	0.4	0	11	0.5	50	0.680415897	0.376332789
47	0.05	5	0.4	0	13	0.5	50	0.680380916	0.377320389
48	0.05	5	0.4	0	15	0.5	50	0.680602735	0.378465144
49	0.05	5	0.4	0	17	0.5	50	0.680849188	0.378289557
50	0.05	5	0.4	0	19	0.5	50	0.679826299	0.375838004
51	0.05	5	0.5	0	1	0.5	50	0.678942634	0.376026014
52	0.05	5	0.5	0	3	0.5	50	0.680794848	0.377877299
53	0.05	5	0.5	0	5	0.5	50	0.678800918	0.372785804
54	0.05	5	0.5	0	7	0.5	50	0.680078235	0.376388496
55	0.05	5	0.5	0	9	0.5	50	0.681700764	0.378631142
56	0.05	5	0.5	0	11	0.5	50	0.678627622	0.374583034
57	0.05	5	0.5	0	13	0.5	50	0.680077625	0.378280979
58	0.05	5	0.5	0	15	0.5	50	0.680910146	0.378168953
59	0.05	5	0.5	0	17	0.5	50	0.682918874	0.382263893
60	0.05	5	0.5	0	19	0.5	50	0.682509712	0.382214478
61	0.05	5	0.6	0	1	0.5	50	0.682707137	0.380686293
62	0.05	5	0.6	0	3	0.5	50	0.680001788	0.376785004
63	0.05	5	0.6	0	5	0.5	50	0.677803784	0.372058889
64	0.05	5	0.6	0	7	0.5	50	0.678109142	0.374541889
65	0.05	5	0.6	0	9	0.5	50	0.68044825	0.377044415
66	0.05	5	0.6	0	11	0.5	50	0.678025255	0.378450847
67	0.05	5	0.6	0	13	0.5	50	0.678017136	0.376889751
68	0.05	5	0.6	0	15	0.5	50	0.683751486	0.384548866
69	0.05	5	0.6	0	17	0.5	50	0.680836321	0.378228222
70	0.05	5	0.6	0	19	0.5	50	0.680154838	0.37898347
71	0.05	5	0.7	0	1	0.5	50	0.678965556	0.375018146
72	0.05	5	0.7	0	3	0.5	50	0.678243437	0.374875908
73	0.05	5	0.7	0	5	0.5	50	0.680898009	0.378838933
74	0.05	5	0.7	0	7	0.5	50	0.679284721	0.375288112
75	0.05	5	0.7	0	9	0.5	50	0.678938814	0.374738178
76	0.05	5	0.7	0	11	0.5	50	0.678481878	0.373742229
77	0.05	5	0.7	0	13	0.5	50	0.682538893	0.381791948
78	0.05	5	0.7	0	15	0.5	50	0.680302653	0.384312411
79	0.05	5	0.7	0	17	0.5	50	0.68205421	0.382363992
80	0.05	5	0.7	0	19	0.5	50	0.681896289	0.380138438
81	0.05	5	0.8	0	1	0.5	50	0.682101354	0.381738871
82	0.05	5	0.8	0	3	0.5	50	0.679358192	0.375072577
83	0.05	5	0.8	0	5	0.5	50	0.68019635	0.376428156
84	0.05	5	0.8	0	7	0.5	50	0.678271288	0.372517889
85	0.05	5	0.8	0	9	0.5	50	0.680383651	0.376887396
86	0.05	5	0.8	0	11	0.5	50	0.680838134	0.377681554
87	0.05	5	0.8	0	13	0.5	50	0.682088884	0.380088863
88	0.05	5	0.8	0	15	0.5	50	0.680154164	0.377891191
89	0.05	5	0.8	0	17	0.5	50	0.682614857	0.381232537
90	0.05	5	0.8	0	19	0.5	50	0.681743399	0.378888188
91	0.05	5	0.9	0	1	0.5	50	0.679434551	0.375320711
92	0.05	5	0.9	0	3	0.5	50	0.680038288	0.377185891
93	0.05	5	0.9	0	5	0.5	50	0.678198293	0.374468973
94	0.05	5	0.9	0	7	0.5	50	0.68181854	0.381388888
95	0.05	5	0.9	0	9	0.5	50	0.681781617	0.380093183
96	0.05	5	0.9	0	11	0.5	50	0.680495738	0.377828833
97	0.05	5	0.9	0	13	0.5	50	0.680898987	0.378148881
98	0.05	5	0.9	0	15	0.5	50	0.678561834	0.372517179
99	0.05	5	0.9	0	17	0.5	50	0.683372873	0.383090881
100	0.05	5	0.9	0	19	0.5	50	0.680114318	0.378385238
101	0.05	5	1	0	1	0.5	50	0.681077498	0.380027556
102	0.05	5	1	0	3	0.5	50	0.680203098	0.37654312
103	0.05	5	1	0	5	0.5	50	0.680203677	0.374104113
104	0.05	5	1	0	7	0.5	50	0.681781819	0.378877325
105	0.05	5	1	0	9	0.5	50	0.680267388	0.377444421
106	0.05	5	1	0	11	0.5	50	0.678198871	0.374285571
107	0.05	5	1	0	13	0.5	50	0.680418068	0.378595028
108	0.05	5	1	0	15	0.5	50	0.681593131	0.37897208
109	0.05	5	1	0	17	0.5	50	0.680077882	0.37746242
110	0.05	5	1	0	19	0.5	50	0.682121888	0.380771188
111	0.05	5	0	0	1	0.5	50	0.679357848	0.372180831
112	0.05	5	0	0	3	0.5	50	0.67898314	0.370446383
113	0.05	5	0	0	5	0.5	50	0.67728115	0.369811388
114	0.05	5	0	0	7	0.5	50	0.678488807	0.370481179
115	0.05	5	0	0	9	0.5	50	0.678408767	0.370101475
116	0.05	5	0	0	11	0.5	50	0.680003677	0.374104113
117	0.05	5	0	0	13	0.5	50	0.675153515	0.363881522
118	0.05	5	0	0	15	0.5	50	0.679773696	0.373811159
119	0.05	5	0	0	17	0.5	50	0.677804981	0.36983747
120	0.05	5	0	0	19	0.5	50	0.675008849	0.373097983
121	0.05	5	0.1	0	1	0.5	50	0.678334131	0.36921612
122	0.05	5	0.1	0	3	0.5	50	0.678978154	0.370151921
123	0.05	5	0.1	0	5	0.5	50	0.678143701	0.368812202
124	0.05	5	0.1	0	7	0.5	50	0.678384645	0.368388864
125	0.05	5	0.1	0	9	0.5	50	0.678537884	0.364838336
126	0.05	5	0.1	0	11	0.5	50	0.678383786	0.368888888
127	0.05	5	0.1	0	13	0.5	50	0.680040011	0.374300042
128	0.05	5	0.1	0	15	0.5	50	0.678433217	0.372821062
129	0.05	5	0.1	0	17	0.5	50	0.680238458	0.376183726
130	0.05	5	0.1	0	19	0.5	50	0.678678287	0.372438208
131	0.05	5	0.2	0	1	0.5	50	0.678797194	0.369397223
132	0.05	5	0.2	0	3	0.5</			

#	sta	max	depth	alpha	colsample	bytes	min	child	weight	subsample	rounds	Accuracy	Kappa
302	0.05	7	0.8	0.3	1	0.5	50	0.67847827	0.36217455				
303	0.05	7	0.8	0.3	1	0.5	50	0.67427481	0.35941272				
303	0.05	7	0.8	0.3	0.5	0.5	50	0.674659026	0.359071908				
304	0.05	7	0.8	0.3	0.5	0.5	50	0.67444345	0.35610098				
305	0.05	7	0.8	0.3	0.5	0.5	50	0.67451533	0.360786371				
306	0.05	7	0.8	0.3	0.5	0.5	11	0.5	0.675251913	0.36351743			
307	0.05	7	0.8	0.3	0.3	0.5	50	0.676441481	0.364478568				
308	0.05	7	0.8	0.3	0.3	0.5	50	0.67488587	0.368937362				
309	0.05	7	0.8	0.3	0.3	0.5	17	0.5	0.67728794	0.3674703345			
310	0.05	7	0.8	0.3	0.5	0.5	19	0.5	0.67802406	0.368635169			
311	0.05	7	0.8	0.3	0.5	0.5	1	0.5	0.678071645	0.354253058			
312	0.05	7	0.8	0.3	0.5	0.5	50	0.675929268	0.36059418				
313	0.05	7	0.9	0.5	0.5	0.5	50	0.672727245	0.355237905				
314	0.05	7	0.9	0.5	0.5	0.5	7	0.5	0.676250838	0.361967647			
315	0.05	7	0.9	0.5	0.3	0.5	50	0.677046334	0.364989437				
316	0.05	7	0.9	0.5	0.3	0.5	11	0.5	0.678109483	0.36900664			
317	0.05	7	0.9	0.5	0.3	0.5	50	0.67893063	0.369483556				
318	0.05	7	0.9	0.5	0.5	0.5	50	0.681172985	0.373818468				
319	0.05	7	0.9	0.5	0.5	0.5	17	0.5	0.675001287	0.362070391			
320	0.05	7	0.9	0.5	0.5	0.5	19	0.5	0.675027085	0.37190209			
321	0.05	7	1	0.5	0.5	0.5	1	0.5	0.673106379	0.354408749			
322	0.05	7	1	0.5	0.5	0.5	3	0.5	0.672993303	0.354761923			
323	0.05	7	1	0.5	0.5	0.5	19	0.5	0.67816887	0.364115098			
324	0.05	7	1	0.5	0.5	0.5	7	0.5	0.673750317	0.35744209			
325	0.05	7	1	0.5	0.5	0.5	0.5	0.5	0.67755175	0.366035633			
326	0.05	7	1	0.5	0.5	0.5	11	0.5	0.678666924	0.371187697			
327	0.05	7	1	0.5	0.5	0.5	13	0.5	0.675301757	0.360774695			
328	0.05	7	1	0.5	0.5	0.5	19	0.5	0.676439285	0.364674937			
329	0.05	7	1	0.5	0.5	0.5	17	0.5	0.67831359	0.368389303			
330	0.05	7	1	0.5	0.5	0.5	50	0.67761611	0.368793126				
331	0.05	8	0	0.5	0.5	0.5	1	0.5	0.670040648	0.344434112			
332	0.05	8	0	0.5	0.5	0.5	3	0.5	0.674212188	0.347687865			
333	0.05	8	0	0.5	0.5	0.5	50	0.67552453	0.349702014				
334	0.05	8	0	0.5	0.5	0.5	7	0.5	0.67597943	0.353855563			
335	0.05	8	0	0.5	0.5	0.5	9	0.5	0.674548872	0.357189492			
336	0.05	8	0	0.5	0.5	0.5	11	0.5	0.67527408	0.360165439			
337	0.05	8	0	0.5	0.5	0.5	13	0.5	0.67628948	0.361880154			
338	0.05	8	0	0.5	0.5	0.5	15	0.5	0.674657561	0.359446078			
339	0.05	8	0	0.5	0.5	0.5	17	0.5	0.67609963	0.361134428			
340	0.05	8	0	0.5	0.5	0.5	50	0.678933347	0.365477236				
341	0.05	8	0.1	0.5	0.5	0.5	1	0.5	0.668867222	0.343206224			
342	0.05	8	0.1	0.5	0.5	0.5	3	0.5	0.671744075	0.350035402			
343	0.05	8	0.1	0.5	0.5	0.5	50	0.67847475	0.348157575				
344	0.05	8	0.1	0.5	0.5	0.5	7	0.5	0.676787023	0.353512113			
345	0.05	8	0.1	0.5	0.5	0.5	9	0.5	0.672765212	0.353419368			
346	0.05	8	0.1	0.5	0.5	0.5	11	0.5	0.674244303	0.353121118			
347	0.05	8	0.1	0.5	0.5	0.5	13	0.5	0.672616198	0.3555212915			
348	0.05	8	0.1	0.5	0.5	0.5	15	0.5	0.676346903	0.362927322			
349	0.05	8	0.1	0.5	0.5	0.5	17	0.5	0.676101474	0.361508344			
350	0.05	8	0.1	0.5	0.5	0.5	19	0.5	0.675307946	0.361724169			
351	0.05	8	0.2	0.5	0.5	0.5	1	0.5	0.66602719	0.344092102			
352	0.05	8	0.2	0.5	0.5	0.5	3	0.5	0.672237921	0.349872817			
353	0.05	8	0.2	0.5	0.5	0.5	5	0.5	0.67409474	0.349670944			
354	0.05	8	0.2	0.5	0.5	0.5	7	0.5	0.676454332	0.348934480			
355	0.05	8	0.2	0.5	0.5	0.5	9	0.5	0.673638705	0.355210288			
356	0.05	8	0.2	0.5	0.5	0.5	11	0.5	0.675341788	0.359105738			
357	0.05	8	0.2	0.5	0.5	0.5	13	0.5	0.674393873	0.358166479			
358	0.05	8	0.2	0.5	0.5	0.5	15	0.5	0.67655528	0.363461721			
359	0.05	8	0.2	0.5	0.5	0.5	17	0.5	0.676782862	0.363544756			
360	0.05	8	0.2	0.5	0.5	0.5	19	0.5	0.677272995	0.365599615			
361	0.05	8	0.3	0.5	0.5	0.5	1	0.5	0.669777018	0.344622594			
362	0.05	8	0.3	0.5	0.5	0.5	3	0.5	0.670645673	0.346319832			
363	0.05	8	0.3	0.5	0.5	0.5	5	0.5	0.672122519	0.350208344			
364	0.05	8	0.3	0.5	0.5	0.5	7	0.5	0.67265377	0.352932488			
365	0.05	8	0.3	0.5	0.5	0.5	9	0.5	0.675228927	0.359786182			
366	0.05	8	0.3	0.5	0.5	0.5	11	0.5	0.676628086	0.363191203			
367	0.05	8	0.3	0.5	0.5	0.5	13	0.5	0.674399022	0.358733503			
368	0.05	8	0.3	0.5	0.5	0.5	15	0.5	0.674304122	0.361416813			
369	0.05	8	0.3	0.5	0.5	0.5	17	0.5	0.675455098	0.362636782			
370	0.05	8	0.3	0.5	0.5	0.5	19	0.5	0.675419269	0.361938289			
371	0.05	8	0.4	0.5	0.5	0.5	1	0.5	0.671553035	0.348343848			
372	0.05	8	0.4	0.5	0.5	0.5	3	0.5	0.67178354	0.349877767			
373	0.05	8	0.4	0.5	0.5	0.5	5	0.5	0.67676593	0.356179396			
374	0.05	8	0.4	0.5	0.5	0.5	7	0.5	0.67303263	0.34786193			
375	0.05	8	0.4	0.5	0.5	0.5	9	0.5	0.676767631	0.353143464			
376	0.05	8	0.4	0.5	0.5	0.5	11	0.5	0.676385852	0.361631188			
377	0.05	8	0.4	0.5	0.5	0.5	13	0.5	0.67772879	0.364593663			
378	0.05	8	0.4	0.5	0.5	0.5	15	0.5	0.675327172	0.361846556			
379	0.05	8	0.4	0.5	0.5	0.5	17	0.5	0.67656123	0.361240161			
380	0.05	8	0.4	0.5	0.5	0.5	19	0.5	0.677388667	0.365034449			
381	0.05	8	0.5	0.5	0.5	0.5	1	0.5	0.67183385	0.352838379			
382	0.05	8	0.5	0.5	0.5	0.5	3	0.5	0.67418812	0.346097548			
383	0.05	8	0.5	0.5	0.5	0.5	5	0.5	0.669775367	0.345349674			
384	0.05	8	0.5	0.5	0.5	0.5	7	0.5	0.673450048	0.353948957			
385	0.05	8	0.5	0.5	0.5	0.5	9	0.5	0.672337377	0.352937007			
386	0.05	8	0.5	0.5	0.5	0.5	11	0.5	0.673485124	0.355273622			
387	0.05	8	0.5	0.5	0.5	0.5	13	0.5	0.675192398	0.360430568			
388	0.05	8	0.5	0.5	0.5	0.5	15	0.5	0.676704176	0.363383652			
389	0.05	8	0.5	0.5	0.5	0.5	17	0.5	0.67649893	0.36142673			
390	0.05	8	0.5	0.5	0.5	0.5	19	0.5	0.677272532	0.365714645			
391	0.05	8	0.6	0.5	0.5	0.5	1	0.5	0.669510389	0.343485055			
392	0.05	8	0.6	0.5	0.5	0.5	3	0.5	0.674093899	0.354256141			
393	0.05	8	0.6	0.5	0.5	0.5	5	0.5	0.671711192	0.345333079			
394	0.05	8	0.6	0.5	0.5	0.5	7	0.5	0.669888315	0.347383379			
395	0.05	8	0.6	0.5	0.5	0.5	9	0.5	0.676383268	0.358426851			
396	0.05	8	0.6	0.5	0.5	0.5	11	0.5	0.674441146	0.360732591			
397	0.05	8	0.6	0.5	0.5	0.5	13	0.5	0.674120667	0.357050589			
398	0.05	8	0.6	0.5	0.5	0.5	15	0.5	0.675380357	0.360877105			
399	0.05	8	0.6	0.5	0.5	0.5	17	0.5	0.675329327	0.361550444			
400	0.05	8	0.6	0.5	0.5	0.5	19	0.5	0.675303366	0.362018183			
401	0.05	8	0.7	0.5	0.5	0.5	1	0.5	0.66635659	0.343899566			
402	0.05	8	0.7	0.5	0.5	0.5	3	0.5	0.671513963	0.344993368			
403	0.05	8	0.7	0.5	0.5	0.5	5	0.5	0.674254249	0.350486793			
404	0.05	8	0.7	0.5	0.5	0.5	7	0.5	0.675221179	0.354744827			
405	0.05	8	0.7	0.5	0.5	0.5	9	0.5	0.				

#	sta	max	depth	alpha	colsample	bytree	min	child	weight	subsample	rounds	Accuracy	Kappa
601	0.05	10	0.5	0.5	1	0.5	50	0.66512549	0.327902763				
602	0.05	10	0.5	0.5	1	0.5	50	0.66144442	0.330766977				
603	0.05	10	0.5	0.5	1	0.5	50	0.667423522	0.336654162				
604	0.05	10	0.5	0.5	7	0.5	50	0.666335753	0.339133363				
605	0.05	10	0.5	0.5	7	0.5	50	0.670418278	0.345935544				
606	0.05	10	0.5	0.5	11	0.5	50	0.672274636	0.349937013				
607	0.05	10	0.5	0.5	13	0.5	50	0.672046419	0.352237684				
608	0.05	10	0.5	0.5	15	0.5	50	0.674169968	0.357491744				
609	0.05	10	0.5	0.5	17	0.5	50	0.675311442	0.360896647				
610	0.05	10	0.5	0.5	19	0.5	50	0.677275875	0.363743445				
611	0.05	10	0.6	0.5	1	0.5	50	0.666667539	0.330346639				
612	0.05	10	0.6	0.5	5	0.5	50	0.67003221	0.34116488				
613	0.05	10	0.6	0.5	9	0.5	50	0.670672848	0.343953603				
614	0.05	10	0.6	0.5	7	0.5	50	0.669547966	0.342178317				
615	0.05	10	0.6	0.5	11	0.5	50	0.668976725	0.341803624				
616	0.05	10	0.6	0.5	13	0.5	50	0.67028292	0.348986807				
617	0.05	10	0.6	0.5	15	0.5	50	0.6729007	0.352455153				
618	0.05	10	0.6	0.5	17	0.5	50	0.672525302	0.353977178				
619	0.05	10	0.6	0.5	19	0.5	50	0.672088175	0.353945918				
620	0.05	10	0.6	0.5	1	0.5	50	0.676176242	0.363275711				
621	0.05	10	0.7	0.5	1	0.5	50	0.66522806	0.329411147				
622	0.05	10	0.7	0.5	3	0.5	50	0.667485259	0.336532602				
623	0.05	10	0.7	0.5	5	0.5	50	0.669262914	0.340277240				
624	0.05	10	0.7	0.5	7	0.5	50	0.670912033	0.345524135				
625	0.05	10	0.7	0.5	9	0.5	50	0.670682996	0.345511702				
626	0.05	10	0.7	0.5	11	0.5	50	0.669284129	0.346879331				
627	0.05	10	0.7	0.5	13	0.5	50	0.67159538	0.352382421				
628	0.05	10	0.7	0.5	15	0.5	50	0.67357501	0.357610711				
629	0.05	10	0.7	0.5	17	0.5	50	0.67585359	0.361777048				
630	0.05	10	0.7	0.5	19	0.5	50	0.675910105	0.361894963				
631	0.05	10	0.8	0.5	1	0.5	50	0.6710623	0.341224094				
632	0.05	10	0.8	0.5	3	0.5	50	0.665986925	0.331350891				
633	0.05	10	0.8	0.5	5	0.5	50	0.669284129	0.33202512				
634	0.05	10	0.8	0.5	7	0.5	50	0.66856535	0.340295076				
635	0.05	10	0.8	0.5	9	0.5	50	0.669090235	0.343309347				
636	0.05	10	0.8	0.5	11	0.5	50	0.67196296	0.351775311				
637	0.05	10	0.8	0.5	13	0.5	50	0.672009618	0.351925272				
638	0.05	10	0.8	0.5	15	0.5	50	0.676175423	0.36086804				
639	0.05	10	0.8	0.5	17	0.5	50	0.674242371	0.357630784				
640	0.05	10	0.8	0.5	19	0.5	50	0.67251123	0.364689274				
641	0.05	10	0.9	0.5	1	0.5	50	0.667535394	0.334767241				
642	0.05	10	0.9	0.5	3	0.5	50	0.665721199	0.331994489				
643	0.05	10	0.9	0.5	5	0.5	50	0.667651514	0.337826231				
644	0.05	10	0.9	0.5	7	0.5	50	0.671705829	0.347654466				
645	0.05	10	0.9	0.5	9	0.5	50	0.669393607	0.343702411				
646	0.05	10	0.9	0.5	11	0.5	50	0.671668492	0.350044764				
647	0.05	10	0.9	0.5	13	0.5	50	0.673751058	0.355623328				
648	0.05	10	0.9	0.5	15	0.5	50	0.67534024	0.361499421				
649	0.05	10	0.9	0.5	17	0.5	50	0.674344419	0.356918451				
650	0.05	10	0.9	0.5	19	0.5	50	0.675834345	0.363031261				
651	0.05	10	0.9	0.5	1	0.5	50	0.6709601	0.339214458				
652	0.05	10	0.9	0.5	3	0.5	50	0.666206061	0.338526526				
653	0.05	10	0.9	0.5	5	0.5	50	0.669811955	0.342030317				
654	0.05	10	0.9	0.5	7	0.5	50	0.667462172	0.339189364				
655	0.05	10	0.9	0.5	9	0.5	50	0.670416553	0.345907507				
656	0.05	10	0.9	0.5	11	0.5	50	0.670830339	0.347030148				
657	0.05	10	0.9	0.5	13	0.5	50	0.670739969	0.349435959				
658	0.05	10	0.9	0.5	15	0.5	50	0.670381472	0.349395699				
659	0.05	10	0.9	0.5	17	0.5	50	0.674168493	0.357217536				
660	0.05	10	0.9	0.5	19	0.5	50	0.67554774	0.362091919				
661	0.05	10	0.9	0.5	1	0.5	50	0.675075707	0.35317853				
662	0.05	10	0.9	0.5	3	0.5	50	0.67799434	0.372769007				
663	0.05	10	0.9	0.5	5	0.5	50	0.680835421	0.377961363				
664	0.05	10	0.9	0.5	7	0.5	50	0.67553545	0.371460834				
665	0.05	10	0.9	0.5	9	0.5	50	0.680071201	0.376339628				
666	0.05	10	0.9	0.5	11	0.5	50	0.67535493	0.375133006				
667	0.05	10	0.9	0.5	13	0.5	50	0.677404551	0.371620722				
668	0.05	10	0.9	0.5	15	0.5	50	0.678812488	0.377118244				
669	0.05	10	0.9	0.5	17	0.5	50	0.681175548	0.379833019				
670	0.05	10	0.9	0.5	19	0.5	50	0.681781633	0.381156033				
671	0.05	10	0.9	0.5	1	0.5	50	0.678197563	0.373233627				
672	0.05	10	0.9	0.5	3	0.5	50	0.686152228	0.379531356				
673	0.05	10	0.9	0.5	5	0.5	50	0.678525433	0.373007934				
674	0.05	10	0.9	0.5	7	0.5	50	0.675644734	0.368828295				
675	0.05	10	0.9	0.5	9	0.5	50	0.680529965	0.375501935				
676	0.05	10	0.9	0.5	11	0.5	50	0.677765961	0.373147591				
677	0.05	10	0.9	0.5	13	0.5	50	0.67803226	0.373074701				
678	0.05	10	0.9	0.5	15	0.5	50	0.67859112	0.376884958				
679	0.05	10	0.9	0.5	17	0.5	50	0.680595468	0.378605207				
680	0.05	10	0.9	0.5	19	0.5	50	0.681253426	0.38005211				
681	0.05	10	0.9	0.5	1	0.5	50	0.675135433	0.366549049				
682	0.05	10	0.9	0.5	3	0.5	50	0.67469693	0.375023037				
683	0.05	10	0.9	0.5	5	0.5	50	0.678827588	0.374505859				
684	0.05	10	0.9	0.5	7	0.5	50	0.678406629	0.372904562				
685	0.05	10	0.9	0.5	9	0.5	50	0.677765961	0.372227019				
686	0.05	10	0.9	0.5	11	0.5	50	0.678620988	0.376770262				
687	0.05	10	0.9	0.5	13	0.5	50	0.680758892	0.377925266				
688	0.05	10	0.9	0.5	15	0.5	50	0.678656336	0.375222878				
689	0.05	10	0.9	0.5	17	0.5	50	0.681252435	0.380081862				
690	0.05	10	0.9	0.5	19	0.5	50	0.680597598	0.378709755				
691	0.05	10	0.9	0.5	1	0.5	50	0.678941867	0.373898772				
692	0.05	10	0.9	0.5	3	0.5	50	0.679739397	0.375785033				
693	0.05	10	0.9	0.5	5	0.5	50	0.677731454	0.371655012				
694	0.05	10	0.9	0.5	7	0.5	50	0.679592479	0.37476615				
695	0.05	10	0.9	0.5	9	0.5	50	0.68087153	0.378698994				
696	0.05	10	0.9	0.5	11	0.5	50	0.677349154	0.372112733				
697	0.05	10	0.9	0.5	13	0.5	50	0.680154078	0.377589050				
698	0.05	10	0.9	0.5	15	0.5	50	0.679736206	0.377118202				
699	0.05	10	0.9	0.5	17	0.5	50	0.679424556	0.379652744				
700	0.05	10	0.9	0.5	19	0.5	50	0.678338322	0.378193677				
701	0.05	10	0.9	0.5	1	0.5	50	0.677764269	0.371682590				
702	0.05	10	0.9	0.5	3	0.5	50	0.679225453	0.372534626				
703	0.05	10	0.9	0.5	5	0.5	50	0.678599124	0.375330335				
704	0.05	10	0.9	0.5	7	0.5	50	0.677388193	0.3708223				
705	0.05	10	0.9	0.5	9	0.5	50	0.680457022	0.378138108				
706	0.05	10	0.9	0.5	11	0.5	50	0.67874221	0.373947077				
707	0.05	10	0.9	0.5	13	0.5	50	0.67893442	0.378947929				
708	0.05	10	0.9	0.5	15								

toys	eta	max_depth	gamma	min_child_weight	min_child_weight	subsample	threads	Accuracy	Map50
1051	0.5	0.6	0.5	1	0.5	0.5	0.688109947	0.340374782	
1052	0.5	0.6	0.5	3	0.5	0.5	0.671137109	0.347852391	
1053	0.5	0.6	0.5	5	0.5	0.5	0.6710262	0.348897055	
1054	0.5	0.6	0.5	7	0.5	0.5	0.67166025	0.351345454	
1055	0.5	0.6	0.5	9	0.5	0.5	0.672010303	0.351974422	
1056	0.5	0.6	0.5	11	0.5	0.5	0.671935403	0.353036242	
1057	0.5	0.6	0.5	13	0.5	0.5	0.671712196	0.353568568	
1058	0.5	0.6	0.5	15	0.5	0.5	0.675192685	0.351665588	
1059	0.5	0.6	0.5	17	0.5	0.5	0.674283838	0.359974611	
1060	0.5	0.6	0.5	19	0.5	0.5	0.674775975	0.35164457	
1061	0.5	0.6	0.5	21	0.5	0.5	0.674528101	0.35345454	
1062	0.5	0.6	0.5	23	0.5	0.5	0.671029247	0.347073914	
1063	0.5	0.6	0.5	25	0.5	0.5	0.675674458	0.340443565	
1064	0.5	0.6	0.5	27	0.5	0.5	0.68531303	0.345881572	
1065	0.5	0.6	0.5	29	0.5	0.5	0.67045159	0.348112891	
1066	0.5	0.6	0.5	31	0.5	0.5	0.671781262	0.352752629	
1067	0.5	0.6	0.5	33	0.5	0.5	0.67260931	0.357056341	
1068	0.5	0.6	0.5	35	0.5	0.5	0.671428101	0.35785938	
1069	0.5	0.6	0.5	37	0.5	0.5	0.67530487	0.361569251	
1070	0.5	0.6	0.5	39	0.5	0.5	0.67110081	0.356959189	
1071	0.5	0.6	0.5	41	0.5	0.5	0.68562103	0.344735667	
1072	0.5	0.6	0.5	43	0.5	0.5	0.67121266	0.347077598	
1073	0.5	0.6	0.5	45	0.5	0.5	0.672232050	0.350570478	
1074	0.5	0.6	0.5	47	0.5	0.5	0.68971458	0.346856555	
1075	0.5	0.6	0.5	49	0.5	0.5	0.670871614	0.345886534	
1076	0.5	0.6	0.5	51	0.5	0.5	0.671743904	0.352141679	
1077	0.5	0.6	0.5	53	0.5	0.5	0.675418922	0.350666638	
1078	0.5	0.6	0.5	55	0.5	0.5	0.67579081	0.354395251	
1079	0.5	0.6	0.5	57	0.5	0.5	0.67309174	0.358626061	
1080	0.5	0.6	0.5	59	0.5	0.5	0.675185351	0.362944776	
1081	0.5	0.6	0.5	61	0.5	0.5	0.68788125	0.345077228	
1082	0.5	0.6	0.5	63	0.5	0.5	0.67500019	0.351029038	
1083	0.5	0.6	0.5	65	0.5	0.5	0.686829343	0.34371404	
1084	0.5	0.6	0.5	67	0.5	0.5	0.68662341	0.343043031	
1085	0.5	0.6	0.5	69	0.5	0.5	0.686013488	0.344607155	
1086	0.5	0.6	0.5	71	0.5	0.5	0.670801591	0.343048682	
1087	0.5	0.6	0.5	73	0.5	0.5	0.673563505	0.356585178	
1088	0.5	0.6	0.5	75	0.5	0.5	0.671972511	0.353676166	
1089	0.5	0.6	0.5	77	0.5	0.5	0.67500019	0.351029038	
1090	0.5	0.6	0.5	79	0.5	0.5	0.674357453	0.361160307	
1091	0.5	0.6	0.5	81	0.5	0.5	0.674020326	0.360912454	
1092	0.5	0.6	0.5	83	0.5	0.5	0.68666241	0.344124614	
1093	0.5	0.6	0.5	85	0.5	0.5	0.686621696	0.345434289	
1094	0.5	0.6	0.5	87	0.5	0.5	0.673030492	0.353902218	
1095	0.5	0.6	0.5	89	0.5	0.5	0.672727272	0.353371768	
1096	0.5	0.6	0.5	91	0.5	0.5	0.675151529	0.352505285	
1097	0.5	0.6	0.5	93	0.5	0.5	0.673978292	0.358184676	
1098	0.5	0.6	0.5	95	0.5	0.5	0.67292608	0.34666254	
1099	0.5	0.6	0.5	97	0.5	0.5	0.674778295	0.345181637	
1100	0.5	0.6	0.5	99	0.5	0.5	0.675683172	0.363863826	
1101	0.5	0.6	0.5	101	0.5	0.5	0.678795758	0.371862028	
1102	0.5	0.6	0.5	103	0.5	0.5	0.66615518	0.333291718	
1103	0.5	0.6	0.5	105	0.5	0.5	0.667767015	0.33913866	
1104	0.5	0.6	0.5	107	0.5	0.5	0.671063806	0.347748117	
1105	0.5	0.6	0.5	109	0.5	0.5	0.676870919	0.34747689	
1106	0.5	0.6	0.5	111	0.5	0.5	0.67185891	0.351489178	
1107	0.5	0.6	0.5	113	0.5	0.5	0.686596386	0.345434289	
1108	0.5	0.6	0.5	115	0.5	0.5	0.676252258	0.362907144	
1109	0.5	0.6	0.5	117	0.5	0.5	0.67288125	0.345192044	
1110	0.5	0.6	0.5	119	0.5	0.5	0.672615203	0.357055874	
1111	0.5	0.6	0.5	121	0.5	0.5	0.686661688	0.340398295	
1112	0.5	0.6	0.5	123	0.5	0.5	0.686903006	0.339489894	
1113	0.5	0.6	0.5	125	0.5	0.5	0.686956384	0.34239884	
1114	0.5	0.6	0.5	127	0.5	0.5	0.687464588	0.339614434	
1115	0.5	0.6	0.5	129	0.5	0.5	0.671518526	0.349404003	
1116	0.5	0.6	0.5	131	0.5	0.5	0.68871458	0.34111368	
1117	0.5	0.6	0.5	133	0.5	0.5	0.67337321	0.355608006	
1118	0.5	0.6	0.5	135	0.5	0.5	0.67329801	0.356257189	
1119	0.5	0.6	0.5	137	0.5	0.5	0.673942107	0.356780317	
1120	0.5	0.6	0.5	139	0.5	0.5	0.672235551	0.356677222	
1121	0.5	0.6	0.5	141	0.5	0.5	0.687180206	0.344695574	
1122	0.5	0.6	0.5	143	0.5	0.5	0.685078628	0.332422773	
1123	0.5	0.6	0.5	145	0.5	0.5	0.67445159	0.352505285	
1124	0.5	0.6	0.5	147	0.5	0.5	0.68754021	0.340162128	
1125	0.5	0.6	0.5	149	0.5	0.5	0.67033982	0.34730128	
1126	0.5	0.6	0.5	151	0.5	0.5	0.670282162	0.34865532	
1127	0.5	0.6	0.5	153	0.5	0.5	0.686848752	0.348384545	
1128	0.5	0.6	0.5	155	0.5	0.5	0.675482822	0.360278918	
1129	0.5	0.6	0.5	157	0.5	0.5	0.67349408	0.358292867	
1130	0.5	0.6	0.5	159	0.5	0.5	0.675330710	0.357671734	
1131	0.5	0.6	0.5	161	0.5	0.5	0.684320343	0.329046045	
1132	0.5	0.6	0.5	163	0.5	0.5	0.686691265	0.335279003	
1133	0.5	0.6	0.5	165	0.5	0.5	0.687097947	0.338818181	
1134	0.5	0.6	0.5	167	0.5	0.5	0.689714505	0.345415499	
1135	0.5	0.6	0.5	169	0.5	0.5	0.672570945	0.351622388	
1136	0.5	0.6	0.5	171	0.5	0.5	0.671849897	0.351495133	
1137	0.5	0.6	0.5	173	0.5	0.5	0.671552001	0.354481297	
1138	0.5	0.6	0.5	175	0.5	0.5	0.671384471	0.352050611	
1139	0.5	0.6	0.5	177	0.5	0.5	0.67223614	0.355437124	
1140	0.5	0.6	0.5	179	0.5	0.5	0.67288125	0.345860168	
1141	0.5	0.6	0.5	181	0.5	0.5	0.685532379	0.332822195	
1142	0.5	0.6	0.5	183	0.5	0.5	0.687180206	0.339873545	
1143	0.5	0.6	0.5	185	0.5	0.5	0.68825226	0.338781912	
1144	0.5	0.6	0.5	187	0.5	0.5	0.68484151	0.3545611	
1145	0.5	0.6	0.5	189	0.5	0.5	0.68913022	0.34440184	
1146	0.5	0.6	0.5	191	0.5	0.5	0.689594123	0.346004062	
1147	0.5	0.6	0.5	193	0.5	0.5	0.671937374	0.351875722	
1148	0.5	0.6	0.5	195	0.5	0.5	0.671687374	0.352791194	
1149	0.5	0.6	0.5	197	0.5	0.5	0.672389479	0.354690674	
1150	0.5	0.6	0.5	199	0.5	0.5	0.672389314	0.358243118	
1151	0.5	0.6	0.5	201	0.5	0.5	0.684050931	0.347603984	
1152	0.5	0.6	0.5	203	0.5	0.5	0.687094767	0.338427157	
1153	0.5	0.6	0.5	205	0.5	0.5	0.689548495	0.343344652	
1154	0.5	0.6	0.5	207	0.5	0.5	0.689714505	0.344115424	
1155	0.5	0.6	0.5	209	0.5	0.5	0.67040346	0.346831193	
1156	0.5	0.6	0.5	211	0.5	0.5	0.689244548	0.345888787	
1157	0.5	0.6	0.5	213	0.5	0.5	0.673314068	0.354987057	
1158	0.5	0.6	0.5	215	0.5	0.5	0.673152001	0.34148482	
1159	0.5	0.6	0.5	217	0.5	0.5	0.674052543	0.358363531	
1160	0.5	0.6	0.5	219	0.5	0.5	0.674056382	0.359229265	
1161	0.5	0.6	0.5	221	0.5	0.5	0.686821696	0.345454281	
1162	0.5	0.6	0.5	223	0.5	0.5	0.686933161	0.337261415	
1163	0.5	0.6	0.5	225	0.5	0.5	0.686932302	0.337651287	
1164	0.5	0.6	0.5	227	0.5	0.5	0.686988022	0.343221238	
1165	0.5	0.6	0.5	229	0.5	0.5	0.67166905	0.351927477	
1166	0.5	0.6	0.5	231	0.5	0.5	0.67205012	0.351762552	
1167	0.5	0.6	0.5	233	0.5	0.5	0.674283507	0.357197309	
1168	0.5	0.6	0.5	235	0.5	0.5	0.67198125	0.357335446	
1169	0.5	0.6	0.5	237	0.5	0.5	0.673600741	0.358070564	
1170	0.5	0.6	0.5	239	0.5	0.5	0.673984388	0.358806062	
1171	0.5	0.6	0.5	241	0.5	0.5	0.68540426	0.332327557	
1172	0.5	0.6	0.5	243	0.5	0.5	0.686819356	0.335601764	
1173	0.5	0.6	0.5	245	0.5	0.5	0.686962495	0.343737827	
1174	0.5	0.6	0.5	247	0.5	0.5	0.685908084	0.342760389	
1175	0.5	0.6	0.5	249	0.5	0.5	0.672857034	0.349547525	
1176	0.5	0.6	0.5	251	0.5	0.5	0.671969927	0.352102105	
1177	0.5	0.6	0.5	253	0.5	0.5	0.67136399	0.351295486	
1178	0.5	0.6	0.5	255	0.5	0.5	0.672532535	0.354797098	

#	sta	max	depth	alpha	colsample	bytes	min	child	weight	subsample	rounds	Accuracy	Kappa
1201	0.05	9	1	0.3	1	0.5	100	0.66525726	0.34978421				
1202	0.05	9	1	0.3	1	0.5	100	0.67088848	0.34223235				
1203	0.05	9	1	0.3	1	0.5	100	0.668400705	0.341107994				
1204	0.05	9	1	0.3	7	0.5	100	0.66971767	0.34320434				
1205	0.05	9	1	0.3	7	0.5	100	0.66678893	0.343748314				
1206	0.05	9	1	0.3	11	0.5	100	0.66585155	0.34769608				
1207	0.05	9	1	0.3	13	0.5	100	0.671480215	0.352287481				
1208	0.05	9	1	0.3	15	0.5	100	0.67248801	0.350843618				
1209	0.05	9	1	0.3	17	0.5	100	0.673484498	0.352038192				
1210	0.05	9	1	0.3	19	0.5	100	0.675114041	0.361882417				
1211	0.05	10	0	0.3	1	0.5	100	0.665380994	0.329078629				
1212	0.05	10	0	0.3	1	0.5	100	0.662388212	0.324841902				
1213	0.05	10	0	0.3	5	0.5	100	0.665193323	0.32226721				
1214	0.05	10	0	0.3	7	0.5	100	0.666820218	0.33663891				
1215	0.05	10	0	0.3	9	0.5	100	0.669734732	0.344258851				
1216	0.05	10	0	0.3	11	0.5	100	0.669550988	0.34450574				
1217	0.05	10	0	0.3	13	0.5	100	0.670227244	0.348588477				
1218	0.05	10	0	0.3	15	0.5	100	0.672047365	0.353020203				
1219	0.05	10	0	0.3	17	0.5	100	0.672009172	0.354016458				
1220	0.05	10	0	0.3	19	0.5	100	0.67462125	0.360542632				
1221	0.05	10	0.1	0.3	1	0.5	100	0.662918012	0.323024719				
1222	0.05	10	0.1	0.3	3	0.5	100	0.66765368	0.338390723				
1223	0.05	10	0.1	0.3	5	0.5	100	0.667769608	0.337074469				
1224	0.05	10	0.1	0.3	7	0.5	100	0.664507412	0.332026207				
1225	0.05	10	0.1	0.3	9	0.5	100	0.666028278	0.346059508				
1226	0.05	10	0.1	0.3	11	0.5	100	0.67068226	0.347811104				
1227	0.05	10	0.1	0.3	13	0.5	100	0.671518892	0.350514015				
1228	0.05	10	0.1	0.3	15	0.5	100	0.672502426	0.353984645				
1229	0.05	10	0.1	0.3	17	0.5	100	0.67242671	0.355058669				
1230	0.05	10	0.1	0.3	19	0.5	100	0.67291281	0.358022932				
1231	0.05	10	0.2	0.3	1	0.5	100	0.666554463	0.323113283				
1232	0.05	10	0.2	0.3	3	0.5	100	0.665607578	0.330962012				
1233	0.05	10	0.2	0.3	5	0.5	100	0.665418819	0.331905211				
1234	0.05	10	0.2	0.3	7	0.5	100	0.669736528	0.342850304				
1235	0.05	10	0.2	0.3	9	0.5	100	0.66833481	0.341615593				
1236	0.05	10	0.2	0.3	11	0.5	100	0.67023095	0.347099387				
1237	0.05	10	0.2	0.3	13	0.5	100	0.670701814	0.350120094				
1238	0.05	10	0.2	0.3	15	0.5	100	0.673146411	0.355246566				
1239	0.05	10	0.2	0.3	17	0.5	100	0.673184333	0.356037348				
1240	0.05	10	0.2	0.3	19	0.5	100	0.674231898	0.360216448				
1241	0.05	10	0.3	0.3	1	0.5	100	0.661620899	0.321212917				
1242	0.05	10	0.3	0.3	3	0.5	100	0.666402215	0.333022704				
1243	0.05	10	0.3	0.3	5	0.5	100	0.667238006	0.336849412				
1244	0.05	10	0.3	0.3	7	0.5	100	0.669212708	0.339333063				
1245	0.05	10	0.3	0.3	9	0.5	100	0.669131514	0.343183615				
1246	0.05	10	0.3	0.3	11	0.5	100	0.669878147	0.343851268				
1247	0.05	10	0.3	0.3	13	0.5	100	0.671592103	0.350895905				
1248	0.05	10	0.3	0.3	15	0.5	100	0.67128991	0.351695665				
1249	0.05	10	0.3	0.3	17	0.5	100	0.672220228	0.360889963				
1250	0.05	10	0.3	0.3	19	0.5	100	0.674718628	0.358541102				
1251	0.05	10	0.4	0.3	1	0.5	100	0.663791508	0.325729509				
1252	0.05	10	0.4	0.3	3	0.5	100	0.667816841	0.336181423				
1253	0.05	10	0.4	0.3	5	0.5	100	0.666137279	0.334247639				
1254	0.05	10	0.4	0.3	7	0.5	100	0.667089962	0.337362369				
1255	0.05	10	0.4	0.3	9	0.5	100	0.668749454	0.3442671924				
1256	0.05	10	0.4	0.3	11	0.5	100	0.669055454	0.34420355				
1257	0.05	10	0.4	0.3	13	0.5	100	0.671917346	0.348358989				
1258	0.05	10	0.4	0.3	15	0.5	100	0.672047494	0.352961228				
1259	0.05	10	0.4	0.3	17	0.5	100	0.674434505	0.358228643				
1260	0.05	10	0.4	0.3	19	0.5	100	0.675421013	0.362230576				
1261	0.05	10	0.5	0.3	1	0.5	100	0.665183103	0.324140987				
1262	0.05	10	0.5	0.3	3	0.5	100	0.665305094	0.330367348				
1263	0.05	10	0.5	0.3	5	0.5	100	0.667613734	0.336891173				
1264	0.05	10	0.5	0.3	7	0.5	100	0.668487338	0.340053562				
1265	0.05	10	0.5	0.3	9	0.5	100	0.66765695	0.341002581				
1266	0.05	10	0.5	0.3	11	0.5	100	0.669606008	0.345218816				
1267	0.05	10	0.5	0.3	13	0.5	100	0.669433085	0.348370992				
1268	0.05	10	0.5	0.3	15	0.5	100	0.67079798	0.3505232				
1269	0.05	10	0.5	0.3	17	0.5	100	0.672841788	0.356287072				
1270	0.05	10	0.5	0.3	19	0.5	100	0.674699785	0.360718358				
1271	0.05	10	0.5	0.3	1	0.5	100	0.669575387	0.33899099				
1272	0.05	10	0.5	0.3	3	0.5	100	0.669672387	0.337200202				
1273	0.05	10	0.5	0.3	5	0.5	100	0.669016242	0.339439565				
1274	0.05	10	0.5	0.3	7	0.5	100	0.669696388	0.3429115608				
1275	0.05	10	0.5	0.3	9	0.5	100	0.669091427	0.3421441414				
1276	0.05	10	0.5	0.3	11	0.5	100	0.668297492	0.3431683				
1277	0.05	10	0.5	0.3	13	0.5	100	0.67001565	0.347574599				
1278	0.05	10	0.5	0.3	15	0.5	100	0.671706241	0.352920113				
1279	0.05	10	0.5	0.3	17	0.5	100	0.673336238	0.358717811				
1280	0.05	10	0.5	0.3	19	0.5	100	0.672919137	0.35666699				
1281	0.05	10	0.7	0.3	1	0.5	100	0.664509093	0.32730594				
1282	0.05	10	0.7	0.3	3	0.5	100	0.665849098	0.332010938				
1283	0.05	10	0.7	0.3	5	0.5	100	0.666819313	0.335695191				
1284	0.05	10	0.7	0.3	7	0.5	100	0.669848629	0.344605517				
1285	0.05	10	0.7	0.3	9	0.5	100	0.672743031	0.350290901				
1286	0.05	10	0.7	0.3	11	0.5	100	0.669878731	0.345289011				
1287	0.05	10	0.7	0.3	13	0.5	100	0.66804886	0.34592752				
1288	0.05	10	0.7	0.3	15	0.5	100	0.673487381	0.3553830476				
1289	0.05	10	0.7	0.3	17	0.5	100	0.671780134	0.358026656				
1290	0.05	10	0.7	0.3	19	0.5	100	0.67675127	0.363039628				
1291	0.05	10	0.8	0.3	1	0.5	100	0.669357408	0.337024324				
1292	0.05	10	0.8	0.3	3	0.5	100	0.669759595	0.331120526				
1293	0.05	10	0.8	0.3	5	0.5	100	0.667934338	0.338253251				
1294	0.05	10	0.8	0.3	7	0.5	100	0.665345398	0.334182217				
1295	0.05	10	0.8	0.3	9	0.5	100	0.669280861	0.344029177				
1296	0.05	10	0.8	0.3	11	0.5	100	0.671888444	0.350647698				
1297	0.05	10	0.8	0.3	13	0.5	100	0.67193334	0.352243201				
1298	0.05	10	0.8	0.3	15	0.5	100	0.672198494	0.353280107				
1299	0.05	10	0.8	0.3	17	0.5	100	0.672097109	0.355404072				
1300	0.05	10	0.8	0.3	19	0.5	100	0.674491953	0.359194146				
1301	0.05	10	0.9	0.3	1	0.5	100	0.666784016	0.332753999				
1302	0.05	10	0.9	0.3	3	0.5	100	0.666478361	0.334075099				
1303	0.05	10	0.9	0.3	5	0.5	100	0.66878199	0.336855321				
1304	0.05	10	0.9	0.3	7	0.5	100	0.66769117	0.338942505				
1305	0.05	10	0.9	0.3	9	0.5	100	0.669319099	0.343753562				
1306	0.05	10	0.9	0.3	11	0.5	100	0.670267749	0.348879578				

[illegible]

#	sta	max	depth	alpha	colsample	bytes	min	child	weight	subsample	rounds	Accuracy	Kappa
1801	0.05	9	0.4	0.5	1	0.5	150	0.66691938	0.32135088	150	0.66691938	0.32135088	
1802	0.05	9	0.4	0.5	1	0.5	150	0.66615101	0.32839322	150	0.66615101	0.32839322	
1803	0.05	9	0.4	0.5	8	0.5	150	0.666518304	0.333194034	150	0.666518304	0.333194034	
1804	0.05	9	0.4	0.5	7	0.5	150	0.666927291	0.33517325	150	0.666927291	0.33517325	
1805	0.05	9	0.4	0.5	17	0.5	150	0.666500069	0.3243432694	150	0.666500069	0.3243432694	
1806	0.05	9	0.4	0.5	11	0.5	150	0.667652387	0.338893744	150	0.667652387	0.338893744	
1807	0.05	9	0.4	0.5	13	0.5	150	0.669473154	0.343622202	150	0.669473154	0.343622202	
1808	0.05	9	0.4	0.5	15	0.5	150	0.667617449	0.341778922	150	0.667617449	0.341778922	
1809	0.05	9	0.4	0.5	17	0.5	150	0.666720992	0.344351514	150	0.666720992	0.344351514	
1810	0.05	9	0.4	0.5	19	0.5	150	0.671024852	0.349645692	150	0.671024852	0.349645692	
1811	0.05	9	0.5	0.5	1	0.5	150	0.662570699	0.322310124	150	0.662570699	0.322310124	
1812	0.05	9	0.5	0.5	5	0.5	150	0.66318411	0.324829277	150	0.66318411	0.324829277	
1813	0.05	9	0.5	0.5	9	0.5	150	0.6648993	0.329882436	150	0.6648993	0.329882436	
1814	0.05	9	0.5	0.5	7	0.5	150	0.666708	0.334991998	150	0.666708	0.334991998	
1815	0.05	9	0.5	0.5	13	0.5	150	0.66632698	0.335353138	150	0.66632698	0.335353138	
1816	0.05	9	0.5	0.5	19	0.5	150	0.6662515	0.338304501	150	0.6662515	0.338304501	
1817	0.05	9	0.5	0.5	13	0.5	150	0.666814495	0.344936886	150	0.666814495	0.344936886	
1818	0.05	9	0.5	0.5	19	0.5	150	0.669932266	0.346526719	150	0.669932266	0.346526719	
1819	0.05	9	0.5	0.5	17	0.5	150	0.670797922	0.348028877	150	0.670797922	0.348028877	
1820	0.05	9	0.5	0.5	19	0.5	150	0.669776014	0.347257775	150	0.669776014	0.347257775	
1821	0.05	9	0.6	0.5	1	0.5	150	0.663370705	0.323120592	150	0.663370705	0.323120592	
1822	0.05	9	0.6	0.5	3	0.5	150	0.66445012	0.331233594	150	0.66445012	0.331233594	
1823	0.05	9	0.6	0.5	9	0.5	150	0.664669521	0.329319307	150	0.664669521	0.329319307	
1824	0.05	9	0.6	0.5	7	0.5	150	0.666100003	0.333748379	150	0.666100003	0.333748379	
1825	0.05	9	0.6	0.5	9	0.5	150	0.66935138	0.341690027	150	0.66935138	0.341690027	
1826	0.05	9	0.6	0.5	11	0.5	150	0.667089369	0.338946026	150	0.667089369	0.338946026	
1827	0.05	9	0.6	0.5	13	0.5	150	0.671821513	0.349358302	150	0.671821513	0.349358302	
1828	0.05	9	0.6	0.5	15	0.5	150	0.670227373	0.346981193	150	0.670227373	0.346981193	
1829	0.05	9	0.6	0.5	17	0.5	150	0.670457488	0.348670882	150	0.670457488	0.348670882	
1830	0.05	9	0.6	0.5	19	0.5	150	0.671783132	0.351720757	150	0.671783132	0.351720757	
1831	0.05	9	0.7	0.5	1	0.5	150	0.656132411	0.315675541	150	0.656132411	0.315675541	
1832	0.05	9	0.7	0.5	3	0.5	150	0.656175114	0.330738411	150	0.656175114	0.330738411	
1833	0.05	9	0.7	0.5	9	0.5	150	0.656518523	0.333326507	150	0.656518523	0.333326507	
1834	0.05	9	0.7	0.5	7	0.5	150	0.656289225	0.333846208	150	0.656289225	0.333846208	
1835	0.05	9	0.7	0.5	9	0.5	150	0.655192891	0.332307332	150	0.655192891	0.332307332	
1836	0.05	9	0.7	0.5	11	0.5	150	0.656281822	0.342968352	150	0.656281822	0.342968352	
1837	0.05	9	0.7	0.5	13	0.5	150	0.656827928	0.3431303452	150	0.656827928	0.3431303452	
1838	0.05	9	0.7	0.5	15	0.5	150	0.656250603	0.343212205	150	0.656250603	0.343212205	
1839	0.05	9	0.7	0.5	17	0.5	150	0.671289357	0.349962188	150	0.671289357	0.349962188	
1840	0.05	9	0.7	0.5	19	0.5	150	0.671631723	0.351999319	150	0.671631723	0.351999319	
1841	0.05	9	0.8	0.5	1	0.5	150	0.65488775	0.327538759	150	0.65488775	0.327538759	
1842	0.05	9	0.8	0.5	3	0.5	150	0.654812123	0.328599445	150	0.654812123	0.328599445	
1843	0.05	9	0.8	0.5	5	0.5	150	0.6568919	0.335140782	150	0.6568919	0.335140782	
1844	0.05	9	0.8	0.5	7	0.5	150	0.65697987	0.338626098	150	0.65697987	0.338626098	
1845	0.05	9	0.8	0.5	9	0.5	150	0.656251088	0.335959696	150	0.656251088	0.335959696	
1846	0.05	9	0.8	0.5	11	0.5	150	0.659773899	0.344188323	150	0.659773899	0.344188323	
1847	0.05	9	0.8	0.5	13	0.5	150	0.659905416	0.345999336	150	0.659905416	0.345999336	
1848	0.05	9	0.8	0.5	15	0.5	150	0.659017145	0.344503532	150	0.659017145	0.344503532	
1849	0.05	9	0.8	0.5	17	0.5	150	0.670379793	0.348682488	150	0.670379793	0.348682488	
1850	0.05	9	0.8	0.5	19	0.5	150	0.672127229	0.352074485	150	0.672127229	0.352074485	
1851	0.05	9	0.8	0.5	1	0.5	150	0.656936074	0.329026053	150	0.656936074	0.329026053	
1852	0.05	9	0.8	0.5	3	0.5	150	0.664319138	0.327784884	150	0.664319138	0.327784884	
1853	0.05	9	0.8	0.5	5	0.5	150	0.65655452	0.334361168	150	0.65655452	0.334361168	
1854	0.05	9	0.8	0.5	7	0.5	150	0.667048158	0.335972875	150	0.667048158	0.335972875	
1855	0.05	9	0.8	0.5	9	0.5	150	0.656063157	0.335151183	150	0.656063157	0.335151183	
1856	0.05	9	0.8	0.5	11	0.5	150	0.667766154	0.340550111	150	0.667766154	0.340550111	
1857	0.05	9	0.8	0.5	13	0.5	150	0.659528217	0.344221658	150	0.659528217	0.344221658	
1858	0.05	9	0.8	0.5	15	0.5	150	0.659590017	0.344133388	150	0.659590017	0.344133388	
1859	0.05	9	0.8	0.5	17	0.5	150	0.670193414	0.347557369	150	0.670193414	0.347557369	
1860	0.05	9	0.8	0.5	19	0.5	150	0.672545564	0.353405698	150	0.672545564	0.353405698	
1861	0.05	9	0.9	0.5	1	0.5	150	0.652485484	0.322846938	150	0.652485484	0.322846938	
1862	0.05	9	0.9	0.5	3	0.5	150	0.658525321	0.338048826	150	0.658525321	0.338048826	
1863	0.05	9	0.9	0.5	5	0.5	150	0.655190564	0.331477702	150	0.655190564	0.331477702	
1864	0.05	9	0.9	0.5	7	0.5	150	0.657738878	0.333948442	150	0.657738878	0.333948442	
1865	0.05	9	0.9	0.5	9	0.5	150	0.656910566	0.3347270781	150	0.656910566	0.3347270781	
1866	0.05	9	0.9	0.5	11	0.5	150	0.667050524	0.338825698	150	0.667050524	0.338825698	
1867	0.05	9	0.9	0.5	13	0.5	150	0.667596338	0.342444844	150	0.667596338	0.342444844	
1868	0.05	9	0.9	0.5	15	0.5	150	0.669281954	0.344863881	150	0.669281954	0.344863881	
1869	0.05	9	0.9	0.5	17	0.5	150	0.67121614	0.34958138	150	0.67121614	0.34958138	
1870	0.05	9	0.9	0.5	19	0.5	150	0.671365718	0.350975493	150	0.671365718	0.350975493	
1871	0.05	10	0	0.5	1	0.5	150	0.654456102	0.318121669	150	0.654456102	0.318121669	
1872	0.05	10	0	0.5	3	0.5	150	0.656254523	0.321728413	150	0.656254523	0.321728413	
1873	0.05	10	0	0.5	5	0.5	150	0.653071205	0.324205243	150	0.653071205	0.324205243	
1874	0.05	10	0	0.5	7	0.5	150	0.657425331	0.334712825	150	0.657425331	0.334712825	
1875	0.05	10	0	0.5	9	0.5	150	0.658924978	0.338726936	150	0.658924978	0.338726936	
1876	0.05	10	0	0.5	11	0.5	150	0.667273515	0.337930545	150	0.667273515	0.337930545	
1877	0.05	10	0	0.5	13	0.5	150	0.668448834	0.341417149	150	0.668448834	0.341417149	
1878	0.05	10	0	0.5	15	0.5	150	0.67023516	0.345925288	150	0.67023516	0.345925288	
1879	0.05	10	0	0.5	17	0.5	150	0.671934243	0.350549438	150	0.671934243	0.350549438	
1880	0.05	10	0	0.5	19	0.5	150	0.670305629	0.348220157	150	0.670305629	0.348220157	
1881	0.05	10	0.1	0.5	1	0.5	150	0.659774328	0.314220347	150	0.659774328	0.314220347	
1882	0.05	10	0.1	0.5	3	0.5	150	0.663751186	0.324032633	150	0.663751186	0.324032633	
1883	0.05	10	0.1	0.5	5	0.5	150	0.652274799	0.322711034	150	0.652274799	0.322711034	
1884	0.05	10	0.1	0.5	7	0.5	150	0.66490366	0.330068388	150	0.66490366	0.330068388	
1885	0.05	10	0.1	0.5	9	0.5	150	0.66291702	0.328231557	150	0.66291702	0.328231557	
1886	0.05	10	0.1	0.5	11	0.5	150	0.66715549	0.338692386	150	0.66715549	0.338692386	
1887	0.05	10	0.1	0.5	13	0.5	150	0.668221646	0.341008251	150	0.668221646	0.341008251	
1888	0.05	10	0.1	0.5	15	0.5	150	0.667891886	0.340491745	150	0.667891886	0.340491745	
1889	0.05	10	0.1										

[illegible]

#	eta	max depth	gamma	colsample	bytree	min child weight	subsample	rounds	Accuracy	Kappa
2401	0.05	8	0.9	0.5	0.5	1	0.5	200	0.66220765	0.320754613
2402	0.05	8	0.9	0.5	0.5	2	0.5	200	0.663232892	0.324494916
2403	0.05	8	0.9	0.5	0.5	3	0.5	200	0.663977827	0.326886417
2404	0.05	8	0.9	0.5	0.5	4	0.5	200	0.663713637	0.327915124
2405	0.05	8	0.9	0.5	0.5	5	0.5	200	0.660770767	0.324443421
2406	0.05	8	0.9	0.5	0.5	6	0.5	200	0.667957528	0.338549430
2407	0.05	8	0.9	0.5	0.5	11	0.5	200	0.66868604	0.336892352
2408	0.05	8	0.9	0.5	0.5	15	0.5	200	0.666252594	0.336345492
2409	0.05	8	0.9	0.5	0.5	17	0.5	200	0.670343122	0.348577547
2410	0.05	8	0.9	0.5	0.5	19	0.5	200	0.667350093	0.340914135
2411	0.05	8	1	0.5	1	0.5	200	0.664470763	0.327054341	
2412	0.05	8	1	0.5	1	0.5	200	0.661554152	0.321723847	
2413	0.05	8	1	0.5	5	0.5	200	0.664357018	0.327852365	
2414	0.05	8	1	0.5	7	0.5	200	0.65949392	0.332572763	
2415	0.05	8	1	0.5	9	0.5	200	0.66632776	0.335643862	
2416	0.05	8	1	0.5	11	0.5	200	0.66526076	0.333087796	
2417	0.05	8	1	0.5	13	0.5	200	0.664320559	0.331825447	
2418	0.05	8	1	0.5	15	0.5	200	0.668411558	0.340517171	
2419	0.05	8	1	0.5	17	0.5	200	0.6694354	0.344502926	
2420	0.05	8	1	0.5	19	0.5	200	0.66903092	0.343419305	
2421	0.05	8	1	0.5	1	0.5	200	0.6621261	0.318757566	
2422	0.05	8	1	0.5	3	0.5	200	0.660339522	0.316047586	
2423	0.05	8	1	0.5	5	0.5	200	0.66173674	0.318476644	
2424	0.05	8	1	0.5	7	0.5	200	0.661934537	0.322012338	
2425	0.05	8	1	0.5	9	0.5	200	0.665305236	0.329792276	
2426	0.05	8	1	0.5	11	0.5	200	0.663903246	0.32805556	
2427	0.05	8	1	0.5	13	0.5	200	0.664470828	0.33050943	
2428	0.05	8	1	0.5	15	0.5	200	0.668297018	0.339374913	
2429	0.05	8	1	0.5	17	0.5	200	0.66894706	0.339857675	
2430	0.05	8	1	0.5	19	0.5	200	0.66852701	0.341810408	
2431	0.05	8	1	0.5	1	0.5	200	0.661595479	0.31728791	
2432	0.05	8	1	0.5	3	0.5	200	0.660721886	0.316682149	
2433	0.05	8	1	0.5	5	0.5	200	0.6621917	0.320189726	
2434	0.05	8	1	0.5	7	0.5	200	0.664282551	0.326210070	
2435	0.05	8	1	0.5	9	0.5	200	0.664280271	0.327708464	
2436	0.05	8	1	0.5	11	0.5	200	0.663929771	0.328177185	
2437	0.05	8	1	0.5	13	0.5	200	0.665117177	0.332118966	
2438	0.05	8	1	0.5	15	0.5	200	0.668524634	0.339497863	
2439	0.05	8	1	0.5	17	0.5	200	0.66894966	0.339693937	
2440	0.05	8	1	0.5	19	0.5	200	0.669042892	0.341810408	
2441	0.05	8	1	0.5	1	0.5	200	0.660911207	0.315585852	
2442	0.05	8	1	0.5	3	0.5	200	0.663637665	0.322247408	
2443	0.05	8	1	0.5	5	0.5	200	0.664739881	0.326745381	
2444	0.05	8	1	0.5	7	0.5	200	0.666422892	0.332033811	
2445	0.05	8	1	0.5	9	0.5	200	0.664606095	0.329311521	
2446	0.05	8	1	0.5	11	0.5	200	0.664057593	0.329570164	
2447	0.05	8	1	0.5	13	0.5	200	0.6672744	0.336897038	
2448	0.05	8	1	0.5	15	0.5	200	0.666137451	0.334718828	
2449	0.05	8	1	0.5	17	0.5	200	0.667766757	0.338861218	
2450	0.05	8	1	0.5	19	0.5	200	0.669393998	0.34276382	
2451	0.05	8	1	0.5	1	0.5	200	0.660117174	0.314180232	
2452	0.05	8	1	0.5	3	0.5	200	0.661742173	0.318821471	
2453	0.05	8	1	0.5	5	0.5	200	0.66371325	0.324044228	
2454	0.05	8	1	0.5	7	0.5	200	0.663980291	0.326114841	
2455	0.05	8	1	0.5	9	0.5	200	0.663030037	0.32561098	
2456	0.05	8	1	0.5	11	0.5	200	0.665226596	0.331627734	
2457	0.05	8	1	0.5	13	0.5	200	0.665001732	0.331789438	
2458	0.05	8	1	0.5	15	0.5	200	0.668335326	0.338497853	
2459	0.05	8	1	0.5	17	0.5	200	0.667994675	0.33941458	
2460	0.05	8	1	0.5	19	0.5	200	0.669169348	0.342742823	
2461	0.05	8	1	0.5	1	0.5	200	0.66589181	0.331142138	
2462	0.05	8	1	0.5	3	0.5	200	0.662463809	0.320257208	
2463	0.05	8	1	0.5	5	0.5	200	0.664851205	0.326480129	
2464	0.05	8	1	0.5	7	0.5	200	0.663638094	0.325124858	
2465	0.05	8	1	0.5	9	0.5	200	0.66367801	0.326383272	
2466	0.05	8	1	0.5	11	0.5	200	0.665494371	0.331677894	
2467	0.05	8	1	0.5	13	0.5	200	0.667881597	0.337495006	
2468	0.05	8	1	0.5	15	0.5	200	0.667198124	0.337307123	
2469	0.05	8	1	0.5	17	0.5	200	0.665342168	0.334189881	
2470	0.05	8	1	0.5	19	0.5	200	0.668335197	0.341306244	
2471	0.05	8	1	0.5	1	0.5	200	0.659245058	0.31393647	
2472	0.05	8	1	0.5	3	0.5	200	0.6615947	0.318607113	
2473	0.05	8	1	0.5	5	0.5	200	0.661025703	0.318884802	
2474	0.05	8	1	0.5	7	0.5	200	0.665456665	0.328732862	
2475	0.05	8	1	0.5	9	0.5	200	0.665606806	0.331144405	
2476	0.05	8	1	0.5	11	0.5	200	0.662198011	0.325235241	
2477	0.05	8	1	0.5	13	0.5	200	0.668450384	0.339027118	
2478	0.05	8	1	0.5	15	0.5	200	0.6632736	0.333647653	
2479	0.05	8	1	0.5	17	0.5	200	0.668070894	0.339072044	
2480	0.05	8	1	0.5	19	0.5	200	0.667843633	0.340183143	
2481	0.05	8	1	0.5	1	0.5	200	0.661022053	0.315964077	
2482	0.05	8	1	0.5	3	0.5	200	0.664294471	0.324038885	
2483	0.05	8	1	0.5	5	0.5	200	0.66015221	0.316944899	
2484	0.05	8	1	0.5	7	0.5	200	0.663978062	0.326171118	
2485	0.05	8	1	0.5	9	0.5	200	0.665116567	0.329348484	
2486	0.05	8	1	0.5	11	0.5	200	0.664078174	0.326898165	
2487	0.05	8	1	0.5	13	0.5	200	0.666708991	0.335510917	
2488	0.05	8	1	0.5	15	0.5	200	0.667840706	0.338729557	
2489	0.05	8	1	0.5	17	0.5	200	0.66785171	0.338679683	
2490	0.05	8	1	0.5	19	0.5	200	0.669321208	0.343714311	
2491	0.05	8	1	0.5	1	0.5	200	0.658905525	0.312186996	
2492	0.05	8	1	0.5	3	0.5	200	0.66374988	0.322657071	
2493	0.05	8	1	0.5	5	0.5	200	0.663897138	0.324832526	
2494	0.05	8	1	0.5	7	0.5	200	0.663637148	0.325682623	
2495	0.05	8	1	0.5	9	0.5	200	0.664399309	0.327727251	
2496	0.05	8	1	0.5	11	0.5	200	0.664283175	0.328940403	
2497	0.05	8	1	0.5	13	0.5	200	0.666176492	0.33057153	
2498	0.05	8	1	0.5	15	0.5	200	0.666363519	0.335534440	
2499	0.05	8	1	0.5	17	0.5	200	0.669418118	0.342064464	
2500	0.05	8	1	0.5	19	0.5	200	0.66875301	0.343318518	
2501	0.05	8	1	0.5	1	0.5	200	0.660379827	0.31511533	
2502	0.05	8	1	0.5	3	0.5	200	0.661250568	0.318191778	
2503	0.05	8	1	0.5	5	0.5	200	0.664959511	0.325326486	
2504	0.05	8	1	0.5	7	0.5	200	0.66584188	0.329923137	
2505	0.05	8	1	0.5	9	0.5	200	0.664736968	0.329284053	
2506	0.05	8	1	0.5	11	0.5	200	0.66579267	0.332431354	
2507	0.05	8	1	0.5	13	0.5	200	0.667101544	0.338091828	
2508	0.05	8	1	0.5	15	0.5	200	0.666364767	0.338037882	
2509	0.05	8	1	0.5	17	0.5	200	0.668145976	0.340581491	
2510	0.05	8	1	0.5	19	0.5	200	0.66801168	0.341989852	
2511	0.05	8	1	0.5	1	0.5	200	0.663034272	0.320947789	
2512	0.05	8	1	0.5	3	0.5	200	0.663257887	0.321842896	
2513	0.05	8	1	0.5	5	0.5	200	0.664434497	0.326293376	
2514	0.05	8	1	0.5	7	0.5	200	0.664347448	0.327846172	
2515	0.05	8	1	0.5	9	0.5	200	0.666177052	0.331733378	
2516	0.05	8	1	0.5	11	0.5	200	0.664262421	0.329840409	
2517	0.05	8	1	0.5	13	0.5	200	0.665707011	0.335538035	
2518	0.05	8	1	0.5	15	0.5	200	0.667112471	0.337501684	
2519	0.05	8	1	0.5	17	0.5	200	0.667200128	0.338144782	
2520	0.05	8	1	0.5	19	0.5	200	0.66851383	0.344593907	
2521	0.05	8	1	0.5	1	0.5	200	0.66238629	0.319589504	
2522	0.05	8	1	0.5	3	0.5	200	0.664963123	0.324750587	
2523	0.05	8	1	0.5	5	0.5	200	0.65924376	0.316821743	
2524	0.05	8	1	0.5	7	0.5	200	0.662390208	0.326188827	
2525	0.05	8	1	0.5	9	0.5	200	0.665305519	0.330227111	
2526	0.05	8	1	0.5						