Predicting Nature of Default using Machine Learning Techniques

Longden, Elaine

Publication date:
2021

Document Version
Early version, also known as pre-print

Department of Finance
Tilburg School of Economics and Management

Predicting Nature of Default using Machine Learning Techniques

Elaine Hu Longden

April 2021

Abstract

This paper presents machine learning techniques to help financial institutions model a loan's nature of default and further incorporate nature of default in the prediction of loss given default models. Nature of default decribes the main reason why the lender puts the borrower in default. The comparison of different techniques show the decision tree approach as the best model, specifically the time it takes to default since a loan's origination is the most important feature in distinguishing different default types. We find loans with longer time to default are more likely to emerge to bankruptcy; whereas loans defaulted shortly after origination are more likely to be sold at a discount, resulting in a material credit loss. We also find that trade finance loans are more likely to receive a specific provision or write-off from the lending bank when they default, possibly due to the significant decrease in collateral valuations when the company is in financial difficulty. The nature of default is also found to be a significant factor in predicting loss given default. The unique insight this paper provides, when compared to similar default and loss studies in the existing literature, lies with its specificity in the loan's nature of default and its association with loss rates.

1. Introduction

Credit risk has always been one of the key components in risk management for financial institutions. Specifically, the default risk related to borrowers' inability to pay back credit obligation has inspired many practitioners in the industry to develop models that can predict and manage such risks. Since Altman's early publication of predicting corporate bankruptcy (Altman E. I. 1968) using financial ratios from company annual reports, there have been many studies focusing on assessing borrower's credit quality and predicting loan's default risk. Kealhofer (2003a), (2003b) and Vasicek (1984) used an extensive historical corporate default database from Moody's to estimate an empirical distribution of default and developed a structural model that predicts corporate's default probability. Jarrow-Lando-Turnbull (1997) used a Markov model to predict the term structure of credit risk spreads as an indication of a borrower's likeliness to default. Duffie-Singleton (1999) assumed randomness in the default events and simulated correlated default times and losses on portfolios of loans, bonds and other credit positions to predict the probability of default of large portfolios. There is an abundant amount of research and statistical models focusing on predicting the probability of default. With the development of machine learning techniques, there are also many recent publications focusing on modelling credit defaults using machine-learning algorithms, especially in the small and medium-sized enterprise (SME) and retail consumer credit space where more default data points are available. Khandani, Kim and Lo (2010) apply machine-learning techniques to build a nonlinear nonparametric model to predict consumer credit card defaults. Kim and Sohn (2010) use support vector machine to predict the default of SMEs with guarantees and found the model is better than the back-propagation neural networks and the traditional logistic regression. Li, Kolehmainen and Niskanen (2016) build a credit default hybrid model for SME lending by combining logistic regression and artificial neural networks. In the space of recovery rate modeling, Acharya et al. (2007) find that recovery rates are lower when the industry is in distress and the collateral assets are not easily redeployable by other industries. Betz et al. (2020) show the resolution time for defaulted loans is more than double in a recession when compared to an expansion economic environment. Keijsers et al. (2018) discover there are common cyclical components in default rate and loss given default of bank loans and the variation in severe or mild losses drives the cyclic behavior of loss given default.

Although most of these papers focus on the prediction of probability of default (PD) or loss given default (LGD), none of them focus on the nature of default or whether the nature of default of a loan affects the recovery rate. Nature of default decribes the main reason why the lender puts the borrower in default. For example, a default could be due to missing payments for more than 90 days; or filing for bankruptcy; or the sale of the loan at a discount; or restructuring of a distressed debt. Different nature of defaults could cause the recovery process and the recovery amount from the defaulted loan to be very different. The economic motivation to predict nature of default would be to better understand the risks related to the loss after a loan's default and better manage the portfolio before default occurs through limiting exposures with certain characteristics that could result in a specific type of default with higher loss rate.

Beyond the economic motivation, there is also a regulatory requirement. In fact, the topic of nature of default has only become a relevant topic in the industry after the publication of the new European Banking Authority (EBA) guidance on the application of the definition of default under the EU Regulation in 2016, where it specifically requires financial institutions to adjust historical data for the obligor's nature of default based on the new default definitions in the guidelines. Additional capital will be required as a margin of conservatism from financial institutions who fail to classify defaults into the specified categories under the new definition of default guidance. There are also different regulatory criteria for different nature of defaults to be allowed to cure (return back to a non-defaulted status). For example, for a distressed restructuring default the borrower needs to stay in a non-defaulted status for at least one year until it is allowed to cure, whereas for other types of default the minimum condition is only three months. Given all the regulations concerning the nature of default, it is essential for financial institutions to classify and model defaulted loans correctly to avoid additional capital required by the regulator due to misclassification of default types. A great amount of efforts has been made by financial institutions to improve the process of collecting default information since the publication of this regulation. The efforts being made ensure the nature of default data will be collected from now but does not solve the problem that the default information lost in the past still cannot be retrieved.

This paper presents different types of classification models that can help financial institutions predict the nature of default of loans. The motivation to experiment with machine learning techniques is to see if their performance is better compared to traditional statistical models. There are existing studies comparing model accuracy between logistic regression and machine learning classifers in the space of credit scoring (Baesens et al., 2015, Khandani et al., 2010); but no similar comparison studies have been done for the classification and prediction of nature of default. In addition, the default data very often contains a large amount of categorical variables and very few numeric variables, modeling using traditional statistical methods such as multinomial logistic regression would require a large number of dummy variables for each category under each variable. The computational time and the variable selection process could be extensive and time-consuming. Machine learning models are faster at handling large numbers of features and also use non-parameteric methods that consider all features in the model without the need for variable selection and any underlying assumption of the data distribution. It would be interesting to see whether machine learning models outperform traditional statistical models in the prediction of a loan's nature of default.

The paper also further explores whether nature of default as a variable can improve the prediction of loan losses when incorporated in the LGD model. We further examine which type of nature of default contributes the most in predicting LGD; for example whether loans with certain default type are more likely to end up in a higher loss. The paper therefore also contributes to the literature of modeling LGD.

The data used in this study is the default dataset obtained from Global Credit Data (GCD) Loss Given Default (LGD) & Exposure At Deafult (EAD) platform[1]. GCD is one of the very few organisations that provides comprehensive historical information for defaulted loans across multiple geographic regions and asset types, especially default data from borrowers that are not public listed companies. The GCD dataset has been frequently used in recent years for studies in LGD (Betz, Kellner & Rosch (2020)), EAD (Thackham & Ma (2019)) and systemic and macroeconomic effects in losses on bank loans (Keijsers, Diris & Kole (2018)). The consolidated data was collected from GCD's members banks from different regions and has a variety of features available with a historical period extending from 1983 to 2020, including the loan's nature of default. All the available features from the loan table in GCD's LGD & EAD platform are used in this study to predict the loan's nature of default.

The empirical estimation framework is set up based on the commonly used machine learning structure, which includes splitting the datasets into subsets for training and validation; performing cross-validation to train the models and finally evaluating the model performance to select the best model. Among the five candidate models, we find decision tree approach to be the best model, outperforming other models consistently using all evaluation criteria on both the training and validation set. We further observe that the feature time to default in the decision tree contributes the most in distinguishing different types of nature of default. The longer time it takes for a loan to default since its origination date, the more likely the nature of default is bankruptcy. This might be due to the fact that the decision to file for bankruptcy by the borrower and the procedure of handling the loan of a bankrupt company by the lending bank usually take much longer, which is different from defaults triggered by a more simple definition such as missing payments for more than 90 days or accrued interests unpaid exceeding a threshold. For loans defaulted shortly after origination, they are more likely to be sold at a material credit loss. This could be explained by credit officers' tendency to assign a high risk rating to borrowers with short history in the bank and defaulted shortly after orgination than loans with longer payment history. As a result, they are also more inclined to transfer the default risks off the book from the high risk loans through fire sale than to wait for the recovery process to finish. Lastly, we observe the inclusion of nature of default as an additional variable to an existing model with other variables slightly increases the prediction of LGD.

The rest of the paper is organized as follows: Section 2 describes the legal framework and background related to the nature of default; Section 3 presents the data for modelling; Section 4 describes in detail the estimation methodologies and the results from each model; Section 5 assesses the impact of nature of default on the prediction of LGD and Section 6 summarises the findings and future areas of study.

2. Legal framework

Since the European Banking Authority's (EBA) publication of the discussion paper on the "Future of the IRB approach" in March 2015, multiple papers detailing the new

[1] Global Credit Data (GCD) is an international "not-for-profit"-association owned by international banks and active in the pooling of historical credit data. https://www.globalcreditdata.org

Internal Rating Based (IRB) requirements and implementation timelines are published subsequently. Financial institutions have been working towards the deadline of 1 January 2021 to adjust the current IRB and standardized models in order to comply with the new EBA standards. As part of the EBA's regulatory review aiming to restore market participant's trust in internal models by reducing unjustified variability in the outcome, one of the key requirements is for banks to comply with the new definition of default. The EBA publication –Guidelines: On the application of the definition of default under Article 178 of Regulation (EU) No 575/2013 (Capital Requirements Regulation – CRR), was published in 2016 and later finalized in January 2017[2] . The Guidelines requires financial institutions to "where possible, adjust the historical data based on the new definition of default according to these guidelines", and " until an adequate time period with homogenous default definition is reached, …, to include an additional margin of conservatism in their rating systems in order to account for the possible distortions of risk estimates resulting from the inconsistent definition of default in the historic data used for modeling purposes.[3]"

Historically, banks follow their internal credit default policies to determine the default status of a counterparty. Under the new regulation, detailed classifications of the definition of default are described in the guidelines. There are generally two categories of default criteria – past due criterion and indications of unlikeliness to pay. In the past due criterion, the definition has always been less ambiguous because the days past due threshold is often set explicitly by local regulators for different asset classes[4]. In the default category of unlikeliness to pay, the definitions are not specifically clarified in the past and the new regulation aims to reduce this inconsistency across financial institutions by introducing specific default categories for unlikeliness to pay. A summary of the past due default and unlikeliness to pay default and their definitions is shown in Table 1.

	Default Type	**Definition of default**
Past due criterion	Counting of days past due	A borrower is in default if there are more than 90 days payment past due on any material credit obligation.
	Technical past due	Default due to loan payment past due triggered by bank's system error, failure of payment transaction order, or time delay of the transaction.
	Sovereign past due	Default due to loan payment past due from central governments, local authorities and public sector entities. A different past due criterion may apply other than the normal 90 days, but not longer than 180 days.

[2] Guidelines on the application of the definition of default under Article 178 of Regulation (EU) No 575/2013 (18/01/2017) https://eba.europa.eu/sites/default/documents/files/documents/10180/1721448/052c260f-da9a-4c86-8f0a-09a1d8ae56e7/Guidelines%20on%20default%20definition%20(EBA-GL-2016-07)_EN.pdf

[3] Chapter 3: Implementation, section 11 of the Guidelines on the application of the definition of default under Article 178 of Regulation (EU) No 575/2013 (18/01/2017)

[4] Basel Committee on Banking Supervision, QIS 3 FAQ: F. Definition of default / loss https://www.bis.org/bcbs/qis/qis3qa_f.htm

	Specific provision	Default due to a write-off or specific provisions put aside by the bank on the balance sheet for the future potential losses from a significant decline in credit quality.
Indication of unlikeliness to pay	Non-accrued status	Interests accrued from unpaid credit obligations and fees exceed the allowed threshold set by accounting rules and cannot be recorded as profit in the income statement, leading to a default status.
	Specific credit risk adjustment	Any specific credit risk adjustments resulting from a significant decline in credit quality such as loss recognized, loss incurred but not reported and credit-impaired exposures.
	Sale of asset	Any credit risk related sale of credit obligations and the economic loss of the sale is more than 5%.
	Distressed restructuring	Restructuring resulting in a diminished financial obligation of more than 1%; or other possible indications of unlikeliness to pay if below the 1% threshold.
	Bankruptcy	Bankruptcy defined in the institution's internal policy taking into account of all relevant legal framework.
	Other indications	Other indications such as source of income no longer available, breaching other credit contracts or any information in external databases to indicate unlikeliness to pay.

Table 1 – Definition of default under EBA guidelines (*Article 178 of Regulation (EU) No 575/2013*). The two main categories of defaults are past due criterion and indication of unlikeliness to pay.

3. Data

GCD's LGD and EAD platform covers default information at loan level from 1983 to 2020. It contains default data collected from 55 member banks across North America, Europe, Asia, Middle East, Oceania and Africa[5]. The loan portfolio covers large corporates, SMEs, banks and financial companies, sovereigns and central banks, public services, private banking and non-trade finance entities i.e., ship, aircraft, real estate finance etc. The available features from the loan level data contains categorical information such as the loan's asset class, region, type of the loan facility, seniority in the hierarchy as well as numeric information such as the time it has taken for the loan to default, to resolve or to have stayed in unresolved. Based on the full list of features available in GCD default dataset (

[5] Current members include ING, NIBC, Abn Amro, RBS, Barclays, UBS, Credit Suisse, J.P. Morgan, Citibank etc. https://www.globalcreditdata.org/about/member

Appendix B), a selective set of features are chosen based on data quality, relevancy and availability[6]. The estimation of the nature of default is performed at loan level with the final selected features in Table 5.

The field "Nature of Default" from GCD's LGD and EAD platform (Table 3) largely corresponds to the EBA regulatory default definition and can be used as the target variable for the prediction of different default types. Table 2 shows that nine out of the eleven definitions of default under the EBA guidelines can be mapped to a specific GCD nature of default.

	Default Definition under EBA Guidelines	**GCD Nature of Default**
Past due criterion	Counting of days past due	90 Days Past Due
	Technical past due	Not available
	Exposures to central government, local authorities and public sector	Available in entity asset class from the entity dataset
	Specific provisions applicable to factoring and purchased receivables	Not available
	Materiality threshold	Considered by member banks in the submission for all default types
Unlikeliness to pay	Non-accrued status	Non-Accrual
	Specific credit risk adjustment	Specific Provision or Charge-off
	Sale of asset	Sold at Material Credit Loss
	Distressed restructuring	Distressed Restructuring
	Bankruptcy	Bankruptcy
	Other indications	Unlikely to Pay

Table 2 – GCD nature of default to EBA definition of default mapping. 9 out of 11 default definitions from EBA can be mapped to a GCD definition. The detailed definition of each nature of default from GCD is in Appendix A.

Nature of default code	**Nature of default type**	**Definition in GCD**
200	90 Days Past Due	Obligor is more than 90 Days past due on any material credit obligation.
210	Unlikely to Pay	Obligor is unlikely to pay its credit obligations in full without recourse by the bank to actions such as realizing security. May include less than 90-days past due payment. In case a more specific default reason applies as well (200, 210 to 260) please choose the more specific default reason.
220	Bankruptcy	Either the Obligor has sought or been placed in bankruptcy protection or the Bank has filed for the Obligor's bankruptcy.
230	Charge-off or Specific Provision	Bank makes a charge-off or account-specific provision resulting from a significant perceived decline in credit quality subsequent to the bank taking on the exposure.

[6] The reasons for exclusion from the full list of variables are in Appendix B.

240	Sold at Material Credit Loss	Sale of a distressed credit obligation resulting in a material economic loss. Applies also to counterparties on financial markets, in case of sale of credit obligation or Closure of the counterparty accounts, without material credit loss (LGD near 0).
250	Distressed Restructuring	A loan which has had a restructuring of its credit obligation, resulting in a loss caused by the material forgiveness or postponement of principal, interest and any associated fee payments.
260	Non-Accrual	A loan which has been classified or placed on non-accrual status because the accrued interests and fees exceed the allowed threshold.

Table 3 – Definition of nature of default under the GCD LGD & EAD platform data. The default categories are filled in by the member banks based on these definitions and applied consistently across all datasets collected from different banks.

The distribution of the target variable nature of default in Figure 1 shows that most of the defaults concentrate in “90 days past due”, “Unlikely to pay” and “Non-Accrual”. These three classes contribute to more than 75% of the total defaults. The same is true when looking at nature of default by region, asset class, loan seniority, dummy indicator variables such as collateral and guarantee; except for asset class – Private Banking, where the highest number of defaults comes from “Specific Provision or Charge-off” (Appendix G).

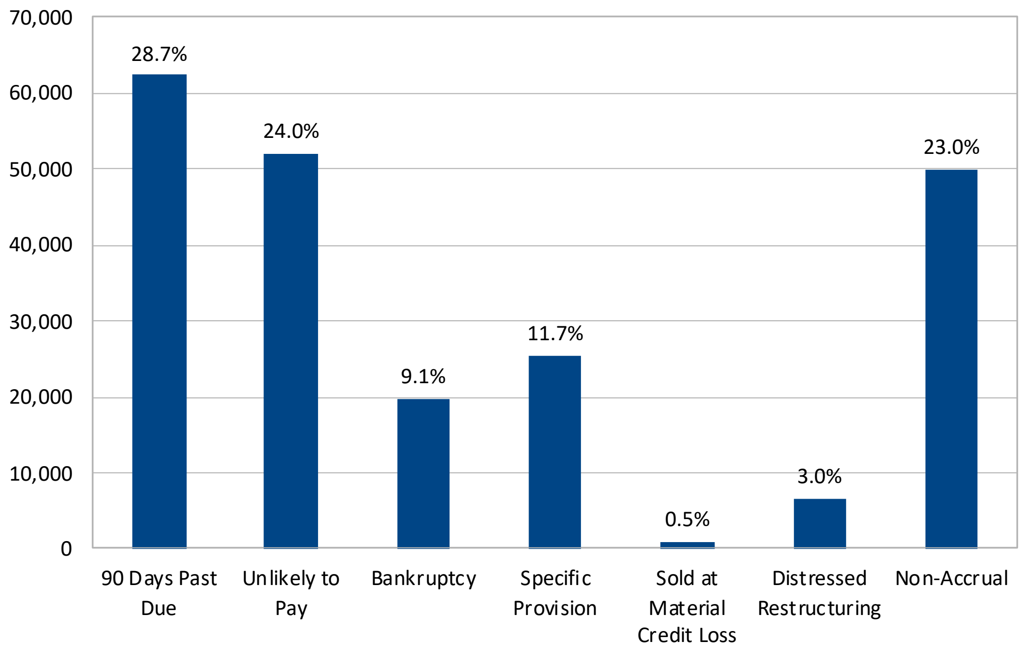


Figure 1 – Nature of default distribution in number of loan and in percentage of total number of loans after removing duplicates.

The three numeric features as shown in Table 4 are all related to the count of days since a default related event, i.e. default, resolved, unresolved. Among all the nature of default types, “Sold at Material Credit Loss” default type has the shortest average time to default since loan origination and “Bankruptcy” has the longest. “Sold at Material Credit Loss” default type also has the shortest average resolution time once in default; whereas “Unlikely to pay” has the longest. For the unresolved loans, “Sold at Material Credit Loss” again has the shorted average time in unresolved status since default whereas “90 days past due” default has the longest.

Nature of default	Numeric Feature	mean	std	min	median	max
90 Days Past Due	Time To Default	845	1,371	0	629	42,733
Unlikely to Pay		736	1,735	0	214	42,028
Bankruptcy		1,633	3,334	0	717	40,317
Specific Provision		1,478	2,832	0	871	41,885
Sold at Material Credit Loss		384	797	0	0	5,894
Distressed Restructuring		1,388	2,596	0	488	40,180
Non-Accrual		736	887	0	639	15,931
90 Days Past Due	Time To Resolution	615	741	0	355	6,698
Unlikely to Pay		800	844	0	534	7,082
Bankruptcy		758	823	0	492	7,307
Specific Provision		703	753	0	430	7,439
Sold at Material Credit Loss		304	373	0	210	2,731
Distressed Restructuring		675	722	0	465	7,120
Non-Accrual		853	839	0	580	8,303
90 Days Past Due	Time In Default Unresolved	298	1,028	0	0	8,419
Unlikely to Pay		108	539	0	0	7,522
Bankruptcy		242	783	0	0	6,940
Specific Provision		81	491	0	0	6,690
Sold at Material Credit Loss		22	227	0	0	3,561
Distressed Restructuring		212	771	0	0	7,014
Non-Accrual		83	462	0	0	6,056

Table 4 – Summary statistics for numeric features – time to default is the days between loan origination date and default date; time to resolution is the days between default date and resolution date; time in unresolved is the count of days after the default date to the reference date when the default data is reported. (Awaiting response from GCD on the large numbers in time to default)

Features	Description
Region	Region of jurisdiction of the loan mapped from feature Country of Jurisdiction.

Facility Asset Class	Asset class i.e., large corporates, SME, banks, FIs etc.
Tradeinf Finance Indicator	Binary dummy variable indicating whether the loan is trade finance facility. Only filled in for trade finance asset classes (Facility Asset Class = 1,2,3,8,9 or 10).
Guarantee Indicator	Binary dummy variable indicating whether the loan has guarantee or not.
Collateral Indicator	Binary dummy variable indicating whether the loan has collaterals or not.
Committed Indicator	Binary dummy variable describing whether a loan is committed or uncommitted. A committed loan refers to a loan that the bank is committed to offer whenever the request is made by the borrower; as opposed to uncommitted loans where banks can refuse to offer the loan.
Facility Type	Type of the loan facility i.e., bridge loan, revolver, term loan etc.
Seniority Code	Describe the hierarchy of the loan i.e., senior, subordinated, pari-passu etc.
Time to default	The number of days it takes for the loan to default.
Time to resolution	The number of days it takes for the loan to resolve entirely.
Time in default unresolved	The number of days the loan is in default but still unresolved.

Table 5 – Final features selected for training different types of machine learning and regression models. The full list of available features and reasons to exclude some features are detailed in Appendix B.

4. Classification techniques for nature of default

Given most of the existing methodologies concerning default prediction in credit risk are based on statistical regression models, this paper will mainly focus on machine learning techniques that can be used to predict nature of default. The following sections introduce four well-known classifiers in supervised learning, namely decision tree (DT), random forest (RF), K nearest neighbor (KNN), and Naïve Bayes (NB) where NB is used as a baseline model as the minimum benchmark to compare with other more complex models. A multinomial logit (MNL) regression model will also be used as a benchmark comparison between the regression approach and the machine learning approach.

Multinomial Logit

Given a set of N loans $L = \{(x_i, y_i)\}_{i=c}^{N,K}$, with input feature $x_i \in \mathbb{R}^N$ and K nature of default classes where each loan i belongs to one nature of default class c; the MNL regression is built to predict the probability of a defaulted loan belonging to a specific nature of default as follows:

$$P(y_i = c) = \frac{e^{\beta_c X_i}}{\sum_{c=1}^{K} e^{\beta_c X_i}}$$

whereby $P(y_i = c)$ is the probability of a defaulted loan belonging to nature of default class c; β_c is the model coefficients for class c; X_i is the input feature for loan i in the regression. The parameter β_c is typically estimated using the maximum likelihood estimation. The regression predicts a probability for each nature of default class and the class with the highest probability will be the final predicted nature of default.

Naïve Bayes

NB classifier is a quick and simple classification algorithm that is often used as a baseline model to compare with other more complex machine learning models. It is built based on the Bayes' theorem and assumes that each feature is independent of each other and each contributes to the outcome equally. The probability of a loan belonging to class c given feature x is the posterior probability and is defined as:

$$P(y_i = c|x_i) \ = \frac{P(x_i|y_i = c)P(y_i = c)}{P(x_i)}$$

where $P(y_i = c)$ is the prior probability of class c calculated as the ratio of number of loans in nature of default class c and total number of loans; $P(x_i)$ is the evidence as the probability of a loan with feature x; and $P(x_i|y_i = c)$ is the probability of a loan with feature x_i given it belongs to nature of default class c . The formula can be expanded to represent multi-features for nature of default class c given feature vector x_1 to x_n:

$$p(y = c|x_1, \dots, x_n) \ = \frac{P(y = c) \prod_{i=1}^{n} P(x_{\mathrm{i}}|y = c)}{P(x_1, \dots, x_n)}$$

NB classifier assigns the class that has the highest posterior probability as the final predicted class. Despite the simple assumption, NB works quite well in many classification problems and produces extremely fast results compared to more sophisticated methods (John & Langley 1995). This is due to the fact that the class independence assumption allows the NB classifer to quickly train high dimensional features with limited training data. All the input probabilities in the NB classifer can be very quickly calculated by simply counting the corresponding number of classes and there is no hyperparameter to tune in the model.

Decision Tree and Random Forest

DT is an intuitive method to classify the target variable nature of default into different categories through sequential decision process. A decision tree starts with the root node containing all data in the training set, the node is then split into sub-nodes called decision node using the selected splitting criteria. Given dataset $D = (x_i, y_i)$ where $i = 1 \dots n, x_i \in \mathbb{R}, y_i \in Y = \{1, \dots k\}$, n is the total number of loans, k is the total number of nature of default classes, for each feature $j = 1 \dots d$, for each splitting value $v \in \mathbb{R}$, the decision tree split the dataset as:

$$I_< = \{i: x_{ij} < v\} \text{ and } I_> = \{i: x_{ij} < v\}$$

whereby $I_<$ and $I_>$ are the left and right node from the split; the estimated probability of each nature of default class k in decision leaf R_j is:

$$P_{jk,<} = \frac{\sum_i \mathbb{I}(y_i = k) \cdot \mathbb{I}(x_i \in I_< , \mathrm{R}_j)}{\sum_i \mathbb{I}(x_i \in I_< , \mathrm{R}_j)} \ and \ P_{jk,>} = \frac{\sum_i \mathbb{I}(y_i = k) \cdot \mathbb{I}(x_i \in I_> , \mathrm{R}_j)}{\sum_i \mathbb{I}(x_i \in I_> , \mathrm{R}_j)}$$

whereby $P_{jk,<}$ and $P_{jk,>}$ denote the probability of the left and the right node; the numerator is the frequency of loans belong to nature of default class k in leaf node R_j and the denominator is the total number of loans in leaf node R_j. The quality of the split is measured by the splitting criteria, which evaluates how efficient and accurate the split of the tree is by assessing the target variable's homogeneity in each node – in this case how many homogeneous nature of default classes there are in each node. The most commonly used splitting criteria for classification problems is entropy, which is defined as:

$$E(j) = \sum_{c=1}^{K} -p_c log_2 p_c$$

where $E(j)$ is the entropy of a leaf node j, K is the total number of nature of default classes, p_c is the probability of default class c. The lower the entropy, the higher the purity and the lower the randomness in the node. Entropy is measured between 0 and 1 and a pure homogeneous node will have 0 entropy. Entropy is calculated for each node in each split in order to select the best split among all available features; then it is weighted by the number of defaulted loans in the node to obtain the weighted average value for the split. The split with the lowest entropy is then selected and the splitting process is repeated until a pure class of nature of default is achieved in a node.

For a tree to stop splitting naturally, all terminal nodes need to belong to one pure class of nature of default. However, this is often not easily achievable and requires a lot of computation time. More importantly, even if it is achieved, the tree might grow too deep and might be overfitting the training data. Therefore, stopping criteria are needed to stop the tree from growing too deep. A decision node becomes a terminal node when the stopping criteria is met, and no further split is needed. The stopping criteria, also known as the hyperparameters of the model, are selected based on whether the criteria can eventually stop the tree from growing. The typical stopping criteria in a decision tree are maximum depths of the tree, minimum samples in a leaf, maximum features, maximum leaf nodes etc. Some stopping criterion is binding when tested alone but might become non-binding and impose no effect on stopping the tree when tested with other criteria, so all criteria needs to be tested in conjunction with each other for them to be effective.

RF classifier is similar to DT and reduces the correlation between samples by randomly splitting the population into multiple samples to build individual decision trees. Each individual tree in RF gives a predicted class and the most predicted class is the final prediction of the model. The rationale behind RF is that each tree is independent and uncorrelated to each other. The advantage of uncorrelated decision tree is that their errors are also uncorrelated. The downside of RF is that if the features have no predictive power, i.e., no features exist to differentiate the classes, then the result is a random guess and the final predicted class from the majority vote will be different every time the model is run. It is therefore essential that the predictions are tested through trials of different hyperparameters to make sure the outcome is reliable.

K-Nearest Neighbour

KNN is a non-parametric algorithm that classifies data points into similar groups by considering the K closest data. It then assigns the most frequent class among its k nearest neighbours. The steps of running a KNN algorithm is to first calculate the distance between the target point, a defaulted loan with a specific nature of default, and every other point in the training dataset; then to initialise a value of k and find the nearest k points to the target point using the selected distance function. Finally, the most frequent class from the k nearest points will be the predicted class for the target point. The distance function and the number of neighbours k are the two hyperparameters that need to be tuned when training the model.

Among the various types of distance functions, some of them such as Euclidean distance and Manhattan distance are suitable when the magnitude between vectors matters more; some of them such as cosine distance are more suitable when the orientation matters more than the magnitude; the others such as Mahalanobis distance is more appropriate for calculating two data points in a multivariate space. The generalisation of all distance function is called Minkowski distance as defined below where x and y represent two random points in the sample; n is the number of pairs and p is the order of the function.

$$d(x,y) = (\sum_{i=1}^{n} |x_i - y_i|^p)^{\frac{1}{p}}$$

For the KNN algorithm, the magnitude between vectors is more important than anything since the classification is drawn based on this distance. The Euclidean distance function (p = 2) which measures a straight line from one point to another point in a two-dimensional space is the most commonly used distance function in KNN. The Manhattan distance (p = 1) is also common but more suitable for higher dimension problems. Given the number of features available in the dataset is not too large, the Euclidean distance function is considered suitable.

Evaluation Criteria for Classification

Accuracy

Accuracy is defined as the ratio of correct predictions to the total number of predictions. It works well when the number of observations is balanced in each class. When there is a high concentration of observations in a particular class, the accuracy tends to be high for the most frequent class and low for the less frequent class even in the case when the classification model randomly allocates observations to each class. This is the case in the default dataset for nature of default classification where the “90 days past due”, “Unlikely to pay” and “Non-Accrual” constitute 76% of the total defaults and the rest only account for a small portion.

To remediate the over-estimation of accuracy in the most frequent class, balanced accuracy score is used instead for datasets with unbalanced classes (Table 6). The metrics from the confusion matrix can be used to explain the difference between the two. Balanced accuracy adjusts the accuracy so that it does not give an artificially high score when the model is simply predicting the most frequent class for all observations.

Confusion Matrix	Actual Positive	Actual Negative
Predicted Positive	True Positive (TP)	False Positive (FP)
Predicted Negative	False Negative (FN)	True Negative (TN)

Table 6 – Confusion matrix

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

$$Balanced\ Accuracy = \frac{Sensitivity + Specificity}{2}$$

Where:

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

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

Log Loss

Log loss, also known as logistic loss, is an evaluation metric that computes the cross-entropy between the true distribution of the class and the distribution from the prediction. It ranges from zero to one and when the predicted distribution exactly matches the true distribution, log loss is zero. It works well for models that predicts a probability of a class occurring rather than directly predicting the class. In a multi-class classification, the log loss is computed as:

$$Log\ loss = -\sum_{c}^{K} y_c \log{(p_c)}$$

where y_c is the true distribution and p_c is the predicted probability of class c. Compared to accuracy, which counts the number of predictions matching the actual class, log loss takes into account the uncertainty of the prediction based on how much it deviates from the actual class. However, log loss can only be calculated for probability-based estimations and not all of the classification algorithms are probability based.

Precision, Recall and F1 Score

Precision is another evaluation metrics derived from the confusion matrix. It ranges from 0 to 1 with the best outcome being 1 (no false positives) and the worst outcome being 0 (no true positives). Precision is defined as the percentage of correctly predicting the positives, meaning a specific type of default in this study, out of all the positives predicted.

$$Precision = \frac{TP}{TP + FP}$$

Another metric closely linked to precision is recall, which is defined as the percentage of correctly predicting the positives out of all the observations that are actually positive. F1 score makes use of both precision and recall and gives not only an indication of how precise the classifier is but also how robust it is. A model can have a high precision but low recall, ignoring a significant number of observations that should have been classified as positive. F1 score takes into consideration of both precision and recall and finds the balance between the two and is the preferred evaluation criteria in this paper.

$$F1\ score = 2 \times \frac{1}{\frac{1}{Precision} + \frac{1}{Recall}}$$

where:

$$Recall = \frac{TP}{TP + FN}$$

In multi-class classification problem, overall accuracy, precision and F1 score are the same when using a micro-averaging approach to aggregate across all classes. Therefore, precision and F1 score are reported as the weighted average of each class instead for the overall classification performance in this paper.

Feature Importance

Feature importance is used to assess the feature that contributes the most in correctly predicting the nature of default class. The higher the feature importance value is, the more important the feature. The importance of a feature is computed as the normalized total reduction of the selected criterion (i.e., entropy) brought by that feature[7]. The mathematical formula to interpret the calculation is:

$$FI(j) = \frac{n_j}{N} \times [C_j - \frac{n_{j,R}}{n_j} \times C_{j,R} - \frac{n_{j,L}}{n_j} \times C_{j,L}]$$

where FI(j) is the feature importance at node j; N is the total number of samples; n_j is the number of samples at node j; $n_{j,R}$ is the number of samples from the right subsequent child node of node j and $n_{j,L}$ is the left child node sample size; $C_{j,R}$ or $C_{j,L}$ is the entropy of the corresponding right or left child node that is used to calculate impurity from the node.

[7] Friedman, J.H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics, 1189-1232.*

5. Predicting nature of default

Estimation framework

Once the target variable nature of default and the features data as shown in Table 5 are prepared, different data samples are constructed by splitting the default population into three subsets – the training set, the test set and the validation set. The training set is used to train the model using different parameters while the test set is used to tune these parameters and to perform evaluations on different models so that the best model can be selected. The validation dataset is an independent dataset that has not been used to train the parameters or tune the models. It is an unseen dataset used to give an unbiased evaluation of the final model performance. While there is no ideal ratio to split between training set and validation set, the rule of thumb is 75%/25% split (Guyon 1997). There are in total 217,112 default data points and 25% of the total default data points are randomly selected and left out for validation purpose, whereas the 75% is then split into training and test set using cross validation method. The duplicated features are removed after the splitting and the final proportion for training and validation set is 73% and 27% respectively (Table 7).

Dataset	# of Loan
Total data points	217,112
After removing duplicates	174,804
Training set (5-fold)	127,356
4 folds to train	101,885
1 fold to test	25,472
Validation set	47,448

Table 7 – Summary count of total data points, training set including the 5-fold training and testing sets; and tbe validation datasets. 5-fold training set is used to select features and the best model; validation set is used to assess model performance.

Cross validation is used to evaluate the model on unseen data by resampling multiple times from the population. This method is applied to the training set data to further split it into cross validation sets. There are different types of cross-validation methods, ranging from the simplest holdout method to the most extreme leave-one-out method. The holdout method splits the population into two datasets and the model is built on the training set and evaluated using the validation set. K-fold method splits the data into k subsets and uses one subset as the validation set and the rest k-1 subsets as the training set. The process is repeated k times and the best model is selected. Leave-one-out method is similar to the K-fold but with k equals to the total number of data points. Each time one data point is left out while the rest is trained until all data points have been left out once.

The most commonly used method is the K-fold cross validation. Given the 75% of the default dataset still has quite a large number of data points, leave-one-out cross validation will take too much computation time. Holdout method only splits the data once into two sets and it doesn't ensure any data points in the data have equal chance of being in the training set and test set, so the results are still dependent on how the data is split. Given the trade-off between computation time and estimation accuracy, K-fold cross validation seems to be the most reasonable method to use. A 5-fold cross validation is chosen by splitting the training set sample equally into 5 subsets, leaving enough samples in each subset to train and validate the models. Table 8 shows the proportions of loans by nature of default in the training set for each fold are similar, ensuring no bias exist in the target variable for each fold while training the models .

	K fold	90 Days Past Due	Unlikely to Pay	Bankruptcy	Specific Provision	Sold at Material Credit Loss	Distressed Restructuring	Non-Accrual
Training	Fold 2+3+4+5	28,745	23,139	10,365	12,489	362	3,155	23,629
	Fold 1	7,338	5,770	2,456	2,898	107	797	6,106
	Fold 1+3+4+5	28,829	23,185	10,252	12,377	381	3,185	23,676
	Fold 2	7,254	5,723	2,569	3,010	88	767	6,060
	Fold 1+2+4+5	28,845	23,362	10,365	12,265	365	3,120	23,563
	Fold 3	7,238	5,836	2,456	3,122	104	832	5,883
	Fold 1+2+3+5	28,948	23,300	10,121	12,232	381	3,193	23,710
	Fold 4	7,135	5,535	2,700	3,155	88	759	6,099
	Fold 1+2+3+4	28,965	23,428	10,181	12,185	387	3,155	23,584
	Fold 5	7,118	5,661	2,640	3,202	82	797	5,971
Validation		13,480	11,334	10,916	5,567	4,597	1,373	181

Table 8 – Number of observations in the training and validation dataset per nature of default group. Training set consists of 5 folds; each time 4 folds are used for training 1 fold is used for testing. The number of observations in each nature of default group are similar in each set.

Since most of the candidate machine leaning models would require one or more initial hyperparameters to run the model, the model parameter control is needed to run different trials of hyperparameters. A hyperparameter is a parameter used to control the learning process of the machine learning algorithm, which is usually constraints or stopping criteria put in place for the learning process to stop. In the initial run, GridSearch is used to automatically find the optimal hyperparameters using 5-fold cross validation. The initial values are shown in Table 9 and they often tend to overfit the training set, showing by the outperformance of the training set compared to the validation set based on the evaluation metrics. Therefore, manual hyperparameter tuning is required to reduce overfitting in the training set. Note that MNL model and the baseline NB model do not have hyper-parameters to tune, so they're not included in Table 9. The optimal hyperparameters from the parameter tuning are used to run the final models and the performance is assessed using the evaluation criteria. The final results are shown in the next section.

Model	Hyper-parameter	Initial value	Dataset	Balanced Accuracy	Log Loss	Precision	Recall	F1 Score
DT	Minimum sample leaf[8]	10	Training	66%	0.94	67%	61%	62%

[8] The minimum number of observations required to be in a leaf node. The split is only implemented if both the left and right branch of the tree reach the minimum sample size.

	Maximum depth[9]	15	Validation	59%	1.77	65%	58%	60%
RF	Number of estimators	500	Training	63%	1.17	65%	60%	61%
	Maximum depth	20	Validation	58%	1.19	65%	58%	60%
	Minimum sample leaf	12						
KNN	Number of neighbours	50	Training	25%	1.43	40%	41%	37%
	Power parameter P	2	Validation	24%	1.68	39%	40%	37%

Table 9 – Initial hyper-parameters in each model and their model performance in training and validation set measured by balanced accuracy, log loss, precision, recall and F1 score using these hyper-parameters.

Results

Table 10 and Figure 2 report the training and validation set balanced accuracy, log loss, precision, recall, F1 score and the confusion matrix of all classifers. Based on these evaluation metrics, we found that DT gives the best performance, and the predictions contain all the seven classes of nature of default. RF also has a relatively good performance compared to the other models but slightly worse than DT. NB stands in the middle and has better performance than MNL and KNN. KNN has the poorest performance and is also the least preferred classifer due to its long computation time and highly sensitive outcome to the number of neighbours chosen.

The main advantage of DT and RF is that the algorithm has the flexibility to adjust for the unbalanced class when training the model. By assigning a higher weight to the less frequent class[10], the predictions are less biased to the most frequent class. This can be seen from the confusion matrix (Figure 2) that the percentage of predictions in the least frequent class - "Sold at Material Credit Loss" and "Distressed Restructuring – is much higher from DT and RF than from the other models. KNN does not produce any predictions and MNL and NB produce very low percentage of the predictions in the least two frequent classes. DT and RF use balanced class weight to ensure the predictions are not heavily skewed to the most frequent class, giving more reliable prediction outcome than the other models.

Model	Hyper-parameter	Final value	Dataset	Balanced Accuracy	Log Loss	Precision	Recall	F1 Score
MNL	-	-	Training	35%	1.28	51%	51%	50%
	-	-	Validation	35%	1.28	51%	52%	50%
NB	-	-	Training	39%	1.51	49%	48%	47%
	-	-	Validation	39%	1.51	49%	48%	48%

[9] The maximum depth of the tree, i.e., root note's depth is 0. Depth larger than 30 is usually not reliable. The tree stops splitting once the maximum depth is reached.

[10] The "balanced" mode in the sklearn decision tree and random forest classifier uses the values of y to automatically adjust weights inversely proportional to class frequencies in the input data as n_samples / (n_classes * np.bincount(y))
https://scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeClassifier.html

DT	Minimum sample leaf	300	Training	55%	1.22	61%	53%	55%
	Maximum depth	20	Validation	53%	1.27	61%	52%	55%
RF	Number of estimators	500	Training	52%	1.47	59%	49%	50%
	Maximum depth	20	Validation	50%	1.48	59%	48%	50%
	Minimum sample leaf	300						
KNN	Number of neighbours	1000	Training	20%	1.57	34%	35%	31%
	Power parameter P	2	Validation	30%	1.57	34%	35%	31%

Table 10 – Final hyper-parameters in each model after parameter tuning. Details of different trials of hyper-parameters for DT, RF and KNN are in Appendix A. Comparison of model performance among the final models show DT has the best performance using the five evaluation metrics.

Looking more specifically at MNL model, Table 11 shows that although most of the variables are significant at 5% level, the pseudo R-squred is low at 22%[11]. The confusion matrix for MNL also shows that most of the predictions concentrate in the most frequent classes – “90 days past due”, “Unlikely to pay” and “Non-Accrual”. The model doesn’t perform well for the low frequency default classes. This is also the case for NB and KNN. The highest precision (true positives) for MNL is in default class “Unlikely to pay”, whereas for NB and KNN it is in “Non-Accrual” and “90 Days Past Due” respectively. The percentage of predictions in each nature of default class from DT and RF is more evenly spread out and the average precision levels are also higher than the other models (Table 10). DT and RF both predicts “Sold at Material Credit Loss” with the highest precision and the average precision are approximately 60%.

The overall performance from the training set and validation set are very similar for all models, indicating there is very little overfitting in the training set and the hyper-parameters selected for the final models are reliable. Details of the different hyper-parameters used in the trails are in Appendix A. The best performance from both the training and the validation set is DT, with the highest balanced accuracy, precision, recall and F1 score; and the lowest log loss among all other classifiers. The most important feature in the DT classifier is Time To Default.

[11] The full MNL regression summary including all variables can be found in Appendix

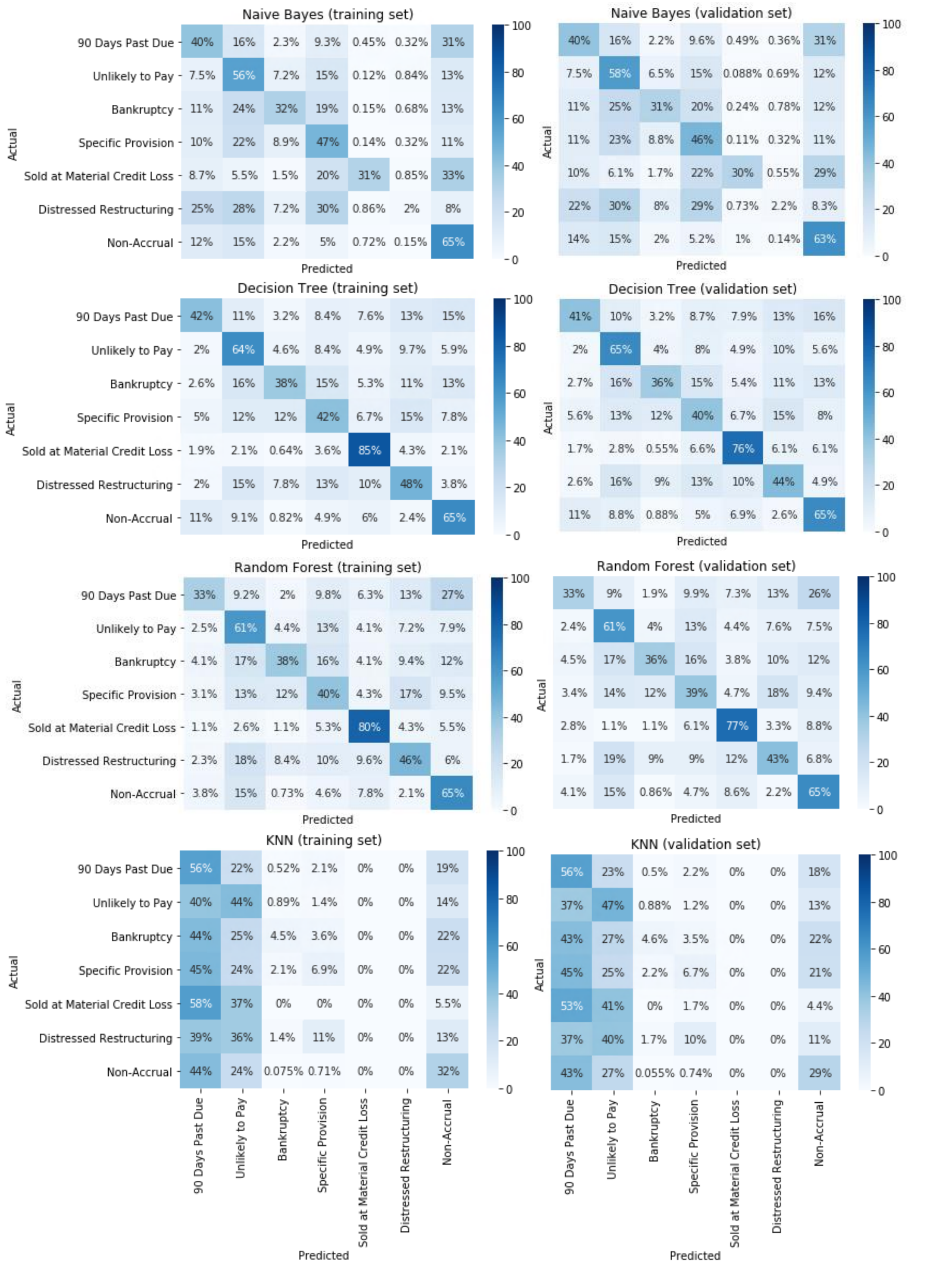


Figure 2 – Confusion matrix per nature of default class for training and validation set; diagonal of the matrix shows the percentage of correctly predicting the class.

The feature importance is calculated as the normalized total reduction of entropy brought by each feature as described in section 4; the higher the feature important value is, the more reduction in entropy thus the more important the feature is. Figure 3 shows that the time it takes for a loan to default since origination and the time to resolve the defaulted loan play more important roles in distinguishing different types of nature of default than other features. The interpretation from this could be that certain types of default take less time to happen and to resolve compared to other types.

Specifically, loans that are sold at material credit loss after default has the shortest average time to default (384 days) since loan origination; they also take the least time to resolve (304 days) as shown in Table 4 and Appendix . This could indicate that if a loan has defaulted very soon after its origination, credit officers are more likely to transfer the risks off the book through fire sale than to keep it on the book and wait out for the recovery and resolution process. The time to default is the longest for loans defaulted due to bankruptcy, on average 4.5 years. For distressed restructuring and default due to a charge-off /write-off or account-specific provision made by banks, the time to default is also longer, approximately 4 years. These types of defaults often indicate there are serious financial problems and significant deterioration in credit quality of the borrower; and they are often consequences of management decisions and require legal procedures and actions. On the contrary, defaults due to violating clearly defined default definitions without the need for management decisions i.e. 90 days past due, unlikely to pay under internal credit policy and non-accrual interests under the accounting rules[12], take shorter time to happen, with approximately 2 years time to default. It is reasonable to see that it would take longer time since loan origination for a company to declear bankruptcy or distressed restructuring, or for the lending bank to decide to put aside provisions for a loan because there is a significant decrease in the collateral value or even no chance of any recovery in the future. The average resolution time for different types of nature of default is approximately 2 years, ranging from 1.7 to 2.3 years, except for defaulted loans sold at material credit loss, which only takes 1 year to resolve.

Note that feature “Time To Default” and “Time To Resolution” are both only observable after the default happens or resolved. Although they can still be used in predicting the nature of default class for defaulted loans, they cannot be used for performing loans as the data is not available. Therefore, it is useful to understand the next important feature in the line, “Trade Finance Indicator”. Trade finance loans are speciliased lending in shipping, aircraft, real estate or a specific project. We observe that while both trade finance and non-trade finance loans have the majority of default types in “90 Days Past Due” and “Unlikely To Pay”, trade finance loans have a higher percentage of default in “Specific Provisoin” (Figure 4). This might be due to the fact that trade finance collaterals i.e. the ships, aircrafts or properties pledged to the bank, are more likely to receive a decrease in valuation when the company is in financial difficulty, which is a unique feature of trade finance loans that could lead to the bank‘s decision to reserve a specific provision or make a write-off.

[12] Banks are required to declare defaults if the unpaid interest and fees exceed a certain amount.

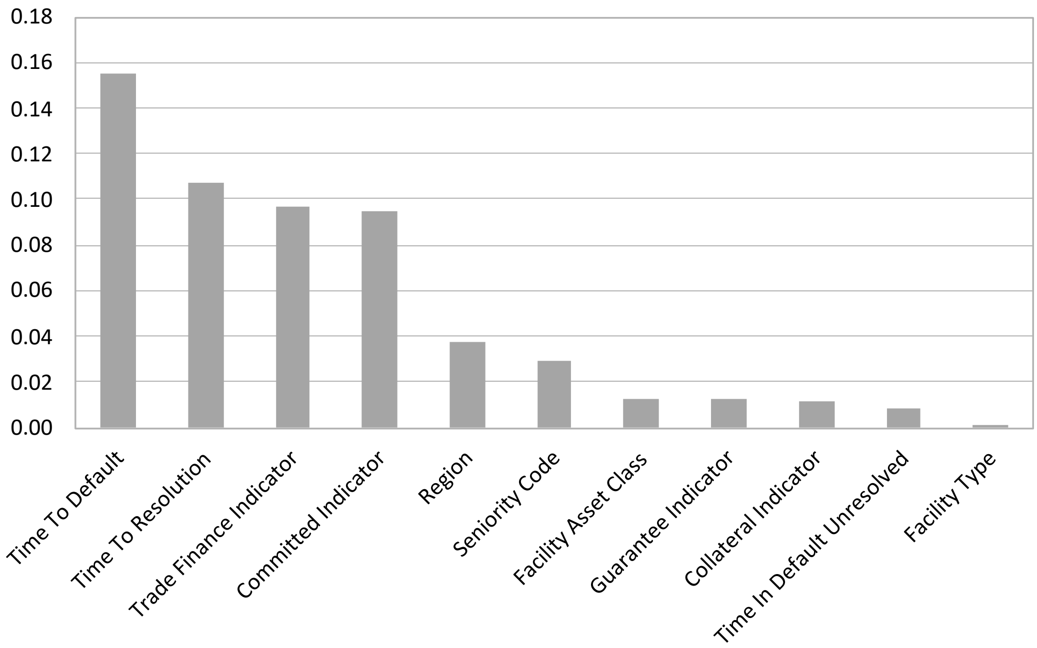


Figure 3 – Feature importance from the final selected model DT, ranked from the most important feature to the least important.

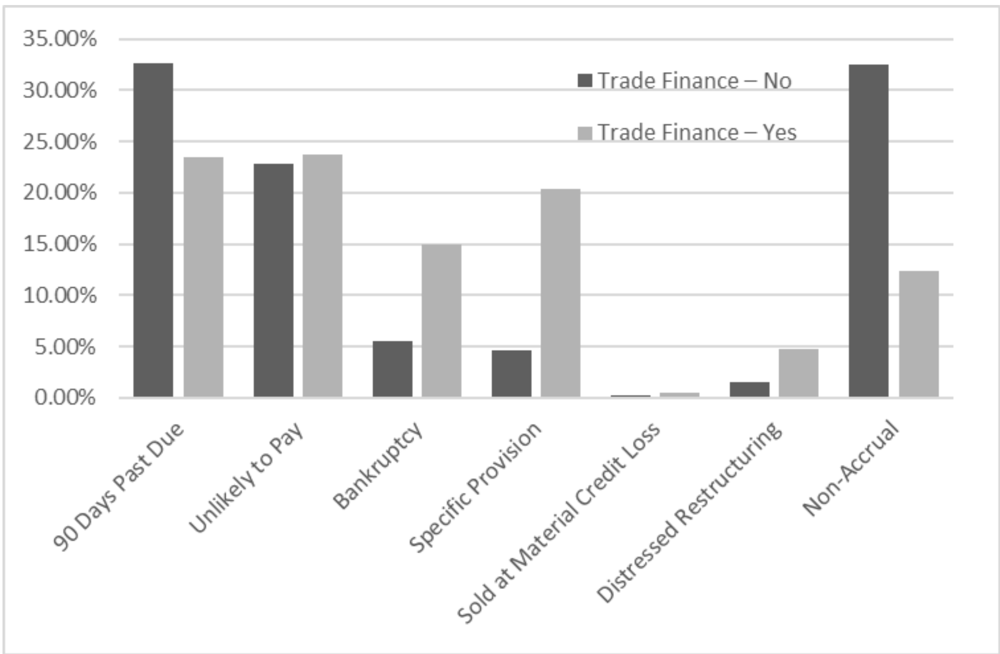


Figure 4 – Distribution of nature of default class for trade finance and non-trade finance loans.

Variable	(1)		(2)		(3)		(4)		(5)		(6)	
Intercept	0.1326	***	0.0543	***	0.0290	***	0.0011	***	0.0008	***	1.7614	***
	(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)	
Collateral Indicator	2.1359	***	1.0175		1.0059		1.5627	***	1.7236	***	1.9077	***
(Yes/No)	(0.00)		(0.10)		(0.65)		(0.00)		(0.00)		(0.00)	
Committed Indicator	0.1749	***	0.1771	***	0.5706	***	0.9552	***	0.7874	***	0.2802	***
(Yes/No)	(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)	
Guarantee Indicator	1.1807	***	1.2219	***	0.7405	***	0.5527	***	0.9661	***	0.8994	***
(Yes/No)	(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)	
Time In Default Unresolved	0.9998	***	1.0000	*	0.9996	***	0.9995	***	0.9999	**	0.9998	***
	(0.00)		(0.01)		(0.00)		(0.00)		(0.01)		(0.00)	
Time To Default	0.9999	***	1.0001	***	1.0001	***	1.0000		1.0001	***	0.9999	***
	(0.00)		(0.00)		(0.00)		(0.09)		(0.00)		(0.00)	
Time To Resolution	1.0002	***	1.0003	***	1.0001	***	1.0000		1.9991	***	1.0004	***
	(0.00)		(0.00)		(0.00)		(0.56)		(0.00)		(0.00)	
Trade Finance Indicator	1.7603	***	4.3939	***	6.1653	***	1.5384	***	4.6594	***	0.4884	***
(Yes/No)	(0.00)		(0.00)		(0.00)		(0.00)		(0.00)		(0.00)	
Facility Asset Type FE	Yes		Yes		Yes		Yes		Yes		Yes	
Facility Type FE	Yes		Yes		Yes*		Yes		Yes		Yes	
Seniority FE	Yes		Yes		Yes		Yes		Yes		Yes	
Region FE	Yes		Yes		Yes		Yes		Yes		Yes	
No. of observations	127,356											
Pseudo R-squared	22.3%											

Signif. Codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' '

* All coefficients are significant at 5% for Facility Type except for "Term Loan" (p-value = 0.11) for Specific Provision

Table 11 – MNL model summary for nature of default class – (1) Unlikely to pay; (2) Bankruptcy; (3) Specific Provision; (4) Sold at Material Credit Loss; (5) Distressed Restructuring; (6) Non-Accrual.

6. Impact on LGD

Nature of default driven by different causes could have an impact on the amount that could be lost after the loan default. In order to assess if nature of default has an impact on the loss given default (LGD) and whether it adds additional value to the prediction of LGD, a regression model is built using the same variables available in the default dataset (Table 5)[13]; then nature of default is added as an additional variable to see if the model performance can be improved. The loss from a defaulted loan can be modeled through different parametric approaches. Here we use a simple linear regression to test if nature of default as a variable is significant in the regression and whether it can improve model performance. The ordinary least square (OLS) full regression results are in Appendix J and Appendix K.

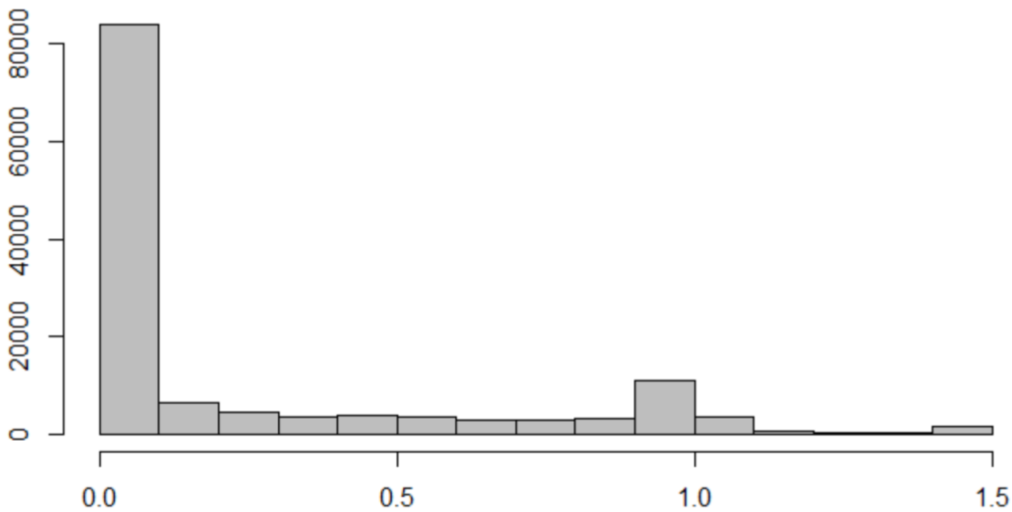


Figure 5 – Histogram of observed LGD with values ranging from 0% to 150%; values above 100% due to additional fees, interests accrued and cost related to the resolution of the defaulted loan.

Figure 5 shows the observed LGD distribution and most of the values concentrate on 0% and 100%. LGD above 100% are likely due to fees and interests accured and also the cost of selling the collaterals or debt collections. Table 12 shows the OLS regression summary for predicting LGD with and without nature of default. All of the nature of default dummy variables are significant when included in the OLS regression, except for “Unlikely to Pay”. The model with nature of default has a higher adjusted R-squared and smaller residual standard error. It is observed that all types of nature of defaults have a positive effect on LGD; defaulted loan “Sold at Material Credit Loss” increases predicted LGD the most, by approximately 24%; followed by “Bankruptcy” with an increase of 13%; whereas “Unlikely to Pay” default increases only 0.6% of the predicted LGD. This finding is consistent with the realised LGD where these two types of nature of default also have the highest realised LGD and “Unlikely to Pay” has the lowest. The results can be intuitively explained based on the definitions of these defaults – “Sold at Material Credit Loss” default is recorded when the loan is sold at a large discount, often more than 10%; “Bankruptcy” is recorded when there is a court proceeding such as the filing of Chapter 11 in the U.S. and appointment of a liquidator

[13] Time_in_default_unresolved is not used in the regression due to too many NAs because only a small portion of loan in the data is unresolved.

in the U.K. These two types of default usually end up with larger loss given the nature of default; whereas “Unlikely to Pay” is a default type determined by individual bank’s internal policy depending on how the banks perceive the credit quality of the borrower and whether there is a significant deterioration of the credit quality. The decision to categorise a borrower as “Unlikely to Pay” does not trigger a legal action against the borrower and sometimes the borrower could return to performing and pays off the loan. Compared to the other nature of default types, “Unlikely to Pay” default often involves credit offiers’ assessment and judgement and is more often a precaution put in place by banks.

The model results also show collateralised or committed loans (such as term loan and revolving credit facility) receive a negative effect in the predicted LGD, meaning the expected loss is smaller when there is collateral or the loan is a committed facility. While the predicted LGD doesn’t decrease if a loan is guaranteed, the amount of increase indicated by their coefficients is smaller compared to non-guaranteed loans. This is the same for trade finance loans. We also find that the longer it takes for a loan to default, the lower the LGD; whereas the longer the resolution time after the default, the higher the LGD. Although both “Time to Default” and “Time to Resolution” are significant variables in the regression, they can only be observed after the default and after the recovery process is completed; so it would be difficult to use them as inputs in the model for performing loans.

Variables	(1)		(2)	
(Intercept)	0.0678	**	0.1720	***
	0.0015		< 2e-16	
Guarantee Indicator (No)	0.1042	***	0.0543	***
	0.0000		0.0012	
Guarantee Indicator (Yes)	0.0444	**	-0.0107	
	0.0093		0.5265	
Collateral Indicator (No)	0.0710	***	0.0748	***
	< 2e-16		< 2e-16	
Collateral Indicator (Yes)	-0.0034		-0.0078	
	0.6245		0.2550	
Committed Indicator (Yes)	-0.0598	***	-0.0598	***
	< 2e-16		< 2e-16	
Trade Finance Indicator (No)	0.0538	***	0.0222	***
	< 2e-16		1.09E-13	
Trade Finance Indicator (Yes)	0.1612	***	0.0934	***
	< 2e-16		0.0000	
Time To Default	-2.94E-06	***	-1.03E-06	.
	4.17E-08		0.0549	
Time To Resolution	0.0001	***	0.0001	***
	< 2e-16		< 2e-16	
Facility Asset Class Fixed Effects	Yes	(3)	Yes	(4)
Facility Type Fixed Effects	Yes	(5)	Yes	(6)
Region Fixed Effects	Yes	(7)	Yes	(8)

Seniority Fixed Effects	Yes	(9)	Yes	(10)
Nature of Default Fixed Effects	Yes		No	
90 Days Past Due	0.0291	***		
	0.0000			
Unlikely to Pay	0.0061			
	0.2231			
Bankruptcy	0.1303	***		
	< 2e-16			
Specific Provision	0.1164	***		
	< 2e-16			
Sold at Material Credit Loss	0.2434	***		
	< 2e-16			
Distressed Restructuring	0.1171	***		
	< 2e-16			
Non-Accrual	0.0712	***		
	< 2e-16			
Degrees of freedom	133193		133200	
Adjusted R-squared	12.70%		11.62%	
Residual standard error	0.3613		0.3635	

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’

(1) OLS regression with Nature of Default; (2) OLS regression without Nature of Default; (3) Facility Asset Class 3,4,6,7,10,11 are insignificant; (4) Facility Asset Class 3,4,7,10 are insignificant; (5)&(6) 13 facility types are insignificant; (7) Region America & Middle East are insignificant; (8) Region Africa & Asia are insignificant; (9) Subordinated and equity loans are insignificant; (10) Subordinated loan is insignificant.

Table 12 – LGD performance comparison from OLS regression with and without nature of default

7. Conclusions

In this paper, we explored various classification models for the prediction of nature of default of a loan. The models were built using the default loan data from GCD's LGD and EAD platform. The classification performance was assessed by multiple evaluation criteria including balanced accuracy, log loss, precision, recall and F2 score. Among all the models experimented for the prediction of nature of default, we find DT classifer gives the best performance; the most important feature in distinguishing all nature of default types from DT is the time it takes for a loan to default since its origination. Loans that take longer time to default are more likely to be bankruptcy default; whereas loans defaulted shortly after origination are more likely to be sold immediately at a discount and end up in the default class of sold at material credit loss. We also find that trade finance loans are more likely to receive a specific provision / write-off type of default from the lending bank, possibly due to the significant decrease in collateral valuation while the company has financial difficulty.

We also further studied whether nature of default is useful in LGD modeling. When comparing LGD models with and without nature of default as additional variable in the OLS regression, we found the model with nature of default performs slightly better than without; and all the dummy variables created from each nature of default class are significant except for one – "Unlikely to Pay". The nature of default class that contributes the most in predicting LGD is "Sold at Material Credit Loss".

The paper starts with explaining the importance of predicting nature of default from a business and regulatory perspective. It then compares different machine learning techniques and found DT classifier outperforms others in predicting nature of default classes. Since the paper only explored the four commonly known machine learning techniques, further studies can be done to investigate other more sophisticated methods such as Support Vector Machine, Neural Network or Stochastic Gradient Descent etc. There is also room to assess the nature of default impact on LGD with a more suitable LGD model such as censored linear regression (Tobit regression) or inflated beta regression given the concentration of observed LGD at 0% and 100%.

Another area of future improvement could be to investigate a more sophisticated solution to reduce prediction bias towards the more frequent classes. Balanced accuracy is used in this paper as a simple way to adjust the bias created by unbalanced class when assessing model performance. A more comprehensive solution could be to under sample the frequent class and over sample the infrequent class when training the model. Given unbalanced class is a very common issue in default risk modeling, further study in this area could help improve accuracy of the model.

References

A.E. Khandani, A.J. Kim and A.W. Lo (2010). Consumer credit-risk models via machine-learning algorithms. *Journal of Banking & Finance, Volumne 34, Issue 11, 2767-2787, November 2010.*

Basel Committee on Banking Supervision, QIS 3 FAQ: F. Definition of default / loss.

B. Keijsers, B. Diris, E. Kole (2018). Cyclicality in losses on bank loans. *Journal of Applied Econometrics, 2018;33: 533-332.*

D. Duffie and K. Singleton (1998). Simulating correlated defaults. *Graduate School of Business, Stanford University.*

E. I. Altman (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *The Journal of Finance, 23(4), 589-609.*

Friedman, J.H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics, 1189-1232.*

Guidelines on the application of the definition of default under Article 178 of Regulation (EU) No 575/2013 (18/01/2017)

G.H. John, P. Langley (1995).) Estimating continuous distributions in Bayesian classifiers. *Presented at the 11th Conference on Uncertainty in Artificial Intelligence (UAI), August 1995, Quebec, Canada.*

H. Degryse, V. Ioannidou, J.M. Liberti, J. Sturgess (2019). How do laws and institutions affect recovery rate for collateral? *Review of Corporate Finance Studies, v9n1, 2020.*

H.S. Kim, S.Y. Sohn (2010). Support vector machines for default prediction of SMEs based on technology credit. *European Journal of Operational Research, 201 (2010) 838-846.*

I. Guyon (1997). A scaling law for the validation-set training-set size ratio. *AT&T Bell Laboratories, Berkeley, California.*

J. Betz, S. Krüger, R. Kellner, D. Rösch (2020). Macroeconomic effects and frailties in the resolution of non-performing loans. *Journal of Banking & Finance, 112 (2020) 105212.*

J. Betz, R. Kellner, D. Rösch (2018). Systematic effects among loss given defaults and their implications on downturn estimation. Europen *Journal of Operational Research, Vol. 271, Issue 4, p1113-1144.*

K. Li, J. Niskanen, M. Kolehmainen, M. Niskanen (2016). Credit default hybrid model for SME lending. *Expert Systems With Applications 61 (2016) 343-355.*

M. Thackham, J. Ma (2019). Exposure at default without conversion factors—evidence from Global Credit Data for large corporate revolving facilities. *Journal of the Royal Statistical Society: Series A (Statistics in Society), Vol. 182, Issue 4, p1267-1286.*

O. Vasicek (1984). Credit Valuation. *White paper, Moody's KMV.*

R.A. Jarrow, D. Lando, and S.M. Turnbull (1997). A Markov model for the term structure of credit risk spreads*. Review of Financial Studies, 10:481–523, 1997*

S. Lessmann, B. Baesens, H. Seow, and L.C. Thomas (2015). Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. *European Journal of Operational Research 247 (2015) 124-136.*

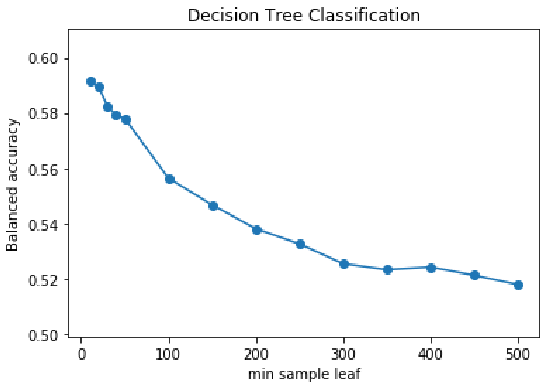
S. Kealhofer (2003a). Quantifying Credit Risk I: Default Prediction, *Financial Analysts Journal, January/February 2003.*

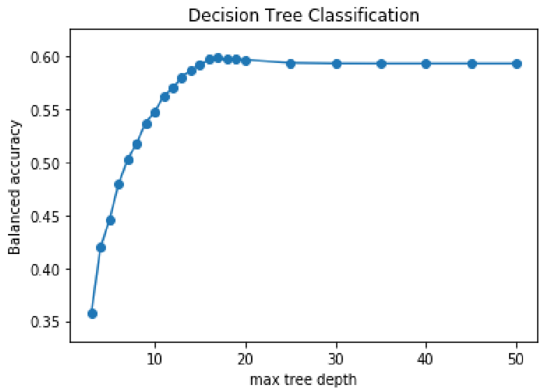
S. Kealhofer (2003b). Quantifying Credit Risk II: Debt Valuation, *Financial Analysts Journal, March/April 2003.*

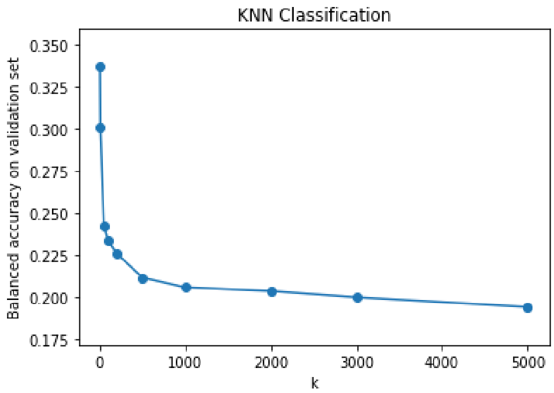
V.V. Acharya, S.T. Bharath, and A. Srinivasan. "Does industry-wide distress affect defaulted firms? Evidence from creditor recoveries." *Journal of financial economics 85.3 (2007): 787-821.*

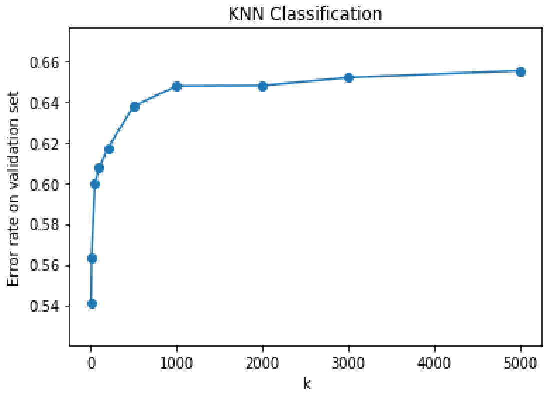
Appendix A

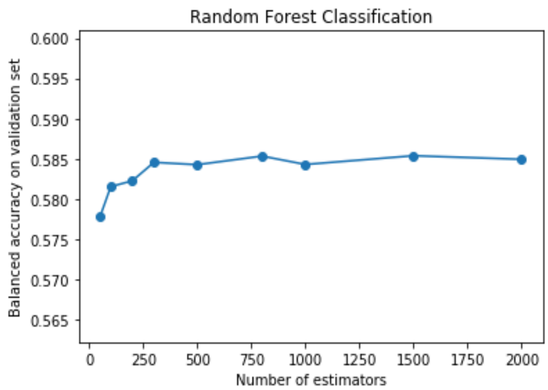
Hyper-parameters tuning with trial values for DT, RF and KNN classifier.

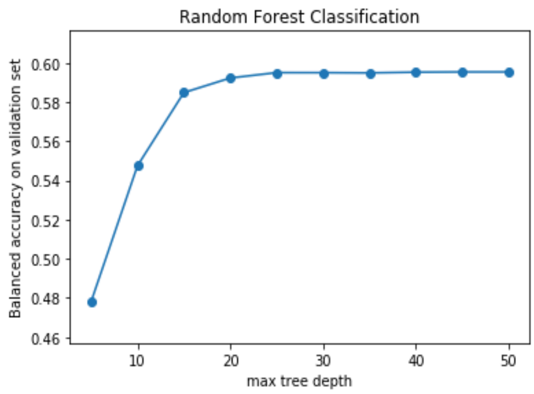


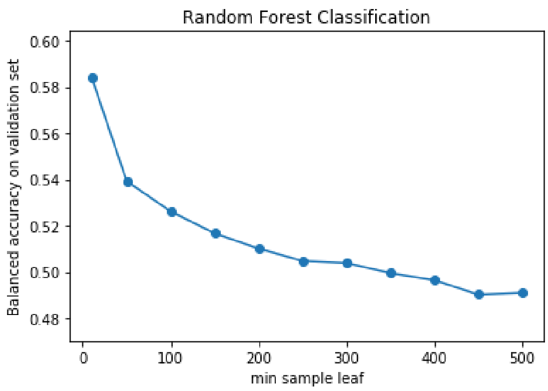












Appendix B

Full list of variables in GCD LGD & EAD platform observed default loan table.

Field name	Data Format	Data Type	Definition	Used for training model?
DA_LOAN_ID	INTEGER	Numeric	Primary key for the loan dataset	Yes
Facility_Asset_Class	INTEGER	Categorical	Asset class i.e. large corporates, SME, banks, FIs etc.	Yes
US_Segment	INTEGER	Categorical	Describes US specific segments. A field calculated by GCD which segments the data between shared segments from US members.	No (only related to US loans)
Product_Code	INTEGER	Categorical	Product Code used to group loans into a segment which can be comparable with the product segment in member bank, however it is defined.	No (only related to certain credit facilities)
Trade_Finance_Indicator	INTEGER	Categorical	Binary dummy variable indicates if facility is a trade finance facility.	Yes
Syndicated_Indicator	INTEGER	Category	Indicates if loan is part of a syndication.	No (only related to syndicated loans)
Lead_Syndicate_Indicator	INTEGER	Category	Indicates whether the Bank is the lead syndicate or Agent Bank.	
Total_Syndicated_Amount	FLOAT	Numeric	Total amount of syndicated loan Input in Syndicated_Currency, Output in EUR	
Syndicated_Currency	CHAR	Categorical	Currency denomination of the Total Syndicated Amount.	
Country_of_Jurisdiction	CHAR	Categorical	Country of the Law applicable to the facility, i.e. the law of the loan contract.	No (country is too granular, region is used instead)
Region	CHAR	Categorical	Region of the loan mapped from Country.	Yes
Guarantee_Indicator	INTEGER	Categorical	Binary variable indicates loan has underlying protection in the form of a guarantee, a CDS or some support from a Key party.	Yes
Collateral_Indicator	INTEGER	Categorical	Binary variable indicates loan has underlying protection in the form of collateral or security.	Yes
Facility_Type	INTEGER	Categorical	Type of the loan facility i.e. bridge loan, revolver, term loan etc.	Yes
Control_Goods_And_Flows	INTEGER	Categorical	Details the control procedures attached to the facility and in force at the time of the Event.	No (too many missing and only relevant to collateral)

Committed_Indicator	INTEGER	Categorical	The contractual obligation for the bank to "make the funds" when the facility is drawn by the client.	Yes
Balloon_Percentage	FLOAT	Numeric	Balloon % of Total Loan. Payment as a portion of the total loan amount.	No (only specific to balloon loan)
Seniority_Code	INTEGER	Categorical	Describe the hierarchy of the loan i.e. senior, subordinated, pari-passu etc.	Yes
Nature_Of_Default	INTEGER	Categorical	Indicates the first reason why the lender has put the Borrower in default (Basel II guidelines) at the Event Date (Default Date).	Yes
Dominion_Of_Funds	INTEGER	Categorical	Defines in a borrowing base transaction if the lender has direct access to payments from the clients to the borrower.	No (only relevant to certain types of loans)
Maturity_Date	DATE	Date	Contractual date of Termination of the facility. The date that was recorded at the time of the default.	No (duplicate with other fields used)
Default_Date	DATE	Date	Date of default of the loan.	No (duplicate with other fields used)
Resolution_Date	DATE	Date	Date of resolution of the loan.	No (duplicate with other fields used)
TIME_TO_DEFAULT	FLOAT	Numeric	Number of Days between origination and default date of the loan.	Yes
TIME_TO_RESOLUTION	FLOAT	Numeric	Number of Days between default and resolution date of the loan.	Yes
TIME_IN_DEFAULT_UNRESOLVED	FLOAT	Numeric	Calculated Number of Days the after the default date to the Reference Date, where the Reference Date is either 30th June for H1 submission year and 31st December for H2 submission year.	Yes
Conversion_Rate, Conversion_Rate_TOTAL_SYNDIC_AMO	FLOAT	Numeric	Currency conversion related fields.	No (duplicate with other fields used)
PV_NET_CUM_AMO fields	FLOAT	Numeric	Cumulative present value of the recovered amount	No (duplicate with other fields used)
Default_Amount fields	FLOAT	Numeric	Fields related to the default amount from the loan.	No (duplicate with other fields used)
NOM_DEFAULT_AMOUNT fields	FLOAT	Numeric	Fields related to the nominal default amount from the loan.	No (duplicate with other fields used)
Recovery rate related fields	FLOAT	Numeric	Fields related to the recovery rate.	No (duplicate with LGD fields)

LGD related fields	FLOAT	Numeric	Fields related to the loss give default.	Yes
Discount_Rate	FLOAT	Numeric	The discount rate used to discount recovered amount back to default date.	No (duplicate with other fields used)

Appendix C

Asset class code and description.

Facility_Asset_Class	Description
1	Small/Medium Enterprises (SME)
2	Large Corporates
3	Banks & Financial Companies
4	Ship Finance
5	Aircraft Finance
6	Real Estate Finance
7	Project Finance
8	Commodities Finance
9	Sovereigns, Central Banks
10	Public Services
11	Private Banking

Appendix D

Seniority code and description.

Seniority_Code	Description	Definition
100	Super Senior	Lender has made agreement with other lenders to “promote” this loan
110	Pari-Passu	By definition a loan is always Pari-Passu unless the lender has made agreements with other lenders to “promote” or “demote” itself
200	Subordinated or Junior	Lender has made agreement with other lenders to “demote” this loan
300	Equity	Only to be used for facilities which are Equity: Facility type 620 (Equity) or 630 (Debt/Equity Hybrids)
-1	Unknown	

Appendix E

Region code and description.

Region_Code	Description
100	Africa
200	America
300	Asia
400	Europe
500	Middle East
600	Oceania

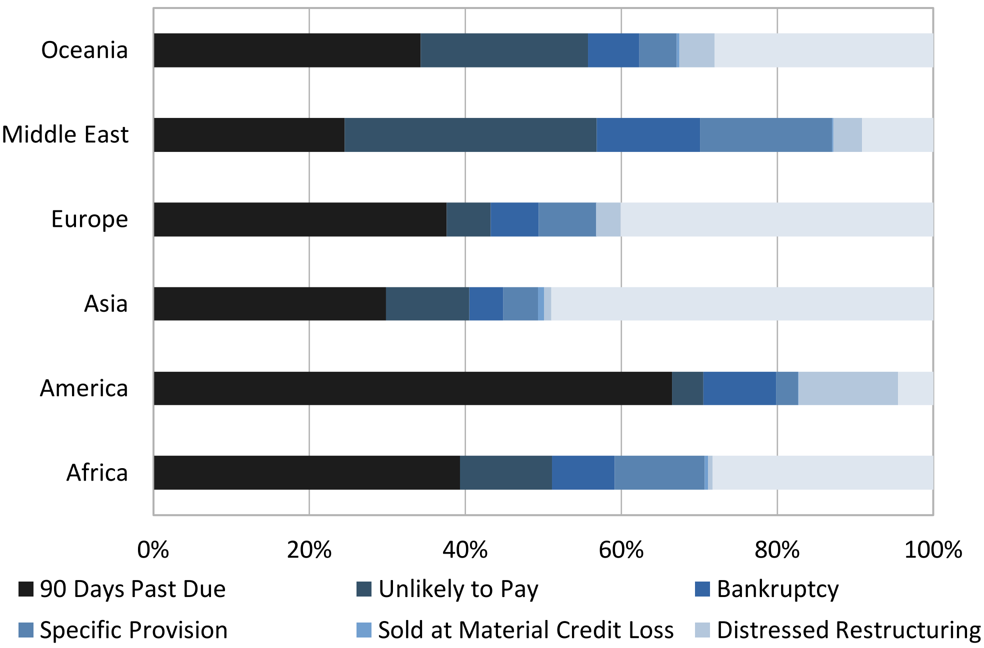
Appendix F

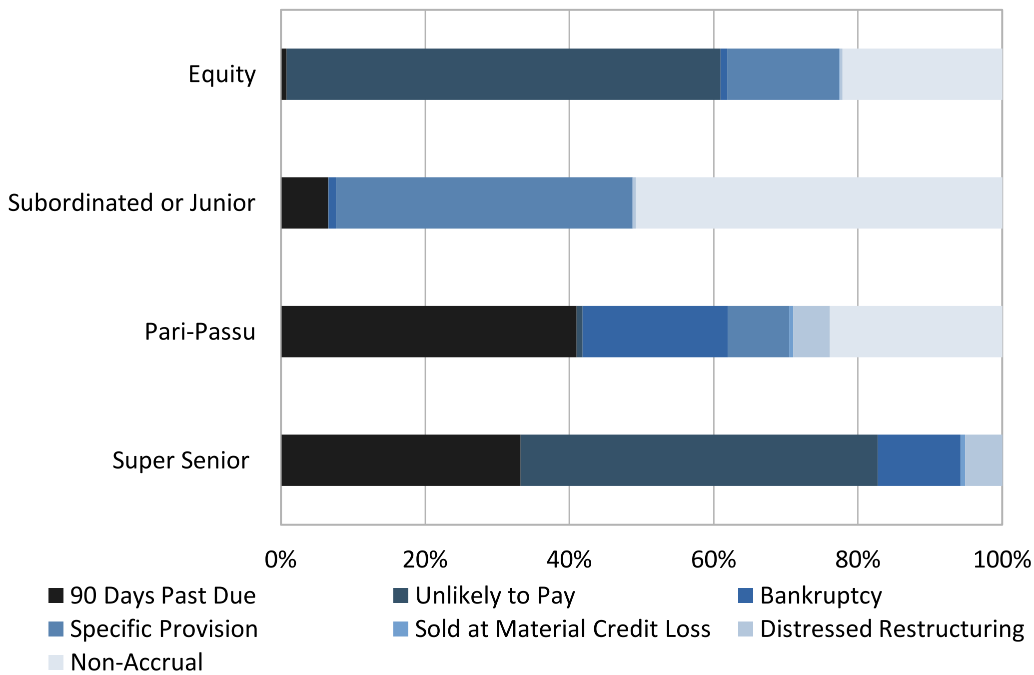
Facility type code and description.

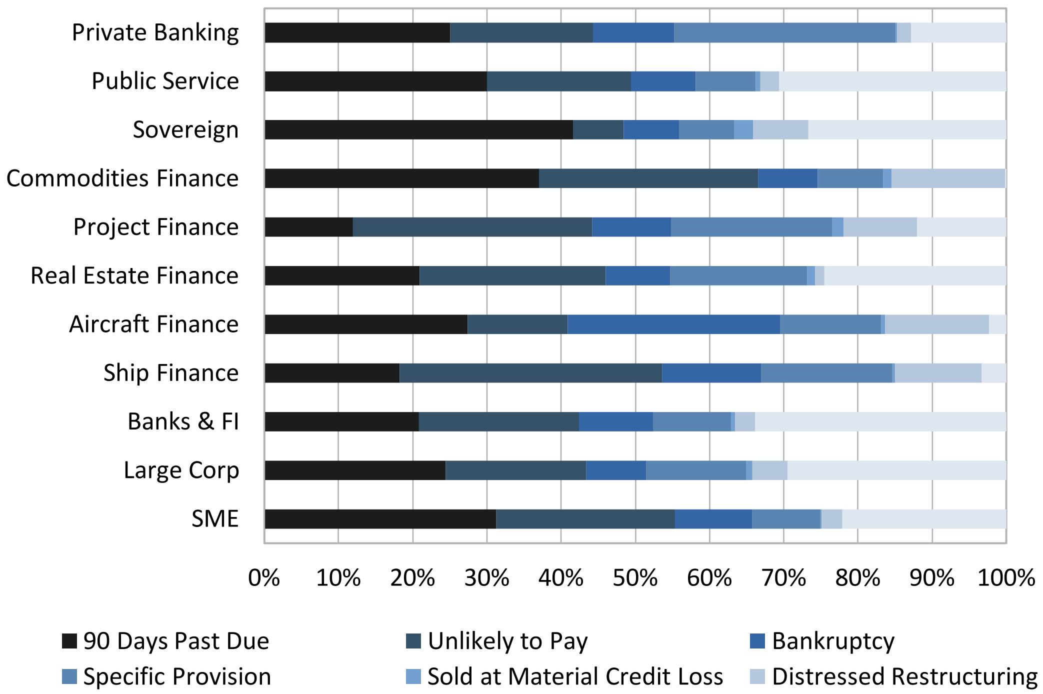
Facility_Type	Description	Facility_Type	Description
100	Bridge Loan	811	Trade Related Payment Guarantee
200	Revolver/Line > 1 year	812	Trade related Documentary Credit (L/C) at Usance
210	Revolver/Line < 1 year	813	Confirmed Export L/C (S & U)
250	Overdraft	820	Contract Bonds
260	Margin Loan	830	Payment Guarantee and Stand By LC's
270	Money Market Lines	840	Repurchase Agreement (REPO)
300	Term Loan	850	Securities Lending
400	Demand Loan	860	Settlement Limit Facility
500	Commercial Paper Backup	865	Vostro / Nostro
620	Equity	870	Intraday Facility
630	Debt/Equity Hybrid	883	FOREX Derivative
700	Capital Lease	884	Interest Rate Derivative
710	Operating Lease	885	Credit Default Swap
800	Receivables Financing	886	Equity Derivative
801	ECA Export Finance	887	Commodities Derivative
802	Transactional Trade Finance	888	Derivative of Mixed or Unknown Type
803	Bill accepted by Bank	890	Bonds in Banking Book
804	Prepayment Finance	891	US Municipalities - General Limited Tax Obligation
805	Pre-export Finance	892	US Municipalities - General Obligation subject to Appropriation
806	Structured Inventory Finance	893	US Municipalities - General Unlimited Tax Obligation
807	Bid or Performance Bond	894	US Municipalities - State General Obligation
8091	Borrowing Base Finance - Revolving Loan	895	US Municipalities - Statutory Pledge
8092	Borrowing Base Finance - Term Loan	896	US Municipalitites - Other Bond Obligation
810	Trade Related Documentary Credit (L/C) at Sight	900	Aggregate Exposure

Appendix G

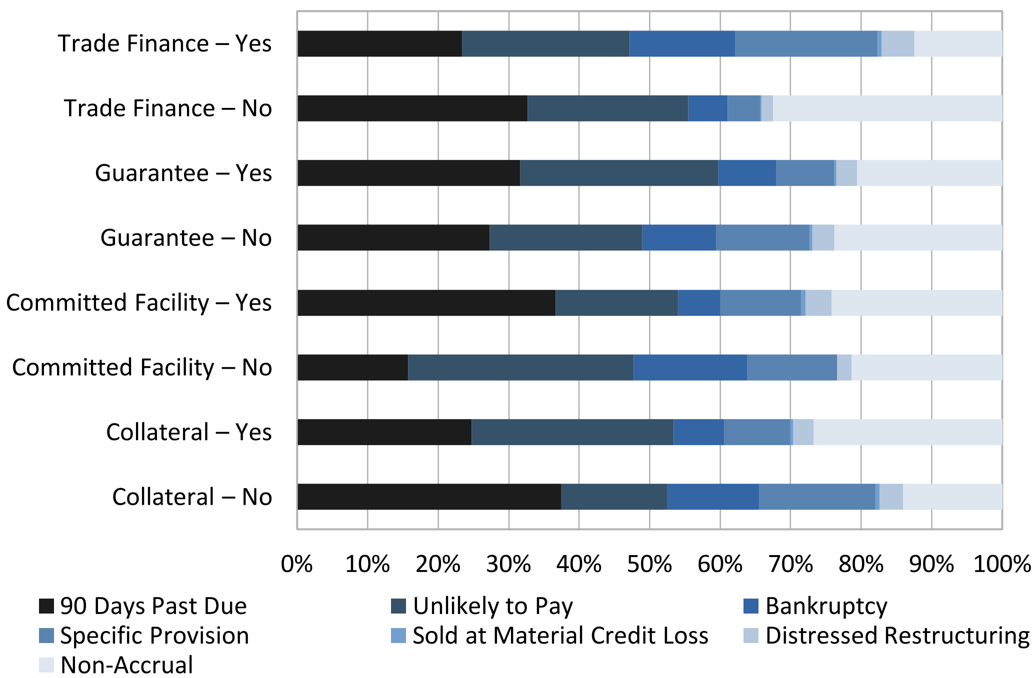
Nature of default in percentage by region, seniority, asset class, and other dummy variables.







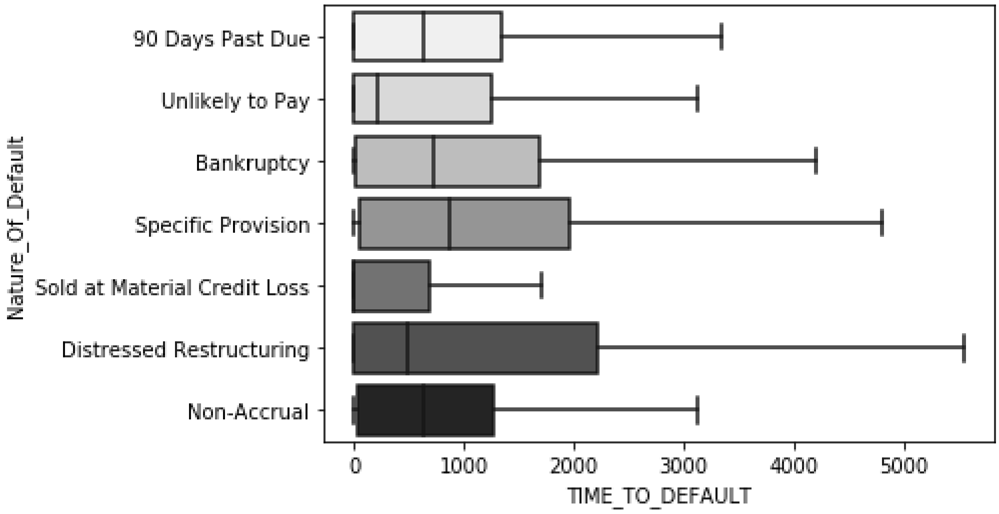
Private Banking
Public Service
Sovereign
Commodities Finance
Project Finance
Real Estate Finance
Aircraft Finance
Ship Finance
Banks & FI
Large Corp
SME
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%
90 Days Past Due
Unlikely to Pay
Bankruptcy
Specific Provision
Sold at Material Credit Loss
Distressed Restructuring

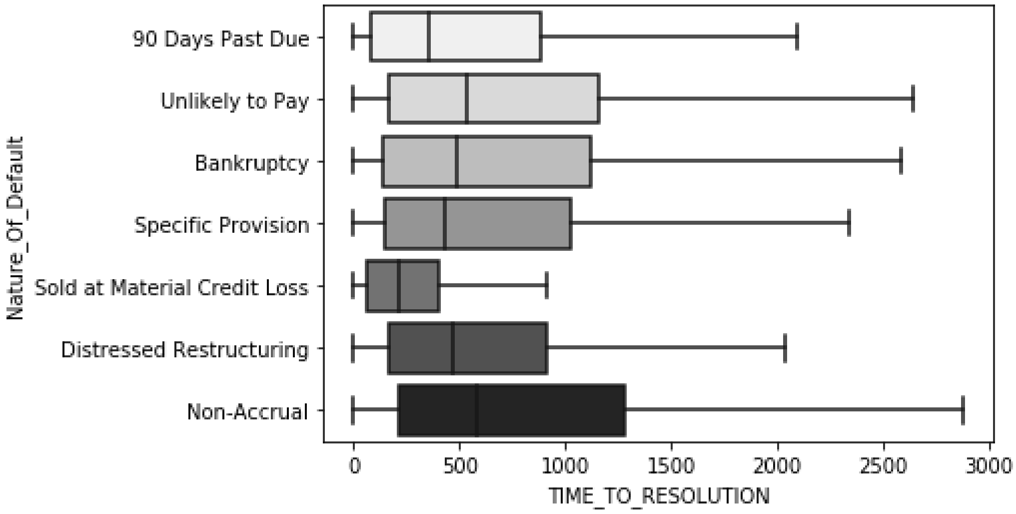


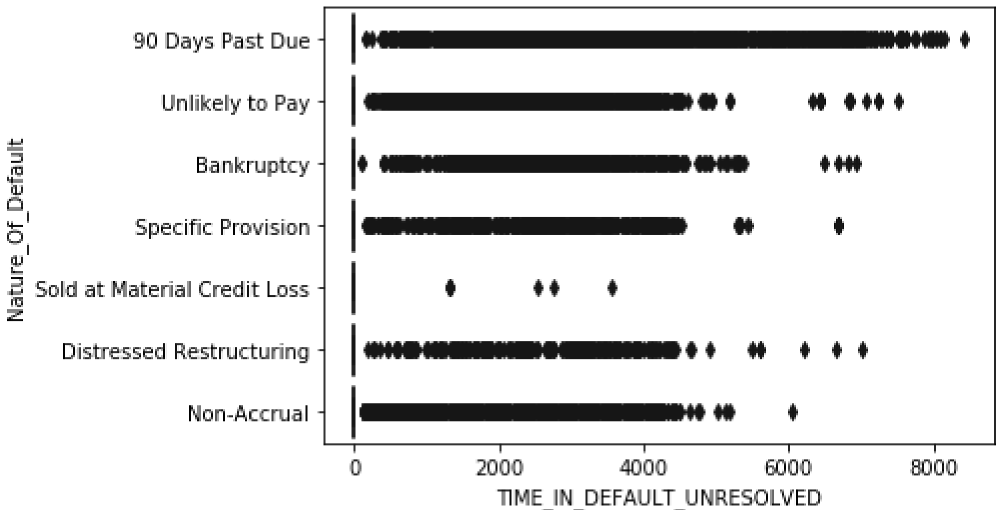
Trade Finance – Yes
Trade Finance – No
Guarantee – Yes
Guarantee – No
Committed Facility – Yes
Committed Facility – No
Collateral – Yes
Collateral – No
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%
90 Days Past Due
Unlikely to Pay
Bankruptcy
Specific Provision
Sold at Material Credit Loss
Distressed Restructuring
Non-Accrual

Appendix H

Box plot for numeric features by nature of default types







Appendix I

MNL model regression full summary table

Nature of Default	Feature Number	Feature	estimate	std.error	p.value
210	1	(Intercept)	0.132553	0.005698	0
210	2	TIME_TO_DEFAULT	0.999938	6.66E-06	1.92E-20
210	3	TIME_TO_RESOLUTION	1.000159	1.18E-05	8.49E-42
210	4	TIME_IN_DEFAULT_UNRESOLVED	0.999831	1.38E-05	1.68E-34
210	5	Facility_Asset_Class2	2.887779	0.00866	0
210	6	Facility_Asset_Class3	2.675488	0.000338	0
210	7	Facility_Asset_Class4	3.077898	0.000217	0
210	8	Facility_Asset_Class5	1.008429	7.44E-05	0
210	9	Facility_Asset_Class6	0.784652	0.006193	0
210	10	Facility_Asset_Class7	4.153567	0.000237	0
210	11	Facility_Asset_Class8	0.795825	7.11E-05	0
210	12	Facility_Asset_Class9	1.722662	1.61E-05	0
210	13	Facility_Asset_Class10	0.880658	2.22E-05	0
210	14	Facility_Asset_Class11	0.849108	0.001468	0
210	15	Trade_Finance_Indicator1	1.760286	0.010836	0
210	16	Guarantee_Indicator1	1.180731	0.014094	4.52E-32
210	17	Collateral_Indicator1	2.135935	0.011309	0
210	18	Facility_Type100	1.8288	7.57E-05	0
210	19	Facility_Type200	1.096842	0.002836	4.34E-233
210	20	Facility_Type210	0.834486	0.001161	0
210	21	Facility_Type240	0.956554	0.000335	0
210	22	Facility_Type250	2.071351	0.007599	0
210	23	Facility_Type260	0.664753	5.80E-06	0
210	24	Facility_Type270	1.186376	3.11E-06	0
210	25	Facility_Type300	1.297976	0.009457	2.04E-167
210	26	Facility_Type400	1.541725	0.000963	0
210	27	Facility_Type500	0.110951	2.56E-06	0
210	28	Facility_Type620	4.25793	3.25E-05	0
210	29	Facility_Type630	2.856379	7.64E-05	0
210	30	Facility_Type700	0.357853	0.000776	0
210	31	Facility_Type710	0.966506	0.00039	0
210	32	Facility_Type800	18.99744	0.000472	0
210	33	Facility_Type801	2.980907	0.00014	0
210	34	Facility_Type802	1.578982	0.000173	0
210	35	Facility_Type803	0.208157	9.90E-06	0
210	36	Facility_Type804	0.281239	3.40E-07	0
210	37	Facility_Type805	1.882357	3.33E-05	0
210	38	Facility_Type806	0.113249	5.44E-06	0
210	39	Facility_Type807	23.5866	1.33E-05	0

210	40	Facility_Type809	0.084205	7.45E-07	0
210	41	Facility_Type810	1.523972	0.000158	0
210	42	Facility_Type811	2.168886	4.69E-05	0
210	43	Facility_Type812	2.824789	8.34E-06	0
210	44	Facility_Type813	27.01543	3.41E-06	0
210	45	Facility_Type820	2.216574	0.000206	0
210	46	Facility_Type830	2.189885	0.001314	0
210	47	Facility_Type840	1.333463	8.41E-07	0
210	48	Facility_Type850	0.154571	1.97E-06	0
210	49	Facility_Type860	6.596203	3.03E-05	0
210	50	Facility_Type865	263.7976	6.27E-06	0
210	51	Facility_Type870	0.114995	1.03E-05	0
210	52	Facility_Type880	7.175303	4.10E-05	0
210	53	Facility_Type881	0.546149	1.01E-07	0
210	54	Facility_Type883	4.975304	4.36E-05	0
210	55	Facility_Type884	11.61092	7.04E-05	0
210	56	Facility_Type885	1.473786	2.53E-06	0
210	57	Facility_Type886	3.290475	5.94E-06	0
210	58	Facility_Type887	0.31883	5.19E-06	0
210	59	Facility_Type888	3.506603	0.000184	0
210	60	Facility_Type890	21.31634	3.29E-05	0
210	61	Facility_Type900	0.428719	1.89E-06	0
210	62	Facility_Type8091	0.077996	1.23E-05	0
210	63	Facility_Type8092	0.000856	7.21E-08	0
210	64	Committed_Indicator1	0.174853	0.011387	0
210	65	Seniority_Code100	3.727528	0.009059	0
210	66	Seniority_Code110	1.096956	0.008303	7.55E-29
210	67	Seniority_Code200	0.512602	0.000453	0
210	68	Seniority_Code300	3.198289	4.07E-05	0
210	69	Region100	0.040892	0.000398	0
210	70	Region200	1.289137	0.008343	1.53E-203
210	71	Region300	0.907484	0.000689	0
210	72	Region400	4.349452	0.008438	0
210	73	Region500	1.081892	5.51E-05	0
210	74	Region600	16.04567	0.000594	0
220	1	(Intercept)	0.054272	0.005284	0
220	2	TIME_TO_DEFAULT	1.000078	5.46E-06	7.04E-46
220	3	TIME_TO_RESOLUTION	1.000295	1.43E-05	2.07E-94
220	4	TIME_IN_DEFAULT_UNRESOLVED	1.000039	1.40E-05	0.005495
220	5	Facility_Asset_Class2	1.706015	0.004953	0
220	6	Facility_Asset_Class3	1.438445	0.000198	0
220	7	Facility_Asset_Class4	1.287645	0.000114	0
220	8	Facility_Asset_Class5	5.025211	0.000148	0

220	9	Facility_Asset_Class6	0.341706	0.003863	0
220	10	Facility_Asset_Class7	2.081579	0.000157	0
220	11	Facility_Asset_Class8	0.285735	2.78E-05	0
220	12	Facility_Asset_Class9	1.176397	2.22E-05	0
220	13	Facility_Asset_Class10	1.144153	1.54E-05	0
220	14	Facility_Asset_Class11	0.392005	0.001768	0
220	15	Trade_Finance_Indicator1	4.393902	0.010493	0
220	16	Guarantee_Indicator1	1.221922	0.006309	1.81E-221
220	17	Collateral_Indicator1	1.017545	0.010614	0.101289
220	18	Facility_Type100	2.603427	7.38E-05	0
220	19	Facility_Type200	1.423636	0.003532	0
220	20	Facility_Type210	1.234399	0.001415	0
220	21	Facility_Type240	0.161901	8.21E-05	0
220	22	Facility_Type250	1.989651	0.008921	0
220	23	Facility_Type260	0.226445	1.17E-06	0
220	24	Facility_Type270	0.238339	1.18E-07	0
220	25	Facility_Type300	0.959245	0.009799	2.17E-05
220	26	Facility_Type400	0.266355	0.000338	0
220	27	Facility_Type500	1.301154	6.79E-06	0
220	28	Facility_Type620	2.450714	2.15E-05	0
220	29	Facility_Type630	0.725017	1.50E-05	0
220	30	Facility_Type700	0.269763	0.000176	0
220	31	Facility_Type710	0.832085	0.000288	0
220	32	Facility_Type800	2.013056	0.000294	0
220	33	Facility_Type801	1.906682	0.000117	0
220	34	Facility_Type802	1.059947	0.000133	0
220	35	Facility_Type803	0.041186	9.79E-07	0
220	36	Facility_Type804	16.20296	3.94E-06	0
220	37	Facility_Type805	1.707593	2.29E-05	0
220	38	Facility_Type806	0.196641	3.26E-06	0
220	39	Facility_Type807	0.178386	1.24E-07	0
220	40	Facility_Type809	9.890026	2.61E-05	0
220	41	Facility_Type810	1.57217	0.000141	0
220	42	Facility_Type811	1.481551	3.32E-05	0
220	43	Facility_Type812	8.74415	1.51E-05	0
220	44	Facility_Type813	0.438746	1.98E-07	0
220	45	Facility_Type820	11.74576	0.000546	0
220	46	Facility_Type830	3.726118	0.001902	0
220	47	Facility_Type840	23.45529	1.63E-05	0
220	48	Facility_Type850	1.185612	1.03E-05	0
220	49	Facility_Type860	0.3103	2.11E-06	0
220	50	Facility_Type865	4.254608	3.73E-07	0
220	51	Facility_Type870	2.310404	7.55E-05	0

220	52	Facility_Type880	2.393271	2.46E-05	0
220	53	Facility_Type881	1.704041	4.48E-07	0
220	54	Facility_Type883	9.11514	4.65E-05	0
220	55	Facility_Type884	14.25578	5.51E-05	0
220	56	Facility_Type885	14.39197	5.16E-06	0
220	57	Facility_Type886	25.75513	9.81E-06	0
220	58	Facility_Type887	0.982911	5.71E-06	0
220	59	Facility_Type888	0.501382	3.19E-05	0
220	60	Facility_Type890	18.40218	2.70E-05	0
220	61	Facility_Type900	248.1075	2.48E-05	0
220	62	Facility_Type8091	0.450349	3.09E-05	0
220	63	Facility_Type8092	0.014152	3.66E-07	0
220	64	Committed_Indicator1	0.17712	0.011018	0
220	65	Seniority_Code100	3.312123	0.009216	0
220	66	Seniority_Code110	1.875228	0.00919	0
220	67	Seniority_Code200	1.597409	0.000636	0
220	68	Seniority_Code300	4.505127	2.35E-05	0
220	69	Region100	0.095547	0.001571	0
220	70	Region200	2.044939	0.007518	0
220	71	Region300	1.734574	0.001017	0
220	72	Region400	3.118466	0.007954	0
220	73	Region500	0.882297	3.77E-05	0
220	74	Region600	6.234551	0.000219	0
230	1	(Intercept)	0.02899	0.004547	0
230	2	TIME_TO_DEFAULT	1.000068	5.36E-06	2.02E-36
230	3	TIME_TO_RESOLUTION	1.000132	1.41E-05	8.01E-21
230	4	TIME_IN_DEFAULT_UNRESOLVED	0.999595	2.01E-05	1.43E-90
230	5	Facility_Asset_Class2	3.104173	0.007131	0
230	6	Facility_Asset_Class3	1.944989	0.000251	0
230	7	Facility_Asset_Class4	1.300182	9.64E-05	0
230	8	Facility_Asset_Class5	1.279558	6.99E-05	0
230	9	Facility_Asset_Class6	0.835483	0.006775	4.25E-155
230	10	Facility_Asset_Class7	2.755446	0.000158	0
230	11	Facility_Asset_Class8	0.417051	2.97E-05	0
230	12	Facility_Asset_Class9	1.253262	1.87E-05	0
230	13	Facility_Asset_Class10	1.268475	1.37E-05	0
230	14	Facility_Asset_Class11	1.395763	0.003139	0
230	15	Trade_Finance_Indicator1	6.165306	0.009948	0
230	16	Guarantee_Indicator1	0.740466	0.004852	0
230	17	Collateral_Indicator1	1.005888	0.012981	0.651093
230	18	Facility_Type100	6.338025	0.000144	0
230	19	Facility_Type200	1.225549	0.004339	0
230	20	Facility_Type210	0.981589	0.001172	1.37E-56

230	21	Facility_Type240	0.017615	1.55E-05	0
230	22	Facility_Type250	1.456778	0.008502	0
230	23	Facility_Type260	0.011445	4.21E-07	0
230	24	Facility_Type270	32.02127	5.82E-06	0
230	25	Facility_Type300	1.017126	0.010738	0.113813
230	26	Facility_Type400	0.80498	0.000457	0
230	27	Facility_Type500	0.780091	5.84E-06	0
230	28	Facility_Type620	1.567499	2.38E-05	0
230	29	Facility_Type630	0.552461	1.76E-05	0
230	30	Facility_Type700	0.418926	0.000356	0
230	31	Facility_Type710	1.120556	0.000515	0
230	32	Facility_Type800	1.586664	0.000193	0
230	33	Facility_Type801	1.216947	9.95E-05	0
230	34	Facility_Type802	0.842133	0.000109	0
230	35	Facility_Type803	0.390671	4.50E-06	0
230	36	Facility_Type804	0.448518	5.00E-07	0
230	37	Facility_Type805	0.053495	8.48E-07	0
230	38	Facility_Type806	0.237568	4.82E-06	0
230	39	Facility_Type807	0.026393	7.05E-08	0
230	40	Facility_Type809	39.87176	0.000189	0
230	41	Facility_Type810	0.51163	0.000107	0
230	42	Facility_Type811	0.98502	4.52E-05	0
230	43	Facility_Type812	3.584521	9.82E-06	0
230	44	Facility_Type813	2.432833	1.31E-06	0
230	45	Facility_Type820	2.339499	0.000346	0
230	46	Facility_Type830	1.17053	0.000988	0
230	47	Facility_Type840	68.8039	1.82E-05	0
230	48	Facility_Type850	0.538785	1.24E-05	0
230	49	Facility_Type860	3.243862	1.74E-05	0
230	50	Facility_Type865	0.519308	5.25E-08	0
230	51	Facility_Type870	21.24033	0.000241	0
230	52	Facility_Type880	0.776867	2.01E-05	0
230	53	Facility_Type881	0.02623	2.48E-08	0
230	54	Facility_Type883	9.887587	3.98E-05	0
230	55	Facility_Type884	12.73472	3.77E-05	0
230	56	Facility_Type885	0.141368	3.62E-07	0
230	57	Facility_Type886	1.242729	5.20E-07	0
230	58	Facility_Type887	0.307472	1.17E-06	0
230	59	Facility_Type888	0.625741	3.47E-05	0
230	60	Facility_Type890	8.123749	1.26E-05	0
230	61	Facility_Type900	10.11682	1.77E-05	0
230	62	Facility_Type8091	1.236963	9.46E-05	0
230	63	Facility_Type8092	0.734109	2.54E-05	0

230	64	Committed_Indicator1	0.570564	0.009509	0
230	65	Seniority_Code100	1.927368	0.009881	0
230	66	Seniority_Code110	1.398986	0.009774	1.39E-258
230	67	Seniority_Code200	0.498954	0.00044	0
230	68	Seniority_Code300	6.173387	2.76E-05	0
230	69	Region100	0.101307	0.000476	0
230	70	Region200	1.748258	0.00841	0
230	71	Region300	2.329808	0.001508	0
230	72	Region400	5.906964	0.009796	0
230	73	Region500	0.703137	3.41E-05	0
230	74	Region600	2.458726	0.000121	0
240	1	(Intercept)	0.001103	0.000904	0
240	2	TIME_TO_DEFAULT	1.00003	1.76E-05	0.090058
240	3	TIME_TO_RESOLUTION	1.00002	3.53E-05	0.563706
240	4	TIME_IN_DEFAULT_UNRESOLVED	0.999453	8.69E-05	3.07E-10
240	5	Facility_Asset_Class2	3.211205	0.000375	0
240	6	Facility_Asset_Class3	1.332383	1.31E-05	0
240	7	Facility_Asset_Class4	5.232685	2.21E-05	0
240	8	Facility_Asset_Class5	1.613381	4.21E-06	0
240	9	Facility_Asset_Class6	2.304797	0.00033	0
240	10	Facility_Asset_Class7	5.79413	2.51E-05	0
240	11	Facility_Asset_Class8	3.87437	5.64E-06	0
240	12	Facility_Asset_Class9	4.215019	2.18E-06	0
240	13	Facility_Asset_Class10	0.841942	4.82E-07	0
240	14	Facility_Asset_Class11	1.435976	6.37E-05	0
240	15	Trade_Finance_Indicator1	1.538418	0.000684	0
240	16	Guarantee_Indicator1	0.552733	0.000205	0
240	17	Collateral_Indicator1	1.562716	0.00077	0
240	18	Facility_Type100	6.079538	3.62E-06	0
240	19	Facility_Type200	1.587117	0.000143	0
240	20	Facility_Type210	1.567965	4.75E-05	0
240	21	Facility_Type240	0.000613	1.02E-08	0
240	22	Facility_Type250	4.069806	0.000152	0
240	23	Facility_Type260	1.377117	2.28E-07	0
240	24	Facility_Type270	0.980629	8.22E-09	0
240	25	Facility_Type300	2.730176	0.000639	0
240	26	Facility_Type400	0.470186	5.77E-06	0
240	27	Facility_Type500	0.832191	6.43E-08	0
240	28	Facility_Type620	10.67936	4.29E-06	0
240	29	Facility_Type630	0.461348	7.36E-07	0
240	30	Facility_Type700	0.566579	1.84E-05	0
240	31	Facility_Type710	2.374239	2.06E-05	0
240	32	Facility_Type800	6.154025	7.84E-06	0

240	33	Facility_Type801	2.091114	1.48E-06	0
240	34	Facility_Type802	0.610139	1.36E-06	0
240	35	Facility_Type803	3.796072	9.09E-07	0
240	36	Facility_Type804	0.973266	4.56E-08	0
240	37	Facility_Type805	0.656044	2.78E-07	0
240	38	Facility_Type806	0.514337	1.08E-07	0
240	39	Facility_Type807	1.003607	1.92E-08	0
240	40	Facility_Type809	16.38997	5.37E-06	0
240	41	Facility_Type810	19.05684	3.44E-05	0
240	42	Facility_Type811	0.806376	2.14E-07	0
240	43	Facility_Type812	1.092477	7.36E-08	0
240	44	Facility_Type813	1.029897	9.42E-09	0
240	45	Facility_Type820	5.209116	6.02E-06	0
240	46	Facility_Type830	3.683693	2.68E-05	0
240	47	Facility_Type840	0.724043	3.86E-08	0
240	48	Facility_Type850	0.665151	2.84E-07	0
240	49	Facility_Type860	6.745921	1.36E-06	0
240	50	Facility_Type865	0.893992	2.05E-09	0
240	51	Facility_Type870	8.262365	5.97E-06	0
240	52	Facility_Type880	3.42282	7.70E-07	0
240	53	Facility_Type881	7.498578	4.89E-08	0
240	54	Facility_Type883	1.563738	2.46E-07	0
240	55	Facility_Type884	14.87012	3.06E-06	0
240	56	Facility_Type885	0.933916	3.30E-08	0
240	57	Facility_Type886	1.110331	3.08E-08	0
240	58	Facility_Type887	0.907406	1.23E-07	0
240	59	Facility_Type888	3.019348	3.86E-06	0
240	60	Facility_Type890	0.96569	1.42E-08	0
240	61	Facility_Type900	18.82441	5.49E-07	0
240	62	Facility_Type8091	1.832324	1.09E-05	0
240	63	Facility_Type8092	0.258147	5.76E-07	0
240	64	Committed_Indicator1	0.955172	0.000786	0
240	65	Seniority_Code100	4.339308	0.000523	0
240	66	Seniority_Code110	3.86856	0.000755	0
240	67	Seniority_Code200	2.094139	1.03E-05	0
240	68	Seniority_Code300	11.6917	4.28E-06	0
240	69	Region100	0.144093	4.57E-06	0
240	70	Region200	3.290717	0.000643	0
240	71	Region300	0.665682	2.38E-05	0
240	72	Region400	1.146615	0.000555	0
240	73	Region500	1.514717	2.78E-06	0
240	74	Region600	7.312851	3.70E-05	0
250	1	(Intercept)	0.000816	0.00558	0

250	2	TIME_TO_DEFAULT	1.000095	6.53E-06	1.15E-47
250	3	TIME_TO_RESOLUTION	1.000122	2.08E-05	4.85E-09
250	4	TIME_IN_DEFAULT_UNRESOLVED	0.999948	1.92E-05	0.006314
250	5	Facility_Asset_Class2	4.170584	0.002148	0
250	6	Facility_Asset_Class3	1.650104	8.79E-05	0
250	7	Facility_Asset_Class4	2.008917	0.000151	0
250	8	Facility_Asset_Class5	2.849651	4.68E-05	0
250	9	Facility_Asset_Class6	0.325268	0.00057	0
250	10	Facility_Asset_Class7	2.285639	6.85E-05	0
250	11	Facility_Asset_Class8	3.32356	7.06E-05	0
250	12	Facility_Asset_Class9	2.883119	7.07E-06	0
250	13	Facility_Asset_Class10	1.725006	6.37E-06	0
250	14	Facility_Asset_Class11	0.522891	0.000345	0
250	15	Trade_Finance_Indicator1	4.659406	0.003909	0
250	16	Guarantee_Indicator1	0.96615	0.002349	1.16E-48
250	17	Collateral_Indicator1	1.723645	0.005136	0
250	18	Facility_Type100	4.07023	2.64E-05	0
250	19	Facility_Type200	1.155714	0.000485	0
250	20	Facility_Type210	0.473325	0.000132	0
250	21	Facility_Type240	0.101507	2.77E-05	0
250	22	Facility_Type250	1.479877	0.001778	0
250	23	Facility_Type260	0.393082	2.03E-06	0
250	24	Facility_Type270	91.50224	4.67E-06	0
250	25	Facility_Type300	1.457044	0.003965	0
250	26	Facility_Type400	0.568986	8.20E-05	0
250	27	Facility_Type500	0.21018	6.53E-07	0
250	28	Facility_Type620	8.091593	8.27E-06	0
250	29	Facility_Type630	0.000354	1.33E-08	0
250	30	Facility_Type700	0.215289	6.60E-05	0
250	31	Facility_Type710	1.250364	9.65E-05	0
250	32	Facility_Type800	1.763032	6.96E-05	0
250	33	Facility_Type801	3.37622	5.35E-05	0
250	34	Facility_Type802	1.289307	4.95E-05	0
250	35	Facility_Type803	0.146231	2.21E-06	0
250	36	Facility_Type804	0.700125	4.46E-07	0
250	37	Facility_Type805	5.066095	1.37E-05	0
250	38	Facility_Type806	0.813348	3.71E-06	0
250	39	Facility_Type807	30.6552	1.25E-05	0
250	40	Facility_Type809	6.759194	7.98E-06	0
250	41	Facility_Type810	0.784266	5.68E-05	0
250	42	Facility_Type811	0.277348	2.36E-06	0
250	43	Facility_Type812	0.131941	2.52E-07	0
250	44	Facility_Type813	0.756211	5.59E-08	0

250	45	Facility_Type820	2.153588	8.56E-05	0
250	46	Facility_Type830	1.491704	0.000366	0
250	47	Facility_Type840	0.435189	1.07E-07	0
250	48	Facility_Type850	0.170624	8.93E-07	0
250	49	Facility_Type860	1.682676	3.27E-06	0
250	50	Facility_Type865	113.1733	6.01E-06	0
250	51	Facility_Type870	0.839112	2.07E-05	0
250	52	Facility_Type880	1.392249	9.31E-06	0
250	53	Facility_Type881	0.518391	8.08E-08	0
250	54	Facility_Type883	0.989168	1.81E-06	0
250	55	Facility_Type884	18.43968	2.05E-05	0
250	56	Facility_Type885	3.279368	1.59E-06	0
250	57	Facility_Type886	55.56001	6.90E-06	0
250	58	Facility_Type887	4.093436	3.89E-06	0
250	59	Facility_Type888	3.56775	6.01E-05	0
250	60	Facility_Type890	43.20465	1.88E-05	0
250	61	Facility_Type900	7.561177	3.00E-06	0
250	62	Facility_Type8091	1.706395	5.50E-05	0
250	63	Facility_Type8092	0.057004	7.20E-07	0
250	64	Committed_Indicator1	0.787372	0.005275	0
250	65	Seniority_Code100	4.239112	0.007289	0
250	66	Seniority_Code110	2.542902	0.007445	0
250	67	Seniority_Code200	1.549595	0.000183	0
250	68	Seniority_Code300	3.923919	8.25E-06	0
250	69	Region100	6.79321	0.001727	0
250	70	Region200	6.380731	0.001732	0
250	71	Region300	11.83894	0.000399	0
250	72	Region400	18.31432	0.004928	0
250	73	Region500	3.874998	1.16E-05	0
250	74	Region600	58.42836	0.000184	0
260	1	(Intercept)	1.761387	0.005615	0
260	2	TIME_TO_DEFAULT	0.999873	8.34E-06	1.69E-52
260	3	TIME_TO_RESOLUTION	1.000418	1.11E-05	2.43E-307
260	4	TIME_IN_DEFAULT_UNRESOLVED	0.999842	1.48E-05	8.66E-27
260	5	Facility_Asset_Class2	1.470226	0.009063	0
260	6	Facility_Asset_Class3	2.134662	0.000308	0
260	7	Facility_Asset_Class4	1.097468	6.64E-05	0
260	8	Facility_Asset_Class5	0.140161	7.88E-06	0
260	9	Facility_Asset_Class6	1.911964	0.007045	0
260	10	Facility_Asset_Class7	1.628299	9.86E-05	0
260	11	Facility_Asset_Class8	0.03808	2.97E-06	0
260	12	Facility_Asset_Class9	1.356201	2.39E-05	0
260	13	Facility_Asset_Class10	1.180061	2.35E-05	0

260	14	Facility_Asset_Class11	1.021085	0.001244	3.56E-63
260	15	Trade_Finance_Indicator1	0.488398	0.011573	0
260	16	Guarantee_Indicator1	0.899427	0.010761	6.83E-23
260	17	Collateral_Indicator1	1.907669	0.011086	0
260	18	Facility_Type100	0.472828	2.91E-05	0
260	19	Facility_Type200	0.860646	0.004854	6.58E-210
260	20	Facility_Type210	0.865779	0.002148	0
260	21	Facility_Type240	0.011128	1.14E-05	0
260	22	Facility_Type250	1.366433	0.006207	0
260	23	Facility_Type260	0.131124	1.48E-06	0
260	24	Facility_Type270	0.065142	4.76E-07	0
260	25	Facility_Type300	0.968506	0.011477	0.005301
260	26	Facility_Type400	0.410271	0.000545	0
260	27	Facility_Type500	0.0206	7.55E-07	0
260	28	Facility_Type620	0.269477	1.43E-05	0
260	29	Facility_Type630	0.093642	5.12E-06	0
260	30	Facility_Type700	0.910838	0.00172	0
260	31	Facility_Type710	0.442104	0.000636	0
260	32	Facility_Type800	1.576002	0.000119	0
260	33	Facility_Type801	0.021801	2.26E-06	0
260	34	Facility_Type802	0.109147	3.38E-05	0
260	35	Facility_Type803	3.71E-05	5.47E-09	0
260	36	Facility_Type804	0.929659	1.81E-06	0
260	37	Facility_Type805	0.171083	1.22E-05	0
260	38	Facility_Type806	0.000214	3.09E-08	0
260	39	Facility_Type807	0.433105	2.76E-07	0
260	40	Facility_Type809	8.527076	0.000157	0
260	41	Facility_Type810	0.265465	4.03E-05	0
260	42	Facility_Type811	0.015656	2.18E-07	0
260	43	Facility_Type812	0.067167	2.62E-07	0
260	44	Facility_Type813	0.20165	6.24E-07	0
260	45	Facility_Type820	0.053958	3.10E-06	0
260	46	Facility_Type830	0.207756	0.000139	0
260	47	Facility_Type840	0.013033	4.39E-08	0
260	48	Facility_Type850	0.589102	8.05E-06	0
260	49	Facility_Type860	2.459581	2.81E-05	0
260	50	Facility_Type865	0.078119	4.57E-08	0
260	51	Facility_Type870	2.677684	0.000111	0
260	52	Facility_Type880	0.507534	7.05E-06	0
260	53	Facility_Type881	898.1187	1.14E-06	0
260	54	Facility_Type883	0.007358	4.12E-07	0
260	55	Facility_Type884	0.571981	2.03E-05	0
260	56	Facility_Type885	1.133849	2.12E-06	0

260	57	Facility_Type886	0.012868	8.14E-08	0
260	58	Facility_Type887	0.01	8.35E-07	0
260	59	Facility_Type888	0.26789	8.92E-05	0
260	60	Facility_Type890	0.323101	2.33E-06	0
260	61	Facility_Type900	0.149771	6.22E-07	0
260	62	Facility_Type8091	0.377816	0.000128	0
260	63	Facility_Type8092	4.74E-05	1.54E-08	0
260	64	Committed_Indicator1	0.280173	0.011512	0
260	65	Seniority_Code100	1.738746	0.011071	0
260	66	Seniority_Code110	2.06763	0.009547	0
260	67	Seniority_Code200	0.823804	0.000514	0
260	68	Seniority_Code300	12.21781	1.66E-05	0
260	69	Region100	0.023659	0.000534	0
260	70	Region200	1.046343	0.008724	2.07E-07
260	71	Region300	0.840673	0.002503	0
260	72	Region400	0.134859	0.007773	0
260	73	Region500	0.274733	7.57E-05	0
260	74	Region600	1.322604	0.000267	0

Appendix J

LGD linear regression without nature of default

```
Call:
lm(formula = CAP_LGD_1 ~ Facility_Asset_Class + Guarantee_Indicator +
    Collateral_Indicator + Committed_Indicator + Region + Facility_Type +
    Seniority_Code + TIME_TO_DEFAULT + TIME_TO_RESOLUTION + Trade_Finance_Indicator,
    data = lgd_data_final)

Residuals:
    Min      1Q  Median      3Q     Max
-0.9300 -0.2329 -0.1395  0.1457  1.4758

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)              1.720e-01  2.085e-02   8.247  < 2e-16 ***
Facility_Asset_Class10  -2.985e-02  2.663e-02  -1.121 0.262307
Facility_Asset_Class11  -1.061e-02  5.419e-03  -1.957 0.050347 .
Facility_Asset_Class2   -3.215e-02  3.247e-03  -9.901  < 2e-16 ***
Facility_Asset_Class3   -7.091e-03  7.485e-03  -0.947 0.343402
Facility_Asset_Class4   -4.854e-03  1.194e-02  -0.407 0.684206
Facility_Asset_Class5   -4.835e-02  1.808e-02  -2.675 0.007473 **
Facility_Asset_Class6   -2.236e-02  3.686e-03  -6.066 1.32e-09 ***
Facility_Asset_Class7    1.252e-02  1.445e-02   0.866 0.386300
Facility_Asset_Class8   -1.101e-01  2.378e-02  -4.631 3.65e-06 ***
Facility_Asset_Class9   -1.266e-01  3.239e-02  -3.909 9.27e-05 ***
Guarantee_IndicatorN     5.430e-02  1.682e-02   3.228 0.001248 **
Guarantee_IndicatorY    -1.073e-02  1.695e-02  -0.633 0.526519
Collateral_IndicatorN    7.480e-02  7.178e-03  10.421  < 2e-16 ***
Collateral_IndicatorY   -7.838e-03  6.886e-03  -1.138 0.255013
Committed_IndicatorY    -5.978e-02  2.586e-03 -23.116  < 2e-16 ***
Region100                1.397e-02  1.094e-02   1.278 0.201406
Region200                1.389e-02  8.414e-03   1.650 0.098851 .
Region300               -1.503e-02  1.252e-02  -1.200 0.230019
Region400               -3.774e-02  8.624e-03  -4.376 1.21e-05 ***
Region500               -5.758e-02  3.178e-02  -1.812 0.070016 .
Region600                3.763e-02  1.645e-02   2.287 0.022177 *
Facility_Type100         1.425e-01  2.442e-02   5.836 5.35e-09 ***
Facility_Type200         7.720e-02  4.340e-03  17.787  < 2e-16 ***
Facility_Type210         3.134e-02  5.651e-03   5.546 2.93e-08 ***
Facility_Type240         3.955e-01  2.400e-02  16.481  < 2e-16 ***
Facility_Type250         6.253e-02  5.021e-03  12.453  < 2e-16 ***
Facility_Type260        -4.187e-03  2.505e-02  -0.167 0.867276
Facility_Type270         5.514e-02  1.286e-01   0.429 0.668121
Facility_Type300        -1.657e-02  3.844e-03  -4.312 1.62e-05 ***
Facility_Type400         4.903e-02  8.420e-03   5.824 5.77e-09 ***
Facility_Type500        -1.755e-01  2.102e-01  -0.835 0.403801
Facility_Type620         3.322e-01  7.869e-02   4.222 2.42e-05 ***
Facility_Type630         3.704e-01  4.479e-02   8.270  < 2e-16 ***
Facility_Type700        -6.025e-02  5.796e-03 -10.395  < 2e-16 ***
Facility_Type710         4.445e-02  8.590e-03   5.175 2.28e-07 ***
Facility_Type800        -8.370e-02  1.457e-02  -5.744 9.28e-09 ***
Facility_Type801        -9.476e-02  1.975e-02  -4.797 1.61e-06 ***
Facility_Type802         8.533e-02  2.446e-02   3.488 0.000486 ***
Facility_Type803        -1.391e-01  7.853e-02  -1.771 0.076520 .
Facility_Type804        -2.658e-01  1.386e-01  -1.918 0.055073 .
Facility_Type805        -1.307e-01  2.878e-02  -4.541 5.61e-06 ***
Facility_Type806        -1.725e-01  6.440e-02  -2.679 0.007393 **
Facility_Type807        -2.627e-01  3.636e-01  -0.723 0.469904
Facility_Type809        -1.507e-01  2.837e-02  -5.313 1.08e-07 ***
Facility_Type8091       -9.667e-02  3.213e-02  -3.008 0.002627 **
Facility_Type8092        1.783e-01  3.645e-01   0.489 0.624699
Facility_Type810        -1.650e-01  5.065e-02  -3.256 0.001128 **
Facility_Type811        -1.961e-01  6.250e-02  -3.138 0.001703 **
Facility_Type812         2.240e-01  1.385e-01   1.617 0.105925
Facility_Type813         5.929e-02  1.821e-01   0.326 0.744797
Facility_Type820         1.493e-01  3.257e-02   4.585 4.55e-06 ***
Facility_Type830         2.805e-01  2.517e-02  11.147  < 2e-16 ***
Facility_Type840        -1.663e-01  7.049e-02  -2.360 0.018286 *
Facility_Type850         2.774e-01  1.098e-01   2.526 0.011552 *
Facility_Type860        -1.114e-01  5.566e-02  -2.001 0.045404 *
Facility_Type865        -9.066e-02  1.215e-01  -0.746 0.455725
Facility_Type870         7.170e-02  4.368e-02   1.642 0.100664
Facility_Type880        -4.203e-02  1.819e-01  -0.231 0.817214
Facility_Type881         6.103e-01  1.097e-01   5.563 2.66e-08 ***
Facility_Type883         8.451e-02  5.644e-02   1.497 0.134292
Facility_Type884         2.228e-02  3.411e-02   0.653 0.513624
Facility_Type885         2.708e-01  1.486e-01   1.823 0.068273 .
Facility_Type886         2.023e-01  1.051e-01   1.925 0.054257 .
Facility_Type887        -1.444e-02  1.214e-01  -0.119 0.905340
Facility_Type888        -1.245e-01  3.081e-02  -4.042 5.31e-05 ***
Facility_Type890         1.348e-01  7.008e-02   1.923 0.054430 .
Facility_Type900         2.716e-01  2.573e-02  10.556  < 2e-16 ***
Seniority_Code100       -6.502e-02  3.480e-03 -18.682  < 2e-16 ***
Seniority_Code110       -4.981e-02  4.216e-03 -11.814  < 2e-16 ***
Seniority_Code200       -1.676e-02  8.117e-03  -2.065 0.038922 *
Seniority_Code300       -3.514e-02  6.865e-02  -0.512 0.608719
TIME_TO_DEFAULT         -1.029e-06  5.360e-07  -1.919 0.054931 .
TIME_TO_RESOLUTION       1.323e-04  1.320e-06 100.253  < 2e-16 ***
Trade_Finance_IndicatorN 2.217e-02  2.984e-03   7.430 1.09e-13 ***
Trade_Finance_IndicatorY 9.344e-02  1.498e-02   6.238 4.46e-10 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.3635 on 133200 degrees of freedom
Multiple R-squared:  0.1167,    Adjusted R-squared:  0.1162
F-statistic: 234.6 on 75 and 133200 DF,  p-value: < 2.2e-16
```

Appendix K

LGD linear regression with nature of default

```
Call:
lm(formula = CAP_LGD_1 ~ Facility_Asset_Class + Guarantee_Indicator +
    Collateral_Indicator + Committed_Indicator + Region + Facility_Type +
    Seniority_Code + TIME_TO_DEFAULT + TIME_TO_RESOLUTION + Trade_Finance_Indicator +
    Nature_Of_Default, data = lgd_data_final)

Residuals:
    Min      1Q  Median      3Q     Max
-0.9835 -0.2338 -0.1329  0.1443  1.5106

Coefficients:
                           Estimate Std. Error t value Pr(>|t|)
(Intercept)               6.780e-02  2.138e-02   3.170  0.00152 **
Facility_Asset_Class10   -2.061e-02  2.648e-02  -0.778  0.43654
Facility_Asset_Class11   -4.083e-03  5.418e-03  -0.754  0.45115
Facility_Asset_Class2    -4.186e-02  3.269e-03 -12.805  < 2e-16 ***
Facility_Asset_Class3    -1.175e-02  7.450e-03  -1.577  0.11489
Facility_Asset_Class4     8.189e-03  1.189e-02   0.689  0.49085
Facility_Asset_Class5    -4.888e-02  1.797e-02  -2.720  0.00653 **
Facility_Asset_Class6     1.245e-03  3.755e-03   0.332  0.74024
Facility_Asset_Class7     1.254e-02  1.437e-02   0.873  0.38288
Facility_Asset_Class8    -1.121e-01  2.366e-02  -4.738 2.16e-06 ***
Facility_Asset_Class9    -1.340e-01  3.220e-02  -4.161 3.17e-05 ***
Guarantee_IndicatorN      1.042e-01  1.695e-02   6.147 7.91e-10 ***
Guarantee_IndicatorY      4.438e-02  1.707e-02   2.599  0.00934 **
Collateral_IndicatorN     7.097e-02  7.183e-03   9.880  < 2e-16 ***
Collateral_IndicatorY    -3.370e-03  6.886e-03  -0.489  0.62452
Committed_IndicatorY     -5.977e-02  2.713e-03 -22.034  < 2e-16 ***
Region100                 2.151e-02  1.098e-02   1.959  0.05014 .
Region200                -1.870e-04  8.458e-03  -0.022  0.98236
Region300                -3.433e-02  1.248e-02  -2.751  0.00595 **
Region400                -5.460e-02  8.633e-03  -6.325 2.54e-10 ***
Region500                -4.494e-02  3.159e-02  -1.422  0.15492
Region600                 2.737e-02  1.640e-02   1.669  0.09516 .
Facility_Type100          1.041e-01  2.431e-02   4.283 1.85e-05 ***
Facility_Type200          7.199e-02  4.348e-03  16.557  < 2e-16 ***
Facility_Type210          2.351e-02  5.643e-03   4.167 3.09e-05 ***
Facility_Type240          4.296e-01  2.395e-02  17.939  < 2e-16 ***
Facility_Type250          5.349e-02  5.035e-03  10.625  < 2e-16 ***
Facility_Type260          4.583e-02  2.499e-02   1.834  0.06668 .
Facility_Type270          3.866e-02  1.278e-01   0.302  0.76234
Facility_Type300         -2.570e-02  3.921e-03  -6.553 5.66e-11 ***
Facility_Type400          6.854e-02  8.386e-03   8.174 3.02e-16 ***
Facility_Type500         -1.750e-01  2.089e-01  -0.838  0.40219
Facility_Type620          3.248e-01  7.822e-02   4.152 3.30e-05 ***
Facility_Type630          3.945e-01  4.457e-02   8.852  < 2e-16 ***
Facility_Type700         -5.899e-02  5.891e-03 -10.014  < 2e-16 ***
Facility_Type710          4.589e-02  8.539e-03   5.374 7.71e-08 ***
Facility_Type800         -9.401e-02  1.452e-02  -6.473 9.66e-11 ***
Facility_Type801         -8.796e-02  1.966e-02  -4.475 7.65e-06 ***
Facility_Type802          5.086e-02  2.436e-02   2.088  0.03680 *
Facility_Type803         -1.791e-01  7.808e-02  -2.294  0.02182 *
Facility_Type804         -3.637e-01  1.378e-01  -2.640  0.00828 **
Facility_Type805         -1.142e-01  2.861e-02  -3.990 6.62e-05 ***
Facility_Type806         -1.644e-01  6.403e-02  -2.567  0.01026 *
Facility_Type807         -3.254e-01  3.614e-01  -0.900  0.36793
Facility_Type809         -1.869e-01  2.822e-02  -6.623 3.53e-11 ***
Facility_Type8091        -1.385e-01  3.198e-02  -4.329 1.50e-05 ***
Facility_Type8092         4.566e-02  3.623e-01   0.126  0.89972
Facility_Type810         -1.273e-01  5.041e-02  -2.525  0.01157 *
Facility_Type811         -2.430e-01  6.215e-02  -3.910 9.25e-05 ***
Facility_Type812          2.459e-01  1.377e-01   1.786  0.07414 .
Facility_Type813          3.243e-02  1.810e-01   0.179  0.85782
Facility_Type820          9.480e-02  3.246e-02   2.921  0.00349 **
Facility_Type830          2.420e-01  2.511e-02   9.636  < 2e-16 ***
Facility_Type840         -2.088e-01  7.009e-02  -2.980  0.00289 **
Facility_Type850          2.243e-01  1.092e-01   2.055  0.03993 *
Facility_Type860         -1.383e-01  5.534e-02  -2.499  0.01247 *
Facility_Type865         -1.402e-01  1.209e-01  -1.160  0.24601
Facility_Type870          5.715e-02  4.342e-02   1.316  0.18814
Facility_Type880         -8.690e-02  1.808e-01  -0.481  0.63067
Facility_Type881          5.768e-01  1.092e-01   5.284 1.27e-07 ***
Facility_Type883          4.112e-02  5.611e-02   0.733  0.46363
Facility_Type884          1.159e-02  3.392e-02   0.342  0.73257
Facility_Type885          2.335e-01  1.477e-01   1.581  0.11377
Facility_Type886          1.539e-01  1.045e-01   1.473  0.14079
Facility_Type887         -6.556e-02  1.207e-01  -0.543  0.58693
Facility_Type888         -1.290e-01  3.063e-02  -4.213 2.52e-05 ***
Facility_Type890          1.218e-01  6.966e-02   1.748  0.08049 .
Facility_Type900          1.880e-01  2.582e-02   7.281 3.32e-13 ***
Seniority_Code100        -5.535e-02  3.677e-03 -15.052  < 2e-16 ***
Seniority_Code110        -4.478e-02  4.331e-03 -10.339  < 2e-16 ***
Seniority_Code200        -8.390e-03  8.106e-03  -1.035  0.30064
Seniority_Code300        -5.507e-02  6.824e-02  -0.807  0.41961
TIME_TO_DEFAULT          -2.938e-06  5.358e-07  -5.484 4.17e-08 ***
TIME_TO_RESOLUTION        1.265e-04  1.331e-06  95.113  < 2e-16 ***
Trade_Finance_IndicatorN  5.375e-02  3.225e-03  16.667  < 2e-16 ***
Trade_Finance_IndicatorY  1.612e-01  1.501e-02  10.744  < 2e-16 ***
Nature_Of_Default200      2.908e-02  4.608e-03   6.310 2.81e-10 ***
Nature_Of_Default210      6.050e-03  4.965e-03   1.218  0.22308
Nature_Of_Default220      1.303e-01  5.786e-03  22.517  < 2e-16 ***
Nature_Of_Default230      1.164e-01  5.382e-03  21.634  < 2e-16 ***
Nature_Of_Default240      2.434e-01  2.365e-02  10.295  < 2e-16 ***
Nature_Of_Default250      1.171e-01  8.190e-03  14.298  < 2e-16 ***
Nature_Of_Default260      7.118e-02  4.330e-03  16.439  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.3613 on 133193 degrees of freedom
Multiple R-squared:  0.1276,    Adjusted R-squared:  0.127
F-statistic: 237.5 on 82 and 133193 DF,  p-value: < 2.2e-16
```

Appendix L

Decision Tree Chart
(Feature number correspond to the decision tree chart)

Feature Number	Feature	Feature Number	Feature
1	Facility_Asset_Class1	38	Facility_Type810
2	Facility_Asset_Class2	39	Facility_Type811
3	Facility_Asset_Class3	40	Facility_Type812
4	Facility_Asset_Class4	41	Facility_Type813
5	Facility_Asset_Class5	42	Facility_Type820
6	Facility_Asset_Class6	43	Facility_Type830
7	Facility_Asset_Class7	44	Facility_Type840
8	Facility_Asset_Class8	45	Facility_Type850
9	Facility_Asset_Class9	46	Facility_Type860
10	Facility_Asset_Class10	47	Facility_Type865
11	Facility_Asset_Class11	48	Facility_Type870
12	Trade_Finance_IndicatorY	49	Facility_Type880
13	Guarantee_IndicatorY	50	Facility_Type881
14	Collateral_IndicatorY	51	Facility_Type883
15	Facility_Type100	52	Facility_Type884
16	Facility_Type200	53	Facility_Type885
17	Facility_Type210	54	Facility_Type886
18	Facility_Type240	55	Facility_Type887
19	Facility_Type250	56	Facility_Type888
20	Facility_Type260	57	Facility_Type890
21	Facility_Type270	58	Facility_Type900
22	Facility_Type300	59	Facility_Type8091
23	Facility_Type400	60	Facility_Type8092
24	Facility_Type500	61	Committed_Indicator1
25	Facility_Type620	62	Seniority_Code100
26	Facility_Type630	63	Seniority_Code110
27	Facility_Type700	64	Seniority_Code200
28	Facility_Type710	65	Seniority_Code300
29	Facility_Type800	66	Region100
30	Facility_Type801	67	Region200
31	Facility_Type802	68	Region300
32	Facility_Type803	69	Region400
33	Facility_Type804	70	Region500
34	Facility_Type805	71	Region600
35	Facility_Type806	72	TIME_TO_DEFAULT
36	Facility_Type807	73	TIME_TO_RESOLUTION
37	Facility_Type809	74	TIME_IN_DEFAULT_UNRESOLVED

```
|--- feature_66 <= 0.50
|   |--- feature_11 <= 0.50
|   |   |--- feature_61 <= 0.50
|   |   |   |--- feature_60 <= 0.50
|   |   |   |   |--- feature_72 <= 496.50
|   |   |   |   |   |--- feature_0 <= 0.50
|   |   |   |   |   |   |--- feature_72 <= 54.50
|   |   |   |   |   |   |   |--- class: 240.0
|   |   |   |   |   |   |--- feature_72 >  54.50
|   |   |   |   |   |   |   |--- feature_71 <= 556.00
|   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |   |--- feature_71 >  556.00
|   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |--- feature_0 >  0.50
|   |   |   |   |   |   |--- feature_72 <= 296.50
|   |   |   |   |   |   |   |--- feature_13 <= 0.50
|   |   |   |   |   |   |   |   |--- feature_72 <= 77.50
|   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |   |   |--- feature_72 >  77.50
|   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |   |--- feature_13 >  0.50
|   |   |   |   |   |   |   |   |--- feature_72 <= 118.50
|   |   |   |   |   |   |   |   |   |--- feature_72 <= 24.50
|   |   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |   |   |   |--- feature_72 >  24.50
|   |   |   |   |   |   |   |   |   |   |--- feature_72 <= 71.50
|   |   |   |   |   |   |   |   |   |   |   |--- truncated branch of depth 2
|   |   |   |   |   |   |   |   |   |   |--- feature_72 >  71.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |   |   |--- feature_72 >  118.50
|   |   |   |   |   |   |   |   |   |--- feature_71 <= 1771.00
|   |   |   |   |   |   |   |   |   |   |--- feature_71 <= 657.00
|   |   |   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |   |   |   |   |--- feature_71 >  657.00
|   |   |   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |   |   |   |--- feature_71 >  1771.00
|   |   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |--- feature_72 >  296.50
|   |   |   |   |   |   |   |--- feature_71 <= 436.50
|   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |   |--- feature_71 >  436.50
|   |   |   |   |   |   |   |   |--- feature_21 <= 0.50
|   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |   |   |--- feature_21 >  0.50
|   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |--- feature_72 >  496.50
|   |   |   |   |   |--- feature_0 <= 0.50
|   |   |   |   |   |   |--- feature_72 <= 1255.50
|   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |--- feature_72 >  1255.50
|   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |--- feature_0 >  0.50
|   |   |   |   |   |   |--- feature_71 <= 9.00
|   |   |   |   |   |   |   |--- feature_72 <= 1310.00
|   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |   |--- feature_72 >  1310.00
|   |   |   |   |   |   |   |   |--- class: 210.0
```

```
|   |   |   |   |   |   |--- feature_71 >  9.00
|   |   |   |   |   |   |   |--- feature_72 <= 1077.50
|   |   |   |   |   |   |   |   |--- feature_71 <= 365.50
|   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |   |   |--- feature_71 >  365.50
|   |   |   |   |   |   |   |   |   |--- feature_21 <= 0.50
|   |   |   |   |   |   |   |   |   |   |--- feature_13 <= 0.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |   |   |   |   |--- feature_13 >  0.50
|   |   |   |   |   |   |   |   |   |   |   |--- truncated branch of depth 2
|   |   |   |   |   |   |   |   |   |--- feature_21 >  0.50
|   |   |   |   |   |   |   |   |   |   |--- feature_71 <= 819.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |   |   |   |   |--- feature_71 >  819.50
|   |   |   |   |   |   |   |   |   |   |   |--- truncated branch of depth 2
|   |   |   |   |   |   |   |--- feature_72 >  1077.50
|   |   |   |   |   |   |   |   |--- feature_72 <= 1638.50
|   |   |   |   |   |   |   |   |   |--- feature_71 <= 1417.50
|   |   |   |   |   |   |   |   |   |   |--- feature_72 <= 1351.00
|   |   |   |   |   |   |   |   |   |   |   |--- truncated branch of depth 2
|   |   |   |   |   |   |   |   |   |   |--- feature_72 >  1351.00
|   |   |   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |   |   |--- feature_71 >  1417.50
|   |   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |   |--- feature_72 >  1638.50
|   |   |   |   |   |   |   |   |   |--- feature_72 <= 2473.50
|   |   |   |   |   |   |   |   |   |   |--- feature_71 <= 669.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |   |   |   |--- feature_71 >  669.50
|   |   |   |   |   |   |   |   |   |   |   |--- truncated branch of depth 2
|   |   |   |   |   |   |   |   |   |--- feature_72 >  2473.50
|   |   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |--- feature_60 >  0.50
|   |   |   |   |--- feature_0 <= 0.50
|   |   |   |   |   |--- feature_68 <= 0.50
|   |   |   |   |   |   |--- feature_71 <= 31.00
|   |   |   |   |   |   |   |--- feature_72 <= 396.50
|   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |--- feature_72 >  396.50
|   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |--- feature_71 >  31.00
|   |   |   |   |   |   |   |--- feature_73 <= 2471.50
|   |   |   |   |   |   |   |   |--- class: 250.0
|   |   |   |   |   |   |   |--- feature_73 >  2471.50
|   |   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |--- feature_68 >  0.50
|   |   |   |   |   |   |--- feature_72 <= 745.00
|   |   |   |   |   |   |   |--- feature_12 <= 0.50
|   |   |   |   |   |   |   |   |--- feature_72 <= 210.50
|   |   |   |   |   |   |   |   |   |--- class: 240.0
|   |   |   |   |   |   |   |   |--- feature_72 >  210.50
|   |   |   |   |   |   |   |   |   |--- feature_71 <= 366.50
|   |   |   |   |   |   |   |   |   |   |--- class: 250.0
|   |   |   |   |   |   |   |   |   |--- feature_71 >  366.50
|   |   |   |   |   |   |   |   |   |   |--- class: 240.0
|   |   |   |   |   |   |   |--- feature_12 >  0.50
|   |   |   |   |   |   |   |   |--- class: 250.0
```

```
|   |   |   |   |   |   |--- feature_72 >  745.00
|   |   |   |   |   |   |   |--- feature_71 <= 213.00
|   |   |   |   |   |   |   |   |--- class: 250.0
|   |   |   |   |   |   |   |--- feature_71 >  213.00
|   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |--- feature_0 >  0.50
|   |   |   |   |   |--- feature_71 <= 1.00
|   |   |   |   |   |   |--- feature_72 <= 1468.00
|   |   |   |   |   |   |   |--- feature_68 <= 0.50
|   |   |   |   |   |   |   |   |--- feature_16 <= 0.50
|   |   |   |   |   |   |   |   |   |--- feature_21 <= 0.50
|   |   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |   |   |--- feature_21 >  0.50
|   |   |   |   |   |   |   |   |   |   |--- feature_72 <= 517.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |   |   |   |--- feature_72 >  517.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |   |--- feature_16 >  0.50
|   |   |   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |   |   |--- feature_68 >  0.50
|   |   |   |   |   |   |   |   |--- feature_72 <= 638.50
|   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |   |--- feature_72 >  638.50
|   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |--- feature_72 >  1468.00
|   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |--- feature_71 >  1.00
|   |   |   |   |   |   |--- feature_21 <= 0.50
|   |   |   |   |   |   |   |--- feature_26 <= 0.50
|   |   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |   |   |--- feature_26 >  0.50
|   |   |   |   |   |   |   |   |--- feature_71 <= 1006.50
|   |   |   |   |   |   |   |   |   |--- feature_71 <= 665.50
|   |   |   |   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |   |   |   |   |--- feature_71 >  665.50
|   |   |   |   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |   |   |   |--- feature_71 >  1006.50
|   |   |   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |   |--- feature_21 >  0.50
|   |   |   |   |   |   |   |--- feature_13 <= 0.50
|   |   |   |   |   |   |   |   |--- feature_72 <= 29.50
|   |   |   |   |   |   |   |   |   |--- feature_73 <= 1920.00
|   |   |   |   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |   |   |   |   |--- feature_73 >  1920.00
|   |   |   |   |   |   |   |   |   |   |--- feature_73 <= 3195.00
|   |   |   |   |   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |   |   |   |   |   |--- feature_73 >  3195.00
|   |   |   |   |   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |   |   |   |--- feature_72 >  29.50
|   |   |   |   |   |   |   |   |   |--- feature_71 <= 161.50
|   |   |   |   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |   |   |   |   |--- feature_71 >  161.50
|   |   |   |   |   |   |   |   |   |   |--- feature_71 <= 1181.50
|   |   |   |   |   |   |   |   |   |   |   |--- truncated branch of depth 3
|   |   |   |   |   |   |   |   |   |   |--- feature_71 >  1181.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |   |   |--- feature_13 >  0.50
```

```
|   |   |   |   |   |   |   |   |--- feature_63 <= 0.50
|   |   |   |   |   |   |   |   |   |--- feature_71 <= 834.00
|   |   |   |   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |   |   |   |   |--- feature_71 >  834.00
|   |   |   |   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |   |   |   |--- feature_63 >  0.50
|   |   |   |   |   |   |   |   |   |--- feature_71 <= 1154.50
|   |   |   |   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |   |   |   |   |--- feature_71 >  1154.50
|   |   |   |   |   |   |   |   |   |   |--- class: 200.0
|   |   |--- feature_61 >  0.50
|   |   |   |--- feature_71 <= 30.50
|   |   |   |   |--- feature_72 <= 1379.50
|   |   |   |   |   |--- feature_13 <= 0.50
|   |   |   |   |   |   |--- class: 230.0
|   |   |   |   |   |--- feature_13 >  0.50
|   |   |   |   |   |   |--- feature_12 <= 0.50
|   |   |   |   |   |   |   |--- feature_21 <= 0.50
|   |   |   |   |   |   |   |   |--- feature_72 <= 323.50
|   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |   |   |--- feature_72 >  323.50
|   |   |   |   |   |   |   |   |   |--- feature_72 <= 607.50
|   |   |   |   |   |   |   |   |   |   |--- class: 250.0
|   |   |   |   |   |   |   |   |   |--- feature_72 >  607.50
|   |   |   |   |   |   |   |   |   |   |--- class: 250.0
|   |   |   |   |   |   |   |--- feature_21 >  0.50
|   |   |   |   |   |   |   |   |--- feature_72 <= 459.00
|   |   |   |   |   |   |   |   |   |--- class: 250.0
|   |   |   |   |   |   |   |   |--- feature_72 >  459.00
|   |   |   |   |   |   |   |   |   |--- class: 250.0
|   |   |   |   |   |   |--- feature_12 >  0.50
|   |   |   |   |   |   |   |--- feature_72 <= 622.50
|   |   |   |   |   |   |   |   |--- feature_21 <= 0.50
|   |   |   |   |   |   |   |   |   |--- feature_18 <= 0.50
|   |   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |   |   |   |--- feature_18 >  0.50
|   |   |   |   |   |   |   |   |   |   |--- feature_72 <= 278.00
|   |   |   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |   |   |   |   |--- feature_72 >  278.00
|   |   |   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |   |   |--- feature_21 >  0.50
|   |   |   |   |   |   |   |   |   |--- feature_72 <= 375.50
|   |   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |   |   |   |--- feature_72 >  375.50
|   |   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |   |--- feature_72 >  622.50
|   |   |   |   |   |   |   |   |--- feature_72 <= 1022.50
|   |   |   |   |   |   |   |   |   |--- feature_72 <= 723.50
|   |   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |   |   |   |--- feature_72 >  723.50
|   |   |   |   |   |   |   |   |   |   |--- feature_72 <= 916.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |   |   |   |   |--- feature_72 >  916.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |   |   |--- feature_72 >  1022.50
|   |   |   |   |   |   |   |   |   |--- feature_18 <= 0.50
|   |   |   |   |   |   |   |   |   |   |--- class: 210.0
```

```
|   |   |   |   |   |   |   |   |   |--- feature_18 >  0.50
|   |   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |--- feature_72 >  1379.50
|   |   |   |   |   |--- feature_12 <= 0.50
|   |   |   |   |   |   |--- feature_21 <= 0.50
|   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |--- feature_21 >  0.50
|   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |--- feature_12 >  0.50
|   |   |   |   |   |   |--- feature_72 <= 2518.50
|   |   |   |   |   |   |   |--- feature_72 <= 1580.50
|   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |   |--- feature_72 >  1580.50
|   |   |   |   |   |   |   |   |--- feature_72 <= 1753.00
|   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |   |   |--- feature_72 >  1753.00
|   |   |   |   |   |   |   |   |   |--- feature_72 <= 1981.50
|   |   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |   |   |   |--- feature_72 >  1981.50
|   |   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |--- feature_72 >  2518.50
|   |   |   |   |   |   |   |--- feature_72 <= 3287.00
|   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |   |--- feature_72 >  3287.00
|   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |--- feature_71 >  30.50
|   |   |   |   |--- feature_0 <= 0.50
|   |   |   |   |   |--- feature_72 <= 369.00
|   |   |   |   |   |   |--- class: 250.0
|   |   |   |   |   |--- feature_72 >  369.00
|   |   |   |   |   |   |--- class: 220.0
|   |   |   |   |--- feature_0 >  0.50
|   |   |   |   |   |--- feature_26 <= 0.50
|   |   |   |   |   |   |--- feature_15 <= 0.50
|   |   |   |   |   |   |   |--- feature_60 <= 0.50
|   |   |   |   |   |   |   |   |--- class: 220.0
|   |   |   |   |   |   |   |--- feature_60 >  0.50
|   |   |   |   |   |   |   |   |--- feature_71 <= 1643.50
|   |   |   |   |   |   |   |   |   |--- feature_12 <= 0.50
|   |   |   |   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |   |   |   |   |--- feature_12 >  0.50
|   |   |   |   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |   |   |   |--- feature_71 >  1643.50
|   |   |   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |   |--- feature_15 >  0.50
|   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |--- feature_26 >  0.50
|   |   |   |   |   |   |--- feature_72 <= 665.50
|   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |   |--- feature_72 >  665.50
|   |   |   |   |   |   |   |--- class: 200.0
|   |--- feature_11 >  0.50
|   |   |--- feature_60 <= 0.50
|   |   |   |--- feature_68 <= 0.50
|   |   |   |   |--- feature_65 <= 0.50
|   |   |   |   |   |--- class: 240.0
|   |   |   |   |--- feature_65 >  0.50
```

```
|   |   |   |   |   |--- feature_72 <= 243.50
|   |   |   |   |   |   |--- class: 250.0
|   |   |   |   |   |--- feature_72 >  243.50
|   |   |   |   |   |   |--- feature_71 <= 694.00
|   |   |   |   |   |   |   |--- class: 250.0
|   |   |   |   |   |   |--- feature_71 >  694.00
|   |   |   |   |   |   |   |--- feature_71 <= 3070.00
|   |   |   |   |   |   |   |   |--- feature_71 <= 1254.50
|   |   |   |   |   |   |   |   |   |--- class: 250.0
|   |   |   |   |   |   |   |   |--- feature_71 >  1254.50
|   |   |   |   |   |   |   |   |   |--- feature_72 <= 728.50
|   |   |   |   |   |   |   |   |   |   |--- class: 250.0
|   |   |   |   |   |   |   |   |   |--- feature_72 >  728.50
|   |   |   |   |   |   |   |   |   |   |--- class: 250.0
|   |   |   |   |   |   |   |--- feature_71 >  3070.00
|   |   |   |   |   |   |   |   |--- class: 250.0
|   |   |   |--- feature_68 >  0.50
|   |   |   |   |--- feature_28 <= 0.50
|   |   |   |   |   |--- feature_0 <= 0.50
|   |   |   |   |   |   |--- feature_72 <= 1406.00
|   |   |   |   |   |   |   |--- feature_71 <= 0.50
|   |   |   |   |   |   |   |   |--- class: 240.0
|   |   |   |   |   |   |   |--- feature_71 >  0.50
|   |   |   |   |   |   |   |   |--- feature_13 <= 0.50
|   |   |   |   |   |   |   |   |   |--- feature_73 <= 276.00
|   |   |   |   |   |   |   |   |   |   |--- feature_10 <= 0.50
|   |   |   |   |   |   |   |   |   |   |   |--- truncated branch of depth 5
|   |   |   |   |   |   |   |   |   |   |--- feature_10 >  0.50
|   |   |   |   |   |   |   |   |   |   |   |--- truncated branch of depth 4
|   |   |   |   |   |   |   |   |   |--- feature_73 >  276.00
|   |   |   |   |   |   |   |   |   |   |--- feature_73 <= 2490.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 220.0
|   |   |   |   |   |   |   |   |   |   |--- feature_73 >  2490.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 220.0
|   |   |   |   |   |   |   |   |--- feature_13 >  0.50
|   |   |   |   |   |   |   |   |   |--- feature_5 <= 0.50
|   |   |   |   |   |   |   |   |   |   |--- feature_72 <= 30.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 220.0
|   |   |   |   |   |   |   |   |   |   |--- feature_72 >  30.50
|   |   |   |   |   |   |   |   |   |   |   |--- truncated branch of depth 2
|   |   |   |   |   |   |   |   |   |--- feature_5 >  0.50
|   |   |   |   |   |   |   |   |   |   |--- feature_72 <= 351.50
|   |   |   |   |   |   |   |   |   |   |   |--- truncated branch of depth 6
|   |   |   |   |   |   |   |   |   |   |--- feature_72 >  351.50
|   |   |   |   |   |   |   |   |   |   |   |--- truncated branch of depth 4
|   |   |   |   |   |   |--- feature_72 >  1406.00
|   |   |   |   |   |   |   |--- feature_13 <= 0.50
|   |   |   |   |   |   |   |   |--- feature_71 <= 1099.00
|   |   |   |   |   |   |   |   |   |--- class: 230.0
|   |   |   |   |   |   |   |   |--- feature_71 >  1099.00
|   |   |   |   |   |   |   |   |   |--- class: 230.0
|   |   |   |   |   |   |   |--- feature_13 >  0.50
|   |   |   |   |   |   |   |   |--- feature_72 <= 2345.00
|   |   |   |   |   |   |   |   |   |--- feature_72 <= 1792.50
|   |   |   |   |   |   |   |   |   |   |--- class: 230.0
|   |   |   |   |   |   |   |   |   |--- feature_72 >  1792.50
|   |   |   |   |   |   |   |   |   |   |--- feature_71 <= 880.50
```

```
|   |   |   |   |   |   |   |   |   |   |   |--- class: 230.0
|   |   |   |   |   |   |   |   |   |   |--- feature_71 >  880.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 230.0
|   |   |   |   |   |   |   |   |--- feature_72 >  2345.00
|   |   |   |   |   |   |   |   |   |--- class: 220.0
|   |   |   |   |   |--- feature_0 >  0.50
|   |   |   |   |   |   |--- feature_71 <= 1.00
|   |   |   |   |   |   |   |--- feature_61 <= 0.50
|   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |--- feature_61 >  0.50
|   |   |   |   |   |   |   |   |--- feature_18 <= 0.50
|   |   |   |   |   |   |   |   |   |--- class: 220.0
|   |   |   |   |   |   |   |   |--- feature_18 >  0.50
|   |   |   |   |   |   |   |   |   |--- class: 220.0
|   |   |   |   |   |   |--- feature_71 >  1.00
|   |   |   |   |   |   |   |--- feature_72 <= 880.00
|   |   |   |   |   |   |   |   |--- feature_73 <= 203.00
|   |   |   |   |   |   |   |   |   |--- feature_72 <= 204.50
|   |   |   |   |   |   |   |   |   |   |--- feature_72 <= 35.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 220.0
|   |   |   |   |   |   |   |   |   |   |--- feature_72 >  35.50
|   |   |   |   |   |   |   |   |   |   |   |--- truncated branch of depth 4
|   |   |   |   |   |   |   |   |   |--- feature_72 >  204.50
|   |   |   |   |   |   |   |   |   |   |--- feature_72 <= 291.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 220.0
|   |   |   |   |   |   |   |   |   |   |--- feature_72 >  291.50
|   |   |   |   |   |   |   |   |   |   |   |--- truncated branch of depth 3
|   |   |   |   |   |   |   |   |--- feature_73 >  203.00
|   |   |   |   |   |   |   |   |   |--- feature_73 <= 2397.00
|   |   |   |   |   |   |   |   |   |   |--- class: 220.0
|   |   |   |   |   |   |   |   |   |--- feature_73 >  2397.00
|   |   |   |   |   |   |   |   |   |   |--- class: 220.0
|   |   |   |   |   |   |   |--- feature_72 >  880.00
|   |   |   |   |   |   |   |   |--- feature_72 <= 1770.50
|   |   |   |   |   |   |   |   |   |--- feature_71 <= 1613.50
|   |   |   |   |   |   |   |   |   |   |--- feature_71 <= 742.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 220.0
|   |   |   |   |   |   |   |   |   |   |--- feature_71 >  742.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 220.0
|   |   |   |   |   |   |   |   |   |--- feature_71 >  1613.50
|   |   |   |   |   |   |   |   |   |   |--- class: 220.0
|   |   |   |   |   |   |   |   |--- feature_72 >  1770.50
|   |   |   |   |   |   |   |   |   |--- class: 220.0
|   |   |   |   |--- feature_28 >  0.50
|   |   |   |   |   |--- feature_72 <= 427.50
|   |   |   |   |   |   |--- feature_72 <= 188.50
|   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |--- feature_72 >  188.50
|   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |--- feature_72 >  427.50
|   |   |   |   |   |   |--- feature_72 <= 860.50
|   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |--- feature_72 >  860.50
|   |   |   |   |   |   |   |--- class: 210.0
|   |   |--- feature_60 >  0.50
|   |   |   |--- feature_72 <= 334.50
|   |   |   |   |--- feature_1 <= 0.50
```

```
|   |   |   |   |   |--- feature_71 <= 3088.00
|   |   |   |   |   |   |--- feature_68 <= 0.50
|   |   |   |   |   |   |   |--- feature_65 <= 0.50
|   |   |   |   |   |   |   |   |--- feature_62 <= 0.50
|   |   |   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |   |   |   |--- feature_62 >  0.50
|   |   |   |   |   |   |   |   |   |--- feature_71 <= 120.00
|   |   |   |   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |   |   |   |   |--- feature_71 >  120.00
|   |   |   |   |   |   |   |   |   |   |--- class: 240.0
|   |   |   |   |   |   |   |--- feature_65 >  0.50
|   |   |   |   |   |   |   |   |--- class: 250.0
|   |   |   |   |   |   |--- feature_68 >  0.50
|   |   |   |   |   |   |   |--- feature_71 <= 7.50
|   |   |   |   |   |   |   |   |--- feature_62 <= 0.50
|   |   |   |   |   |   |   |   |   |--- feature_18 <= 0.50
|   |   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |   |   |   |--- feature_18 >  0.50
|   |   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |   |   |--- feature_62 >  0.50
|   |   |   |   |   |   |   |   |   |--- feature_72 <= 62.50
|   |   |   |   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |   |   |   |   |--- feature_72 >  62.50
|   |   |   |   |   |   |   |   |   |   |--- feature_13 <= 0.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 230.0
|   |   |   |   |   |   |   |   |   |   |--- feature_13 >  0.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 230.0
|   |   |   |   |   |   |   |--- feature_71 >  7.50
|   |   |   |   |   |   |   |   |--- feature_0 <= 0.50
|   |   |   |   |   |   |   |   |   |--- feature_71 <= 1401.50
|   |   |   |   |   |   |   |   |   |   |--- class: 250.0
|   |   |   |   |   |   |   |   |   |--- feature_71 >  1401.50
|   |   |   |   |   |   |   |   |   |   |--- class: 240.0
|   |   |   |   |   |   |   |   |--- feature_0 >  0.50
|   |   |   |   |   |   |   |   |   |--- feature_13 <= 0.50
|   |   |   |   |   |   |   |   |   |   |--- feature_71 <= 579.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 240.0
|   |   |   |   |   |   |   |   |   |   |--- feature_71 >  579.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 240.0
|   |   |   |   |   |   |   |   |   |--- feature_13 >  0.50
|   |   |   |   |   |   |   |   |   |   |--- feature_71 <= 1232.00
|   |   |   |   |   |   |   |   |   |   |   |--- class: 230.0
|   |   |   |   |   |   |   |   |   |   |--- feature_71 >  1232.00
|   |   |   |   |   |   |   |   |   |   |   |--- class: 230.0
|   |   |   |   |   |--- feature_71 >  3088.00
|   |   |   |   |   |   |--- feature_71 <= 3958.50
|   |   |   |   |   |   |   |--- class: 250.0
|   |   |   |   |   |   |--- feature_71 >  3958.50
|   |   |   |   |   |   |   |--- feature_72 <= 165.50
|   |   |   |   |   |   |   |   |--- class: 250.0
|   |   |   |   |   |   |   |--- feature_72 >  165.50
|   |   |   |   |   |   |   |   |--- class: 250.0
|   |   |   |   |--- feature_1 >  0.50
|   |   |   |   |   |--- feature_61 <= 0.50
|   |   |   |   |   |   |--- feature_72 <= 190.50
|   |   |   |   |   |   |   |--- feature_72 <= 90.50
|   |   |   |   |   |   |   |   |--- class: 240.0
```

```
|   |   |   |   |   |   |   |--- feature_72 >  90.50
|   |   |   |   |   |   |   |   |--- class: 230.0
|   |   |   |   |   |   |--- feature_72 >  190.50
|   |   |   |   |   |   |   |--- class: 240.0
|   |   |   |   |   |--- feature_61 >  0.50
|   |   |   |   |   |   |--- class: 230.0
|   |   |   |--- feature_72 >  334.50
|   |   |   |   |--- feature_71 <= 2876.00
|   |   |   |   |   |--- feature_62 <= 0.50
|   |   |   |   |   |   |--- feature_65 <= 0.50
|   |   |   |   |   |   |   |--- feature_71 <= 1.50
|   |   |   |   |   |   |   |   |--- feature_61 <= 0.50
|   |   |   |   |   |   |   |   |   |--- feature_18 <= 0.50
|   |   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |   |   |   |--- feature_18 >  0.50
|   |   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |   |   |--- feature_61 >  0.50
|   |   |   |   |   |   |   |   |   |--- feature_12 <= 0.50
|   |   |   |   |   |   |   |   |   |   |--- feature_72 <= 1906.50
|   |   |   |   |   |   |   |   |   |   |   |--- truncated branch of depth 3
|   |   |   |   |   |   |   |   |   |   |--- feature_72 >  1906.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |   |   |   |--- feature_12 >  0.50
|   |   |   |   |   |   |   |   |   |   |--- class: 230.0
|   |   |   |   |   |   |   |--- feature_71 >  1.50
|   |   |   |   |   |   |   |   |--- feature_63 <= 0.50
|   |   |   |   |   |   |   |   |   |--- feature_71 <= 524.50
|   |   |   |   |   |   |   |   |   |   |--- feature_0 <= 0.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 230.0
|   |   |   |   |   |   |   |   |   |   |--- feature_0 >  0.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |   |   |   |   |--- feature_71 >  524.50
|   |   |   |   |   |   |   |   |   |   |--- feature_72 <= 509.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 230.0
|   |   |   |   |   |   |   |   |   |   |--- feature_72 >  509.50
|   |   |   |   |   |   |   |   |   |   |   |--- truncated branch of depth 3
|   |   |   |   |   |   |   |   |--- feature_63 >  0.50
|   |   |   |   |   |   |   |   |   |--- class: 250.0
|   |   |   |   |   |   |--- feature_65 >  0.50
|   |   |   |   |   |   |   |--- class: 250.0
|   |   |   |   |   |--- feature_62 >  0.50
|   |   |   |   |   |   |--- feature_68 <= 0.50
|   |   |   |   |   |   |   |--- feature_67 <= 0.50
|   |   |   |   |   |   |   |   |--- feature_71 <= 183.50
|   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |   |--- feature_71 >  183.50
|   |   |   |   |   |   |   |   |   |--- feature_13 <= 0.50
|   |   |   |   |   |   |   |   |   |   |--- class: 230.0
|   |   |   |   |   |   |   |   |   |--- feature_13 >  0.50
|   |   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |--- feature_67 >  0.50
|   |   |   |   |   |   |   |   |--- feature_72 <= 825.50
|   |   |   |   |   |   |   |   |   |--- class: 230.0
|   |   |   |   |   |   |   |   |--- feature_72 >  825.50
|   |   |   |   |   |   |   |   |   |--- class: 250.0
|   |   |   |   |   |   |--- feature_68 >  0.50
|   |   |   |   |   |   |   |--- feature_5 <= 0.50
```

```
|   |   |   |   |   |   |   |   |   |--- feature_72 <= 1309.50
|   |   |   |   |   |   |   |   |   |   |--- feature_72 <= 838.50
|   |   |   |   |   |   |   |   |   |   |   |--- feature_12 <= 0.50
|   |   |   |   |   |   |   |   |   |   |   |   |--- truncated branch of depth 3
|   |   |   |   |   |   |   |   |   |   |   |--- feature_12 >  0.50
|   |   |   |   |   |   |   |   |   |   |   |   |--- truncated branch of depth 2
|   |   |   |   |   |   |   |   |   |   |--- feature_72 >  838.50
|   |   |   |   |   |   |   |   |   |   |   |--- feature_72 <= 1028.00
|   |   |   |   |   |   |   |   |   |   |   |   |--- class: 250.0
|   |   |   |   |   |   |   |   |   |   |   |--- feature_72 >  1028.00
|   |   |   |   |   |   |   |   |   |   |   |   |--- class: 230.0
|   |   |   |   |   |   |   |   |   |--- feature_72 >  1309.50
|   |   |   |   |   |   |   |   |   |   |--- feature_72 <= 2281.00
|   |   |   |   |   |   |   |   |   |   |   |--- feature_0 <= 0.50
|   |   |   |   |   |   |   |   |   |   |   |   |--- class: 230.0
|   |   |   |   |   |   |   |   |   |   |   |--- feature_0 >  0.50
|   |   |   |   |   |   |   |   |   |   |   |   |--- truncated branch of depth 2
|   |   |   |   |   |   |   |   |   |   |--- feature_72 >  2281.00
|   |   |   |   |   |   |   |   |   |   |   |--- class: 250.0
|   |   |   |   |   |   |   |   |--- feature_5 >  0.50
|   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |--- feature_71 >  2876.00
|   |   |   |   |   |   |--- feature_72 <= 461.50
|   |   |   |   |   |   |   |--- class: 250.0
|   |   |   |   |   |   |--- feature_72 >  461.50
|   |   |   |   |   |   |   |--- feature_72 <= 802.00
|   |   |   |   |   |   |   |   |--- class: 250.0
|   |   |   |   |   |   |   |--- feature_72 >  802.00
|   |   |   |   |   |   |   |   |--- class: 250.0
|--- feature_66 >  0.50
|   |--- feature_71 <= 0.50
|   |   |--- feature_5 <= 0.50
|   |   |   |--- feature_0 <= 0.50
|   |   |   |   |--- feature_72 <= 578.50
|   |   |   |   |   |--- feature_12 <= 0.50
|   |   |   |   |   |   |--- feature_11 <= 0.50
|   |   |   |   |   |   |   |--- feature_72 <= 200.00
|   |   |   |   |   |   |   |   |--- class: 240.0
|   |   |   |   |   |   |   |--- feature_72 >  200.00
|   |   |   |   |   |   |   |   |--- class: 240.0
|   |   |   |   |   |   |--- feature_11 >  0.50
|   |   |   |   |   |   |   |--- class: 240.0
|   |   |   |   |   |--- feature_12 >  0.50
|   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |--- feature_72 >  578.50
|   |   |   |   |   |--- feature_72 <= 1092.00
|   |   |   |   |   |   |--- feature_72 <= 747.00
|   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |--- feature_72 >  747.00
|   |   |   |   |   |   |   |--- class: 240.0
|   |   |   |   |   |--- feature_72 >  1092.00
|   |   |   |   |   |   |--- feature_21 <= 0.50
|   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |--- feature_21 >  0.50
|   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |--- feature_0 >  0.50
|   |   |   |   |--- feature_60 <= 0.50
```

```
|   |   |   |   |   |--- feature_21 <= 0.50
|   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |--- feature_21 >  0.50
|   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |--- feature_60 >  0.50
|   |   |   |   |   |--- feature_21 <= 0.50
|   |   |   |   |   |   |--- feature_62 <= 0.50
|   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |   |--- feature_62 >  0.50
|   |   |   |   |   |   |   |--- feature_15 <= 0.50
|   |   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |   |   |--- feature_15 >  0.50
|   |   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |--- feature_21 >  0.50
|   |   |   |   |   |   |--- feature_12 <= 0.50
|   |   |   |   |   |   |   |--- feature_72 <= 547.50
|   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |--- feature_72 >  547.50
|   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |--- feature_12 >  0.50
|   |   |   |   |   |   |   |--- class: 260.0
|   |   |--- feature_5 >  0.50
|   |   |   |--- feature_72 <= 394.50
|   |   |   |   |--- class: 240.0
|   |   |   |--- feature_72 >  394.50
|   |   |   |   |--- class: 240.0
|   |--- feature_71 >  0.50
|   |   |--- feature_62 <= 0.50
|   |   |   |--- feature_61 <= 0.50
|   |   |   |   |--- feature_60 <= 0.50
|   |   |   |   |   |--- feature_71 <= 882.50
|   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |--- feature_71 >  882.50
|   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |--- feature_60 >  0.50
|   |   |   |   |   |--- feature_0 <= 0.50
|   |   |   |   |   |   |--- feature_10 <= 0.50
|   |   |   |   |   |   |   |--- feature_72 <= 272.00
|   |   |   |   |   |   |   |   |--- feature_5 <= 0.50
|   |   |   |   |   |   |   |   |   |--- feature_72 <= 91.50
|   |   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |   |   |--- feature_72 >  91.50
|   |   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |   |--- feature_5 >  0.50
|   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |--- feature_72 >  272.00
|   |   |   |   |   |   |   |   |--- feature_21 <= 0.50
|   |   |   |   |   |   |   |   |   |--- feature_11 <= 0.50
|   |   |   |   |   |   |   |   |   |   |--- feature_72 <= 758.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |   |   |   |--- feature_72 >  758.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |   |   |--- feature_11 >  0.50
|   |   |   |   |   |   |   |   |   |   |--- feature_71 <= 974.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |   |   |   |--- feature_71 >  974.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 260.0
```

```
|   |   |   |   |   |   |   |   |--- feature_21 >  0.50
|   |   |   |   |   |   |   |   |   |--- feature_71 <= 730.50
|   |   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |   |   |--- feature_71 >  730.50
|   |   |   |   |   |   |   |   |   |   |--- feature_72 <= 609.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |   |   |   |--- feature_72 >  609.50
|   |   |   |   |   |   |   |   |   |   |   |--- truncated branch of depth 2
|   |   |   |   |   |   |--- feature_10 >  0.50
|   |   |   |   |   |   |   |--- feature_72 <= 259.00
|   |   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |   |   |--- feature_72 >  259.00
|   |   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |--- feature_0 >  0.50
|   |   |   |   |   |   |--- feature_21 <= 0.50
|   |   |   |   |   |   |   |--- feature_72 <= 1215.50
|   |   |   |   |   |   |   |   |--- feature_13 <= 0.50
|   |   |   |   |   |   |   |   |   |--- feature_72 <= 54.50
|   |   |   |   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |   |   |   |   |--- feature_72 >  54.50
|   |   |   |   |   |   |   |   |   |   |--- feature_72 <= 161.00
|   |   |   |   |   |   |   |   |   |   |   |--- truncated branch of depth 2
|   |   |   |   |   |   |   |   |   |   |--- feature_72 >  161.00
|   |   |   |   |   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |   |   |   |--- feature_13 >  0.50
|   |   |   |   |   |   |   |   |   |--- feature_71 <= 1246.50
|   |   |   |   |   |   |   |   |   |   |--- feature_71 <= 579.50
|   |   |   |   |   |   |   |   |   |   |   |--- truncated branch of depth 2
|   |   |   |   |   |   |   |   |   |   |--- feature_71 >  579.50
|   |   |   |   |   |   |   |   |   |   |   |--- truncated branch of depth 3
|   |   |   |   |   |   |   |   |   |--- feature_71 >  1246.50
|   |   |   |   |   |   |   |   |   |   |--- feature_71 <= 2769.50
|   |   |   |   |   |   |   |   |   |   |   |--- truncated branch of depth 3
|   |   |   |   |   |   |   |   |   |   |--- feature_71 >  2769.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |   |   |--- feature_72 >  1215.50
|   |   |   |   |   |   |   |   |--- feature_72 <= 1491.50
|   |   |   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |   |   |   |--- feature_72 >  1491.50
|   |   |   |   |   |   |   |   |   |--- feature_71 <= 759.50
|   |   |   |   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |   |   |   |   |--- feature_71 >  759.50
|   |   |   |   |   |   |   |   |   |   |--- feature_72 <= 2160.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |   |   |   |   |   |--- feature_72 >  2160.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |   |--- feature_21 >  0.50
|   |   |   |   |   |   |   |--- feature_12 <= 0.50
|   |   |   |   |   |   |   |   |--- feature_71 <= 609.50
|   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |   |--- feature_71 >  609.50
|   |   |   |   |   |   |   |   |   |--- feature_71 <= 1294.00
|   |   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |   |   |--- feature_71 >  1294.00
|   |   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |--- feature_12 >  0.50
|   |   |   |   |   |   |   |   |--- feature_72 <= 517.50
```

```
|   |   |   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |   |   |   |--- feature_72 >  517.50
|   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |--- feature_61 >  0.50
|   |   |   |   |--- feature_11 <= 0.50
|   |   |   |   |   |--- feature_60 <= 0.50
|   |   |   |   |   |   |--- feature_0 <= 0.50
|   |   |   |   |   |   |   |--- feature_72 <= 533.00
|   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |--- feature_72 >  533.00
|   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |--- feature_0 >  0.50
|   |   |   |   |   |   |   |--- feature_72 <= 1095.50
|   |   |   |   |   |   |   |   |--- feature_21 <= 0.50
|   |   |   |   |   |   |   |   |   |--- feature_71 <= 1338.50
|   |   |   |   |   |   |   |   |   |   |--- feature_71 <= 911.50
|   |   |   |   |   |   |   |   |   |   |   |--- truncated branch of depth 2
|   |   |   |   |   |   |   |   |   |   |--- feature_71 >  911.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |   |   |--- feature_71 >  1338.50
|   |   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |   |--- feature_21 >  0.50
|   |   |   |   |   |   |   |   |   |--- feature_71 <= 337.00
|   |   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |   |   |--- feature_71 >  337.00
|   |   |   |   |   |   |   |   |   |   |--- feature_72 <= 396.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |   |   |   |--- feature_72 >  396.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |--- feature_72 >  1095.50
|   |   |   |   |   |   |   |   |--- feature_71 <= 1279.50
|   |   |   |   |   |   |   |   |   |--- feature_71 <= 457.50
|   |   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |   |   |--- feature_71 >  457.50
|   |   |   |   |   |   |   |   |   |   |--- feature_72 <= 1552.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |   |   |   |--- feature_72 >  1552.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |   |--- feature_71 >  1279.50
|   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |--- feature_60 >  0.50
|   |   |   |   |   |   |--- feature_26 <= 0.50
|   |   |   |   |   |   |   |--- feature_73 <= 60.50
|   |   |   |   |   |   |   |   |--- feature_72 <= 492.00
|   |   |   |   |   |   |   |   |   |--- feature_72 <= 182.50
|   |   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |   |   |--- feature_72 >  182.50
|   |   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |   |--- feature_72 >  492.00
|   |   |   |   |   |   |   |   |   |--- feature_12 <= 0.50
|   |   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |   |   |--- feature_12 >  0.50
|   |   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |--- feature_73 >  60.50
|   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |--- feature_26 >  0.50
|   |   |   |   |   |   |   |--- feature_71 <= 1279.50
```

```
|   |   |   |   |   |   |   |   |   |--- feature_71 <= 850.50
|   |   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |   |   |--- feature_71 >  850.50
|   |   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |   |--- feature_71 >  1279.50
|   |   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |--- feature_11 >  0.50
|   |   |   |   |   |--- feature_60 <= 0.50
|   |   |   |   |   |   |--- feature_1 <= 0.50
|   |   |   |   |   |   |   |--- feature_12 <= 0.50
|   |   |   |   |   |   |   |   |--- feature_71 <= 2468.00
|   |   |   |   |   |   |   |   |   |--- feature_72 <= 637.50
|   |   |   |   |   |   |   |   |   |   |--- feature_71 <= 1030.50
|   |   |   |   |   |   |   |   |   |   |   |--- truncated branch of depth 2
|   |   |   |   |   |   |   |   |   |   |--- feature_71 >  1030.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |   |   |   |--- feature_72 >  637.50
|   |   |   |   |   |   |   |   |   |   |--- feature_5 <= 0.50
|   |   |   |   |   |   |   |   |   |   |   |--- truncated branch of depth 2
|   |   |   |   |   |   |   |   |   |   |--- feature_5 >  0.50
|   |   |   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |   |   |--- feature_71 >  2468.00
|   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |   |--- feature_12 >  0.50
|   |   |   |   |   |   |   |   |--- feature_21 <= 0.50
|   |   |   |   |   |   |   |   |   |--- feature_71 <= 1231.50
|   |   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |   |   |   |--- feature_71 >  1231.50
|   |   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |   |   |--- feature_21 >  0.50
|   |   |   |   |   |   |   |   |   |--- class: 210.0
|   |   |   |   |   |   |--- feature_1 >  0.50
|   |   |   |   |   |   |   |--- class: 250.0
|   |   |   |   |   |--- feature_60 >  0.50
|   |   |   |   |   |   |--- feature_5 <= 0.50
|   |   |   |   |   |   |   |--- feature_21 <= 0.50
|   |   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |   |   |   |--- feature_21 >  0.50
|   |   |   |   |   |   |   |   |--- class: 240.0
|   |   |   |   |   |   |--- feature_5 >  0.50
|   |   |   |   |   |   |   |--- feature_12 <= 0.50
|   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |--- feature_12 >  0.50
|   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |--- feature_62 >  0.50
|   |   |   |--- feature_72 <= 81.00
|   |   |   |   |--- class: 240.0
|   |   |   |--- feature_72 >  81.00
|   |   |   |   |--- feature_11 <= 0.50
|   |   |   |   |   |--- feature_0 <= 0.50
|   |   |   |   |   |   |--- feature_15 <= 0.50
|   |   |   |   |   |   |   |--- feature_13 <= 0.50
|   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |--- feature_13 >  0.50
|   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |--- feature_15 >  0.50
|   |   |   |   |   |   |   |--- class: 260.0
```

```
|   |   |   |   |   |--- feature_0 >  0.50
|   |   |   |   |   |   |--- feature_12 <= 0.50
|   |   |   |   |   |   |   |--- feature_21 <= 0.50
|   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |   |--- feature_21 >  0.50
|   |   |   |   |   |   |   |   |--- class: 260.0
|   |   |   |   |   |   |--- feature_12 >  0.50
|   |   |   |   |   |   |   |--- class: 200.0
|   |   |   |   |--- feature_11 >  0.50
|   |   |   |   |   |--- feature_10 <= 0.50
|   |   |   |   |   |   |--- feature_1 <= 0.50
|   |   |   |   |   |   |   |--- class: 240.0
|   |   |   |   |   |   |--- feature_1 >  0.50
|   |   |   |   |   |   |   |--- class: 230.0
|   |   |   |   |   |--- feature_10 >  0.50
|   |   |   |   |   |   |--- class: 260.0
```