Final Project: Finding Fraud Faster

Executive Summary

The Problem Statements

The task is to develop and compare machine learning models to predict loan default for a major financial institution. The goal is to accurately identify which loans are likely to default, ensuring regulatory compliance and providing explainable predictions.

Key Findings:

1) Feature Importance: Through analysis, it was found that features such as **interest rate**, **last payment amount**, and **loan term** were consistently important across various models, indicating their significant impact on predicting loan default.
2) Factors that contribute to the likelihood of default include longer loan terms, higher interest rates, higher funds, higher installments, higher inquiries made in the last six months, higher total credit revolving balance, and a higher number of credit lines. Additionally, if the founder amount interval is less than 5000 or the last payment amount is less than 2000, the likelihood of default is higher. However, if the annual income exceeds a certain threshold (represented by 100000), the likelihood of default decreases significantly. Furthermore, as the total late fee increases from 0 to 18, the likelihood of default increases, but after 18, the likelihood of default decreases significantly.
3) Increasing the False Positive Rate (FPR) from 2% to 5% in the best-performing Stacking Classifier model results in significant changes: Recall increases from 28% to 49%, indicating improved sensitivity in detecting loan defaults. Precision decreases from 71% to 63%, maintaining reasonable accuracy but allowing more false alarms. The threshold decreases from 0.62 to 0.37, leading to a more lenient classification approach. At 2% FPR, the model prioritizes precision, suitable for conservative risk management. At 5% FPR, it becomes more liberal, aiming to capture more defaults but may raise more false alarms, ensuring no significant risks are overlooked.

Model Performance & Interpretation:

Based on the values of accuracy, AUC-ROC, precision, recall, and F1-score, the **stacking classifier model** performs best as a prediction model based on the test data set.

Treated "default" as the positive class (class 1) and "current" as the negative class (class 0).

- ✓ Accuracy: The model's predictions are overall correct in 88% of cases.
- ✓ Precision: Accuracy of positive predictions, indicating that when the model predicts a default, it's correct 68% of the time.
- ✓ Recall: Sensitivity to detecting positives, meaning the model identifies 40% of actual defaults.
- ✓ F1-score: The balance between precision and recall reflects a moderate trade-off between correctly identifying defaults and avoiding false alarms.
- ✓ AUC-ROC: The model's discriminatory power, with a high value of 0.9, showcases its ability to differentiate between default and non-default cases.
- ✓ AUC-PR: The stacked classifier model achieved a moderate AUC-PR value of 0.61, indicating better performance in identifying loan defaults while reducing false alarms.

Actionable Recommendations:

1) Use PR Curve and ROC AUC to operate the model. A higher area under the PR curve means better accuracy in pinpointing loan defaults while keeping false positives low. Our final model achieved a moderate AUC-PR value of 0.61, better in identifying loan defaults while reducing false alarms. ROC AUC helps assess the model's ability to separate actual defaults from non-defaults. Discriminatory power, with a value of 0.9, showcases the model's ability to differentiate between default and non-default cases. Using these curves, evaluate the model's ability to predict loan defaults and make informed decisions. Ensure model accurately identifies defaults and minimizes false alarms, supporting risk management efforts and decision-making processes.
2) By setting a threshold for the predicted probability of loan default that achieves a 5% FPR, the firm can effectively manage risk while ensuring they are not overly conservative in their predictions. Operating at a 5% false positive rate (FPR) means that out of all instances classified as "current" (0), 5% of them will be mistakenly classified as "default" (1) by the model. To operate at a 5% false positive rate (FPR) using the best-performing model, the firm needs to adjust the threshold to 0.37. The Recall is 0.49, meaning the model identifies 40% of actual defaults. The Precision is 0.63, meaning that 63% of the instances predicted as positive are actually positive.

Model Report

Methodology

1. Data Exploration and Preprocessing

1) Exploratory Data Analysis (EDA)

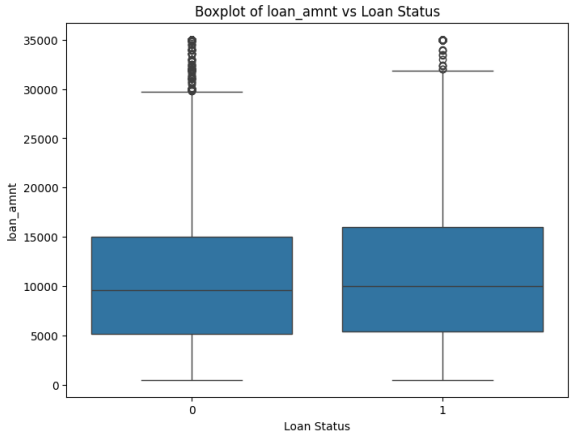
- Feature screen
  - ✓ The data set contained about 29k+ record
  - ✓ 50 numerical and categorical variables
  - ✓ The distribution of target variable is current account for 85% and default account for 15%
  - ✓ Some mistyped data, such as interest rate, term, employment length, which were recorded as an object
  - ✓ Some categorical variables get 70% unique values like id, member_id, url, emp_title, desc, title
  - ✓ Some categorical variables or numerical variables has one value like collections_12_mths_ex_med, policy_code, application_type, chargeoff_within_12_mths, tax_liens
  - ✓ Some categorical variables has more than 20% missing values like desc, mths_since_last_delinq, mths_since_last_record, next_pymnt_d
  - ✓ Some time stamp issue_d, last_pymnt_d,next_pymnt_d, last_credit_pull_d

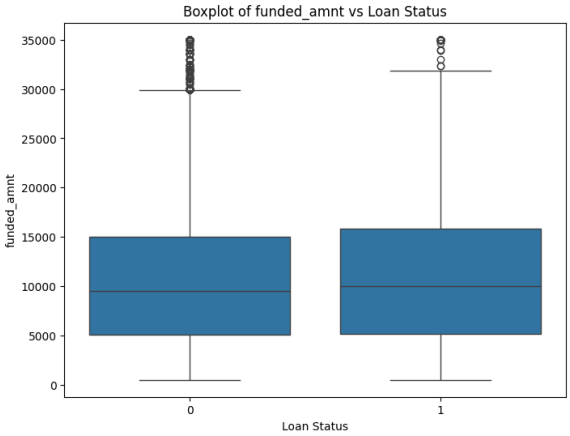
- Numerical variables analysis vs Target:

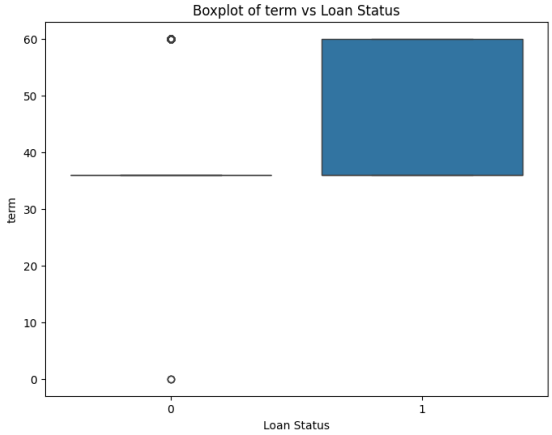
  The boxplots can be helpful in understanding how different features vary across different loan statuses.

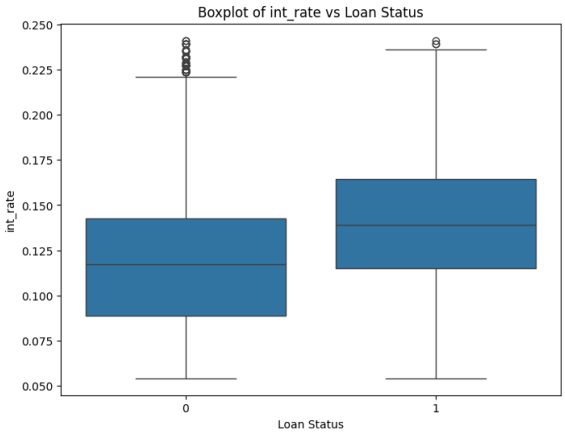
  - ✓ Based on Numerical analysis, we observe that default will be more likely to happen when the loan has
  - ✓ Higher amount of the loan applied for by the borrower
  - ✓ Higher total amount committed to that loan
  - ✓ 60-month term
  - ✓ Higher Interest rate
  - ✓ Higher dti
  - ✓ The monthly payment owned by the borrower

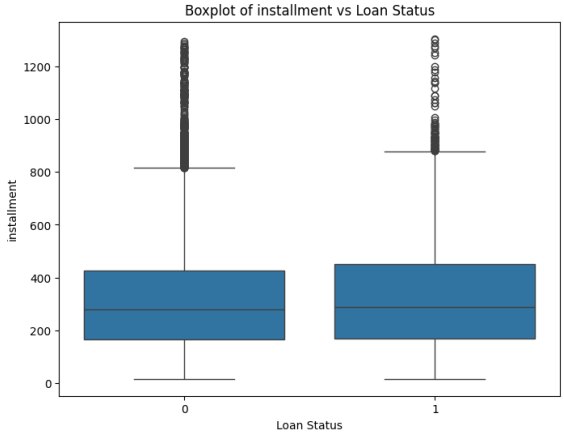
- ✓ Lower FICO range low
- ✓ Lower FICO range high
- ✓ Lower Last payment

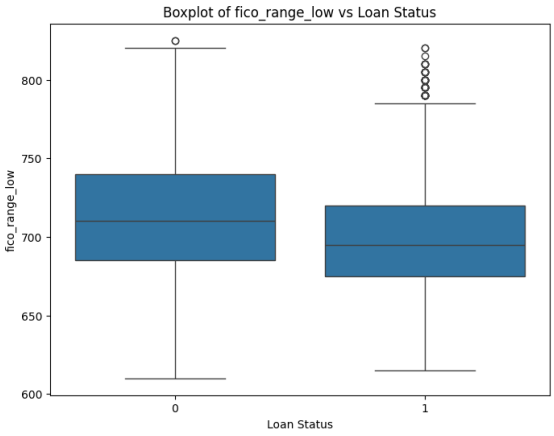


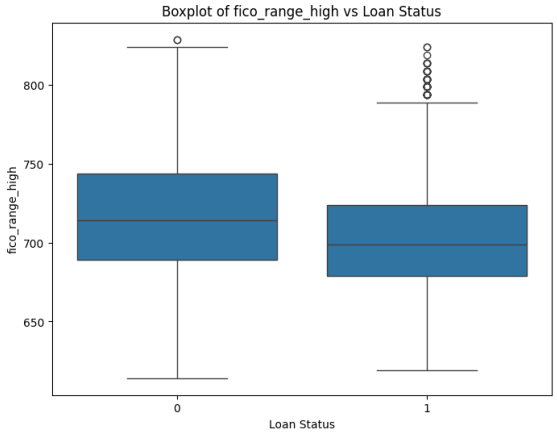


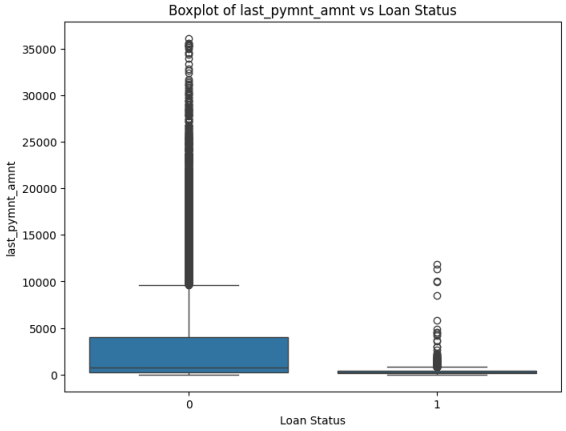


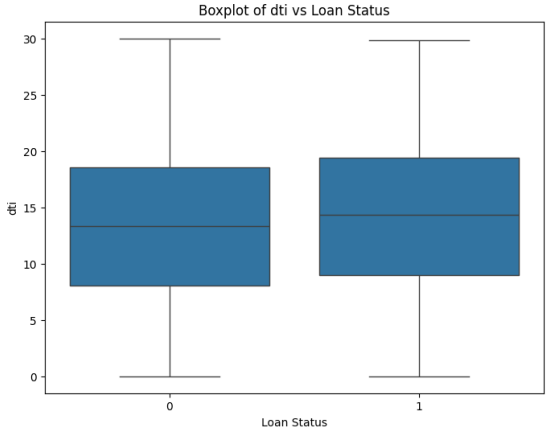








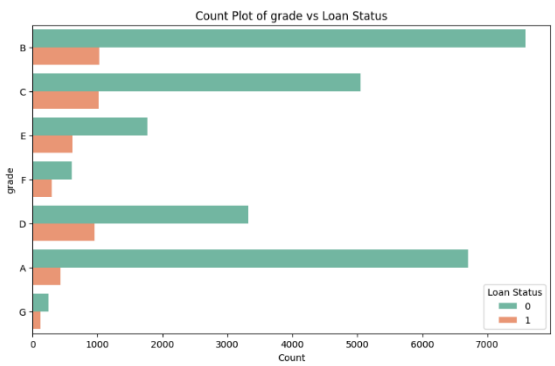


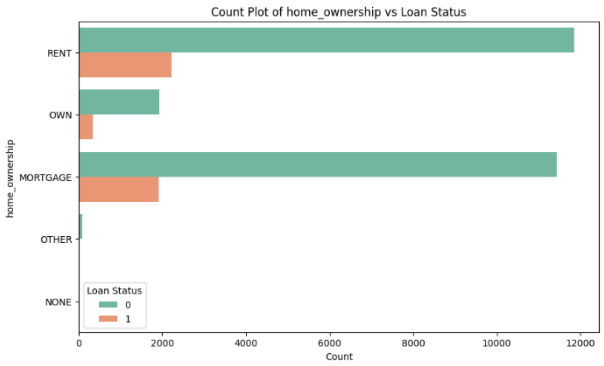


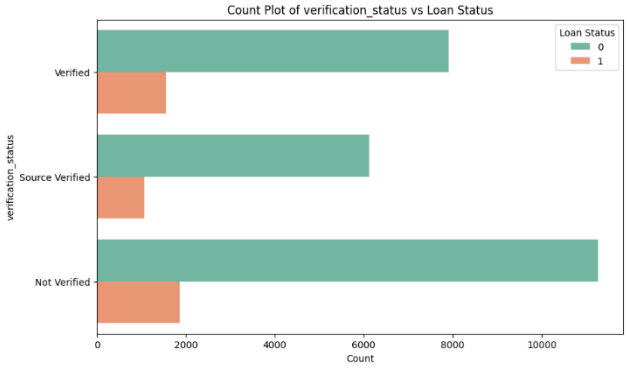
- Categorical variables analysis vs Target

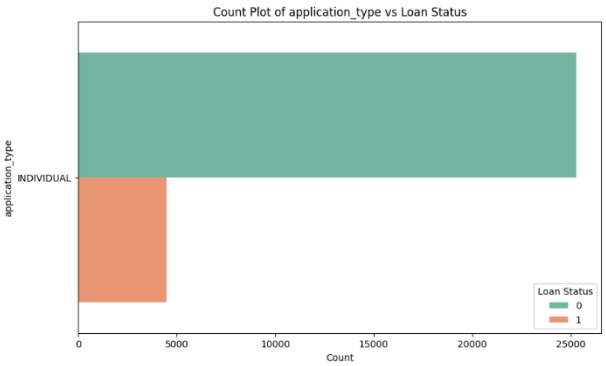
  Bar plots help understand the distribution of each categorical feature concerning different loan statuses, providing insights into potential relationships between features and loan status outcomes. Based on categorical analysis, we observe that default will be affected by

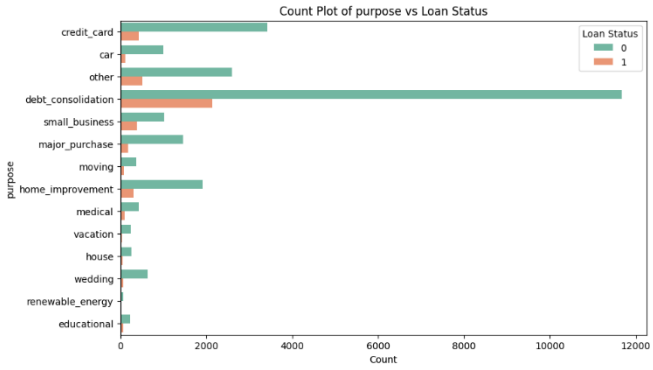
  - ✓ Grade
  - ✓ Home ownership status
  - ✓ Verification status
  - ✓ Application type
  - ✓ Purpose











2) Data Preprocessing

- Changed the wrongly typed values to the correct type to avoid confusion.

  After feature screening, the Term = “36 months” is converted from string to integer, the Interest rate = 16% is converted from string to decimal, and the Employment length = “10 years” is converted to integer.

- Address missing values, encode categorical variables, and standardize numerical features to prepare the data for modeling.

Data Transformation	Description
Impute Missing Values	The missing values in the dataset will be replaced with the mean value along each column using SimpleImputer with strategy=' mean'.
Encode Categorical Variables	Categorical variables will be encoded into one-hot encoded format using OneHotEncoder from the scikit-learn library.
Standardize Numerical Features	Numerical features will be standardized using the StandardScaler function from scikit-learn to bring them to a common scale.

2. Model Development

1) Model Training:

Developed models using Logistic Regression, Random Forest, and GBM/XGBoost, Neural Network and SKLEARN stacking ensemble model on the training data.

2) Parameter Tuning:

Optimize model parameters to enhance performance for each model.

3. Global Model Explanations

1) Model Comparison

1. Performance Metrics:

Based on the values of accuracy, AUC-ROC, precision, recall, and F1-score, the **stacking classifier model** performs best as a prediction model based on the test data set.

Treated "default" as the positive class (class 1) and "current" as the negative class (class 0).

- ✓ Accuracy: The overall correctness of the model's predictions shows that it's accurate in 88% of cases.
- ✓ Precision: Accuracy of positive predictions, indicating that when the model predicts a default, it's correct 68% of the time.
- ✓ Recall: Sensitivity to detecting positives, meaning the model identifies 40% of actual defaults.
- ✓ F1-score: The balance between precision and recall reflects a moderate trade-off between correctly identifying defaults and avoiding false alarms.
- ✓ AUC-ROC: The model's discriminatory power showcases a strong ability to differentiate between default and non-default cases, with a high value of 0.9.

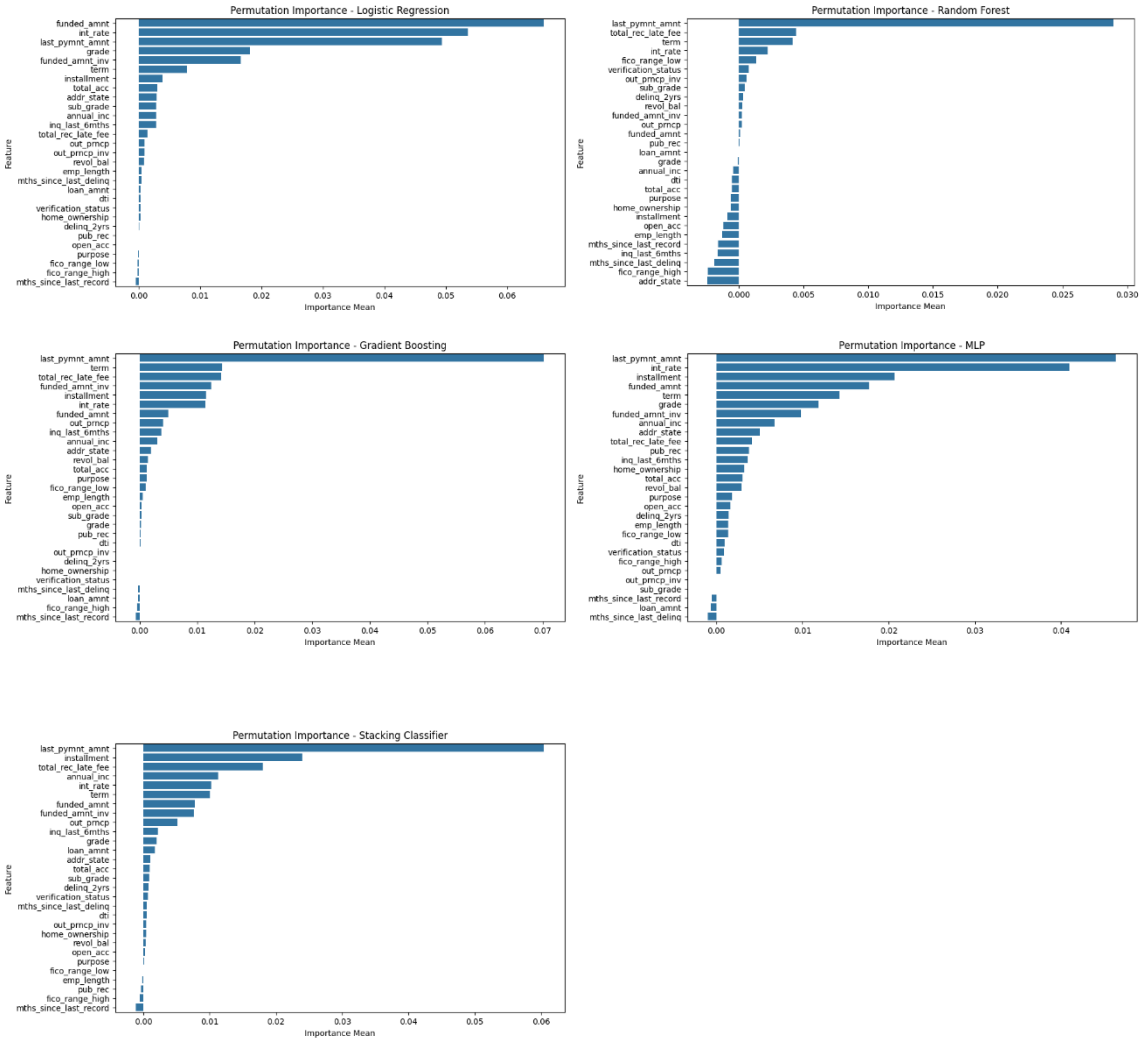
	Train set					Test set				
Model	**Accuracy**	**Precision**	**Recall**	**F1-score**	**AUC-ROC**	**Accuracy**	**Precision**	**Recall**	**F1-score**	**AUC-ROC**
Logistic Regression	0.87	0.66	0.3	0.41	0.88	0.87	0.62	0.28	0.39	0.86
Random Forest	0.98	1	0.87	0.93	1	0.86	0.61	0.18	0.27	0.84
Gradient Boosting	0.9	0.74	0.49	0.59	0.93	0.88	0.63	0.4	0.49	0.89
Neural Network	1	1	1	1	1	0.84	0.43	0.37	0.4	0.81
Stacking Classifier	0.93	0.89	0.6	0.72	0.96	0.88	0.68	0.4	0.51	0.9

2) Feature Importance

- ✓ Permutation importance is a method used to assess the significance of features in machine learning models by measuring the effect of shuffling their values on the model's performance metric.
- ✓ The differences in feature importance across various models stem from their inherent algorithmic characteristics, their ability to capture different types of patterns in the data, and how they leverage features during the learning process. For instance, logistic regression is a linear model that assumes a linear relationship between features and the target variable. At the same time, random forests and gradient-boosting machines can capture non-linear relationships through decision trees. Neural networks are highly flexible and can learn complex patterns from data, potentially capturing intricate feature interactions. Stacking classifiers combine multiple models, potentially leveraging different aspects of each base model's predictions.
- ✓ In the prediction models, it's observed that three features—Interest rate, last payment amount, and loan term—consistently rank among the top six essential features across all prediction models. This consistency suggests that these features play a crucial role in determining the model's predictive performance and are likely to significantly impact the

outcome of the predictions. Therefore, understanding the influence of these features through permutation importance analysis can provide valuable insights into how they contribute to the overall predictive power of the models.

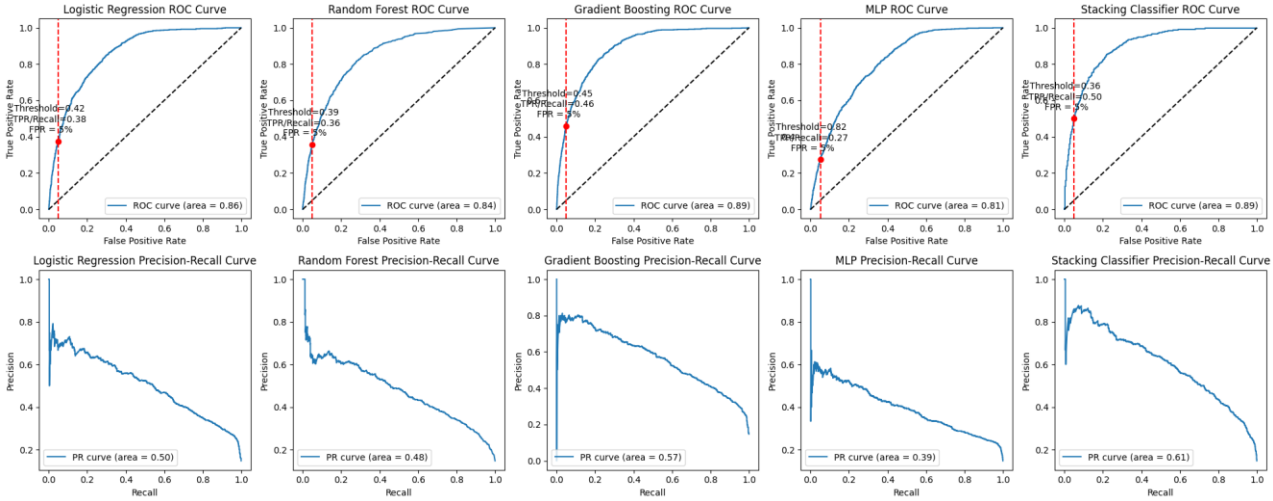
	Feature		
Appear 5 times	int_rate	last_pymnt_amnt	term
Appear 3 times	total_rec_late_fee	installment	
Appear 2 times	funded_amnt	grade	funded_amnt_inv
Appear once	fico_range_low	verification_status	verification_status



3) Roc Curve & PR Curves on the Test set

Roc Curve: higher AUC-ROC values indicate better model performance in predicting loan default, making them valuable tools for model evaluation and selection in the financial institution's context. Specifically:

- ✓ Logistic Regression (0.86): Effective in differentiating between default and current loans.
- ✓ Random Forest (0.84): Slightly lower discriminative power compared to logistic regression but still strong.
- ✓ Gradient Boosting (0.89): Superior performance in distinguishing between default and current loans.
- ✓ Neural Network (0.81): Reasonable discriminative power, slightly lower than other models.
- ✓ Stacking Classifier (0.9): Excellent discriminative power and overall performance, ranking 90% of default loans higher than current loans.



PR Curves: the precision-recall curve provides valuable insights into the performance of a binary classification model, particularly in scenarios with imbalanced class distribution. The relationship between recall and precision is negative. It's kind of a trade-off based on business demand.

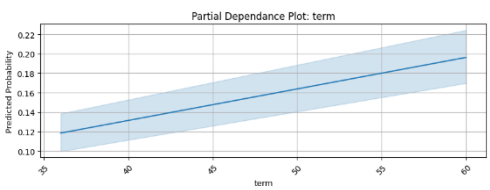
- ✓ A high recall (sensitivity) means that the model effectively identifies most default cases, minimizing the risk of overlooking potential defaults. This is crucial for risk management and ensuring the financial institution can proactively address loan defaults.
- ✓ A high precision (positive predictive value) ensures that the instances classified as defaults are highly likely to be true defaults, reducing unnecessary interventions and false alarms. It helps the institution allocate resources effectively by focusing on cases with a higher probability of default.

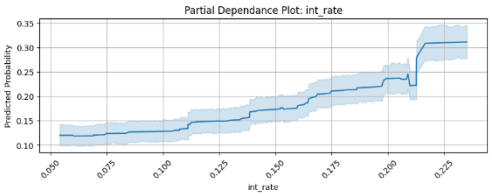
- ✓ A higher AUC-PR value indicates better performance in identifying loan defaults while reducing false alarms. The stacked classifier model achieved the highest AUC-PR value of 0.61, representing moderate precision across recall levels.

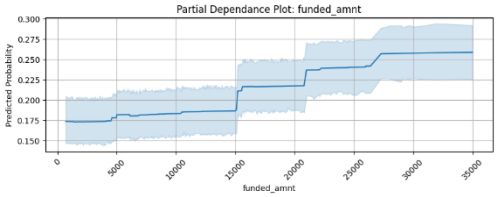
4) Partial Dependence Plot

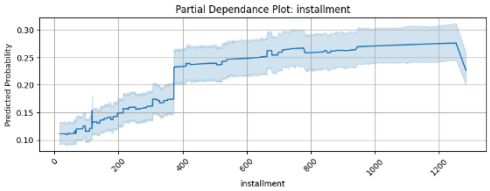
Based on the above steps, the best-performing model is the Stacking Classifier model.
The pdp_plot_numeric generates partial dependence plots (PDPs) for numeric features, which provide insights into how changing the values of numeric features impacts the model's predictions.

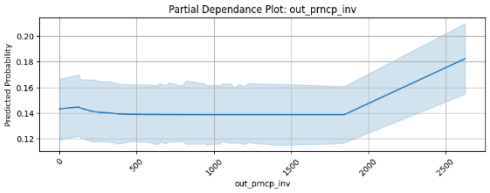
- ➢ Numeric features
  - ✓ Longer loan terms, higher interest rates, higher funds, higher installments, higher inquiries in the last six months, higher total credit revolving balance, and total number of credit lines are associated with a higher likelihood of default.

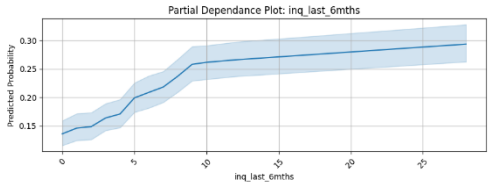


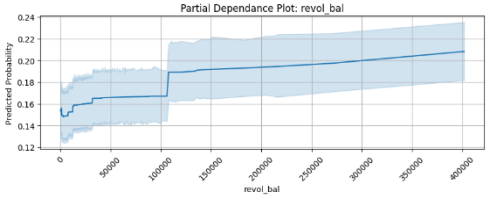


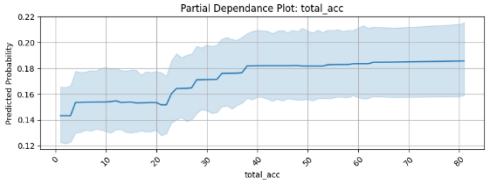




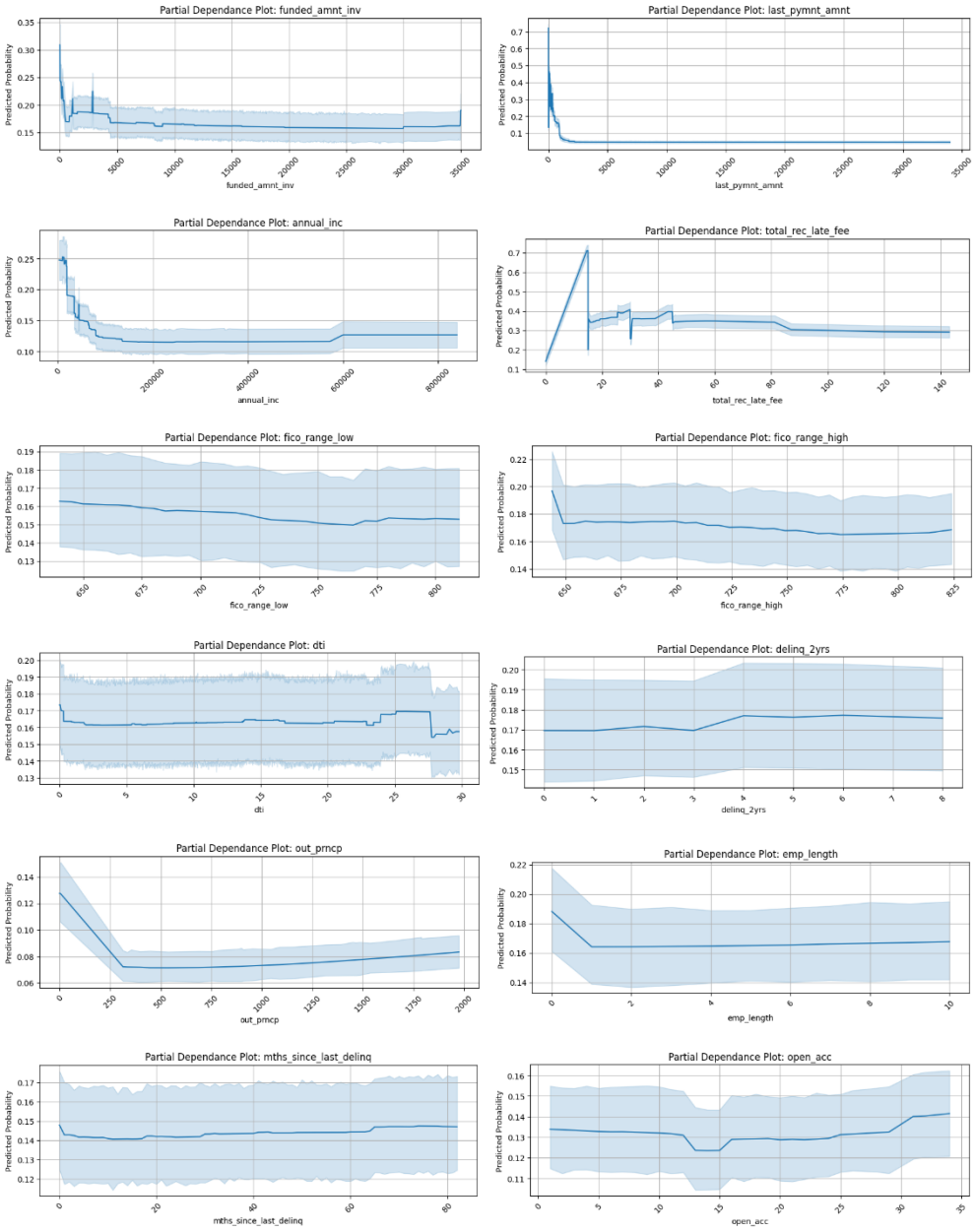


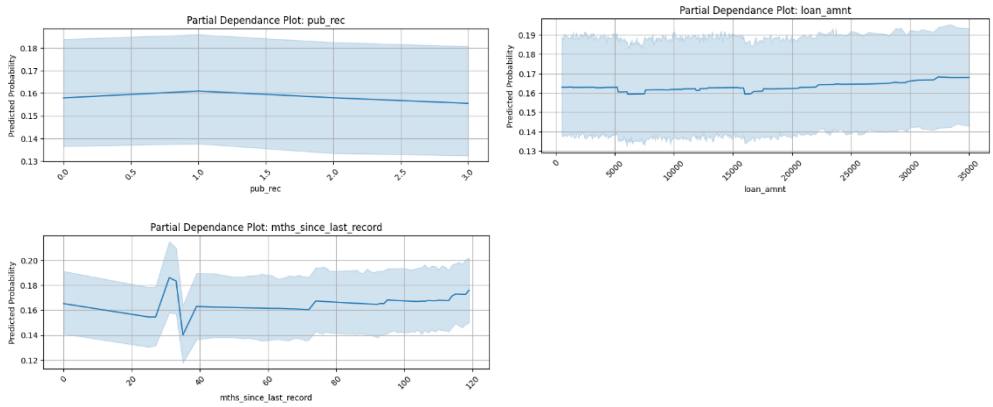






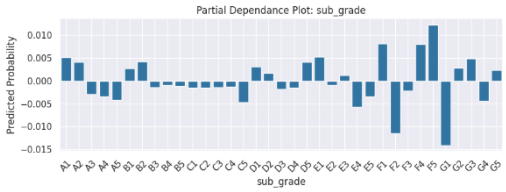
- ✓ When the founder amount interval is less than 5000, it has a higher likelihood of default
- ✓ When the last payment amount is less than 2000, it is more likely to default.
- ✓ When the annual income exceeds 100000, the likelihood of default decreases significantly.
- ✓ As the total rec late fee increases from 0 to 18, it is more likely to default. After 18, the likelihood of default decreases a lot.



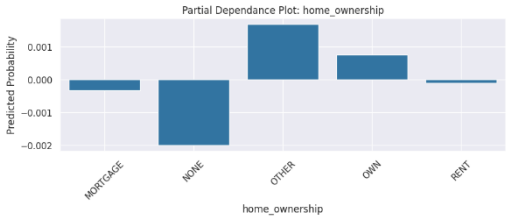


- Categorical features
  - ✓ D, F, G shows positive results, suggesting that grade with DFG is associated with a relatively higher risk of default, according to the model.

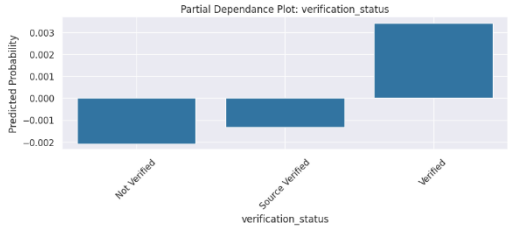
  - ✓ Subgrades with A1,A2,B1,B2,D1,D2,D5,E1,F1,F4,F5,G2,G3,G5 is associated with a relatively higher risk of default.



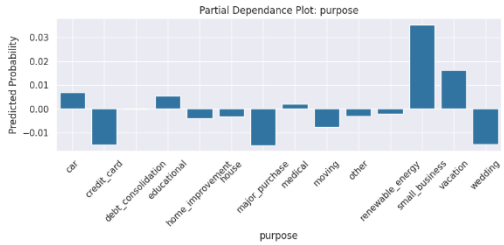
  - ✓ Homeownership type is other and own is associated with a relatively higher risk of default.



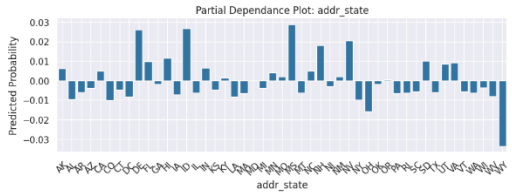
✓ Verified status is associated with a relatively higher risk of default.



✓ The loan purpose of small businesses, vacations, and cars is associated with the top 3 highest default risks.

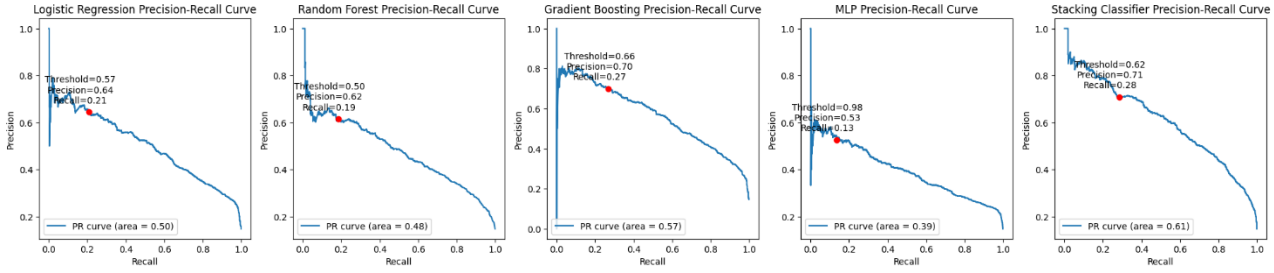


✓ Some states, such as DE, ID, MS, have a higher risk of default than others.

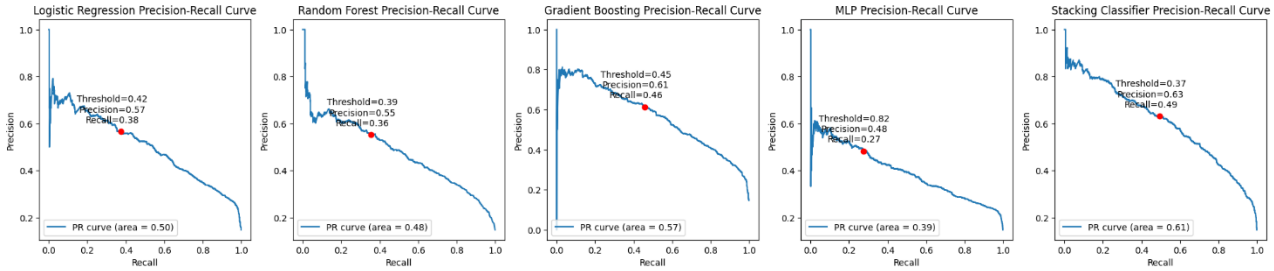


5) Operational Strategy at 2% and 5% FPR

➢ At 2%



- At 5%



- Increase FPR from 2% to 5%, the value of recall increases, the precision decreases, and the threshold decreases.

2%	Recall	Precision	Threshold
Logistic Regression	0.21	0.64	0.57
Random Forest	0.19	0.62	0.5
Gradient Boosting	0.27	0.7	0.66
Neural Network	0.13	0.53	0.98
Stacking Classifier	0.28	0.71	0.62

5%	Recall	Precision	Threshold
Logistic Regression	0.38	0.57	0.42
Random Forest	0.36	0.55	0.39
Gradient Boosting	0.46	0.61	0.45
Neural Network	0.27	0.48	0.82
Stacking Classifier	0.49	0.63	0.37

To be more specific, for the best-performing model, the Stacking Classifier:

- ✓ Recall:

  At FPR 2%, the Recall is 0.28, meaning that the model correctly identifies 28% of all actual positive cases (e.g., loan defaults). This indicates that the model captures a relatively small portion of the true positives.
  At FPR 5%, the Recall increases to 0.49, indicating that the model improves in identifying actual positive cases. The Recall nearly doubles, suggesting that the model becomes more sensitive to detecting loan defaults as the FPR increases.

- ✓ Precision:

  At FPR 2%, the Precision is 0.71, indicating that 71% of the instances predicted as positive are indeed positive (true positives). This suggests that the model has a relatively high precision at this FPR, meaning that when it predicts a default, it's likely to be correct.

At FPR 5%, the Precision decreases to 0.63, meaning that 63% of the instances predicted as positive are actually positive. Although the Precision decreases, it remains relatively high, indicating that the model still maintains a reasonable level of accuracy in its positive predictions.

- ✓ Threshold:

At FPR 2%, the Threshold is 0.62, which represents the probability threshold used by the model to classify instances as positive or negative. This indicates that the model sets a relatively high threshold for classifying instances as positive, resulting in a higher Precision but lower Recall.

At FPR 5%, the Threshold decreases significantly to 0.37, indicating that the model becomes more lenient in classifying instances as positive. This leads to an increase in Recall but a decrease in Precision, as the model becomes more aggressive in labeling instances as positive.

- ✓ Implications:

At a 2% False Positive Rate (FPR), the model is cautious and aims to make fewer mistakes by focusing on accurately predicting loan defaults. It might miss some defaults (lower Recall), but it's usually correct when it does expect a default (higher Precision). This cautious approach is suitable for situations where avoiding unnecessary alarms is crucial, like in conservative risk management practices.

At a 5% False Positive Rate (FPR), the model becomes more liberal and tries to catch as many loan defaults as possible, even if it means making some mistakes (lower Precision). It's more sensitive and captures more defaults (higher Recall) but may also raise more false alarms. This approach is practical when the priority is to identify all potential loan defaults, even if it means tolerating more false alarms, to ensure that no significant risks are overlooked.

4. Local Explanations

Used pipeline_explainer to compute the SHAP values for each feature, providing insights into how each feature contributes to the model's prediction for that instance. Features contributed positively or negatively to the prediction and to what extent.

1) TP-top 10 true positive, loan default = 1, predictive loan = 1

The consistent appearance of the "last payment amount" feature in the top 10 true positive predictions suggests that this feature plays a crucial role in the model's ability to correctly classify instances as positive (in this case, as loan defaults). The value of the last payment amount is less than 220.

For example, for the first record (record 0) in the top True Positive predictions, the model's predicted probability is 0.971. Additionally, the top three features contributing to this prediction, along with their SHAP values, are as follows:

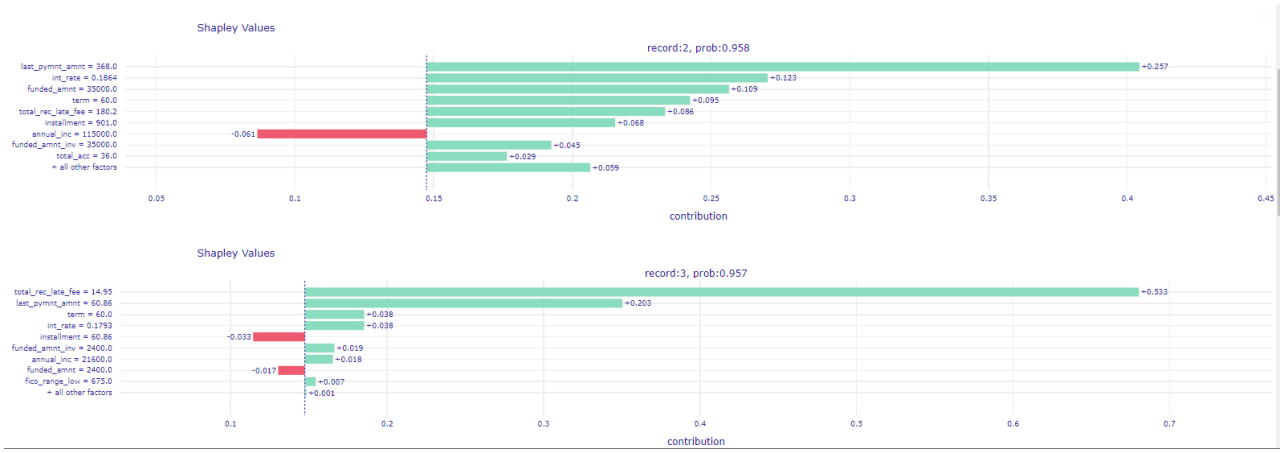
- Last Payment Amount: SHAP Value: +0.417

This indicates that a higher last payment amount positively contributes to the model's prediction of a positive outcome (loan default in this case). According to the model, a larger last payment amount is associated with a higher likelihood of default.

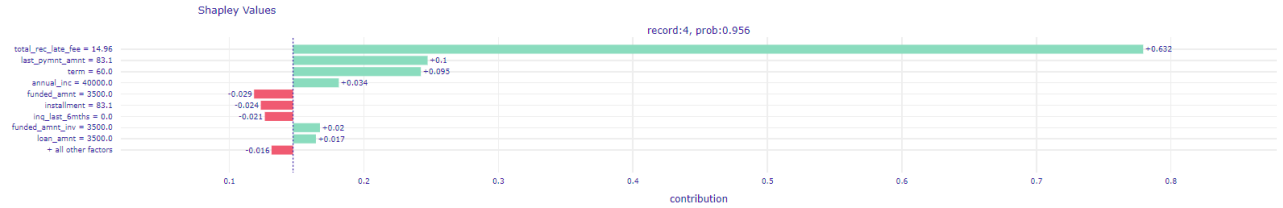
Interest Rate:

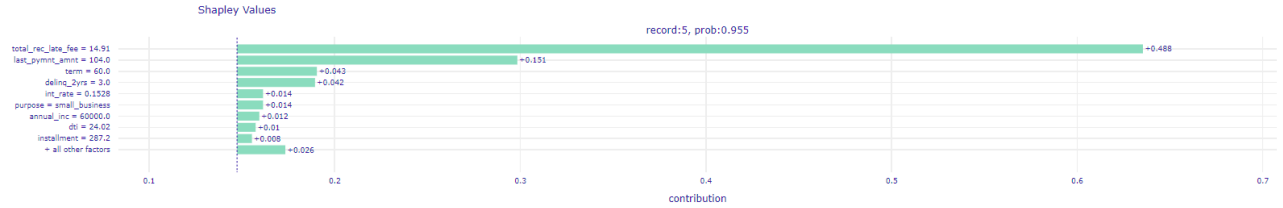
- Interest rate: SHAP Value: +0.180

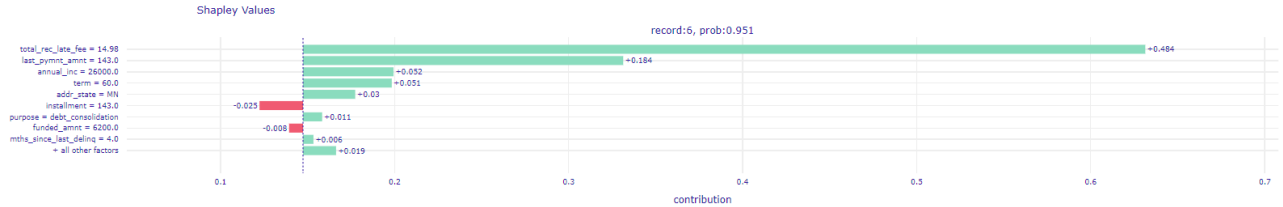
A higher interest rate also positively contributes to the model's prediction of a positive outcome. This suggests that loans with higher interest rates are more likely to be classified as defaults by the model.

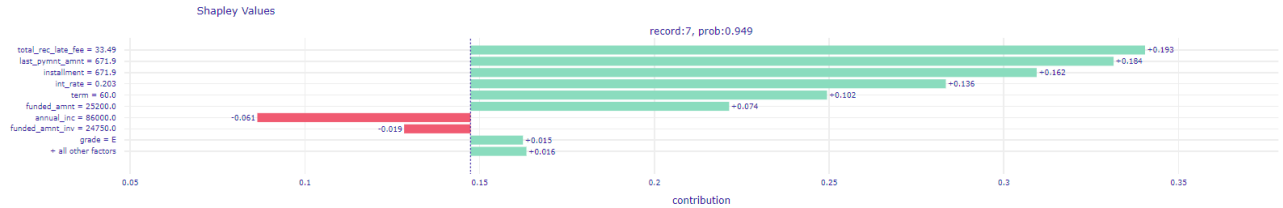

Shapley Values
record:2, prob:0.958
last_pymnt_amnt = 368.0
int_rate = 0.1864
funded_amnt = 35000.0
term = 60.0
total_rec_late_fee = 180.2
installment = 901.0
annual_inc = 115000.0
funded_amnt_inv = 35000.0
total_acc = 36.0
+ all other factors
+0.257
+0.123
+0.109
+0.095
+0.086
+0.068
-0.061
+0.045
+0.029
+0.059
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
contribution

Shapley Values
record:3, prob:0.957
total_rec_late_fee = 14.95
last_pymnt_amnt = 60.86
term = 60.0
int_rate = 0.1793
installment = 60.86
funded_amnt_inv = 2400.0
annual_inc = 21600.0
funded_amnt = 2400.0
fico_range_low = 675.0
+ all other factors
+0.533
+0.203
+0.038
+0.038
-0.033
+0.019
+0.018
-0.017
+0.007
+0.001
0.1
0.2
0.3
0.4
0.5
0.6
0.7
contribution


Shapley Values
record:4, prob:0.956
total_rec_late_fee = 14.96
last_pymnt_amnt = 83.1
term = 60.0
annual_inc = 40000.0
funded_amnt = 3500.0
installment = 83.1
inq_last_6mths = 0.0
funded_amnt_inv = 3500.0
loan_amnt = 3500.0
+ all other factors
+0.632
+0.1
+0.095
+0.034
-0.029
-0.024
-0.021
+0.02
+0.017
-0.016
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
contribution


Shapley Values
record:5, prob:0.955
total_rec_late_fee = 14.91
last_pymnt_amnt = 104.0
term = 60.0
delinq_2yrs = 3.0
int_rate = 0.1528
purpose = small_business
annual_inc = 60000.0
dti = 24.02
installment = 287.2
+ all other factors
+0.488
+0.151
+0.043
+0.042
+0.014
+0.014
+0.012
+0.01
+0.008
+0.026
0.1
0.2
0.3
0.4
0.5
0.6
0.7
contribution


Shapley Values
record:6, prob:0.951
total_rec_late_fee = 14.98
last_pymnt_amnt = 143.0
annual_inc = 26000.0
term = 60.0
addr_state = MN
installment = 143.0
purpose = debt_consolidation
funded_amnt = 6200.0
mths_since_last_delinq = 4.0
+ all other factors
+0.484
+0.184
+0.052
+0.051
+0.03
-0.025
+0.011
-0.008
+0.006
+0.019
0.1
0.2
0.3
0.4
0.5
0.6
0.7
contribution


Shapley Values
record:7, prob:0.949
total_rec_late_fee = 33.49
last_pymnt_amnt = 671.9
installment = 671.9
int_rate = 0.203
term = 60.0
funded_amnt = 25200.0
annual_inc = 86000.0
funded_amnt_inv = 24750.0
grade = E
+ all other factors
+0.193
+0.184
+0.162
+0.136
+0.102
+0.074
-0.061
-0.019
+0.015
+0.016
0.05
0.1
0.15
0.2
0.25
0.3
0.35
contribution

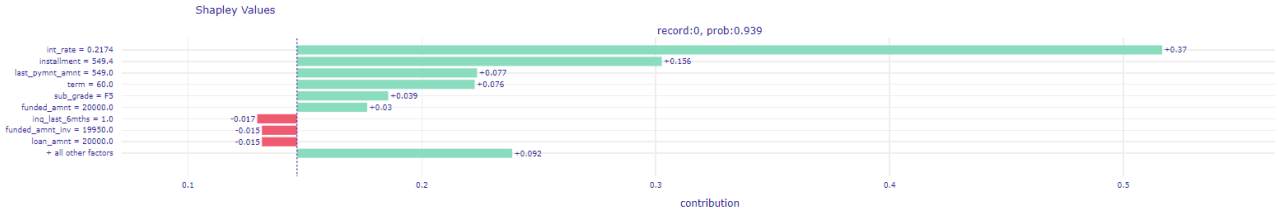
2) TP-top 10 false positive, loan default =0, predictive loan = 1

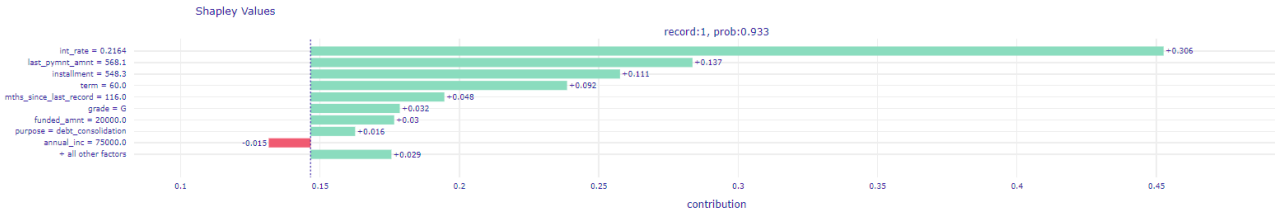
- ✓ The consistent appearance of the "interest rate" feature in the top 10 false positive predictions suggests that this feature plays a significant role in the model's misclassification of instances as positive (loan defaults) when they are actually negative (non-defaults).In summary, the repeated presence of the "interest rate" feature in the top false positive predictions indicates its significance in influencing the model's misclassifications.

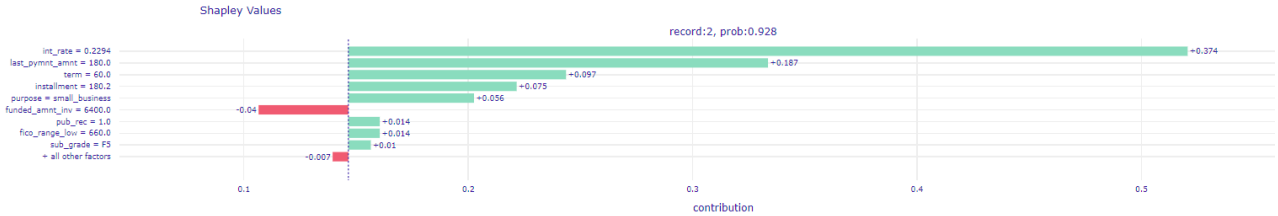
- ✓ The potential reason for FP:

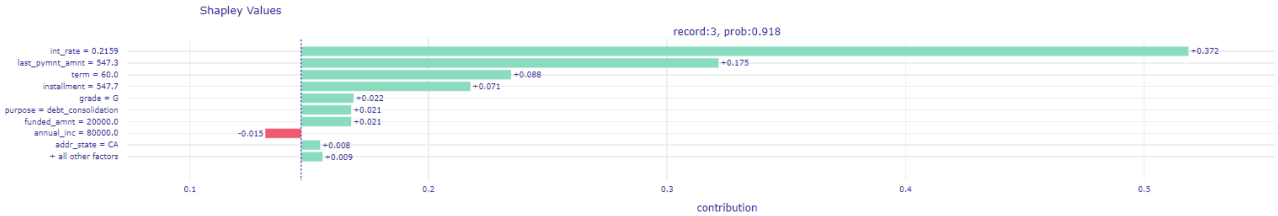
  1) While high-interest rates may intuitively signal higher default risk, the model might be overly reliant on this feature and incorrectly classify instances with high-interest rates as positive (loan defaults) when they are actually negative (non-defaults). This could be due to a lack of balance in the model's consideration of other relevant features.

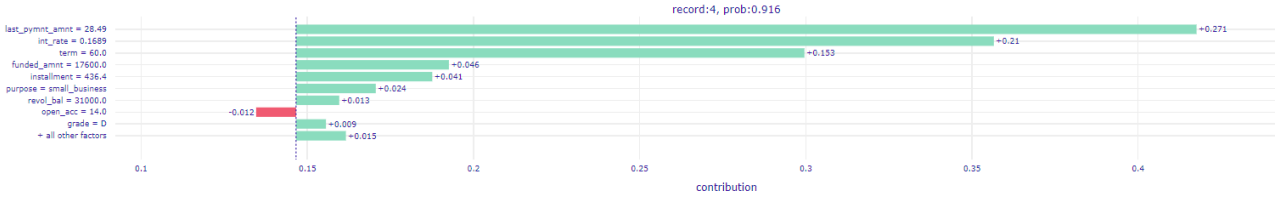
  2) Data Imbalance: There is a significant imbalance in the dataset with fewer positive instances (loan defaults) compared to negative instances (non-defaults), the model may struggle to distinguish between the two classes effectively. In such cases, features like "interest rate" might disproportionately influence predictions, leading to false positives

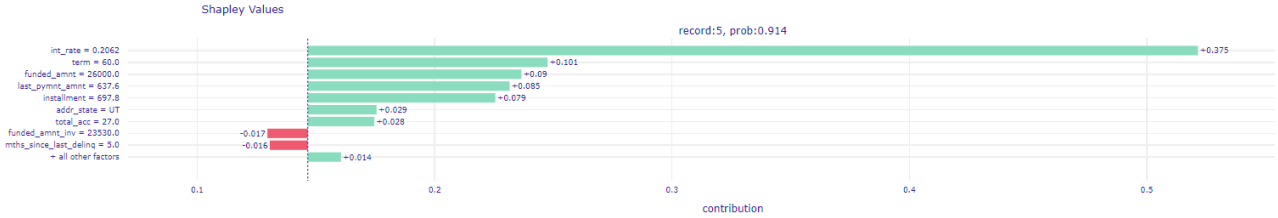

Shapley Values
record:0, prob:0.939
int_rate = 0.2174
installment = 549.4
last_pymnt_amnt = 549.0
term = 60.0
sub_grade = F5
funded_amnt = 20000.0
inq_last_6mths = 1.0
funded_amnt_inv = 19950.0
loan_amnt = 20000.0
+ all other factors
+0.37
+0.156
+0.077
+0.076
+0.039
+0.03
-0.017
-0.015
-0.015
+0.092
0.1
0.2
0.3
0.4
0.5
contribution

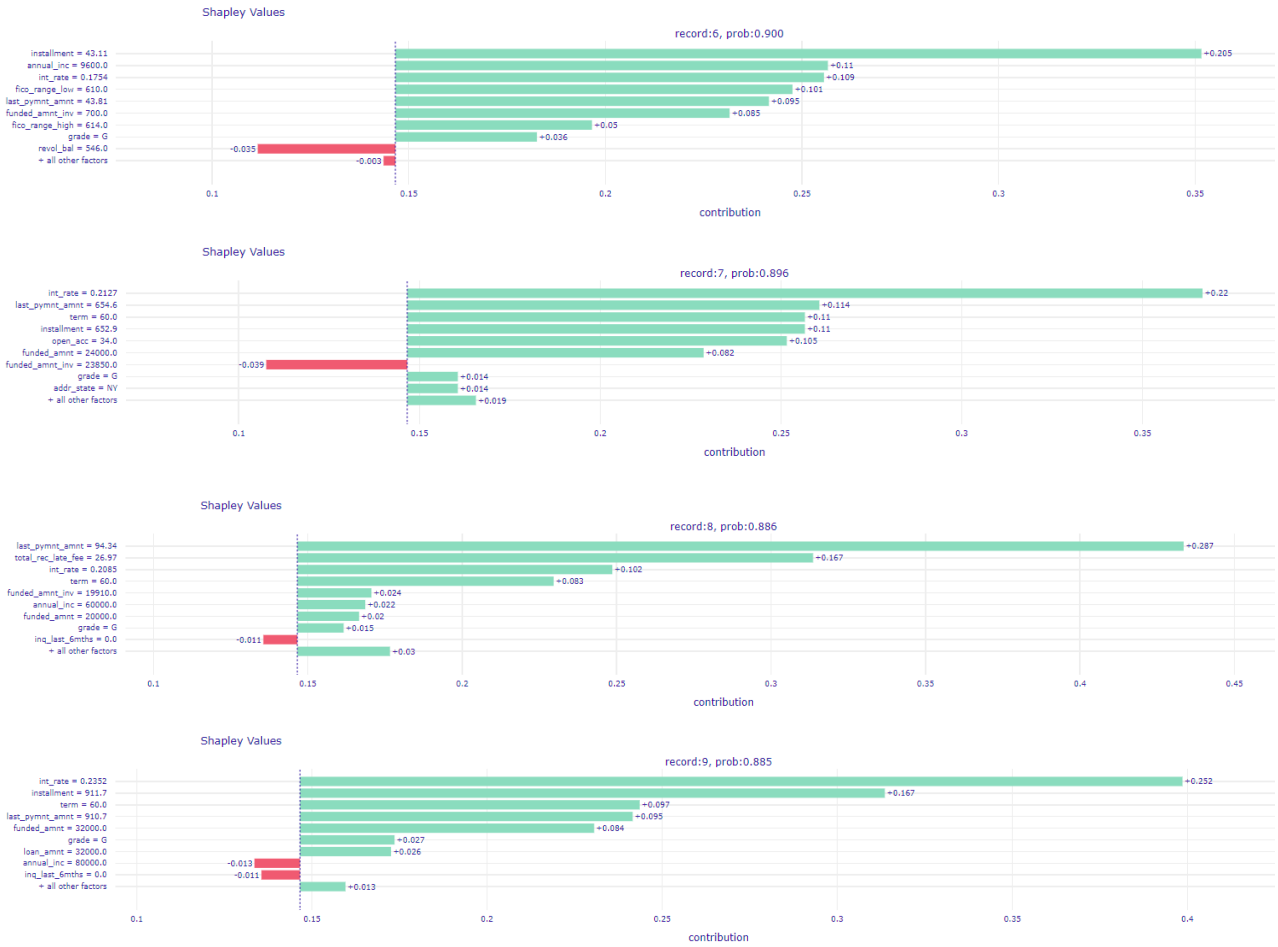

Shapley Values
record:1, prob:0.933
int_rate = 0.2164
last_pymnt_amnt = 568.1
installment = 548.3
term = 60.0
mths_since_last_record = 116.0
grade = G
funded_amnt = 20000.0
purpose = debt_consolidation
annual_inc = 75000.0
+ all other factors
+0.306
+0.137
+0.111
+0.092
+0.048
+0.032
+0.03
+0.016
-0.015
+0.029
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
contribution


Shapley Values
record:2, prob:0.928
int_rate = 0.2294
last_pymnt_amnt = 180.0
term = 60.0
installment = 180.2
purpose = small_business
funded_amnt_inv = 6400.0
pub_rec = 1.0
fico_range_low = 660.0
sub_grade = F5
+ all other factors
+0.374
+0.187
+0.097
+0.075
+0.056
-0.04
+0.014
+0.014
+0.01
-0.007
0.1
0.2
0.3
0.4
0.5
contribution


Shapley Values
record:3, prob:0.918
int_rate = 0.2159
last_pymnt_amnt = 547.3
term = 60.0
installment = 547.7
grade = G
purpose = debt_consolidation
funded_amnt = 20000.0
annual_inc = 80000.0
addr_state = CA
+ all other factors
+0.372
+0.175
+0.088
+0.071
+0.022
+0.021
+0.021
-0.015
+0.008
+0.009
0.1
0.2
0.3
0.4
0.5
contribution


record:4, prob:0.916
last_pymnt_amnt = 28.49
int_rate = 0.1689
term = 60.0
funded_amnt = 17600.0
installment = 436.4
purpose = small_business
revol_bal = 31000.0
open_acc = 14.0
grade = D
+ all other factors
+0.271
+0.21
+0.153
+0.046
+0.041
+0.024
+0.013
-0.012
+0.009
+0.015
0.1
0.15
0.2
0.25
0.3
0.35
0.4
contribution


Shapley Values
record:5, prob:0.914
int_rate = 0.2062
term = 60.0
funded_amnt = 26000.0
last_pymnt_amnt = 637.6
installment = 697.8
addr_state = UT
total_acc = 27.0
funded_amnt_inv = 23530.0
mths_since_last_delinq = 5.0
+ all other factors
+0.375
+0.101
+0.09
+0.085
+0.079
+0.029
+0.028
-0.017
-0.016
+0.014
0.1
0.2
0.3
0.4
0.5
contribution



3) TP-top 10 false negative, loan default = 1, predictive loan = 0

The consistent appearance of the "last payment amount" feature in the top 10 false negative predictions, alongside consistently negative scores, suggests a notable pattern that warrants further investigation.

The prevalence of large last payment amounts among the false negative predictions indicates that the model may be misinterpreting these large values. While intuitively, large last payments might suggest financial stability and a lower risk of default, in these instances, they might be misleading the model into predicting non-default (negative) outcomes incorrectly.


Shapley Values
record:0, prob:0.031
last_pymnt_amnt = 2300.0
installment = 212.3
emp_length = 0.0
revol_bal = 21680.0
inq_last_6mths = 3.0
grade = A
total_rec_late_fee = 0.0
purpose = major_purchase
addr_state = CO
+ all other factors
-0.109
-0.008
+0.007
+0.006
+0.005
-0.005
-0.004
-0.003
-0.002
-0.004
0.04
0.06
0.08
0.1
0.12
0.14
0.16
contribution
Shapley Values
record:1, prob:0.034
last_pymnt_amnt = 2000.0
total_rec_late_fee = 40.9
int_rate = 0.0938
installment = 425.3
purpose = small_business
addr_state = OH
revol_bal = 126300.0
grade = A
annual_inc = 90000.0
+ all other factors
-0.158
+0.105
-0.051
+0.036
+0.03
-0.022
+0.014
-0.013
-0.013
-0.041
−0.05
0
0.05
0.1
0.15
0.2
0.25
contribution
record:2, prob:0.035
last_pymnt_amnt = 1764.0
installment = 164.9
term = 36.0
int_rate = 0.1149
total_rec_late_fee = 0.0
annual_inc = 37000.0
purpose = credit_card
inq_last_6mths = 0.0
funded_amnt_inv = 5000.0
+ all other factors
-0.076
-0.017
-0.011
-0.009
-0.008
+0.007
-0.005
-0.004
+0.003
+0.009
0.06
0.08
0.1
0.12
0.14
0.16
contribution
Shapley Values
record:3, prob:0.037
last_pymnt_amnt = 4842.0
term = 60.0
int_rate = 0.0962
funded_amnt = 20000.0
annual_inc = 96000.0
loan_amnt = 20000.0
installment = 421.2
total_rec_late_fee = 0.0
addr_state = NY
+ all other factors
-0.127
+0.056
-0.024
+0.02
-0.02
-0.019
+0.017
-0.012
-0.006
+0.005
0
0.05
0.1
0.15
0.2
contribution
record:4, prob:0.038
last_pymnt_amnt = 11800.0
annual_inc = 115000.0
funded_amnt = 35000.0
term = 60.0
int_rate = 0.164
installment = 858.6
purpose = credit_card
loan_amnt = 35000.0
fico_range_low = 745.0
+ all other factors
-0.142
-0.041
+0.037
+0.028
+0.022
+0.019
-0.016
-0.009
-0.009
+0.002
0
0.05
0.1
0.15
0.2
contribution
Shapley Values
record:5, prob:0.039
last_pymnt_amnt = 181.0
int_rate = 0.0542
loan_amnt = 6000.0
funded_amnt_inv = 5950.0
term = 36.0
grade = A
installment = 181.0
inq_last_6mths = 0.0
total_rec_late_fee = 0.0
+ all other factors
-0.069
+0.093
+0.045
-0.039
-0.036
-0.025
-0.02
-0.015
-0.014
-0.029
0.1
0.15
0.2
0.25
contribution


Shapley Values
record:6, prob:0.040
int_rate = 0.0617
installment = 701.5
grade = A
annual_inc = 98000.0
total_rec_late_fee = 0.0
term = 36.0
loan_amnt = 23000.0
funded_amnt = 23000.0
last_pymnt_amnt = 701.5
+ all other factors
-0.044
-0.018
-0.017
-0.017
-0.013
-0.011
-0.006
-0.019
+0.031
+0.007
0.1
0.11
0.12
0.13
0.14
0.15
0.16
0.17
0.18
contribution
Shapley Values
record:7, prob:0.040
installment = 219.2
last_pymnt_amnt = 274.4
int_rate = 0.0699
purpose = major_purchase
loan_amnt = 7100.0
grade = A
term = 36.0
total_rec_late_fee = 0.0
inq_last_6mths = 0.0
+ all other factors
-0.075
-0.047
-0.028
-0.022
-0.022
-0.016
+0.057
+0.027
+0.012
+0.007
0.06
0.08
0.1
0.12
0.14
0.16
0.18
0.2
contribution
Shapley Values
record:8, prob:0.041
last_pymnt_amnt = 10020.0
installment = 385.5
funded_amnt = 16000.0
loan_amnt = 16000.0
term = 60.0
annual_inc = 78000.0
grade = D
verification_status = Verified
total_rec_late_fee = 0.0
+ all other factors
-0.19
-0.016
-0.01
-0.006
+0.048
+0.035
+0.013
+0.006
+0.006
+0.007
−0.05
0
0.05
0.1
0.15
0.2
contribution
Shapley Values
record:9, prob:0.042
last_pymnt_amnt = 1370.0
total_rec_late_fee = 0.0
funded_amnt = 20680.0
annual_inc = 75000.0
addr_state = NV
revol_bal = 50670.0
int_rate = 0.1349
funded_amnt_inv = 19010.0
term = 60.0
+ all other factors
-0.095
-0.013
-0.009
-0.004
-0.003
-0.003
+0.01
+0.004
+0.003
+0.005
0.04
0.06
0.08
0.1
0.12
0.14
0.16
contribution