POC for Bank of China Using Financial Crime Alerts Insight with Watson

FOR INTERNAL USE ONLY

IBM Watson Financial Services

Supervised Machine Learning Analysis

- **Method – Random Forests**
  - Decision Trees based
  - Ensemble supervised learning
  - Predict likelihood of an alert being escalated

- **Process**
  - Preparing Input Data
  - Generating Model
  - Testing Model
  - Validating Performance

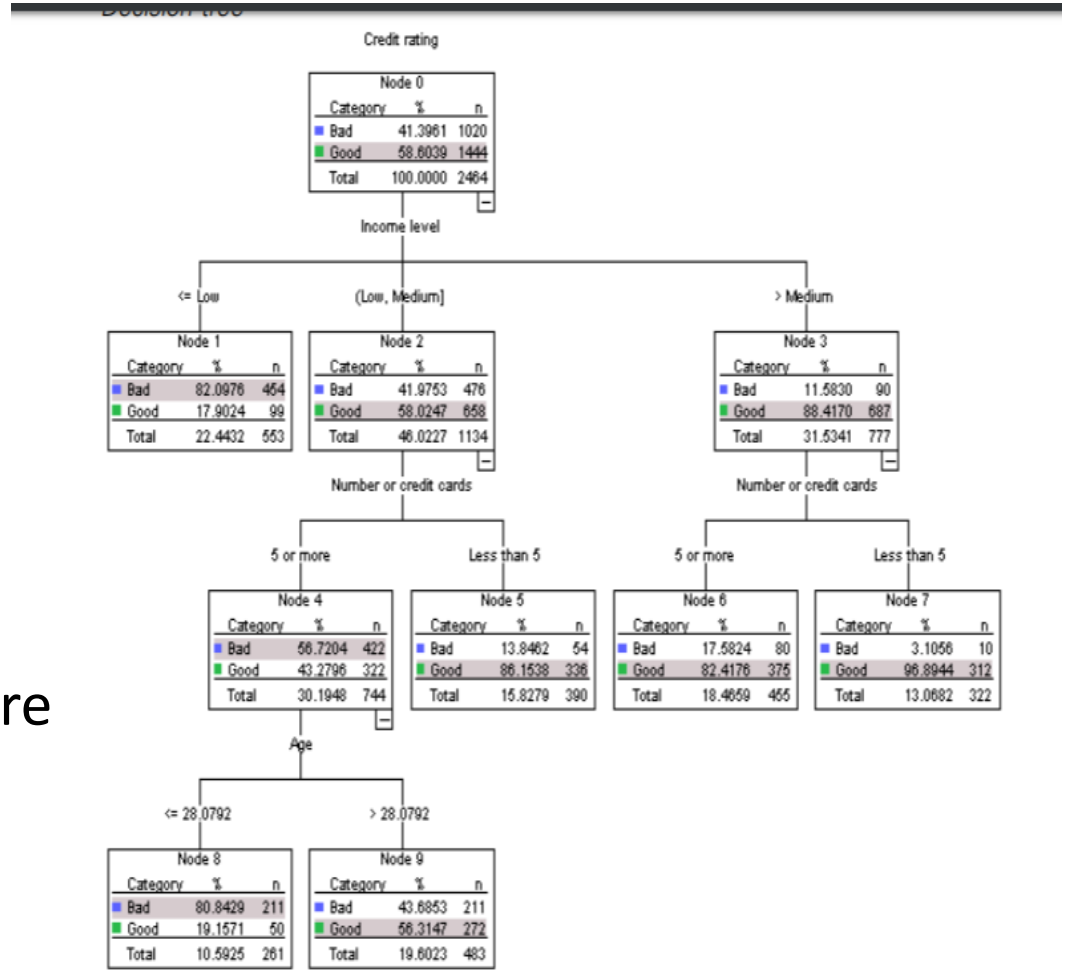
ML – Random Forests Method

❑Decision Trees

- Less sensitive to outliers
- Normalization is not required
- Capture non-linear relationship

❑Ensemble Learning

- Allow automatic feature selection
- Handle missing data
- Apply bagging and cross-validation



ML - Preparing Input Data

- **Create features for the selected aggregator (account)**
  - Given the input data files specified by FCAI data requirements, the FCAI calculate a set of derived statistics to capture transactional behavior for each account
  - Additional features for each account can be added as needed
  - Additional aggregators can be added

- **Generate target variable for the selected aggregator (account)**
  - For each alert, the disposition information is extracted.
  - Based on the subject account for each alert, generate a mapping target variable (SAR or NonSAR) for each account

ML - Generating Model

- **Splitting the input data into training and validation sets**
  - Using different accounts to ensure blind testing
  - The default setting is 80/20 split
  - The split will be based on the target variable distribution so that the percentage of SAR and nonSAR within both data sets stays the same

- **Generate model using training data set**
  - Since the percentage of SAR is very small, over/under sampling of training data set is applied
  - The default supervised ML model in FCAI is Random Forrest
  - The model will be saved in the supervisedML_FPTP.rda file

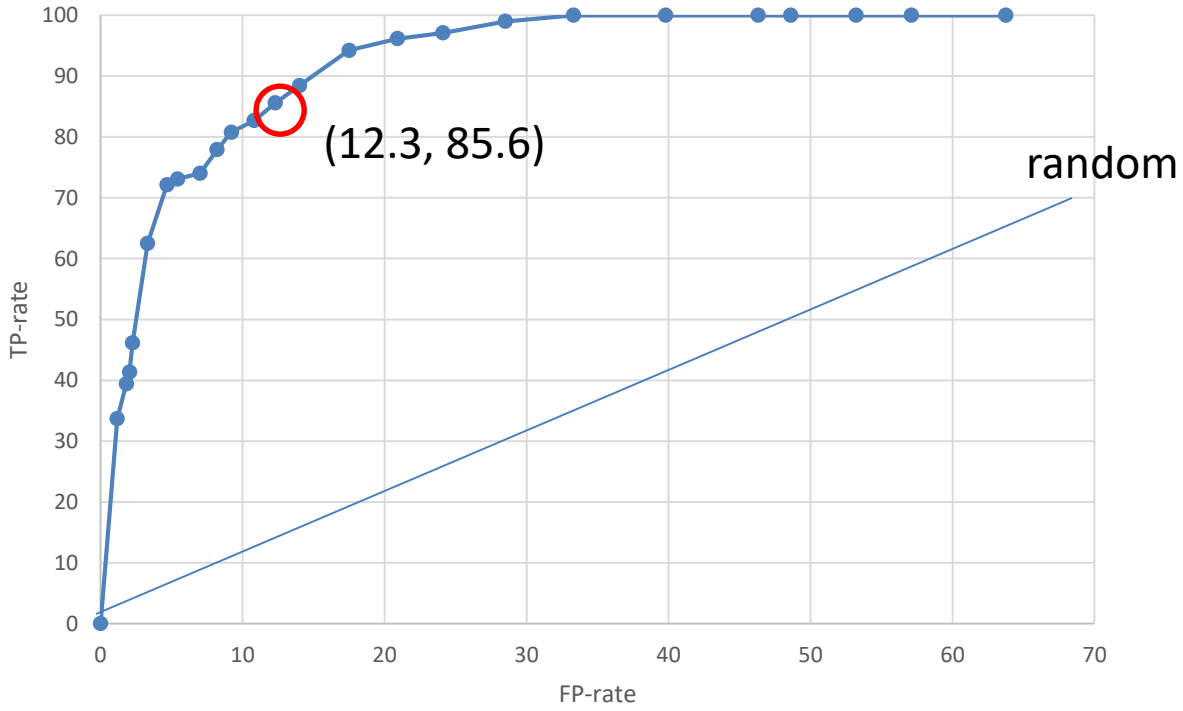
ML - Testing Model

- **Predict SAR/NonSAR for the testing data set**
  - Using the model created based on training data set, generate SAR/nonSAR prediction for each record in the testing data set
  - The output includes prediction probability which can be used as threshold for adjusting tradeoff between False Positive and True Positive rate

- **Evaluate performance for the testing data set**
  - Provide confusion matrix showing the summarized performance
  - Generate ROC curve for selecting optimal setting
  - Extract importance features for overall contribution
  - Explain the output through supporting evidence

ML - Validating Performance

- **Evaluating Model Performance Using FP-rate and TP-rate**
  - **ROC (Receiver Operating Characteristic) Curve**



TP-rate = True Positive rate (model correctly predict SAR)
FP-rate = False Positive (model incorrectly predict SAR when it should be nonSAR)

ML - Validating Performance

❑ Account Level Prediction Results

- Confusion Matrix for validation data with 0.5 cutoff

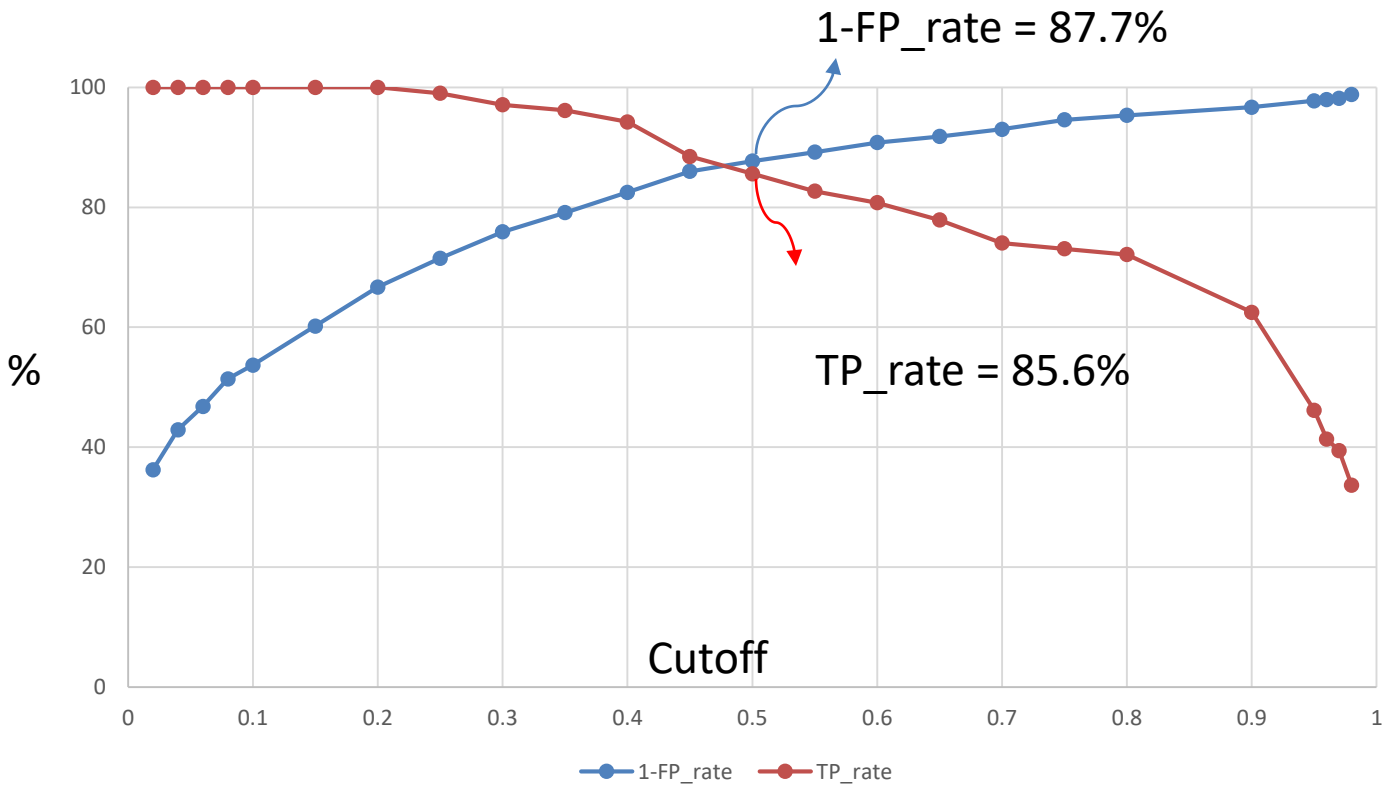
	Ground False	Ground Truth	
	nonSAR	**SAR**	**subtotal**
Predict_nonSAR	**12,039 (TN)**	**15 (FN)**	**12,054 (87%)**
Predict_SAR	**1,691 (FP)**	**89 (TP)**	**1,780(12.9%)**
subtotal	**13,730 (99%)**	**104 (1%)**	**13,834**

Ground False = non-escalated alert/no SAR filed
Ground Truth = escalated alerts/SAR filed
TP = True Positive = model correctly predict SAR
FP = False Positive = model incorrectly predict SAR when it should be nonSAR
TN = True Negative = model correctly predict nonSAR
FN = False Negative = model incorrectly predict nonSAR when it should be SAR

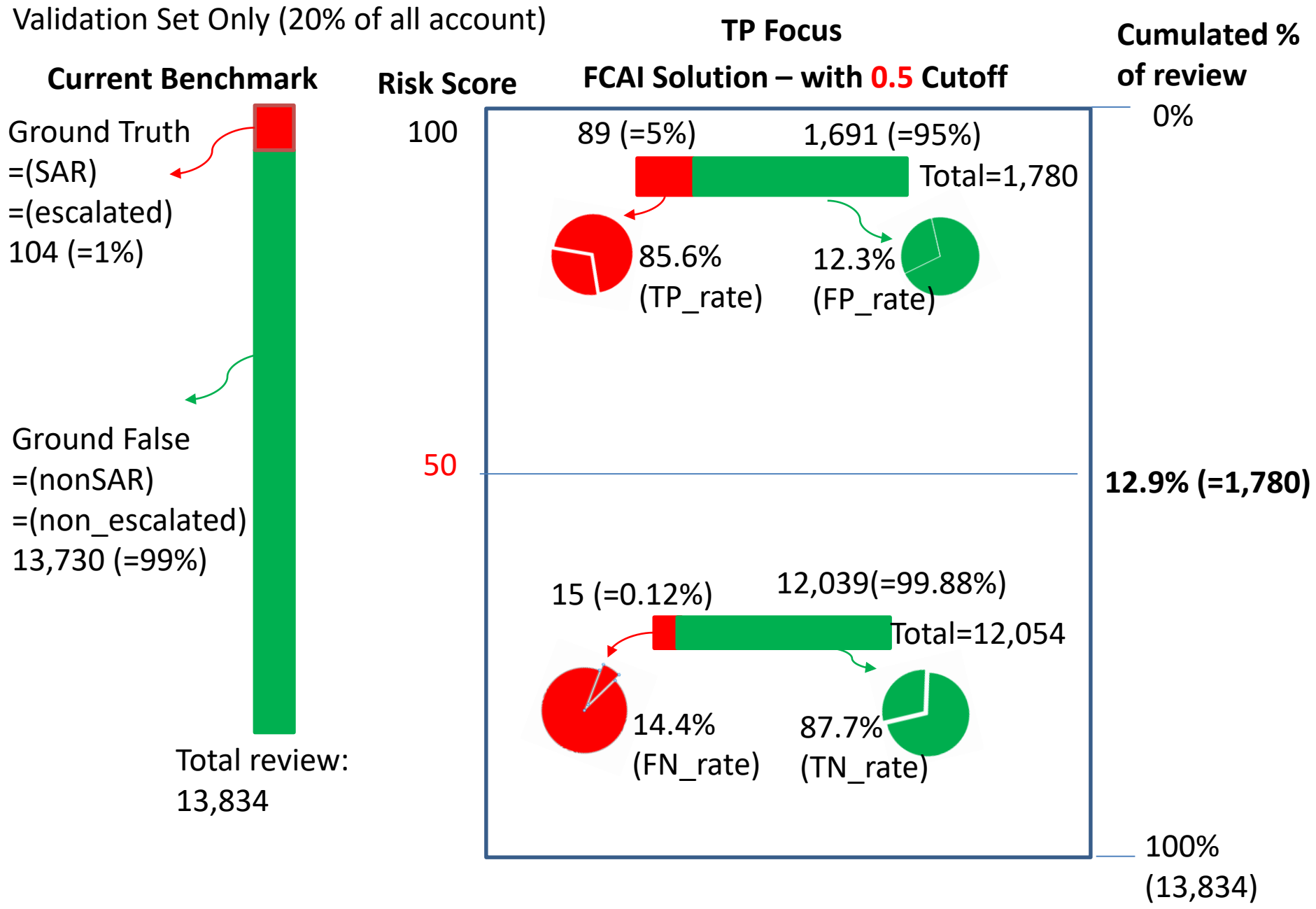
Identify 85.6% (89/104) of the SAR with 12.3% (1,691/13,730) of the FP, by reviewing top 12.9% (1,780/13,834) of the cases

ML - Validating Performance

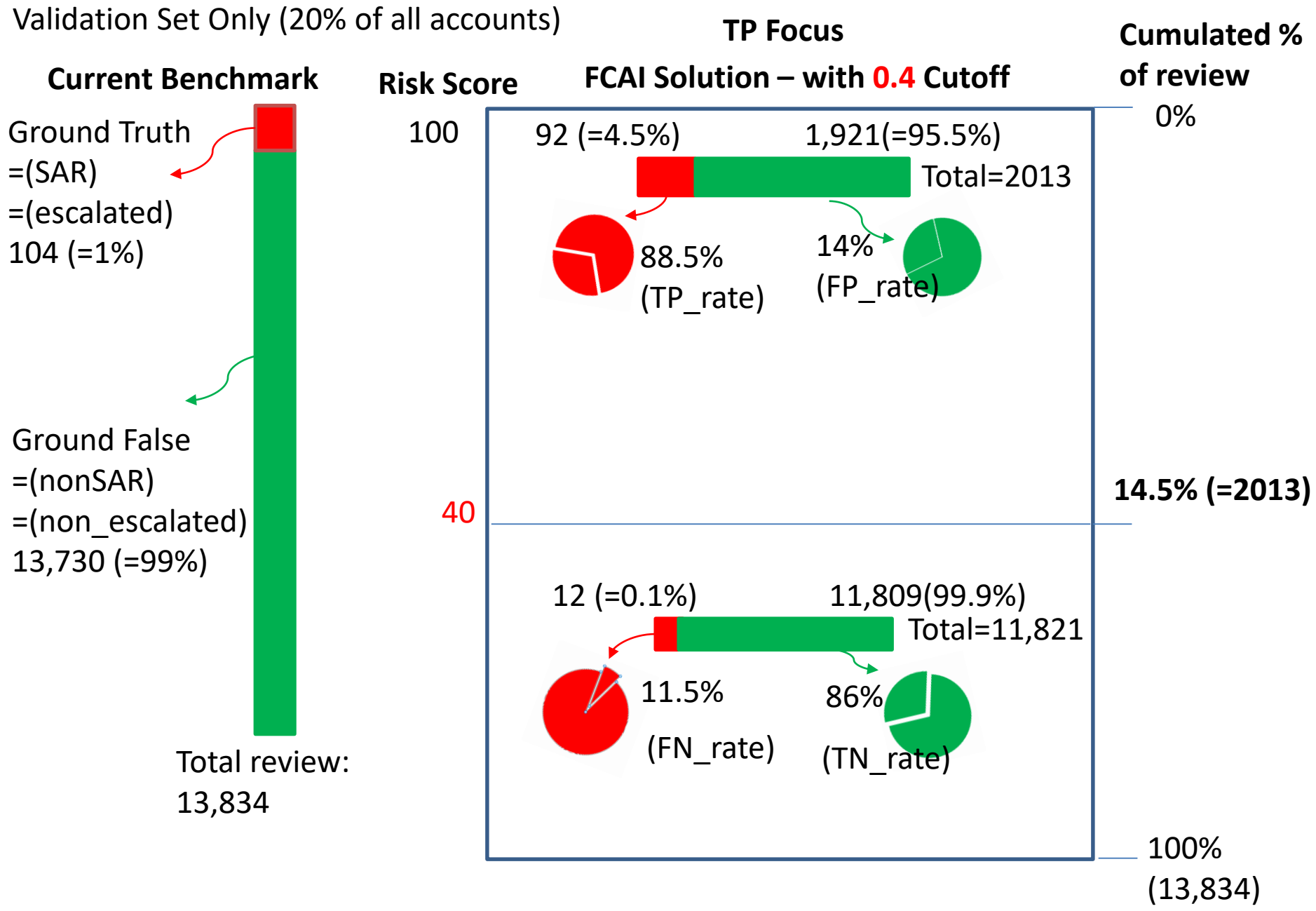
- **Accuracy Plot (prediction that are correct)**
  - **TP-rate and TN-rate (=1-FP-rate) vs Scoring Cutoff**



ML Performance Landscape – Account Level



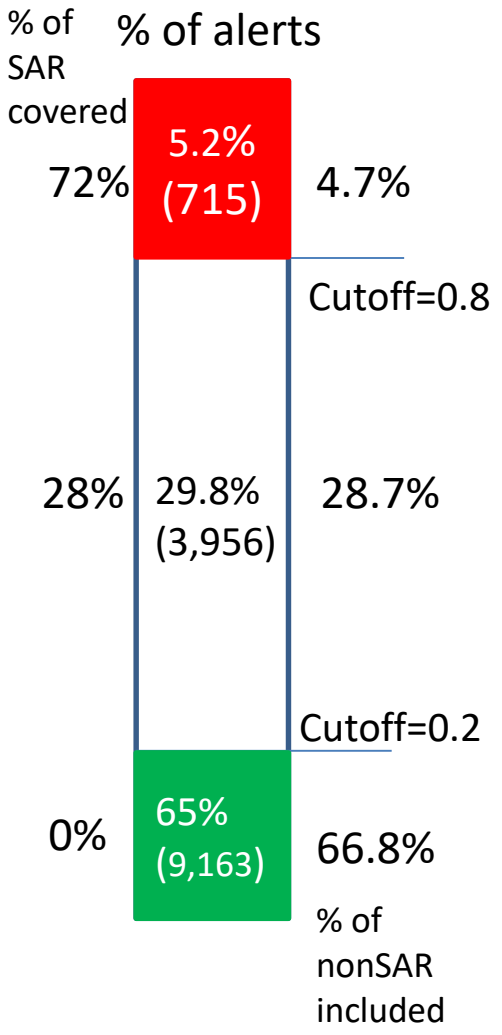
ML Performance Landscape – Account Level



ML Performance Landscape – Account Level

Validation Set Only (20% of all accounts)

Overall SAR/NonSAR ratio = 104:13,730 = 1:132



# of SAR covered	% of SAR covered	# of nonSAR covered	% of nonSAR covered	SAR/Non SAR ratio	Cumulated % of Review
75	72%	640	4.7%	1:8.5	0% – 5.2%
14	13.5%	1,051	7.7%	1:75.1	5.2% – 12.9%
9	8.7%	713	5.2%	1:79.2	12.9% – 18.1%
6	5.8%	2,163	15.8%	1:360	18.1% – 34%
0	0%	9,163	66.8%	NA	34% – 100%

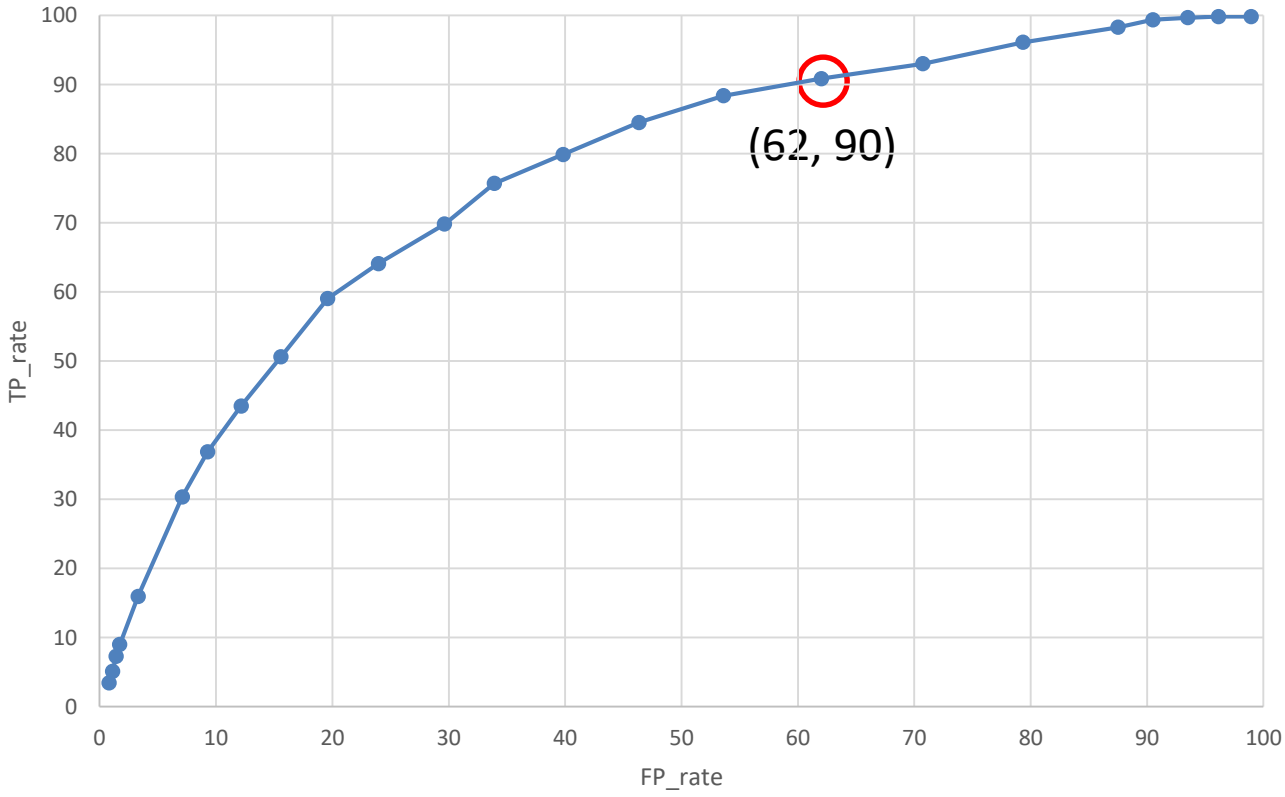
(13,834)

Ensemble Model for Alert Level Score

- **Method –**
  - Random Forests
  - Simple Average
  - Weighted Average

Ensemble - Validating Performance

- **Evaluating Model Performance Using FP-rate and TP-rate**
  - **ROC (Receiver Operating Characteristic) Curve**



Alert - Validating Performance

❑ Alert Level Prediction Results

- Method : ensemble with random forest
- Confusion Matrix with 0.25 cutoff
- Validation Data only

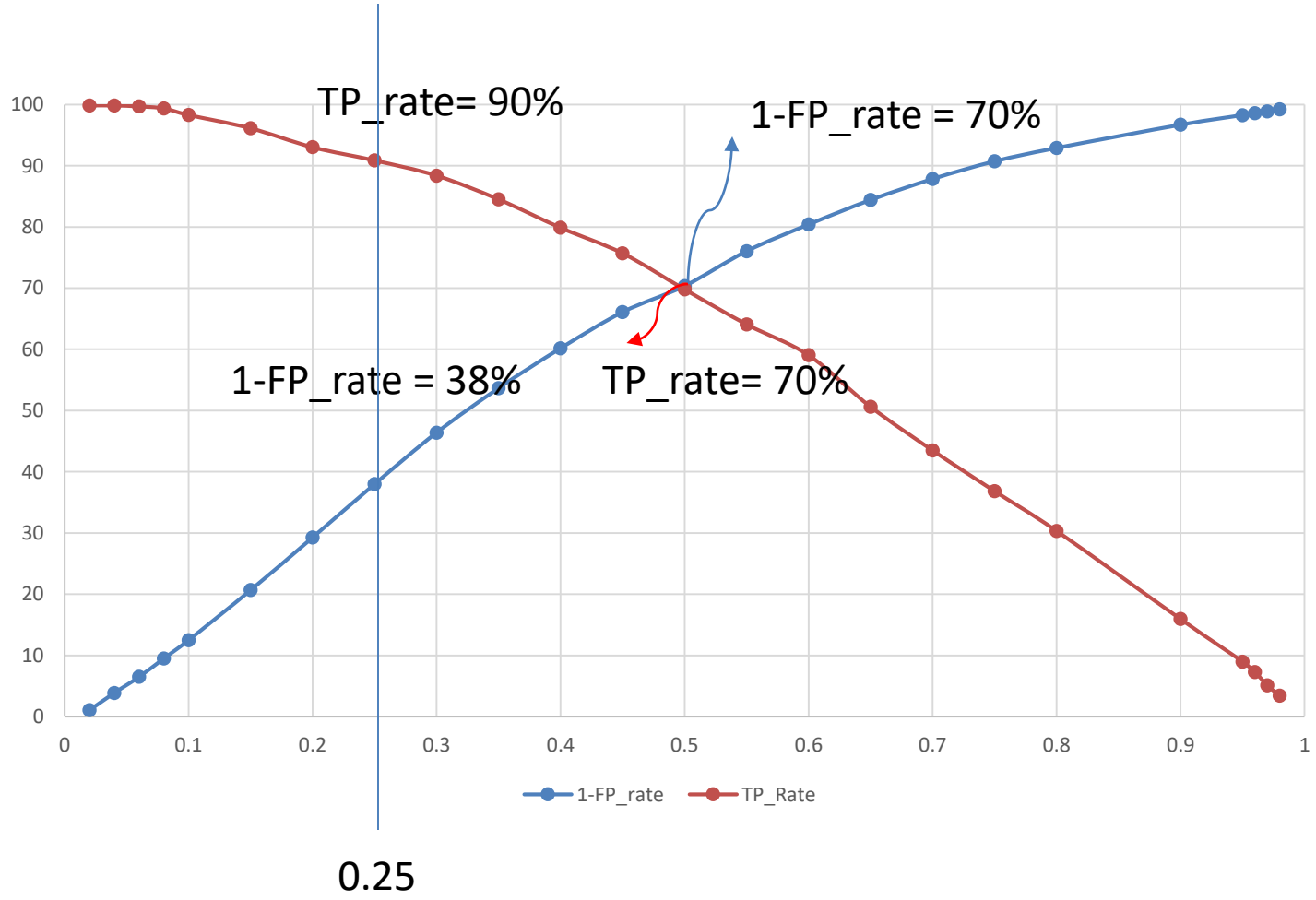
Ground False Ground Truth

	nonSAR	SAR	subtotal
Predict_nonSAR	**7,232(TN)**	**59 (FN)**	**7,291(37%)**
Predict_SAR	**11,807 (FP)**	**587 (TP)**	**12,394(63%)**
subtotal	**19,039 (96.7%)**	**646 (3.3%)**	**19,685**

Identify 90.1% (587/646) of the TP, with 62% (11,807/19,039) of the FP, by reviewing top 63% (12,394/19,685) of the cases

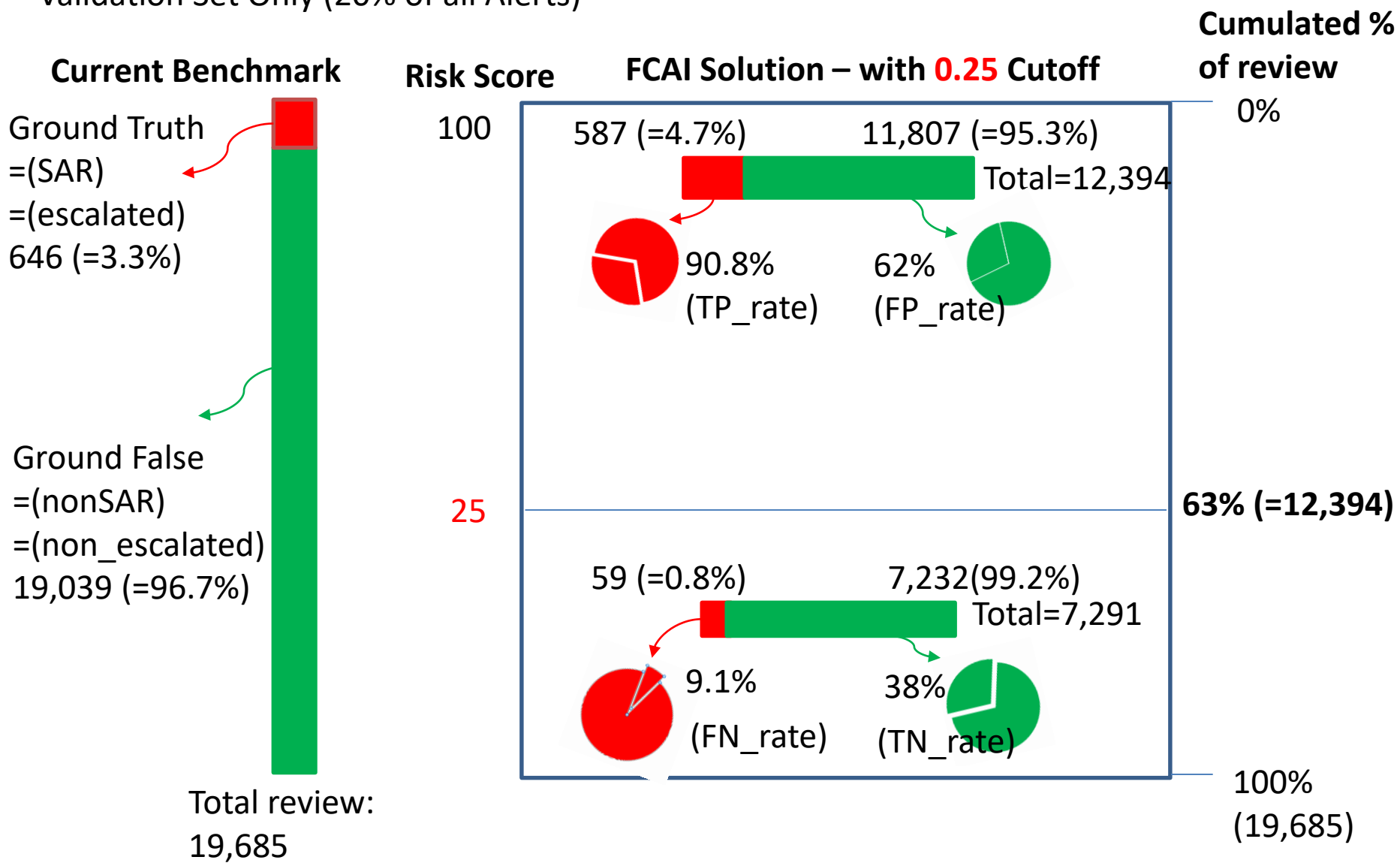
Ensemble - Validating Performance

- Accuracy Plot (prediction that are correct)
  - TP-rate and TN-rate (=1-FP-rate) vs Scoring Cutoff



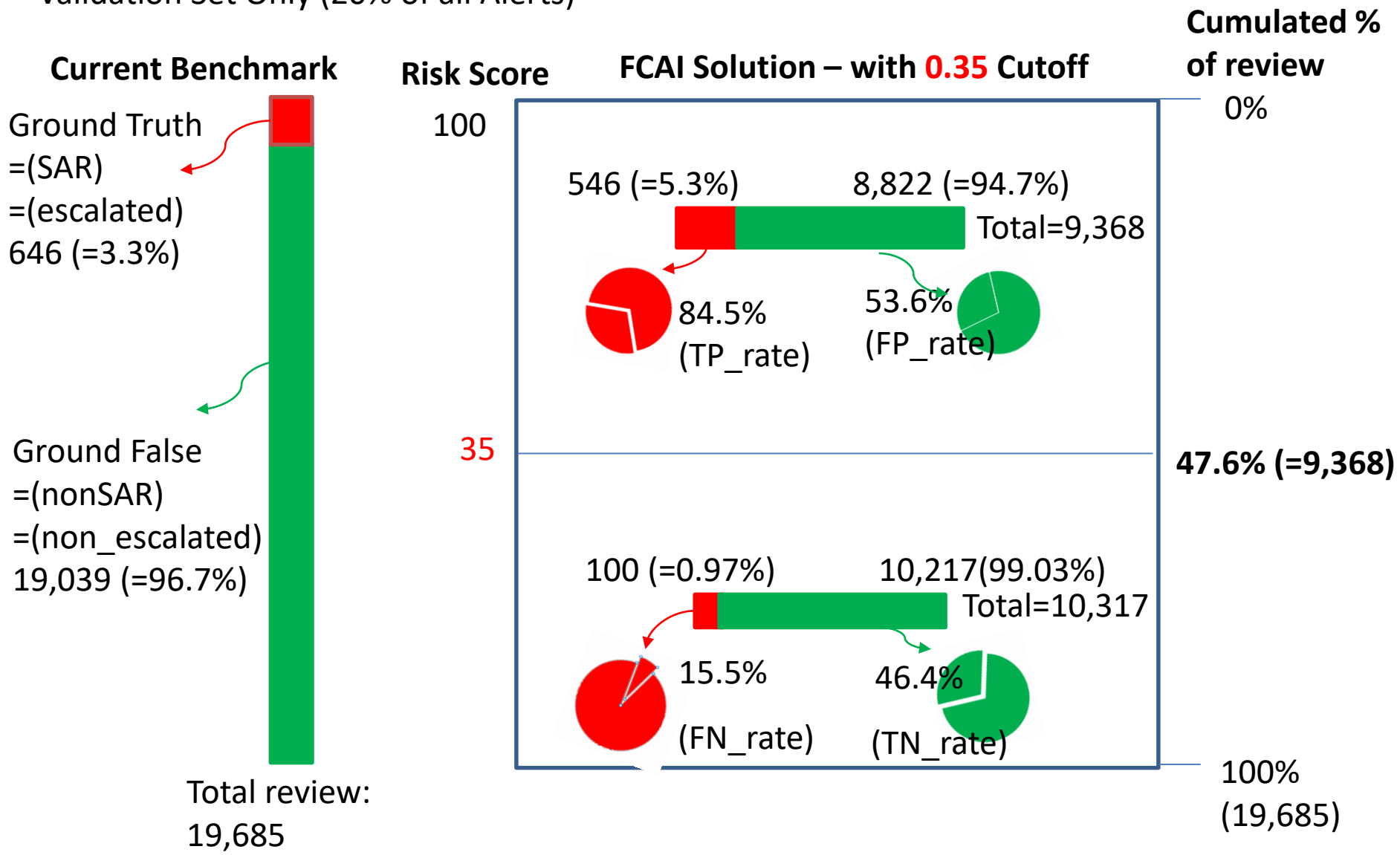
Ensemble Performance Landscape – Alert Level

Validation Set Only (20% of all Alerts)



Ensemble Performance Landscape – Alert Level

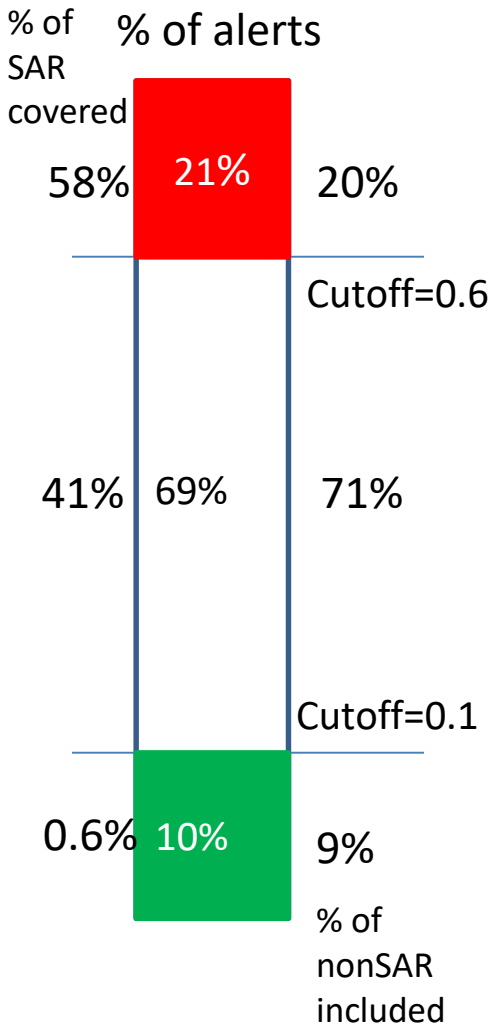
Validation Set Only (20% of all Alerts)



Ensemble Performance Landscape – Alert Level

Validation Set Only (20% of all Alerts)

Overall SAR/NonSAR ratio = 646 : 19,039 = 1:29.5



# of SAR covered	% of SAR covered	# of nonSAR covered	% of nonSAR covered	SAR/Non SAR ratio	Cumulated % of Review
375	58%	3,732	20%	1:10	0%
					21%
141	22%	3,850	20%	1:27	
					41%
105	16%	7,524	40%	1:72	
					80%
21	3%	2,125	11%	1:101	
					90%
4	0.6%	1,808	9%	1:452	
					100%

(19,039)

Enhanced BoC Insights

- **Utilize overall importance features and threshold value as preconfigured insights**
- **Provide historical population statistics**

Supervised Machine Learning Module (examples)

- LOW: The subject account has the following key transaction attributes : average_monthly_TXN_count < 1.8 and max_yearly_TotalAmount < 10,810 which corresponds to 2.5% of True Positive. The population statistics for average_monthly_TXN_count and max_yearly_TotalAmount is 30 and 10,000.

- HIGH: The subject account has the following key transaction attributes : average_monthly_TXN_count >= 1.8 and max_yearly_TotalAmount >=81,200 which corresponds to 54.5% of True Positive. The population statistics for average_monthly_TXN_count and max_yearly_TotalAmount is 30 and 10,000.

- HIGH: The subject account has the following key transaction attributes : average_monthly_TXN_count >= 1.8 and max_yearly_TotalAmount > 10,810 and max_Daily_TXNamount >= 500_which corresponds to 21.4% of True Positive. The population statistics for average_monthly_TXN_count, max_yearly_TotalAmount, and max_Daily_TXNamount is 30, 10,000 and 800.