Airspace Technology Demonstration 2 (ATD-2)

Predictive Analytics for ATD-2

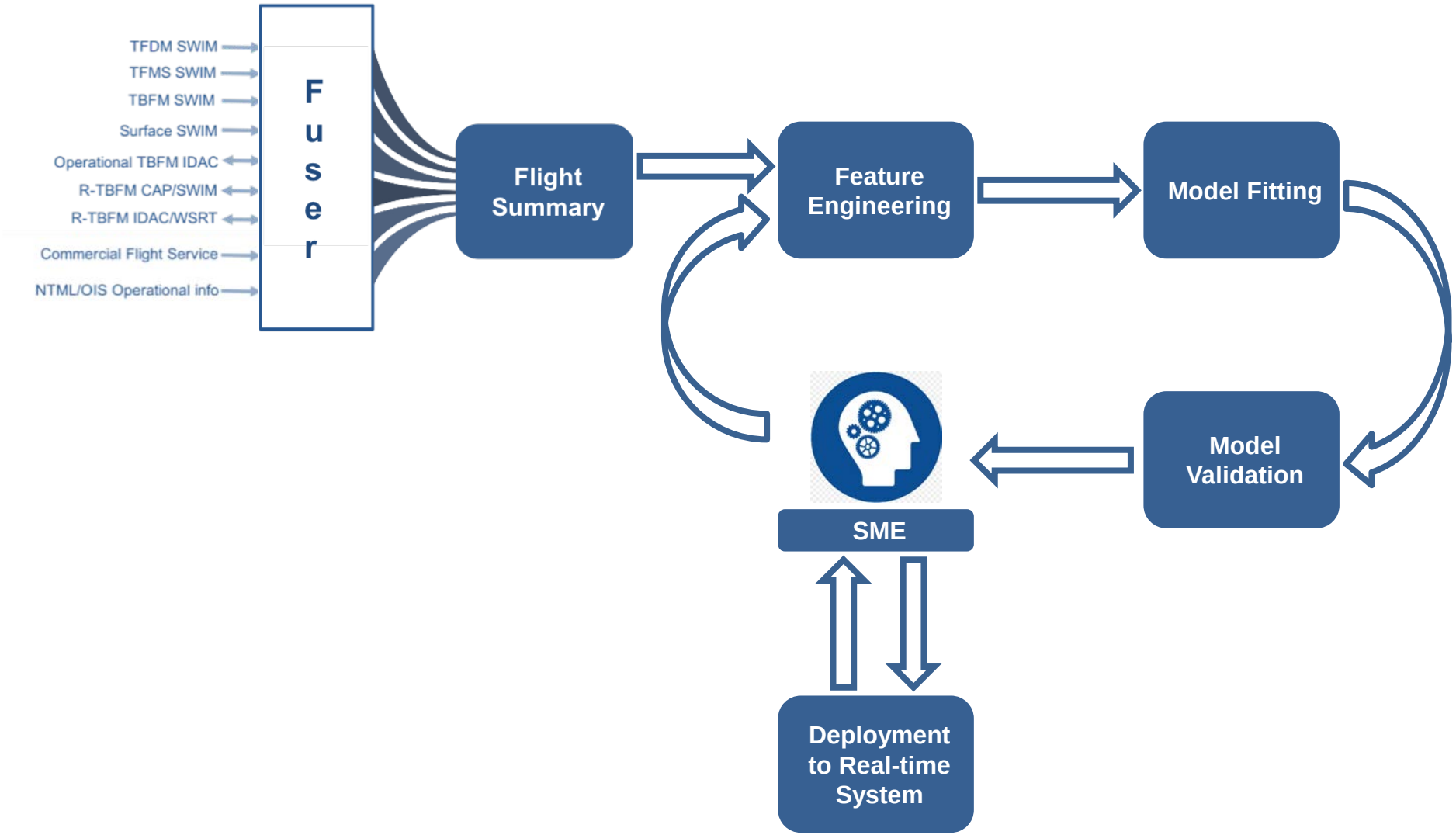
May 22, 2019

Objective

Leverage high quality data and predictive analytics to improve understanding and performance of IADS system

- Develop predictive analytics use cases that are relevant to FAA and operators
- Iterative process between data scientists and Subject Matter Experts (SME) to gain new insights
- Implementation in Python Scikit-learn allows for data scientists to focus on feature engineering and model validation
- Interested in data available in real-time system to fit models that have predictive and ultimately prescriptive capabilities

ATD-2 Predictive Analytics Workflow



Gate Conflict Use Case

Predicting gate conflicts can benefit both FAA and operators

- Providing ramp controllers with early notice of gate conflicts allows them to build a plan
- Providing FAA with early notice of gate conflicts supports the TMC in the decision whether or not to surface metering
- Understanding the different factors that cause gate conflicts could provide strategies to avoid them

Feature Engineering: Bank Level Metrics

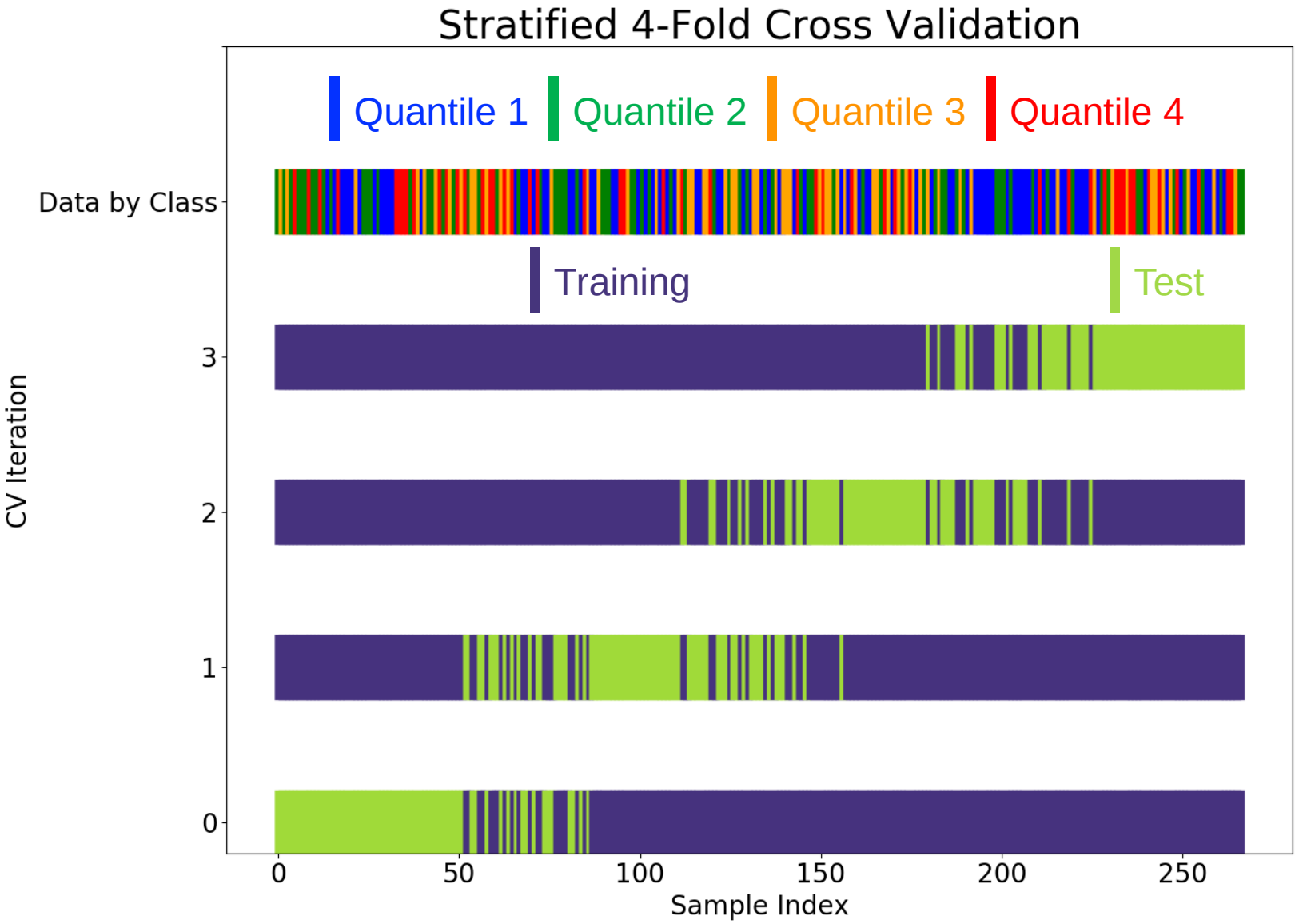
Bank Level Features

Regression Target

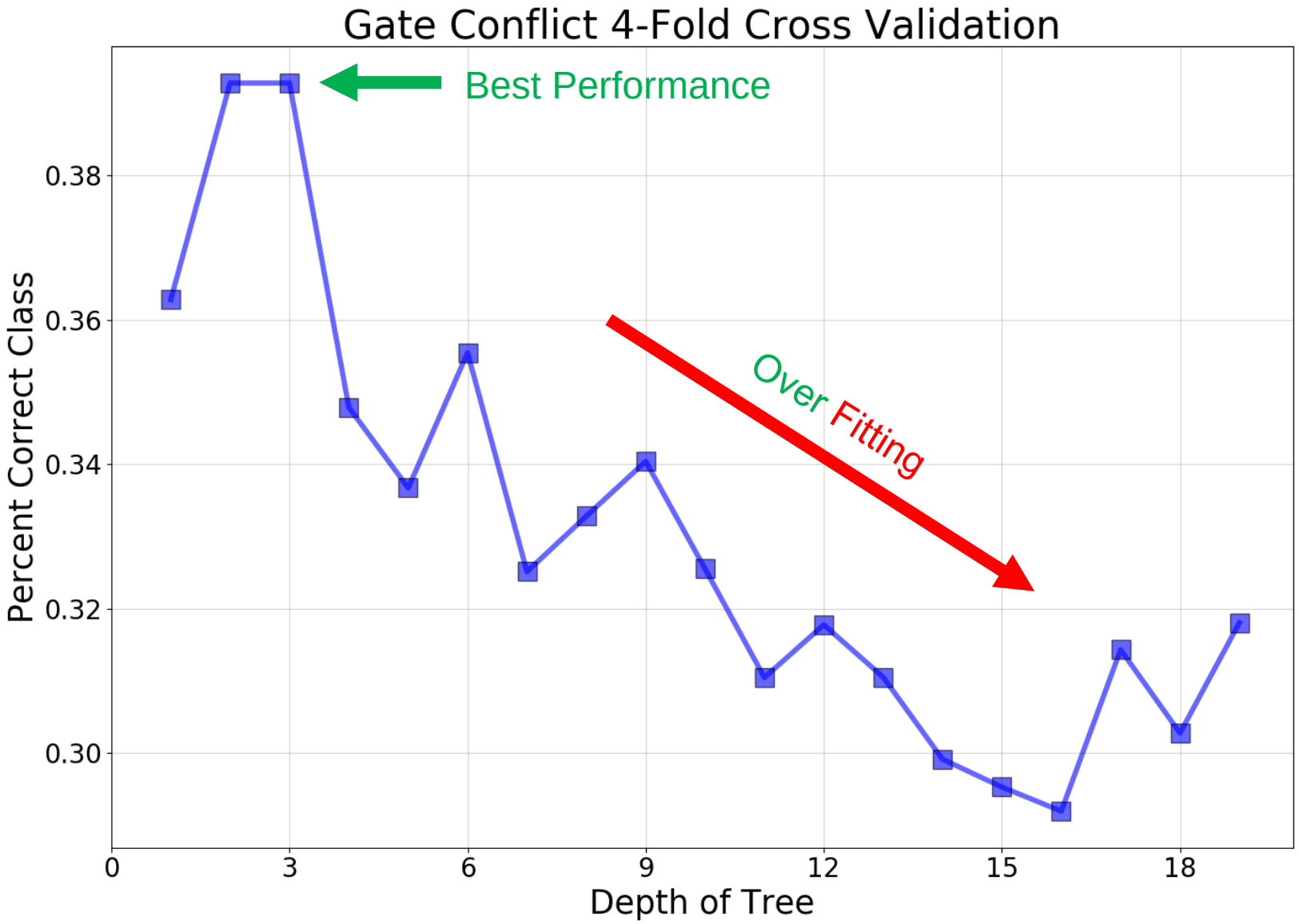
Classification Target

Date	Bank	Count Departure	Count Arrival	Difference in Dep and Arv Bank Start (bank_overlap)	Departure Gate Hold (total_actual_gate_hold)	...	Count Gate Conflict	Quantile Gate Conflict
2018-06-24	2	78	75	11.1	29.9		9	Quantile 3
2018-06-25	2	86	75	21.3	34.7		9	Quantile 3
2018-06-26	2	92	85	19.5	24.0		4	Quantile 1
2018-06-27	2	96	88	24.5	51.5		6	Quantile 2
2018-06-29	2	98	86	8.9	39.2		5	Quantile 1
2018-06-30	2	79	73	13.0	66.7		6	Quantile 2

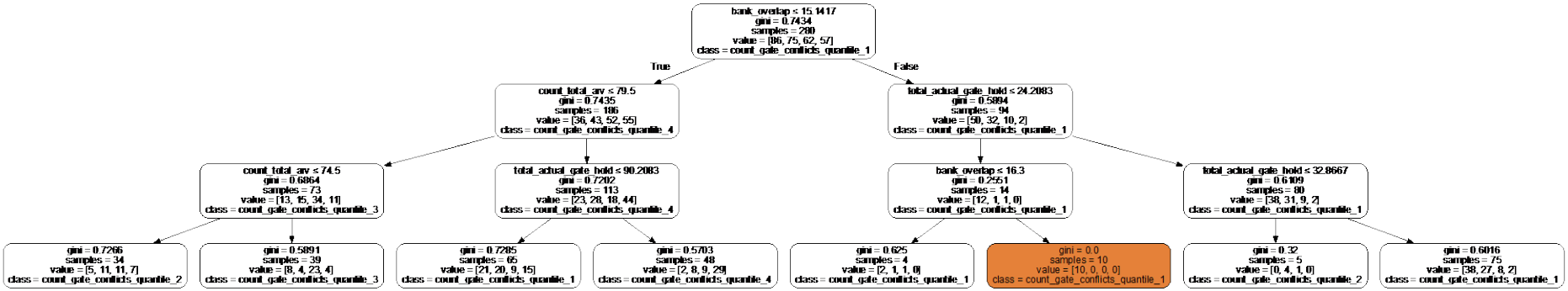
Gate Conflict: Stratified 4-Fold Cross Validation



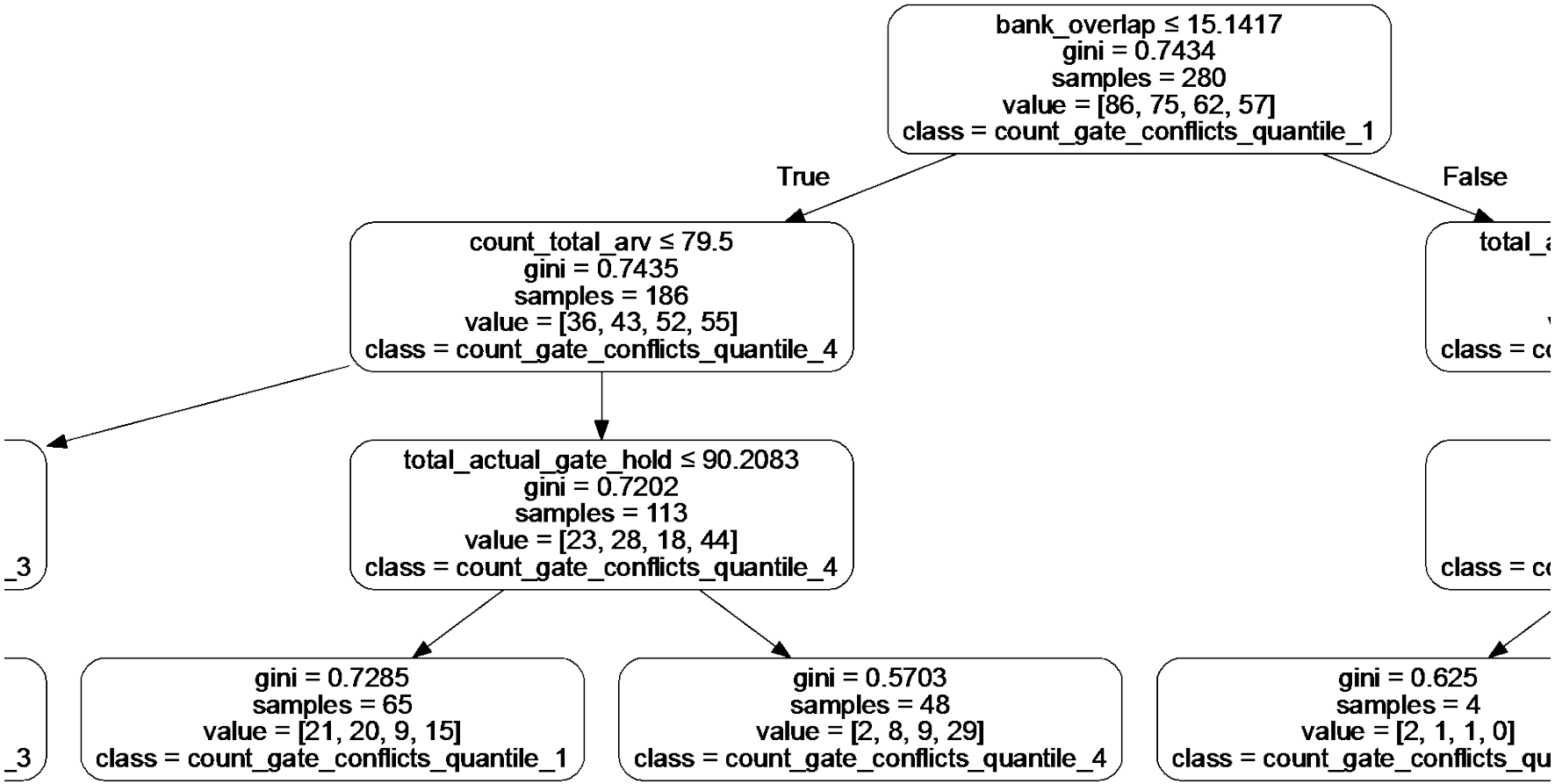
Gate Conflict Decision Tree: Hyperparameter Tuning and Validation



Gate Conflict Decision Tree



Gate Conflict Decision Tree



Wrap-up

- High quality data is the foundation of predictive analytics
- Selecting and building features that best represent the problem is a critical step in the process
- Hyperparameter tuning in combination with cross validation to achieve the best performance
- Models are trained by data scientist and then evaluated by SME in iterative process
- Deployment to real-time system is necessary to achieve impact across the NAS