Airspace Technology Demonstration 2 (ATD-2)

Predictive Analytics for ATD-2

May 22, 2019
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

- Flight Summary
- Feature Engineering
- Model Fitting
- Model Validation
- SME
- Deployment to Real-time System

Data Sources:
- TFDM SWIM
- TFMS SWIM
- TBFM SWIM
- Surface SWIM
- Operational TBFM IDAC
- R-TBFM CAP/SWIM
- R-TBFM IDAC/WSRT
- Commercial Flight Service
- NTML/OIS Operational info

5/22/2019
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

<table>
<thead>
<tr>
<th>Date</th>
<th>Bank</th>
<th>Count Departure</th>
<th>Count Arrival</th>
<th>Difference in Dep and Arv Bank Start (bank_overlap)</th>
<th>Departure Gate Hold (total_actual_gate_hold)</th>
<th>Count Gate Conflict</th>
<th>Quantile Gate Conflict</th>
</tr>
</thead>
<tbody>
<tr>
<td>2018-06-24</td>
<td>2</td>
<td>78</td>
<td>75</td>
<td>11.1</td>
<td>29.9</td>
<td>9</td>
<td>Quantile 3</td>
</tr>
<tr>
<td>2018-06-25</td>
<td>2</td>
<td>86</td>
<td>75</td>
<td>21.3</td>
<td>34.7</td>
<td>9</td>
<td>Quantile 3</td>
</tr>
<tr>
<td>2018-06-26</td>
<td>2</td>
<td>92</td>
<td>85</td>
<td>19.5</td>
<td>24.0</td>
<td>4</td>
<td>Quantile 1</td>
</tr>
<tr>
<td>2018-06-27</td>
<td>2</td>
<td>96</td>
<td>88</td>
<td>24.5</td>
<td>51.5</td>
<td>6</td>
<td>Quantile 2</td>
</tr>
<tr>
<td>2018-06-29</td>
<td>2</td>
<td>98</td>
<td>86</td>
<td>8.9</td>
<td>39.2</td>
<td>5</td>
<td>Quantile 1</td>
</tr>
<tr>
<td>2018-06-30</td>
<td>2</td>
<td>79</td>
<td>73</td>
<td>13.0</td>
<td>66.7</td>
<td>6</td>
<td>Quantile 2</td>
</tr>
</tbody>
</table>
Gate Conflict: Stratified 4-Fold Cross Validation

Stratified 4-Fold Cross Validation

Data by Class

CV Iteration

Sample Index

Quantile 1 | Quantile 2 | Quantile 3 | Quantile 4

Training

Test
Gate Conflict Decision Tree: Hyperparameter Tuning and Validation

Gate Conflict 4-Fold Cross Validation

Best Performance

Over Fitting
Gate Conflict Decision Tree

- **bank_overlap ≤ 15.1417**
  - **gini = 0.7434**
  - **samples = 280**
  - **value = [86, 75, 62, 57]**
  - **class = count_gate_conflicts_quantile_1**

- **count_total_arv ≤ 79.5**
  - **gini = 0.7435**
  - **samples = 186**
  - **value = [36, 43, 52, 55]**
  - **class = count_gate_conflicts_quantile_4**

- **total_actual_gate_hold ≤ 90.2083**
  - **gini = 0.7202**
  - **samples = 113**
  - **value = [23, 28, 18, 44]**
  - **class = count_gate_conflicts_quantile_4**

- **gini = 0.7285**
  - **samples = 65**
  - **value = [21, 20, 9, 15]**
  - **class = count_gate_conflicts_quantile_1**

- **gini = 0.5703**
  - **samples = 48**
  - **value = [2, 8, 9, 29]**
  - **class = count_gate_conflicts_quantile_4**

- **gini = 0.625**
  - **samples = 4**
  - **value = [2, 1, 1, 0]**
  - **class = count_gate_conflicts_quantile_1**
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