Your Flight is Boarding Now!

Vincenzo Tursi
Rosaria Silipo
KNIME
1. Which conditions are more likely to cause delays?

2. Given those conditions, can we predict delays?
This is the last Call. Your Flight is Boarding Now!

Predict Flight Departure Delay in US Airports
Selected Airport: **ORD**

Airline Dataset:

Airport Codes & Cities:
https://www.world-airport-codes.com/

Aircraft Maintenance Data:
http://registry.faa.gov/aircraftinquiry/NNum_Results.aspx

Twitter:
https://twitter.com/

CSV

US Weather Information:
https://www.ncdc.noaa.gov/cdo-web/datasets/

US Holiday Calendar:

Radar Images:
http://vortex.plymouth.edu/rcm-u.html

Sentiment Analysis

CSV

Storm Watch Database (Reports):
https://www.ncdc.noaa.gov/stormevents/
What are we Dealing with? Interactive Exploration

JavaScript-based Nodes

Wrapped Metanode
What are we Dealing with? Interactive Exploration

JavaScript-based Nodes

Wrapped Metanode
What are we Dealing with? Interactive Exploration

JavaScript-based Nodes

Wrapped Metanode
What are we Dealing with? Interactive Exploration

JavaScript-based Nodes

Wrapped Metanode
What are we Dealing with? Interactive Exploration

JavaScript-based Nodes

Wrapped Metanode
What are we Dealing with? Interactive Exploration

JavaScript-based Nodes

Wrapped Metanode
Wrapped Metanode View

Arrival Delay
To the WebPortal!!!
To the WebPortal!!!
To the WebPortal!!!
To the WebPortal!!!
To the WebPortal!!! - LIVE DEMO -
To the WebPortal!!! - LIVE DEMO -

Data Visualization
1. Sunburst Chart
2. Line Plot and Binned Chart
3. Box Plot

Sunburst Chart
0.408%
To the WebPortal!!! - LIVE DEMO -

Data Visualization
1. Sunburst Chart
2. Line Plot and Stacked Chart
3. BoxPlot

Sunburst Chart

© 2017 KNIME AG. All Rights Reserved.

3.02%
To the WebPortal!!! - LIVE DEMO -

Avg Arrival Delay by day of month

Avg Arrival Delay by day of month

May 2007
Avg Arrival Delay by day of month

May 2007

Average arrival delay

day of month

1.00 5.00 10.00 15.00 20.00 25.00 31.00

-11.65 0.00 1.00 2.00 3.00 4.00 5.00 6.00 7.00 7.49
Box Plot
May 2007

© 2017 KNIME AG. All Rights Reserved.
To the WebPortal!!!
To the WebPortal!!!
To the WebPortal!!!

Data Visualization

Your own JavaScript!!!
Model Selection via WebPortal

This workflow creates a model selection on the web portal, where data analysts can choose a training and test set between the datasets of 2006 to 2012, the target variable and the algorithms, which should be considered for the model comparison. The results are then displayed in a ROC Curve. For further comparison two algorithms can be chosen to compare them via a Lift Chart, before the data analysts can decide for a final algorithm.
Model Training for Predictive Analytics – for End User

This workflow creates a model selection on the web portal, where data analysts can choose a training and a test set between the datasets of 2006 to 2012, for further comparison two algorithms can be chosen to compare them via a Lift Chart, before the data analyst can decide for a final algorithm.

VIDEO: https://youtube.be/cFfPRUhz82w
TAG: guided analytics, analytics, WebPortal, KNIME Web Portal, Machine Learning, Model Selection, ROC Curve, Lift Curve, Model Evaluation

Mail notification on completion

Start
Model Training for Predictive Analytics – for End User

Model Comparison and Model Selection

This example leads you through several steps to create, compare and select a model to predict the departure delay from Chicago’s airport.

Please choose a year for training
- 2007

Please choose a year for testing
- 2008

Next
## Model Training for Predictive Analytics – for End User

### Model Comparison and Model Selection

This example leads you through several steps to create, compare and select a model to predict the departure delay from Chicago’s airport.

#### Please select the target variable

- `DepDelay`

#### Please select the algorithms for model comparison

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Node</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision Tree</td>
<td>![Decision Tree Icon]</td>
</tr>
<tr>
<td>Gradient Boosted Trees</td>
<td>![Gradient Boosted Trees Icon]</td>
</tr>
<tr>
<td>Random Forest</td>
<td>![Random Forest Icon]</td>
</tr>
<tr>
<td>Logistic Regression</td>
<td>![Logistic Regression Icon]</td>
</tr>
<tr>
<td>Neural Network</td>
<td>![Neural Network Icon]</td>
</tr>
<tr>
<td>Current Model</td>
<td>![Current Model Icon]</td>
</tr>
</tbody>
</table>

Showing 1 to 6 of 6 entries

© 2017 KNIME AG. All Rights Reserved.
Model Training for Predictive Analytics – for End User
Model Training for Predictive Analytics – for End User

Model Comparison and Model Selection

This example leads you through several steps to create, compare and select a model to predict the departure delay from Chicago's airport.

On this page you see the results displayed in a ROC Curve and two Lift Charts for model comparison. Please select one of the models as the new productive model.
Model Training for Predictive Analytics – for End User
Model Training for Predictive Analytics – for End User
Model Training for Predictive Analytics – for End User

Download the best Model

This example leads you through several steps to create, compare and select a model to predict the departure delay from Chicago's airport.

On this page you can download the best performing model in .table format.

Download the Output File

Download the Output File

[Back] [Discard] [Next]
Summary

• Data visualization matters!
  – New JS based nodes offer large variety of visualization techniques
  – Composite views with JS node views and input Quickforms
  – Running on KNIME Analytics Platform and Web Portal

Data Visualization and Interactive Exploration with KNIME

• Guided Analytics from a web browser to guide also the end users through a data science cycle

Guided Analytics
Thank you!

On behalf of KNIME we thank you for flying with us today!
The KNIME® trademark and logo and OPEN FOR INNOVATION® trademark are used by KNIME.com AG under license from KNIME GmbH, and are registered in the United States. KNIME® is also registered in Germany.