Machine-learning opportunities within the airline industry

March, 2017

Rod Fontecilla Ph.D.
VP and Global Lead Analytics
Unisys Survey Results

• Improving the passenger experience from check-in, baggage handling, immigration, shopping and on-the-go has been a challenge

• The majority (59%) of respondents to a survey conducted by Unisys indicates that it has plans to invest in advanced data analysis solutions over a period ranging from one to five years

• The biggest interest is in applying these Data Analytics solutions to improve the flow of passengers (27%), to track assets and airport operations (22%), geolocation and signaling systems in airports (22%), and Purchase of passengers in shops (15%).

• The world's leading airports have been modernizing in recent years, from the adoption of tablets and applications to improving people's experience to enhancing security

• There is an increase, in the use of real-time data analyzes as a way of understanding where travelers connect and the overall traffic patterns, even to increase sales
Travel & Transportation Data Analytics Capabilities

IMPROVING AIRLINE CUSTOMER EXPERIENCE & ENHANCING EMPLOYEE EMPOWERMENT
THROUGH UNISYS ADVANCED DATA ANALYTICS

CUSTOMER TOUCH POINTS
Call center, sales office, airport & on-board contacts, claims, marketing & surveys, etc.

Web and mobile app navigation, cookies, web analytics, email bounces, etc.

DIGITAL
OFFERS, ORDERS & SERVICES
Flight search, Bookings, Ancillaries, Loyalty, Payments, Servicing, CN&I & Boarding

SOCIAL MEDIA
Facebook, Twitter, Google+, LinkedIn, Instagram

PERSONA-DRIVEN
A personalized and proactive experience for customers

SENSOR INSTRUMENTED APPS
Context relevant, right on time content and information

BEACON BUDDY
Geolocation awareness, improving airport navigation and communication with the customer

DATA LAKE
Data is collected and analyzed for a holistic view of the customer.
- Segmentation
- Churn
- Sentiment Analysis
- Predictive Analysis

EXTERNAL DATA FEED
Flight ops., Cargo ops., Airport ops., Weather, B2B, etc.

THE NEXT GENERATION IN DATA USER EXPERIENCE

CUSTOMER EXPERIENCE MANAGEMENT
OPERATIONS MANAGEMENT
CONTINUOUS PROCESS IMPROVEMENT

OPERATIONAL HUB

ENHANCED CUSTOMER SERVICE
ACCURATE AND ON-TIME CONTENT & INFORMATION

LEADERSHIP EMPOWERMENT
- ENHANCED EFFICIENCY/EFFECTIVENESS
- BETTER VISIBILITY FOR INFORMED LEADERSHIP DECISIONS

TRAVEL & TRANSPORTATION ANALYTICS POCs

CUSTOMERS 360
Recognize all airline’s passengers including those that are not members of the loyalty program. Assess passengers’ current and predict their future value to the airline and propensity to spend on specific products and services. Discover the most popular travel patterns for top customers.

PASSENGER SERVICE DISRUPTION & RECOVERY
Predict the probability of flight delays and lost connections based on travel data, origin and destination, connecting airports and weather conditions. Discover the causal factors of flight delays. Enable proactive re-accommodation minimizing the impact of disruptions on the passengers.

AIRPORTS PASSENGER TRAFFIC MANAGEMENT
Analyze, predict and optimize passenger traffic through the airport using geolocation instrumented mobile app based on beacons or other technology. Track the passengers that are registered mobile app users, enabling personalized, context relevant offers, information and notifications.

PREDICTIVE ON-TIME FREIGHT TRANSPORT
Predict freight success factor of time sensitive goods at the moment of booking. Get to know the causal factors of efficient and less efficient routes. Discover cross-airline performance. Develop the ability to route time sensitive goods in the most successful manner.

PREDICTIVE CYBER SECURITY
Watch over payment transactions on digital channels to quickly detect anomalies, patterns and relationships. Increase the ability to detect suspicious activity, avoid airline losses and improve security to their customers.

© 2017 Unisys Corporation. All rights reserved.
Unisys Machine Learning Helps Achieve Business Insights and Customer Understanding

Machine learning …ability to learn without being explicitly programmed

Business Problem

Data

+ Data Scientists

SMEs

Machine Learning

Predictive Model

Business Insights and Business Value

© 2017 Unisys Corporation. All rights reserved.
Unisys Analytics Roadmap to Achieve Business Value

Enhance Our Client’s Business
- Data-driven discovery and innovation
- Digital transformation via customer analytics
- Use of Orthogonal Data
- Enhanced business decision making
- Massive data ingestion
- Hyperscale real-time analytics

Ingest
- Structured & Unstructured Data
- Massive Data Ingestion

Aggregate
- Create Insights

Analyze
- Create Business Value

Transform
- New Business Models
- Automating workflows
- Change Management
- Prescriptive Analytics

BI & Visualization
- Predictive Analytics
Travel and Transportation POCs

• Three Areas Using Machine Learning Algorithms
  – Air Freight
    • Build predictive classification models to determine if shipments will be
delivered on time as promised
  – Loyalty Program
    • Build visualizations and analytics based on airline loyalty data
    • Predict the probability of spending within 30/60/90 days
  – Missed Connecting Flights
    • Build predictive models on connecting flights to determine if passengers
will lose connections tomorrow based on historical data from airline routes
and tomorrow’s weather predictions
Case 1: Air Freight Data Objectives

- Discover the causation factors to identify and sustain efficient routes, and improve processes on less efficient routes
- Ability to route high value products as efficiently and successfully as possible
- Ability for airlines to charge more on high efficiency routes

Predicting on time freight transport
  - The primary value proposition for air freight is consistent speedy transport of time sensitive goods
  - The industry measures this through consistent standards and average percentage success is about 80%
  - All collection and analysis today is post execution
  - The objective is to inject advanced analysis to predict success rate upfront at the time of booking based on set parameters
Exploration by Customer Group

• Next, the data was divided according to customer groups
  – Top Partner – consists of CUST89 only, which accounts for approximately half of the total records
    • Classification models were built, and tested to be 92.77 to 93.16% accurate
  – Other Partners – consists of all customers except CUST89 (later analyses defined Other Partners as all customers excluding Top Partner and Significant Partners)
    • Classification models were built, and tested to be 88.6 to 90.32% accurate
  – Significant Partners – consists of customers with at least 1% of total records, excluding CUST89
    • Classification models were built, and tested to be 88.52 to 89.76% accurate
Exploration by Route

• The factors ‘origin’ and ‘destination’ were combined to produce ‘route’
• Classification models were built and run using route as a factor
• Model accuracy ranged from 91.99 to 92.38%
• In addition, frequent routes were found, and historical success rates of the frequent routes were determined

• Demo: Datifex
  https://demo1.datifex.com/demo_reports/routing/bin.js/unisys_arcs.html
Case 2: Loyalty Data

- Objectives: perform analytics on Customer Loyalty System data received from a UIS client airline; predict the probability of spending within 30/60/90 days; predict what will be purchased and classify the level of spending

- CLS tracks activity for all registered frequent flyers

- CLS can also track activity for identified airline customers such as registered web or mobile users

- CLS can also track anonymous passenger by analytical means (de-anonymise process identifying “unique” passengers)

- The possibility to use CLS data for the purpose of passenger de-anonymization should be examined
### Analytics from CLS Data: 3 Key Views

<table>
<thead>
<tr>
<th>Customer</th>
<th>Partner</th>
<th>Airline</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Value to the airline of individual customers</td>
<td>• Which partner(s) have the greatest share of wallet for customers</td>
<td>• Most popular routes for top groups of customers</td>
</tr>
<tr>
<td>• Calculated value for an individual customer that indicates likelihood to spend within a defined timeframe</td>
<td>• Which partner(s) have the longest relationship with customers</td>
<td>• Cabin class spread for top groups of customers</td>
</tr>
<tr>
<td>• Likely category of spend for an individual customer</td>
<td>• Are there correlations between partners and popular routes?</td>
<td>• Percent of roundtrip versus one-way flights for top groups of customers</td>
</tr>
<tr>
<td>• Likely spend amount (high, medium, low) for an individual customer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Aggregated views showing where top customers spend their money, and showing popular routes, hotels, and trends</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Results

• 88% Accuracy on determining who would spend in the next 30/60/90 days

• Fairly Good Accuracy on those customers classified as “High” spenders

• Low spenders were very difficult to predict due to the sporadic nature of their spending (one or two flights a year)
Actual Purchase Rate vs. Predicted Rate

- **Purchase Rate by Percentile of Ranked "Predicted Purchase Rate"**
  - Actual Purchase Rate
  - Predicted Rate

- **Credit by Percentile of Ranked "Predicted Credit"**
  - Actual Avg. Credit
  - Predicted Avg. Credit
Case 3: Predicting Lost Connections

- Ability to quickly accommodate affected passengers
- Airline will have a view of the economic impact of flight disruption
- Airline will be prepared to deploy staff in a timely manner
- Airline will receive data in advance to decide if it is worth re-accommodating the passengers on another airline
- Possibility to fly the most valuable passengers through other airline hubs if there is a high probability of losing a connection on the original passenger booking
- Determine the probability of losing connections in hub airports for specific airlines for tomorrow’s flights, utilizing weather predictions for the next day and past on-time performance of airline routes
- If the probability is above a threshold value, the airline reservation system will provide the number of potentially affected passengers
Lost Flight Connections - Insights

US Bureau of Transportation Statistics
National Flight Operations Data

Lost connection

On-time arrival performance

Weather’s Share of Delayed Flights

© 2017 Unisys Corporation. All rights reserved.
Prediction Models – Initial Results

• Using source data from the US Department of Transportation on flights and flight delays, data sets of connecting flights were built

• Models were build using available variables including equipment information, flight schedules, total flight time, airport locations, flight date (month, day of month, day of week)

• Delta flights were analyzed; Minneapolis St. Paul International Airport was used as the hub airport for the connecting flights

• Daily weather data was included in the model

• Initial model accuracy ranged from 85 – 90% depending on the combination of variables and type of model used
Additional Variables

• To improve accuracy, hourly weather data was added

• Historical hourly weather data from the National Oceanic and Atmospheric Administration was included for the hub airport MSP and the following fifteen origin airports: ATL, DTW, ORD, IAD, SFO, JFK, LAX, SEA, FCA, PHL, SLC, DEN, DFW, MCI and MCO

• Additional variables were created and used
  – Change in air pressure - calculated at certain intervals prior each flight
  – Days to the next holiday
  – Route frequency
Model Results

• Several models were explored, including rpart, bagging, naïve bayes, and boosting (C4.5)

• Bagging was determined to have the best accuracy (98%)

• Bagging is a good choice for this dataset because it creates multiple decision trees and averages the results. This helps achieve better accuracy with datasets that contain variance. Since bagging creates multiple decision trees, many variables are considered and become part of the model
Summary

• Using Machine Learning as a Service we are helping airline industry identify new revenue channels, identify efficiencies and decrease costs

• These models are reusable for any airline in the world

• Ability to ingest massive amount of data from different sources and correlate them to get business insights is available to all as a service

• Further models are being developed to solve other business problems airlines are experiencing
Thank you

Contact Us http://www.unisys.com/offerings/advanced-data-analytics