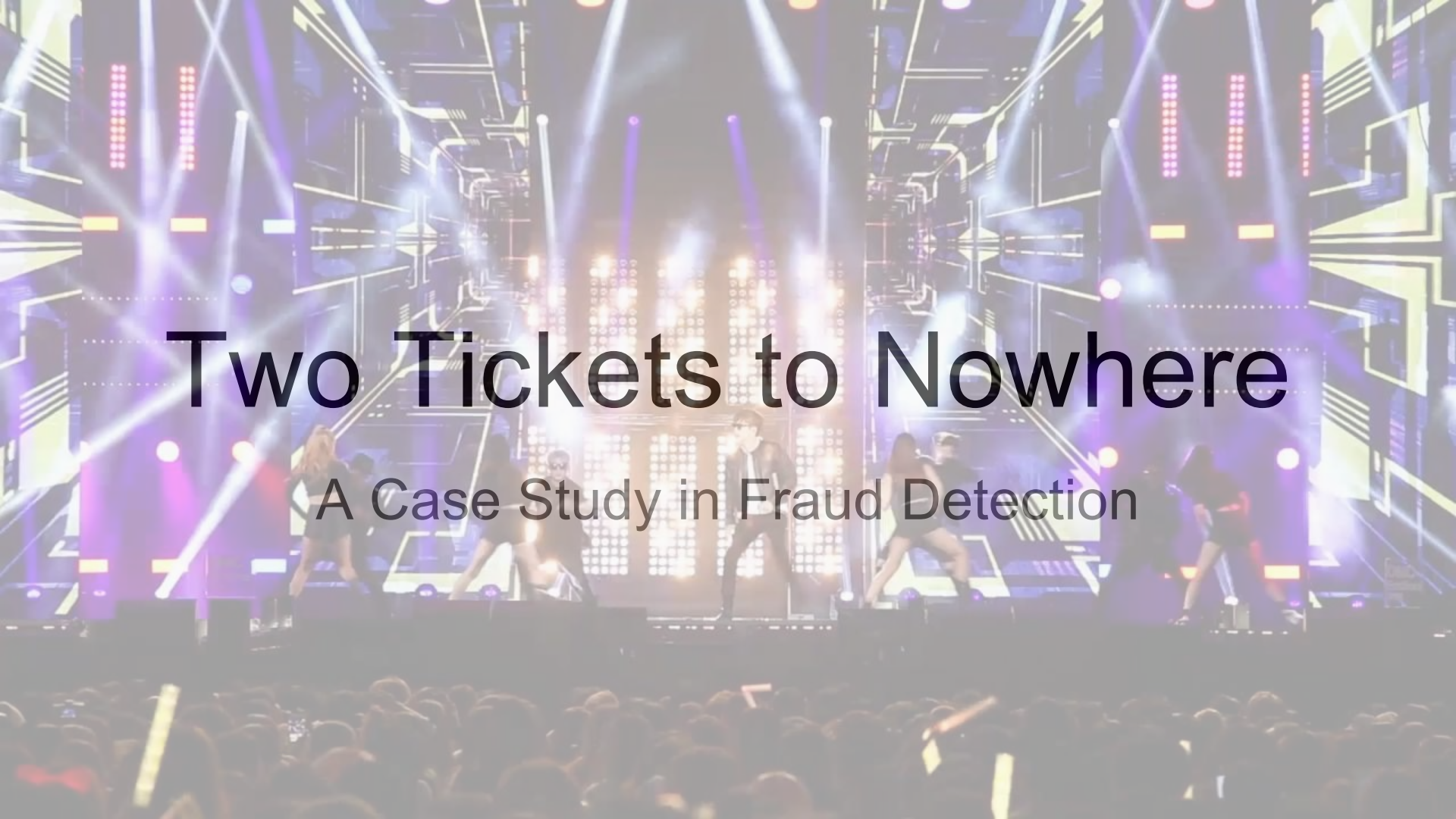


Two Tickets to Nowhere

A Case Study in Fraud Detection



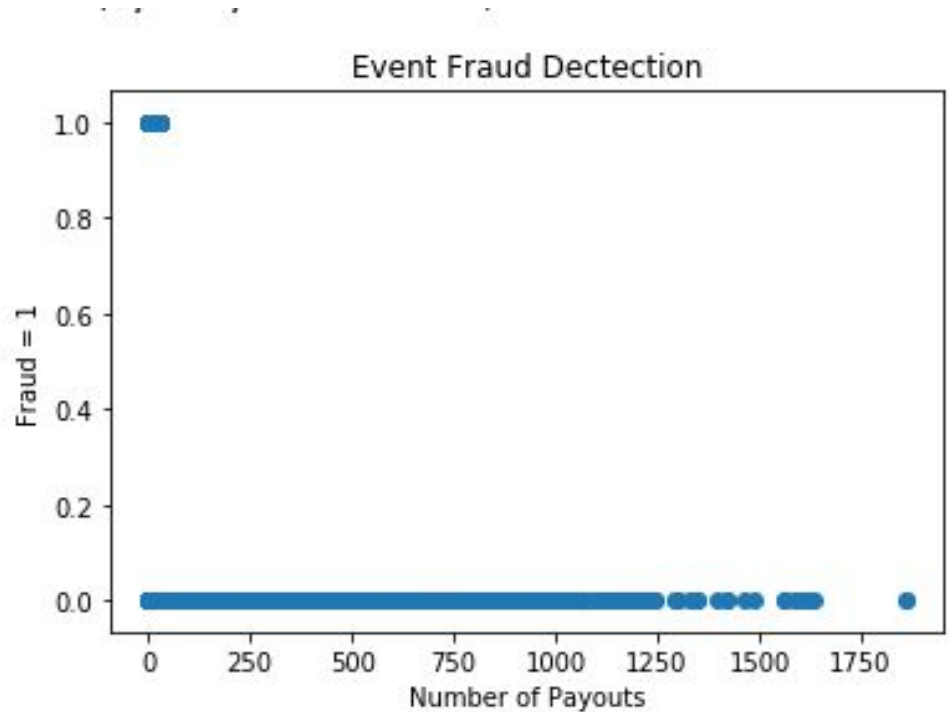
Heartbreak. Disappointment....

What
happens
when
customers
buy tickets
for fake
events?



EDA and Feature Selection

- Looked at features that helped identify events with high risk of fraud
- Ticket_types feature contained tickets_sold-->pulled this out as a feature in our model
- Other features:
 - Sale duration
 - User age
 - Name length
 - GTS
 - Number of payouts



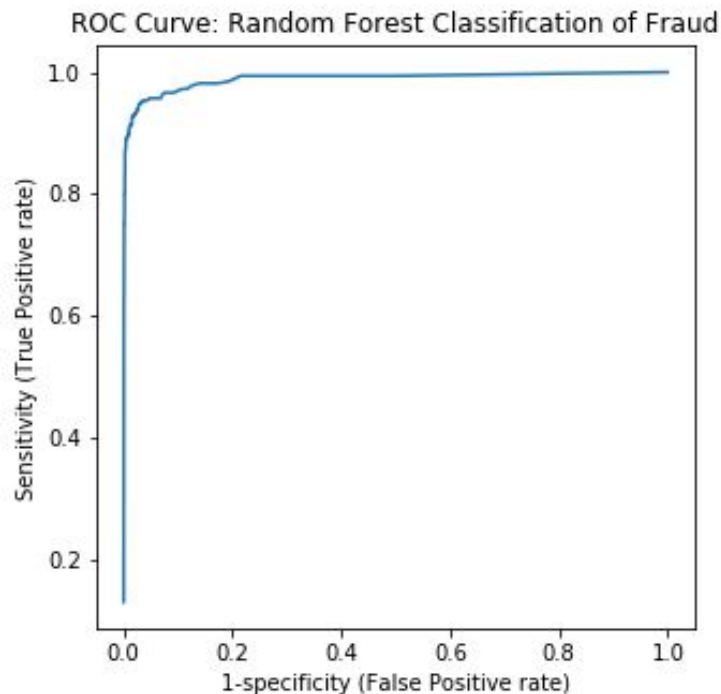
Model Performance:

- Compared Logistic Regression, Random Forest, and Gradient Boosting models using cross validation
 - Random Forest had the highest performance
 - Max_features: log2
 - N_estimators: 400
- Final Model Metrics:

ROC-AUC score: 0.938

F1 score: 0.917

Accuracy score: 0.986



MongoDB and AWS

1. Hosted our prediction script and process of storing on MongoDB with AWS
2. When pulling live data, add predicted probability before we add the response to MongoDB
 - a. This way, when we were developing our web app, we didn't need to call the model at all, we instead just pulled from Mongo
3. Also added features that were most important in deciding each event had low, medium, or high risk of fraud
 - a. Tree Interpreter helped us pull these features



Two Tickets to Nowhere Dashboard

<http://13.56.168.145:3333/>

Future Steps

- Enable Fraud Team to Flag items on Dashboard for Follow-Up and Clear
- Implement a visual to show the risk of fraud
 - Sort the events by level of risk
- Add contact information of events to website
- As time goes on, improve model to catch fraud events that get by our model

Thank you!

Karina Patel, Marcus Steele, Jinglie Zhou, Paul Knowles

Questions?

Cost -

MAY
SALES
814,477

AVG.

2611.

COSS:

0,123-

30.5%

Profit

net

Avg Food Cost

7,700,000

200

7700

Feeds: 7806

Sale duration

• www.ports
FLAGGED FOR THIS
SEASON

1/2

1/2

1/2

1/2

1/2