CABLE-TEC EXPO® 2018

VISION OF THE FUTURE

SCTE·ISBE

SCTE·ISBE CABLE-TEC EXPO 2018

ATLANTA, GA
OCTOBER 22-25

2018 Fall Technical Forum
SCTE·ISBE • NCTA • CABLELABS

DIGITIZING THE CUSTOMER EXPERIENCE

Predicting Service Impairments from STB Errors

Justin Watson, Comcast

Dr. Roger Brooks & Andrew Colby, Guavus

SCTE·ISBE

SCTE·ISBE CABLE-TEC EXPO 2018 | ATLANTA, GA OCTOBER 22-25

COMCAST

guavus a Thales company

Proof of Concept Goals

Identify imminent STB issues **before** they occur

How? Predict incidents based on STB error streams

Deliver insights so preemptive actions can be taken

Potential Benefits

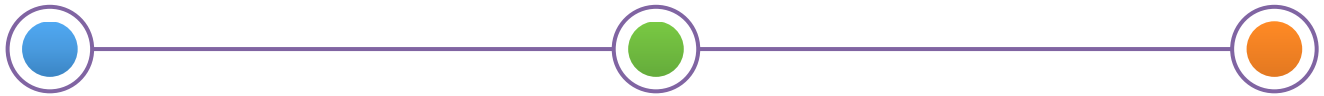
- Reduced care calls
  - Cost savings per call
  - More accurate resolution
  - Faster ID
- Reduced avoidable truck rolls
  - Significant cost savings per truck roll
  - Customer waiting
  - Customer not home
- Happier customers
  - Remote resolution
  - Customer not even aware of problem
  - Higher NPS

Ultimate Goals

Save Operational Costs

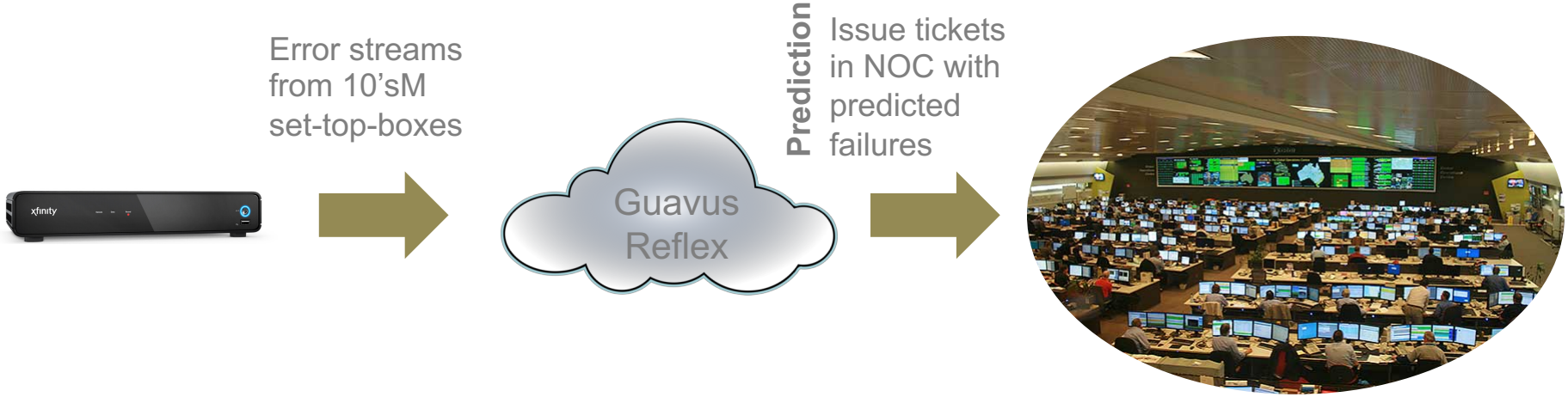
Increase Network Reliability

Improve Customer Experience



Reduce Support Events

Hypothesis: Use X1 errors to predict customer incidents



How? Address the technical problems of subscribers who are likely to call, before they call

1) Identify who will call:

Aggregated Call Prediction

- How many of my subscribers are at risk for calling within the next hours?
- What are the risk drivers?
- *Data: Based on correlation of X1 errors, outage with calls.*

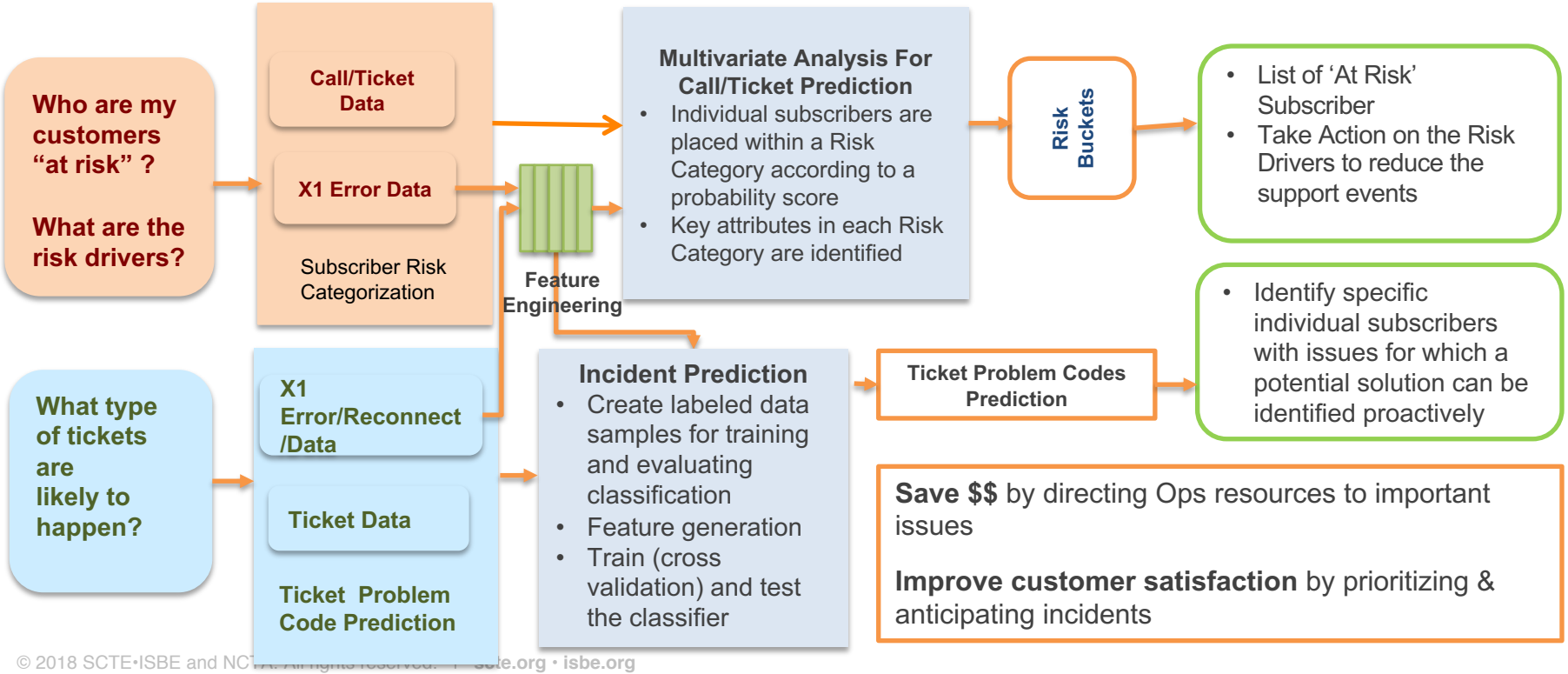
2) Predict problem codes:

Subscriber Ticket Problem Code Prediction

- Which individual subscribers are having problems which can be handled without a support event?
- *Data: Based on correlation of X1 errors, outage, reconnects with ticket problem codes*

Focus my repair efforts on the subscriber segments with high potential for calling

Understand the interconnections between symptoms and incidents



Analytics Used

Challenges

1000's of error codes

- 1000-2000 error codes
- Types of errors + meta data
  - Firmware, type, s/w version
- Error code by itself is not predictive
- Need to take it in context & consider combo of events
- History of errors from your STB and the history of problems from ticket data

Ticket data varies

- Tickets are generated by agents - can be subjective
- May/may not be correct assessment
- Multiple tickets from 1 customer call
- No taxonomic structure
- 2 different description for same problem
- Limits accuracy

10-40% will not call

- 10-40% will not call even though they have a problem
- Silent sufferers
  - Not home
  - TV not on
  - Self resolution
  - Wait to see if fixes itself
- Causes false positives
  - Algorithm predicts customer will call and they don't

Use Machine Learning (supervised and unsupervised) to:

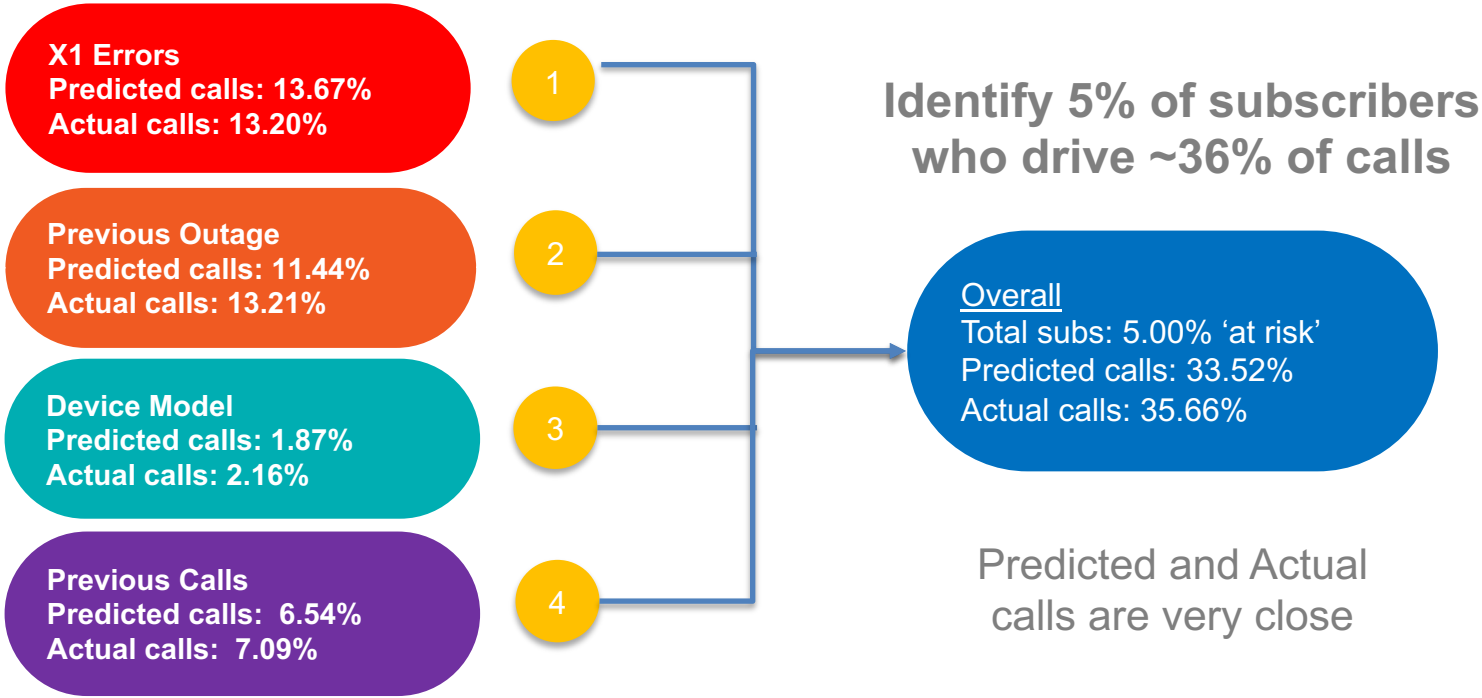
- Group ticket cause descriptions together using clustering methodology
- Classify the error codes into families
- Identify high risk subscribers (subs with high propensity to call in)
- Identify potential calls saved if certain feature/s removed
- Predict ticket type / problem code

Use risk bucketing to aggregate calls and predict tickets

- Train data model with all calls and tickets with aggregated features
- Training Period: 7days. Evaluation period: 7days.
- All subscribers with errors in last 7 days are used for prediction.
- Hourly prediction

Identify "At Risk" subs each hour: Top 5% of total subs with errors and their risk drivers.

Subs who will call: Identify risk drivers associated with ‘at risk’ subs



Calls potentially saved if risk drivers removed

1. ~14% Tickets have strong correlation with XRE and RDK Errors
2. ~11% Tickets have strong correlation with Outage events

Subs "at Risk" of Ticketing for Given Risk Drivers

Risk Drivers	Call saved if risk driver is removed	Fraction of subs impacted	Call propensity impacted subs
A	42%	91%	0.092
B	31%	95%	0.063
C	14%	100%	0.038
D	12%	100%	0.048

3. ~2% calls have strong correlation with certain device models
4. ~7% calls have strong correlation with Previous calls

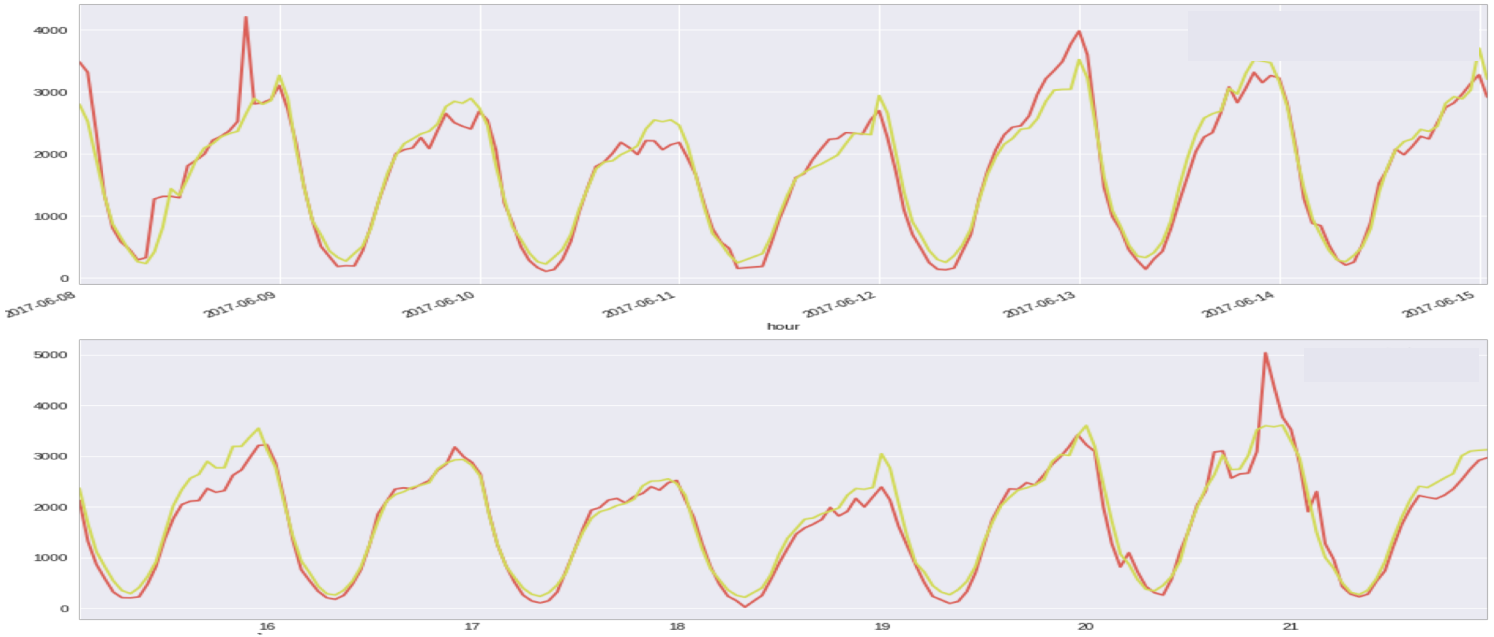
Call Prediction: Results

Risk Buckets: Predicted and Actual Calls vs. Total Subs with Errors (hourly predicted result aggregated for 7 days)

Risk Bucket	Predicted Calls	Actual Calls	Subs w/Errors
0	8%	6%	35%
1	15%	14%	25%
2	19%	18%	20%
3	24%	26%	15%
4	34%	36%	5%
total	100%	100%	100%

Subscribers in Risk Bucket 4, while accounting for only 5% of the total subscribers with set-top errors, drive 36% of actual technical calls.

Aggregated Call Prediction

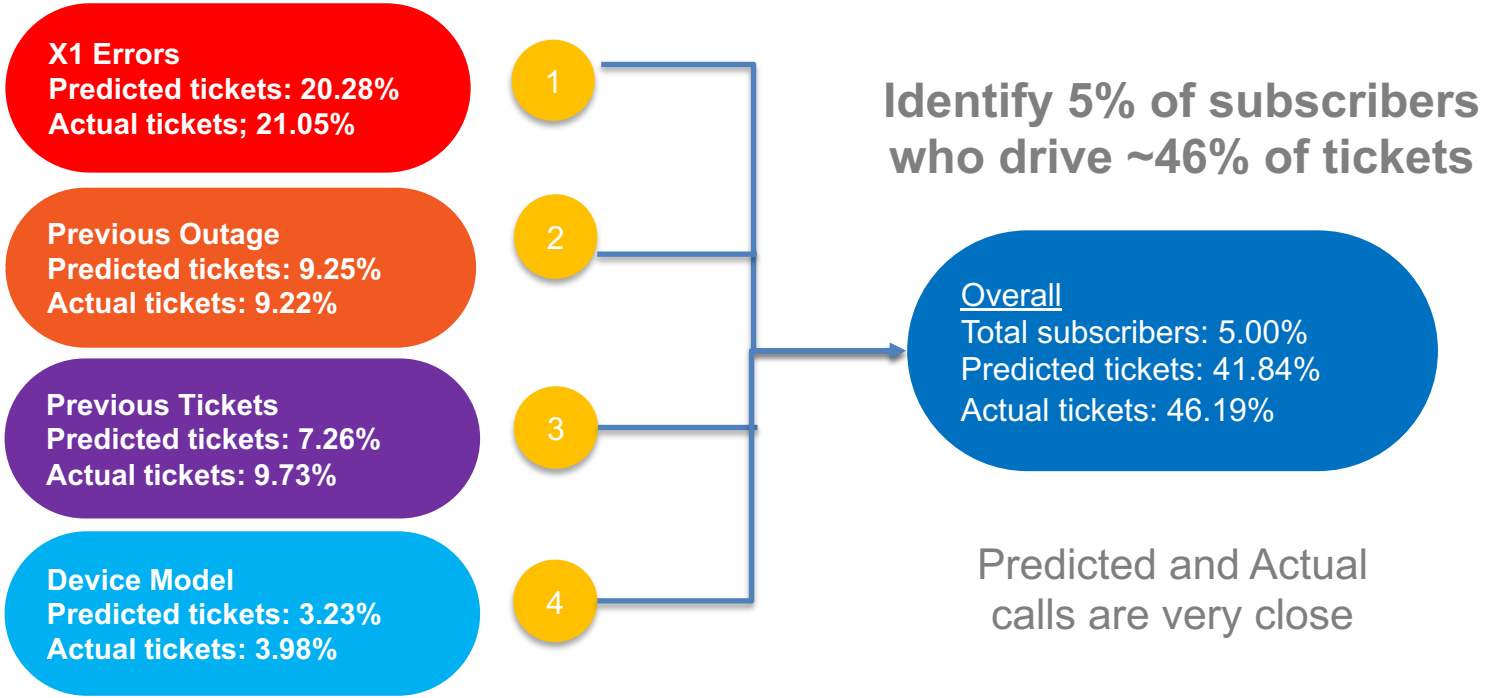


Actual

Predicted

- Predicted and Actual calls are very close suggesting that the model is highly accurate
- Call spikes trend are also captured

Subs who will generate a ticket: Identify risk drivers associated with ‘at risk’ subscribers



Tickets saved if risk drivers removed

1. ~20% Tickets have strong correlation with XRE and RDK Errors
2. ~9% Tickets have strong correlation with Outage events

Extracting Subs "at Risk" of Ticketing for Given Risk Drivers

Risk Drivers	Tickets saved if risk driver is removed	Fraction of subs impacted	Ticket propensity impacted subs
E	53%	92%	0.128
F	17%	100%	0.039
G	14%	100%	0.043
H	13%	100%	0.039

3. ~7% Tickets have strong correlation with Previous Tickets
4. ~3% Tickets have strong correlation with certain device models

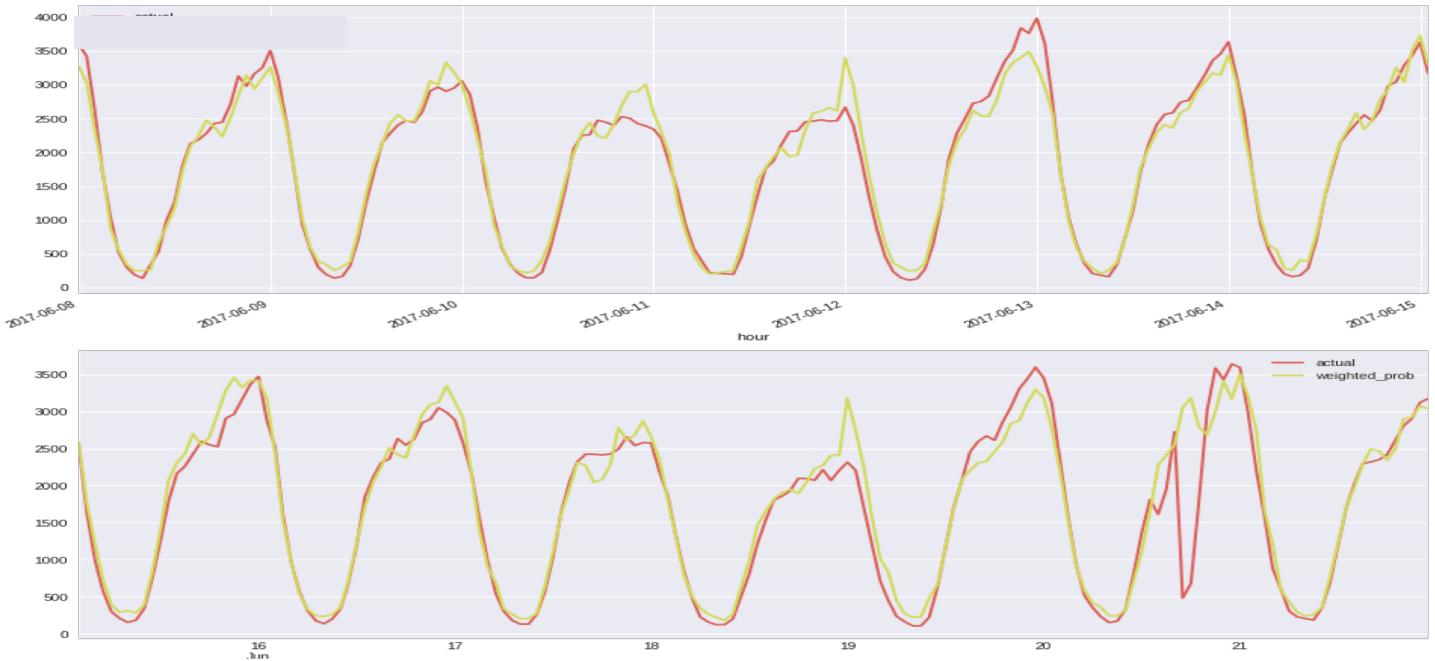
Ticket Prediction Results

Risk Buckets: Predicted and Actual Tickets vs. Total Subs with Errors (hourly predicted result aggregated for 7 days)

Risk Bucket	Predicted Tickets	Actual Tickets	Subs w/Errors
0	6%	5%	35%
1	15%	13%	25%
2	18%	16%	20%
3	19%	20%	15%
4	42%	46%	5%
total	100%	100%	100%

Subscribers in Risk Bucket 4, while accounting for only 5% of the total subscribers with set-top errors, drive 46% of actual technical tickets

Aggregated ticket forecast



Actual

Predicted

- Predicted and Actual Tickets are very close suggesting that the model is highly accurate
- Ticket spikes trend are also captured

Ticket Problem Code Prediction

How? Address the technical problems of subscribers who are likely to call, before they call

1) Identify who will call:

Aggregated Call Prediction

- How many of my subscribers are at risk for calling within the next hours?
- What are the risk drivers?
- *Data: Based on correlation of X1 errors, outage with calls.*

Focus my repair efforts on the subscriber segments with high potential for calling

2) Predict problem codes:

Subscriber Ticket Problem Code Prediction

- Which individual subscribers are having problems which can be handled without a support event?
- *Data: Based on correlation of X1 errors, outage, reconnects with ticket problem codes*

Problem Code Predictions

13 Ticket Classes, 3 Clustered sets, Total true positive tickets predicted (year)

Predicted ticket classes, problem codes	Precision	Recall	True positives/year
Class 1, code A	98%	21%	Set 1: 57.4%
Class 2, code B	91%	38%	
Class 3, code C	100%	74%	
Class 4, code D	100%	28%	
Class 5, code E	98%	50%	
Class 6, code F	90%	27%	
Class 7, code G	81%	32%	
Class 8, code H	86%	5%	
Class 9, code J, K	99%	20%	Set 2: 33.0%
Class 10, code L, M	83%	37%	
Class 11, code N, P	92%	37%	
Class 12, code Q, R	73%	4%	
Class 13, codes S, T, U	96%	35%	Set 3: 9.6%

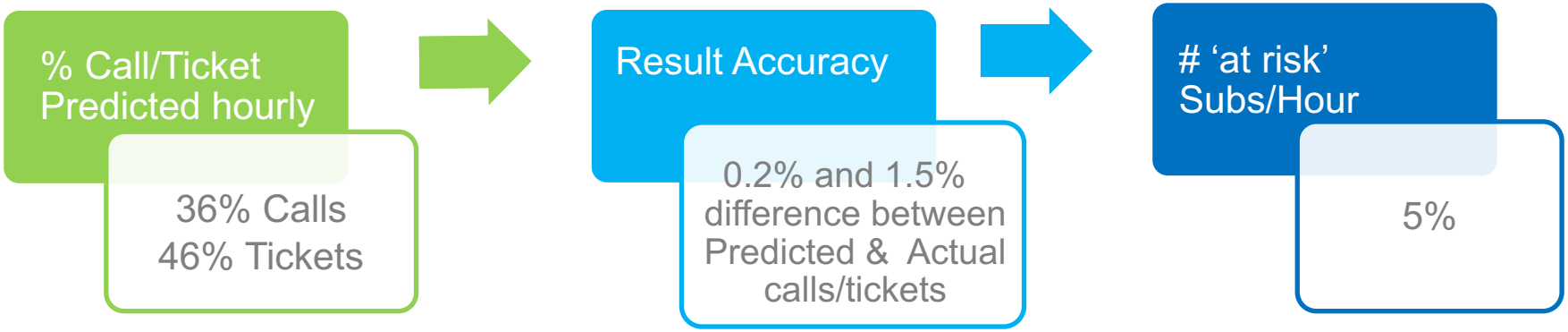
Some subscriber problems can be proactively solved

- Identify those ticket classes for which accurate predictions can be made in regards to appropriate resolutions
- Proactive resolution of issues will decrease customer care costs and increase customer satisfaction hence NPS score

Sub Id	Risk score	Potential ticket problem code	Possible resolution codes
XXX	0.44	Incorrect Boot file	Provisioned modem: 80.96% Customer equipment: 9.21% SIK to Customer: 2.32% Comcast – excluding voicemail 1.71% Reconfigured: 0.12%

Conclusion

Proof of Concept Summary



- **Identified the 5% of subs with error messages 'at risk' of calling care in the next hour.**
- Identified the **5% of subscribers** who **drive 36% of technical calls**
- Identified the **5% of subscribers** who **drive 46% of technical tickets**
- Identified **Risk drivers** associated with 'at risk' subs.

- Potential to predict many **true positive tickets** per year **related to the 13 predicted problem code classes.**

Potential benefits

- Call Prediction
- Guavus can identify the 5% of subscribers who drive 36% of technical tickets.

- Ticket Prediction
- Guavus can identify the 5% of subscribers who drive 46% of technical tickets.

- Predict numerous **true positive tickets** per year
- Identify which risk drivers lead to calls and tickets
- Predict ticket problem codes

Potential benefits

Save Operational Costs

- Reduced care calls
- Reduced avoidable truck rolls
  - ~13% of calls lead to truck rolls
- **Potential for $ millions in savings**

Increase Network Reliability

- Faster ID of problem
- More accurate resolution
- Remote resolution
- **Increased uptime**

Improve Customer Experience

- Happier customers
- Problems resolved proactively
- **Higher NPS**

Thank You!

Justin Watson, Comcast
Justin_Watson@comcast.com

Dr. Roger Brooks, Guavus
Roger.brooks@guavus.com

Andrew Colby, Guavus
Andrew.Colby@guavus.com