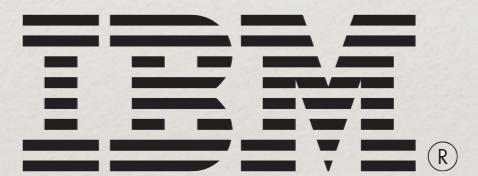
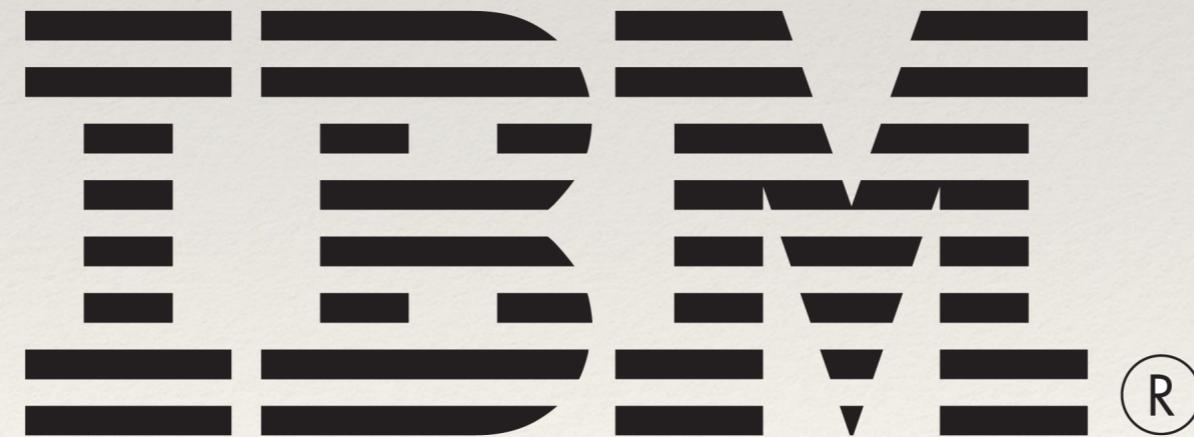


Lloyd Montgomery, The SEGAL Group

Sentiment Analysis as a tool to Define Customer Personas



Motivation



Motivation

**Customers who
Need Help**



**Support Tickets to
Coordinate Help**



**Support Analysts
Administering Help**



Motivation

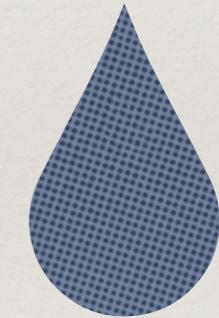
**Customers who
Need Help**



**Support Tickets to
Coordinate Help**



**Support Analysts
Administering Help**



Motivation

**Customers who
Need Help**



**Support Tickets to
Coordinate Help**



**Support Analysts
Administering Help**



Motivation

**Customers who
Need Help** → **Support Tickets to
Coordinate Help** ← **Support Analysts
Administering Help**



**Support Tickets to
Coordinate Help**



**Support Analysts
Administering Help**



Motivation

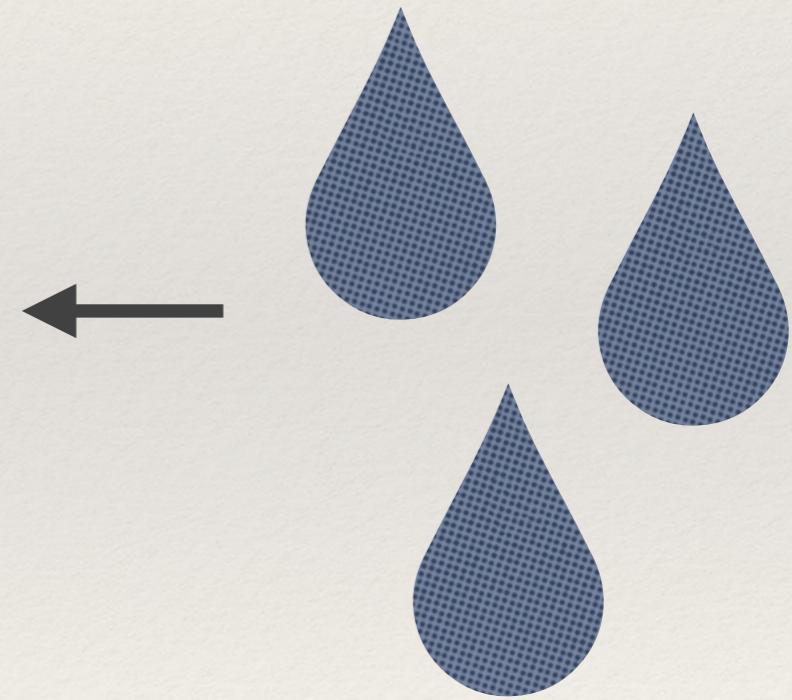
**Customers who
Need Help**



**Support Tickets to
Coordinate Help**

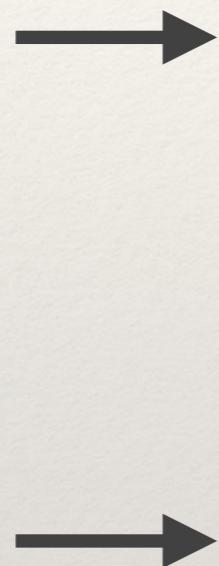


**Support Analysts
Administering Help**



Motivation

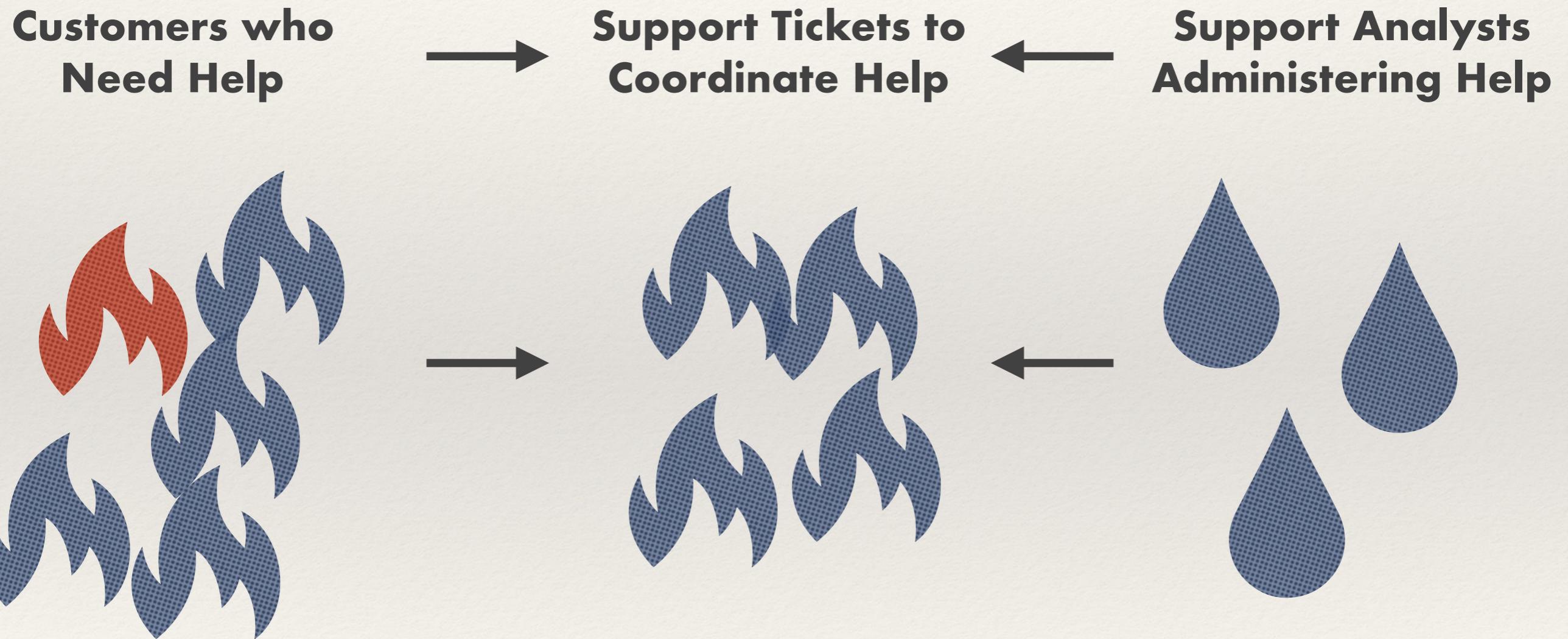
**Customers who
Need Help** → **Support Tickets to
Coordinate Help** ← **Support Analysts
Administering Help**



Motivation



Motivation



Research Questions

- RQ 1. What are the features of a support-ticket model to best describe a customer escalation?**
- RQ 2. Can ML techniques that implement such a model assist in escalation management?**

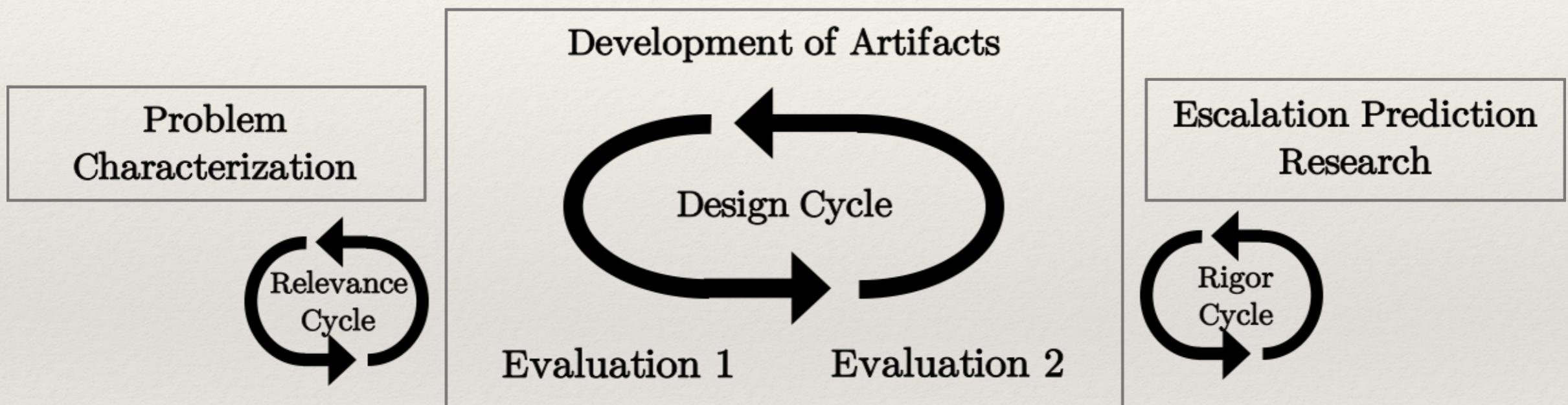
Results!

Context-based features from the perspective of the customer can capture elements of escalations

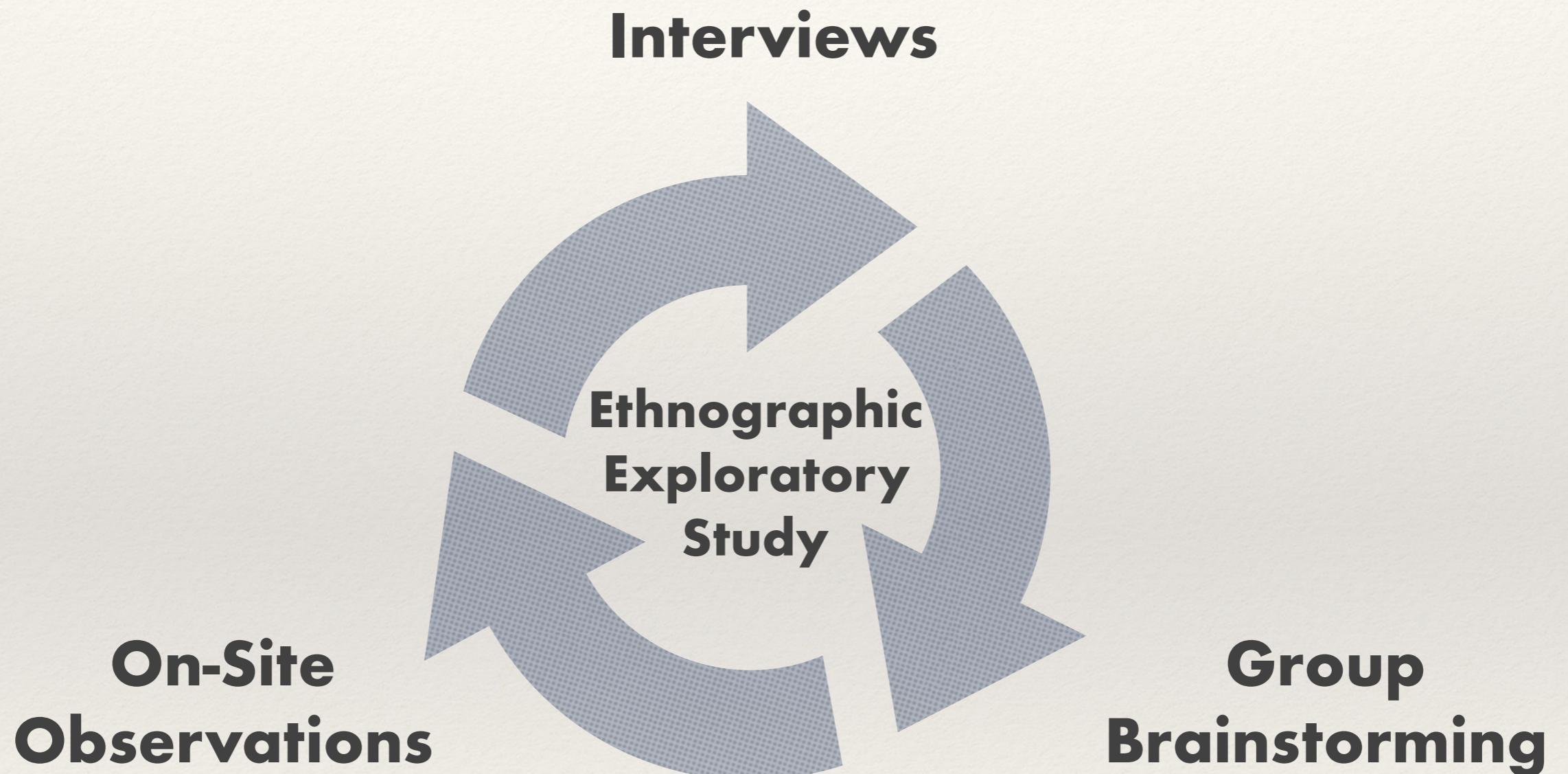
Yes! Machine learning applied to support ticket data can predict customer escalations

Methodology

Design Science Research



RQ1: Problem Characterization



Related Work

**Customer
Relationship
Management**

**Support Ticket
Automated
Categorization**

**Escalation
Prediction**

(Bruckhaus & Madhavji)

**Feature
Engineering**

Feature Engineering

**Support Analysts
know their
Customers**

**Feature Engineering Facilitates
the Transfer of Domain-
Specific Knowledge**

RQ1: Support Ticket Model Features

Basic Attributes

- Number of Entries
- Days Open
- Escalation Type
- Support Ticket Ownership Level

Customer Perception of Process

- Number of Support people
- Number of Increases in Severity
- Number of Decreases in Severity
- Number of Sev4/Sev3/Sev2 to Sev1 Transitions

Customer Perception of Time

- Time until First Contact
- Average Support Response Time
- Difference in Average vs Expected Response Time
- Days Since Last Contact

Customer Profile

- Number of Closed Support Tickets
- Number of Closed Escalations
- Escalation to Support Ticket Ratio
- Expectation of Support Response Time
- Number of Open Support Tickets
- Number of tickets opened in the last X months
- Number of tickets closed in the last X months
- Number of Escs opened in the last X months
- Number of Escs closed in the last X months
- Expected support response time given the last X months

RQ1: Support Ticket Model Features

Basic Attributes

- Number of Entries
- Days Open

Customer Perception of Time

- Time until First Contact
- Average Support Response Time
- Difference in Average vs Expected Response Time

Customer Perception of Process

- Number of Support people
- Number of Increases in Severity
- Number of Decreases in Severity
- Number of Sev4/Sev3/Sev2 to Sev1 Transitions

Customer Profile

- Number of Closed Support Tickets
- Number of Closed Escalations
- Escalation to Support Ticket Ratio
- Expectation of Support Response Time

RQ1: Support Ticket Model Features

Basic Attributes

- Number of Entries
- Days Open

Customer Perception of Time

- Time until First Contact
- Average Support Response Time
- Difference in Average vs Expected Response Time

Customer Perception of Process

- Number of Support people
- Number of Increases in Severity
- Number of Decreases in Severity
- Number of Sev4/Sev3/Sev2 to Sev1 Transitions

Customer Profile

- Number of Closed Support Tickets
- Number of Closed Escalations
- Escalation to Support Ticket Ratio
- Expectation of Support Response Time

RQ1: Support Ticket Model Features

Basic Attributes

- Number of Entries
- Days Open

Customer Perception of Time

- Time until First Contact
- Average Support Response Time
- Difference in Average vs Expected Response Time

Customer Perception of Process

- Number of Support people
- Number of Increases in Severity
- Number of Decreases in Severity
- Number of Sev4/Sev3/Sev2 to Sev1 Transitions

Customer Profile

- Number of Closed Support Tickets
- Number of Closed Escalations
- Escalation to Support Ticket Ratio
- Expectation of Support Response Time

RQ1: Support Ticket Model Features

Basic Attributes

- Number of Entries
- Days Open

Customer Perception of Time

- Time until First Contact
- Average Support Response Time
- Difference in Average vs Expected Response Time

Customer Perception of Process

- Number of Support people
- Number of Increases in Severity
- Number of Decreases in Severity
- Number of Sev4/Sev3/Sev2 to Sev1 Transitions

Customer Profile

- Number of Closed Support Tickets
- Number of Closed Escalations
- Escalation to Support Ticket Ratio
- Expectation of Support Response Time

RQ1: Support Ticket Model Features

Basic Attributes

- Number of Entries
- Days Open

Customer Perception of Time

- Time until First Contact
- Average Support Response Time
- Difference in Average vs Expected Response Time

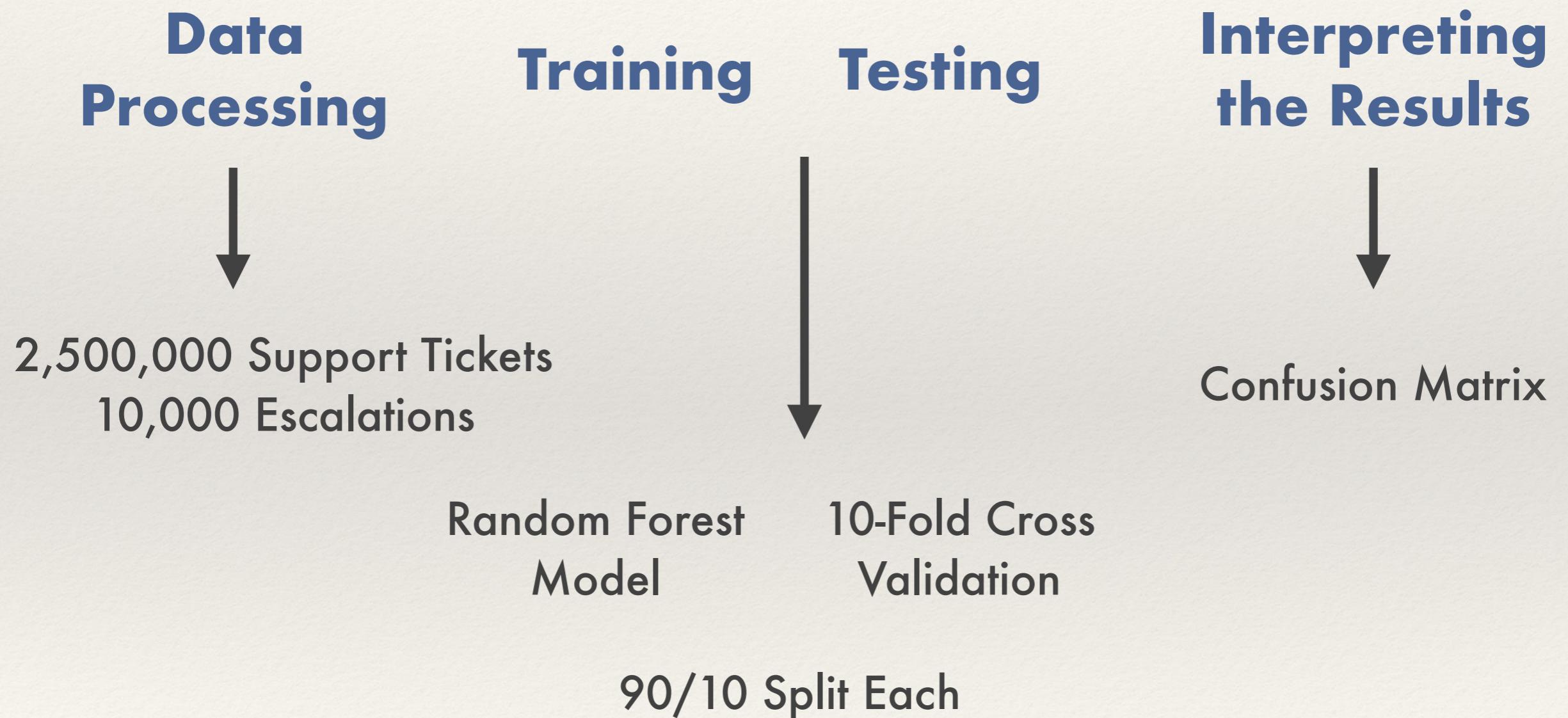
Customer Perception of Process

- Number of Support people
- Number of Increases in Severity
- Number of Decreases in Severity
- Number of Sev4/Sev3/Sev2 to Sev1 Transitions

Customer Profile

- Number of Closed Support Tickets
- Number of Closed Escalations
- Escalation to Support Ticket Ratio
- Expectation of Support Response Time

RQ2: Machine Learning Model



RQ2: Machine Learning Model

Actual	Total	Predicted as	
		Escalation - No	Escalation - Yes
Escalation - No	2,557,730	2,072,496 (TN) 81.03%	485,234 (FP) 18.97%
Escalation - Yes	10,199	2,046 (FN) 20.06%	8,153 (TP) 79.94%

Accuracy
80.49%

Recall
79.94%

Precision
1.65%

Summarization
80.79%

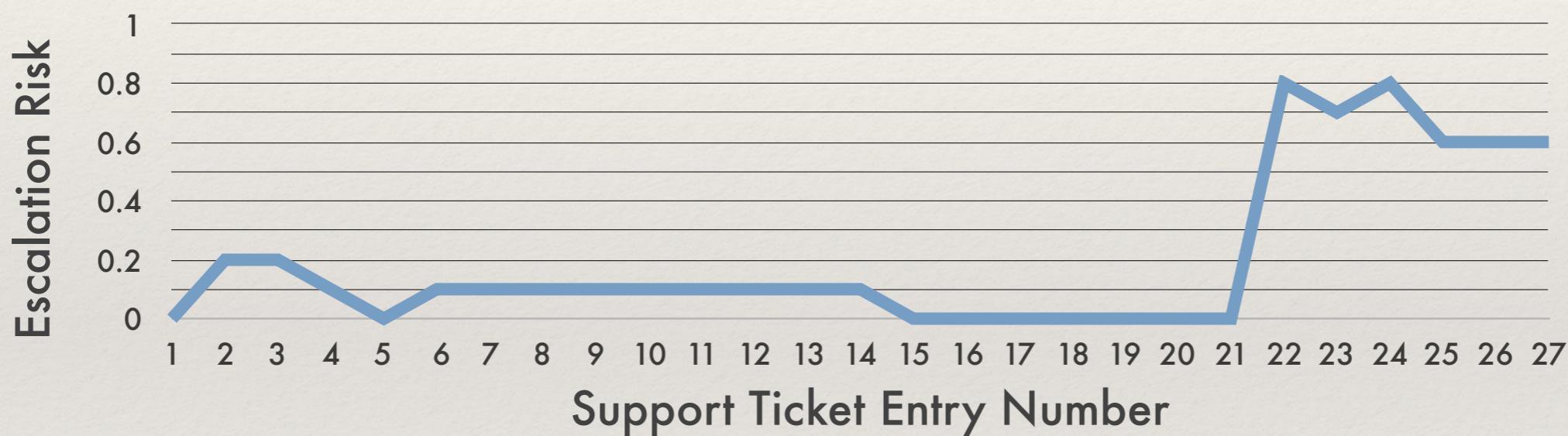
Model Evaluation 1

**In-depth Analysis of Model
Behaviour**

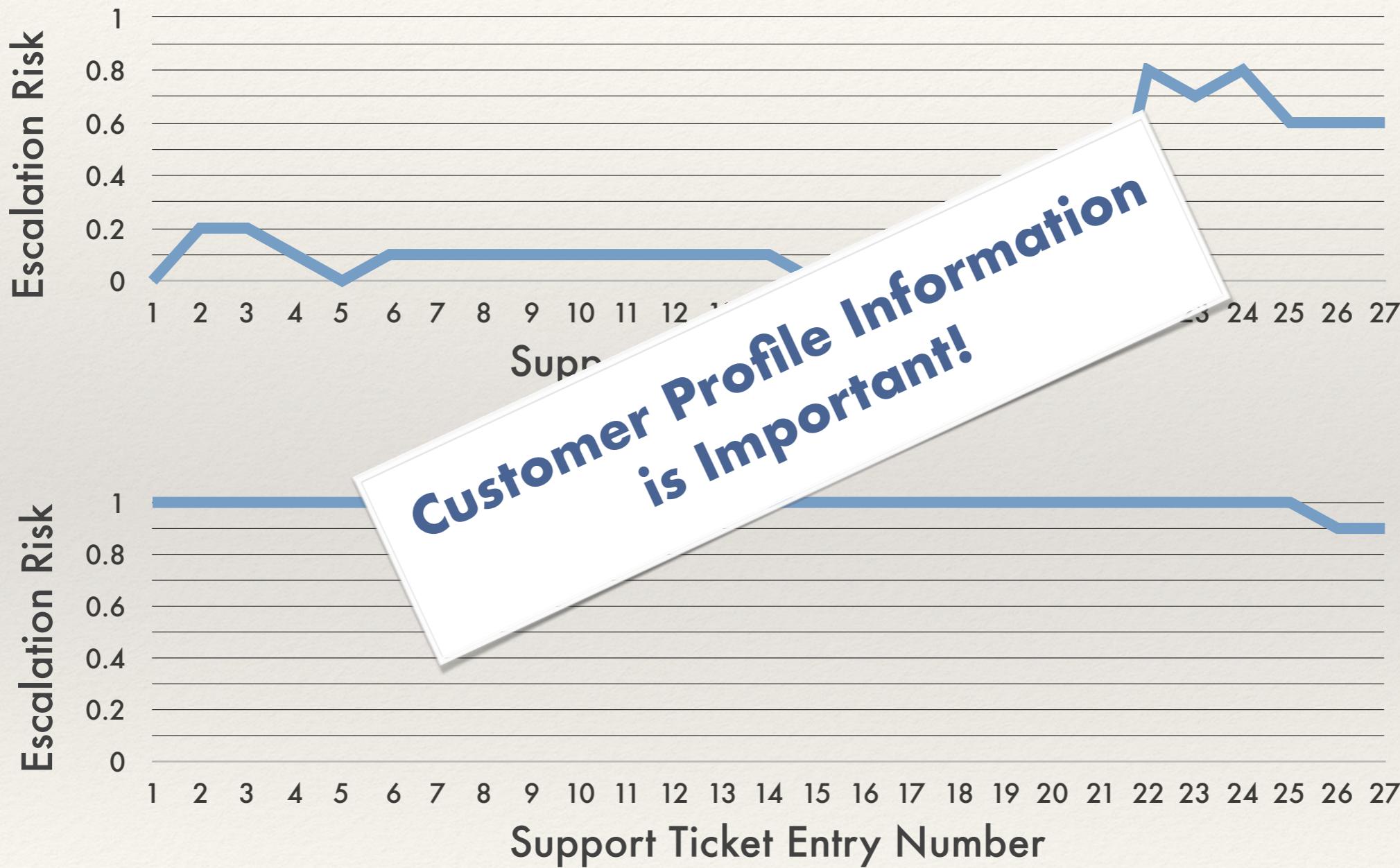
2 Hour Review of Escalations

Graphing Escalation Risk for Support Tickets

Model Evaluation 1



Model Evaluation 1



Model Evaluation 2: ECrits Tool

The screenshot shows a web-based application titled "PMR Summarization and Tracking". The top navigation bar includes tabs for "Overview", "Subscribed", and "Lloyd". The main content area displays a table of PMR entries. The table has columns for Escalation Risk, Manual Risk, Severity, PMR Info, Customer, Problem, Next Action, and Subscribe. Each row contains a set of risk values (e.g., 100, 77, 0, 35, 11), a severity value (e.g., 3, 2, 2, 2, 1), and a customer name (e.g., Doubletrax, Cisent Tenive, Teletri Group, Mathcode, True Media). The problem column contains placeholder text, and the next action column contains actions like "Aliquamer" and "Imperdiet". The subscribe column features red "X" and green "+" icons.

Escalation Risk	Manual Risk	Severity	PMR Info	Customer	Problem	Next Action	Subscribe
100 +70	100	3	06562,589,548 O16/10/11 MILLS, LOIS	Doubletrax	Lorem ipsum dolor sit amet, consectetur adipiscing elit,,	Sed ullamcorper posuere congue. Vestibulum vel nisi rutrum.	X
77 +27	50	2	24244,706,659 O16/10/05 CAIN, RICK	Cisent Tenive	Duis vehicula, odio et maximus ullamcorper, nisi risus consequa ,gravida lectu for accessibility.	Aliquamer	+
0 -15	20	2	17474,018,690 O16/10/10 WHITE, PHIL	Teletri Group	Vivamus pharetra a ori tincidunt., ,	Imperdiet	X
35	35	2	09801,428,800 O16/10/15 BARNETT, JEANNETTE	Mathcode	Morbi feugiat mauris ac lacus tincidunt faucibus. Praesent vehicula est vel nibh lobortis, id pulvinar orci elementum.	Quisque aliquam consequat enim et semper.	+
11 -44	0	1	58627,066,105 O16/10/02 SMITH, JANE	True Media	Etiam et purus vitae velit accumsan semper vel at nunc. Pellentesque ,non congue leo, et sollicitudin purus.	Praesent convallis aliquam laoreet.	+

Model Evaluation 2: ECrits Tool

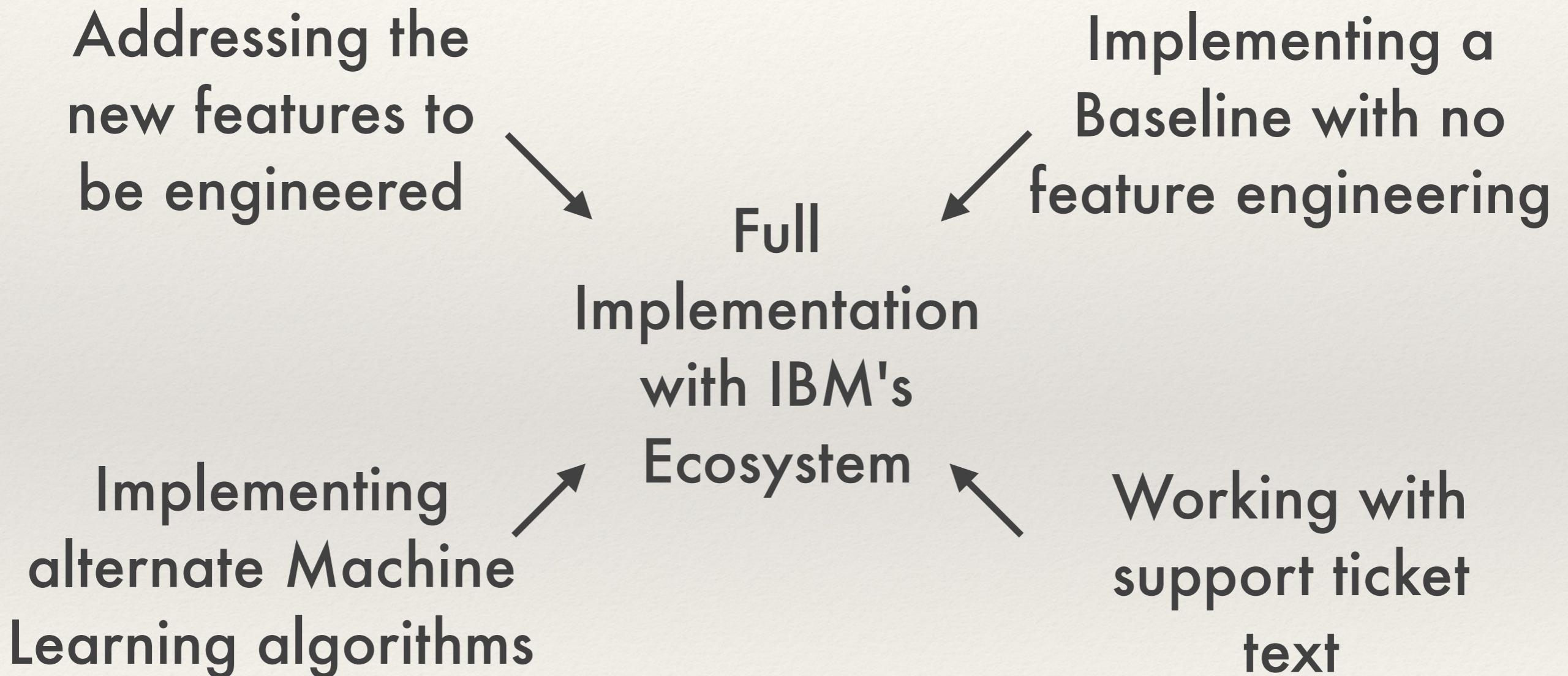
The screenshot shows a web-based application titled "PMR Summarization and Tracking" with a timestamp "(Last Refreshed: Thu Oct 27 2016 18:02:05)". The interface is divided into several sections:

- Overview**: A sidebar listing "Subscribed" companies: Doubletrax (Teletri Group).
- Basic Information**: PMR ID: 06562,589,548 O16/10/11; Severity: 3; Customer: Doubletrax (ID: 5993542); Risk of Escalation: 100%; IBM Owner: MILLS, LOIS; Opened: 2016-10-11 (10 days ago). Includes a red "Unsubscribe" button.
- Problem Description**: Problem: **Lore ipsum dolor sit amet, consectetur adipiscing elit,**; Next Action: **Sed ullamcorper posuere congue. Vestibulum vel nisi rutrum.** (time) **(time)**. Includes a "Next Action..." input field and a blue "Update" button.
- Predictive Model Features**: Last Updated: Invalid Date - NaN:NaN. Includes sections for **Perception of Process** (Involved Support: 4, Severity Increases: 0; Severity 1 Increases: 0, Severity Decreases: 0) and **Perception of Time** (First Contact Delay: 18, Avg Response Time: 92; Last Customer Contact: 2016-03-29 11:15:09, Actual vs Expected Resp Time: -48 (less than expected)).
- Customer Profile / History**: Number of PMRs: 234, Crit Ratio: 2%; Number of CritSits: 7, Expected Resp Time: 140.
- History**: A list of messages:
 - +vitae ultricies tortor non egestas nunc. Integer laoreet nunc liqua, et aliquet est maximus. Send to: Kyle Hi, Dolor sit amet, consectetur adipiscing elit. Etiam nec facilisis eros. Proin eu aliquam nisl. Sed varius cursus eros sit amet lacinia. Nulla eget turpis velit. Morbi consectetur lacinia lorem. Curabitur tempus, purus et accumsan egestas, libero lectus pellentesque sapien, laoreet dignissim lectus quam eu dolor. Vestibulum et mattis purus, fermentum sem. Etiam sit amet consequat libero, scelerisque posuere liqua.
 - +MILLS, LOIS Nulla elementum, mauris quis mollis dapibus, purus nunc rutrum nisi, nec facilisis tellus elit vitae justo. Vivamus non ante ut libero commodo bibendum in non lectus. Suspendisse id laoreet orci. Nunc ex metus, accumsan ac bibendum sed, tempus a eros. Curabitur at velit eros. In convallis lectus non malesuada condimentum. MILLS, LOIS, 2016-03-16 13:00:03
 - +ANDREWS, LELAND Hi Leland, Cras consequat est vel nisi condimentum elementum. Donec fringilla quis eros quis dignissim. Nam in nisl pellentesque, eleifend nulla tempor, molestie lacus. Phasellus faucibus sapien magna, quis fermentum magna pretium aliquam. Mauris sed est eget augue efficitur ornare. In dictum enim quis ipsum aliquam, eu luctus odio pulvinar. Fusce malesuada eros sed feugiat tempor, mauris lectus aliquet lacus, ac.
 - Send to: Kyle Hi, Vestibulum porttitor nisi sed mi egestas luctus. In varius lacinia nunc nec egestas. unknown author, 2016-03-21 12:43:50
 - Send to: Kyle Hi, Pellentesque euismod viverra libero, sed tincidunt libero malesuada vel. Morbi est nulla, dapibus ut mauris nec, varius varius est. Sed bibendum tortor sed nulla malesuada, vel congue nibh rutrum. unknown author, 2016-03-24 11:59:16
 - Send to: Kyle Morbi dolor tellus, lobortis quis nibh a, vestibulum vehicula felis. unknown author, 2016-03-25 18:30:36
 - +MILLS, LOIS Quisque non risus neque. Nunc gravida velit ipsum, a blandit eros ullamcorper eget. Vivamus vulputate elit ut velit hendrerit, ut lacinia risus viverra. Nam pretium tempus consectetur. MILLS, LOIS, 2016-03-29 11:15:09
- A comment input field labeled "Comment..." and a blue "Comment" button.

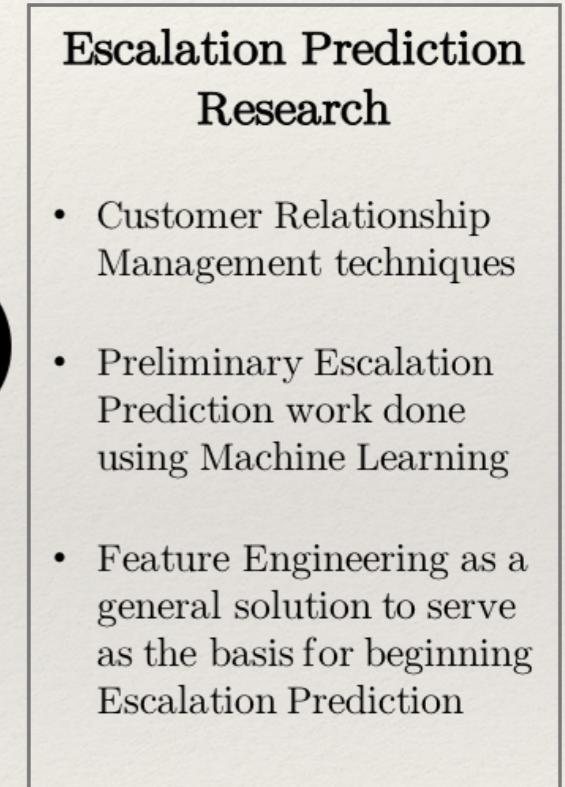
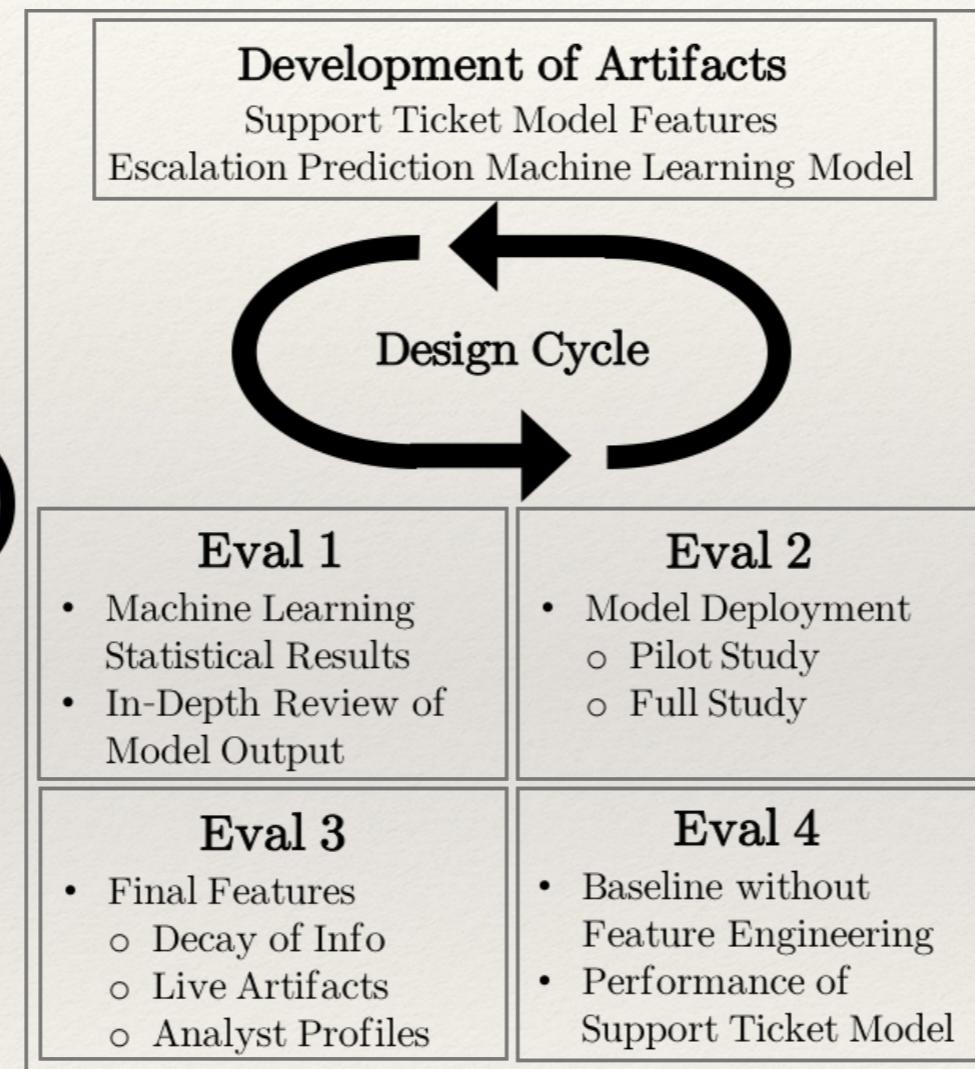
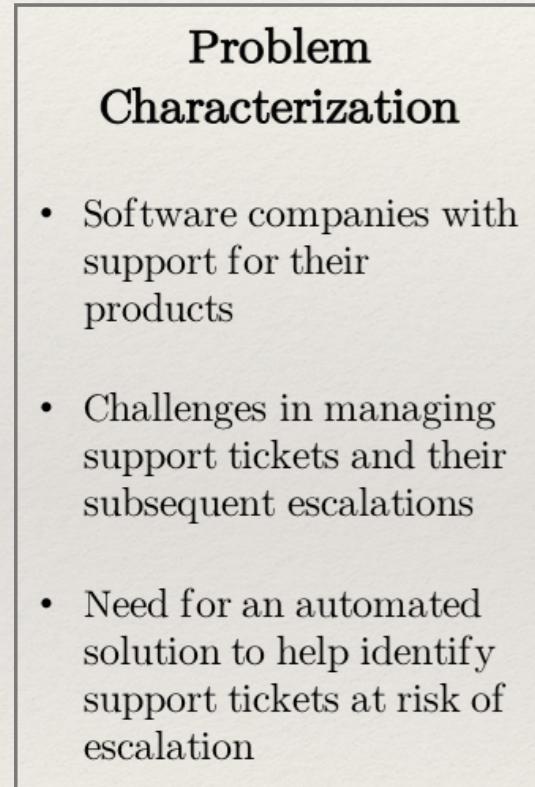
Conclusions

1. Machine Learning can be used to predict escalations against at-risk support tickets
2. Feature Engineering can capture and harness the knowledge of Support Analysts
3. Design Science is a useful methodology when undertaking the complex task of conducting research with industry

Future / In-Progress Work



Extending the Research



Improved Algorithm

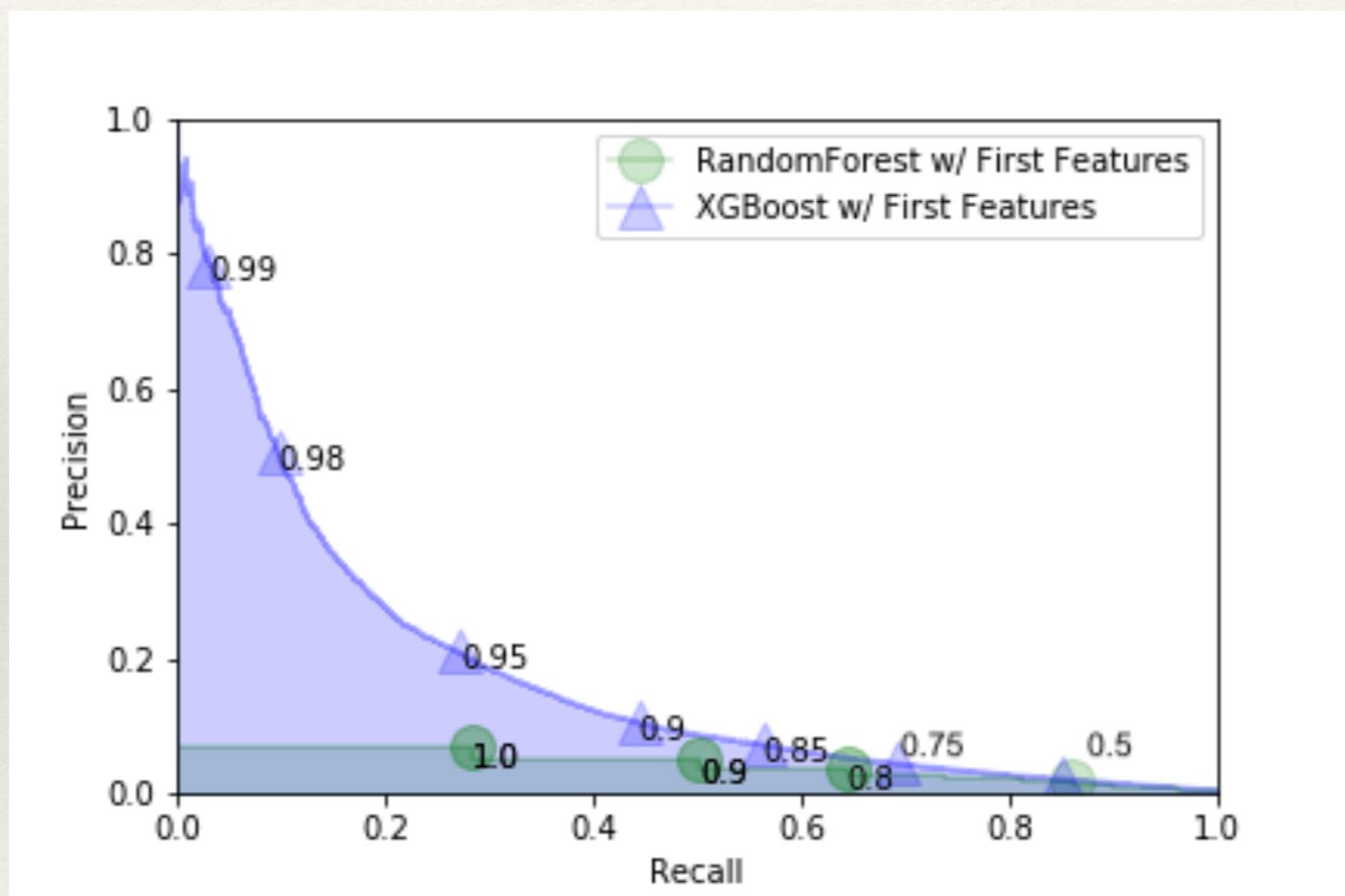
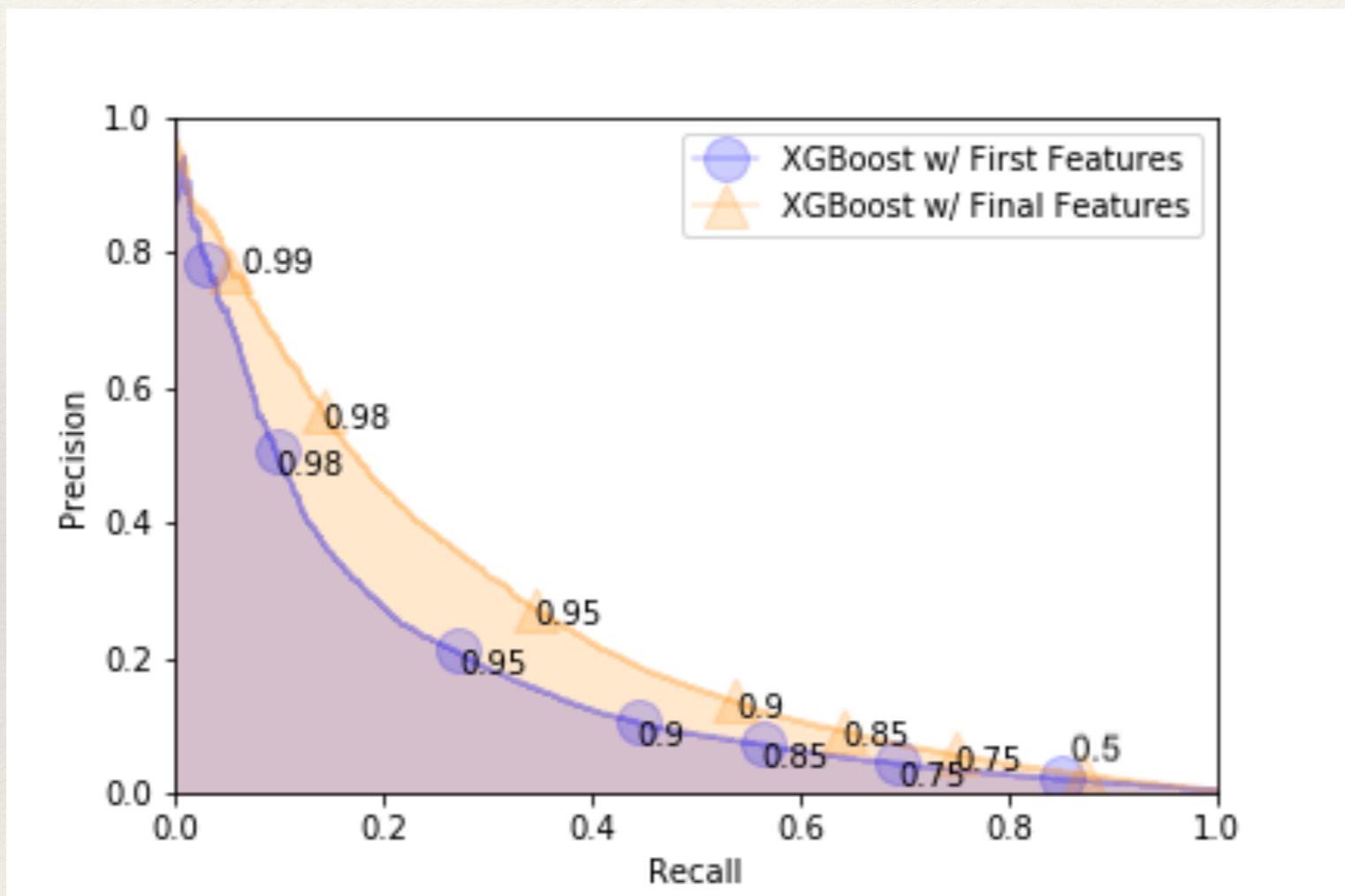


Table 1 Support Ticket Model Features with Stages of Development

Category	Feature	Description	Created or Improved During		
			Problem Characterization	Eval 1	Eval 2
Basic Attributes	Number of entries	Number of events/actions on the PMR	✓		
	Days open	Days from open to close (or CritSit)	✓		
	PMR ownership level	Level of Support (L0 - L3) that is in charge of the PMR, calculated per entry			✓
Customer Perception of Process	Number of support people in contact with customer	Number of support people the customer is communicating with	✓		
	Number of increases in severity	Number of times the Severity increase	✓		
	Number of decreases in severity	Number of times the Severity decrease	✓		
	Number of sev4/sev3/sev2 to sev1 transitions	Number of changes in Severity from 4, 3, or 2, straight to 1	✓		
Customer Perception of Time	Time until first contact	Minutes before the customer hears from IBM for the first time on this PMR	✓		
	Average support response time	Average number of minutes of all the support response times on this PMR	✓		
	Difference in average vs expected response time	(Expectation of support response time) minus (Average support response time)	✓		
	Days since last contact	Number of days since last contact, calculated per entry			✓
	Difference in customer vs analyst expected response time	(Expectation of support response time) minus (Expected analyst response time)			✓
	Decay of Information * Live Artifacts †				
Customer Profile	Number of open PMRs * †	Number of PMRs owned by customer that are open		✓	
	Number of closed PMRs *	Number of PMRs owned by customer that are closed	✓	✓	
	Number of open CritSits * †	Number of CritSits owned by customer that are open		✓	
	Number of closed CritSits *	Number of CritSits owned by customer that are closed	✓	✓	
	Open CritSit to PMR ratio †	(Number of open CritSits) / (Number of open PMRs)		✓	
	Closed CritSit to PMR ratio	(Number of closed CritSits) / (Number of closed PMRs)	✓		
	Expectation of analyst response time	Average of all “Average support response time” of all PMRs owned by a customer	✓		
Support Analyst Profile	Number of open PMRs * †	Number of PMRs owned by customer that are closed			✓
	Number of closed PMRs *	Number of PMRs owned by the analyst that are closed			✓
	Number of open CritSits * †	Number of CritSits owned by the analyst that are open			✓
	Number of closed CritSits *	Number of CritSits owned by the analyst that are closed			✓
	Open CritSit to PMR ratio †	(Number of open CritSits) / (Number of open PMRs)			✓
	Closed CritSit to PMR ratio	(Number of closed CritSits) / (Number of closed PMRs)			✓
	Expected analyst response time	Average of all “Average support response time” of all PMRs owned by an analyst			✓

* in the last N weeks, where N = ∞ , 12, 24, 36, and 48

Improved Features



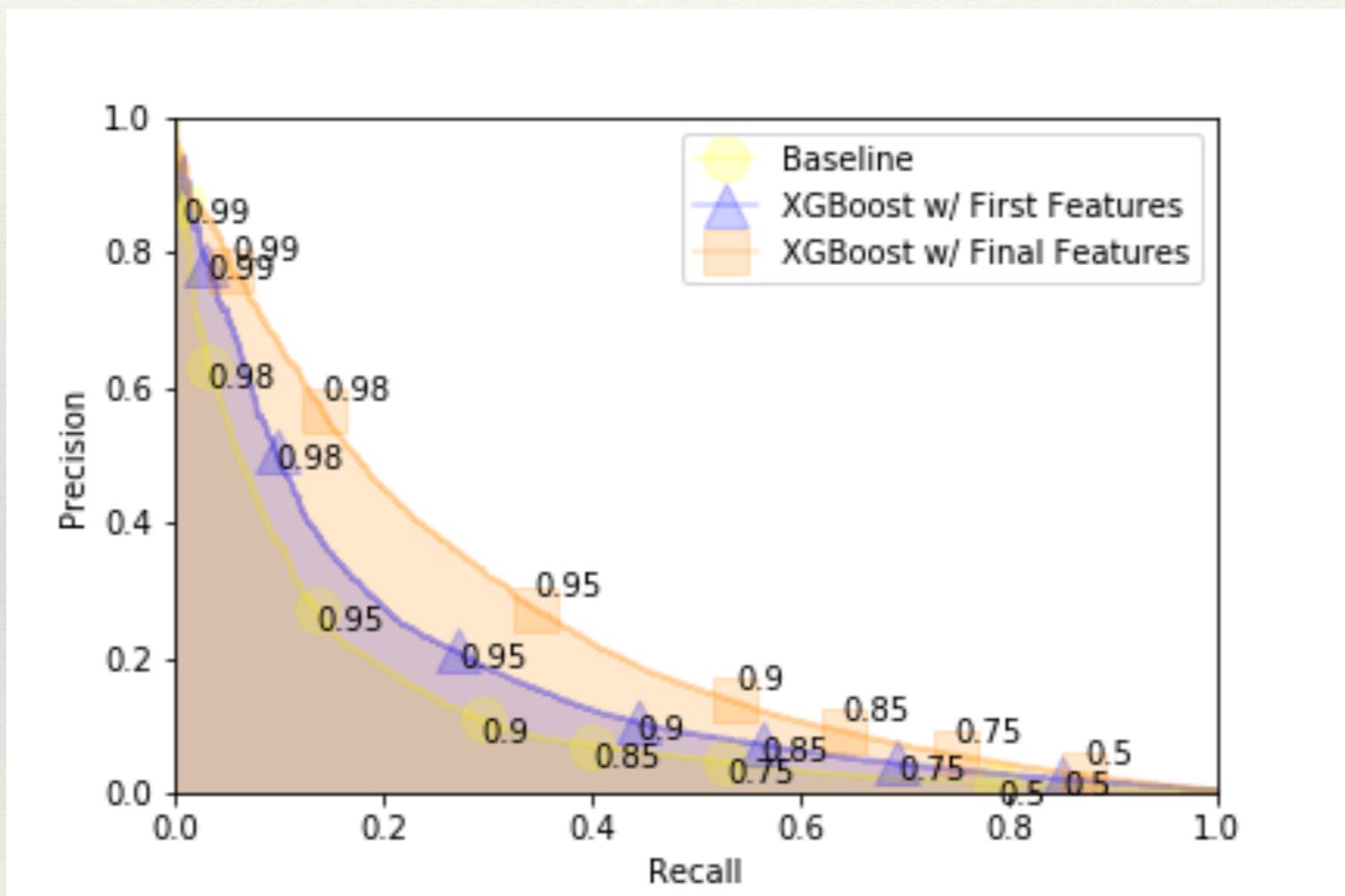
Improved Results

Actual	Total	Predicted as	
		Escalation - No	Escalation - Yes
Escalation - No	2,532,745	2,242,064 (TN) 88.52%	290,681 (FP) 11.48%
Escalation - Yes	9,536	1,205 (FN) 12.64%	8,331 (TP) 87.36%

Accuracy 87.94% **Recall** 87.36% **Precision** 2.79% **Summarization** 88.23%

Accuracy 80.49% **Recall** 79.94% **Precision** 1.65% **Summarization** 80.79%

Comparison to Baseline



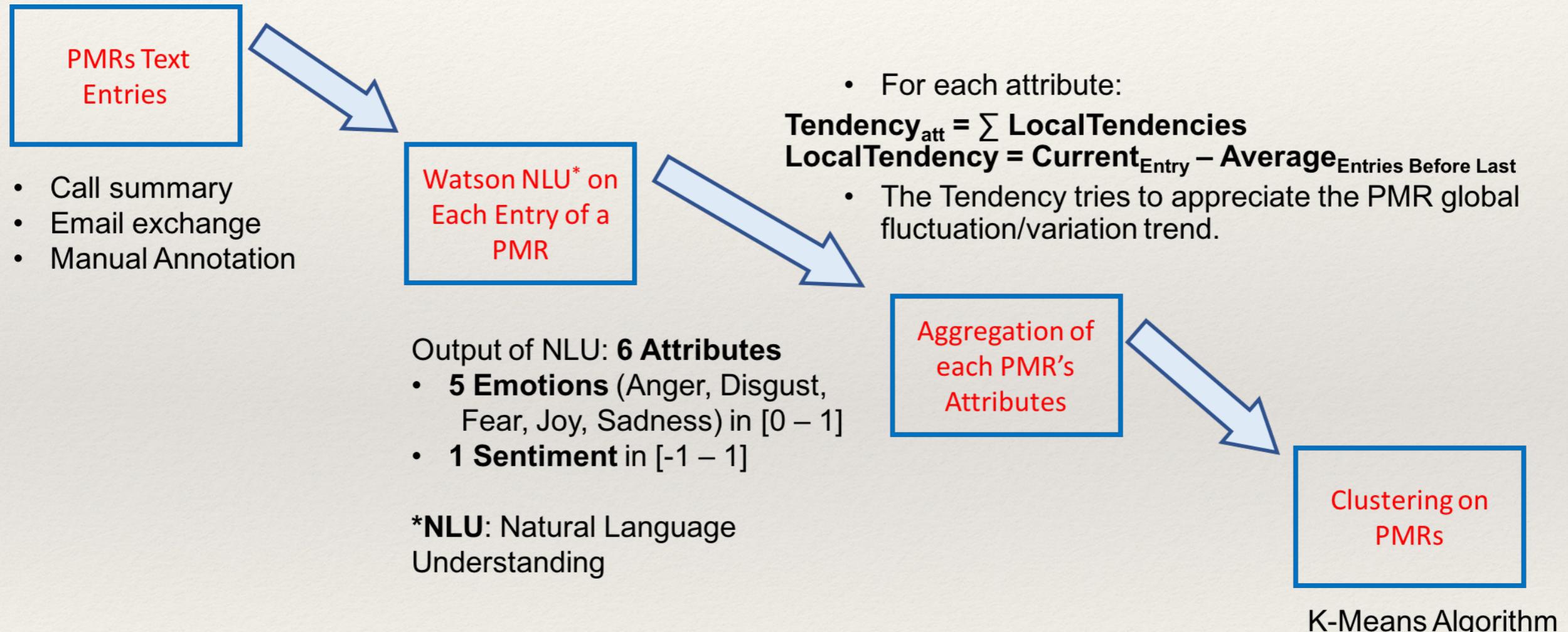
Research Questions

RQ 1. Are the emotions of customers significantly different during support tickets that escalate versus during support tickets that do not escalate?

RQ 2. Are the trends in the emotions of customers significantly different during support tickets that escalate versus during support tickets that do not escalate?

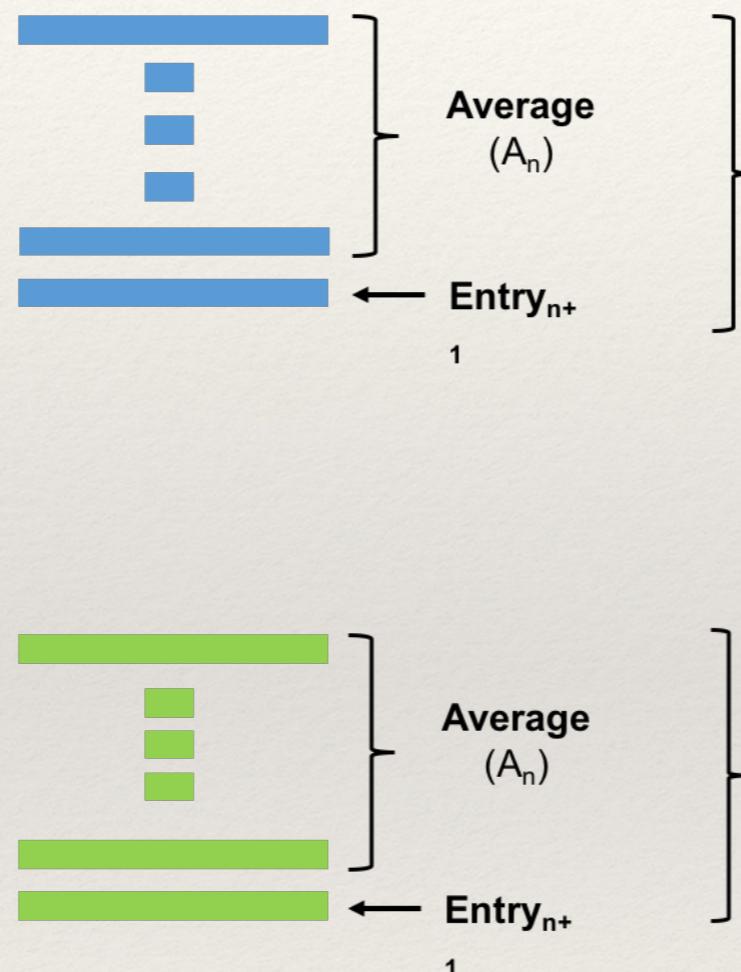
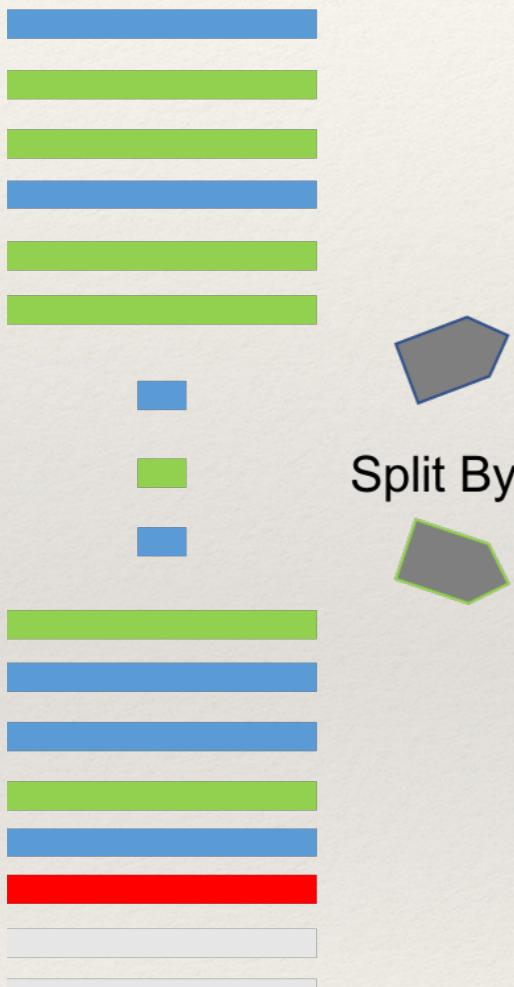
RQ 3. Can these differences in emotions be utilized to assist support analysts in understanding which customers are likely to escalate their support tickets?

Watson Natural Language Understanding

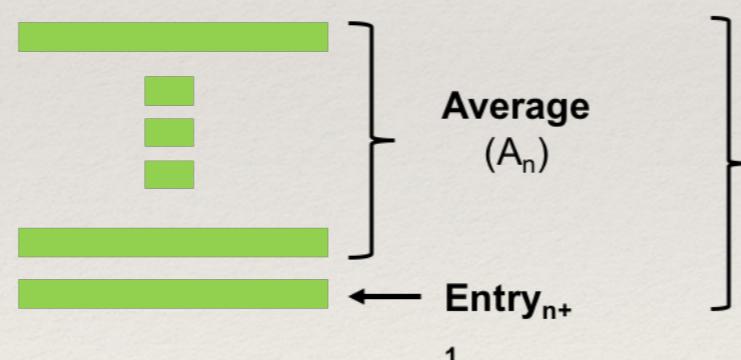


Tendency

PMR Entries



LocalTendency = Entry_{n+1} - A_n
On customer entries



LocalTendency = Entry_{n+1} - A_n
On support entries

█ Customer

█ Support

█ Crit

Difference Testing: NLU Emotions

Customer Crit Analysis						
	# PMRs	# CRs	Avg CRs/PMR			
	208	3253	15.64			
	Anger	Disgust	Fear	Joy	Sadness	Sentiment
Skew	3.21942	5.62226	3.72404	1.59225	0.95461	-0.01712
Kurtosis	15.31188	38.71298	22.66040	1.42776	-0.42217	-0.96636
D'Agostino-Pearson p-value	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Shapiro-Wilk p-value	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Customer NonCrit Analysis						
	# PMRs	# CRs	Avg CRs/PMR			
	94	324	3.45			
	Anger	Disgust	Fear	Joy	Sadness	Sentiment
Skew	3.65583	7.06570	4.30747	1.33707	0.92829	-0.14786
Kurtosis	17.36818	69.35238	25.57274	0.33010	-0.52496	-0.96109
D'Agostino-Pearson p-value	0.00000	0.00000	0.00000	0.00000	0.00000	0.00098
Shapiro-Wilk p-value	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Customer Crit vs NonCrit	Anger	Disgust	Fear	Joy	Sadness	Sentiment
Mann-Whitney p-value 2-tailed	0.017498009	0.002043608	0.11755407	0.82824515	0.64600566	0.006840317
p <	0.05	0.005				0.01

Difference Testing: NLU Emotions

Support Crit Analysis						
	# PMRs	# CRs	Avg CRs/PMR			
	240	6015	25.06			
	Anger	Disgust	Fear	Joy	Sadness	Sentiment
Skew	3.60804	6.85180	4.24604	1.31043	1.11244	-0.29744
Kurtosis	19.50157	60.04080	31.06732	0.31019	-0.07140	-0.80999
D'Agostino-Pearson p-value	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Shapiro-Wilk p-value	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
Support NonCrit Analysis						
	# PMRs	# CRs	Avg CRs/PMR			
	113	580	5.13			
	Anger	Disgust	Fear	Joy	Sadness	Sentiment
Skew	3.26043	7.19992	4.10390	1.11259	1.37023	-0.54942
Kurtosis	16.34400	74.23851	27.82842	-0.31806	0.52460	-0.34665
D'Agostino-Pearson p-value	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Shapiro-Wilk p-value	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Support Crit vs NonCrit	Anger	Disgust	Fear	Joy	Sadness	Sentiment
Mann-Whitney p-value 2-tailed	0.39195	0.54057	0.43421	0.85616	0.00369	0.00000
p <					0.005	0.001

Difference Testing: Tendency

Customer Crit Analysis						
	# PMRs	# CRs	Avg CRs/PMR			
	208	3253	15.64			
	Anger	Disgust	Fear	Joy	Sadness	Sentiment
Skew	0.36150	-1.90652	-3.31477	-1.56725	-0.68063	-0.74375
Kurtosis	4.73224	30.05383	27.92736	8.64653	6.13592	7.26335
D'Agostino-Pearson p-value	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Shapiro-Wilk p-value	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Customer NonCrit Analysis						
	# PMRs	# CRs	Avg CRs/PMR			
	94	324	3.45			
	Anger	Disgust	Fear	Joy	Sadness	Sentiment
Skew	-2.39825	-2.35297	-2.98650	-0.26061	-0.41145	0.34563
Kurtosis	15.82966	12.03922	21.48420	0.99622	1.90652	0.89205
D'Agostino-Pearson p-value	0.00000	0.00000	0.00000	0.07476	0.00014	0.07392
Shapiro-Wilk p-value	0.00000	0.00000	0.00000	0.09992	0.00072	0.19200
Customer Crit vs NonCrit	Anger	Disgust	Fear	Joy	Sadness	Sentiment
Mann-Whitney p-value 2-tailed	0.00172	0.09817	0.22529	0.27127	0.79452	0.00149
p <	0.005					0.005

Difference Testing: Tendency

Support Crit Analysis						
	# PMRs	# CRs	Avg CRs/PMR			
	240	6015	25.06			
	Anger	Disgust	Fear	Joy	Sadness	Sentiment
Skew	0.26215	-2.13930	-3.34973	-1.62834	-0.71857	-0.77493
Kurtosis	5.18491	29.62222	29.09388	8.70218	6.81533	7.98232
D'Agostino-Pearson p-value	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Shapiro-Wilk p-value	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Support NonCrit Analysis						
	# PMRs	# CRs	Avg CRs/PMR			
	113	580	5.13			
	Anger	Disgust	Fear	Joy	Sadness	Sentiment
Skew	-2.50322	-2.57349	-3.06617	-0.19679	-0.48804	0.37181
Kurtosis	16.60973	14.69686	23.25213	1.15703	2.52528	1.36107
D'Agostino-Pearson p-value	0.00000	0.00000	0.00000	0.02566	0.00000	0.00278
Shapiro-Wilk p-value	0.00000	0.00000	0.00000	0.06440	0.00002	0.04266
Support Crit vs NonCrit	Anger	Disgust	Fear	Joy	Sadness	Sentiment
Mann-Whitney p-value 2-tailed	0.00039	0.06832	0.27152	0.48716	0.81915	0.00033
	0.001					0.001

Preliminary ML Results

Random Classifier

```
print('Random Classifier')
print('Precision: {:.2f}%'.format(sum(y) / len(y) * 100))
print('Recall    : 50.00%')
print('Summ.    : 50.00%)
```

```
Random Classifier
Precision: 67.98%
Recall    : 50.00%
Summ.    : 50.00%
```

Gaussian Naive Bayes

```
n_run_m_cv_train_test_output(GaussianNB(), 10, 10, name='GaussianNB')
```

```
GaussianNB: 10-Fold, Avg of 10 Runs
Precision: 81.16% (+/- 13.19)
Recall    : 42.63% (+/- 20.19)
```

Preliminary ML Results

Logistic Regression

```
n_run_m_cv_train_test_output(LogisticRegression(), 10, 10, name='LogisticRe
```

LogisticRegression: 10-Fold, Avg of 10 Runs
Precision: 67.94% (+/- 4.62)
Recall : 95.87% (+/- 14.87)

SVM

```
n_run_m_cv_train_test_output(svm.SVC(), 10, 10, name='SVM - SVC')
```

SVM - SVC: 10-Fold, Avg of 10 Runs
Precision: 67.99% (+/- 1.53)
Recall : 100.00% (+/- 0.00)

```
n_run_m_cv_train_test_output(svm.NuSVC(), 10, 10, name='SVM - NuSVC')
```

SVM - NuSVC: 10-Fold, Avg of 10 Runs
Precision: 69.86% (+/- 11.60)
Recall : 80.97% (+/- 18.76)

```
n_run_m_cv_train_test_output(svm.LinearSVC(), 10, 10, name='SVM - LinearSVC')
```

SVM - LinearSVC: 10-Fold, Avg of 10 Runs
Precision: 67.72% (+/- 5.26)
Recall : 95.85% (+/- 16.58)

Suggestions?

