



# Order from Chaos: Using Natural Language Processing to Find Patterns in Unstructured Text Survey Responses

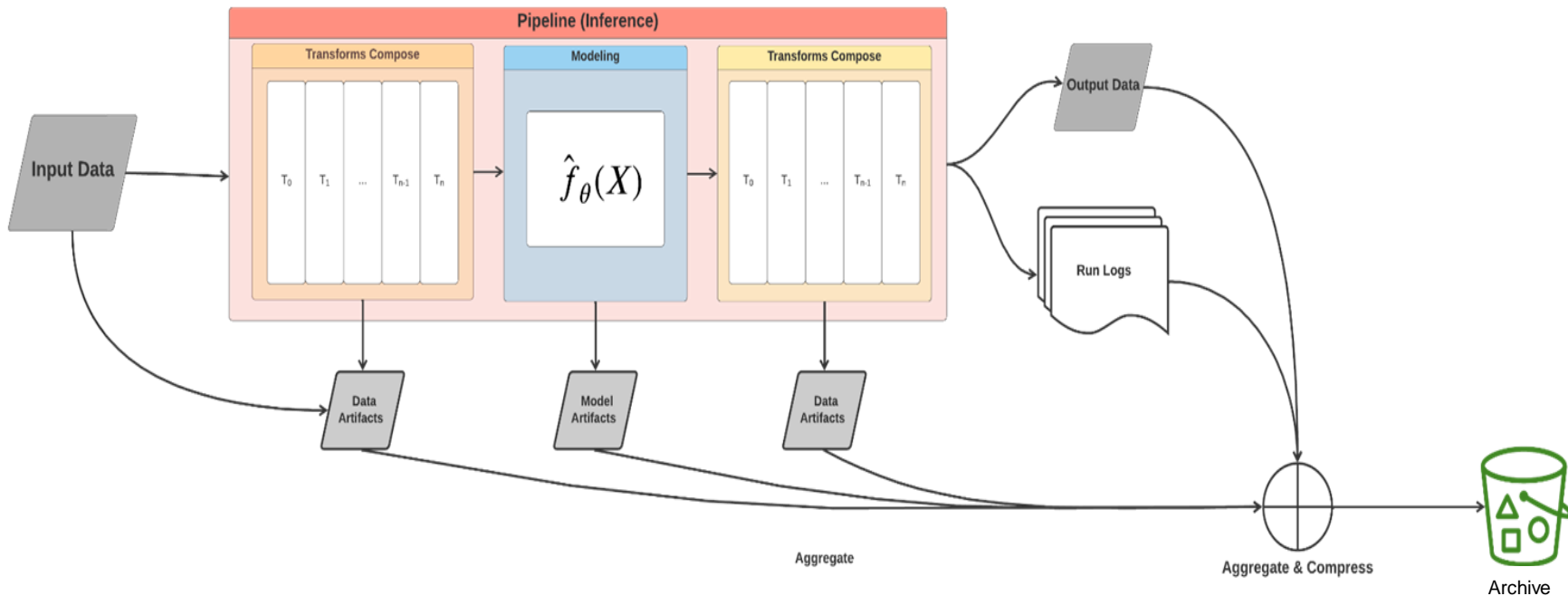
William Skorupski & Karl Konz  
FedCASIC presentation  
April 6, 2022

# Introduction

---

- Leverage DRC's Artificial Intelligence Engine to analyze unstructured text using Natural Language Processing (NLP)
- Use open-ended comments as a rich source of data to add context to quantitative analyses
- Review Two Research Studies using NLP
  1. Automate detection of comments implying urgent response
  2. Estimate explanatory models for evaluating theme comments' prediction of Net Promotor Scores

# DRC Artificial Intelligence Engine



# Context

---

- Analyses were conducted for a major automotive manufacturer who administers Customer Loyalty Surveys to customers who have purchased a vehicle or had service work done at retail locations (literally millions of comments over several years)
- Likert-type questions to gauge customer loyalty, satisfaction, likelihood of recommendation (Net Promotor Score), and open-ended comments for additional feedback

# Natural Language Processing

---

- General term for a family of modeling approaches to understand and quantify different aspects of unstructured text responses
- Extract meaning, translate text, evaluate semantics and themes, and identify patterns
- Sentiment analysis to understand comment valence and a deeper context for “customer satisfaction”

# Outcomes

- For correct identification of valence or prediction of NPS, we evaluated overall Accuracy, Recall, and Precision:

Outcome	Formula	Notes
Accuracy	$\frac{TP + TN}{N}$	Proportion of correct classifications
Recall	$\frac{TP}{TP + FN}$	Proportion of positive cases correctly classified
Precision	$\frac{TP}{TP + FP}$	Proportion of positive classifications which were accurate

# Study 1: “Hot Comments”

---

- Testing dataset: 35,000+ surveys were coded by expert human scorers to identify themes in the open-ended responses
- Specific goal: identify unhappy customers so that immediate action can be taken
- Manual processing requires too much time, makes it impossible to act promptly

# Methods

---

- Hot comments identified via a ***transformer model***; focus on specificity vs. sensitivity
  - Transformer models evaluate sequence-to-sequence unstructured text strings for context, meaning
  - Identify comments which indicate recommended follow-up



# Comment Processing



## Overview

The purpose of the AI-Driven Comment Processing was to automate hot comment processing by using an AI-Engine to automatically screen comments for “alerts”



## Benefits

- High Specificity for screening out “non-alert” comments  
Accuracy: 97.2%; 98.7% recall; 98.3% precision
- Automatically screen 80% of total comments in less than 1 hour of processing time (up to 9,000 comments an hour)
- Significantly reduce manual steps, labor hours, and cycle time associated with processing hot comments
- Incorporation of hot comment detection leads to getting alerts in front of retailers as quickly as possible, allowing them to act quickly

# Study 2: Explain NPS

---

- Respondents are asked the likelihood that they would recommend [retailer] to others
- Specific goal: identify thematic aspects of customer experience that predict high NPS
- Determine best practices for customer interactions, use in feedback loop to retail locations

# Methods

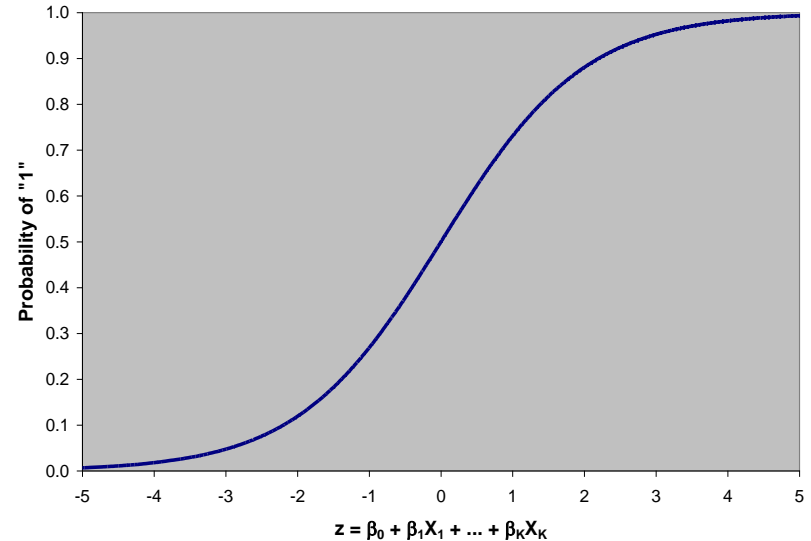
---

- Binary NPS values subjected to a **logistic regression** with several thematic independent variables as predictors
- Predictions based on detection of thematic valence of responses, to deliver feedback on “what works” in terms of customer service

# Logistic Regression

$$P_i = \frac{e^{z_i}}{1 + e^{z_i}}, \text{ where}$$
$$z_i = b_0 + b_1 X_{i1} + \dots + b_K X_{iK}$$

Probability of high NPS (expected value) is a function of the predictors (X) and model coefficients (b)



79.6% overall accuracy  
-- 86.6% Recall  
-- 87.2% Precision

# Results – “What works”

---

- Knowledgeable, thorough retailers
- Reliability of the experience during and across visits
- Personal connections

# Summary

---

- NLP approaches very successful in identifying valence and polarity of unstructured text comments
- Predictions of NPS allows continuous, automated feedback for enhanced customer service



**Thanks!**

[wskorupski@datarecognitioncorp.com](mailto:wskorupski@datarecognitioncorp.com)