Master Modeler Competition Round 1

Joey Hodson and Laird Stewart
Emory University
March 1, 2021

U of SC | South Carolina

ERASE CHILD TRAFFICKING

Agenda

- ❖ Introduction
- ❖ Data
- ❖ Model
- ❖ Conclusion

Introduction

Joey Hodson

Laird Stewart

Objectives

"Leverage social media to reach individuals in need of assistance for leaving a human trafficking situation."

A. Determine the types of posts that are most successful

B. Estimate the value of each type of engagement (likes, etc.)

C. Evaluate post attributes (use of celebrities, etc.) that tend to drive more positive post results

Agenda

Data Cleaning

01 *Cleaning* - missing values, data types, duplicates, etc..

02 *Filtering* - subsetting data to highlight information

Data Manipulation

01 *Hot Encoding* - qualitative variables to binary

02 *Time Matching* - utilizing time as a variable of data

03 *Custom Features* - combining content for additional value
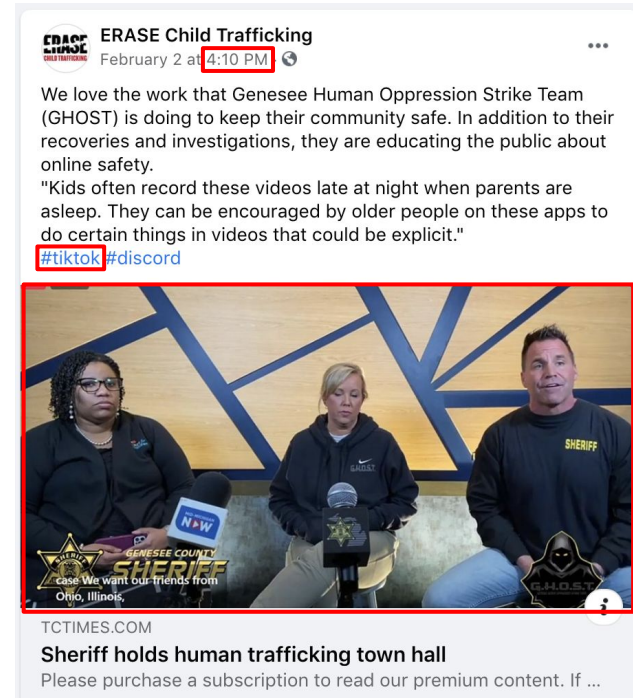
04 Aggregation - merging disparate data sets

05 *Dimensionality Reduction* - for model training purposes

Post Attribute Analysis

❖ Sentiment Analysis

❖ Length

❖ Endorsements

❖ Media Form (video/image/etc.)

❖ Post Type (informational, personal, etc.)

❖ Time

❖ Hashtag Use

❖ Etc...

**ERASE Child Trafficking**
February 2 at 4:10 PM

We love the work that Genesee Human Oppression Strike Team (GHOST) is doing to keep their community safe. In addition to their recoveries and investigations, they are educating the public about online safety.
"Kids often record these videos late at night when parents are asleep. They can be encouraged by older people on these apps to do certain things in videos that could be explicit."
#tiktok #discord

TCTIMES.COM
Sheriff holds human trafficking town hall
Please purchase a subscription to read our premium content. If ...

Agenda

Metric Weighting

1. Engagement Rate by Reach (ERR)
2. Engagement Rate by Post (ERP)
3. Engagement Rate by Impressions (ERI)
4. Daily Engagement Rate (DER)
5. Engagement Rate by Views (ERV)
6. Factored Engagement Rate (FER)
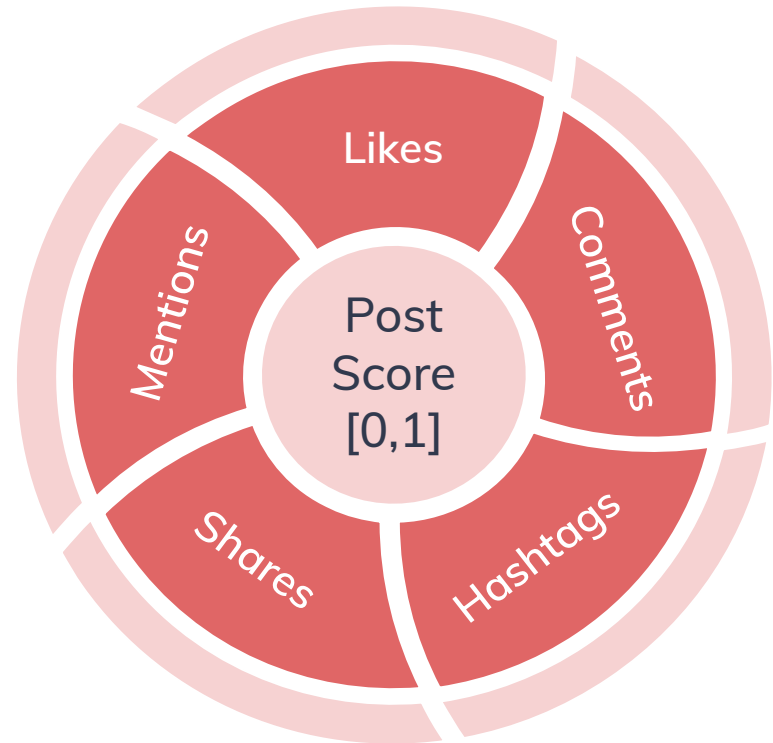
Machine Learning Model Development

Model objective:
Determine which attributes have the greatest impact on Post Score [0,1]

Model inputs:
Post attributes, Post Score [0,1]

Model output:
Coefficients of attributes

Resources:
Python 3.9, Scikit-learn, etc...

Agenda

Conclusion

With this analysis and modeling we hope to improve ERASE Child Trafficing's ability to increase awareness of human and child trafficking, reduce the level of trafficking, and reach individuals in need of assistance for leaving a human trafficking situation.