

Social Media as a Service Assurance Tool

Legal Notice

- ❑ Please note that the Twitter data provided for this project is only to be accessed or downloaded by SheHacks participants
- ❑ This data is only to be retained for the duration of the SheHacks event
- ❑ At the conclusion of the event, participants must delete all copies of the original and derived data

Project Overview

The purpose of this project is to develop automated processes to detect and analyze problems with Bell Services using Social Media.

Some examples of insights that might be valuable include:

1. Sentiment analysis to determine whether comments are positive, negative or neutral
2. Classification of the comments based on natural clustering found within the data
3. Spam detection/Troll Busting to identify comments that are not relevant
4. Anomaly detection to determine whether large service disruptions can be identified

Of course, these are only suggestions! You may find other (or additional) types of insights in the data.

Data Acquisition

Two files are available for analysis:

- bell_support.txt includes actual Twitter data related to Bell's support account (@Bell_Support)
- bell.txt includes actual Twitter data related to Bell's general account (@Bell)

To obtain the url and logon credentials for the server on which these files are stored, please email peter.paton@bell.ca with the text "SheHacks 2019 File Transfer Request" in the subject line. In response, you'll receive an email with the required information.

Assessment

Projects will be evaluated based on the following criteria:

1. Creativity/Originality (30%)
2. Technical Complexity (30%)
3. Presentation/Visualization/Wow-effect (20%)
4. Insights discovered (20%)

Tool Options

- Participants are welcome to select and use their own preferred tools
- For participants new to machine learning but experienced with Python, the following frameworks and libraries might be a useful place to start:
 - NLTK (Natural Language Toolkit)
 - Scikit-learn
 - Keras
 - Statsmodel
 - Textblob
 - Seaborn (visualization)
- Students who would prefer to use R (instead of Python) can leverage the following packages and libraries:
 - Tm or OpenNLP
 - Tsoutlier
 - Caret or Keras for R
 - Forecast
 - Ggplot (visualization)

Are you new to Machine Learning?

Participants who are new to the field are encouraged to attend Bell's presentation at 11 am.