

Project Charter: Spotify Twitter Analysis



<https://twitter.com/Spotify>

Background

Spotify is aware of the enormous benefits in increasing discussion volume, increasing the number of likes and increasing the number of reshares, leading to a bigger audience and more potential customers. Spotify already have a tool for Twitter Sentiment Analysis, but it is currently lacking for Spotify's business.

The project is targets to provide a better, more efficient tool.

This kind of support tool needs to stay relevant and find strategies to advance the business, analysing the customers general opinions on the company or on company specific products.

The tool aims to collect the customer "sentiment" better and constantly, considering that customer "sentiment" is essential to win the competition with other similar service providers.

This proposal focuses on the following types of Sentiment Analysis:

- ✓ Standard Sentiment Analysis: categorizes subjective details into positive, neutral and negative.
Examples:
 - positive: "I love Spotify"
 - neutral: "Spotify is a streaming platform"
 - negative: "Spotify sucks"

This analysis provides information about the customers' reactions to new functions, promotions, artists and posts in general.

- ✓ Aspect-Sentiment Analysis: focuses on certain aspects of an offered product by analysing the customers' feedback about it.

This type of analysis gathers information about certain aspects of Spotify. A set of functions or aspects of the App, used to get information about the product, are selected and for each of them a detailed customer analysis is performed as to obtain information about the user satisfaction and about possible improvements.

Scope

Two parties are involved in the project. The supplier HW Analytics and the client Spotify. The project aims at developing a sentiment analysis tool for Spotify's Twitter page, supporting Spotify's business decisions. Nowadays social media plays an essential role in public discourse; hence Spotify is aware that Sentiment Analysis represents for its business an essential tool for obtaining information, other than playing a central role in the interactions with customers.

The tool is a standalone application. There is no graphical user interface, as the reports are created automatically on a weekly basis. Spotify does not have to execute anything; you will receive a weekly report via email. A set of functions or aspects of the tool, which is executed automatically, are selected and for each of them a detailed customer analysis is performed as to obtain information about the user satisfaction and about possible improvements.

At the completion of the project a full Sentiment Analysis Tool will be released. The tool consists of an automatic script, written in Python, that will be scheduled on a weekly or monthly basis, delivering a detailed report. The tool provides these three services:

1. **Brand Monitoring:** All texts about Spotify are scraped and stored weekly in the company database. Then follows data analysis making it possible to observe certain trends and the overall reaction to actions. This will be achieved by performing language analysis and categorizing the tweets.
2. **Customer Support Analysis:** Twitter interactions between the brand Spotify and its customers is monitored and analysed weekly, giving a detailed report of the effectiveness of support and the customers' satisfaction.
3. **Market Research:** Analysis of music and technology trends can be run, enabling to make contracts with upcoming musical artists and to implement the newest technologies before competitors do so. This analysis is done by monitoring and analysing the most popular hashtags.

Data are automatically scraped and analysed weekly and the analysis outcomes are reported in a document file, and this file is sent by email.

Goals

Spotify will receive weekly reports (textual and graphic) providing a detailed report of their twitter page including comments, interactions, and the overall reactions of its followers.

1. **Goal:** Spotify's Twitter posts within a period of 1 week are scraped and analysed. All responses are categorized into positive, negative, or neutral. For each category, the most common words are identified and the most important (most likes, retweets) reactions can be consulted.
2. **Goal:** Tweets containing Spotify's name are scraped and analysed. A report containing interesting observations, even graphically (graphs and diagrams), is created.
3. **Goal:** The Twitter interaction between Spotify and its customers is analysed and a detailed report is created.
4. **Goal:** Relevant Tweets (most interactions: likes, retweets, comments) that contain the words "music", "streaming", "Spotify" are analysed and detailed textual reports are created, containing suggestions on possible actions to take.

Goals 1 and 2 are implemented, Goals 3 and 4 will be implemented in a second phase.

The opportunity to test and measure Goal 1 will be offered, delivering a POC.

The POC will be based on a Python notebook which connects to the Twitter API and then scrapes Spotify's tweets from the last 6 months and categorises the responses into positive, negative, or neutral. Each category the most common words are identified. Important reactions according to number of likes and retweets are identified. The analysis produces a textual report in the form of a PDF document.

Metrics

Many brands pursue the path of Sentiment Analysis. One example is Uber. A [case study](#) shows that they were able to achieve 100% growth in discussion volume, 90% growth in number of likes and 250% increase in number of shares due to paying attention to Sentiment Analysis during the process of developing a new app.

Personnel involved

- HW Analytics:
 - Project Manager: Hannes Wiedenhofer
 - Data Scientist 1: Christian Ronaldo (web scraping)
 - Data Scientist 2: Gianluigi Chiellini (python)
 - Data Scientist 3: Giorgio Buffon (data visualization)
 - Developer 1: Tony Halford (python)
 - Developer 2: Rob Iommi (web services)
 - Account Manager: Andrea Nedved
 - System Administrator: Pavel Pirlo
 - Lawyer: Leonardo Messi
- Spotify:
 - Data Administrator: Bruce Araya
 - Business Contact: Tom Dickinson
 - CTO: Danny Petrucci
 - Data Management Analyst: John Carey

Key Stakeholders

| | |
|-----------------|--------------------------------|
| Client | Spotify |
| Sponsor | Social Media Manager (Spotify) |
| Project manager | Hannes Wiedenhofer |

Project Milestones

| Milestone | Date |
|--|------------|
| M1: All comments categorised, Textual reports finished | 25.03.2021 |
| M2: Graphical report for Goal 2 finished | 25.03.2021 |
| M3: Graphical report for Goal 3 finished | 20.04.2021 |
| M4: Textual report for Goal 4 finished | 20.04.2021 |

Project Budget

Estimated Project Budget: 70.000€

Data Architecture

All data is obtained from Twitter using its API and [Tweepy](#). After cleaning and organizing the data is stored on our servers. The analysis process uses the prepared data. This is done on a weekly basis. The result of the analysis will be reported in different documents, each one stored on our server. Reports are automatically created and sent to Spotify via email. We are responsible for the data maintenance and privacy on our servers. We respect the GDPR norms. We apply the policy to remove old data from our servers before new reports are created.

Data shall be modelled as here reported:

| | date and time | userid | comment |
|-----|------------------------|---------------------|---|
| 0 | 2020-12-16 00:18:55 | 1339002027121184768 | Take a second to kick back and let @GrangerSmi... |
| 1 | 2020-12-15 22:04:34 | 1338968219172794369 | @Prerana38094169 @karolg Nope 😊 https://t.co/7... |
| 2 | 2020-12-15 21:43:57 | 1338963030118440960 | Drain Gang, check in. @BladeeCity and @mechato... |
| 3 | 2020-12-15 17:41:27 | 1338901999929135107 | RT @SpotifyCanada: Going #OnTheRecord about hi... |
| 4 | 2020-12-15 15:28:03 | 1338868430389141512 | Put. The. Kettle. On.\n\nA new holiday special... |
| ... | ... | ... | ... |
| 145 | 2020-12-03 15:02:23 | 1334513315976974339 | @taylorswift13 @jackantonoff We're all august ... |
| 146 | 2020-12-03 14:32:53 | 1334505892872523776 | @scorpions You rocked us like a hurricane 🙌❤️ |

Constraints, Assumptions, Risks and Dependencies

| | |
|-------------------------------|--|
| Constraints | <ul style="list-style-type: none"> • If Spotify's Twitter page is inactive, the tool will not have any data to work with, therefore resulting in empty reports. • The tool does not include a graphical user interface. • Twitter Dev Account can only be set up on March 3rd because of vacation of SysAdmin • Requirement from Spotify: Data Quality verification must be done by March 11th • Project must be finished by July, 2nd |
| Assumptions | <ul style="list-style-type: none"> • A reference person within Spotify is identified. This person should respond to questions from our site, within 3 working days. • We are dealing with sensitive data (algorithmic decision making), so GDPR directives must be followed. • We will be immediately notified if Spotify makes changes to their Twitter page (name changes, change of availability). |
| Risks and Dependencies | <ul style="list-style-type: none"> • Spotify must provide an email address for receiving the weekly reports. Spotify is responsible for the privacy and security of their inbox. |

Approval Signatures

| | | |
|----------------|-----------------|-----------------|
| _____ | _____ | _____ |
| Project Client | Project Sponsor | Project Manager |