DATA 607 Presentation - Sentiment Analysis

David Simbandumwe

2021-10-04

Introduction

In my last position one of my responsibilities was to advocate for our digital services with our customers. One of the tools we used was email campaigns and mass mailings. My team was responsible for monitoring the responses to our advertising activity and forwarding any complaints to our compliance team to remediation. This was a manual process that took a substantial amount of my team time.

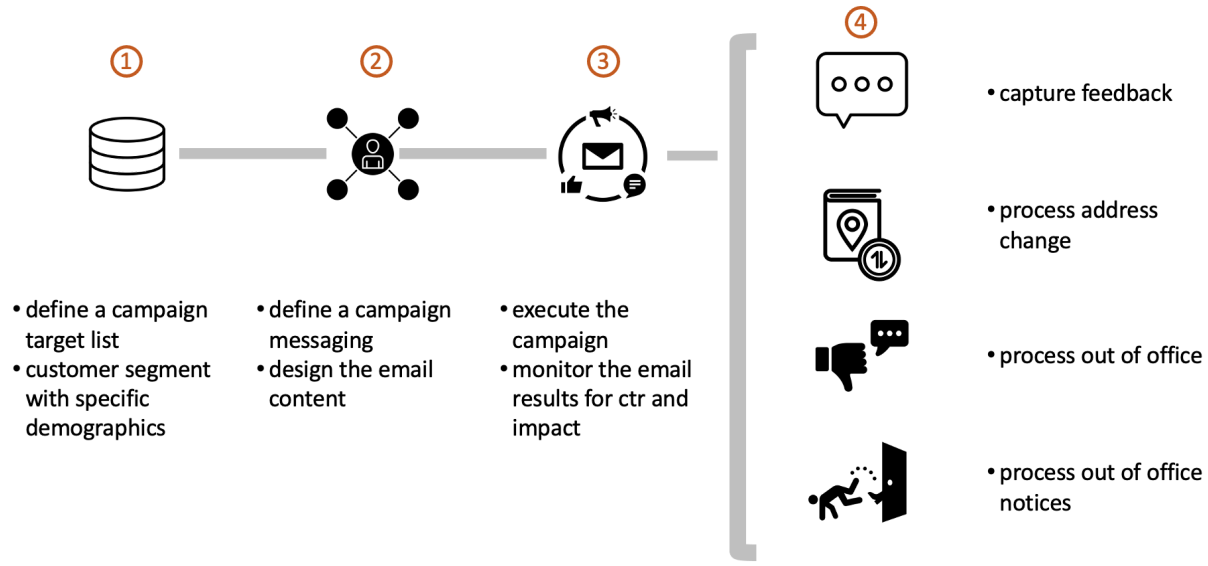

Figure 1: Advertising process diagram

process the file data

read all text emails and process the

```
files <- list.files(email_path )
sentiment_df <- tibble(
  email = character(),
  sentiment = double(),
  method = character()
)
```

```
email_df <- tibble(
  email = character(),
  event = character(),
  related_email = character()
)


for (i in files) {

  fileText <- GetFile(i,email_path)
  fromEmail <- GetToEmail(fileText)

  # get sentiment
  s_df <- GetSentiment(fileText, fromEmail)
  s_df <- s_df %>%
    mutate (
      email = fromEmail
    ) %>%
    select (email, sentiment, method)
    sentiment_df <- union(s_df, sentiment_df)




  # get other emails
  rel_df <- GetRelatedEmails(fileText, fromEmail, known_emails)
  rel_df <- tibble(rel_df) %>%
    rename (related_email = str) %>%
    mutate (
      email = fromEmail
    )
    email_df <- union(email_df, rel_df)



}
```

```
## Joining, by = "word"
## Joining, by = "word"
## Joining, by = "word"
## Joining, by = "word"
## Joining, by = "word"
## Joining, by = "word"
## Joining, by = "word"
## Joining, by = "word"
```

Conclusion

After processioning the 8 emials in the folder we generated the following sentiment scores for each email.

```
# sentiment analysis
sentiment_df
```

```
## # A tibble: 8 x 3
##   email                               sentiment method
##   <chr>                                   <dbl> <chr>
## 1 john.brown@cnt1.foodtalkdaily.com          -2 AFINN
## 2 sam@cnt1.foodtalkdaily.com                 -1 AFINN
## 3 amy@aol.com                                10 AFINN
## 4 grandpa@aol.com                            10 AFINN
## 5 jeff.smith@yahoo.com                       -3 AFINN
## 6 addam.madison@yahoo.com                    -6 AFINN
## 7 braddyh@google.com                         -1 AFINN
## 8 bradh@google.com                           -1 AFINN
```

The 8 emails contained the following secondary emails. Processing the text in the email we can map the new email address to an event.

```
# email output
email_df
```

```
## # A tibble: 4 x 3
##   email                               event          related_email
##   <chr>                               <chr>          <chr>
## 1 bradh@google.com                    change address bob.smith@cuny.com
## 2 braddyh@google.com                  change address gram.olsky@nyu.com
## 3 sam@cnt1.foodtalkdaily.com          other          other_email@gmail.com
## 4 john.brown@cnt1.foodtalkdaily.com   other          rapha@gmail.com
```