Customer Feedback

Promothesh Chatterjee[1]

Customer Feedback Analysis Using R

This write-up explains a script that analyzes customer feedback data. The script performs various tasks including loading necessary libraries, generating sample data, extracting themes and topics from the feedback, and fitting a regression model to understand the impact of these topics on customer satisfaction. Let's break down each step of the script.

Step 1: Load Necessary Libraries

```
library(dplyr)
library(lubridate)
library(readr)
library(quanteda)
library(topicmodels)
library(tidytext)
library(ggplot2)
```

These libraries are essential for data manipulation (`dplyr`), date handling (`lubridate`), reading data (`readr`), text analysis (`quanteda`, `tidytext`), topic modeling (`topicmodels`), and data visualization (`ggplot2`).

Step 2: Generate Sample Data

```
set.seed(123)
customer_ids <- 1:76
dates <- sample(seq(as.Date('2024-01-01'), as.Date('2024-06-30'), by="day"),
76, replace = TRUE)
```

This section generates sample data with customer IDs and random dates between January 1, 2024, and June 30, 2024.

```
feedbacks <- c(
  "The camera quality is outstanding, but the battery drains quickly when
using it.", # 8
  "I love the sleek design, but the phone feels a bit slippery without a
case.", # 7
  "Great performance overall, but the phone heats up during heavy gaming
sessions.", # 6
  "The user interface is very intuitive, though I would appreciate more
customization options.", # 9
  "Excellent build quality, but it would be nice if the screen was more
scratch-resistant.", # 8
  "Fast charging is a lifesaver, but the phone could use better thermal
management.", # 7
  "I miss the expandable storage option, especially for such a high-end
device.", # 5
  "Superb sound quality, but the fingerprint sensor is hit or miss at
times.", # 6
  "The battery life is fantastic, but the phone feels a bit heavy.", # 7
  "Customer service was very helpful when I had an issue with the software.",
# 9
  "The display is stunning, but I wish it was slightly larger.", # 8
```

```
  "Having dual SIM support is great, but the second SIM slot should support
5G.", # 5
  "The operating system is smooth and fast, but it lacks some essential apps
out of the box.", # 6
  "Good phone overall, but I feel like the price is a bit high for the
features offered.", # 5
  "I appreciate the water resistance, but the charging port cover feels
flimsy.", # 6
  "Great speaker quality, but the maximum volume could be a bit higher.", # 7
  "I love the new design, but it feels slippery without a case.", # 6
  "Fast processor, but I noticed some lag when switching between apps.", # 5
  "The built-in apps are useful, but they take up a lot of storage space.", #
5
  "Battery life is good, but it drains quickly when using certain apps.", # 5
  "The camera is amazing in daylight, but it struggles in low light
conditions.", # 7
  "I like the facial recognition feature, but it's not very reliable in low
light.", # 6
  "The screen resolution is top-notch, but the refresh rate could be
higher.", # 8
  "The phone is very durable, but the glass back is prone to fingerprints.",
# 6
  "The software updates are infrequent, which is disappointing for such an
expensive phone.", # 4
  "I love the color options available, but they seem to scratch easily.", # 6
  "The phone's performance is great, but the storage capacity is limited.", #
6
  "The UI is clean and minimalistic, but it lacks some customization
features.", # 5
  "Great value for money, but the lack of a headphone jack is inconvenient.",
# 5
  "The camera quality is top-notch, but the app could be more user-
friendly.", # 7
  "I like the fast charging feature, but the battery life could be better.",
# 6
  "The phone's build quality is excellent, but it feels a bit bulky.", # 6
  "I appreciate the fast processor, but the RAM management could be
improved.", # 6
  "The sound quality is superb, but the volume buttons are hard to press.", #
7
  "The display is bright and clear, but it could be more power-efficient.", #
7
  "The phone's design is sleek, but the camera bump is quite large.", # 8
  "The battery lasts all day, but the phone gets warm during extended use.",
# 7
  "The fingerprint sensor is fast, but it doesn't work well with wet
fingers.", # 6
  "The phone's software is very user-friendly, but it has a lot of
bloatware.", # 6
  "The build quality is very solid, but the phone is a bit on the heavier
```

```
side.", # 6
  "The camera performs exceptionally well, but the video stabilization could
be improved.", # 7
  "The phone's battery life is excellent, but it takes a while to fully
charge.", # 7
  "The display is vibrant, but the bezels are a bit thick.", # 6
  "The phone's performance is smooth, but it occasionally lags when
multitasking.", # 6
  "The camera takes great pictures, but the shutter speed is a bit slow.", #
7
  "The phone's design is elegant, but it shows fingerprints easily.", # 7
  "The battery life is impressive, but the phone heats up during long
calls.", # 7
  "The phone's interface is intuitive, but it could use more personalization
options.", # 8
  "The display is sharp, but it could be more responsive to touch.", # 6
  "The phone feels sturdy, but the buttons are hard to press.", # 5
  "The sound quality is great, but the speaker placement is awkward.", # 7
  "The battery lasts long, but the charging port is flimsy.", # 4
  "The phone's camera is excellent, but the app crashes sometimes.", # 5
  "The design is sleek, but the phone is slippery without a case.", # 6
  "The phone's performance is great, but the screen scratches easily.", # 5
  "The battery life is good, but the phone is quite heavy.", # 5
  "The sound quality is superb, but the volume could be louder.", # 7
  "The display is clear, but the brightness could be better.", # 6
  "The phone's interface is user-friendly, but it lacks some advanced
features.", # 6
  "The camera takes beautiful photos, but the low-light performance is
poor.", # 5
  "The phone's build quality is solid, but it feels bulky.", # 6
  "The battery lasts a long time, but the phone heats up during gaming.", #
5,
  "The display is bright, but it drains the battery quickly.", # 5,
  "The phone's design is stylish, but the camera protrudes too much.", # 7,
  "The camera quality is excellent, but the app is not very intuitive.", # 6,
  "The battery life is impressive, but it takes too long to charge.", # 7,
  "The phone feels durable, but the back panel is prone to scratches.", # 6,
  "The sound quality is fantastic, but the speaker placement is not ideal.",
# 7,
  "The display is vivid, but the refresh rate could be higher.", # 7,
  "The phone's performance is outstanding, but it has limited storage.", # 7,
  "The camera performs well, but the app interface needs improvement.", # 6,
  "The battery lasts all day, but the phone gets warm during extended use.",
# 7,
  "The design is elegant, but it shows fingerprints easily.", # 7,
  "The phone's interface is smooth, but it could be more customizable.", # 7,
  "The display is sharp, but the bezels are a bit thick.", # 6,
  "The phone feels sturdy, but the buttons are hard to press." # 5
)
```

```
# Corresponding satisfaction scores
satisfaction_scores <- c(
  8, 7, 6, 9, 8, 7, 5, 6, 7, 9, 8, 5, 6, 5, 6, 7, 6, 5, 5, 5,
  7, 6, 8, 6, 4, 6, 6, 5, 5, 7, 6, 6, 7, 7, 8, 7, 6, 6, 6, 7,
  7, 6, 6, 7, 7, 7, 8, 6, 6, 7, 7, 4, 5, 6, 5, 5, 7, 6, 6, 5,
  6, 5, 6, 7, 6, 7, 6, 7, 7, 7, 6, 5, 7, 4, 5, 6
)
```

This code block contains a list of 76 feedback comments along with corresponding satisfaction scores.

```
feedback_data <- data.frame(
  customer_id = customer_ids,
  feedback = feedbacks,
  date = dates,
  satisfaction_score = satisfaction_scores
)
```

Here, we create a dataset combining customer IDs, feedback, dates, and satisfaction scores.

Step 3: Extract Themes and Topics

Custom Stopwords

```
custom_stopwords <- c("phone", "phones", "phone's", "quality")
all_stopwords <- c(stopwords("en"), custom_stopwords)
```

We define custom stopwords related to the word “phone” and combine them with default English stopwords.

Text Preprocessing

```
corpus <- corpus(feedback_data$feedback)
tokens <- tokens(corpus, remove_punct = TRUE, remove_numbers = TRUE,
remove_symbols = TRUE, remove_separators = TRUE) %>%
  tokens_tolower() %>%
  tokens_remove(all_stopwords)
dfm <- dfm(tokens)
```

This section converts the feedback comments into a corpus and tokenizes the text, removing punctuation, numbers, symbols, and stopwords.

Topic Modeling

```
lda_model <- LDA(dfm, k = 5, control = list(seed = 1234))
topics <- tidy(lda_model, matrix = "beta")
```

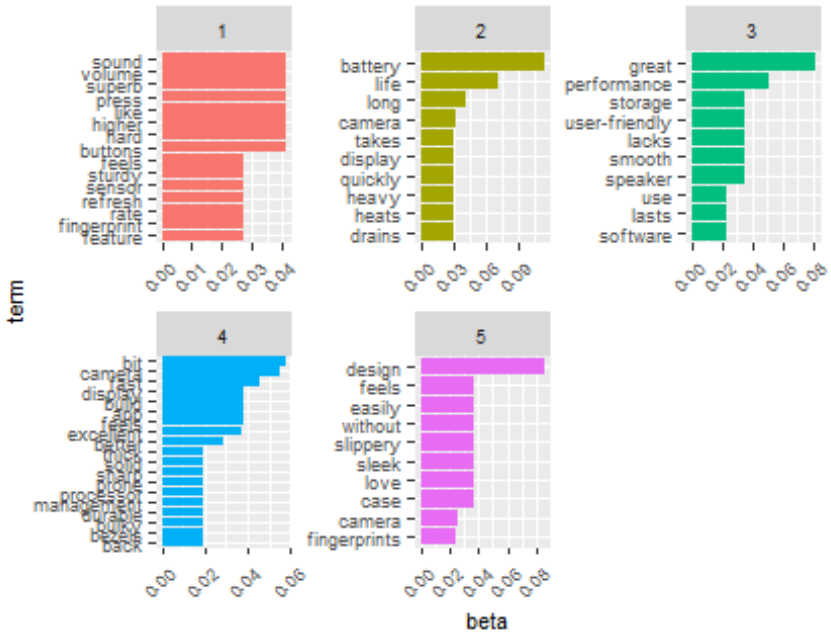
We perform Latent Dirichlet Allocation (LDA) to extract 5 topics from the text data. The `tidy` function converts the LDA model output into a tidy format.

Top Terms for Each Topic

```
top_terms <- topics %>%
  group_by(topic) %>%
  top_n(10, beta) %>%
```

```
  ungroup() %>%
  arrange(topic, -beta)

top_terms %>%
  mutate(term = reorder_within(term, beta, topic)) %>%
  ggplot(aes(term, beta, fill = as.factor(topic))) +
  geom_col(show.legend = FALSE) +
  facet_wrap(~ topic, scales = "free") +
  coord_flip() +
  scale_x_reordered() +
  theme(
    text = element_text(size = 8),
    axis.text.x = element_text(angle = 45, hjust = 1),
    plot.margin = margin(1, 1, 1, 1, "cm")
  )
```



This code block identifies the top 10 terms for each topic and visualizes them using `ggplot2`.

Step 4: Use Insights to Direct Product Development Efforts

Regression Analysis

```
topic_probabilities <- posterior(lda_model)$topics
colnames(topic_probabilities) <- paste0("Topic_",
1:ncol(topic_probabilities))
regression_data <- cbind(feedback_data, topic_probabilities)
```

```r
model <- lm(satisfaction_score ~ ., data = regression_data[, -c(1, 2, 3)])
summary(model)
```

```
##
## Call:
## lm(formula = satisfaction_score ~ ., data = regression_data[,
##     -c(1, 2, 3)])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.4695 -0.5910 -0.1277  0.6194  2.5984
##
## Coefficients: (1 not defined because of singularities)
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   6.4059     0.3200  20.017   <2e-16 ***
## Topic_1       0.1933     0.4681   0.413    0.681
## Topic_2       0.1385     0.4329   0.320    0.750
## Topic_3      -0.6333     0.4309  -1.470    0.146
## Topic_4      -0.2839     0.4228  -0.671    0.504
## Topic_5           NA         NA      NA       NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.066 on 71 degrees of freedom
## Multiple R-squared:  0.07045,    Adjusted R-squared:  0.01808
## F-statistic: 1.345 on 4 and 71 DF,  p-value: 0.2618
```

We extract topic probabilities for each document and combine them with the original data. Then, we fit a regression model to quantify the impact of each topic on customer satisfaction.

Avoiding Multicollinearity

```r
model <- lm(satisfaction_score ~ Topic_1 + Topic_2 + Topic_3 + Topic_4, data
= regression_data)
summary(model)
```

```
##
## Call:
## lm(formula = satisfaction_score ~ Topic_1 + Topic_2 + Topic_3 +
##     Topic_4, data = regression_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.4695 -0.5910 -0.1277  0.6194  2.5984
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   6.4059     0.3200  20.017   <2e-16 ***
## Topic_1       0.1933     0.4681   0.413    0.681
## Topic_2       0.1385     0.4329   0.320    0.750
```

```
## Topic_3      -0.6333     0.4309  -1.470    0.146
## Topic_4      -0.2839     0.4228  -0.671    0.504
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.066 on 71 degrees of freedom
## Multiple R-squared:  0.07045,    Adjusted R-squared:  0.01808
## F-statistic: 1.345 on 4 and 71 DF,  p-value: 0.2618
```

To avoid multicollinearity, we exclude one topic from the regression model and fit it again. The summary provides insights into which topics significantly impact customer satisfaction.

This comprehensive analysis helps in understanding customer feedback, identifying key themes, and using these insights to inform product development and improvements.