

BEYOND STAR RATINGS

**USING TEXT ANALYSIS TO EXTRACT USER'S TEXTUAL OPINION
TOWARD LEADING E-COMMERCE SAAS - SHOPIFY**

April 2024

Huei-Wen Cheng

Yu-Ting Tseng

TABLE OF CONTENTS

01

Introduction

- 1. Background
- 2. Problem statement

02

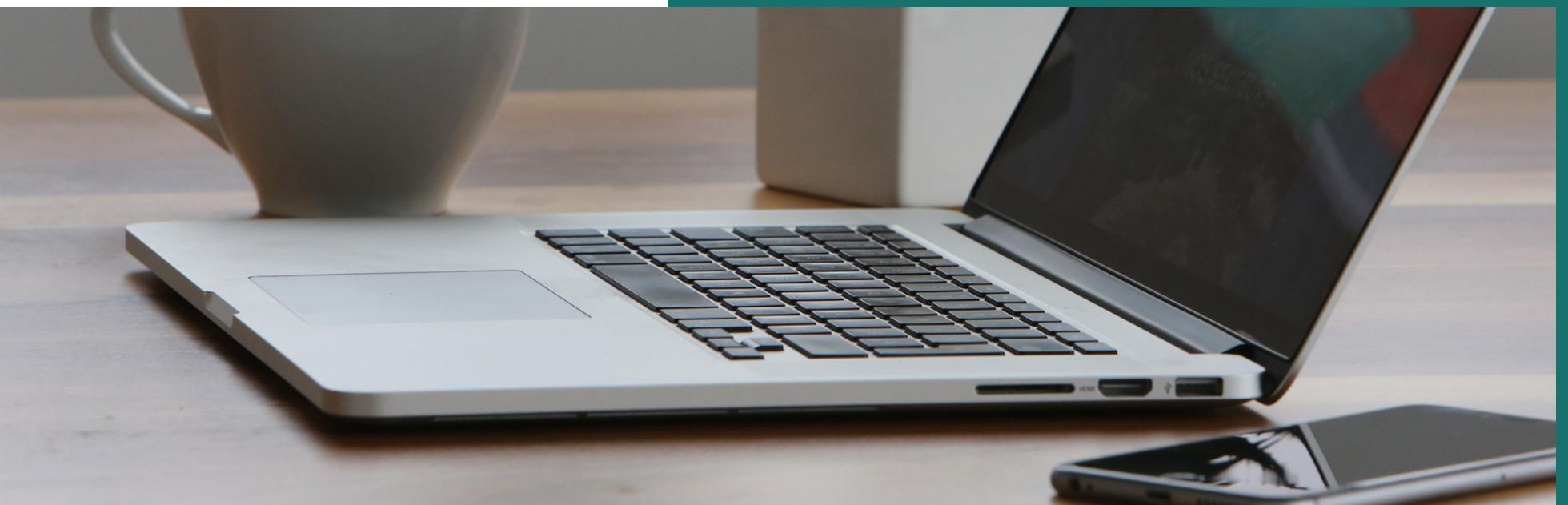
Methods & Process

- 1. Method
- 2. Process

03

Conclusion

- 1. Experiential result
- 2. Challenges
- 3. Limitation
- 4. Future advancement



shopify

INTRODUCTION

- Background
- Problem Description



BACKGROUND

THE CHALLENGE OF NAVIGATING E-COMMERCE REVIEWS

01

Lack of efficient navigation of relevant information

Business owners now have to consult various sources when selecting an e-commerce platform.

02

Lack of filtering and categorizing based on key concerns

Considering online reviews is important, but most review systems lack the ability to filter information.

03

Lack of good user experience

Users have to manually search through a vast amount of reviews to find relevant information, such as user interface, payment options, and pricing.



15 years helping Canadian businesses
choose better software

Search for sof

Shopify Reviews

★★★★★ 4.5 (6,282) Write a Review!

LOCAL PRODUCT



VIEW PROFILE



Marco

Owner in US

Health, Wellness & Fitness, 2–10 Employees

Used the Software for: 1+ year

Reviewer Source ?

Alternatives Considered:

Square Point of Sale



Rey

Owner in US

Consumer Electronics, Self Employed

Used the Software for: 1+ year

Reviewer Source ?

Shopify works!

★★★★★ 5.0 4 weeks ago

Comments: Extremely happy with problems. Very very pleased.

+ Pros:

Ease of use, never have had one safe, secure, people are able to na

- Cons:

Kinda expensive to keep up with t

Ease of use and profession

★★★★★ 5.0 4 weeks ago

+ Pros:

The ease of use. I literally learned code writing background

- Cons:

The price is a little bit high if you really cheap

PROBLEM DESCRIPTION

CATEGORIZING USER COMMENTS
FOR EFFICIENT DECISION-MAKING

We aim to identify recurring topics or specific features highlighted in user comments related to Shopify, finding what are the functions or issues the users discuss most frequently in their comments.



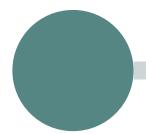
.METHOD & PROCESS

- Process
- Method



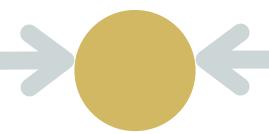
PROCESS

WHAT WE DID



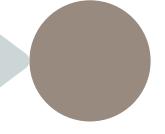
Data Collection

- Tools: Python
- Source: Capterra (Business Software Review Site)



Data Cleaning

- Tools: Excel



Data Analysis

- Tools: RapidMiner

DATA COLLECTION APPROACH

Scrape user reviews for Shopify from Capterra website

```
: from bs4 import BeautifulSoup
from webdriver_manager.chrome import ChromeDriverManager
from selenium.webdriver.common.by import By
from selenium import webdriver

: # Retrieve post content
browser = webdriver.Chrome()

# Page URL to get data
url = 'https://www.capterra.ca/reviews/83891/shopify'

# Navigate to the page
browser.get(url)
```

Set up environment

- Import required libraries
- Initialize Chrome webdriver
- Define target URL

DATA COLLECTION

```
# Use Selenium to retrieve the page source
page_source = browser.page_source

# Pass web page source to BeautifulSoup
soup = BeautifulSoup(page_source, 'html.parser')

# Find all elemets with <div class="col-lg-7">
review_divs = soup.find_all('div', class_='col-lg-7')

# Iterate through each comment div and extract content
all_reviews = []
for review_div in review_divs:
    review_content = review_div.get_text(strip=True)
    all_reviews.append(review_content)

# The all_reviews list now contains all comment content
print(all_reviews)
```

Retrieve page source

- Use Selenium to retrieve the page source
- Pass page source HTML to BeautifulSoup
- Find <div> elements with class = "col-lg-7"
- Extract and store revies text content

DATA COLLECTION

["Shopify is the best way to start ecommerce5.02 years agoComments:I ought to say this is the stage that I might want to use for my online business clients. Presumably that they are working really hard structure. I trust Shopify start a respectable client assistance that we can call and allow us to pose inquiries as a paying client.Pros:So natural to simply import all my client's items and the day after that large number of items where on Google's most memorable page that astounded me. Extremely simple to arrangement and use. I use it for my client's internet based shop.Cons:Where do I started? – Only a couple of free site topic! truly?? indeed, even we're burning through \$30/month we need to pay like \$300 to get a respectable subject! – Although the vast majority (counting myself) thought there was such countless instructional exercises on web however there isn't sufficient. A large portion of them are being made by the more established rendition and doesn't match the new look. – Google's page speed test is continuously showing lower speed contrasting with different stages like Wordpress. – No client assistance which isn't adequate", "Good platform, great value4.02 months agoNewComments:Overall, Shopify has been a great app for my company, and we've only ever used it. It just has boosted our sales so much with its clean interface and easy product tracking functionalities.Pros:I like that Shopify is by far the platform that makes it easiest for you to manage your products and shop overall, it also gives great analytics details of what your customers look at before they make a purchase.Cons:The thing I disliked most about Shopify is the lack of ease in customizing the User Experience of your shop, it is doable, but it is not super easy for the average person to figure out.", "If you want to have problems, choose this platform2.0last monthNewComments:catastrophic, I spent 3 months building a site that I will likely have to redo elsewhere!Pros:I really see none in English Or perhaps if you are going to learn coding since their so-called experts constantly leave you alone with problems.Cons:If you want a site with just three pages and no interesting features, it might be okay, but as soon as you add the slightest functionality, nothing works anymore. It's a very poorly functioning mess. I made the mistake of thinking, after seeing an advertisement, that I would leave WordPress for a Shopify site and not have to deal with technical aspects anymore, what a mistake! Since I've been with them, I've had n,

DATA COLLECTION

```
import openpyxl
```

```
base_url = 'https://www.capterra.ca/reviews/83891/shopify'  
total_pages = 252  
all_reviews = []  
  
# Create a new Excel workbook  
workbook = openpyxl.Workbook()  
worksheet = workbook.active  
  
# Set titles for worksheets  
worksheet['A1'] = 'Reviews'
```

Create a new Excel workbook using openpyxl

DATA COLLECTION

```
row = 2

for page_num in range(1, total_pages+1):
    page_url = base_url if page_num == 1 else f"{base_url}?page={page_num}"

    browser.get(page_url)

    page_source = browser.page_source

    soup = BeautifulSoup(page_source, 'html.parser')

    review_divs = soup.find_all('div', class_='col-lg-7')

    # Fetch all comments on the current web page
    # Strip is used to clear whitespace in the middle of the text
    for review_div in review_divs:
        review_content = review_div.get_text(strip=True)
        all_reviews.append(review_content)

    # Write comments to Excel worksheet
    worksheet.cell(row=row, column=1, value=review_content)
    row += 1

print(f"Comments for page {page_num} have been successfully retrieved")
```

Iterate through multiple pages

- Use a loop to iterate through each page URL
- Pass page source HTML to BeautifulSoup
- Find <div> elements with class = "col-lg-7"
- Extract and store reviews text content

DATA COLLECTION

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	Reviews																
2	Don't Shop Around, Shop-ify!	5.04 years ago	Comments:	Overall, Shopify has been good to me. They made me a Partner because of the – you know I can't actually recall what the criteria was to become or													
3	Good platform, great value	4.0 last month	NewComments:	Overall, Shopify has been a great app for my company, and we've only ever used it. It just has boosted our sales so much with its clean interface a													
4	If you want to have problems, choose this platform	2.03 weeks ago	NewComments:	catastrophic, I spent 3 months building a site that I will likely have to redo elsewhere! Pros:I really see none in English C													
5	Shopify works!	5.02 weeks ago	NewComments:	Extremely happy with it. So glad we changed from word press which had lots of problems. Very very pleased.Pros:Ease of use, never have had one single pro													
6	Ease of use and professional look	5.02 weeks ago	NewPros:	The ease of use. I literally learned the ins and outs of shopify in a few weeks without having a code writing backgroundCons:The price is a little bi													
7	Shopify is good for Retail online and offline ecommerce	5.0 last month	NewComments:	Overall we are happy with shopify , just think the cost maybe abit steep. every new app we add in we have extra cos													
8	Good but difficult to customize for our needs	4.0 last month	NewComments:	In general, we are happy with our website, but it would be nice to have more control on what changes can be madeit can get e													
9	Best e-commerce platform out there	5.02 months ago	NewPros:	I have used Shopify for over a year now and I have moved all my clients e-commerce sites to Shopify now. I develop my own themes via sho													
10	My Shopify Experience	5.0 last month	NewPros:	I love the function of having to detail my products in an orderly manner and infuse specifications too.Cons:I haven't yet figured that out. For now, there ar													
11	Amazing app	5.0 last month	NewPros:	I love the way it makes me connect to million of people in the work and it gives you the chance selling to customers that you don't need to know themCons:Shopify i													
12	Shopify for the win!	4.04 years ago	Comments:	I love that shopify will scale with my business and that there are various pricing plans that give you different options and shipping discounts. It's easy to use													
13	Shopify is easy to launch and manage	5.04 years ago	Comments:	My overall experience has been positive using Shopify as my shopping cart for my store. I felt the transition from Magento was pretty simp													
14	Why I chose Shopify (but also use other platforms, too)	5.04 years ago	Comments:	I'm a fan - their customer service is excellent, and I've never had any issues with downtime or anything like that.Pros:I ha													
15	The Best E-commerce Platform	5.03 years ago	Comments:	I'm very satisfied with Shopify. It's really the best in its class, and I have recommended it regularly to friends and family members who are looking													
16	They promise the moon and provide a crumbling chunk of space rock	1.03 years ago	Comments:	I couldn't move away from this platform fast enough.I lost 7 months of effort and about \$3000 cash dealin													
17	Best e-commerce Website Hosting Platform	5.03 years ago	Comments:	Our clients are extremely happy with Shopify as a website hosting platform and have seen positive increases in their business from h													
18	Shopify Fit Our Needs Perfectly And Continues To Impress	5.02 years ago	Comments:	Until the end of 2020 we had a legal forms business, selling paper and electronic forms both in a storefront retail setti													
19	Shopify, Great for Running Your Buisness	4.02 years ago	Comments:	Shopify has overall been very wonderful for my business and myself. I am able to track manage ship all products to customers without :													
20	OK but MUCH room for improvement	4.05 years ago	Comments:	2014 =great2018 =terribleI would encourage rep. to contact me.Pros:When I first signed up for Shopify 4-5 years ago I thought it was the													
21	Amazing eCommerce platform for novices and developers alike!	5.05 years ago	Pros:	I've been running an eCommerce shop for 5 years on Shopify. Easy setup and easy import of products from a CSV file. Th													
22	Simple and Easy E-commerce platform!	5.04 years ago	Comments:	When we launched our web store back in 2016 we looking for something simple yet robust. After reviewing with several e-platform soft													
23	Works reasonably well out of the box	4.02 years ago	Comments:	Shopify is a good platform to get you up and running very quickly. It is one big ecosystem that sometimes works in your favour but unfortu													
24	The best for e-commerce	5.0 last year	Comments:	Impressive e-commerce facilities. The best choice to build an online store.Pros:Of the big 3 website builders (Shopify, Squarespace, Wix), Shopify is hands													
25	Let Shopify Power Your E-Commerce Business	5.09 months ago	Comments:	My online business has changed dramatically thanks to Shopify. Because of how simple it is to use, I was able to build a beautifu													
26	Shopify Review	5.03 years ago	Comments:	It has truly been a very wonder asset to my company as well as other companies I am associated with. Its ease to not only display products, get the word out, and													
27	Avoid its terrible now!!!	1.04 months ago	Comments:	Terrible...sad...exhausting...poor...depressing...suicidal....Pros:Ease of setting up a store without coding.Cons:NO support staff anymore they cancel													
28	Small Business Owner's best friend	5.05 months ago	Pros:	Ease of use, robust security, feature rich, clean interface, the app ecosystem.Drag and drop feature. Easy set up of the store by a non technical per													
29	Shopify is a trap	1.08 months ago	Comments:	Using Shopify can be summed up like this: an ongoing and demoralising feeling of I've been tricked and can't do my job without buying yet another plugin an													
30	Streamlined Convenience: Embracing Ecommerce with Shopify	4.09 months ago	Comments:	Shopify has truly revolutionized my online business. With its user-friendly interface, extensive range of customizability, and reliabl													
31	Accelerating Digital Growth with Shopify Website Development and Marketing	5.012 months ago	Comments:	I have had a great experience with Shopify website development and marketing. It has been a game changer for our busines													
32	Servis is good	5.0 last year	Comments:	Although I have listed the minuses quite a bit, it is recommended for a quick start to e-commerce.It speeds up your work, it doesn't let you down.Pros:It's really imp													
33	High cart flexibility and low cost	4.0 last year	Comments:	You can later extend the required functions with third-party applications developed by others. There are numerous applications available, rangin													
34	JAM Shopify Review	5.02 years ago	Comments:	Well, around the beginning of the pandemic, I noticed many stores at physical locations closing. People were panicking with uncertainty of where to look r													
35	Shopify is the Yahoo Stores for a New Generation	3.02 years ago	Comments:	We are only now setting up the store and getting into using the software so my opinion may change over time (although I dou													

Retrieved 6,163 reviews

DATA CLEANING APPROACH

Using Excel to clean the data



Removing
blank values

Filtering out
non-English
reviews

Removing
meaningless
words

- Originally, there were 6,163 records of reviews.
- After removing reviews containing only ratings but null comments or textual feedback, there were 5,235 left.

Based on the results from the preliminary runs: we remove some words such as “years,” “agopros,”, “agocomments.” We also tried out remove “Shopify.”

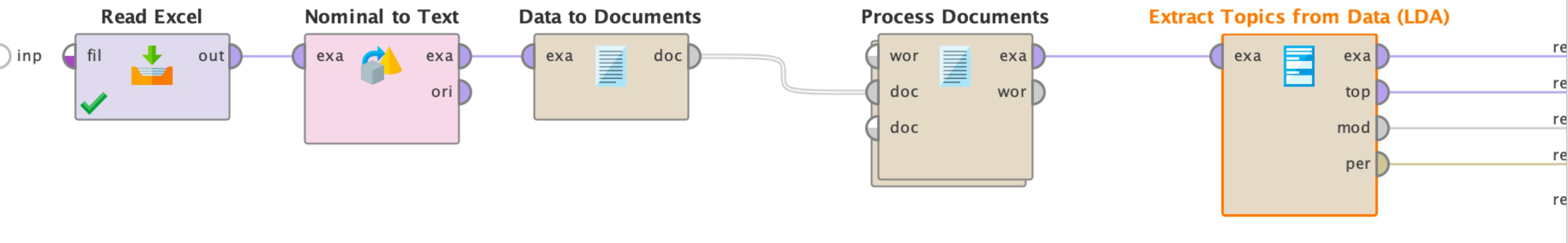
METHOD - LDA

What is LDA (Latent Dirichlet Allocation)?

- LDA is a popular topic modeling technique used in natural processing.
- It helps discover abstract “topics” that occur in a collection of documents.
- Each document is seen as a mixture of various topics, and each topic is seen as a mixture of words.

PROCESS IN RAPIDMINER

Process



01 Read Excel

Input file.

03 Data to Documents

Transform the data (records of texts) to multiple documents. Each record (reviews in our datasets) is transformed to individual documents.

02 Nominal to Text

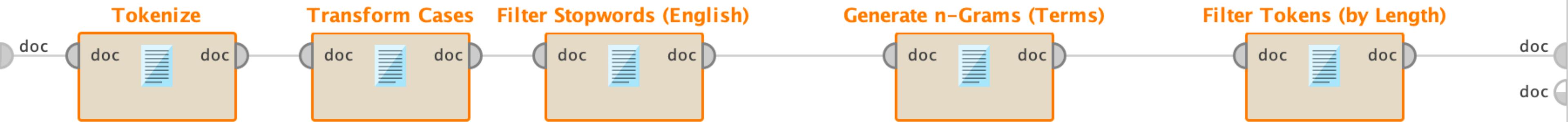
Transform nominal to text for further text processing.

04 Process Documents

Pre-process the data for the topic mining in next step. The operator has layers where we will further explore in the next slide.

PROCESS IN RAPIDMINER

Process Documents



04 Process Documents

4.1 Tokenize

Split the text of documents into tokens.

4.2 Transform Cases

Unify the cases of texts (lower case).

4.3 Filter Stopwords

Remove the English stopwords from the built-in dictionary in Rapidminer.

4.4 Generate n-Grams (Terms)

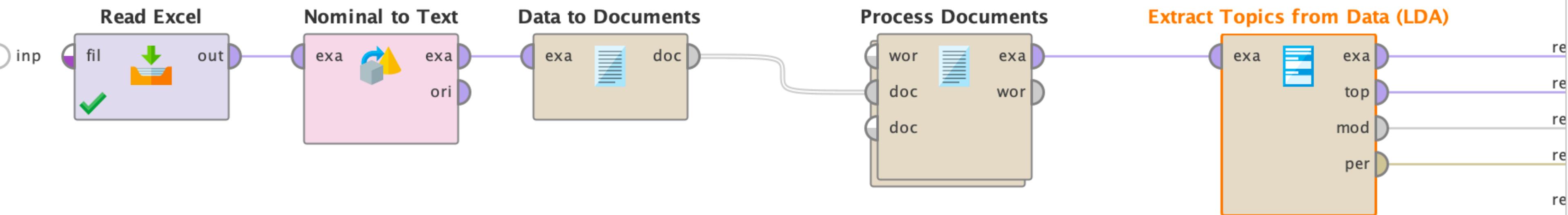
Generate n consecutive numbers of tokens. If set to 1, all tokens (words) stand independently. If set to 2 or more, some phrases can be identified, e.g. easy_use.

4.5 Filter Tokens (by Length)

Filter tokens by setting the minimum and maximum number of characters each token include.

PROCESS IN RAPIDMINER

Process



05

Extract Topics from Data (LDA)

Apply LDA method to identify topics among documents by setting the desired number of topic and number of top words per topic.

Parameters

Extract Topics from Data (LDA)	
text attribute	text
number of topics	4
show optimization settings	<input type="checkbox"/>
iterations	1000
top words per topic	10
stopword language	english

.CONCLUSION

- Experiential results
- Challenges
- Limitations
- Future Advancement



EXPERIENTIAL RESULTS

ANSWERING OUR QUESTION

What are the recurring topics, functions or specific features discussed most frequently in user comments?

Example of Results Generated by RapidMiner

- Number of topics: 4
- Top words per topic: 10

- Tokenize: non-letters vs linguistic tokens
- Pruning: TF-IDF, Percentile, Absolute, etc.

Row No.	topicId	word	weight
1	0	easy	1734
2		store	1728
3		great	1068
4		online	1036
5		platform	888
6		e-commerce	881
7		website	869
8		online_store	770
9		business	764
10		months	761
11	1	easy	2237
12		easy_use	1596
13		platform	1497
14		use	1170
15		great	1145
16		i_love	1077
17		love	1016
18		ecommerce	884
19		store	723
20		business	598
21	2	business	1473
22		easy	1342
23		software	1323
24		overall	1152
25		great	1145
26		review	765
27		store	744
28		experience	666
29		website	663
30		online	648
31	3	easy	1282
32		great	1097
33		store	981
34		days	775
35		online	737
36		business	722
37		platform	622
38		shop	573
39		free	522
40		e-commerce	470

EXPERIENTIAL

RESULTS

ANSWERING OUR QUESTION

What are the recurring topics, functions or specific features discussed most frequently in user comments?

Functions/Features (appeared in several results)

Topic 1:

- inventory
- shipping
- orders
- cart

Topic 2:

- client
- customer
- support

Topic 3:

- app
- plugins
- code
- solution

Topic 4:

- facebook
- wordpress
- instagram

[more results](#)

Opinionated high frequency words

customization_limited

easy_use

easy_use_affordable

bit_limited love

I_love_simplicity

user_friendly

great

easy

easy_use_quick

good helpful

excellent_support incredibly_easy

easy_understand

fantastic

simple

easy_fast



SOME **CHALLENGES** IN OUR WORKING PROCESS

Besides our lack of experience with web-crawling programming and the trial and error learning associated with the operators required for this process, which were relatively manageable, some other challenges emerged...

01

Retrieve perfectly cleaned and organized data

Due to some formatting problems inherent from the website structure within the data retrieved, there were some non-meaningful words that were difficult to remove.

02

Identify optimized parameters to generate meaningful results

Setting different values in each operator, we may get very different outcomes. We trialed different settings in different operators (such as tokenization, pruning options, number of topics, top words per topics, etc.) and compare and contrast to identify some topics from the results.

SOME **LIMITATIONS** IN OUR CURRENT APPROACH

Using RapidMiner to perform LDA topic modeling causes some limitations within RapidMiner's Process Documents operator, which is a crucial step in the LDA workflow

01

Only allow the selection of TF-IDF as the vector creation method

TF-IDF stands for "Term Frequency, Inverse Document Frequency." Since it can be used to score the important words in a document based on the frequency of the terms, highly frequent meaningless words tend to dominate the document's score. It is hard for meaningful information to stand out.

02

Restrict the potential performance of LDA model

Other vector creation methods, such as Word Embedding and Bag-of-Words, which capture semantic relationships and context, are not available within the Process Documents operator. This can impact the quality of generated topics.

POSSIBLE ADVANCEMENT IN THE **FUTURE**

IF TO TAKE ONE STEP FURTHER

Developed a machine learning framework to automatically classify comments and understand their intent

01

Unsupervised Topic Discovery

- Use Python for flexible and powerful topic modeling capabilities
- Leverage external libraries (Gensim, scikit-learn) for more flexibility, which support different vector representations

02

Labeling and Supervised Topic Classification

- Manually label a subset of data using the topics discovered in Phase 1 as labels
- Build a multi-class classification model using deep learning (e.g., CNN, RNN)
- Train on labeled data with topics from Phase 1 as labels



THANK YOU

FOR YOUR NICE ATTENTION

