

# Automatic User Preferences Elicitation: A Data-Driven Approach

Tong Li<sup>1</sup>, Fan Zhang<sup>2</sup>, Dan Wang<sup>1</sup>

<sup>1</sup> Beijing University of Technology, Beijing, China

<sup>2</sup> Institute of Software Chinese Academy of Sciences, China

24<sup>th</sup> REFSQ @ Utrecht, The Netherlands

22nd, March, 2018



# Outline

---



- Background and Motivation
- Related Work
- Proposal
- Evaluation Plan
- Conclusion and Future Work

# Background and Motivation



Develop a particular type of software application



(Start-up)  
company

What features have  
been developed  
for this type of  
applications?

Survey existing  
applications!

More  
than 2  
million  
apps

What features are  
most liked/disliked  
by users

Look into  
application  
reviews!

Numerous  
reviews

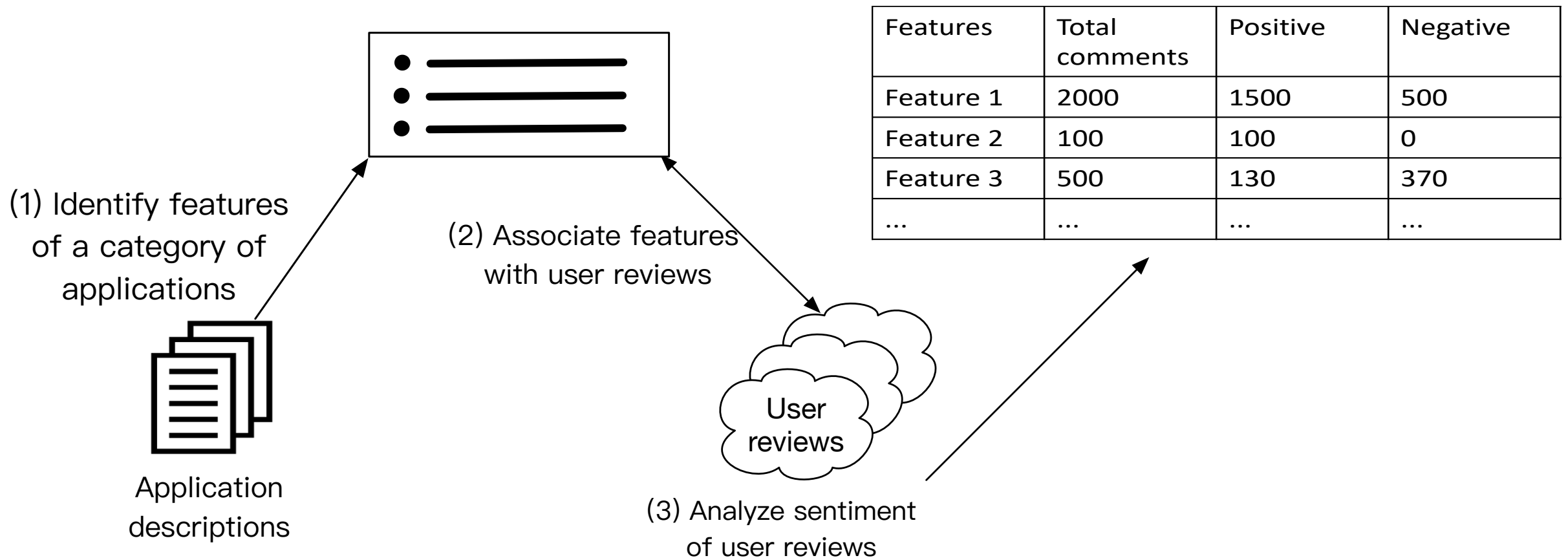
# Related Work

---



- Research on mining user reviews [Carreo2013,Guzman2014]
  - Mining features from user reviews
  - Sentiment-based preference analysis
- Research on mining application descriptions[Hariri2013]
  - Clustering-based feature extraction
  - Association rule-based feature recommendation

# Proposal



# Feature Identification

---



- A Clustering-Based Method
  - Generate clusters (categories): *doc2vec* + density-peak
  - A collocation finding algorithm for identifying features
- Topic Modeling-Based Method

## Description

ZAFUL, your fashion story, creates outstanding styles for adorable you. It has become one of the most popular fashion brands in the world, and more and more young customers have chosen Zaful as their first choice to shape a brilliant recognition or thoughts on fashion. Let you get inspired anytime, anywhere by seeing our adorable, good-quality swimwear, bikinis, one-pieces, tees, blouses, sweaters, sweatshirts, hoodies, jackets, two-pieces outfits, rompers, jumpsuits,

## Description

Poshmark is the leading marketplace to buy and sell fashion.

\*\*\*Shop over 5,000 brands in every size—including plus size, petites, and juniors—at up to 70% off retail

\*\*\*Sell and turn your closet into cash. List your items for free!

Shop. Sell. Style. Party.

### SHOP

Whatever you're shopping for, Poshmark's got it—for much less than at discount clothing retailers and outlets like Nordstrom Rack and TJ Maxx. With over a million dollars worth of new and gently used merchandise added to the marketplace daily, you will find amazing fashion and beauty finds from the hottest brands like Louis Vuitton, Michael Kors, MAC Cosmetics, Nike, Brandy Melville and more!

### STYLE

Need ideas for what to wear to your next date or that upcoming wedding? Do you hate getting dressed in the morning? Poshmark has millions of outfit ideas for you. Shop the personal style and closets of other Poshers today, or request to be styled by your favorite Stylist!

### SELL: CONSIGN YOUR CLOSET

Do you have items in your closet that didn't work out? Maybe a designer handbag, statement shoes, or even your old wedding dress you wore once and won't ever wear again? Poshmark is the #1 online consignment solution for a girl to sell her unwanted clothing in order to buy new things she's coveting, providing an alternative to Goodwill and other consignment stores like Buffalo Exchange. 1 in 50 American women have their closets open for sale on Poshmark—you could be next!

### PARTY

Party on the go, no matter where you're at! Join the thousands of women who come together three times a day, seven days a week to participate in our themed virtual Posh Parties. Create new listings, share and shop together! Have Tory Burch, Forever 21, or Zara on the brain? There's a party for you.

Shop online. @zafulshopping

# Associate features with User Reviews

---



- *word2vec* for producing word embedding
  - Train a neural network model
  - quantify and categorize semantic similarities between words



# Sentiment Analysis

---



- Train a sentiment classifier based on
  - Lexical evidence
  - Syntactic structure
  - Semantic dependency

# Evaluation Plan

---

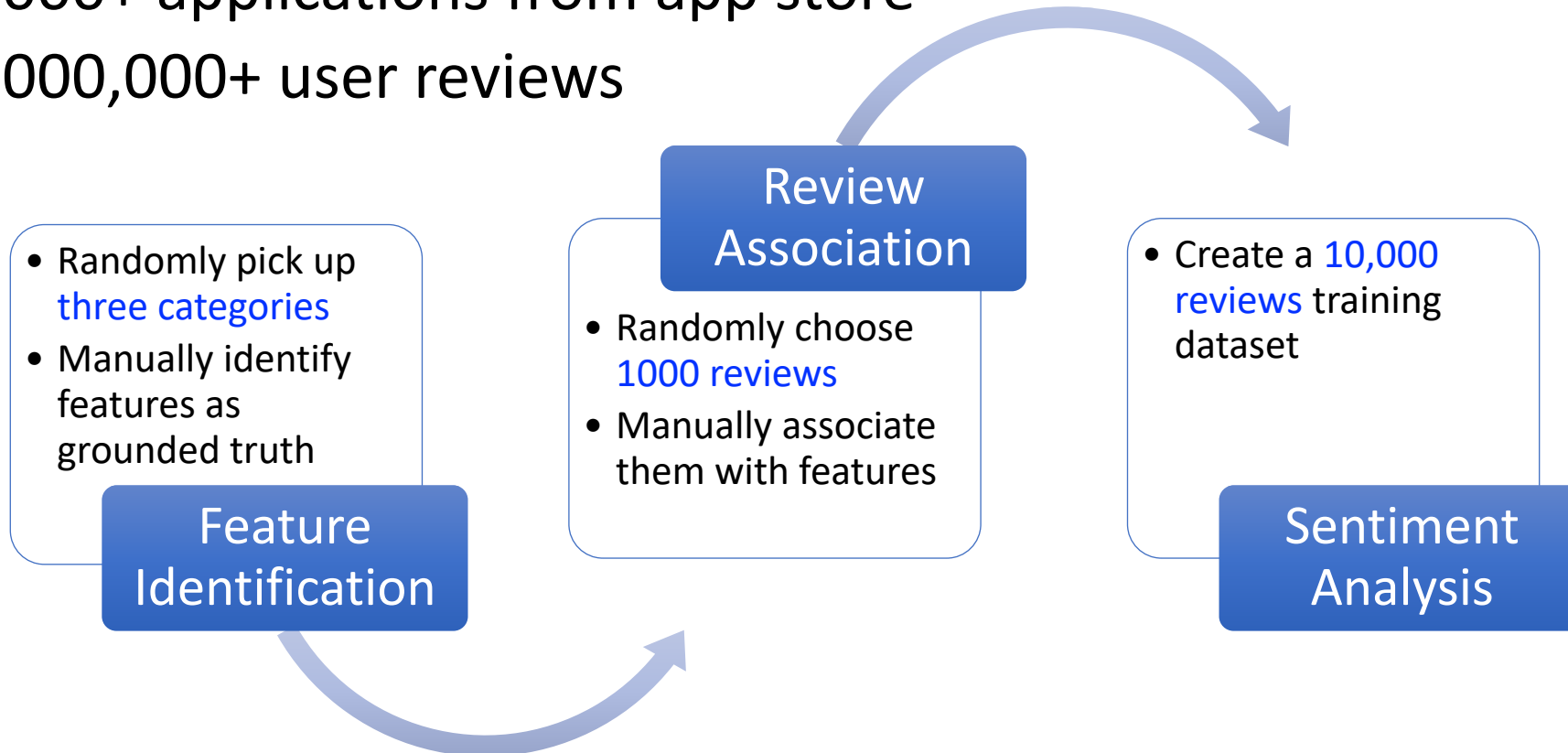


- **RQ1.** To what extent can the topic modelling-based method and the clustering-based method respectively extract features of a category of software applications from the unstructured descriptions?
- **RQ2.** To what extent can the word2vec method associate user reviews with previously identified features?
- **RQ3.** To what extent can our proposal accurately classify sentiments of user reviews?
- **RQ4.** Whether software companies can benefit from our approach and would like to adopt it?

# Evaluation Plan



- Data collection
  - 5,000+ applications from app store
  - 1,000,000+ user reviews



# Conclusions and Future work

---



- A research preview about a data driven user preference elicitation approach
- Methods for filtering useless information from application descriptions
- Syntactic templates for feature extraction
- Effective visualization algorithms



北京工業大學

BEIJING UNIVERSITY OF TECHNOLOGY

THANK YOU!

Contact: [litong@bjut.edu.cn](mailto:litong@bjut.edu.cn)

