

Abstract

Research on mining user reviews in mobile application stores has noticeably advanced in the past few years. The majority of the proposed techniques rely on classifying the textual description of user reviews into different categories of technically informative user requirements and uninformative feedback. Relying on the textual attributes of reviews, however, often produces high dimensional models. This increases the complexity of the classifier and can lead to overfitting problems. We propose a novel approach for application review classification. The proposed approach is based on the notion of semantic role labeling, or characterizing the lexical meaning of text in terms of semantic frames. Semantic frames help to generalize from text (individual words) to more abstract scenarios (contexts). This reduces the dimensionality of the data and enhances the predictive capabilities of the classifier. As a result, the proposed approach can be used to generate lower dimensional and more accurate models in comparison to text classification methods.

Motivation

App store reviews contains contain substantial amounts of up-to-date technical information that can be leveraged by software developers to create better software products.

(It wont load any of my notifications when I click on them.)



It wont load any of my notifications when I click on them





Browse Classify All and Export

MARC 1.0 - Mobile App Review Classifier Conclusions Import Reviews iOS App ID: 31063399 Number of Reviews: 100 - Import The Bag-of-Frames approach Imported Reviews 100 Reviews Imported I use this instead of iMessage generates less complex models. Love this app! Easy to use ! No problems so far Great way to stay in touch with friends overseas!! Can you stop your stupid update ?? Every hour we find a new update oooowww , what's your problem , it's s Great updates but please add a passcode for WhatsApp to improve privacy further 🗟 552 unique frames vs 1592 unique This application is my favorite l love this app. Must get....Please add to Apple Watch words. Classification Settings 60% reduction in space and time Default Training Bag of Frames Support V. Machine Custom Training requirements. 📃 Bag of Words Kernel: PolyKernet Remove Stopwords Naive Bayes Stemmina 4. Less chances of over-fitting. **Classify Single Review** Select a review from above list No need for pre-processing. Classification Type Text Ӿ Bug Report The app was working fine, updated to this version and now I can't send messages verified that my internet is working, I can send messages with different apps, but w. Classify Selected Review

https://github.com/seelprojects/MARC

——— College of ——— Engineering



Classify All Reviews and Export

Output Folder

Mining User Requirements from Application **Store Reviews Using Frame Semantics**

Nishant Jha and Anas Mahmoud

Division of Computer Science and Engineering, Louisiana State University

Traditional Approach and its Limitations

Text is represented as an unordered vector of words. Limitations of Bag-of-Words approach:

- 1. 60 70% accuracy.
- 2. Too many classification features (words).
- 3. Huge and sparse word x document matrix.
- 4. Over-fitting

"It keeps crashing every time I delete a comment and other comments disappear too"

BOW – Bag of Words

<keep, crash, time, delete, comment, disappear, too>

The FrameNet Project

The FrameNet is a lexical database of English. Housed and maintained by the International Computer Science Institute in Berkeley, California, the FrameNet project provides a massive machine readable database of manually annotated sentences based on the theory of Frame Semantics.

- 1. Semantic Frame (SF): A schematic representation of a situation (events, actions) involving various elements.
- 2. Frame element (FE): A participant entity or a semantic role in the action described by the frame.
- 3. Lexical units (LU): The words that evoke different frame elements.











Our Approach

- Uses the FrameNet Project.
- Reduces the number of classification features.
- Rely on words' meaning and context, known as Frames





Papers

MARC: A Mobile Application Review Classifier N. Jha and A. Mahmoud, Requirements Engineering: Foundation for Software Quality (REFSQ: Tools), accepted, 2017

Mining user requirements from application store reviews using frame semantics N. Jha and A. Mahmoud, Requirements Engineering: Foundation for Software Quality (REFSQ), accepted, 2017

Source Internal Data from Data from Total

Table 2. The performance of NB and SVM over the BOF and the BOW representations of the data in Table 1

Classifier BOF + NBBOF + SVNBOW + NHBOW + SV

Classifie BOF + NBOF + SBOW + BOW +



Dataset

We used two external datasets and an internal dataset obtained from different sources to reduce potential sampling bias.

					5	
	$\operatorname{sampled}$	discarded	Bugs	Req.	Others	Total
data	705	3	170	65	467	702
m[1]	725	8	318	199	200	717
m[2]	1500	2	854	537	107	1498
	2930	13	1342	801	774	2917

Table 1. The dataset used in our analysis





Classification Accuracy & Test Dataset

	Bug Reports			User Requirements			
	р	r	F_1	р	r	F_1	
	0.80	0.83	0.81	0.70	0.69	0.70	
Λ	0.84	0.88	0.86	0.73	0.75	0.74	
3	0.81	0.77	0.79	0.71	0.73	0.72	
Μ	0.78	0.93	0.85	0.83	0.69	0.75	

Table 3. A test set of app reviews sampled from three apps

Source	Bugs	Req.	Others	Total
Google Chrome	125	26	91	242
Facebook	56	7	32	95
Google Maps	108	17	50	175
Total	289	50	173	512

Results

Table 4. The performance of the different classifiers over the test set (Table 3)

	Bug Reports			User Requirements		
r	р	r	F_1	р	r	F_1
ΙB	0.85	0.92	0.88	0.41	0.73	0.53
VM	0.94	0.99	0.96	0.62	0.96	0.75
NB	0.84	0.71	0.77	0.28	0.62	0.39
SVM	0.78	0.97	0.86	0.45	0.68	0.54









