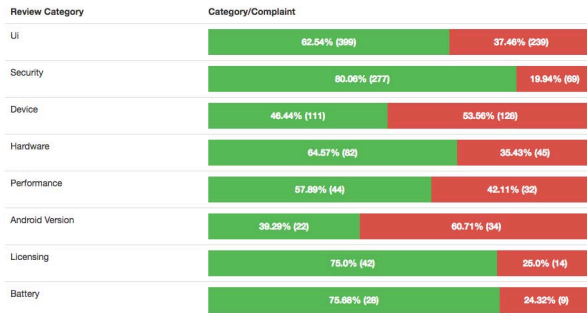


Table 1: Taxonomy of Review Categories

Review Category	Description
Device	mentions a specific mobile phone device (<i>i.e. Galaxy 6</i>).
Android Version	references the OS version. (<i>i.e. Marshmallow</i>).
Hardware	talks about a specific hardware component.
App Usability	talks about ease or difficulty in using a feature.
UI	mentions an UI element (<i>i.e. button, menu item</i>).
Performance	talks about the performance of the app (<i>slow, fast</i>).
Battery	references related to the battery (<i>i.e. drains battery</i>).
Memory	mentions issues related to the memory (<i>i.e. out of memory</i>).
Licensing	references the licensing model of the app (<i>i.e. free, pro version</i>).
Price	talks about money aspects (<i>i.e. donated 5\$</i>).
Security	talks about security/lack of it.
Privacy	issues related to permissions and user data.
Complaint	the users reports or complains about an issue with the app

Reviews Analysis for com.achep.acdisplay.csv**Figure 1: Analysis Results for the AcDisplay App.**

green) and the ones that are (in red) are shown, thus the developer can quickly notice which categories are problematic for the users, e.g. in Figure 1 it is easy to see that the most troublesome category is *Android version*. Subsequently, the developer can decide to only look at reviews belonging to this category and understand what problems the users are reporting. For example, the developer might filter the reviews further and look at only the ones that are classified as *Android version* and *Performance*. Another use case is for the developer to focus on each mobile specific topic and understand if the developers report problems (*Complaint*) or are in general happy with that particular aspect of the application.

Classified Reviews

The file com.achep.acdisplay.csv contains 2403 reviews according to the selected categories. Currently reviews 1-20 are shown.

Rating	REVIEW	Select Categories
4	Great app just missing some The reason I say missing is because I want a built in app password. Anyway this is an amazing application	
3	I like the app but it brings up the "use fingerprint because security policy" thing sometimes. If you can get it to turn the screen off by root or something (or however you can fix this) 5 stars	
1	Can not uninstall???? I used it for a while and it was very helpful but now it's integrated inside my onepus 2 and I wanted to uninstall it but I can't. Coz I can not remove it from device administration. Everytime I touch the tick, it takes me back to settings and I can do nothing about it. So this app is permanently stuck on my phone	
3	Good app BUT... Every time when I restart my phone I have to approve permissions to activate it.. Pls improve it...	

Figure 2: Classification Results for the AcDisplay App.**4 EVALUATION**

We evaluated our tool prototype from both a quantitative and preliminary qualitative perspective. We analysed the performance of the ML models using a 10-fold cross-validation approach and we obtained an average precision of 0.836, an average recall of 0.763

and an average F1 score of 0.793. Our tool shows good accuracy for automatically classifying reviews according to the fine grained taxonomy we defined. Finally we asked three external evaluators to analyse the reviews of 3 open source apps using our tool and Excel and report their experience. Overall the feedback from the preliminary user study was positive, the evaluators reported that they needed less time to analyse the reviews using our tool as opposed to Excel and that they found it *helpful*. The full details of the evaluation are available in the online repository of the tool.

5 CONCLUSIONS

We present a novel tool, AUREA, that is able to classify and filter mobile apps reviews based on fine grained mobile specific categories. Through our work we want to help developers analyse the reviews of their apps in less time, better comprehend what issues users are reporting and plan their change tasks accordingly. The evaluation showed that our tool obtains high precision and recall in classifying reviews and our study participants found it helpful for analyzing reviews. **We plan as future work to extend our dataset, experiment with different summarisation techniques and conduct a more comprehensive qualitative evaluation.**

ACKNOWLEDGMENTS

We acknowledge the Swiss National Science foundation's support for the project SURF-MobileAppsData (SNF Project No. 200021-166275) and the partial support of the CHOOSE organization to attend the conference.

REFERENCES

- [1] N. Chen, J. Lin, S. C. H. Hoi, X. Xiao, and B. Zhang. Ar-miner: Mining informative reviews for developers from mobile app marketplace. In *Proceedings of the 36th International Conference on Software Engineering, ICSE 2014*, pages 767–778, New York, NY, USA, 2014. ACM.
- [2] A. Ciurumelea. URR. <https://github.com/adelinac/urr/>. [Online; accessed 31-January-2018].
- [3] A. Ciurumelea, A. Schaufelbühl, S. Panichella, and H. C. Gall. Analyzing reviews and code of mobile apps for better release planning. In *2017 IEEE 24th International Conference on Software Analysis, Evolution and Reengineering (SANER)*, pages 91–102, Feb 2017.
- [4] A. Di Sorbo, S. Panichella, C. V. Alexandru, J. Shimagaki, C. A. Visaggio, G. Canfora, and H. C. Gall. What would users change in my app? summarizing app reviews for recommending software changes. In *Proceedings of the 2016 24th ACM SIGSOFT International Symposium on Foundations of Software Engineering, FSE 2016*, pages 499–510, New York, NY, USA, 2016. ACM.
- [5] E. Guzman, M. El-Haliby, and B. Bruegge. Ensemble methods for app review classification: An approach for software evolution (n). In *2015 30th IEEE/ACM International Conference on Automated Software Engineering (ASE)*, pages 771–776, Nov 2015.
- [6] E. Guzman and W. Maalej. How do users like this feature? a fine grained sentiment analysis of app reviews. In *2014 IEEE 22nd International Requirements Engineering Conference (RE)*, pages 153–162, Aug 2014.
- [7] C. Jacob and R. Harrison. Retrieving and analyzing mobile apps feature requests from online reviews. In *Proceedings of the 10th Working Conference on Mining Software Repositories, MSR '13*, pages 41–44, Piscataway, NJ, USA, 2013. IEEE Press.
- [8] W. Martin, F. Sarro, Y. Jia, Y. Zhang, and M. Harman. A survey of app store analysis for software engineering. *IEEE Transactions on Software Engineering*, PP(99):1–1, 2017.
- [9] D. Pagano and W. Maalej. User feedback in the appstore: An empirical study. In *2013 21st IEEE International Requirements Engineering Conference (RE)*, pages 125–134, July 2013.
- [10] scikit-learn developers. Gradient Boosting Classifier. <http://scikit-learn.org/stable/modules/generated/sklearn.ensemble.GradientBoostingClassifier.html>. [Online; accessed 14-June-2017].
- [11] L. Villarroel, G. Bavota, B. Russo, R. Oliveto, and M. Di Penta. Release planning of mobile apps based on user reviews. In *Proceedings of the 38th International Conference on Software Engineering, ICSE '16*, pages 14–24, New York, NY, USA, 2016. ACM.