

Automated Microloan Risk Assessment using Crowdsourcing and Social Networks

Tofiq Hasanov, Motoyuki OZEKI, Natsuki OKA

Kyoto Institute of Technology, Japan

Abstract

The task of automated risk assessment is gaining attention in the light of recent spike of micro-loan popularity. The industry requires a method for fast processing of large numbers of applicants for short term small loans in real time. Obviously manual checking is not a viable option here since number of applications is supposed to be very large. In cooperation with a micro-loan company in Azerbaijan, we have researched an option of automated risk assessment using crowdsourcing. The main idea behind this approach is the fact that a lot of information about a particular applicant can be retrieved from his social networks. Our suggested approach can be divided into three main parts: First, information about the applicant is collected on social networks, such as LinkedIn, Facebook and Twitter. Of course, this happens only if the applicant gives us a permission to do so. In the next step, this data is being processed by a program, which extracts relevant information segments from this data. Such segments, for example, could be applicant's professional connections, his job history, recommendations, number of followers on Twitter, interests on Facebook, etc. Finally, in the third step, these information segments are evaluated using crowdsourcing. We have tested several possible crowdsourcing methods. First, we have distributed the task of evaluation of information segments to a group of people, hired for this job, similar to Amazon Mechanical Turk. In the second approach, we have tried to evaluate information segments inside the social networks. To do that, we have been automatically posting statuses on social networks about some information segments and evaluated community response by counting likes and shares. For example, by posting status "Do you think that a person who has worked in ABC Company is more likely to return loan? Please like this post if you agree", or by posting voting posts of similar nature and then counting number of likes or votes and comparing to other posts, we were able to estimate public opinion. Both crowdsourcing approaches had been effectively solving the task, but each had its pros and cons. Once evaluated, each information segment was then given a weight factor, which was optimized using available loan return test data, provided to us by a company. We have then tested our system on the validation set of 400 applicants. We have been able to show that the resulting solution gives 92.5% correct assessment, with 6.6% false negative and 11.4% false positive.