

Data-Driven Secret Santa

Giorgi Kvernadze

UUCS-18-004

School of Computing
University of Utah
Salt Lake City, UT 84112 USA

May 10, 2018

Abstract

In 2016, the Bank of Georgia and the Georgian Post conducted a nation-wide secret santa for the residents of the country Georgia. For 2017, they wanted to repeat the campaign while using the collected data from the previous year to improve the experience of the participants as well as the overall quality of the campaign. There were three main objectives in the project: 1) Extract information from the user-provided text fields. This was done through a combination of ad hoc linguistic and probabilistic methods. 2) Building a recommendation system that recommends items from local businesses in order to boost local economy. This was accomplished through combining item-item collaborative filtering with the demographic information of the users. And 3) creating a matching algorithm that maximizes the number of participants who both sent and received a gift. We trained a hybrid model with a Random Forest and Logistic Regression. The results were: 1417 extracted keywords, over 450 different local businesses connected with the users and increasing the percentage of people who sent out as well as received a gift by 14% compared to the previous year.