

Goodreads Report 1

Matthew D. Branson, James R. Brown

Department of Computer Science

Missouri State University

Springfield, USA

branson773@live.missouristate.edu

Abstract

Index Terms

sentiment analysis, natural language processing, Goodreads, domain adaptation, classification, information retrieval

I. PROJECT DESCRIPTION

Using the UCSD Goodreads Book Graph datasets, we propose to examine book reviews and the users who wrote them. The dataset contains information about 2,360,655 books (1,521,962 works, 400,390 book series, 829,529 authors); 876,145 users; 228,648,342 user-book interactions in users' shelves (including 112,131,203 reads and 104,551,549 ratings). We are currently in an exploratory analysis phase, using a PyQt6 application for dataset downloading and normalization into a SQL database. Our preparation can be observed at <https://github.com/mdb42/goodreads>.

II. LITERATURE REVIEW

III. BACKGROUND

IV. METHODS

V. TASK ASSIGNMENTS