Proceedings of the 15th Conference on Natural Language Processing (KONVENS 2019)

Identification of Reading Absorption in User-Generated Book Reviews

Piroska Lendvai, Simone Rebora, Moniek Kuijpers
Digital Humanities Lab
University of Basel, Switzerland
piroska.r@gmail.com

Abstract

We introduce a language processing approach to detect reading absorption expressed in book review texts in English from an online social reading platform. Such texts are opinionated, subjective, variable length self-narratives. We describe our corpus annotated with absorption categories that were defined in empirical aesthetics, based on which we performed supervised, sentence level, binary classification of the presence or absence of reading absorption, using text-based, distributional features.

1 Introduction

Our study aims to contribute to affect recognition research by introducing the affective state of absorption during reading fiction. Our goal is to computationally process user-generated book reviews from an online social reading community, to identify passages that textually express reading absorption. The reviews belong to the genre of self-narratives that report about individual experiences in a non-elicited way, typically serving multiple user intents such as providing one or more of evaluation, recommendation, feedback, as well as influencing and socializing. The reviews often do not merely contain mentions of evaluative sentiment toward (components of) the book, but rather also express complementary aspects in terms of engagement of the reader, for example immersive experiences ("*I was glued to my kindle.*"; *"It stayed with me, even when I wasn't reading it"*), transportation to the fictional world (*"i felt like am living inside it"*), altered sense of time during reading (*"it is almost 800 pages that just fly by"*), emotional engagement (*"I cried reading the last 30 pages"*), and others.

Computational analysis of detecting reader absorption has so far been largely unaddressed, except for our initial text similarity approach to detect absorption in the story world (Rebora et al., 2018). In this study, we are interested in Natural Language Processing (NLP) based modeling of the specific affective state of reading absorption. Our corpus construction is currently ongoing: the experiments were based on 200 reviews, from which we generated the first processing resources, i.e., distributional language models and supervised classifiers, to benefit researchers in computational linguistics, literature and social sciences studies.

2 Corpus and Labeling

Our current corpus consists of 200 English review texts which we collected from a social reading platform. These reviews pertained to books from different literary genres (romance, fantasy, thriller) that we pre-selected based on high star-ratings on the platform and the presence of trigger words. The data were balanced for amount of review per book.

We trained five annotators for labeling absorption in terms of a taxonomy of roughly 40 fine-grained absorption labels, grouped under broad concepts such as Attention, Transportation, Emotional Engagement, Mental Imagery, Disconnection from reality, etc. taken from Kuijpers et al. (2014) and Bálint et al. (2016). The annotators could also mark up when users explicitly signaled the lack of absorption (e.g. *"I struggled to get through a lot of the pages*" or *"None of the characters really mattered to me"*), to make them distinct from expressions of the presence of absorption.

The annotators worked on the review level using Brat[1] and could assign labels to text segments of arbitrary length. For the current study, we aggregated all annotators' labels into a generic absorption category: the *Abs* label was assigned if at least one of five annotators judged some part of a sentence as explicitly expressing some type of absorption

[1] https://brat.nlplab.org/

Sent	**Text**	Fine-grained label (Support)	**Binary label**	Justification
1	FUNNY.	-	nonAbs	generic evaluation
2	Funny funny funny and sexy as hell.	-	nonAbs	generic evaluation
3	I don't only like the Heroine,	-	nonAbs	generic sentiment
4	I LOVE her.	-	nonAbs	generic sentiment
...	...	...	...	...
12	*I want to be Molly when I grow up.*	Wishful identification (4)	Abs	Wish of having the same characteristics as protagonist
13	I loved her backstory and why she is the way she is.	-	nonAbs	generic sentiment/evaluation
14	Her Career Secret frustrated me at times.	-	nonAbs	generic sentiment
15	*Most of the time, I was like, "JUST TELL HIM."*	Participatory response (1)	Abs	Intention to intervene, at times by addressing the characters directly
16	*But I got why she felt she couldn't.*	Emotional understanding (2)	Abs	Emotional or cognitive understanding of the character's feelings or perspective
17	This is a great book.	-	nonAbs	generic evaluation

Table 1: Corpus excerpt with fine-grained absorption annotations (column 2) and the binarized target labels to classify (column 3). *Abs*: reading absorption or its lack is expressed, *nonAbs*: no absorption or its lack is expressed.

or the lack of it. Sentence-level segmentation was obtained using the Spacy package[2] that worked best for our user generated text type. In the pilot annotation round, the average review length was 25 sentences (stdev $\pm$28), inter-annotator agreement on the sentence level was 0.59 (Fleiss' Kappa).

To illustrate our data and classification task, a review excerpt is presented in Table 1, in which e.g. sentence *"I want to be Molly when I grow up."* was judged as *Wishful identification* by four annotators.

3 Reading Absorption Identification

The current dataset is imbalanced, as only 13% of the instances have the target class *Abs* (660 vs 4,327 sentences).We used a random undersampling method[3] during training to account for it. Sentences were stripped of punctuation, tokens were lowercased and stemmed (mean normalized sentence length: 15$\pm$13 tokens), and represented in terms of a count vector (length: 6,064) as well as a sentence embedding vector (length: 100). We generated the sentence embedding representation using the *sent2vec* tool[4] that we retrained on 2.45 million unlabeled social reading narratives collected from the online platform.

Next, we performed classification experiments using two classical machine learners with no optimization: logistic regression (class_weight=balanced) and random forest, in 5-fold cross-validation. The feature sets representing the sentences were tested in isolation and in combination. The results are presented in Table 2 and show that good precision is difficult to achieve for the target class in the current setup: the best F-scores are .42 using the large bag of words count vector or using all features. We are currently growing the corpus and consolidating the still evolving labeling scheme, after which we will be able to test more advanced data representation and learning approaches, and evaluate classification on the fine-grained absorption labels.

LR	**Abs**	**nonAbs**	**RF**	**Abs**	**nonAbs**
cv	P: 0.30	0.94	cv	P: 0.30	0.93
	R: 0.70	0.75		R: 0.62	0.78
	F: 0.42	0.84		F: 0.40	0.85
s2v	P: 0.24	0.94	s2v	P: 0.26	0.96
	R: 0.73	0.65		R: 0.81	0.64
	F: 0.37	0.77		F: 0.39	0.77
all	P: 0.29	0.95	all	P: 0.25	0.95
	R: 0.76	0.72		R: 0.79	0.65
	F: 0.42	0.82		F: 0.38	0.78

Table 2: Classification results by Logistic Regression (LR) and Random Forest (RF) in terms of Precision, Recall, and F-score per class, averaged from 5-fold cross-validation. Features: CountVectorizer (cv) and sent2vec embeddings (s2v).

[2]https://spacy.io

[3]https://github.com/scikit-learn-contrib/imbalanced-learn

[4]https://github.com/epfml/sent2vec

References

Bálint, K., Hakemulder, F., Kuijpers, M., Doicaru, M., Tan, E. S. (2016). Reconceptualizing foregrounding. Scientific Study of Literature, 6(2), 176-207.

Kuijpers, M., Hakemulder, F., Tan, E.E. and Doicaru, M.M. (2014). Exploring absorbing reading experiences. Developing and validating a self-report scale to measure story world absorption. Scientific Study of Literature, 4(1): 89122.

Rebora, S., Lendvai, P. and Kuijpers M. (2018). Reader experience labeling automatized: Text similarity classification of user-generated book reviews. In: EADH 2018 Conference, Book of Abstracts.