Analysis Of Aggregated Emotions On Product Reviews

Nitin cs14Btech11023
Sujit cs14btech11034
Ibrahim cs14btech11041

About Paper

Title : Prediction of Helpful reviews using Emotion Extraction

By:

Lionel Martin and Pearl Pu **Human** Computer Interactions Group School of Computer and Communication Sciences Swiss Federal Institute of Technology (EPFL) CH–1015, Lausanne, Switzerland {lionel.martin, pearl.pu}@epfl.ch

- Contribution
  - Extracted the emotionality from the review text and applied supervised classification method to derive the emotion-based helpful review prediction.
- Why does it matter ?
  - Generally for newer products (or sometimes even older products) we don’t have helpfulness / unhelpfulness score for a review.
- Claim
  - This Evaluation Framework shows that emotion-based methods are outperforming the structure-based approach, by up to 9%.

About Dataset

- Amazon Product Review (http://jmcauley.ucsd.edu/data/amazon/)
- Dataset is 5-core i.e. each product and user has at least 5 reviews
- Product Categories
  - Electronics
  - Health Personal Care
  - Clothing Shoes Jewelry
  - Cell Phones Accessories

About Dataset (Format and Assumptions)

- Format of the Data (JSON)
  - asin - ID of the product, e.g. 0000013714
  - Helpful - helpfulness rating of the review, e.g. $2/3$
  - Review Text - text of the review
  - Overall - rating of the product
  - Summary - summary of the review
  - Review Time - time of the review
- Assumption :
  - The time when review was written is within a month's neighbourhood of purchase of the product.

Steps Taken To Analyse Each Product Category

- Find if there exists sentiments in the category
- If sentiments found; correlate those sentiments with time of purchase
- Predict the best month to purchase the product
    - Overall Rating
    - Sentiment Analysis (opinion Mining)
    - Helpfulness of the Review
- Analyse the effect of emotions aggregated over all the months
    - Using GALC (Geneva Affect label Encoder)
        - It contains 39 categories of strong emotional words

Steps Taken To Analyse Each Product Category

- Study the relation b/w emotions and ranks over all the months
    - Extract Emotions and check which category of emotions (emotional words) affected the rank of that products for a particular month.
    - Answer the question: Does combining consecutive months give better result?
        - If Yes, then how many months should be combined.

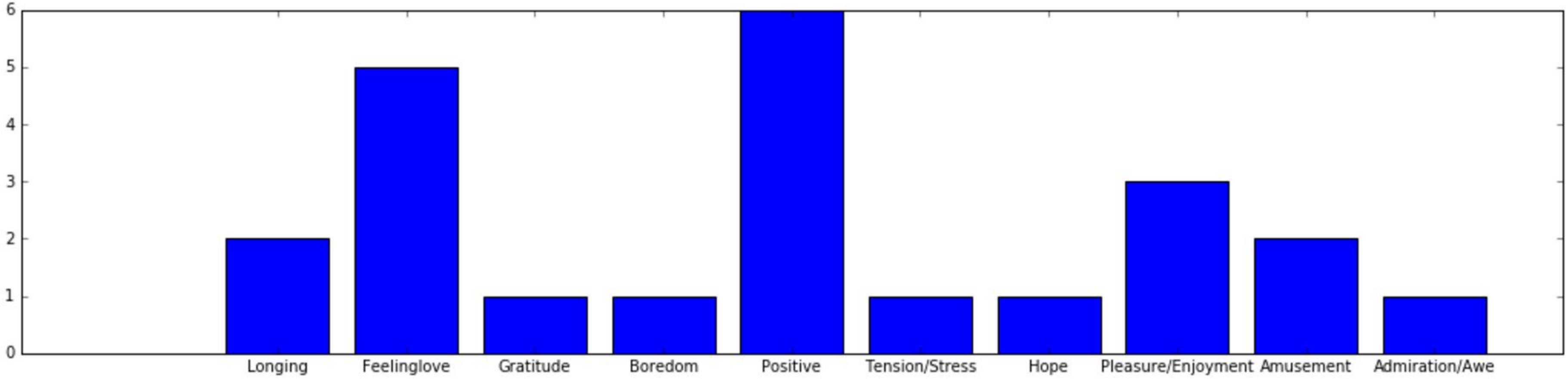
Model for Rank Prediction

- Model
  - Map (To find emotions which influence) score of each product’s review to it's time of purchase (month)
    - Give score for each month based on following.
    - Determine helpfulness of each review (w = fraction positive i.e (r+ve)/(r-ve + r+ve)) and use it to give weights to sentiment, rating, emotional scores.
    - Separate hyper-parameters for sentiment analysis and rating scores.
    - Aggregate over all the reviews.
    - Normalise and predict Rank according to scores.

Analyse the Aggregated Emotions w.r.t Best Buy months predicted

- For multiple set of products use best months and worst months to find pattern in emotions and find if emotions are a good indicator if people will buy a product or not.
  - Using emotions from different products best and worst months.
  - Plot Histograms for best and worst buy months of different products.

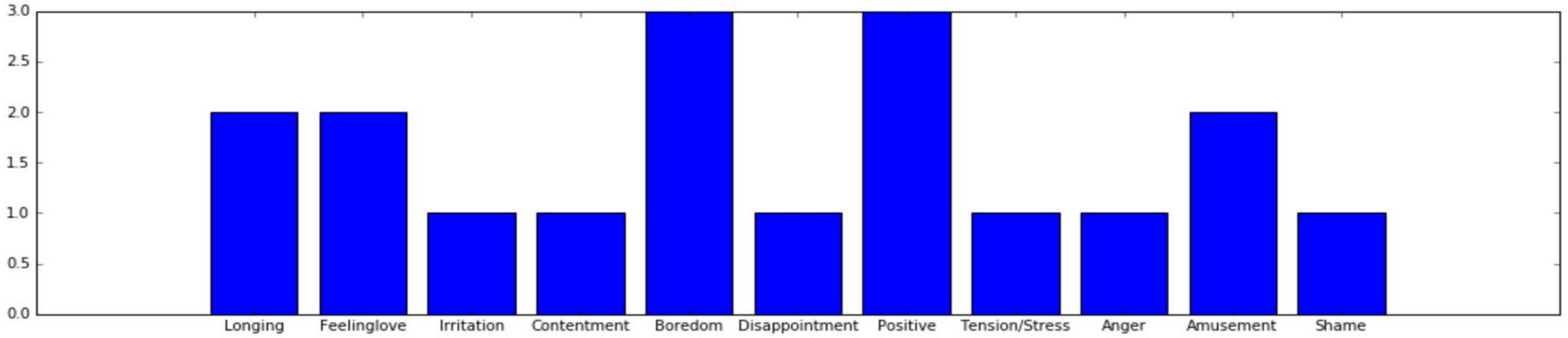
Emotions in the predicted best month (i.e with highest score) to sell/buy a product.



Y axis - Cumulative Score for emotions. (for clothes and wearables DataSet)

X - axis - Prominent emotions in the given months.

Emotions in the predicted month with least score to sell/buy a product.



- Common emotions can be ignored e.g “**longing**”, “**boredom**”, to get better **Feature set** to train a **Machine learning Model** for better prediction.

Best Predicted using structured Approach (Not Machine Learning)

Month's in which People like buying clothes and wearables:

1. Feb
2. Dec
3. Jan
4. Aug
5. Apr

References And Tools

- Prediction of Helpful Reviews using Emotions Extraction
- Manning, Christopher D., Mihai Surdeanu, John Bauer, Jenny Finkel, Steven J. Bethard, and David McClosky. 2014. The Stanford CoreNLP Natural Language Processing Toolkit In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pp. 55-60
- http://sentiment.vivekn.com/docs/api/
- http://text-processing.com/docs/sentiment.html
- (GALC) Geneva Affect Label Coder. (specifies 39 emotion based classes).