Selected Papers
Sentiment, Emotions, Others

Contents

1 Identifying Expressions of Emotion in Text

[Aman and Szpakowicz(2007)]
`https://ccc.inaoep.mx/~villasen/bib/Identifying%20Expression%20of%20Emotion%20in%20Text.pdf`

- Describes annotation task of identifying emotion category, emotion intensity and the words/phrases that indicate emotion in text.

- Data was collected from blogs. Emotional data are expected to be personal texts, such as diaries, email etc.

- Related works

  - MPQA corpus contains 10000 sentences collected from news. They were annotated by emotions, opinions and sentiments.
  - Appraisal Framework is a functional theory of language used for conveying attitudes, judgements and emotions.
    - These two don't deal with emotion exclusively which is the focus of this paper.
  - Automatic classification of basic emotions (by Ekman) ANGRY, DISGUSTED, FEARFUL, HAPPY, SAD, POSITIVELY SURPRISED, NEGATIVELY SURPRISED by Alm et al. Manually annotated data. Targeted for text to speech system for expressive rendering of stories.

- Annotation :

  - Picked 6 basic emotions. *Happiness, sadness, anger, disgust, surprise and fear*
  - Choose seed words for these emotions and categorized blog posts. *Not sure what they did if a document contains seed words for multiple emotions.*
  - 4 annotators
  - Added 2 extra emotion categories. *Mixed emotion and no emotion*
  - Annotated category and intensity (high, medium, low)
  - Identify spans of text that contain content of emotions.
  - Emotion is often conveyed by phrases.
  - IO encoding for In or Outside emotion indicator in word level. (Paper : Predicting Movie Sales from Blogger Sentiment)

- Classification :

  - **General Inquirer and Wordnet-Affect** for emotional lexicons. (EMOT in GI)
  - Each sentence was represented by 14 features. Each feature was to indicate the number of words from a certain category.
  - 6 Classes.

per se : By itself, intrinsically

2 Are Influential Writers More Objective? An Analysis of Emotionality in Review Comments

[Martin et al.(2014)Martin, Sintsova, and Pu]
`http://wwwconference.org/proceedings/www2014/companion/p799.pdf`

- Insight about influence can be gained from analyzing the affective content of the text as well as affect intensity.
- Extract *emotionality* of 68k hotel reviews to investigate how the influencers behave, especially their choice of words.
- Helpfulness prediction using emotion lexicons.
- Models were built around 20 emotional categories.
- 3 questions to answer.
  - How influential users behave in social media? What kinds of comments they are likely to construct? What is the emotionality of these comments?
  - Beyond positive and negative emotionality, is there a more elaborate emotion model to further identify emotionality? What is the method to detect these emotions?
  - What will be the major differences of emotionality between influential and non-influential users? Is our finding restricted to one type of dataset or can it be generalized to several others?

  Answers will help to build a model to predict helpfulness.
- Ekman defined 6 and then 7 emotions (Happiness, Sadness, Surprise, Fear, Disgust, Anger. Added **Contempt** later
- Plutchik associated emotions by **pair**. *Joy and sadness, Trust and Disgust, Fear and Anger, Surprise and Anticipation.* These families could be expressed at different intensity defining three layers.
-

3 From Once Upon a Time to Happily Ever After: Tracking Emotions in Novels and Fairy Tales

[Mohammad(2011)]
`http://hnk.ffzg.hr/bibl/acl2011/LaTeCH/pdf/LaTeCH-201114.pdf`

- How sentiment analysis can be used in tendem with effective visualizations to quantify and track emotions in both individual books and across very large collections.
- Introduced the **concept of emotion word density** and using Brothers Grimm fairy tales as example, we show how **collections of text can be organized for better search**.
- Using the Google Books Corpus we show how to determine an entity's emotion associations from co-occurring words.
- Compared emotion words in fairy tales and novels and showed that **fairy tales have a much wider range of emotion word densities than novels**.
- **Applications :** Search (darkest books, text snippets with suspense), Social Analysis (How books have portrayed different people over time), Comparative analysis of literary works, genres and styles (Do women authors use a different distribution of emotion words than their male counterparts), Summarization (Automatically generate summaries that capture the different emotional states of the characters in a novel), Analyzing Persuasion Tactics (Emotional words and their role in Persuasion)
- **Discussed :** How we use a large word emotion association lexicon to create a simple emotion analyzer. Visualizations that help track and analyze the use of emotion words in individual texts and across very large collections.

4 Affective Content Analysis in Comedy and Horror Videos by Audio Emotional Event Detection

[Xu et al.(2005)Xu, Chia, and Jin]
`http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1521500`

- Affective contents contain emotional factors that are related to the users' attention, evaluation and memories of the content. Modeling the affective contents depends on genres.
- Worked on comedy and horror films to extract the affective content by detecting a set of so-called **Audio Emotional Event (AEE)** such as laughing, horror sounds, etc.
- Used AEE as a clue to locate corresponding video segments.
- Dataset contains 40-minutes of comedy video and 40-minutes horror film.
- Affective content analysis puts much more emphasis on the audience's reactions and emotions. Like semantic analysis, the affective content analysis is also challenging due to the gap between low-level perceptual features and high-level understanding.
- Affective content doesn't require deep understanding of the contents. Rather it emphasizes on the factors that affect viewers attention, valuation and memory.

5 Predicting Emotional Word Ratings using Distributional Representations and Signed Clustering

`http://wwbp.org/papers/affnorms17eacl.pdf`

- Model can determine word-level ratings to unrated words using signed clustering of vector space word representations. It can determine words valence and arousal. Valence and Arousal can determine a words position on the circumplex model of affect. Worked with 3 languages.

  Related Works :

  - All affective states are represented as a linear combination of valence and arousal (Posner et all).
  - Creating rating database is time consuming and costly. Have chances of biases or halo effects.
  - Automatic expansion experiments started with the intuition that, words similar in reduced semantic space will have similar ratings. Bestgen and Vince computed the rating by averaging k-nearest neighbours from low dimentional semantic space. **Downside: antonyms are also semantically similar.**
  - Orthographic similarity has shown to slightly improve results.
  - *Graph methods inspired by label propagation*, *Adjective intensity prediction* used distributional methods. But these works are restricted to discrete categories and relative ranking within each semantic property.
  - Polarity is assigned to words in order to determine if a text is subjective and its sentiment, which is slightly different to word-level affective norms e.g., 'sunshine' is an objective word (neural polarity), but has a positive affective rating.

  Data :

  - Annotated in lab. Scale is 1 to 9 for both arousal(1:calm, 9:excited) and valence(1:negative, 9:positive)
  - Worked on ANEW corpus (affective norms of words). English, Dutch, Spanish.

  Method :

  - Two steps.
  - First step represents the words in a semantic space with reduced dimensionality.
  - Antonyms stay closer. So signed spectral clustering distorts the word representations.
  - English and Spanish vectors were made from Gigaword corpora. For dutch, Tulkens et. all (320 dimension). No stemming or lowercasing. Just tokenization.
  - Used signed spectral clustering (Sedoc 2016)

6 Readers vs. Writers vs. Texts: Coping with Different Perspectives of Text Understanding in Emotion Annotation

https://sigann.github.io/LAW-XI-2017/papers/LAW01.pdf

- ..

7 To Sing Like a Mockingbird

[Gatti et al.(2017)Gatti, Özbal, Stock, and Strapparava]
`http://www.aclweb.org/anthology/E/E17/E17-2048.pdf`

- Generates parody of a song by. It can identify the characterizing words and concepts related to a novel text, which are taken from the daily news. These concepts are then used as seeds to appropriately replace part of the original lyrics of a song using metrical, rhyming and lexical constraints.
- Related Works :
  - Poetry generation faces the same challenge in the computational creativity problem that they have to combine lexical, semantic and phonetic features in an unified framework.
  - Greene et. al's poetry generation model allowed users to control the meter and rhyme scheme. It later used in translation.
  - Toivanen proposed to generate poems by replacing words in an existing poem with morphologically compatible words that are semantically related to a target domain. No works with content control, inclusion of phonetic features and syntactic information.
  - Ozbal et. al proposed BRAINSUP that generates creative sentences by making heavy use of syntactic information to enforce well-formed sentences and to constraint the search for a solution. It provides an extensible framework in which various forms of linguistic creativity can easily incorporated.
  - *Valentino* (Guerini) slants existing textual expressions to obtain more positively or negatively valenced versions by using WordNet semantic relations and SentiWordNet. The slanting is done my modifying, adding or removing single words from existing sentences. Insertion and deletion of words is performed by utilizing Google Web IT 5-grams corpus to extract information bout the modifiers of terms vased on their parts of speech.
  - Studies focused on humor generation studies lexical substitution.
  - Stock and Strapparava generate acronyms based on lexical substitution via semantic field opposition, rhyme, rhythm and semantic relations provided by WordNet. It was limited to the generation of noun phrases.
  - Pun generator by Velitutti(2009) that modifies familiar expressions with word substitution. Velitutti(2013) propose an approach based on lexical substitution to introduce adult humor in SMS texts. A 'taboo' word is injected in an existing sentence to make it humorous.
- Used the 100 song corpus developed by Strapparava and Mihalcea(2012) and enriched the annotation by adding new tags like ¡bridge¿, ¡chorus¿ etc.
- The parody generation process is divided into four basic steps.
  - Retrieving the daily news and identifying the most characterizing words of each news piece. (Nouns, verbs, Named entities).
  - Finding new concepts and words evoking the initial text. (Synonyms from Thesaurus, WordNet and WikiData)
  - Generating parodies by replacing words inside the chorus of a song with these concepts, according to musical and linguistic constraints. (Replace if (i) two words are rhythmically similar, (ii) both have same number of syllables (iii) both have the same parts of speech) Used rhythm information from CMU punctuation dictionary.)
  - Producing a final output file MIDI files using Vocaloid.
- Evaluated by 3 crowdflower annotators. Checked on 3 criteria. (i) Grammatically correct? (ii) More related to the headline? (iii)Funny?

8 BRAINSUP: Brainstorming Support for Creative Sentence Generation

[Özbal et al.(2013)Özbal, Pighin, and Strapparava]
`http://www.aclweb.org/anthology/P13-1142`

- Framework that supports creative sentence generation by enabling users to force several words to appear in the sentences and to control the generation process across several semantic dimensions.
- Semantic dimensions are *emotions, domain-relatedness, color, phonetic properties.*
-

9 EMOBANK: Studying the Impact of Annotation Perspective and Representation Format on Dimensional Emotion Analysis

[Buechel and Hahn(2017)]
`http://aclweb.org/anthology/E17-2092`

- EMOBANK is a corpus of 10k English sentences balancing multiple genres, which we annotated with dimensional emotion metadata in the Valence-Arousal- Dominance (VAD) representation format.
- Related Works
  - Models of emotion are divided into two categories. Categorical (Valence, Arousal and Dominance -VAD) and dimensional (Wheel of Emotion by Plutchik, Siz basic emotions by Ekman).
  - VAD annotated corpora are surprisingly rare in number and small in size, and also tend to be restricted in reliability. ANET, for instance, comprises only 120 sentences designed for psychological research (Bradley and Lang, 2007), while Preotiuc-Pietro et al. (2016) created a corpus of 2,895 English Face- book posts relying on only two annotators. Yu et al. (2016) recently presented a corpus of 2,009 Chinese sentences from various online texts.
  - For categorical emotions there are no common practices. In psychology, there is still no consensus on a set of fundamental emotions. So comparison between two sets is difficult. Using VAD model provides that opportunity.
  - Earlier works followed two schemes (VAD and emotion categories). There is no accurate mapping between two. This corpus is annotated bi-representationally. So it contains both schemes.
  - While it is common for more basic sentiment analysis systems in NLP to map the many differ- ent possible interpretations of a sentence's affec- tive meaning into a single assessment ("its senti- ment"), there is an increasing interest in a more fine-grained approach where emotion expressed by writers is modeled separately from emotion evoked in readers. An utterance like "Italy de- feats France in the World Cup Final" may be com- pletely neutral from the writer's viewpoint (presumably a professional journalist), but is likely to evoke rather adverse emotions in Italian and French readers (Katz et al., 2007).
  - IIA differs from the perspective of writer and readers.
- Corpus Design
  - Addressed several genres and domains
  - Each sentence were rated based on reader and writers perspective (bi-perspective). Because, studies show that IAA differs between two perspectives.
  - Used Crowdflower to get annotations. JEMAS (a lexicon based tool for VAD prediction) was used to generate gold scores by 3 student annotators expert in linguistic.
  - For each sentence, five annotators generated VAD ratings. Thus, a total of 30 ratings were gathered per sentence (five ratings for each of the three VAD dimensions and two annotation perspectives, WRITER and READER).
- Analysis
  - Used Pearson correlation and MAE for evaluating inter annotator agreement.
  - The READER results in significantly higher correlation, but also higher error than WRITER (p ¡ .05 for Valence in r and for all dimensions in MAE using a two-tailed t-test).

10 A quantitative analysis of gender differences in movies using psycholinguistic normatives

[Ramakrishna et al.(2015)Ramakrishna, Malandrakis, Staruk, and Narayanan]
`http://aclweb.org/anthology/D15-1234`

- Investigates the differences between male and female character depictions in movies, based on patterns of language used. Specifically they use an automatically generated lexicon of linguistic norms characterizing gender ladenness.
- Investigated the depictions of the genders in feature films, through the analysis of their respective dialogues.
- For analyzing the dialogues, they proposed a metric of **language gender ladenness**, a number representing a normative rating of the perceived feminine or masculine association of language (Palvio et al., 1968). This metric originally meant to provide an indication of gender-specificity of individual words, with extreme values assigned to highly stereotypical concepts. Generating this rating for male and female character dialogues and comparing the character gender with this rating of ”language gender” should allow us to observe stereotypical behavior.
- Expanded the gender ladenness lexicons by Clark and Paivio, 2004 using work of (Malandrakis et al., 2013). Its basically a system that uses semantic similarity to assign scores to new lexicons.
- Used Movie DiC dataset that contains 619 movie scripts parsed from IMSDb. Had to clean the data to map character names found in IMDb and DiC scripts. Character names were mapped using Jaro-Winkler distance (variation of traditional edit distance). Assigned gender labels to each character using NamSor Applied Onomastics, 2015. Had to tag manually for 75 movies.

11 Supervised Learning of Universal Sentence Representations from Natural Language Inference Data

[Conneau et al.(2017)Conneau, Kiela, Schwenk, Barrault, and Bordes]
`https://arxiv.org/pdf/1705.02364.pdf`

- In this paper, we show how universal sentence representations trained using the supervised data of the Stanford Natural Language Inference dataset can consistently outperform unsupervised methods like SkipThought vectors (Kiros et al., 2015) on a wide range of transfer tasks. Much like how computer vision uses ImageNet to obtain features, which can then be transferred to other tasks, our work tends to indicate the suitability of natural language inference for transfer learning to other NLP tasks.
- Studied a sentence encoder model trained on large corpus and subsequently transfered to other tasks. **But**, two questions need to be solved.
  - What is the preferable neural network architecture?
  - How and what task should such a network be trained?
- Experiments
  - LSTMs and GRUs with mean and max pooling
  - A self attentive network that incorporates different views of the sentence
  - A hierarchical convolutional network that can be seen as a tree-based method that blends different levels of abstraction.
- Bi-LSTM worked better than the others.

12 The State of the Art in Semantic Representation

[Abend and Rappoport(2017)]
http://www.cs.huji.ac.il/~arir/semantic-reps.pdf

- **Semantic Representations approaches** : 1) Argument structure 2) External (extra-textual criteria) 3) Vector Space Models
- **Semantic Contents**: Events, Predicates and Arguments, Core and Non-core Arguments, Semantic Roles, Co-reference and Anaphora, Temporal Relations, Spatial Relations, Discourse Relations, Logical structure, Inference and Entailment.
- **Semantic Schemes and Resources**: Semantic Role Labeling, AMR (predicate argument relations) , UCCA, UDS, The Prague Dependency Treebank (PDT) Tectogrammatical Layer (PDT-TL), CCG-based Schemes, HPSG-based Schemes, OntoNotes

  [Very descriptive paper]

References

[Abend and Rappoport(2017)] Omri Abend and Ari Rappoport. 2017. The state of the art in semantic representation. In *Proc. of ACL*.

[Aman and Szpakowicz(2007)] Saima Aman and Stan Szpakowicz. 2007. Identifying expressions of emotion in text. In *Proceedings of the 10th International Conference on Text, Speech and Dialogue*. Springer-Verlag, Berlin, Heidelberg, TSD'07, pages 196–205. http://dl.acm.org/citation.cfm?id=1776334.1776363.

[Buechel and Hahn(2017)] Sven Buechel and Udo Hahn. 2017. Emobank: Studying the impact of annotation perspective and representation format on dimensional emotion analysis. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*. Association for Computational Linguistics, Valencia, Spain, pages 578–585. http://www.aclweb.org/anthology/E17-2092.

[Conneau et al.(2017)Conneau, Kiela, Schwenk, Barrault, and Bordes] Alexis Conneau, Douwe Kiela, Holger Schwenk, Loïc Barrault, and Antoine Bordes. 2017. Supervised learning of universal sentence representations from natural language inference data. *CoRR* abs/1705.02364. http://arxiv.org/abs/1705.02364.

[Gatti et al.(2017)Gatti, Özbal, Stock, and Strapparava] Lorenzo Gatti, Gözde Özbal, Oliviero Stock, and Carlo Strapparava. 2017. To sing like a mockingbird. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*. Association for Computational Linguistics, Valencia, Spain, pages 298–304. http://www.aclweb.org/anthology/E17-2048.

[Martin et al.(2014)Martin, Sintsova, and Pu] Lionel Martin, Valentina Sintsova, and Pearl Pu. 2014. Are influential writers more objective?: An analysis of emotionality in review comments. In *Proceedings of the 23rd International Conference on World Wide Web*. ACM, New York, NY, USA, WWW '14 Companion, pages 799–804. https://doi.org/10.1145/2567948.2579242.

[Mohammad(2011)] Saif Mohammad. 2011. From once upon a time to happily ever after: Tracking emotions in novels and fairy tales. In *Proceedings of the 5th ACL-HLT Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities*. Association for Computational Linguistics, Stroudsburg, PA, USA, LaTeCH '11, pages 105–114. http://dl.acm.org/citation.cfm?id=2107636.2107650.

[Özbal et al.(2013)Özbal, Pighin, and Strapparava] Gözde Özbal, Daniele Pighin, and Carlo Strapparava. 2013. Brainsup: Brainstorming support for creative sentence generation. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, Sofia, Bulgaria, pages 1446–1455. http://www.aclweb.org/anthology/P13-1142.

[Ramakrishna et al.(2015)Ramakrishna, Malandrakis, Staruk, and Narayanan] Anil Ramakrishna, Nikolaos Malandrakis, Elizabeth Staruk, and Shrikanth Narayanan. 2015. A quantitative analysis of gender differences in movies using psycholinguistic normatives. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Lisbon, Portugal, pages 1996–2001. http://aclweb.org/anthology/D15-1234.

[Xu et al.(2005)Xu, Chia, and Jin] M. Xu, L. T. Chia, and J. Jin. 2005. Affective content analysis in comedy and horror videos by audio emotional event detection. In *2005 IEEE International Conference on Multimedia and Expo*. pages 4 pp.–. https://doi.org/10.1109/ICME.2005.1521500.