Visual Analysis of Meetup Topics

Alejandro Ardila, Jashua Thejas Arul Dhas, Joshua Demusz, Justin Lewis, Khanh Tran, Tri Nguyen

Introduction

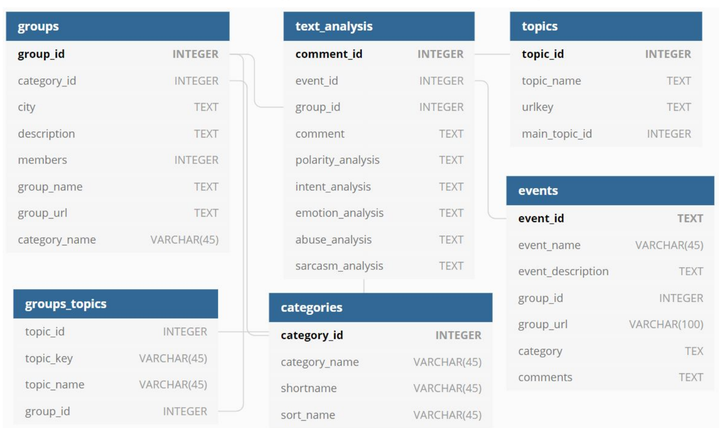
Event marketing is an attractive research opportunity due to emerging technologies that are narrowing the gap between offline social interactions and online event hosting. According to the latest Event Marketing Benchmarks and Trends report, over-performing companies and organizations invest between **20% to 50%** of the marketing budget in hosting live events. Thus, hosting an event is one of the most effective channels for accomplishing business goals. Thanks to data available from event-based social networks like Meetup, **event organizers** can gain valuable insights about their audiences in order to grow and improve their events.

Approach

In order to assist Meetup event planners in improving and growing their events, we have delivered a web application that provides analytical insight into Meetup topics that would be difficult to discover manually. To support this tool we developed a model to cluster similar Meetup event themes, providing an innovative way to categorize events. Utilizing this model as well as the existing Meetup organizational structure, users are able to visualize analytical insights into event comments and topic clusters of their selection.

Dataset

Our data extraction engine uses the Meetup API to collect all forms of data:



- Categories of topic: 2,833
- Events description: 18,821
- Public meetup groups : 12,530
- Group/topics relationship: 19,362
- Topics: 4,123

Fig. DB schema of dataset

Architecture

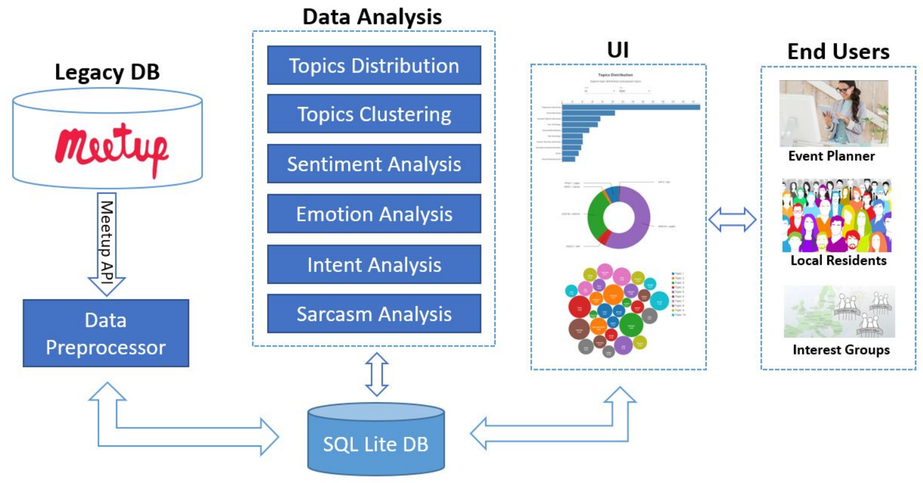


Fig. The design of our application

Topic Clustering

Since Meetup categories are high-level and Meetup groups & events are specific, an intermediary organizational level could provide new analytical insight.

Meetup event planners could find value in this extra tier of granularity, as they may be interested in analyzing topics rather than broad categories or specific events.

Fig. The figure to the left depicts what kinds of topics could be extracted from the data for the 'Tech' category.

The topic extraction model was created using **LDA** (latent Dirichlet allocation), which is a generative statistical model that clusters similar topics. Mathematically, the total probability of the model presented:

$$P(W, Z, \theta, \varphi, \alpha, \beta) = \prod_{i=1}^{K} P(\varphi_i; \beta) \prod_{j=1}^{M} P(\theta_j; \alpha) \prod_{t=1}^{N} P(Z_{j,t}|\theta_j) P(W_{j,t}|\varphi Z_{j,t})$$

This model takes Meetup event names and descriptions as input, clusters frequently occurring words, and outputs weighted topics.

Topic Distribution Analysis

Leveraging the results of the topic extraction model, we have provided an interactive visualization that assists the user in analyzing the weighted distribution of clustered topics underneath a selected category.

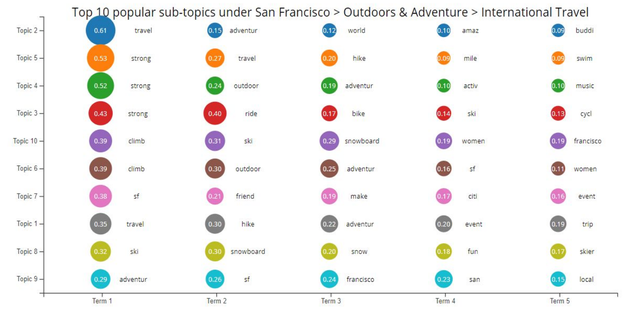


Fig. A bubble matrix is used to visualize the topics generated by the topic extraction model. The size of the bubble corresponds to the weight of the topic.

User Interface: Meetup event planners are able to drill down further into topics that are of interest to them, and grasp metrics such as how many members are encompassed in a particular topic. The user could use this information to determine which topics are popular and unpopular for a given location.

Comment Analysis

We use text analysis to extract and aggregate valuable information from opinion-based comments. Applying this technique to Meetup's comments data helps event organizers understand how their events impact attendees, as well as how attendees feel about an event in a statistical way.

Internally, our text analysis model uses **LSTM** (Long Short Term Memory) algorithms, an artificial recurrent neural network architecture, to classify a text blob's sentiments into various emotional states and feelings:

- Sentiment
- Intent
- Abuse
- Emotion
- Sarcasm

User Interface: For a particular location, the user can select a category, cluster, group, or event in which they want to analyze. The user can select to view metrics for following qualities based on this analysis:

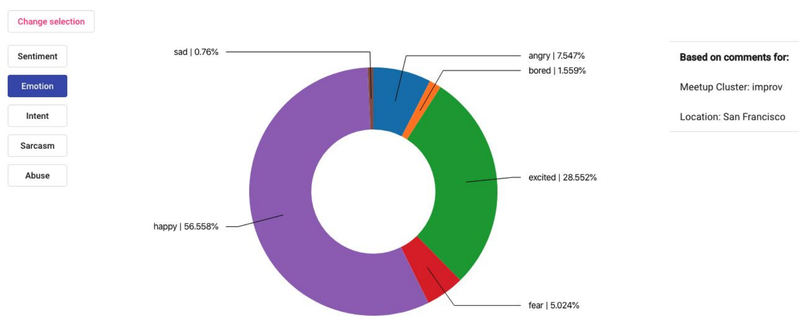
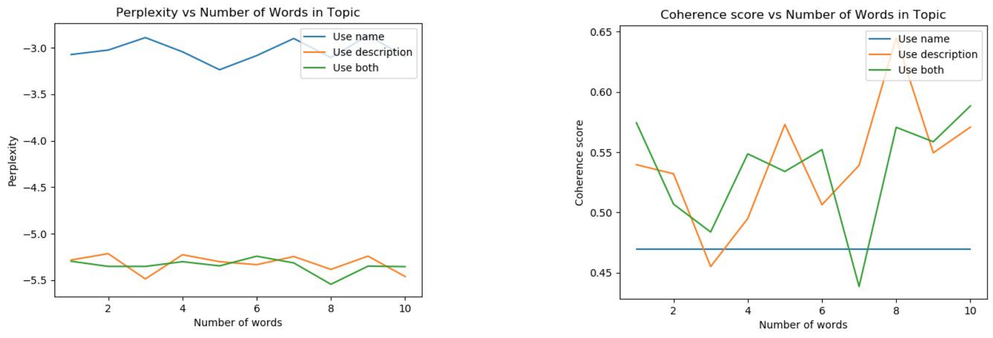


Fig. Breakdown of **emotion** analysis for improv cluster

Event planners can leverage this UI in order to quantify attendee opinion over four different levels of granularity. This ability is not offered by Meetup.com or similar tools.

Topic Extraction Model Parameter Experiment

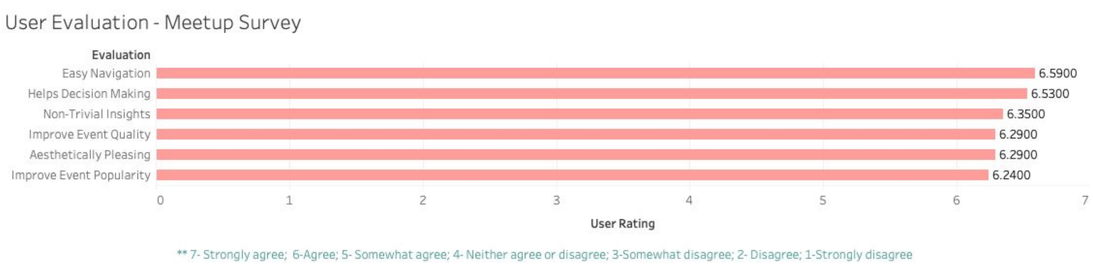
To tune the topic extraction model, various parameter combinations were tested to determine the configuration that optimized **perplexity** and **coherence score**.



A **lower** perplexity and **higher** coherence score is favorable for an LDA model. Thus, the model performs better when it uses both event name and event description as the input corpus.

The number of words in a topic did not correlate with the models' performance.

User Evaluation



Users have praised the tool's ease to navigate as well as it's ability provide non-trivial insight to guide an event planner's decision making. Specifically, this tool expands upon current methods since it provides novel data-driven insights such as new organizational topics/clusters as well as comment analysis metrics.

On the contrary, users less frequently thought that the tool would directly increase the popularity of a Meetup event.