Evaluation of Conflicts among Reddit posts using Clustering Techniques

Srinidhi Gopalakrishnan, Nanditha Sundararajan

gopalakrishnan.s@husky.neu.edu, sundararajan.n@husky.neu.edu

Problem Definition

- Reddit is a large online discussion platform where posts related to a variety of subjects are created and shared among users
- Some Subreddits generate conflicts among communities due to their negative sentiment
- Given features of posts created in Subreddits, we aim to:
  - Group them into clusters
  - Detect magnitude of conflicts directed towards Subreddits called 'Conflict Score'
- Assist Reddit Post-Control Community in controlling network traffic and in preventing conflicts

Existing Methods and Challenges

- Community interaction and Conflicts on the web**
  - Manual Tagging of Subreddits and analysis of interaction between clusters
  - Early Detection of conflicts between communities using LSTMs
- Challenges:
  - Manual Tagging – time consuming and involves human intervention
  - Bias in tagging and hard classification of subreddits leads to inaccurate results
  - All features important to defining subreddit posts are not considered in manual tagging

Proposed Method

- Perform Dimensionality Reduction using PCA, Factor Analysis, SVD and UMAP to capture features of posts in lower dimensions
  - **PCA** – Uses orthogonal transformation to convert observations of correlated variables to linearly uncorrelated variables
  - **Factor Analysis** – Describes variability among observed, correlated variables in terms of lower number of unobserved variables called **factors**
  - **SVD** – Decomposes data into matrices that can be reduced to explain data in lower dimensions
  - **UMAP** – Nonlinearly embeds high dimensional data with fuzzy topological structure into low dimensional manifold
- Apply various clustering techniques to group subreddits and select algorithm which gives well-defined and meaningful clusters
  - **K-means++** – Partitions data into K clusters in Euclidian space with cluster centroids placed far away from each other
  - **Spectral Clustering** – Represents data as graph and clusters points based on graph connectivity
  - **Gaussian Mixture Model** – A soft clustering method that fits Gaussian components on data points
  - **Hierarchical Clustering** – A bottom-up agglomerative approach to group similar clusters
- Define conflict score (x, y) = 1 - mean(c), c = 1 if post from x does not create conflict in y, -1 otherwise
- Construct weighted directed graph G with vertices as clusters and edges with weight as conflict score between clusters

Data Description & Experimental Setup

- Number of Data points: 500,000

Feature name	Feature Description
Source Subreddit	Subreddit where link originates
Target Subreddit	Subreddit where link ends
Timestamp	Time of post
Link_sentiment	Label indicating conflict between source and target subreddit. -1 if there is conflict,1 otherwise
LIWC features	Features representing fraction of words relating to family, money, technology, etc. from Linguistic Inquiry and Word Counts Dictionary (LIWC)

- **Algorithm**:

  Step 1: Sample data to 10,000 rows

  Step 2: Normalize sampled data and check distribution and correlation of features

  Step 3: Perform feature selection based on results from Step 2

  Step 4: Perform dimensionality reduction using above mentioned techniques

  Step 5: Cluster on the low dimensional data; calculate number of points in each cluster, Silhouette score, Davies Bouldin score

  Step 6: Select clustering algorithm that gives most uniform and well-defined clusters

  Step 7: Interpret cluster meaning through feature distribution

  Step 8: Construct graph with vertices as cluster labels and edges as calculated conflict score

Results & Discussion

Dim Redctn Algorithm	Clustering Algorithm	Number of clusters	Silhouette Score	Davies Bouldin Score	Notes
PCA	Kmeans++	7	0.1912	1.6601	Low Silhouette,hard to interpret
PCA	GMM	7	0.1983	1.6237	Low Silhouette,hard to interpret
Factor Analysis	Hierarchical Clustering	5	0.4001	0.8126	Unbalanced number of points in clusters
SVD	Spectral Clustering	4	0.4028	0.8856	Hard to interpret reduced features
UMAP	Kmeans++	5	0.4172	0.8044	Good Silhouette, easy to interpret reduced features

- **UMAP** followed by **K-means++** gives **well-defined, interpretable clusters** based on LIWC feature distribution and **good Silhouette score**
- Per cluster, **dominant features** were identified based on deviation of mean from complete data and used to define cluster labels
- Cluster which talk about **Monetary and Materialistic subjects**, though in a positive sense, **triggers conflicts** or are **misinterpreted by other clusters**
- Clusters which talk about **Monetary and Materialistic subjects** and **Work-politics,** are found to cause **conflicts among themselves and other clusters**
- **Cluster 2** which has **negative emotions** as dominant feature, seems to create conflicts irrespective of source of posts, thus **Post-Control team is recommended to monitor its activities more closely**
- Posts relating to **Religion or Social Causes** are **non-controversial** among reddit communities

Takeaway Points & Future Work

- Use of clustering algorithms coupled with dimensionality reduction provide improvement over traditional tagging methods
- Latent communities were discovered to be crucial in identifying conflict magnitude and paths
- **Future Work:** Extract more text-based features using Reddit APIs, application of new clustering algorithms like **Affinity Propagation, Mean Shift** and dimensionality reduction algorithms like **T-SNE**

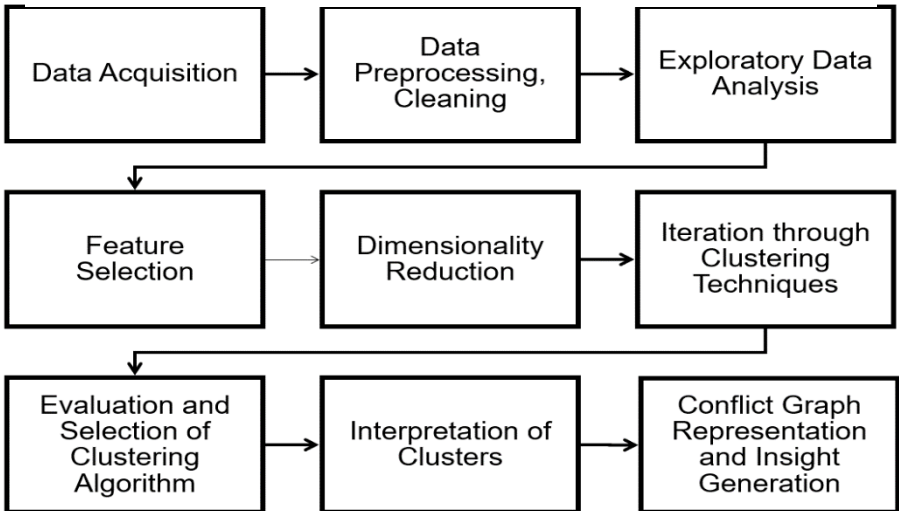Fig1: Diagrammatic representation of workflow
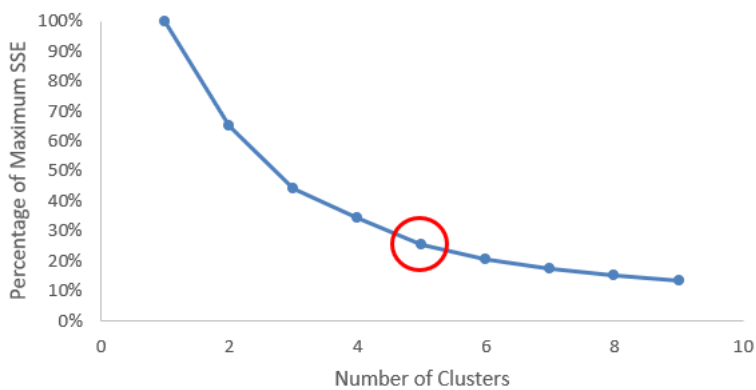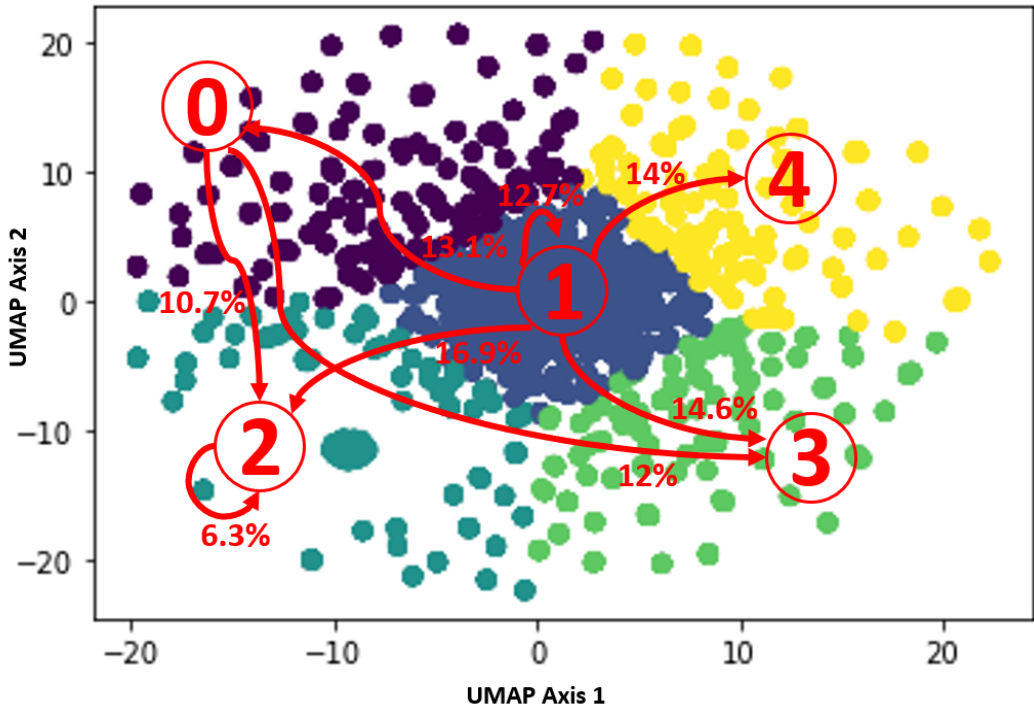

Fig 2: K-means Scree plot (K=5)

Fig3: Final cluster characteristics

Cluster	# Subreddits	Dominant Features
0	1688	Home, Money, Posemo
1	3674	Swear, Work, Achiev, Money
2	1950	Negemo, Percept, Bio, Relativ
3	1278	Religion
4	1410	Relativ, Social

Fig4: Graphical representation of Conflict scores

**[S. Kumar, W.L. Hamilton, J. Leskovec, D. Jurafsky. Community Interaction and Conflict on the Web, World Wide Web Conference, 2018.]