Multi-Modal Topic Modeling of Data and Summarization of Events and Reviews

Prajakta Sonone[1], Prof. A.V.Deorankar[2]

P.G. Scholar, Department of Computer Science and Engineering, Government College of Engineering, Amravati,

Maharashtra, India

Associate Professor, Department of Information Technology Government college of Engineering, Amravati,

Maharashtra, India

ABSTRACT: There has been big amount of growth of events over the internet in recent years. Google is the prime source of knowledge for any event happening over the internet. Google contains all the repositories of information. Some networking sites such as face book, Twitter contains millions of posts, tweets, images etc happening every day over each and every day. While on E-commerce sites, description related products; their reviews are available for the customers for analyzing various products. Various E-commerce websites like Amazon, Ebay, Flipkart are very popular sites among the users. To model this huge amount of multi-modal data having both textual and visual contents, visualization and analysis of multi-modal data is presented in this paper. While dealing with multi-modality, study of semantic relationship between the images and text data is crucial part. This model also helps to study semantic relationship between them effectively. In topic category some topics are not represented with the help images so this model also helps to point out those topics. On some e-commerce shopping sites fake reviews, advertises, spam spreading information is posted. So we have processed the reviews dataset for deciding the overall quality of products. We have collected some reviews and images from Amazon to summarize, analyze and visualize those reviews for better understanding. For event analysis the multi-modal data containing documents are downloaded from Google for mapping events data. Use of JSON parsers for processing gives the fast and quick results of search in the proposed model.

KEYWORDS: Multi-modality, Parsing, Topic Model, visualization.

I. INTRODUCTION

1.1 Motivation

Visualizing and analyzing the, information on internet is interesting job for analyzer because; it becomes easier when information is sorted and classified under specific cluster. Data mining techniques like clustering and classification are the basic steps for analyzing and visualizing the information. Information here is referred as the text and visual data. Text and visual data together considered as multi-modal data. Mapping the visual and textual contents together is difficult task because the parsing of images with relevant text takes time. Information available on social sites as well as E-commerce site contains huge repository of data, images, posts, reviews, product related information, news, advertises etc.

So we can use the multi-modality contents available on these sites for information visualization and analysis. Generally for processing multi-modal data datasets are not available for personal or public use. So creating datasets for these contents is also a complex task. In event detection for particular topics can be categorized into two types as visual representative topics and non-representative topics. This system studies this category detection In E-commerce websites like Amazon, Flipkart, EBay etc. The information related to products, their images, Reviews from the

customers who buy those products is available. We can search and sort the products based on the reviews like good or excellent product. For Mapping the text and images together we have taken the data downloaded from Google for processing. For displaying the results on user input, For showing experimental results for text and data, we have taken the Customer's reviews dataset from Amazon.com, one of the leading E-commerce website. We got the Amazon dataset in multiple pieces The behavioral analysis is carried on reviews dataset. While multi-modal mapping of data is done on document containing images and text for particular event topics.

In this paper related work is presented in Literature survey, in second part overall modules and algorithms are presented and for experimental work the snapshots of presented methods is shown in this paper. The study of Existing work has been described in details which will gives scope for research work and finally we presented conclusion and future score.

II. LITERATURE SURVEY

A lot of work has been carried out in area of event tracking and topic detection. Among them most of the methods are based on single modality information or multi-modality information. However, these models studies visual and non-visual modalities in isolation to model the multimedia event data for social media analysis.

Diakopolas et al. have propose work for studying event visualization and social event analysis by using the twitter tweets related to particular event. Extracting information from large datasets and crawling the dataset information is included in this work for social event analysis.

Makkonen et al. propose the model for extracting meaningful semantics such as names, tags, time references Based on a single clustering partition. A similarity metric have been proposed for these events.

Hierarchical Hidden Markov model has been proposed by Xie et al. over the low-level audio-visual features for discovering the location and time based i.e. spatio-temporal patterns. For finding the clusters of text, the Latent Semantic Analysis is used.

Non-negative Matrix Factorization framework was proposed by Lin et al. by using multi-relational structures for modelling the image stream data including images and short tags form social media events.

Zhang and Xu propose a CO-PMHT model to track event using cross domain knowledge and obtain summary over time from social media events.

Michele Merler in 2012 proposes a Semantic model vectors representation. In this work video event detection have been studied, which combines semantic model vectors and other static or dynamic visual descriptors by extracting the information from various frames in videos.

Topic models that are widely used for the topic modelling includes Latent Dirchlet Allocation (LDA) and probabilistic Latent Semantic analysis. These topics are extended further by introducing Supervised Latent Dirchlet allocation (SLDA).

Yang in 2015 proposes a novel cross domain feature learning framework based on stacked denoising auto-encoder. This algorithm helps to maximize correlations among various modalities and helps to extract semantic features at the same time.

Al Sumait et al. propose online LDA method, which further extends Gibbs Sampling method, which derives topic-word distribution at next time slice.

Hong et al. propose a topic model, which is time-dependent and can be used for considering multiple text sources. However, these models fail to properly model the multi-model data.

Therefore, Corr-LDA was proposed to capture correlations between image and annotations. The mm-LDA can be used for multi-model information modelling which includes textual corpora and visual topics.

III. PROPOSED ALGORITHMS AND MODULES

2.1 Proposed Algorithm

In the proposed method, mapping of images and text data is done by taking the reviews of products as text data and images related to particular product as visual contents. The Algorithm CrossSpace maps these contents together by searching exact set of keywords followed by most relevant image.

Cross Space:

Step1: Initialize, Document D containing W_d as word in topic And V_d as associated images.

i.e. visual contents.

Step 2: For Textual topic representation Initialize document D for topic distribution.

Step 3: Crawl the web page for generating effective summary of web pages using TF-IDF

Step 4: Use the Joint Topic Model for images and text data modelling.

Step 5: Stop

2.2 Modules

1. Dataset Collection:

The two types of datasets are taken for processing the multi-modal data. First type dataset is taken for multi-modality mapping. This dataset contains the collection of some documents which contains the events happening over the world downloaded from Google which contains images as well as text data in the document. These events are related to particular topics. Second type data set is taken for suspicious dense block detection. For suspicious block detection the set of reviews and Meta dataset is taken for reviews categorization.

2. Crawling of Events from Web for summarization:

This module contains the web page crawling for generating summary of events on Google. For this any url's of events on Google is taken for crawling. For generating effective summary the Term- Frequency and Inverse document Frequency algorithm is used for removing stops words for fetching useful words in the web document. It also calculates the frequency scores of top rated words
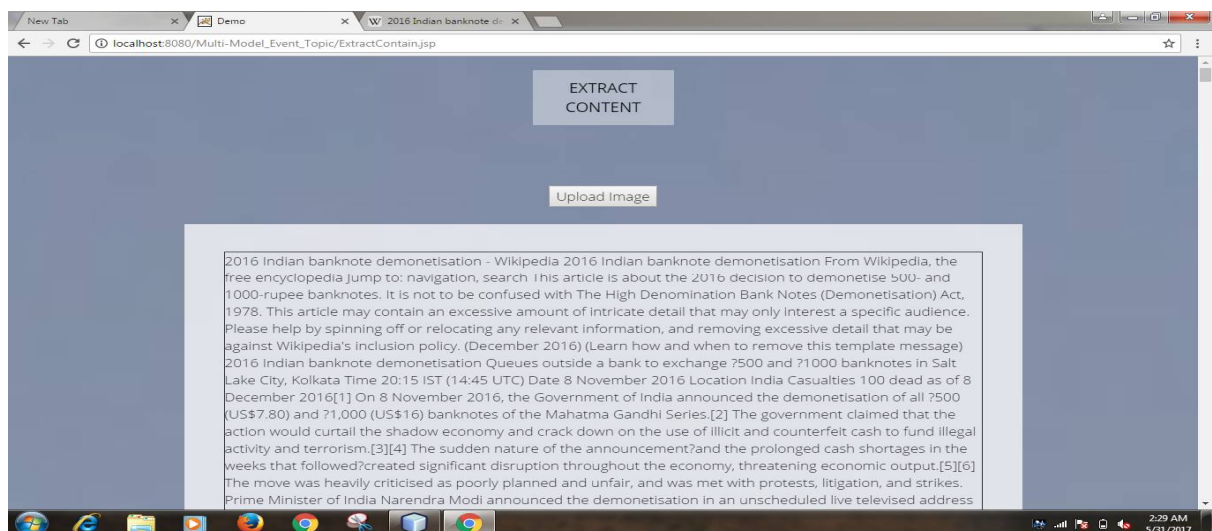
Fig.1 Crawl page for processing the summary of online web events
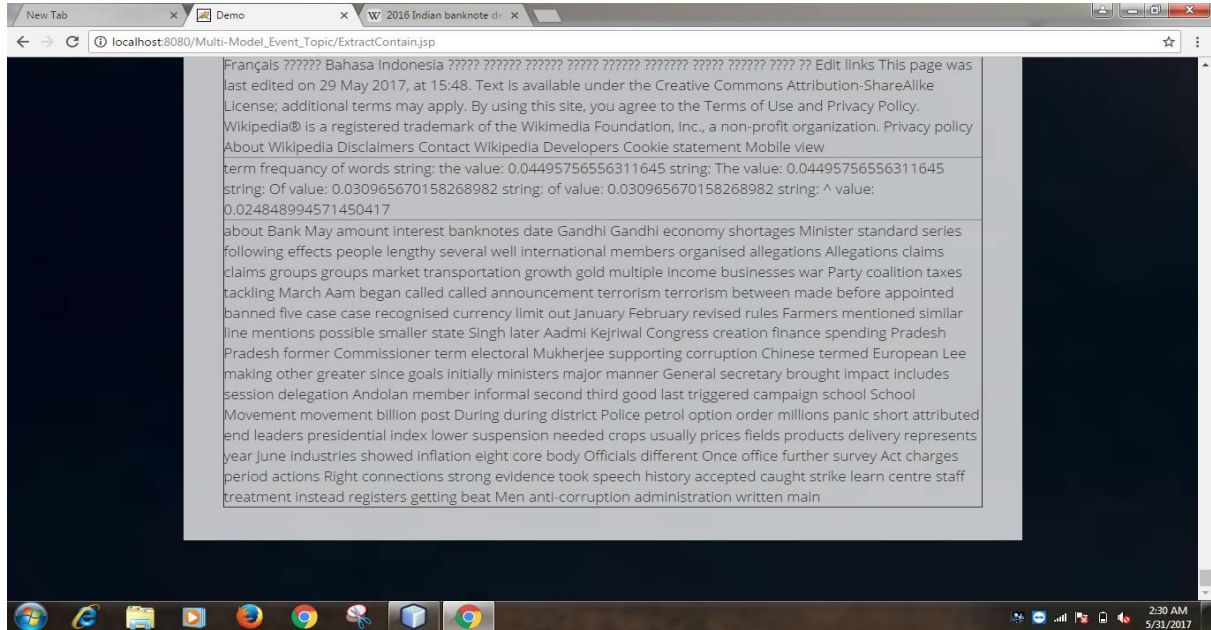
Fig.2 TF-IDF frequency computation

3. Multi-modality Mapping:

In this section the documents containing visual representative and textual topics are taken for mutli-modality mapping. The multi-modal Latent Dirichlet Allocation algorithm is used for mapping these contents together.
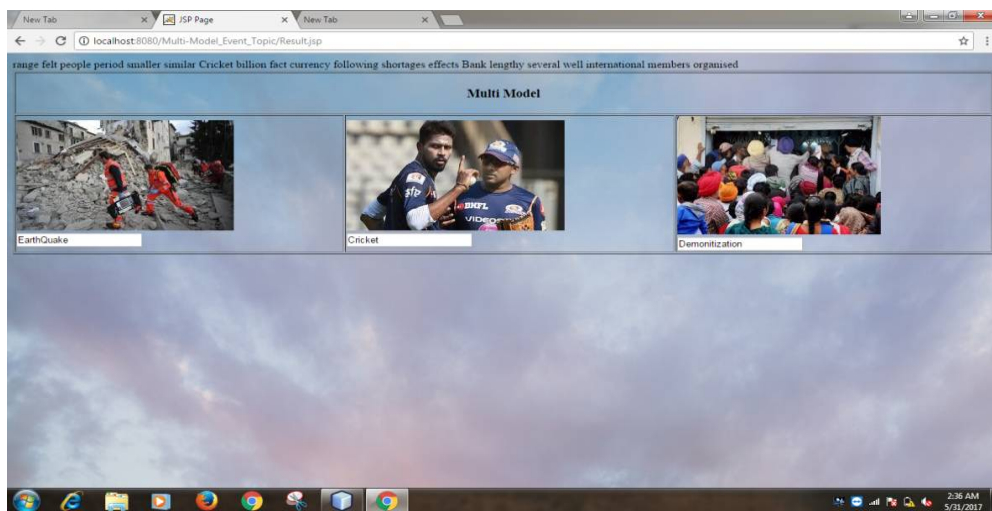

Fig .3 Experimental results after applying proposed methods for text and image modeling.

4. Semantic Mapping:

The representation of topic considering semantic relationship between textual and visual topics is also taken into consideration in this system.
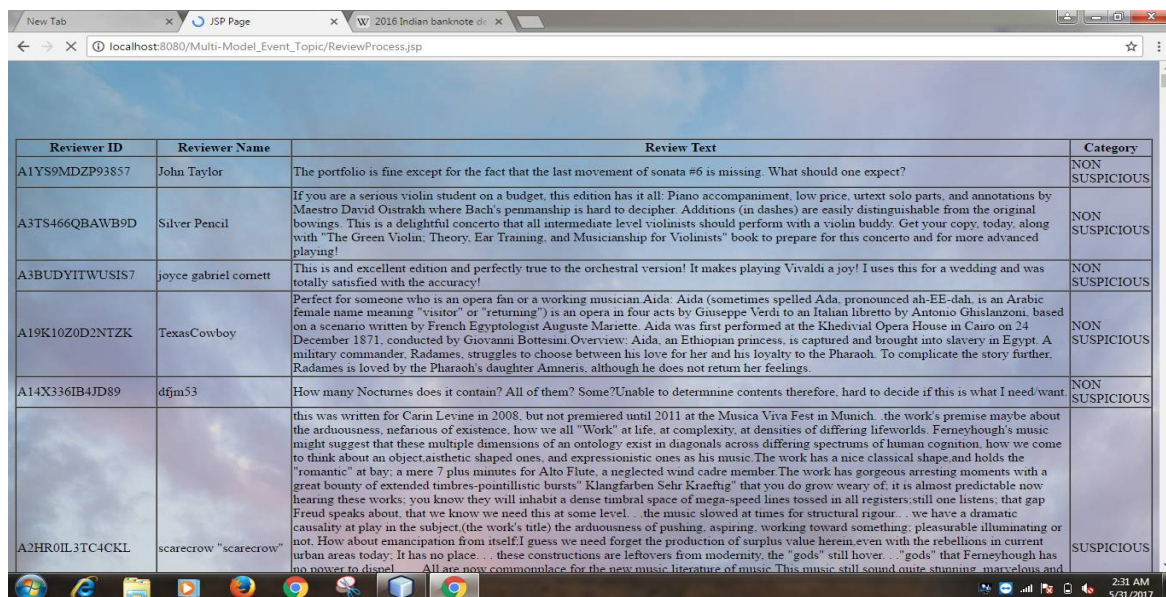
5. Process the dataset by writing parsers with suitable algorithm:

In this system suitable parsers are written using Google Gson library to parse the json objects and load it into the Java objects. These java objects are temporarily being held in the database for further operations.

6. Display, Search Sort the reviews:

Meta files which consist of product names: Reviews file consist of customers' reviews that have purchased the products from Amazon and it also contain some images of the products. The textual dataset has the raw json format. We are searching any free text in our language and applying the Cross Space Algorithm on it.
For example: This is best product I have ever seen. So the Cross Space algorithms takes the parameters as text, analyze each words what the user gives, and check its relevancy with the dataset and the concerned images. Then it creates a collection of results internally and finally displays the images and text results. Also we can
apply searching and sorting on it.

Fig. 4 Suspicious behavior of reviews

7. Suspicious Block Detection:

In this module the reviews are categorized into suspicious and non-suspicious type. For carrying this differentiation sentiments classification as well as dictionary of sentiments is created and dense blocks are considered as suspicious reviews after analyzing the reviews in dataset.

IV. COMPUTATIONAL RESULTS

The time required for overall parsing images takes 4-5 seconds. Because JSON parsers used for fetching the images from text data set and reviews from reviews data set efficiently parse the contents.

1. List of Visual representative topics after modelling topic for Demonetization, Earthquake and Cricket

1. Demonetization – {people, range, bank, Currency}
2. Cricket – {cricket}
3. Earthquake – {No topics detected as visual topics are not found in input document}

2. List of Visual representative topics after modelling topic for Demonetization, Earthquake and Cricket

1. Demonetization – {effects, shortages, several, period}
2. Cricket – {international, members, organised, billion, well}
3. Earthquake – {range, felt, smaller}

3. Computing Behaviour of reviews

 The reviews are detected as suspicious and non-suspicious after analysing the dense block at first and then by analysing sentiments in the sentences. Sentiment analysis is an integral part of aspect ranking mechanism as it finds out sentiments of aspects from product reviews, without which there cannot be existence of aspect ranking. Sentiment analysis, also called opinion mining, is the field of study that analyzes people's opinions, sentiments, evaluations, appraisals, attitudes, and emotions towards entities such as products, services, organizations, individuals, issues, events, topics.
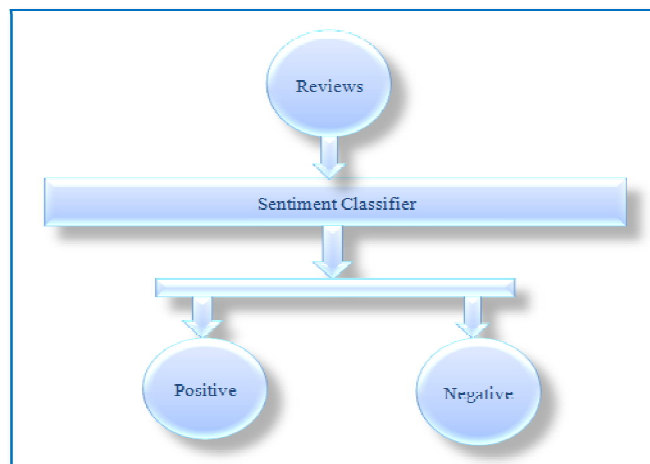
Fig. 5 Sentiment Analysis

V. CONCLUSIONS

    In this paper, multimodal data visualization analysis and summarization methods are depicted. It is suitable approach for visualizing any social media event or any document over the internet regarding to some specific topic. The JSON parser gives faster execution of the search results for mapping the visual and textual contents together. In addition, for future research we can explore whether the visualization and tracking performance can be improved by using the different domains like Flickr, YouTube and Google News for dataset collection of social media event for its detailed analysis.

REFERENCES

[1]   Shengsheng Qian, Tianzhu Zhang, Changsheng Xu, and Jie Shao, "Multi-modal event topic model For social event analysis" IEEE transactions on multimedia, vol. 18, no. 2, february 2016

[2]   M. Merler, B. Huang, L. Xie, G. Hua, and A. Natsev, "Semantic model vectors for complex video event recognition," IEEE Trans. Multimedia, vol. 14, no. 1, pp. 88–101, Feb. 2012.

[3]   T. Zhang and C. Xu, "Cross-domain multi-event tracking via CO-PMHT," ACM Trans. Multimedia Comput. Commun. Appl., vol. 10, no. 4, pp. 31:1–31:19, 2014.

[4]   X. Yang, T. Zhang, C. Xu, and M. S. Hossain, "Automatic visual concept learning for social event understanding," IEEE Trans.Multimedia, vol. 17, no. 3, pp. 346–358, Mar. 2015. N.

[5]   Diakopoulos, M. Naaman, and F. Kivran-Swaine, "Diamonds in the rough: Social media visual analytics for journalistic inquiry," in Proc. IEEE Symp. Vis. Analytics Sci. Technol., Oct. 2010, pp. 115–122.

[6]   X. Wu, C.-W. Ngo, and A. G. Hauptmann, "Multimodal news story clustering with pairwise visual near-duplicate constraint," IEEE Trans. Multimedia, vol. 10, no. 2, pp. 188–199, Feb. 2008.

[7]   I. Kalamaras, A. Drosou, and D. Tzovaras, "Multi-objective optimization for multimodal visualization," IEEE Trans. Multimedia, vol. 16, no. 5, pp. 1460–1472, Aug.

[8]   W. Liu, T. Mei, and Y. Zhang, "Instant mobile video search with layered audio-video indexing and progressive transmission," IEEE Trans.

[9]   C. Kang, S. Xiang, S. Liao, C. Xu, and C. Pan, "Learning consistent feature representation for cross-modal multimedia retrieval," IEEE Trans. Multimedia, vol. 17, no. 3, pp. 370–381, Mar. 2015.

[10] S. Qian, T. Zhang, R. Hong, and C. Xu, "Cross-domain collaborative learning in social multimedia," in Proc. ACM MM, 2015, pp. 99–108.

[11] T. Zhang, C. Xu, G. Zhu, S. Liu, and H. Lu, "A generic framework for video annotation via semi-supervised learning," IEEE Trans. Multimedia, vol. 14, no. 4, pp. 1206–1219, Aug. 2012.

[12] X. Yang, T. Zhang, and C. Xu, "Cross-domain feature learning in multimedia," IEEE Trans.Multimedia, vol. 17,

[13]   M. Jiang, P. Cui, and C. Faloutsos, "Suspicious Behavior Detection:Current Trends and Future Directions," in IEEE Intelligent Systems, no. 1, pp. 31–39, 2016.

[14] Meng Jiang, Alex Beutel, Peng Cui, Bryan Hooi, Shiqiang Yang Senior Member, IEEE, Christos Faloutsos,IEEE Trans. knowledge and Data Engineering, Manuscropt ID "Spotting Suspicious Behaviors in Multimodal Data: A General Metric and Algorithms".

[15] M. Jiang, P. Cui, A. Beutel, C. Faloutsos, and S. Yang, "Catchsync: catching synchronized behavior in large directed graphs," in ACM SIGKDD, pp. 941-950, 2014.

[16] X. Yang, T. Zhang, C. Xu, and M. Yang, "Boosted multifeature learning for cross-domain transfer," MCCAP, vol. 11, no. 3, pp. 35:1–35:18, 2015.

[17]   S. Qian, T. Zhang, C. Xu, and M. S. Hossain, "Social event classification via boosted multimodal supervised latent Dirichlet allocation," TOMCCAP, vol. 11, no. 2, pp. 27:1–27:22, 2014.

[18]   Y.-R. Lin, H. Sundaram, M. D. Choudhury, and A. Kelliher, "Discovering multirelational structure in social media streams," TOMCCAP, vol. 8, no. 1, p. 4, 2012.

[19] L. AlSumait, D. Barbará, and C. Domeniconi, "On-line LDA: Adaptive topic models for mining text streams with applications to topic detection and tracking," in Proc. ICDM, 2008, pp. 3–12.

[20] J. Makkonen, H. Ahonen-Myka, and M. Salmenkivi, "Simple semantics in topic detection and tracking," Inform. Retrieval, vol. 7, no. 3–4, pp. 347–368, 2004. L.

[21] Xie et al., "Discovering meaningful multimedia patterns with audio-visual concepts and associated text," in Proc. ICIP, 2004, pp.

[22] Y.-R. Lin, H. Sundaram, M. D. Choudhury, and A. Kelliher, "Discovering multirelational structure in social media streams," TOMCCAP, vol. 8, no. 1, p. 4, 2012.

[23] L. Page, S. Brin, R. Motwani, and T. Winograd, "The pagerank citation ranking: Bringing order to the web," in 1999.

[24] J. Sun, H. Qu, D. Chakrabarti, and C. Faloutsos, "Neighborhood formation and anomaly detection in bipartite graphs," in IEEE ICDM, p. 8, 2005.

[25] Y Gong and X. Liu, "Generic text summarization using relevance measure and latent semantic analysis," in Proc. SI