iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 7 **Issue:** VI **Month of publication:** June 2019

DOI: http://doi.org/10.22214/ijraset.2019.6251

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 7 Issue VI, June 2019- Available at www.ijraset.com

Review on Text Analytics an Approach to Artificial Intelligence

Vaishnavi D. Fate[1], Prof. A. B. Deshmukh[2], Prof. H. N. Watane[3]
[1, 2, 3]*Department of Information Technology, Sipna C.O.E.T, Amravati, Maharashtra, India*

Abstract: ***Text analytics supports organizations in managing unstructured information, identifying connections and relationships in information, and in extracting relevant entities to improve knowledge management activities. For the past decade, the number of text messages sent daily has increased by more than 7%. Younger generations overwhelmingly prefer texting to phone calls. And this is just scratching the surface, as there are many other types of textual data: support tickets, insurance application forms, healthcare records, product descriptions, and many others. Extracting meaning out of this text is an incredibly complicated task since texts may have different contexts and formats. Textual data is usually referred to as unstructured data because it doesn't have a clear storage format or a predefined data model. Sure, you can put a sentence into an Excel cell. But how would that help you to analyze it? The applications of text analysis are far and wide, from simple automation to advanced interactions between the person inputting the data and the system they interact with. A rudimentary example of that is a chatbot. This paper focuses on how text analytics is a new approach to Artificial Intelligence.***
Keywords: Text Analytics, Artificial Intelligence, Natural Language Processing.

I. INTRODUCTION

Text is one of the traditional ways of communication between people. With the growing availability of text data in electronic form, handling and analysis of text by means of computers gained popularity. Handling text data with machine learning methods brought interesting challenges to the area that got further extended by incorporation of some natural language specifics. As the methods were capable of addressing more complex problems related to text data, the expectations become bigger calling for more sophisticated methods, in particular a combination of methods from different research areas including information retrieval, machine learning, statistical data analysis, data mining, natural language processing, semantic technologies. Automatic text analysis become an integral part of many systems, pushing boundaries of research capabilities towards what one can refer to as an artificial intelligence dreams never ending learning from text aiming at mimicking ways of human learning. Dunja Mladenic, Marko Grobelnik [1]
Text mining, sometimes alternately referred to as text data mining, roughly equivalent to text analytics, refers to the process of deriving high-quality information from text. High-quality information is typically derived through the divining of patterns and trends through means such as statistical pattern learning. Arturas Kaklauskas, Mark Seniut [2] Text analytics is the process of transforming unstructured text documents into usable, structured data. Text analysis works by breaking apart sentences and phrases into their components, and then evaluating each part's role and meaning using complex software rules and machine learning algorithms. Text analytics forms the foundation of numerous natural language processing (NLP) features, including named entity recognition, categorization, and sentiment analysis. In broad terms, these NLP features aim to answer four questions:

A. Who is talking?
B. What are they talking about?
C. What are they saying about those subjects?
D. How do they feel?

Machine learning for natural language processing and text analytics involves using machine learning algorithms and "narrow" artificial intelligence (AI) to understand the meaning of text documents. These documents can be just about anything that contains text: social media comments, online reviews survey responses, even financial, medical, legal and regulatory documents. The role of machine learning and AI in natural language processing (NLP) and text analytics is to improve, accelerate and automate the underlying text analytics functions and NLP features that turn unstructured text into useable data and insights. Natural Language Processing (NLP) is "ability of machines to understand and interpret human language the way it is written or spoken". The objective of NLP is to make computer/machines as intelligent as human beings in understanding language. The ultimate goal of NLP is to the fill the gap how the humans communicate(natural language) and what the computer understands(machine language).

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 7 Issue VI, June 2019- Available at www.ijraset.com

II. LITERATURE REVIEW

When we talk about applying artificial intelligence methods on text data, what we have in mind is a whole range of methods and problems that in some way involve analysis of text data. Many of these problems have been addressed in the area of Text Mining. To be more concrete, we will briefly look into some example tasks that have been addressed in our group during the last twenty years by applying artificial intelligence methods on text data. These include:

A. Visualization of text available in news articles, visualization of named entities over time, visualization of document corpus, visualization of Web pages.
B. Triplet extraction from text, document representation using semantic graph, document summarization.
C. Text enrichment, contextual question answering.
D. Semi-automatic ontology construction from document corpus, ontology extension.
E. Knowledge extraction from text, text mining combined with social network analysis. Arturas Kaklauskas, Mark Seniut [2]

Analysis of Academic Libraries' Facebook Posts: Text and Data Analytics analyzed a dataset of academic libraries posts on Facebook. It applied a text and data analytics approach to a dataset collected from the Facebook posts of academic libraries at the top 100 English-speaking universities, as listed by the 2014 Shanghai World University Rankings. The dataset is from a two-year posting history of 18,333 unique posts, 113,621 likes, and 3401 comments. Less than a quarter of the libraries had more than 2000 post-related likes, and only seven received more than 100 comments on their postings. Content analysis identified the most prevalent single word (unigrams), bigrams (two-word sequences), and trigrams (three-word sequences) in high and low engagement content. Sultan M. Al-Daihani, Alan Abrahams [3]

Broadcast interviews have been the subject of study over some considerable time examining both news interviews as well as celebrity and talk show interviews. These studies have ranged from the collection and analysis of large amounts of interviews and subject using content analysis approaches, through to examining the media interview (both news and non- news) as a locally produced, on-going, and participant negotiated formalized interactional event Broadcast interview hosts are increasingly adopting hybrid forms of interview through the utilization of interview techniques from different genres within the one interview. Methods that can visually represent interviews in their entirety have the potential to assist in tracking and tracing genre shifts within a single interview. In this study we examine traditional genres and hybrid forms of broadcast interviewing using visual text analytic software Discursis. Discursis provides visual representations of whole interviews at-a- glance as well as the ability to focus into particular sections for closer analysis. Drawing on a corpus of 101 interviews from a single television program, this study examines if Discursis can meaningfully visually represent forms of interviewing genres and highlight where shifting techniques are used within a single interview. Daniel Angus, RichardFitzgerald [4]

Enterprise adoption of information technology (IT) innovations has been a topic of tremendous interest to both practitioners and researchers. The study of technological, managerial, strategic, and economic factors as well as adoption processes and contexts has led the field to become a rich tapestry of many theoretical and conceptual foundations. This paper provides a comprehensive multi-disciplinary classification and analysis of the scholarly development of the enterprise-level IT innovation adoption literature by examining articles over the past three decades (1977–2008). We identify 472 articles and classify them by functional discipline, publication, research methodology, and IT type. The paper applies text analytic methods to this document repository to (1) identify salient adoption determinants and their relationships, (2) discover research trends and patterns across disciplines, and (3) suggest potential areas for future research in IT innovation adoption at the enterprise level. Rahul C. Basole, C. David Seuss [5]

In the GB Railways, two systems are in operation today to exploit these benefits: in the Close Call System and the Confidential Incident Reporting and Analysis System.

The GB railways collect about 150,000 text-based records each year on potentially dangerous events and the numbers are on the increase in the Close Call System. The huge volume of text requires considerable human effort to its interpretation. This work focuses on visual text analysis techniques of Close Call records to extract safety lessons more quickly and efficiently. This paper treats basic steps for visual text analysis based on an evaluation test using a pre-constructed test set of 150 Close Call records for ‘‘Trespass”, ‘‘Slip/Trip hazards on site” and ‘‘Level crossing”.

The results demonstrate that visual text analysis can be used to identify the risks in a small-scale test set but differences in language use by different cohorts of people interferes with straightforward risk identification in larger sets. It also demonstrates how new possibilities open up to develop interactive visualizations tools that allow data analysts to use text analysis techniques for risk analysis. Esteban, Peter Hughes[6]

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 7 Issue VI, June 2019- Available at www.ijraset.com

III. PROPOSED WORK

Text analytics is used to answer business questions and to optimize day-to-day operational efficiencies as well as improve long-term strategic decisions in automotive, healthcare, and finance sector. Techniques like categorization, entity extraction, and sentiment analysis are used to identify insights, patterns, and trends in large volumes of unstructured data. Text analytics is an increasingly important task in marketing, as it can provide insights around a company's text data. This study focus on a developing Intelligence Platform that processes, analyzes and provides insights around a company's text data – i.e., surveys, call logs, social media posts, message boards, comments, etc. There are three key needs for using text analytics for market research 1 Built for surveys- A text analytics solution should be able to handle survey specifics, such people's responses that fall into the 'Nothing' category (no clear answer to the question), or 'Other' (different answer to most). 2 Value-based pricing model. A text analytics solution needs a pricing model based on the value it delivers rather than the number of transactions. 3 Easy to use A text analytics solution must be able to be operated by an in-house data analyst (with little skills in NLP) who should be able to do occasional tweaks. Infotools[7]

Proposed Intelligence Platform helps businesses work better with text. This could be as part of a decision support system, or an analytics offering, or predictions. Text is language. To understand text is to understand meaning, in a way, it is to understand the nature of being human, by being able to untangle how we communicate. To understand the meaning of a piece of text, you need to understand three separate things simultaneously. One: The syntax (what words are related to what other words, and how they act on each other), the semantics (what the words mean, and any modifications that are made), and the context (what this syntax and semantics mean coming from this person, in this place, at this time, talking about this topic).

IV. APPLICATION AND FUTURE SCOPE

Text Analytics is used to answer business questions and to optimize day-to-day operational efficiencies as well as improve long-term strategic decisions in automotive, healthcare, and finance sector. here are few real-life application of text mining.

1) *Risk Management:* Inadequate risk analysis accounts for the biggest reasons for failure in any industry. However, text analytics helps us resolve the issue of robust risk analysis.
2) *Knowledge Management:* Managing large amount of data volumes often makes finding specific information, on short notice, a difficult task. Here, knowledge management software based on text analytics offers a clear and reliable solution for the "info-glut" problem.
3) *Prevention of Cybercrime:* The random availability of data on the internet and the consequential exchanges often bear the brunt of cybercrimes. The unidentified criminal soon becomes untraceable. Enterprise and law enforcement or intelligence agencies make use of text mining techniques to analyze and observe the source and nature of data extraction.
4) *Customer Care Service:* Text analytics are widely used for customer care applications. Acceptance of text analytics software ensures improve customer experience using different sources of valuable information such as surveys, trouble tickets, and customer call notes for optimized quality, effectiveness, and speed in resolving problems.
5) *Contextual Advertising:* Text mining has given a new start to digital advertising. Companies are using text mining as the core engine for contextual retargeting for better results. Also, as compared to the traditional cookie-based approach, contextual advertising provides better accuracy, and completely safe, as it preserves the user's privacy.
6) *Business Intelligence:* Companies are using Text Mining techniques to maintain and support decision making. Text mining helps in faster and better analysis.
7) *Spam Filtering:* Spam emails are a pain area for most internet service providers, accounting for the higher cost of service management and hardware\software updating. spam is an entry point for viruses and impacts productivity. Text mining techniques are implemented to improve the effectiveness of statistical-based filtering methods.
8) *Social Media Data Analysis:* The social media which is a potential source of unstructured data is considered as a valuable source of information for market and customer intelligence. Many companies are using text analytics to analyze or predict customer needs and evaluate the perception of their brand. Text analytics can address both the issues analyzing large amount of unstructured data, extracting opinions, emotions and sentiment and their relations with brands and products.

V. CONCLUSION

Text analytics is an increasingly important task in marketing, as it can provide insights around a company's text data. In this study focus on a developing Intelligence Platform that processes, analyzes and provides insights around a company's text data. AI will provide powerful frameworks for decision support systems, and will eventually go so far as to prescribe a course of action.

©IJRASET: All Rights are Reserved

REFERENCE

[1] Dunja Mladenic and Marko Grobelnik, "Automatic Text Analysis by Artificial Intelligence", Jozef Stefan Institute, Artificial Intelligence Laboratory, Jamova 39, 1000 Ljubljana, Slovenia.

[2] Arturas Kaklauskas, Mark Seniut, Dilanthi Amaratunga, Irene Lill, Andrej Safonov, Nikolai Vatin, Justas Cerkauskas, Ieva Jackute, Agne Kuzminske, Lina Peciure, 'Text Analytics for Android Project", 4th International Conference on Building Resilience, Building Resilience 2014, 8-10 September 2014, Salford Quays, United kingdom.

[3] Sultan M. Al-Daihani, Alan Abrahams," Analysis of Academic Libraries' Facebook Posts: Text and Data Analytics", Department of Information Studies, Kuwait University, Kuwait Business Information Technology, Virginia Tech, United States.

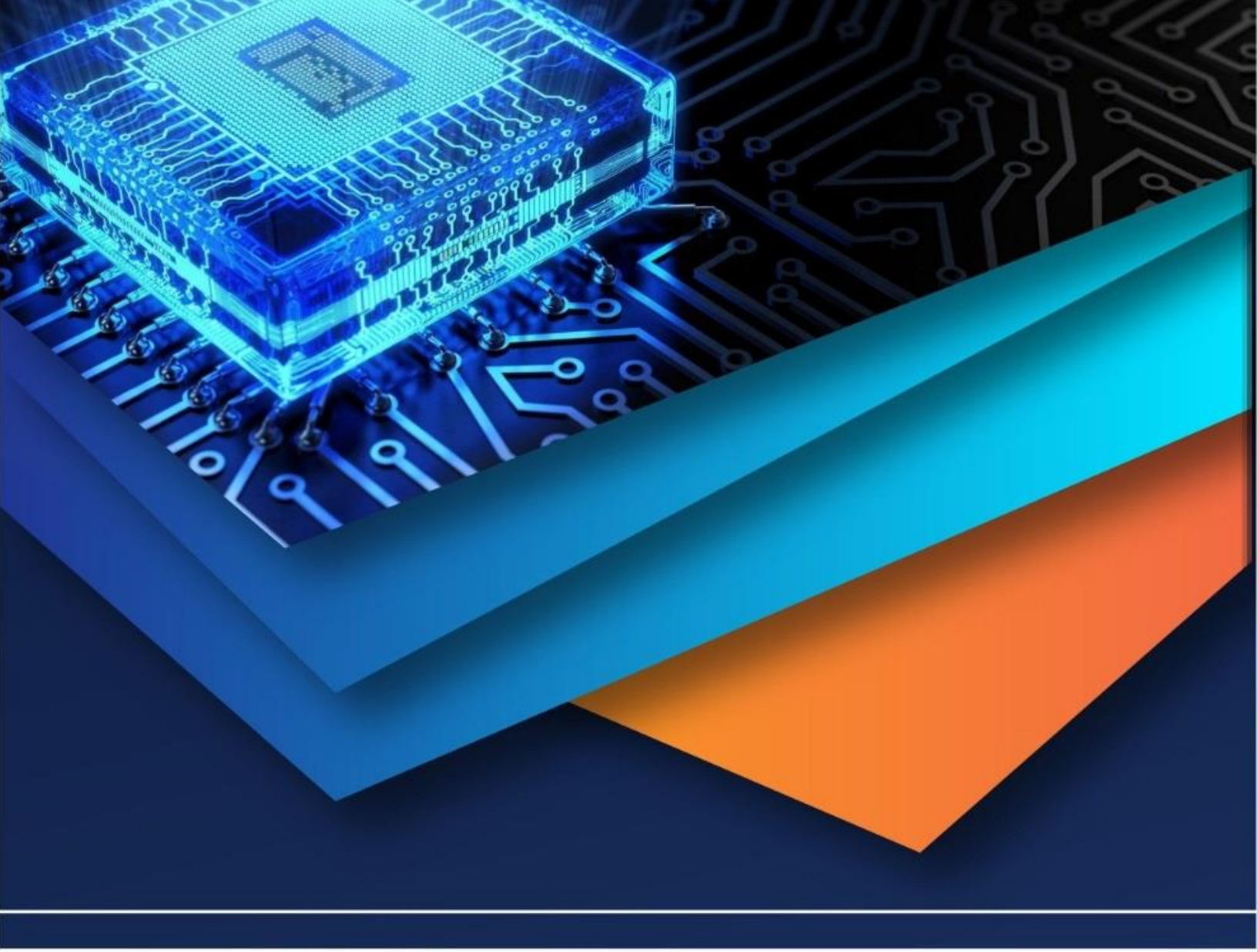
[4] Daniel Angus, RichardFitzgerald, ChristinaAtay, JanetWiles," Using visual text analytics to examine broadcast interviewing", School of Information Technology and Electrical Engineering, The University of Queensland, Brisbane, Australia

[5] Rahul C. Basole, C. David Seuss, William B. Rouse, "IT innovation adoption by enterprises: Knowledge discovery through text analytics", Northern Light Group, One Constitution Center, Boston MA 02129, United States.

[6] Miguel Figueres-Esteban, Peter Hughes, Coen van Gulijk," Visual analytics for text-based railway incident reports", University of Huddersfield, Institute of Railway Research, Queensgate, Huddersfield, UK.

[7] case study, "Text Analytics for Market Research", Infotools, an international market research technology company.

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)