Terna Engineering College

Computer Engineering Department

Program: Sem VIII

Course: Natural Language Processing

Roll No. 50	**Name:** AMEY THAKUR
Class: BE COMPS B	**Batch:** B3
Subject: NLP	**Date:** 14/04/2022

REPORT CASE STUDY ON

"A State of Art for Semantic Analysis of Natural Language Processing"

Dastan Hussen Maulud, Subhi R. M. Zeebaree, Karwan Jacksi,
Mohammed A. M.Sadeeq, Karzan Hussein Sharif

INTRODUCTION

People access social media nowadays because of the intense tendency to share their viewpoints and express their thoughts in a very casual fashion. It is important to consider the sentiment of the passage clearly. Research into word meanings is required. Most Semantic Analysis systems have been developed only in English and European languages. Natural language is the most potent form of communication. In computer science and artificial intelligence, voice is how machines communicate with humans (natural language). There is a wide range of algorithms, including sorting, clustering, text mining, and so on. NLP is this widespread nowadays and involves Chabot and text clustering. Sentiment analysis is a type of text classification where the subjectivity of the statements is essential. Opinion mining is the extraction of opinions from a large number of sources. Sentiment analysis and opinion mining are interchangeable throughout this document. It collects opinion words through the NLP. Sentiment analysis is described as detecting the sentiment of people about a specific topic and its features.

Semantic Analysis

In NLP, it is studied how to use NLP strategies to users' emotions and decide what users are expressing through them. Culture may affect this area differently, and This could be misinterpreted if it has been taken too literally. “This new gadget is bad!” Although it was evident that the title alludes to the user’s dislike of the gadget, the title might endorse the gadget to a particular age group of the community. The sentiment analysis will determine the time at which you express your opinion. To gather statements on a time axis can provide a better insight into peoples' feelings, Facebook and Twitter both provide challenges and opportunities for social movements. On the positive side, it allows people to express themselves freely. The records can be carefully observed for a specified time to study trends. The data will provide a preponderance of the evidence that supports the researcher's hypothesis.

Natural Language Processing

NLP is amongst the most complicated techniques in the world of artificial intelligence, text mining findings are inputs for NLP. NLP's capacity is that humans can speak words. It is the method of converting natural language output (spoken or written) into usable results. NLP is an exciting challenge because it requires a computer and human interaction to implement it. NLP is a field of computer studies concerned with studying and understanding the link between computers and the human language. These tools help developers design practical tech applications. Several areas of interest have been established in NLP. Therefore, the core areas' most important activities concentrate on mining named persons, extracting knowledge from texts, translating texts between languages, summarising written works, inferring answers by inference algorithms, and classifying and clustering papers.

Ontology

Ontology is the philosophy of being, which is concerned with ‘what is, i.e., the nature of reality and which is concerned with structural aspects of existence as such (Crotty, 1998) or whether it is possible to discover about the cosmos (Snape & Spencer, 2003) (Snape & Spencer, 2003). The SAGE Online Dictionary of Social

Science Methods (2006) describes ontology as "a concept concerned with the existence of, and relationships between, different aspects of society such as social actors, cultural norms and social structures." Ontology is the study of reality and our convictions about things. Ontology is the essence of the universe and the existence of truth, but it also determines what can be said about it. Bryman (2008) defines social ontology as the philosophical significance ascribed to social actors' social world. They assume what people think is deciding whether there is a truth that exists separately from human conceptions and interpretations and whether there is an everyday reality or several context-specific ones.

An overview of most recent NLP techniques

Ref	Techniques	Goal of the work	Tool	Significant Result	Accuracy
2020	Semantic Analysis	They are building interactive interfaces to annotate online documents.	Haskell X-SAIGA	The viability of accessing NL data using an event-based semantic-web triple store was demonstrated, which contained thousands of facts.	92%
2020	Semantic Analysis	Described the natural language semantic processing of ECA algorithm and research progress on preprocessing technology	Word2Vec, GloVe	XLNet on the GLUE dataset are significantly better than BERT	--
2020	Topic Modelling	Check for duplicate words in the language	Python	Improve the accuracy and recall rate	--
2020	Semantic Analysis	Transgender community	Python	Among the 300 Reddit comments,	95%

		Sentiment Analysis from Social Media data		72 are categorised as negative posts, while 85 are annotated to be positive. The remaining 48% are considered neutral sentiments. The Cohen's Kappa score is over 0.8 across all classes	
2020	Tokenization	Developing text preprocessing system	Python	All algorithms are given in the solution to produce correct results with a small amount of input data.	--
2020	Named Entity Recognition	ANswering the questions that are asked by the user's	Python	Overall Average over the entire dataset was 64% correct answer prediction	69,69%
2019	Generalized Searching	Answering questions efficiently	bAbI-10k	Single Supporting Facts 90%, two Supporting Facts 90%, and three Supporting Facts 80%, Yes/No questions 95%, List/Sets 60%.	95% for Yes/No questions, overall 83%
2019	Natural Language Generation	Utilize NLP techniques with semantic technologies.	OWL, API	The recall rate of the system, based on impressions of 69 out of 11, is about 90%. A system that is accurate to within 0.862% is very close to precise.	86%

2018	Automated Extraction	Outlined a framework for the various types of legal metadata and then based it on traffic laws.	Python	Depending on whether or not a penalty will be levied for premature or inaccurate annotation, we can obtain a precision of 97.2% to 89.5%.	Between 85.5% and 94.9%.
2018	Semantic Analysis	Construct a score of 76 different languages' similarities based on a publicly available corpus.	t-SNE	English is in the top rank with nine other languages while Chinese is at the bottom	--
2018	Supervised machine learning	Accurately parse natural language, interpret ambiguous input.	POS Tagger	Resolving coordination ambiguity, resolving lexical ambiguity, resolving semantic ambiguity is Yes, and low in terms of user interaction.	--
2017	Ontology	Translate a natural language command into an object-oriented program.	Java	Number of Questions 120, Processed Questions 116, Correct Answers 109, Precision 0.91	94%
2017	Ontology	Translate a text from one language to another.	OWL, XML, API, Java	Highly helpful! Particularly on the disambiguation and understanding measures	--

2017	Semantic Analysis	A cloud service corpus would establish a vocabulary for representing natural language. A new webpages removal algorithm is proposed	de-duplication algorithm, PageRank, Simhash	Actually repeated documents 50, Detected repeated documents 41, Detected correctly repeated documents 46, Recall rate 92%	95%
2017	Tokenization	A question answering system that converts questions in natural language	Oracle, Microsoft SQL	Presents pragmatic ways to translate database queries into natural language that users can search.	92%
2017	Named Entity Recognition	Proposed how to analyze and answer questions	DBpedia, WordNet	The F-measure of 93.43%, an average precision of 92.73%, and over 500 questions successfully administered	94.15%
2016	Ontology	Formulation and analysis of natural language software requirements specifications.	OntoGen	Software Requirements: Sample 15, Correct 10, Incorrect 04, Missing 1, Rec % 0.66 7.	71%
2016	Supervised machine learning	Evaluation techniques that use sentiment analysis and focuses on the need to address NLP open challenges. Without solving	Java, notepad ++	Require training yes, use training data to classify No, probabilistic approach No, driving factor Centroid vector, similarity metric: Vector distance,	Centroid 100%

		NLP challenges, ML techniques will not make significant advances.		strength Classify on vector distance, weakness Sensitive to noise, support for streaming data No.	
2015	Semantic Analysis	Increase the accuracy of classifying with NLP by including Word Sense Disambiguation and semantics.	WordNet	Random Forest: F-measure rate of 91%, average precision of 92%, Decision Tree: F-measure rate of 89%, average precision of 89% over 7086 classified tweets.	Random Forest: 90%, D.T:90%

CONCLUSION

In recent years, both academia and business companies have drawn interest in the IoT. It is now an important component of our lives. It has the potential to link almost everything in our world to everything else. IoT systems are complex in design and have limited capacities for storage and retrieval. The integration of cloud computing with IoT would offer multiple advantages to numerous IoT applications. We have discussed in this article the state-of-the-art cloud infrastructure, including cloud features, architecture, and benefits. The topic also centred on numerous technologies for IoT that would be expanded across the Cloud. Challenges of cloud IoT deployment and transparent problems are also discussed. In general, this paper's purpose was to include an overview to summarise up-to-date research contributions on cloud computing and the IoT and its applications in our environment and illustrate potential research directions and genuine concerns regarding the integration with the IoT of cloud computing.