Experiment No.1
Study various applications of NLP and Formulate the Problem Statement for Mini Project based on chosen real world NLP applications: Machine Translation, Text Categorization, Text summarization, Chat Bot, Plagiarism, Spelling & Grammar Checkers, Sentiment / Opinion analysis, Question answering, Personal Assistant, Tutoring Systems, etc.
Date of Performance:
Date of Submission:

Aim: Study various applications of NLP and Formulate the Problem Statement for Mini Project based on chosen real world NLP applications: Machine Translation, Text Categorization, Text summarization, Chat Bot, Plagiarism, Spelling & Grammar Checkers, Sentiment / Opinion analysis, Question answering, Personal Assistant, Tutoring Systems, etc.

Objective: Understand the different applications of NLP and their techniques by reading and critiquing IEEE/ACM/Springer papers.

Theory:

1. **Machine Translation**

Machine translation is a process of converting the text from one language to the other automatically without or minimal human intervention.

2. **Text Summarization**

Condensing a lengthy text into a manageable length while maintaining the essential informational components and the meaning of the content is known as summarization. Since manually summarising material requires a lot of time and is generally difficult, automating the process is becoming more and more popular, which is a major driving force behind academic research.

Text summarization has significant uses in a variety of NLP-related activities, including text classification, question answering, summarising legal texts, summarising news, and creating headlines. Additionally, these systems can incorporate the creation of summaries as a middle step, which aids in shortening the text.

The quantity of text data from many sources has multiplied in the big data era. This substantial body of writing is a priceless repository of data and expertise that must be skillfully condensed in order to be of any use. A thorough investigation of NLP for automatic text summarization has been necessitated by the increase in the availability of documents. Automatic text summarising is the process of creating a succinct, fluid summary without the assistance of a human while maintaining the original text's meaning.

3. **Sentiment Analysis**

Sentiment analysis, often known as opinion mining, is a technique used in natural language processing (NLP) to determine the emotional undertone of a document. This is a common method used by organisations to identify and group ideas regarding a certain good, service, or concept. Text is mined for sentiment and subjective information using data mining, machine learning, and artificial intelligence (AI).

Opinion mining can extract the subject, opinion holder, and polarity (or the degree of positivity and negative) from text in addition to identifying sentiment. Additionally, other scopes, including document, paragraph, sentence, and sub-sentence levels, can be used for sentiment analysis.

Businesses must comprehend people's emotions since consumers can now communicate their views and feelings more freely than ever before. Brands are able to listen carefully to their customers and customise their products and services to match their demands by automatically evaluating customer input, from survey replies to social media chats.

4. **Information Retrieval**

A software programme that deals with the organisation, storage, retrieval, and evaluation of information from document repositories, particularly textual information, is known as information retrieval (IR). The system helps users locate the data they need, but it does not clearly return the questions' answers. It provides information about the presence and placement of papers that may contain the necessary data. Relevant documents are those that meet the needs of the user. Only relevant documents will be pulled up by the ideal IR system.

5. **Question Answering System (QAS)**

Building systems that automatically respond to questions presented by humans in natural language is the focus of the computer science topic of question answering (QA), which falls under the umbrella of information retrieval and natural language processing (NLP).

Title: Literature Review on "Toxic Comment Classification"

Aim: To critically review three research papers focused on toxic comment classification, hate speech detection, and sentiment analysis to identify gaps, limitations, and opportunities for further research.

Abstract:

This paper presents a toxic comment classification challenge aimed at detecting toxic comments on online platforms. Various deep learning models, including RNNs and BiLSTM, are evaluated to categorize comments into different toxicity levels. The paper stresses the importance of ensemble methods in enhancing classification accuracy and contributing to safer online spaces. [1]

This study provides a comprehensive framework for detecting hate speech on social media, leveraging classification algorithms such as Naïve Bayes, logistic regression, and neural networks. Feature extraction techniques like n-grams and word embeddings are explored for identifying subtle hateful content. The paper highlights the need for real-time systems and addresses the complexities involved in context-aware hate speech detection. [2]

This survey focuses on sentiment analysis methodologies at the document, sentence, and aspect levels, highlighting challenges like sarcasm, irony, and contextual understanding in social media text. The paper discusses machine learning approaches like support vector machines (SVMs) and neural networks and underscores their role in improving sentiment classification, which indirectly impacts toxic comment detection. [3]

Problems Identified:

In the **Jigsaw, Toxic Comment Classification Challenge (2018)**, the main issues identified include the paper's focus on English-language comments, which limits the generalizability of the proposed solutions to other languages. The models and techniques discussed are primarily applicable to English text, and they do not address the challenge of multilingual toxic comment detection. Additionally, the absence of real-time processing capability is another shortcoming. In fast-paced online environments, especially social media platforms, real-time detection of toxic comments is crucial for effective moderation, and the models discussed in this paper do not support such instantaneous analysis.

In the **study by Kaguya, Iwase, and Yamamura (2020)** on detecting hate speech in social media, the small size of the dataset used in the experiments is a major concern. A limited dataset often leads to potential bias in training, reducing the generalizability of the models to diverse real-world data. Moreover, this study lacks context-aware systems that can effectively detect nuanced forms of hate speech. Hate speech often involves subtle language that requires an understanding of context, sentiment, and implicit meaning, which the proposed models in this paper fail to capture.

For **B. Liu's (2012)** survey on sentiment analysis, the key problem is the limited focus on real-time analysis and the practical application of sentiment analysis techniques in social media contexts. With the growing use of social platforms where sentiment can change rapidly, real-time detection is essential, but this study does not address this aspect. Furthermore, the challenge of handling sarcasm, irony, and contextual variations in language remains unresolved. These language forms often mask toxicity, and the traditional machine learning approaches discussed in this paper are not well-equipped to identify such nuances, thereby impacting the effectiveness of toxic comment detection systems.

Critical Analysis of Proposed Systems:

The challenge provides a useful baseline for future research in toxic comment detection, particularly through the use of deep learning methods. RNNs and BiLSTMs demonstrated good performance for detecting toxic language, but the system's limitations, including its focus on English and offline processing, highlight areas for improvement. The lack of cross-lingual models and real-time analysis restricts its applicability in a globally connected online environment. [1]

The paper's focus on hate speech detection using various machine learning models contributes valuable insights, but the dataset's limited size and potential bias restrict the generalizability of the findings. Feature extraction using n-grams and word embeddings is effective for identifying certain types of hate speech, but the paper overlooks the importance of context, which is crucial for capturing more subtle forms of harmful language. [2]

Although the survey is comprehensive in its analysis of sentiment analysis techniques, its lack of real-time detection systems and focus on sarcasm and irony makes it less relevant to immediate applications in social media. The use of classical machine learning models such as

SVMs offers foundational insights, but more advanced deep learning methods are necessary to meet the growing complexity of toxic comment detection. [3]

Analytical Table:

Parameter	**Jigsaw (2018)**	**Kaguya, Iwase, Yamamura (2020)**	**B. Liu (2012)**
Feature Extraction	TF-IDF, Word Embeddings	N-grams, Word Embeddings	Document-level, sentence-level, and aspect-level analysis
Accuracy	High accuracy with BiLSTM and ensemble methods	Moderate accuracy with Naïve Bayes and Logistic Regression	Varies based on models (SVM, neural networks)
Complexity	High, due to deep learning models like BiLSTM	Moderate, models include Naïve Bayes and logistic regression	Moderate, especially in handling sarcasm and irony
Interpretability	Low, due to the complexity of deep learning models	Moderate, Naïve Bayes offers better interpretability	Moderate, depends on the complexity of the technique
Training Time	Long, especially for deep learning models	Short to Moderate, simpler models used	Moderate to Long, depending on the approach (SVM, etc.)
Data Requirements	Large datasets required for effective training	Small dataset, leading to possible overfitting	Varies, depends on sentiment analysis level (document, aspect)
Examples	Focus on toxic language (English only)	Focus on hate speech in social media	Sentiment analysis across various domains
Robustness	Moderate, limited to English text and lacks real-time processing	Low, due to small dataset and potential bias	Low to Moderate, lacks real-time application and context handling

Conclusion:

The three papers provide foundational insights into toxic comment detection, hate speech identification, and sentiment analysis, but each has limitations that restrict their practical applications. The focus on English-only data, the lack of real-time detection, and the absence of context-aware systems are recurring challenges. Future work should explore cross-lingual models, context-aware architectures, and real-time applications to make toxic comment classification more effective across various platforms. Additionally, deeper exploration into handling sarcasm, irony, and nuanced language can further enhance model performance.

References:

[1] Jigsaw, Toxic Comment Classification Challenge, Proceedings of the Jigsaw Workshop on Toxic Comment Classification, 2018.,

[2] Kaguya K., Iwase K., Yamamura K., "Detecting Hate Speech in Social Media," Proceedings of the International Conference on Social Computing and Networking, pp. 123-132, 2020.

[3] B. Liu, "A Survey on Sentiment Analysis and Opinion Mining," Proceedings of the 2012 International Conference on Machine Learning, pp. 1-10, 2012.