

**Senior teacher, Ergasheva Nigora Kurbanovna,
student Tilavaldiyeva Nargiza Anvarjon qizi**

COMPUTATIONAL LINGUISTICS USES STATISTICAL AND MACHINE LEARNING TECHNIQUES TO EXTRACT PATTERNS

Abstract:

This study delves into the application of statistical and machine learning techniques in computational linguistics. By leveraging these methods, computational linguists can effectively extract patterns from language data, leading to advancements in natural language processing and automated linguistic tasks.

Key Words: Computational Linguistics, Statistical Techniques, Machine Learning, Language Data, Pattern Extraction, Natural Language Processing.

Introduction:

Computational linguistics relies on statistical and machine learning techniques to uncover hidden patterns within language data. By utilizing these methods, researchers can automate tasks like part-of-speech tagging, sentiment analysis, and machine translation, ultimately enhancing our understanding of language structure and semantics.

Materials and Methods:

- **Data Collection:** Linguistic data was gathered from various sources, including corpora, text datasets, and online resources.
- **Statistical Techniques:** Employed methods such as n-grams analysis, part-of-speech tagging, and sentiment analysis to extract patterns from text data.
- **Machine Learning Algorithms:** Utilized supervised and unsupervised learning approaches, as well as deep learning models like neural networks for text analysis and language modeling.

Conclusion:

The integration of statistical and machine learning techniques in computational linguistics has revolutionized the field by enabling researchers to extract intricate patterns and structures from language data. These advancements

have paved the way for the development of sophisticated NLP applications, making language processing tasks more efficient and accurate.

Statistical Techniques:

- Corpus Linguistics: Statistical analysis of large collections of textual data to identify trends and patterns.
- n-grams: Analyzing sequences of n consecutive words to predict language patterns.
- Part-of-Speech Tagging: Assigning grammatical categories to words based on statistical patterns.
- Statistical Parsing: Using statistics to build syntactic structures of sentences.
- Sentiment Analysis: Analyzing sentiment in text data using statistical methods.

Machine Learning Techniques:

- Supervised Learning: Training models on labeled data to make predictions or classify text.
- Unsupervised Learning: Discovering patterns and structures in text data without explicit labels.
- Deep Learning: Leveraging neural networks to process sequential and textual data for tasks like language modeling, text classification, and machine translation.
- Word Embeddings: Representing words as dense vectors in a continuous vector space to capture semantic relationships.

By harnessing the power of statistical and machine learning techniques, computational linguists can automate linguistic tasks, build robust NLP applications, and gain deeper insights into the structure and nuances of human language.

References:

1. Manning, C. D., Raghavan, P., & Schütze, H. (2008). Introduction to information retrieval. Cambridge University Press.



2. Jurafsky, D., & Martin, J. H. (2019). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition*. Pearson.

3. Jurafsky, D., & Martin, J. H. (2020). *Speech and language processing*. Prentice Hall.