Abstract

Aspect-Based Sentiment Analysis (ABSA) is a specialized task within Natural Language Processing (NLP) that focuses on extracting opinions tied to specific aspects of products, services, or entities. Unlike traditional sentiment analysis, which determines the overall sentiment of a review, ABSA seeks to provide fine-grained sentiment detection at the aspect level. This task becomes more challenging in low-resource languages like Hindi due to its rich morphology, flexible word order, and limited availability of annotated corpora.

This thesis presents a comprehensive study of ABSA in the context of Hindi documents. A novel and balanced dataset, TU-HSA (Tezpur University - Hindi Sentiment Analysis), has been developed to address the scarcity of annotated data. Alongside this, publicly available datasets, IITP-I and IITP-II from the Indian Institute of Technology Patna, are used for benchmarking. The research explores traditional machine learning classifiers as well as deep learning techniques for sentiment polarity detection and aspect-level sentiment analysis.

A machine learning-based ABSA framework is proposed, which includes three core tasks: aspect term extraction, aspect category detection, and sentiment polarity classification. The framework supports both one-word and multi-word aspect term detection. Experimental results demonstrate the effectiveness of the proposed approach in handling the complexity of Hindi reviews, and performance is evaluated using standard metrics such as accuracy, precision, recall, and F1-score.

Further, a deep learning-based method is presented for aspect term extraction, which significantly improves the accuracy of identifying opinion targets. An

aspect category-based sentiment aggregation model is also proposed to compute sentiment polarity at the document level by combining sentence-wise aspect category sentiments.

This work not only contributes valuable annotated resources but also proposes technically sound and scalable methods for ABSA in Hindi. The outcomes of this research are expected to benefit sentiment-driven applications such as opinion mining, customer feedback systems, and policy analysis in Hindi and other Indian languages.

Keywords: Hindi, aspect-based sentiment analysis, aspect term, machine learning, deep learning, performance metrics, aspect category detection.