AWS Lambda and SageMaker: Real-Time Solutions for Machine Learning

Himanshu Gupta

Lead Software Developer, CAT, FINRA, USA

Abstract*: Natural Language Processing (NLP) has emerged as a transformative power during the era of significant data. The vast and diverse vocabulary generated by big data allows for the development of advanced NLP models through machine learning algorithms and distributed computing methods. The collaboration of NLP and big data has led to the creation of strong language models such as BERT and GPT, enhancing NLP's capability to comprehend content and provide insights across a range of applications, including reasoning, machine translation, automated responses, and personalized NLP text generation.*

Integrating NLP within a big data framework addresses numerous challenges across various industries. Utilizing data collection, business intelligence can provide real-time insights, while collaboration can be enhanced through improved networking. Sentiment analysis aids in refining product and market research by enabling organizations to understand customer opinions and preferences. Cloud services like AWS Lambda simplify serverless execution of NLP workflows, facilitating cost-effective, scalable, and event-driven processing of large datasets. Moreover, AWS SageMaker streamlines the development, training, and deployment of machine learning models, enabling the efficient creation and operationalization of advanced NLP solutions on a large scale.

This study demonstrates the efficacy of NLP in processing large datasets, particularly for sentiment analysis using the MapReduce framework, alongside the operational advantages offered by AWS Lambda and SageMaker.

In conclusion, this document emphasizes the opportunities and challenges that arise from the integration of NLP with big data and cloud-native tools, providing insights into their combined capabilities. This potential can be harnessed in various ways, fostering a deeper understanding of both language and data analysis.

Keywords: AWS Lambda, SageMaker , Natural Language Processing , MapReduce