**Name of student:** Deep Salunkhe

**Roll Number:** 21102A0014

**Branch:** CMPN

**Name of article or IEEE paper:** Linear Discriminant Analysis for Large-Scale data : Application on Text and Image data

**Year of publication:** Dec 2016 @ [2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA)](https://ieeexplore.ieee.org/xpl/conhome/7835817/proceeding)

**Review**

The paper "Linear Discriminant Analysis for Large-Scale Data: Application on Text and Image Data" by Nassara Elhadji Ille Gado, Edith Grall-Maës, and Malika Kharouf addresses the challenge of handling high-dimensional data, especially in the context of text and image datasets, where the number of features can reach hundreds of thousands or more.

The authors propose an efficient algorithm called Fast-LDA to tackle the computational complexity associated with Linear Discriminant Analysis (LDA) for large-scale data. The proposed approach leverages random projection for dimensionality reduction, aiming to reduce the computational burden while maintaining discriminative features that preserve class separability.

The paper begins with an introduction outlining the significance of handling large-scale datasets due to the curse of dimensionality, emphasizing the necessity to reduce the data space to improve computational efficiency and algorithm performance. It discusses common methods for dimensionality reduction, highlighting principal component analysis (PCA) and random projection (RP).

The core of the paper explains LDA fundamentals, emphasizing the objective to find a projection matrix that maximizes the ratio between the between-class scatter and the within-class scatter. It then introduces the Fast Approximate SVD algorithm as part of the proposed approach, emphasizing its role in achieving a low-rank approximation of the original space, thereby facilitating efficient LDA computation.

The proposed Fast-LDA method combines feature extraction using Fast Efficient SVD with classical LDA in the reduced feature space. The paper includes detailed algorithmic steps and notations for Fast Efficient SVD and the overall Fast-LDA method.

Experimental results are presented using real-world datasets for text and image classification. The evaluation compares Fast-LDA with other methods such as LDA/QR, NovRP, and SRDA, showcasing the accuracy rates and computational times across different training set sizes and dataset dimensions. Results indicate that Fast-LDA achieves competitive accuracy while significantly reducing computational time, especially in scenarios involving large training sets and high-dimensional data.

The conclusion summarizes the contributions of the proposed approach, highlighting its stability and efficiency in both high and low-dimensional spaces. The authors acknowledge the support received for this work and provide a list of references for further reading on related methods and concepts.

In summary, the paper introduces a novel approach, Fast-LDA, addressing the challenges of handling large-scale data for discriminant analysis. It offers a comprehensive overview of the methodology, experimental results, and implications for handling high-dimensional datasets in the context of text and image classification.