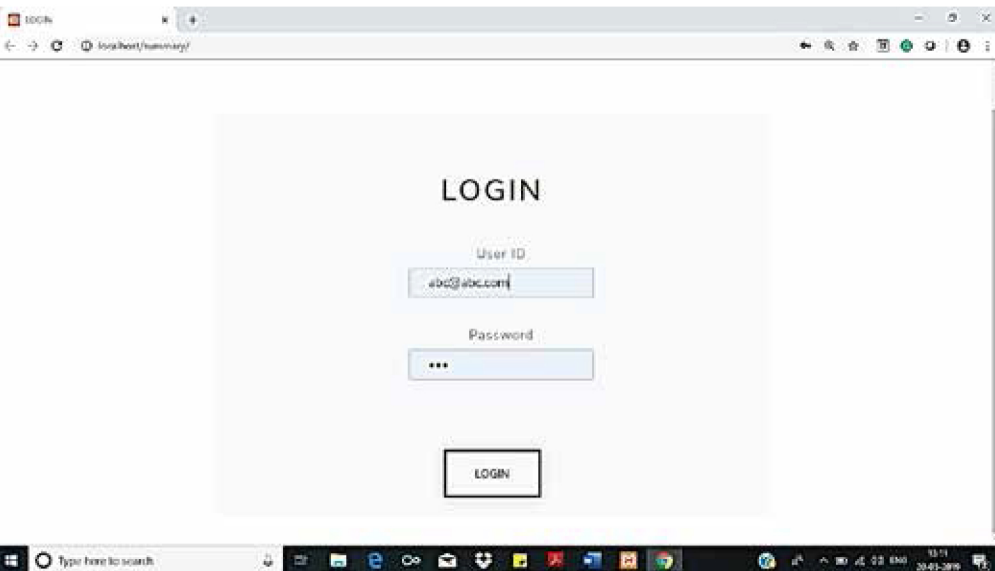


A theme provides brief information about a document. Effective and efficient theme extraction can help users sift through the documents with ease, given the context of information explosion in recent times. There are two ways to design a theme for a document: Abstractive approach and Extractive approach. Abstractive summary approach gathers information by inferring the content and building the new summary. Extractive summary methods identify the important topics for a given document, and design the theme using sentences from the topics identified without altering the original text.

In Clustering Algorithms-based methods, the text is organised in groups. In Lexical Chains-based approach, the semantic relationship between words computed is used to find cohesive links and the sentences with higher links are used to build the theme for the document. But in these approaches the anaphora problem still persists. The proposed method uses synthesis of weighted algorithms LDA and PLSI and overcomes the anaphora problem and generates summary that is more coherent.



Shilpa  
CSE

# 68 PROJECT

THEME  
DESIGNING  
FOR TEXT  
DOCUMENT  
USING  
WEIGHTED  
ALGORITHMS