Multi label text classification (MLTC) or text categorization is an important subfield in natural language processing (NLP) it allows to classify documents ,not in one category like single class (binary or multi-class) classification but in several predefined topics simultaneously, So, the goal of MLTC is to learn from set of TRAINING DOCUMENTS, where each DOC belongs to one or more labels at the same time.

Many problems are associated with MLTC AS the huge number of labels combination, unbalanced data, the non-exploitation of relationships between labels(correlation and diversity),and so on.

BUT SOME HOW SOME COMBINATIONARE IMPOSSIBLE IN REALITY,WE have to propose a solution to exclude them.like in music categorization ,we can t classify a piece as classical and hard rock.

several approaches have been proposed for mltc, the very first one Binary Relevance (BR) treats multiclassification as a multiple independent binary classification

Several issues …

Talk about **classical solutions** in mltc.

Transform the problem in single-label classification tasks and then combine the results **(BR, Ranking via Single-Label Learning, Label Powerset (LP), Pruned Sets)** or adapts the existing ML methods to handle multi-label data directly.

Recently, MLTC, has seen a lot of progress through the appearance of methods based on embeddings layers for the data representation and sequence to sequence models.

*Word embedding technologies, a set of language modeling and feature learning techniques in natural language processing (NLP), are now used in a wide range of applications.*

This paper deals with multi-label document classification by deep neural networks.

In this paper , a deep neural network is proposed, which based on the transformation of document at a specific embeddings layer that preserve relation between set of features and the set of labels.

The rest of this paper is organized as follows. In Section II, definition of multi-label learning is given and previous works in this subfield are reviewed. In Section III, our method is presented. In Section IV, experiments on real-world multi-label learning datasets are reported. Finally in Section VI, the main contribution of this paper is summarized.