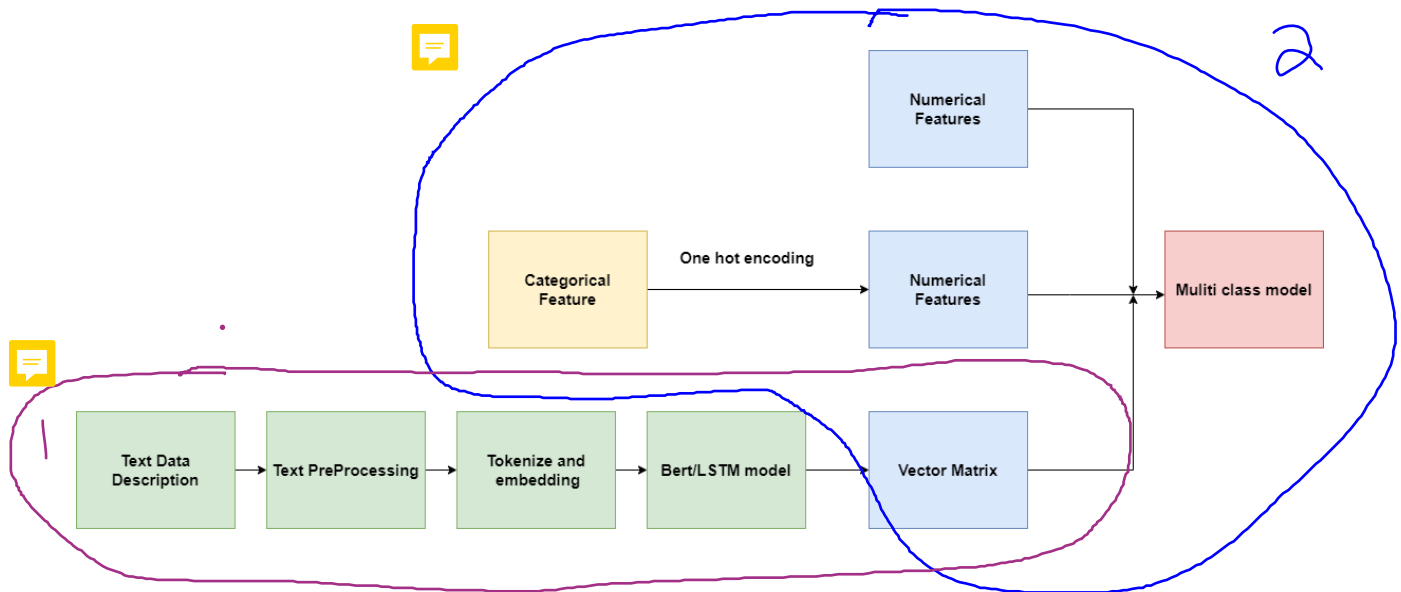


# Classification Model on PartnerShip Data:

Data in our problem statement contain both Text Data and some numerical data. In order to combine them to build a multi-class model we have to convert each feature/column to numeric format.

The suggested method is shown in the diagram:



1. Numeric data:  
Numeric data will be directly passed to the multi-class model
2. Categorical data:  
All the categorical data will be converted into each new dummy column by one hot encoding.
3. Text Data:  
Text data present in the will undergo the following steps:
  1. Text Preprocess(Remove stop words, lemmatize, remove description with fewer tokens/words)
  2. Tokenize text present in each word
  3. Perform Embedding:  
There are different text embedding techniques like  
Short Explanation [Link](#):
    1. [TF-IDF](#)
    2. [WordToVec](#)
    3. [Glove method](#)
    4. [BERT\(Bidirectional Encoder Representations from Transformers\)](#)

5. Video: [https://www.youtube.com/watch?v=gFwr2KxZ1pc&list=PLZoTAE\\_LRMXVN\\_NrHSKv36Lr3\\_156yCo6Nn&index=15](https://www.youtube.com/watch?v=gFwr2KxZ1pc&list=PLZoTAE_LRMXVN_NrHSKv36Lr3_156yCo6Nn&index=15)