Change dataset:

* Samples for use cases:
  + A minimum count of 1000 sentences per use case needs to be created in order to have a reliable model that functions for variant inputs.
  + Repetition of similar sentences over different use cases should be avoided.
  + Redundancy of input statements should be avoided.
  + For similar sentence structure at least 10 variations should be fed to the dataset.
* Variations to the sentences:
  + Variety of sentence structures needs to be added to the dataset.
  + Larger dataset and different use of words will lead to broader problem tackling capability.

Annotation:

* Time consumption:
  + Number of statements is directly proportional to the time taken.
  + Each word in each statement needs to be tagged to a particular entity.
  + Manual labor requirement is high for annotation purpose.
* Tags:
  + All the entities or tags that are to be extracted are to be pre decided.
  + Variety for each tag needs to be introduced in terms of:
    - Sentence structure
    - Word vocabulary (Synonyms)

Note:

* As we are using LSTM layers so sentence sequences will matter a lot.
* The word dependencies on its neighbors is highly valued.
* Wrong sentence structure will lead to high error rates.
* Hyper parameter tuning for training purposes is dependent on:
  + Size of dataset
  + Variations in dataset
  + Hardware used for training