

## Proposed Approach

- We observed that the dictionary  $D$  did not vary much between the iterations and hence the solution was not getting updated. In order to overcome this problem, we decided to add minimal supervision in the form of a seed dictionary. The seed dictionary contains hindi-english word aligned pairs.
- We build a dictionary in an unsupervised manner (explained above) and then start the self learning process. After every iteration we select a subset from the seed dictionary and for each pair in that subset we make the corresponding entry in the dictionary = 1.
- The hyper-parameter here is the subset size. If the subset size is too small then the dictionary wouldn't vary much and if the subset size is too big then the solution may not converge leading to divergence. By adding the supervision after each iteration, the dictionary gets updated significantly and the solution also gets updated.
- We incorporated this approach in the model with and without CSLS and found that the results are better than the previous models.