

## **Approach Followed**

First of all, the dataset was cleaned through pre processing techniques such as tokenization, lemmatization, removal of stop words, removal of hyperlinks(if any) and removal of special symbols and numbers. Then, a similarity is drawn in between the two datasets and it is found that Technical Skills are those where the first letter of all words of the 'skill' are capitalized. Which is why, a function is created to put a flag on skills who are having all first letters of the words in them as capital, and subsequently classify those skills as technical skills(1). Lastly, some of those technical skills are displayed on the screen. Hard Skills are 12990 in total.