Assignment for Week 3

1. Train a custom embedding model for Game Of Thrones data. You can use complete books as your training corpus. The tasks will include - Parsing pdfs, Preprocessing, Training.

2. Normal training would give you vectors for individual words - Can you preprocess your training data in a way that you get one vector for Named Entities for multi word Named Entities like **Jon Snow**, **Daenerys Targaryen** etc ? Use this preprocessed text to train your new embedding model

3. Show some interesting relationships between vectors obtained from your embedding model. Ex : **Winterfell** and **Casterly Rock** should be ideally similar

4. Once you have your embedding model, you can get a vector for one word. Each word can now be treated as a vector. Now imagine a sentence having 'n' words. Can you think of operations(linear or nonlinear) on these vectors to get the embedding. Create a function that can take various operations as input and use those operations on the individual word vectors to get the final sentence vector.