Name: - Udbhav Solanki

Id: - 0873798

Things done in project: -

- In this project I have develop word embedding utilities in Julia using APIs like Word2Vector and text Analytics. And these utilities are used I applications like Compute similar words, Create a group of related words, Feature for text classification, Document clustering, Natural language processing.
- Word2Vec is a model for the creation/forming of word embeddings. It is a modern way of representing the word, where each word is represented by a vector (Array of numbers based on Embedding Size). Vectors are nothing but neuron weights, so we would have 100 weights if we set neurons of size 100 and those weights are our Word Embeddings or simply Dense Vector.
- Word2Vec seeks a relationship (semantic or syntactic) between the words that our conventional TF-IDF or frequency-based approach did not make possible. Each hot-encoded word gets a point in a dimensional space when we train the model, where it learns and groups the words with a similar meaning.
- Word2Vec is a technique for constructing such an integration. Two techniques (both involving Neural Networks) can be used to obtain it: Skip Gram and Typical Bag Of Words (CBOW)
- The method of drawing meaning from written correspondence is Text Analytics. Text analytics may be done manually, but it is an inefficient operation. Therefore, text analytics software has been developed to find meaning in vast quantities of text using text mining and natural language processing algorithms. Text analytics software may also provide an early warning of trouble, since it reveals what consumers are complaining about.