**Computation Argumentation Assignment 2**

**Group Name :** Artificial Minds

**Members:** Vikrant Singh, Akshay Pandey, Amit Kumar, Barun Kumar

**Submitted to** : Milad Alshomary; Maximilian Spliethöver

**Approach** :

1. Created a **dataframe** df for getting the train data and df2 for test data.
2. Used **nltk.pos\_tag** to get the part of speech tag for each sentences and then appended that result into the data frame as an added feature.
3. Tags that were initially considered : **Noun** , **Verbs** and **verb** **Phrase** and verb tags like **VBD,VBZ,VBG,VBN** , then **preposition** , **conjunction** , **modal** auxiliary , **determiners** , **adjectives** and **adverbs**.
4. Other point that was considered , was the usage of ‘**that’** word , as most like when ‘**that’** word is present , its likely to be a **claim.**
5. Then after collecting these features , it was appended in the dataframe.
6. Then , **Word2Vec** embedding was used as an extra layer , to be passed onto the classifier.
7. **Word2Vec** was used as the embedding preserves the semantic nature of the words of sentences.
8. At the end , **BiLSTM** classfier was used , as most of the research papers showed that biLSTM gives better accuracy .
9. **Two input layers** were passed to the classifier , one being the word embedding and other the features extracted using nltk POS TAG
10. Some of the features were not used at the end , as they were inhibiting the classifier accuracy .
11. A dictionary is created to store the data in the json format .

**Approach tried but didn’t submit –**

1. Paper[Context-Independent Claim Detection for Argument Mining] suggested that , When a sentence to subjected to a parse tree , claims have some common CFG Spanning features . S->NP VP  
VP-> VBP SBAR …. So on , so we tried using that and created the tree using nltk lib , but at the end , it didn’t helped our accuracy .