

Reflective Diary Template

Group Number: 146

Full Name	Zilong Zheng
Role in group:	member
Role responsibilities:	<ol style="list-style-type: none">1. Create tokenization function.2. Create porter stemmer function3. Extract the abstract, title and authors.4. Find the first 200 meaningful bigrams5. Create the output file6. The part of report
Contribution to Group:	<ol style="list-style-type: none">1. Create tokenization function.2. Create porter stemmer function3. Extract the abstract, title and authors.4. Find the first 200 meaningful bigrams5. Create the output file6. The part of report
Learnings from Project	How to use nltk library and learn how to calculate the frequency of word in over file.
Learning Techniques	Regular expression NLTK library Porter stemmer Find the bigram
What went well	Tokenization the whole paper.
What went wrong?	
Resolution to solve issues:	
Overall Conclusion: perfect. We are very happy to work together. Very satisfied with the assignment How would you do it, if asked to do it again? YES	

Full Name	Songyuan Liu
Role in group:	member
Role responsibilities:	<ol style="list-style-type: none"> 1. Download the information from pdf and transfer them into the TXT file. 2. Extract the body part. 3. Calculate the frequency of words in over file. 4. Remove context-independent and context-dependent stopwords, rare tokens and tokens(length <3). 5. Create the output file. 6. The part of report
Contribution to Group:	<ol style="list-style-type: none"> 1. Download the information from pdf and transfer them into the TXT file. 2. Extract the body part. 3. Calculate the frequency of words in over file. 4. Remove context-independent and context-dependent stopwords, rare tokens and tokens(length <3). 5. Create the output file. 6. The part of report
Learnings from Project	How to use nltk library and learn how to calculate the frequency of word in over file.
Learning Techniques	Regular expression NLTK library Porter stemmer Find the bigram
What went well	Count the frequency of the words in over file
What went wrong?	
Resolution to solve issues:	
Overall Conclusion: Happy cooperation, clear division of work How would you do it, if asked to do it again? Yes	