

VNIT Final Year Project Tracking

Last Updated: 6th March, 2020

Time Period	Task	Status	Remarks
PROJECT PHASE 1			
VII Semester - Sessional 1	1. Getting started	All completed	
	2. Performing Named Entity Recognition (NER) using spaCy and semi-structured statement extaction using textacy on wikipedia text		
VII Semester Sessional 2	1. Using financial reports	All completed	
	2. Training NER model on custom data		
	3. Comapring Stanford NLP library with spaCy for our task		
	4. Applying Coreference Resolution and demonstrating the spaCy pipeline and textacy output		
VII Semester End Sem	1. Using dependency parse tree to extract semi-structured statements due to issues faced by available methods	All completed	
	2. Demonstrate the work done through a web application with interface (where user can upload a document and write the subject) to end project phase 1		
PROJECT PHASE 2			
VIII Semester Sessional 1	1. Using Stanford openIE to extract factual information (semi-structured information)	All completed	
	2. Using verbs only, determining sentences that indicate positive growth or negative growth through various approaches and comparing them		
VIII Semester Sessional 2	1. Using verbs and phrases, determining sentences that indicate positive growth or negative growth using this approach	Working on these tasks	Extracted date on which document is created using the metadata associated with the document. Additionally performed NER to get dates present in the document. The date which is mentioned the first is assumed to be the date on which the document is published. (Both the above dates are compared, and in case of conflict, the date written in the metadata is chosen)
	2. Extracting timestamp of the documents (date on which the document was published)		

VNIT Final Year Project Tracking

Last Updated: 6th March, 2020

[illegible]