Weeks Report (2018/02/15 – 2018/02/28)
Issued by: Sijin Cheng
2018-02-28

Time Slot 1

What has been done in the past two weeks?
Last week, our focus has been to make the thesis project start and get a clearer picture of what to deliver.

In short:
1. Has finished my thesis plan
2. Has contacted the examiner, Professor Arne, and got the approved for my thesis plan in WExUpp
3. Has a discussion by Slack with Mr. Falah that delivered a clear overview of his expectations
4. Has a video meeting with Mr. Falah on the details regarding the search query and the ranking principles
5. Has set up Python environment on Anaconda
6. Has preprocessed the Patent files with Scikit-learn: (tokenize, remove stopwords and punctuations, stemming, and build tf-idf matrix, clustering with KMeans)

What is planned to be done in the next two weeks?
Focus: Write the first draft of introduction part and theory part, find the appropriate cluster algorithm

In short:
1. Compare different cluster algorithms and choose the appropriate one
2. Find the method to consider weight into while clustering
3. Version 1 of the introduction part will be finished
4. Conclude the key techniques and write the theory part

Which problems exist?
We currently not have any problems with our project.

Time report
Is attached to the next page of this document.

SIJIN CHENG

Master thesis

TOTAL TIME (h)	45
hours needs to be spent to fulfill 800 hours	755

First Month: 2018.2

Week 1 (~-2.21):				24	Week 2 (2.21-12.28):				21	>=40
Date	Time (h)	Activity	Description of activity	participants	Date	Time (h)	Activity	Description of activity	participants	
1.28-2018	2	education	prepare for the first me	Me	22.2-2018	1	meeting	the ranking principles	Me, Mr. Falah	
1.29-2018	1	meeting	discuss the scope of th	Me, Mr. Falah	22.2-2018	3	environment	set up python on Anacor	Me	
31.1-2018	0.5	meeting	discuss the process of t	Me, Prof. Marco	23.2-2018	2	package decis	learn different packages	Me	
1.2-2018	3	education	literature research	Me	23.2-2018	3	education	scikit-learn: learn and pra	Me	
1.2-2018	2	writing	draft of thesis report	Me	26.2-2018	1	education	search more info of Pate	Me	
1.2-2018	2	writing	initial time plan	Me	26.2-2018	4	coding	process one patent abstr	Me	
2.2/2.15-2018	3	email	discuss the application	Me, Prof. Marco, Pro	27.2-2018	2	education	learn clustering algorithm	Me	
15.2-2018	3	education	literature research	Me	27.2-2018	4	coding	tf-idf matrix, KMeans	Me	
15.2-2018	2	writing	revised thesis report	Me	28.2-2018	1	writing	weeks report	Me	
15.2-2018	1	writing	revised time plan	Me						
19.2-2018	1.5	slack	how to access the data	Me, Mr. Falah						
21.2-2018	1	writing	organized the custome	Me						
21.2-2018	2	education	prpare for the meeting	Me						