



Final project proposal

Type	Undergraduate
Title	<i>Text clustering</i>
Supervisor	Prof. Igor Trajkovski
e-mail	trajkovski@finki.ukim.mk
Department / Group	
Intelligent Systems	
Topic(s)	
Natural Language Processing, Machine Learning	
Project can start from	01.04.2014
Project duration	4 months
Short description	
<p>The objective of clustering is to partition an unstructured set of objects into clusters (groups). One often wants the objects to be as similar to objects in the same cluster and as dissimilar to objects from other clusters as possible. Clustering has been used in many different areas and there exist a multitude of different clustering algorithms for different settings.</p> <p>This thesis should explore a ways in which text documents can be clustered. Whether adding background knowledge from ontologies can be added to the model in order to improve the quality of the clustering should be explored in this thesis.</p>	
Results and assessment	
<p>The outcome of this project should be a written report and oral presentation. The elements of the report should include: introduction, review of literature, procedures used for collecting and analyzing data, findings, summary and conclusion.</p>	
Other (additional) information	
<p>http://charuaggarwal.net/text-cluster.pdf http://people.aifb.kit.edu/aho/pub/hotho_etal_techreport425.pdf http://www.time.mk/trajkovski/</p>	