



Final project proposal

Type	Undergraduate
Title	<i>Text classification</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>This thesis should explore a ways in which documents can be automatically classified based on their contents. Automatic classification of data is one of the main applications of machine learning. With the help of already classified data a model for the most likely class can be learned.</p> <p>Whether adding background knowledge from ontologies can be added to the model in order to improve the classification accuracy, 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://www.charuaggarwal.net/text-class.pdf http://nmis.isti.cnr.it/sebastiani/Publications/TM05.pdf http://nltk.googlecode.com/svn/trunk/doc/book/ch06.html http://www.time.mk/trajkovski/</p>	