



# Syllabus przedmiotu

w ramach projektu „@kademia na rzecz rozwoju ICT – studia podyplomowe”

Rok akademicki 2010/2011

Nazwa przedmiotu	<b>Text Mining / Opinion Mining</b>
Kod przedmiotu	TXM

## 1. Opis

Nazwa kierunku	Metody Informatyczne w Badaniach Społecznych	Kod kierunku	MIBS
Jednostka prowadząca	Wydział Informatyki, Warszawa		
Prowadzący przedmiot	Mgr Aleksander Wawer		
Semestr studiów	Pierwszy		
Język wykładowy	Polski		
Plan godzinowy	Wykłady	12	
	Laboratoria	14	
Warunki zaliczenia	zaliczenie		

## 2. Wymogi

Założenia i cele	Poznanie zastosowań technik text miningowych do eksploracji danych tekstowych.
------------------	--



	<p>Przedstawienie metod automatycznego przetwarzania tekstu w zastosowaniach związanych z analizą nastawienia i opinii, a zwłaszcza recenzjach produktów i usług.</p>
Treści programowe	<p><b>1. Podstawy automatycznego przetwarzania tekstów pod kątem analizy treści i opinii</b></p> <ul style="list-style-type: none"><li>• Wstępne przetwarzanie. Angielski, języki fleksyjne.</li><li>• Teoria i główne podejścia do ilościowej analizy tekstów.</li><li>• Sposoby tworzenia reprezentacji numerycznej tekstów, modele wektorowe semantyki, ontologie</li></ul> <p><b>2. Narzędzia i techniki ilościowego badania tekstów</b></p> <ul style="list-style-type: none"><li>• Podstawowe sposoby obliczania nastawienia w tekście, wymiary semantyczne i meta-semantyczne</li><li>• Uczenie maszynowe w automatycznej klasyfikacji opinii oraz kategoryzacji tekstów</li></ul> <p><b>3. Wizualizacja, analiza dyskursu</b></p> <ul style="list-style-type: none"><li>• Przybliżanie treści dużych zbiorów dokumentów technikami statystycznymi</li><li>• Techniki automatycznego streszczania tekstów</li><li>• Metody generowania i analizy sieci konceptów, interpretacja wyników</li></ul>
Wykaz literatury	<p><b>Literatura podstawowa:</b></p> <ol style="list-style-type: none"><li>1. "Opinion mining and sentiment analysis". B. Pang, L. Lee. Foundations and Trends in Information Retrieval 2(1-2), pp. 1–135, 2008.</li><li>2. "Inżynieria lingwistyczna. Komputerowe przetwarzanie tekstów w języku naturalnym". A. Mykowiecka. ISBN: 83-89244-54-3. Wydawnictwo Polsko-Japońskiej Wyższej Szkoły Technik Komputerowych.</li></ol> <p><b>Literatura uzupełniająca</b></p> <ol style="list-style-type: none"><li>1. "Thumbs up? Sentiment Classification using Machine Learning Techniques". B. Pang and L. Lee and S. Vaithyanathan. Proceedings of the 2002 EMNLP Conference.</li><li>2. "Is Sentiment a Property of Synsets? Evaluating Resources for Sentiment Classification using Machine Learning". A. Wawer. LREC 2010</li><li>3. "Studying Complex Discursive Systems: Centering Resonance Analysis of Communication". S. Corman, T. Kuhn R. D. McPhee, K. Dooley. Human Communication Research, v28 n2 p157-206 Apr 2002.</li><li>4. "The Measurement of Meaning". C. E. Osgood, G. J. Suci, P. H. Tannenbaum.</li></ol>



KAPITAŁ LUDZKI  
NARODOWA STRATEGIA SPÓJNOŚCI



UNIA EUROPEJSKA  
EUROPEJSKI  
FUNDUSZ SPOŁECZNY



	<p>1971.</p> <p>5. "The content analysis guidebook". K. A. Neuendorf. 2002</p> <p>6. "Taming Text with the SVD". Albright R.</p> <p><a href="ftp://ftp.sas.com/techsup/download/EMiner/TamingTextwiththeSVD.pdf">ftp://ftp.sas.com/techsup/download/EMiner/TamingTextwiththeSVD.pdf</a></p>
Kontakt do prowadzącego przedmiot	e-mail <a href="mailto:alex@zetema.pl">alex@zetema.pl</a>
Kontakt do Koordynatora merytorycznego kierunku	Mgr inż. Radosław Nielek, -mail: <a href="mailto:nielek@pjawstk.edu.pl">nielek@pjawstk.edu.pl</a>