

Codul și denumirea disciplinei	<i>S.03.A.013 Sisteme de căutare a informației</i>
Titularul disciplinei	<i>prof. univ. dr. hab Popescu Anatol</i>
Total ore	<i>60</i>
Prelegeri	<i>30</i>
Lucrări de laborator	<i>30</i>
Puncte de credit	<i>8</i>
Destinația	<i>Ciclul II, Masterat, Studenții anului II de la specialitatea Informatica Aplicată.</i>

Obiectivele de formare în cadrul disciplinei

La nivel de cunoaștere și înțelegere: Utilizarea metodelor de căutare și extragere a informației în domeniul dobândirii datelor (data mining) permite reducerea timpului de rezolvare a problemelor informaționale, mărește precizia dobândirii informației nestructurate și oferă soluții practice admisibile acolo unde soluții exacte sunt imposibil de abordat. Metode și tehnici de “text mining” trebuie să devină unele din instrumente curente în cercetări. Scopul disciplinei constă în dezvoltarea deprinderilor studenților de a utiliza tehnologii “text mining” pentru rezolvarea diverselor probleme cu caracter științific, informațional sau tehnic, precum și aprofundarea cunoștințelor în domeniul programării la calculator.

La nivel de aplicare și înțelegere: Studentul trebuie să fie capabil să aplice programe (pachete) standarte și să elaboreze programe noi (de exemplu în baza limbajului Perl și R) pentru rezolvarea problemelor din domeniul “text mining”.

Conținutul cursului: Text Patterns. A Simple Concordance. Extracting Sentences. Quantitative Text Summaries. Hashes for organisation of text data. Probability and text sampling. Conditional probability. Mean and variance of random variables. Applying information Retrieval to text mining. The Term Document Matrix Applied to N texts. Document Similarity. Concordance Lines and Corpus Linguistics. Sorting Concordance Lines. Colocation and Concordance lines. Multivariate Techniques with text. Text Clustering. Algorithm Clustering K-Means. Decision Trees and overfitting.

Cunoștințe anterioare necesare: Analiza probabilistico-statistică a textului. Procesarea limbajelor naturale. Information Retrieval.

Evaluarea finală: Examen scris sau oral, în nota finală luând în considerație atestările și activitățile curente în conformitate cu regulamentul în vigoare.

Bibliografie: Roger Bilisoly. Practical text mining with Perl. Willey. 2008. 295p. Cristopher D Meaning. Prabhacar Ragharan. Hinrich Schutze. An introduction to Information Retrieval. Cambridge University Press. 2009. 544p. Randal L. Schwartz & Tom Christianseu. Learning Perl. O’ REILLY. 1997. 267p.