

Classification:

<http://nlp.stanford.edu/IR-book/html/htmledition/contents-1.html>

Clustering:

http://home.deib.polimi.it/matteucc/Clustering/tutorial_html/hierarchical.html

LSI:

<http://nlp.stanford.edu/IR-book/html/htmledition/contents-1.html>

<http://www1.se.cuhk.edu.hk/~seem5680/lecture/LSI-Eg.pdf>

Link Analysis:

http://www.mohaisen.net/manar/classes/sna_lecture4_measures.pdf (slides 4 - 18)

<http://www.math.cornell.edu/~mec/Winter2009/RalucaRemus/Lecture3/lecture3.html>

Crawling:

http://www.cis.uni-muenchen.de/~yeong/Kurse/ss09/WebDataMining/kap8_rev.pdf

Document Summarization:

Graph based:

<https://www.jair.org/papers/paper1523.html>

Lexical Chain based:

<http://www.aclweb.org/anthology/W97-0703>

Wordnet based:

<http://www.lrec-conf.org/proceedings/lrec2004/pdf/342.pdf>

ROUGE:

<http://research.microsoft.com/en-us/um/people/cyl/download/papers/rouge-working-note-v1.3.1.pdf>