



## UNIVERSITI TEKNOLOGI MARA

### CSC770: NATURAL LANGUAGE PROCESSING

<b>Course Name (English)</b>	NATURAL LANGUAGE PROCESSING <b>APPROVED</b>
<b>Course Code</b>	CSC770
<b>MQF Credit</b>	3
<b>Course Description</b>	This course gives students the understanding of theory and practice of natural language processing (NLP) - the creation of computer programs that can understand, generate, and learn natural language. Natural language understanding as a vehicle to introduce the three major subfields of NLP: syntax (which concerns itself with determining the structure of a sentence), semantics (which concerns itself with determining the explicit meaning of a single sentence), and pragmatics (which concerns itself with deriving the implicit meaning of a sentence when it is used in a specific discourse context).
<b>Transferable Skills</b>	Demonstrateability to identify and articulate self skills, knowledge and understanding confidently and in a variety of contexts
<b>Teaching Methodologies</b>	Lectures, Blended Learning, Demonstrations, Presentation, Small Group Sessions , Directed Self-learning
<b>CLO</b>	<p>CLO1 Explain the potentials and limitations of the use of ordinary language in computing systems.</p> <p>CLO2 Write grammars for syntactic processing and integrate semantic and pragmatic aspects.</p> <p>CLO3 Analyze basic issues in this area and present the main techniques needed to obtain successful performance in application areas such as database query answering, text generation, semi-formal specifications and front ends and others</p>
<b>Pre-Requisite Courses</b>	No course recommendations
<b>Topics</b>	
<b>1. Introduction</b> 1.1) Knowledge in Speech Language Processing (SLP) 1.2) Ambiguity 1.3) Models and Algorithm 1.4) Language, Thoughts and Understanding History	
<b>2. Regular Expressions &amp; Automata</b> 2.1) Regular Expressions 2.2) Finite-state methods 2.3) Regular Languages and FSAs	
<b>3. N-Grams</b> 3.1) Counting 3.2) Simple N-Gram 3.3) Training and Test Sets 3.4) Evaluation 3.5) Smoothing 3.6) Interpolation 3.7) Backoff	
<b>4. Part-of-Speech Tagging</b> 4.1) Parts of speech (POS) 4.2) Tagsets 4.3) POS Tagging 4.4) Rule-based tagging 4.5) HMMs and Viterbi algorithm	

<b>5. Formal Grammars</b> 5.1) Context-free grammar 5.2) Grammars for English 5.3) Treebanks 5.4) Dependency grammars
<b>6. Syntactic Parsing</b> 6.1) Parsing with CFGs 6.2) Bottom-up, top-down 6.3) Ambiguity 6.4) CKY parsing
<b>7. Statistical Parsing</b> 7.1) Probabilistic Context Free grammars (PCFG) 7.2) Probabilistic CKY parsing
<b>8. Lexical Semantics</b> 8.1) Concepts about word meaning 8.2) Computational areas by enabling resource 8.3) Computational areas by enabling technology
<b>9. Computational Lexical Semantics</b> 9.1) Supervised Word-Sense Disambiguation (WSD) 9.2) Supervised Learning Algorithms 9.3) Evaluation of WSD 9.4) Similarity Metrics
<b>10. Question Answering &amp; Summarization</b> 10.1) Web-based Question Answering 10.2) Information Retrieval
<b>11. Machine Translation</b> 11.1) Language Similarities and Divergences 11.2) Classic MT Approaches 11.3) Modern Statistical MT 11.4) Evaluation

Assessment Breakdown	%
Continuous Assessment	70.00%
Final Assessment	30.00%

Details of Continuous Assessment	Assessment Type	Assessment Description	% of Total Mark	CLO
	Assignment	Assignment #1=5%, Assignment #2=10% Assignment #3=10%	25%	CLO1
	Quiz	Quiz and Participation in class	5%	CLO1 , CLO2 , CLO3
	Test	TEST 1= 20% TEST 2 =20%	40%	CLO1 , CLO2

Reading List	Recommended Text	<ul style="list-style-type: none"> <li>JURAFSKY D. ET. AL 2014, <i>SPEECH AND LANGUAGE PROCESSING AN INTRODUCTION TO NATURAL LANGUAGE PROCESSING</i>, 2ND EDITION Ed., PEARSON EDUCATION [ISBN: 978-933251841]</li> </ul>
	Reference Book Resources	<ul style="list-style-type: none"> <li>Richard M Reese 2015, <i>Natural Language Processing with Java</i>, Packt Publishing - ebooks Account ( [ISBN: 978-178439179]</li> <li>Grant S. Ingersoll, Thomas S. Morton, Andrew L. Farris 2013, <i>Taming Text: How to Find, Organize, and Manipulate It</i>, 1st Edition Ed., Manning Publications</li> <li>Alexander Clark, Chris Fox, Shalom Lappin 2012, <i>The Handbook of Computational Linguistics and Natural Language Processing</i>, 1 edition Ed., Wiley-Blackwell [ISBN: 978-11183471]</li> <li>Ela Kumar 2011, <i>Natural Language Processing</i>, I K International Publishing House [ISBN: 978-938057877]</li> <li>Steven Bird, Ewan Klein, Edward Loper 2009, <i>Natural Language Processing with Python</i>, O'Reilly Media [ISBN: 978-059651649]</li> <li>Jurafsky, D. &amp; Martin, J.H. 2008, <i>Speech and Language Processing</i>, 2 Ed., Prentice-Hall [ISBN: 978-01318732]</li> </ul>
Article/Paper List	This Course does not have any article/paper resources	
Other References	This Course does not have any other resources	