

|                            |  |
|----------------------------|--|
| <b>MODELLO D (inglese)</b> |  |
| <b>General Information</b> |  |
| Academic subject           | Natural Language Processing                  |
| Degree course              | Master of Science in Computer Science (LM18) |
| Curriculum                 | Artificial Intelligence                      |
| ECTS credits               | 6  |
| Compulsory attendance      | No   |
| Language                   | No, but attendance is highly recommended     |

|                        |                 |                         |        |
|------------------------|-----------------|-------------------------|--------|
| <b>Subject teacher</b> | Name Surname    | Mail address            | SSD    |
|                        | Marco de Gemmis | marco.degemmis@uniba.it | INF/01 |

|                             |                  |            |
|-----------------------------|------------------|------------|
| <b>ECTS credits details</b> | Type of activity | SSD        |
| Basic teaching activities   | 4 T1 credits     | ING-INF/05 |
|                             | 2 T2 credits     | ING-INF/05 |

|                       |  |
|-----------------------|--|
| <b>Class schedule</b> |  |
| Period                | II semester                                |
| Year                  | I year                                     |
| Type of class         | Lectures – Laboratory and guided exercises |

|                        |   |
|------------------------|---|
| <b>Time management</b> |   |
| Hours                  | 62 (teaching and lab.) + 88 (personal work) = 150 |
| Hours of lectures      | 32  |
| Tutorials and lab      | 30  |

|                          |                             |
|--------------------------|-----------------------------|
| <b>Academic calendar</b> |                             |
| Class begins             | 1 <sup>st</sup> March, 2021 |
| Class ends               | 4 <sup>th</sup> June, 2021  |

|  |   |
|--|---|
| <b>Syllabus</b>  |   |
| Prerequisites/requirements   | Formal Prerequisites: None<br>Cultural requirements: basic knowledge in Probability Theory, basic knowledge in Discrete Mathematics and Calculus, basic programming skills  |
| Expected learning outcomes (according to Dublin Descriptors) (it is recommended that they are congruent with the learning outcomes contained in A4a, A4b, A4c tables of the SUA-CdS) | <p><u>Knowledge and understanding</u><br/>The learner will acquire knowledge about:</p> <ul style="list-style-type: none"> <li>- fundamentals of natural language processing (NLP);</li> <li>- the theoretical, methodological and operational aspects of NLP with particular reference to the main levels of linguistic analysis;</li> </ul> <p><u>Applying knowledge and understanding</u><br/>The learner will be able to:</p> <ul style="list-style-type: none"> <li>- apply knowledge on NLP foundations and tools to address problems related to the management of text documents;</li> <li>- exploit NLP techniques to the design of complex systems managing unstructured data (e.g. text mining applications).</li> </ul> <p><u>Making informed judgements and choices</u><br/>The learner will be able to:</p> <ul style="list-style-type: none"> <li>- take decision about the design of AI systems managing unstructured data (e.g. dialoguing agents, AI system with NLP interface);</li> <li>- find, evaluate and adapt known solutions and tools to</li> </ul> |

|   |  |
|---|--|
|   | <p>design NLP pipelines for specific problems (e.g. sentiment analysis or opinion mining)</p> <p><i>Communicating knowledge and understanding</i><br/>The class gives the students the ability to communicate - both at a technical level and at a scientific dissemination level - the choices on the techniques adopted in an NLP pipeline designed to solve a certain problem.</p> <p><i>Capacities to continue learning</i><br/>The class gives the students the ability to follow the advances in the area of NLP, by learning both from books and scientific papers appearing in conference proceedings or journals.</p>                                     |
| Contents  | <p><u>Main Topics</u></p> <ul style="list-style-type: none"> <li>• Introduction to NLP</li> <li>• Introduction to Computational Linguistics</li> <li>• Lexical Level</li> <li>• Part-of-Speech (PoS) tagging</li> <li>• Syntactic level: Parsing</li> <li>• Introduction to Lexical Semantics</li> <li>• NLP tools and applications (e.g. text mining, text categorization)</li> </ul>   |
| <b>Course program</b>   |  |
| Bibliography  | <p>1) C. Manning and H. Schutze, Foundations of Statistical Natural Language Processing. MIT press, 2000.</p> <p>2) D. Jurafsky and J. Martin, Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. Prentice-Hall Series in Artificial Intelligence, 2000. ISBN: 0130950696.</p> <p>3) Steven Bird, Ewan Klein, and Edward Loper. Natural Language Processing with Python. <a href="http://www.nltk.org/book">http://www.nltk.org/book</a><br/>This book is made available under the terms of the Creative Commons Attribution Noncommercial No-Derivative-Works 3.0 US License.</p> |
| Notes   | <p>In addition to the recommended bibliography, scientific papers will be suggested by the lecturer.</p> <p>The teaching material (slides of lectures and papers) will be made available via e-learning platform of the Department of Computer Science: <a href="http://elearning.di.uniba.it">http://elearning.di.uniba.it</a></p>  |
| Teaching methods  | Lectures, exercises in the classroom   |
| Assessment methods (indicate at least the type written, oral, other)  | <p>Written: Students will be given an NLP problem to solve (e.g. design of an appropriate solution for a text mining task).</p> <p>In case of problems related to COVID-19 emergency, the assessment method will be oral.</p>  |
| Evaluation criteria (Explain for each expected learning outcome what a student has to know, or is able to do, and how many levels of achievement there are. | <p>The determination of the grade takes into account the following aspects:</p> <ol style="list-style-type: none"> <li>1) ability to select appropriate NLP techniques for the given problem;</li> <li>2) the completeness of the solution;</li> <li>3) ability to justify the adopted solution among several options.</li> </ol> <p>To pass the exam, students must be able to propose a solution that meet at least aspect 1). The grade above the minimum are attributed to students whose solution also satisfies the other aspects.</p>   |
| Further information   |  |