**PhD Physics course at Bari University (XXXIV Cycle)**

|  |  |
| --- | --- |
| **Title** | Big Data modelling and machine learning techniques: an introduction |
| **Proponent** | Nicola Amoroso |
| **# CFU**  **(1 CFU = 8 hours)** | 2 CFU |
| **Schedule** | 8 lessons of two hours. |
| **Brief Summary of the course** | The increasing availability of high dimensional and heterogeneous data samples (big data) makes urgent the development of a scientific background including data science and machine learning techniques, with applications in many fields.This course provides an introduction to the fundamental concepts in machine learning and state-of-the-art machine learning algorithms.  We will cover the most popular supervised learning algorithms including random forests, artificial neural networks, support vector machines and deep learning. We will also mention the basic algorithms for unsupervised learning, such as clustering and complex networks.  After explaining the necessary steps for automated learning, we will discuss: hypothesis space, overfitting, bias and variance, tradeoffs between representational power and learnability, evaluation strategies and cross-validation. The course will be accompanied by hands-on problem solving with programming in R and some tutorial sessions. |
| **Programme** | * Introduction: Basic definitions, bias, variance and cross-validation. * Random Forests. * Artificial Neural Networks. * Support Vector Machines. * Clustering and Networks. * Deep Learning. * Case study. * The use of computational facilities. |
| **Recommended texts** | Introduction to Machine Learning - EthemAlpaydin - MIT Press 2010  Deep Learning – Ian Goodfellow et al. – MIT Press 2016 |
| **Assessment methods** | T.B.A. |