Course Name	**Scalable Machine Learning**		**التعلم الآلي القابل للتضخم**	
Course Information	**Course Code**	**Course No.**	**Credit Hour**	**Prerequisite(s)**
	0911-1672	672	3 (3-0-6)	**Data Processing and Analytics**
Course Track	☐ Program Core		☒ Electives	

Course Description

Most real-world machine learning work involves very large data sets that go beyond the CPU, memory and storage limitations of a single computer. Crafting an enterprise-scale cost-efficient machine learning solution for Big data to uncover insights and value from an organization's data or to incorporate data-driven decisions into their process is a challenge. In this course, students will learn the machine learning techniques to explore, analyze, and leverage data, and they will be introduced to tools and algorithms that can be used to create machine learning models, and to scale those models up to big data problems. The key objective is how to use these technologies in manipulating, storing, analyzing and visualizing big data to create intelligent solutions in social networks, health, telecommunication, finance and many other fields. Major topics may include noise handling, feature extraction, selection, and learning algorithms in developing such systems; machine learning techniques in big data analytics; processing structured, unstructured and semi-structured data; open source tools like KNIME and Spark with examples that guide through the basic analysis of Big Data; data mining with NoSQL, cloud databases, and enterprise data analytical solutions among other cutting edge big data applications using emerging technologies. The student will gain hands-on experience through a guided project on how to build a data analytics solution using different tools and methods learned during the course.

Course Outcomes. After the completion of this course, the student will be able to:

1. **Describe** the basic principles of machine learning and big data. **[A]**
2. **Develop** a data strategy and process for how data will be generated, collected, and consumed. **[C]**
3. **Identify, apply and evaluate** appropriate technology and machine learning techniques to process structured and unstructured data to solve sample business case **[D, E]**
4. **Apply** appropriate machine learning techniques for Big data analytics in practical applications. **[E, F]**

Assessment Policy (TC)	**Assignments**	10%	**Quiz**	10%	**Capstone Project**	20 %
	Midterm	20%	**Final**	40%		
Textbook	1. Nataraj Dasgupta, "Practical Big Data Analytics: Hands-on techniques to implement enterprise analytics and machine learning using Hadoop, Spark, NoSQL and R", Packet Publishing, 2018, ISBN:9781783554393. 2. Md. Rezaul Karim, Md. Mahedi Kaysar, Large Scale Machine Learning with Spark, Packt Publishing, 2016, ISBN-13: 978-1785888748.					
References	EMC Education Services, "Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data", Willy, 2015, ISBN: 9781118876138.					