Investigation of online courses in machine learning

Adriaan Louw (53031377)

February 28, 2018

Contents

1 Abstract

An investigation into the state of online courses in machine learning

2 Method

I selected 5 online courses. Mostly preferring courses with well know universities. This list is by no means exhaustive. I then tabulated each topic that each course offered.

3 Results

3.1 "Intro to Machine Learning" from Udacity (Stanford University)

This course contains popular topics which include Decision Trees, Regression, Support Vector Machines and clustering. It is also only one of two courses that include Principal Component Analysis (PCA). Th course also covers the use of regression.

It does not however include anything on Neural networks, Perceptrons. The course does cover Naive Bayes but does not seem to cover other topics like Baysian networks. (UDACITY, 2018)

3.2 "Learning from Data" from Caltech

This course includes Support Vector Machines, Regression, Neural Networks and Kernel methods. Linear modeling is also included. The study of the VC dimension is also included. This is only found in one other course.

This course is the leanest of all the courses with regards to topics taught. There is no mention of Naive Bayes, Baysian Networks, Perceptrons, Clustering or Decision Trees. (California Institute of Technology, 2018)

3.3 "Machine Learning" from edX (Georgia Institute of Technology)

With this course we have popular topics which include Decision Trees. Support Vector Machines, Neural Networks, Regression, Kernel Methods, Bayesian Learning and Clustering. It is also the only course to include Markov Decision Processes. VC Dimensions are also included. Also, Instance Based Learning can only be found in this course.

The only popular feature this course is lacking are Perceptrons.(Class Central, 2018)

3.4 "Machine Learning" from edX (Columbia University)

This course has most of the popular topics including Decision Trees, Regression, Support Vector Machines Perceptrons, Bayesian Learning, Kernal Methods and Clustering. It is also one of only 2 courses that include PCA and it is the only course that teaches Hidden Markov Models. This course does lack Neural Networks. (edX, 2018)

3.5 "Introduction to Machine Learning" from NPTEL (Indian Institute of Technology Madras)

This course also includes Decision Trees, Support Vector Machines, Regression, Neural Networks, Perceptrons, Bayesian Learning and Clustering. The only popular topic missing is Kernel Methods. (NPTEL, 2014)

4 Conclusion

From the online courses under discussion we can see that not all topics are equally popular. The only 2 topics were present in each course. Regression and Support Vector machines. Regression usually included at least linear regression, but in some courses Logistic Regression was also covered. Two topics were also found in 4 of our courses. These are Decision Trees and Clustering. To round off the popular topics, we have the topic that were included in at least 3 courses. That would be Kernel Methods, Neural Networks and Bayesian Learning. Some unpopular topics include Hidden Markov Models and PCA. While topics like Q-Learning or Natural Language Processing did not feature at all. From this small dataset it seems like most courses agree on which topics are most important.

References

California Institute of Technology. (2018). *Learning from data.* Retrieved from http://work.caltech.edu/telecourse.html

Class Central. (2018). *edx:machine learning.* Retrieved from https://www.class-central.com/course/edx-machine-learning-8995

edX. (2018). *Machine learning.* Retrieved from https://www.edx.org/course/machine-learning-columbiax-csmm-102x-3

NPTEL. (2014). *Introduction to machine learning.* Retrieved from https://onlinecourses.nptel.ac.in/noc18$_c$s26/*preview*

UDACITY. (2018). *Intro to machine learning.* Retrieved from https://www.udacity.com/course/intro-to-machine-learning--ud120