IS643

Unit 6 Assignment

Alexander Satz

May 3, 2017

<https://www.youtube.com/watch?v=1WmqqfXNFZ4&list=PLaZufLfJumb8Nv9lOK2IVwaF21cM5rkoP&index=6>

Automated Machine Learning In The Wild

This keynote lecture focused on how noise can affect the outcome of machine learning algorithms. The speaker’s company attempts to determine the likelihood that a user will click on an ad, and thereby determine if the client should purchase available add space on the users web browser. The likelihood is determined by the cookie history of the user, and so one could envision a recommender system being employed. However, the speaker’s company appears to use logistic regression instead, in combination with stochastic gradient descent. The mathematical details were not covered.

Qualitative problems, including the use of ‘bots’ to search web pages were discussed, and how this ‘noise’ could lead to spurious results. Indeed, the bots are more predictable than the humans. Also, many web page clicks are accidental, and these accidental clicks are again more readily predicted by models than purposeful clicks.

A directed comparison between a recommender model and the logistic regression (with penalties) was not provided, which for myself would have been interesting. I presume that the regression analysis is faster.