

Hacking Models with Python

Although there are several data mining tools in python, you can use them to deal with almost every kind of data (numeric, text, image, audio, ...) you would meet. Besides, there are also lots of modeling tools in python, you can use them to build the FIRST LIGHTING MODEL to solve the problems.

However, if you want to solve problems deeply, most of time, you need to write down the CUSTOMIZED OBJECTIVE FUNCTIONS of your models and solve them by yourself. Instead of using fast modeling tools, you need to know more about the essential things in modeling:

- What is a model ?
- How models solve your problems ?
- What is the connection between models and data ?
- What is the important data ? important model ?

I deeply believe ... The more the connection between models and data you know, the deeper the problem you can solve.

Last year, I delivered a talk about “Hacking Modes with R” in COSCUP 2013 and MLDM Monday. I summarized the merits and demerits of lots of different kinds of models and try to teach myself and audience how to hack a model and how to evaluate the controller of hacking models. Unfortunately, it's a talk in Chinese and it's with R.

During this two years, I learned much deeper concepts and skills of hacking models in several cases and projects. This year, I am trying to summarize more materials about metric learning and some duality relations in the real applications with Python.

Outline:

- What is Modeling ?
- Data, Model, Evaluators
- Direct Problem: Data + Evaluators => Model
- Inverse Problem: Data + Models => Evaluator
- Hacking Models with Metric Learning
- Data as a Model & Model as a Data
- Duality between Dimension Reduction and Clustering

Reference about “Hacking Modes with R” :

https://bitbucket.org/c3h3/coscup2013_r

http://talks.c3h3.org/coscup2013_r/HackingModelsWithR.slides.html

https://www.youtube.com/watch?v=qwG3e_wjbZ4

<https://www.youtube.com/watch?v=yXPqNUjU-1Q>