

Naruaki TOMA <tnal@ie.u-ryukyu.ac.jp>  
宛先: "slab34-dm@ie.u-ryukyu.ac.jp" <slab34-dm@ie.u-ryukyu.ac.jp>  
返信先: slab34-dm@ie.u-ryukyu.ac.jp  
[slab34-dm:22] Level 4: (線形)回帰問題についての検討

2013年5月16日 17:38

=====  
Level 4: 回帰問題についての検討

べ切: 5/23, 次回実験にて報告してもらいます  
提出方法: Wiki

機械学習の代表的な使い道の一つである「(線形)回帰問題」では、  
「サンプル(入力N次元) + 実数値(出力1次元)」で用意された  
データセットにおいて、その関係性を関数により近似する(表現する)  
ことを目的としている。うまく近似できれば、未知のデータ  
についても適切な出力値を推定することができるだろう。  
回帰において入力と出力の関係性を線形関数(一次関数)で  
表現しようと試みる(=予測モデルを構築する)手法が線形回帰である。

具体的な(線形)回帰の手法は置いて(=何らかの方法で実現するとして)、  
「(線形)回帰問題」について以下の点について検討し、報告せよ。  
参照した文献/Webサイト等があれば参考文献として示すこと。

(1) (自分が推定したい)具体的な回帰問題について述べよ。  
予測したい出力(連続値)は何だろう?  
その出力をうまく予測する回帰を実現する(予測モデルを構築する)には、  
どのような入力(特徴ベクトル)を用意したらよいだろうか?

(2) 分類(classification)においては出力がクラスであり、  
テストデータに対する予想が正しいか否かという形で分類器の質を  
評価することができた。回帰(regression)においては出力は連続値  
であるが、どのように回帰モデルの質を評価したらよいだろうか?

(option) 次元の呪い [1] を鑑みるに、多数の特徴を最初から盛り込んだ  
高次元の特徴ベクトルとして入力を用意するだけでは必ずしも適切  
なモデルを学習できるとは限らない。「なるべく適切な回帰モデル」  
を構築(学習)するために、どのような手段が考えられるか検討せよ。

[1] 怪奇!! 次元の呪い - 識別問題、パターン認識、データマイニングの初心者のために、情報処理, Vol.34, No.5, No.6.  
前編: <http://goo.gl/c14Tm>  
後編: <http://goo.gl/7w2mn>

----  
Naruaki Toma  
E-mail: [tnal@ie.u-ryukyu.ac.jp](mailto:tnal@ie.u-ryukyu.ac.jp), Tel: 098-895-8830  
<http://www.eva.ie.u-ryukyu.ac.jp/~tnal/>

--  
You received this message because you are subscribed to the Google Groups "slab34-dm" group.  
To unsubscribe from this group and stop receiving emails from it, send an email to [slab34-dm+unsubscribe@ie.u-ryukyu.ac.jp](mailto:slab34-dm+unsubscribe@ie.u-ryukyu.ac.jp).  
To post to this group, send email to [slab34-dm@ie.u-ryukyu.ac.jp](mailto:slab34-dm@ie.u-ryukyu.ac.jp).  
Visit this group at <http://groups.google.com/a/ie.u-ryukyu.ac.jp/group/slab34-dm/?hl=en-US>.