Machine Learning Homework 3

1.

一張含有 文字, 字型, 行, 螢幕擷取畫面 的圖片

自動產生的描述

2.

一張含有 圖表, 行, 繪圖, 螢幕擷取畫面 的圖片

自動產生的描述

3.

一張含有 文字, 螢幕擷取畫面, 字型, 數字 的圖片

自動產生的描述

4.

一張含有 文字, 螢幕擷取畫面, 字型, 數字 的圖片

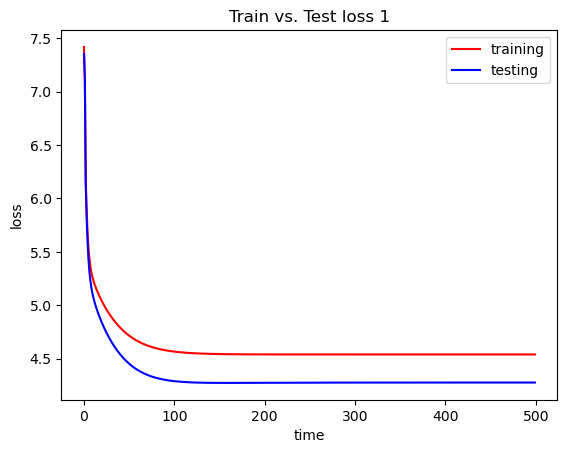
自動產生的描述 一張含有 文字, 字型, 螢幕擷取畫面, 數字 的圖片

自動產生的描述

Beta：

一張含有 文字, 字型, 白色, 印刷術 的圖片

自動產生的描述



一張含有 文字, 字型, 螢幕擷取畫面, 數字 的圖片

自動產生的描述 一張含有 文字, 字型, 螢幕擷取畫面, 數字 的圖片

自動產生的描述

Beta：

一張含有 文字, 字型, 白色, 印刷術 的圖片

自動產生的描述

一張含有 文字, 螢幕擷取畫面, 圖表, 繪圖 的圖片

自動產生的描述

一張含有 文字, 字型, 螢幕擷取畫面, 數字 的圖片

自動產生的描述 一張含有 文字, 字型, 螢幕擷取畫面, 數字 的圖片

自動產生的描述

Beta：

一張含有 文字, 字型, 白色, 印刷術 的圖片

自動產生的描述

一張含有 文字, 螢幕擷取畫面, 圖表, 繪圖 的圖片

自動產生的描述

一張含有 文字, 字型, 螢幕擷取畫面, 數字 的圖片

自動產生的描述 一張含有 文字, 字型, 螢幕擷取畫面, 數字 的圖片

自動產生的描述

Beta：

一張含有 文字, 字型, 白色, 印刷術 的圖片

自動產生的描述

一張含有 文字, 螢幕擷取畫面, 圖表, 陳列 的圖片

自動產生的描述

一張含有 文字, 字型, 螢幕擷取畫面, 數字 的圖片

自動產生的描述 一張含有 文字, 字型, 螢幕擷取畫面, 數字 的圖片

自動產生的描述

Beta：

一張含有 文字, 字型, 白色, 印刷術 的圖片

自動產生的描述

一張含有 文字, 螢幕擷取畫面, 圖表, 陳列 的圖片

自動產生的描述

一張含有 文字, 字型, 螢幕擷取畫面, 數字 的圖片

自動產生的描述 一張含有 文字, 字型, 螢幕擷取畫面, 數字 的圖片

自動產生的描述

Beta：

一張含有 文字, 字型, 白色, 印刷術 的圖片

自動產生的描述

一張含有 文字, 螢幕擷取畫面, 圖表, 繪圖 的圖片

自動產生的描述

一張含有 文字, 字型, 螢幕擷取畫面, 數字 的圖片

自動產生的描述 一張含有 文字, 字型, 螢幕擷取畫面, 數字 的圖片

自動產生的描述

Beta：

一張含有 文字, 字型, 白色, 印刷術 的圖片

自動產生的描述

一張含有 文字, 螢幕擷取畫面, 圖表, 繪圖 的圖片

自動產生的描述

一張含有 文字, 字型, 螢幕擷取畫面, 數字 的圖片

自動產生的描述 一張含有 文字, 字型, 螢幕擷取畫面, 數字 的圖片

自動產生的描述

Beta：

一張含有 文字, 字型, 白色, 印刷術 的圖片

自動產生的描述

一張含有 文字, 螢幕擷取畫面, 陳列, 圖表 的圖片

自動產生的描述

一張含有 文字, 字型, 螢幕擷取畫面, 數字 的圖片

自動產生的描述 一張含有 文字, 字型, 螢幕擷取畫面, 數字 的圖片

自動產生的描述

Beta：

一張含有 文字, 字型, 白色 的圖片

自動產生的描述一張含有 文字, 螢幕擷取畫面, 圖表, 行 的圖片

自動產生的描述

一張含有 文字, 字型, 螢幕擷取畫面, 數字 的圖片

自動產生的描述 一張含有 文字, 字型, 螢幕擷取畫面, 數字 的圖片

自動產生的描述

Beta：

一張含有 文字, 字型, 白色, 印刷術 的圖片

自動產生的描述

一張含有 文字, 螢幕擷取畫面, 圖表, 行 的圖片

自動產生的描述

5.

一張含有 文字, 螢幕擷取畫面, 字型, 數字 的圖片

自動產生的描述

Disscusion:

1.

根據以上的結果，可以判斷(β0, β1, β2, β3)和(β0’, β1’, β2’, β3’)算出來的結果相同。由於題目沒有規定random取值(weighting)的範圍，因此我取0-0.001較符合我的regression model。比如當我超出以上的random範圍時，train和test的loss不一定能夠找到最好的值。如下圖我們可以發現它在經過一段時間便趨近平緩，並未達到最好的結果。由此我理解到learning rate的設定和weighting factor很有關係，當weight-ing factor的初始值距離minium很遠的話，learing rate的初始值也要跟著改變。

一張含有 文字, 螢幕擷取畫面, 行, 繪圖 的圖片

自動產生的描述

而換句話說，當初使條件設定不良的話，gradient descend不一定能找到一個最小的值。而Least Square Method的目標是最小化預測值和實際值之間差的平方和，可以直接通過解矩陣方程式來找到係數，也就代表可以直接通過計算得到最優解，不需要迭代過程。以上便是兩者的差異。

2.

以下三張圖為我所找到的每個feature的best fitting model。首先我在程式中跑了九張圖，也就是每個feature對每個model的圖，然後透過觀察去判斷。X2、X3分別選擇含有2次項的model，(X2為model 2，X3為model 3)，因為我從圖上觀察到它們的分布比較像2次曲線。而X4選擇model 1的原因是根據model的公式，由於model 2和model 3都含有2次項，因此X4在那個model的權重也會因此改變，也就會讓它相比model 1來的差

從以下的圖觀察，首先我發現房價和屋齡的關係呈現U型，如預想的一樣，屋齡越小的房子價格越高，但比較特別的是在屋齡30~40的區段，房價有回升的跡象，我猜測是因為能維持如此久的房子本身用料就不差，所以這個區段的房子本來就屬於高價房。而房價與距離捷運站遠近也有一定的關聯性，可以看到距離捷運越近的房價很高，且有許多房屋集中於接近捷運站的地方，代表交通是房價的一個重要考量之一。而便利商店的數量和房價的關係相較之下比較不是最主要的原因，但也可以觀察到便利商店的數量和房價也是呈正比關係，也就代表便利商店的數量越多也會影響房價，便利商店的數量多寡可能代表都市化的程度，畢竟都市的便利商店數量一定大於鄉村，而都市的房價也大於鄉村。

|  |  |
| --- | --- |
| 一張含有 文字, 螢幕擷取畫面, 圖表, 行 的圖片  自動產生的描述 | 一張含有 文字, 螢幕擷取畫面, 圖表, 行 的圖片  自動產生的描述 |
|  |  |
| 一張含有 文字, 螢幕擷取畫面, 圖表, 行 的圖片  自動產生的描述 | |
|  | |