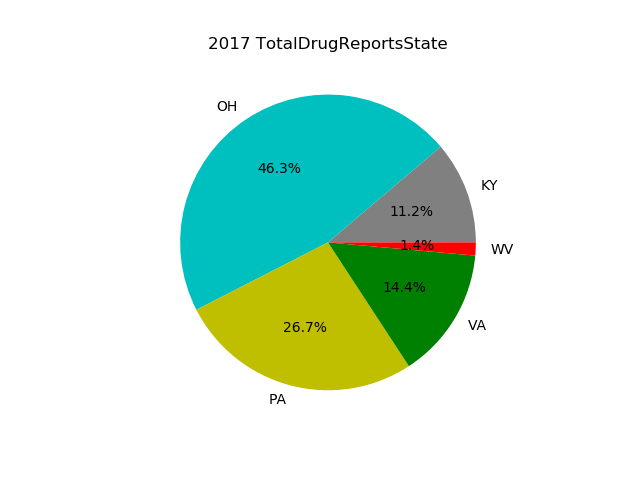
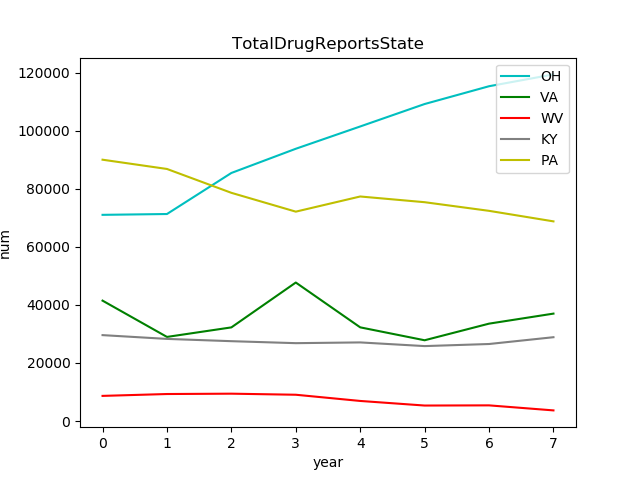


图[1] 2010年各州TotalDrugReportsState占比数饼图



图[2] 2017年各州TotalDrugReportsState占比数饼图

根据图[1]和图[2]，我们发现这7年，OH州药物事件占比数猛增，各州药物事件占比数变化很大，而如果各州影响因子均相同且独立（即五个州同时符合某一线性或非线性的变化），这一药物事件占比数的饼图中，占比数从大到小的排名位次不应该发生改变。



图[3] 各州TotalDrugReportsState量随时间变化图

图[3]中，我们能够更明显看出这一变化，造成变化的因子使这几个州的毒品事件数发生改变。而影响各州的这一变化因子可能包括人口原因，经济原因，地理原因，社会组成原因，政策原因等。由于我们使用的数据集MCM\_NFLIS\_Data中并未包含这些原因，因此我们把这些原因带来的影响使用超参数h来表达。

根据这个数据集，我们能够用来建模的数据有：

TotalDrugReportsCounty TotalDrugReportsState

我们把它们命名为x1,x2。

超参数h作用于x2上，使得:next(x2) = h(x2)。

我们假设：

对于每个COUNTY当年增加的TotalDrugReportsCounty（我们把它设为y），它前一年的DrugReports，所在STATE前一年的DrugReports都对其有作用，这一作用因子我们设为w。

我们可以得出这样一个式子:

y = w\*x + b(b为偏置项bias)

这里超参数h产生的效果我们可以近似于w2,b2。

根据我们上面的三张图可以知道，由于h对于每个州产生的效果不同，我们把x2删去，增加新的x2,x3,x4,x5,x6,作为每个州的 TotalDrugReportsState放入我们的式子中，此时h的作用被等效于为w2,w3,w4,w5,w6,b2,b3,b4,b5,b6。

为了引用非线性关系，我们引入sigmoid作为激活函数，并在后面再增加一个激励因子s，s的作用是控制我们的映射关系落在sigmoid函数的线性区还是非线性区。

x1:前一年的TotalDrugReportsCounty

x2~x6:当前state前一年的TotalDrugReportsState，若不是当前state补0

y:该county当年增加的DrugReports

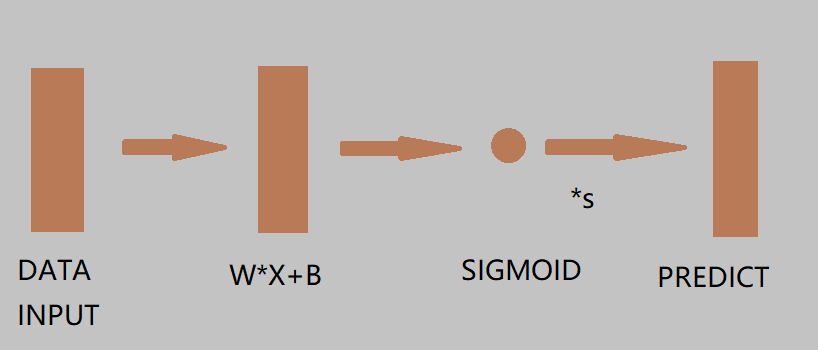
w,b:权重和偏置项

s:激励因子

我们使用了这样一个式子来计算损失:

loss = (predict\_y – y)^2

通过随机梯度下降[1]来最小化损失。



图[4]网络流程图

我们使用了一个浅层的神经网络去训练这个式子，图[4]为网络的流程图，使用了Adam优化器作为优化器。设置学习率为0.01,训练了20000次后得到了很有趣的结果。

"w": [[-0.859485387802124], [-1.393712043762207], [0.4642651379108429], [-1.547942042350769], [-0.49598830938339233], [-1.3133535385131836]],

"b": [1.0060161352157593],

"s": [[10.99660587310791]]

在激励因子s为正数的情况下，我们可以从权值中得到以下分析:

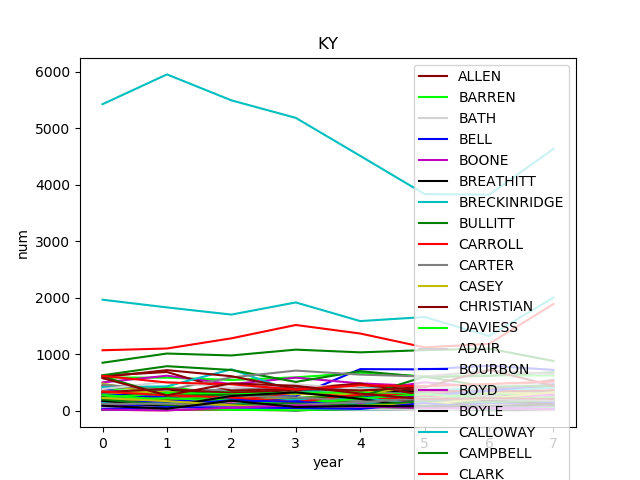
w1 = [-0.859485387802124]，代表当前一年毒品事件数增高时，对下一年毒品事件数的增长的抑制也会增高。

而w2~w6 = [-1.393712043762207], [0.4642651379108429], [-1.547942042350769], [-0.49598830938339233], [-1.3133535385131836]分别代表['VA', 'OH', 'PA', 'KY', 'WV']州对其county的激励因子，除了w3为正激励，其他均为负激励，小幅波动的增长依靠偏置项b来维持。OH州的h因子有利于毒品事件的增长。

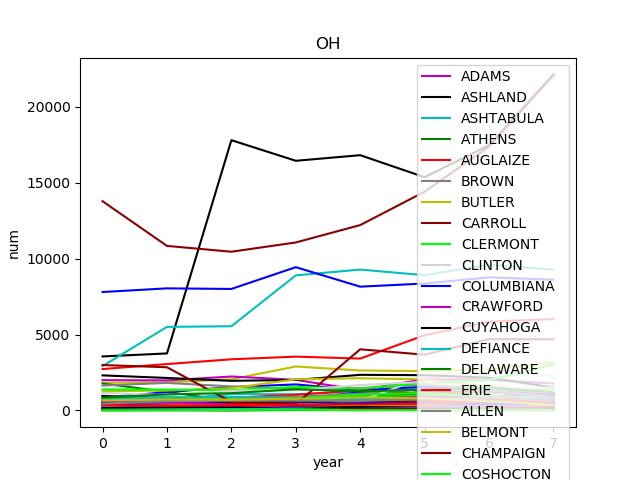
这样来看，我们有理由相信OH州已经开始使用特定阿片类药物。

我们可以看出，该州的毒品事件增长速度会随着增长时间而减少，但这个时间明显不是短期(从图[3]可以较为直观的看出)，在短时间内OH州的增长不会停止。而随着OH州总毒品事件数量的进一步增加，w3对其的激励将会进一步加大，这直接使OH州的毒品事件数量猛增。这一正反馈的结果可能产生严重的毒品危机，这是美国政府所需要担忧的。

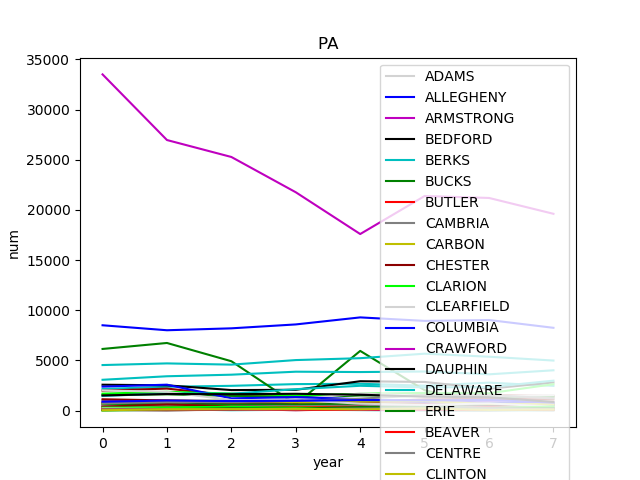
从我们的模型可以看出，其他州的毒品事件还处于一个动态平衡之中（即抑制因子和偏置项互为对抗），而OH州的毒品事件数在一定条件下来看，已经失控，阻止毒品继续泛滥已经刻不容缓。



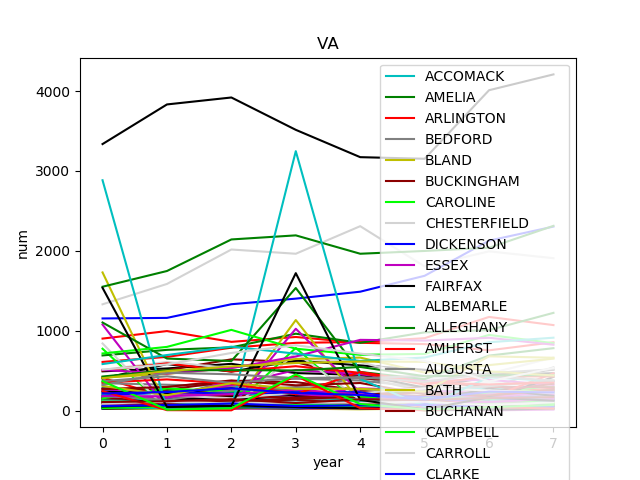
图[5] KY州各个county的毒品事件数随时间变化折线图



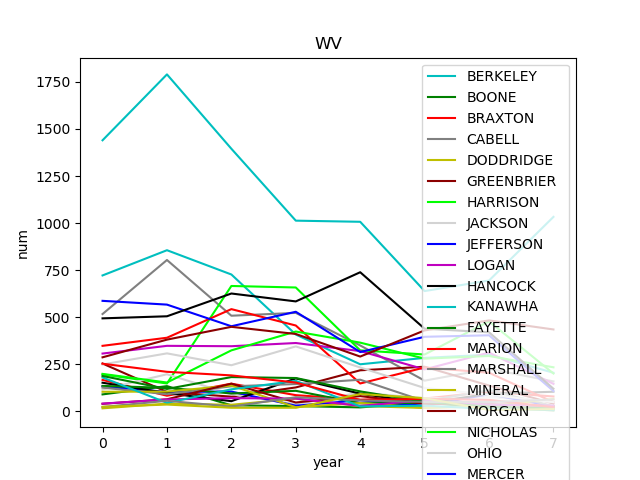
图[6] OH州各个county的毒品事件数随时间变化折线图



图[7] PA州各个county的毒品事件数随时间变化折线图



图[8] VA州各个county的毒品事件数随时间变化折线图



图[9] WV州各个county的毒品事件数随时间变化折线图

现在我们把各个county独立的看待，由图[5]到图[9]可以看出，各个州毒品事件数量和该州主要城市毒品事件数量的关联极大，而从图[6]可以看出，毒品泛滥最为严重的OH州中，其毒品事件增长量主要由CUYAHOGA和HAMILTON提供。结合前面的模型，我们可以更为精准的说：OH州的CUYAHOGA和HAMILTON很可能已经开始使用特定阿片类药物。

Reference

[1]Paras. Stochastic Gradient Descent[J]. Optimization, 2014.