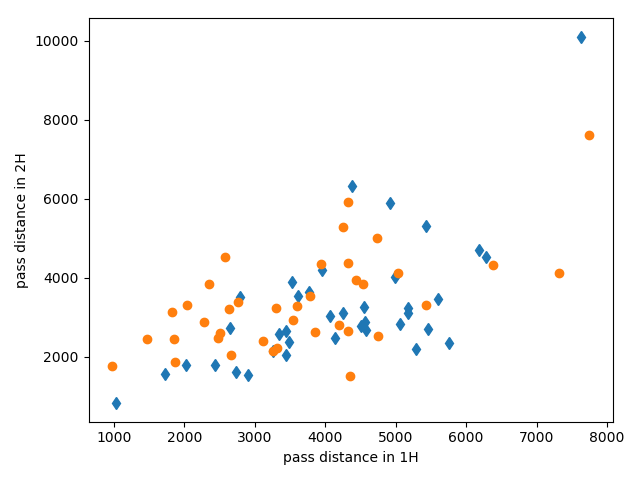
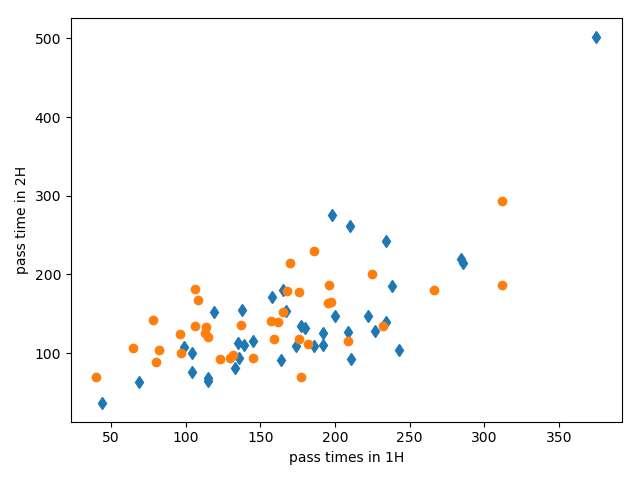
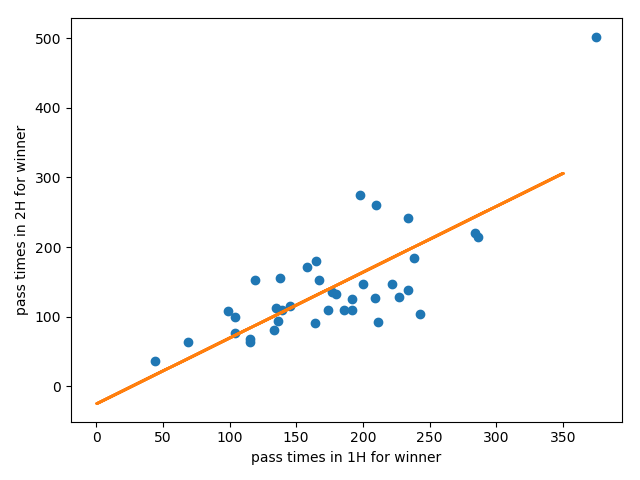
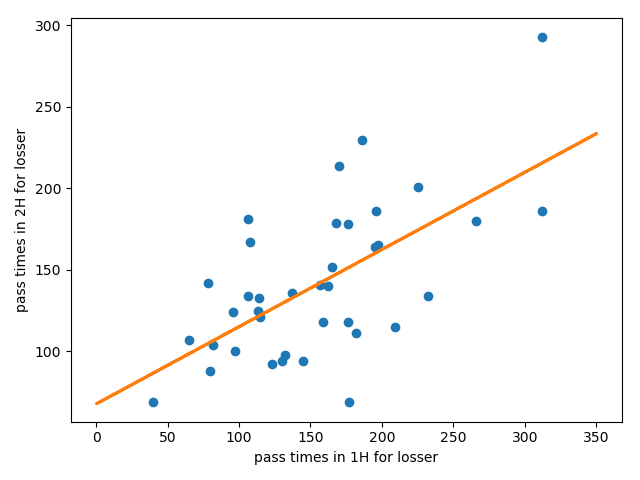
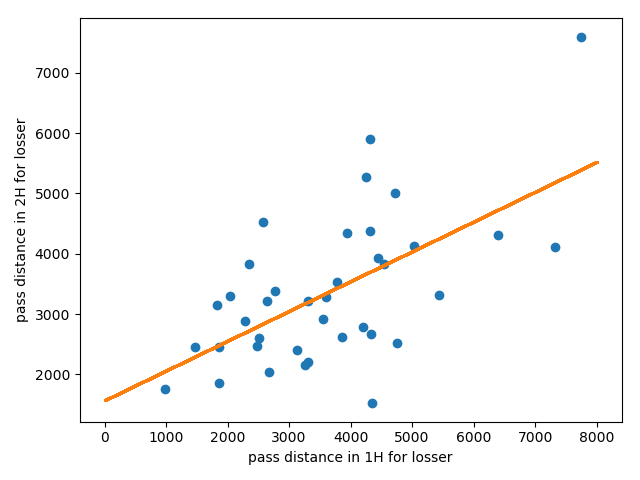
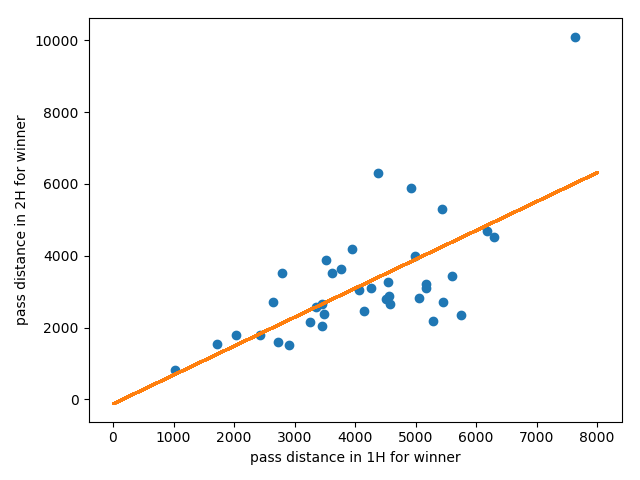
****

**赢家与败家传球次数的上半场与下半场的线性拟合图**

** **

**赢家的斜率0.946 败者斜率0.4739明显小于赢家**

**赢家与败家传球距离的上半场与下半场的线性拟合图**

****

**赢家斜率0.80 败者斜率0.496**

**由数据可以知道，赢家在比赛上下场次中发挥差不多，而败者往往在下半场中发挥不如上半场，由此**

**猜想 比赛结果和心态有关**

**所以，接下来计算每场比赛Huskies队伍的心态**

**心态等于 = 失误率/压力**

**失误率 = 失误事件次数/总事件次数**

**压力计算： 我方发生消极类事件，或 对方发生积极类事件我方总压力加三十（一共30名队员），我方发生积极事件，或 对方发生消极事件 我方压力减三十，如若压力与失误率正相关很大，则心态不好，相反若相关系数很小，则心态较好**

**/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\***

**定义消极类，积极类**

**一，积极类：**

**EventType：Shot**

**EventSubType：Goal kick，Shot，Free kick shot，Penalty，Smart pass**

**二，消极类：**

**EventType：Foul，Offside**

**EventSubType：Touch，Foul，Corner，Late card foul，Protest，Hand foul， Violent Foul，  Out of game foul，Time lost foul，Simulation**

**三，竞争类：**

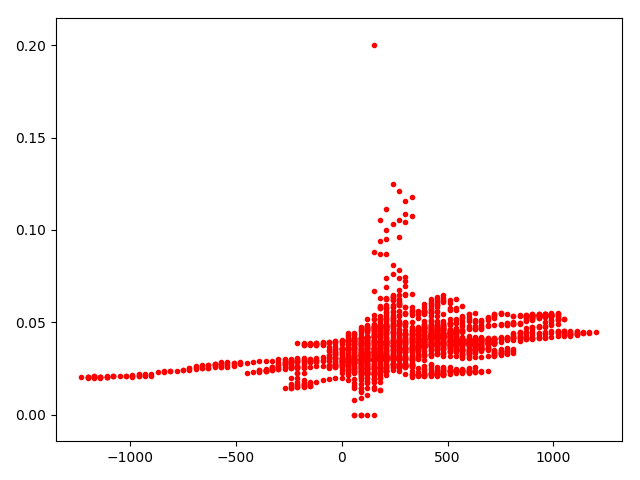
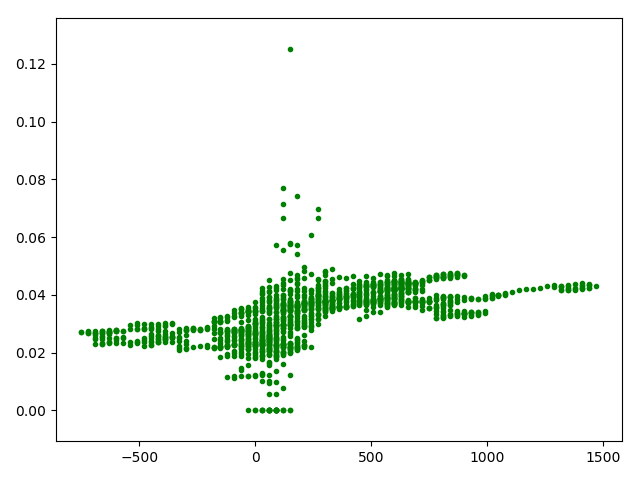
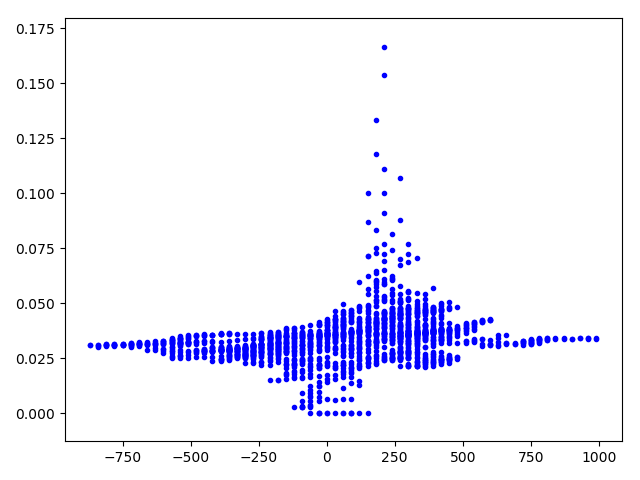
**EventType：Free Kick，Duel**

**EventSubType：Air duel，Head pass，Ground loose ball duel，Simple pass，Launch，High pass，Ground defending duel，Free kick cross， Whistle**

**四，中立事件：**

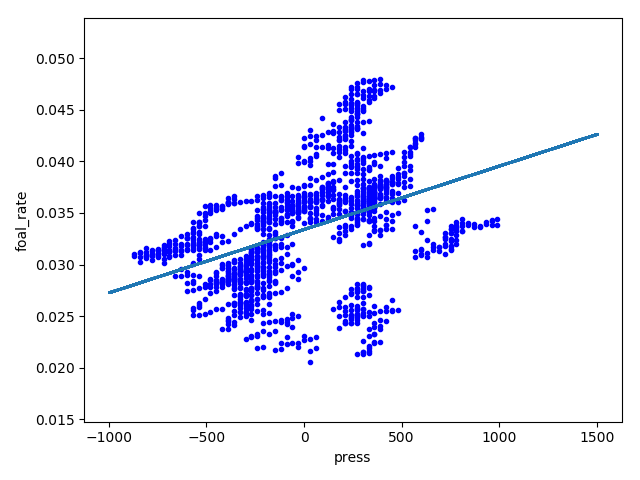
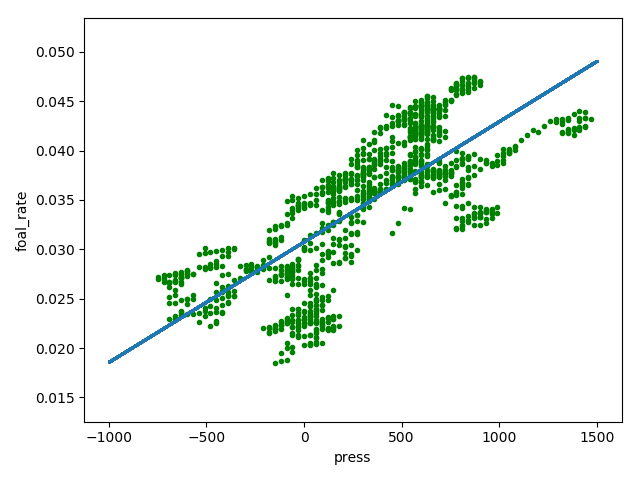
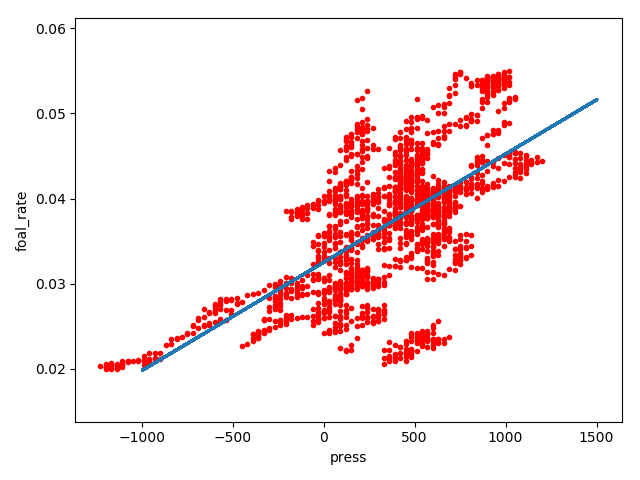
**EventType：Free Kick，Others on the ball，Goalkeeper leaving line，Save  attempt，Substitution，Interruption，Pass**

**\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/**



**胜场的压力-失误率图 平场的压力-失误率图 败场的压力-失误率图**

**我们现在分析一下数据的产生原理，初始设置的初始压力为150，假如前几个事件我们就发生失误了，那么此时的失误率就比较高甚至为1（第一个事件就是消极事件），在比赛靠前的时间段里面出现的数据表现的不太可靠，所以我们试着去掉每场比赛刚开始的50个数据。**

**胜场的压力-失误率图 平场的压力-失误率图 败场的压力-失误率图**

**斜率6.15146464e-06 斜率1.21966911e-05 斜率1.27035477e-05**

**这样处理后数据得到了可观的改善，可以看出失误率与压力是成正相关的，由于每增加一次消极事件，压力增加30，换算过来将斜率乘以三十倍就可以得到每发生一次消极事件，失误率的增量。如此，由压力失误率的斜率可以发现胜场时候胜场次的斜率相比于败场或者平场的时候要小一个数量级，也就是说，失误率随压力的波动增量不大，换句话说就是心态更为平和，胜不骄败不累**

**# ------------------------------------------------------------------------------------------**

**微分方程模型：**

**假设如下参数均是连续且光滑的**

**我们忽略掉普通事件，假设只有积极的事件和消极事件，这样做的后果就是大大的放大了失误率（因为分母大大减小了）**

**/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\***

**数据说明：**

**这里记总事件T = FT + PT**

**FR foal rate失误率**

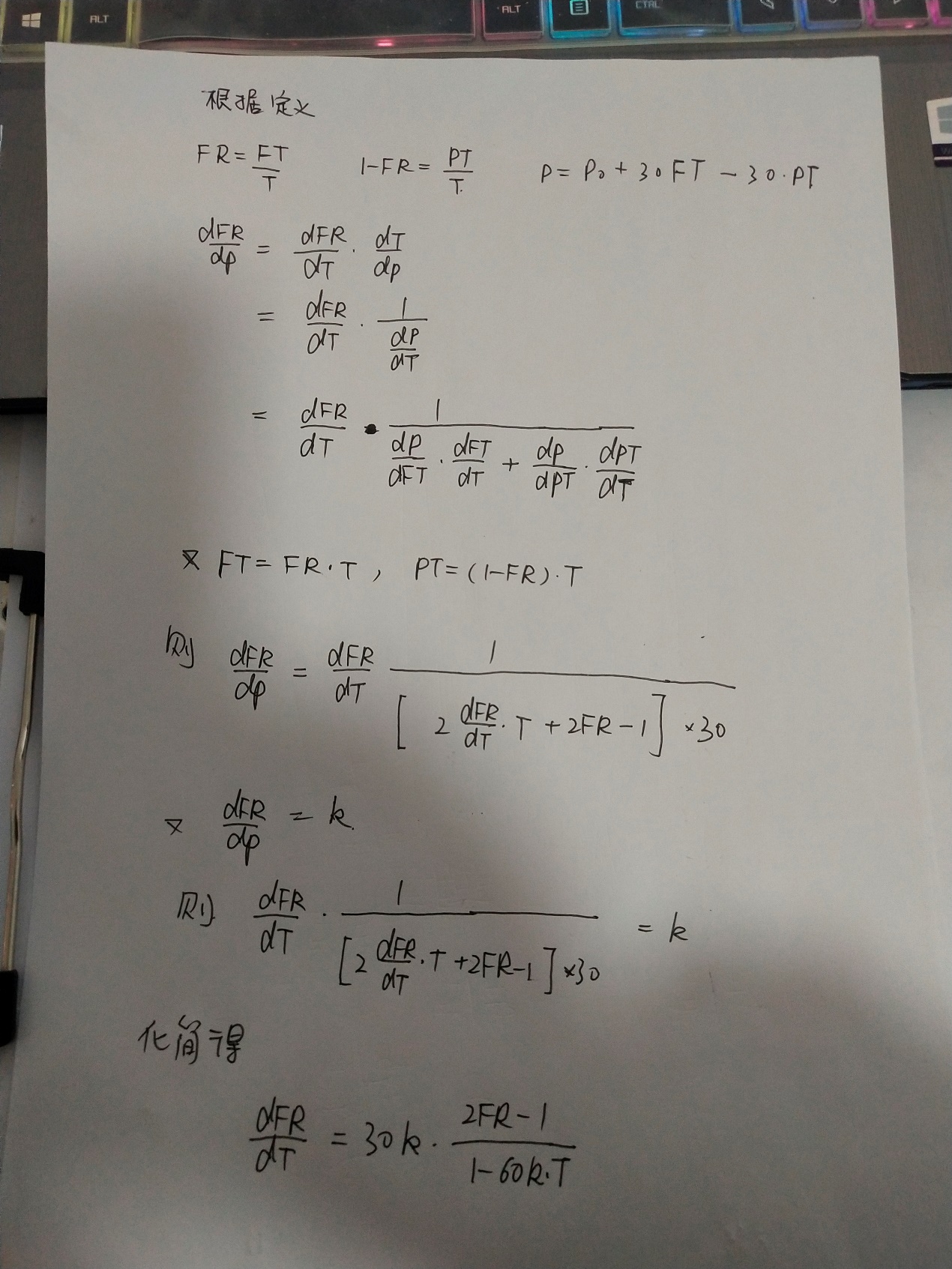
**P press压力**

**FT foal time消极事件次数**

**T vent Time 事件总数**

**PT positive events time 积极事件总数 k 失误率—压力的斜率即心态**

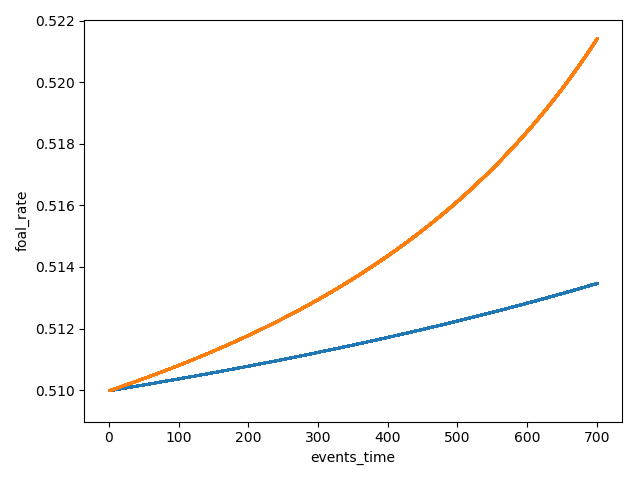
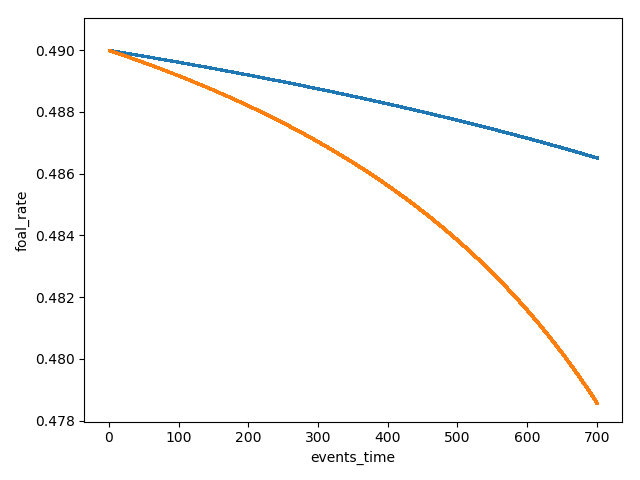
**则 根据定义**



**通过分析知： k的量级很小，在10^-5到10^-6，因此总能保证在比赛时间内，（1-k\*T）是近似为1的**

**DFR/dT = 30k\*(2FR – 1)，当失误率超过0.5的时候，失误率随场次的增加而增加，这个时候心态较好队伍收到的影响较小，比较能在优势下保持胜局，如果心态不好，有可能会被对手翻盘。**

**用四阶龙格库塔法画出的曲线（如下）**

** **

**在失误率超过0.5情况下的图像 在失误率低于0.5的情况下的图像**

由此可以进一步说明在心态更好的时候，失误率的波动不大，队伍可以更好的发挥出真实水平。

**进一步分析内在原因**

**用神经网络来制定评分权重，标准，传球，传球距离等等对输赢的影响**

**1.1模型二**

**1.1.1 初步确定直观影响输赢的因素**

**定义labels=[‘上半场传球总次数’，‘下半场传球总次数’， ‘上半场传球总距离’，‘下半场传球总距离’， ‘不同传球类型的次数’]**

**定义教师信号 赢球率：(我方进球数目-对方进球数目)/( 我方进球数目+对方进球数目)**

**1.1.2 建立神经网络模型进行训练**

**/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\***

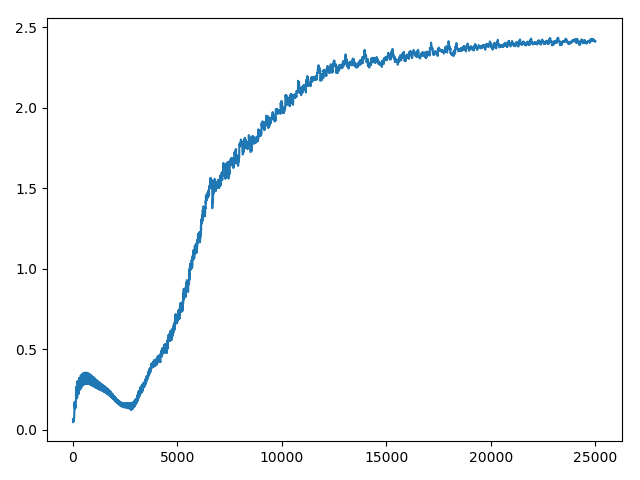
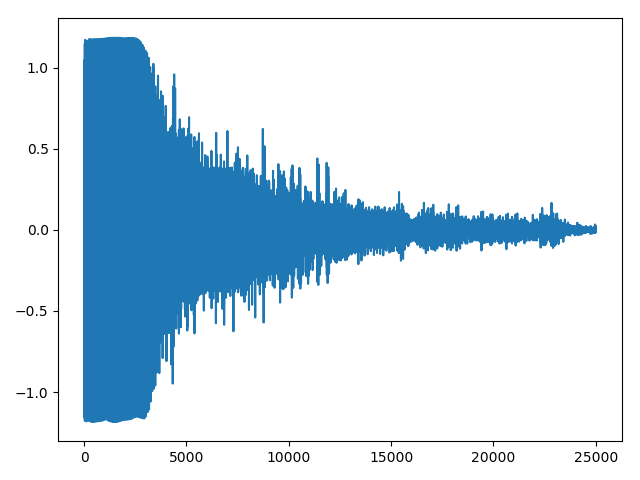
**之前遇到问题，训练无法收敛，后来将数据归一化后收敛，归一方法为**

**Data = (x-Data\_mean)/derta Derta是均方差**

**这样做可以将某个特征的数据变为均值为0，方差为1的训练集，从而使收敛速度加快**

**\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/**

**我们监测神经网络中其中的一个权值，以及误差变化如下：**

** **

**输出层第一个神经元到隐层第 误差在学习25000次后敛**

**一个神经元的权值变化 逐渐向0收**

**由此神经网络学习25000次后得到的最优数据**

**[129.0776809068127, 77.9958890234888, 5314.844617894659, 1868.8745314494613, 13.726945784017213,186.53819822598442,0.9711863561572205,10.432631199340484, 1.14508868180727, 2.4371190073318174, 1.7708923457618795]**

**即对应：**

**[129,78,5315,1869,13,187,1,10,1,2,2]**

**上半场传球次数129，下半场传球次数78，上半场传球距离5314.8，下半场传球距离1868.8， head pass 13，simple pass186.5，launch pass 1，high pass10，hand pass1，smart pass2 cross pass2次可以使赢球率为0.99516474151444023**

**收集huskie数据,与模型结果进行对比**

Huskies数据求完均值后四舍五入后

[150, 125, 3609, 2966, 15, 230, 6, 15, 3, 2 ,3]

对比后

# --------------------------------------------------------

此处应该与神经网络训练结果对比列表放在附录二

**上半场传球次数，下半场传球次数，上半场传球距离，下半场传球距离， head pass，simple pass，launch pass ，high pass，hand pass，smart pass cross pass**

**Huskies\_data = [[222, 147, 5597, 3447, 13, 329, 1, 15, 4, 2, 5],**

**[104, 76, 2430, 1782, 10, 150, 4, 14, 1, 0, 1],**

**[209, 115, 4753, 2522, 4, 294, 6, 13, 7, 0, 0],**

**[176, 178, 4728, 5001，17, 296, 6, 16, 6, 3, 10],**

**[196, 186, 4317, 4380, 16, 342, 8, 8, 4, 2, 2],**

**[192, 126, 4550, 3263, 14, 267, 1, 22, 1, 7, 6],**

**[170, 214, 4246, 5273, 28, 326, 5, 10, 3, 0, 12],**

**[186, 109, 4563, 2887, 11, 261, 1, 14, 6, 1, 1],**

**[80, 88, 1865, 1861, 11, 139, 2, 8, 3, 1, 4],**

**[195, 164, 4538, 3828, 13, 313, 5, 18, 1, 2, 7],**

**[115, 68, 2731, 1608, 11, 145, 9, 12, 1, 3, 2],**

**[69, 63, 1724, 1554, 6, 114, 2, 6, 0, 1, 3],**

**[114, 133, 2767, 3377, 18, 199, 11, 15, 3, 0, 1],**

**[200, 147, 5175，3227, 11, 297, 8, 24, 1, 1, 5],**

**[119, 152, 2786, 3521, 20, 215, 12, 12, 4, 6, 2],**

**[44, 36, 1030, 816, 6, 62, 4, 5, 2, 1, 0],**

**[211, 93, 5293, 2187, 18, 267, 6, 9, 2, 0, 2],**

**[177, 135, 4069, 3035, 18, 275, 3, 14, 1, 1, 0],**

**[135, 113, 3349, 2584, 18, 192, 17, 12, 4, 0, 5],**

**[136, 94, 3261, 2155, 13, 179, 2, 24, 10, 1, 1],**

**[232, 134, 5435, 3309, 19, 300, 17, 19, 5, 3, 3],**

**[145, 94, 3257, 2155, 11, 205, 2, 10, 5, 1, 5],**

**[115, 121, 2509, 2600, 11, 185, 2, 28, 4, 5, 1],**

**[167, 153, 3614, 3532, 23, 272, 4, 14, 4, 1, 2],**

**[139, 110, 3485, 2374, 27, 179, 7, 22, 8, 2, 4],**

**[106, 181, 2574, 4520, 9, 241, 5, 21, 3, 3, 5],**

**[174, 109, 4513, 2784, 11, 241, 2, 22, 1, 2, 4],**

**[113, 125, 2639, 3216, 22, 180, 5, 22, 2, 1, 6],**

**[96, 124, 2285, 2887, 24, 168, 11, 9, 0, 1, 7],**

**[180, 132, 4254, 3103, 20, 265, 6, 14, 0, 3, 4],**

**[158, 171, 3526, 3878, 17, 282, 4, 10, 5, 5, 6],**

**[40, 69, 979, 1754, 4, 87, 2, 13, 1, 2, 0],**

**[99, 108, 2651, 2714, 11, 177, 4, 12, 0, 0, 3],**

**[209, 127, 5175, 3098, 15, 274, 11, 26, 4, 2, 4],**

**[243, 104, 5756, 2351, 14, 301, 1, 19, 9, 1, 2],**

**[192, 110, 4583, 2669, 25, 232, 10, 24, 6, 4, 1],**

**[138, 155, 3767, 3624, 14, 255, 3, 14, 4, 0, 3],**

**[108, 167, 2355, 3837, 19, 229, 3, 16, 2, 5, 1]]**

**训练结果 [129, 78, 5315, 1869, 13, 187, 1, 10, 1, 2, 2]**

**------------------------------------------------------------------#**

**随后我们将训练结果与Huskies的每场数据求欧氏距离(欧氏距离越小，则和神经系统预测最优值相近)后归一化（每个元素除以最大值），然后以 0.4，0.6为阈值做截断函数，**

**[1, -1, 1, -1, -1, 1, -1, 1, -1, 0, 0, -1, -1, 1, -1, -1, 1, 1, 0, 0, 1, 0, -1, 0, 0, -1, 1, -1, -1, 1, -1, -1, -1, 1, 1, 1, 0, -1]**

**与原来的真实情况做对比（列表）**

**[1, 0, -1, -1, -1, 1, -1, 0, -1, -1, 1, 0, -1, 1, 1, 0, 1, 1, 0, 0, -1, -1, -1, 0, 1, -1, 1, -1, -1, 1, 1, -1, 0, 0, 1, 1, 0, -1]**

**除去两个列表中tie的情况即元素为0（我们认为平场情况往往含有运气的成分，难以预测）**

**将上下两个列表的对应元素相减**

**[0, -1, 2, 0, 0, 0, 0, 1, 0, 1, -1, -1, 0, 0, -2, -1, 0, 0, 0, 0, 2, 1, 0, 0, -1, 0, 0, 0, 0, 0, -2, 0, -1, 1, 0, 0, 0, 0]**

**为0的即为预测准确，为正负二的预测我们成为不成功的预测，因为他们不是因为tie而难以区分，这样一来我们成功预测了24场比赛，预测失败了4场，其他10场是因为平局，难以预测运气成分，进一步证明了我们给出的训练结果具有一定的普适性。达到这样的结果会帮助Huskies在绝大多数的比赛中赢的比赛或者与对手打个不相上下**

**其他因素分析**

通过分析matches.csv从大局上（不计场内细节）分析影响输赢的因素

Side

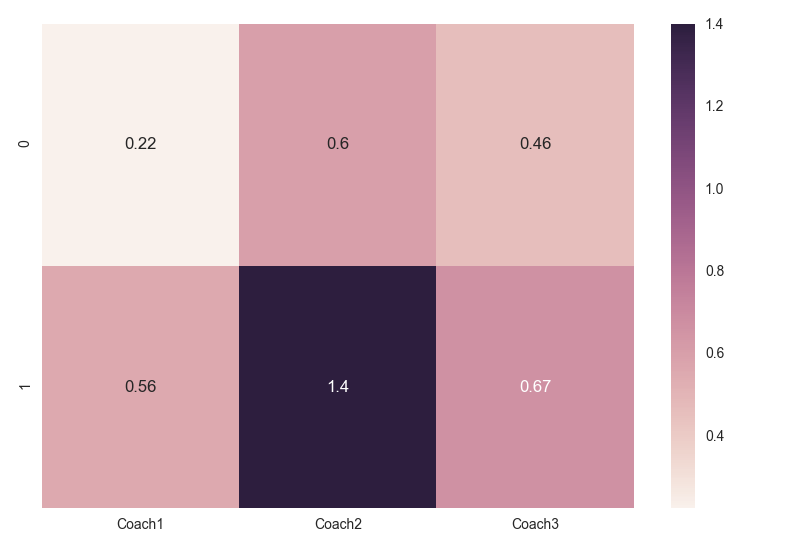
own factor own score

Coach

Win/loss/tie

opponent factor opponent score

**coach**



0 是away，1 是home

通过对matche.csv表格中的数据分析，我们可以得出如下结论，教练三人在主场的平均场次得分均大于客场，说明我方习惯于打顺风赛，在主场表现好（猜想），并且在球队主场和客场两种情况下，三人的平均场次得分由高到低排列，分别为教练二，教练三，教练一。在教练二的带领下均分明显高于教练一二

**优劣分析**

**模型一的优点，我们使用的神经网络模型训练的结果精度高，在比赛的时候能够较为准确的预测输赢，以及给出具有一定普适性的改善方向。劣势 神经网络的隐层数量的选取比较难以确定，隐层数量过少，预测不准确，隐层数量增加可能会过拟合，使预测失去意义。我们通过不断的手动调整隐层数，观察误差变化，以及其收敛速度，最终确定为14层左右，具有较好的预测意义。**

**模型2的优势我们创新性的量化了心态对比赛的影响，分析了影响比赛的内在因素。**

**模型2的的劣势，心态的微分方程模型是在假设失误率是关于事件连续的假设，而实际是离散的，微分方程只具有指导意义，不具有精确计算的能力。而且我们定义的心态在数量级上很小，说明心态只能在很小程度上影响比赛。影响力甚至没有随机事件的影响大，调整心态只能小范围降低失误率，更多的是依赖于队伍的实力**