**实训总结：**

**2018.07.02**

**了解赛制；**

**常规赛每支球队要打满82场比赛：2\*15+4\*4+6\*4+4\*3=82；**

**理解老师给的例子。**

**ELO等级分制度**（英语：Elo rating system）是指由[匈牙利](https://zh.wikipedia.org/wiki/%E5%8C%88%E7%89%99%E5%88%A9)裔[美国](https://zh.wikipedia.org/wiki/%E7%BE%8E%E5%9B%BD)[物理学家](https://zh.wikipedia.org/wiki/%E7%89%A9%E7%90%86%E5%AD%A6%E5%AE%B6)Arpad Elo创建的一个衡量各类对弈活动水平的评价方法，是当今对弈水平评估的公认的权威方法。被广泛用于[国际象棋](https://zh.wikipedia.org/wiki/%E5%9B%BD%E9%99%85%E8%B1%A1%E6%A3%8B)、[围棋](https://zh.wikipedia.org/wiki/%E5%9B%B4%E6%A3%8B)、[足球](https://zh.wikipedia.org/wiki/%E8%B6%B3%E7%90%83)、[篮球](https://zh.wikipedia.org/wiki/%E7%AF%AE%E7%90%83)等运动。网络游戏[英雄联盟](https://zh.wikipedia.org/wiki/%E8%8B%B1%E9%9B%84%E8%81%AF%E7%9B%9F)、[魔兽世界](https://zh.wikipedia.org/wiki/%E9%AD%94%E5%85%BD%E4%B8%96%E7%95%8C)内的竞技对战系统也采用此分级制度。

对ELO的理解：  
Elo rating system  
  
这个系统用来计算玩家的相对技术等级，通常用于博弈类游戏，比如国际象棋，围棋 \*注1。  
  
Elo,并不是一个缩写，而是系统发明者的姓，它本来是个改良的国际象棋积分系统，现在被广泛用于许多游戏。比如帝国时代3， 星际，魔兽争霸3，unreal,guild wars等等。  
  
  
组建模型  
1 Elo假设每个玩家在每盘游戏中的表现是一个正态分布的随机变量，尽管选手在不同的游戏中发挥可能差异很大，但每位选手在一段时间内表现的平均值变化很小。elo用随机变量的平均值来代表选手的真正水平。(这句话基本没什么意义)   
  
2 elo用胜 平 负来评价选手在某一场游戏中的表现，赢就是这场你发挥的比对手好，反之就是不好。(赢就加分，输就扣分，平手不得分),这是因为你不可能从某一手棋的好坏来评定一个选手的等级，下了一手好棋，或一串手筋就给2000分的评定，显然很难有说服力。  
  
这两条就是大致的模型构建，当然是简化了的。。为了简化计算elo有一个true skill值给每个选手，当然这个值还有很多优化的算法，跟魔兽没什么关系。有兴趣的同学可以参照xbox live online的一个算法[[true skill](http://research.microsoft.com/mlp/apg/trueskill.aspx)]。  
  
  
数学公式  
正题了， 做好笔记。。  
  
Ra: A选手当前rating Rb： B选手当前rating  
  
Sa: 实际胜负值， 胜 = 1 平=0。5 负=0  
  
Ea：预期A选手的胜负值，EA = 1/(1+10^[(Rb-Ra)/400])  
Eb: 预期B选手的胜负值 EB = 1/(1+10^[(Ra-Rb)/400])  
Ea+Eb=1  
E值也是预估的胜率，所以Ea+Eb =1,  
  
K：每场比赛能得到的最大rating，魔兽里k=32 \*注2  
  
R'a A选手一场比赛之后的rating  
  
R'a = Ra + K(Sa-Ea)  
  
举例讲解  
A队1500分，B队1600分，预估A队的胜负值Ea = 1/(1+10^[（1600-1500）/400]）=0.36  
  
假设A队赢了，实际胜负值为Sa=1  
  
A队最终得分为 R'a = 1500 + 32\*(1-0.36) = 1500+20.5 = 1520, 赢20分 B队输20分。  
  
假设B队赢了，预估B队胜负值Eb = 0.64  
B队最终得分为 R'b = 1600 + 32\*(1-0.64) = 1600 + 11.52 = 1612, 赢12分，A队输12分。  
  
文字这就是为什么你赢高分队分数多，输给低分的输的也多，赢低分的分数很少。  
  
其实K值就是这个方程的极限，所以理论上你最多可以赢一个队伍32分，实际上29-30已经差不多了，赢了不得分也是有可能的。  
  
平局在wow竞技场里也有可能发生， 最后2个玩家同时死亡，则都不得分。  
  
国际象棋rating的一些分析  
这部分跟wow没有特别大的关系，拿来参考以下。  
  
美国国际象棋联合会制定的评分及称号的标准，也是一1500作为基础分:  
  
\* 2400 及以上 超级大师  
\* 2200 - 2399 大师  
\* 2000 - 2199 专家  
\* 1800 - 1999 Class A  
\* 1600 - 1799 Class B  
\* 1400 - 1599 Class C  
\* 1200 - 1399 Class D  
\* 1000 - 1199 Class E  
全世界有19743人 2200以上，1868人2400-2499，563人2500-2599， 123人2600-2699，18人2700-2799，只有4人打到过2800。其中包括棋王卡斯帕罗夫。  
  
电脑深蓝如果评分的话估也计超2800。  
  
根据这个，算一下国际象棋有多少player pool，wow arena 有多少player pool就能知道wow 竞技场大致冲能到什么分数，目前看来2600是极限了，那也是得在人很多的服务器大组上，当然中国玩家冲到2800我也不会意外的。

运行相关的代码。

**2018.07.03**

**上午：**

**假设数据已经获取的情况下，阅读相关论文，拓展思路。**

**找到了关于比赛的几篇篇论文。对其中的算法进行学习和分析。**

**其中的一片伦分有很大的借鉴意义。谈到了如下的方法。**

**特征值的提取：**

Boruta算法是一个基于随机森林分类算法的包装器，寻找全相关属性。

了解Boruta算法。  
**分类模型：**

支持向量机和模糊支持向量机。

**对于支持向量机的深度理解。找相关 的论文，阅读论文。**

**与一同学讨论该算法和模型，相比于ELO等级制度哪个更合适。讨论结果：需要正确率来说话。**

**2018.7.04**

**上午：**

**阅读和分析ELO等级制度的代码。看懂代码。**

**目的是：**

**试图寻找ELO等级制度方法运用到篮球比赛的缺陷，结果找到了：**

**下午：**

**放假**

**2018.7.05**

**上午：**

**考试**

**下午：**

**组内讨论了数据除了、模型选择以及分工的问题。**

**暂时决定四人一小组，选择两个模型进行试验，最后对比两者的正确率。**

**开始整理本组模型的思路。寻找合适有效的模型。**

**深入了解了逻辑分布和SVM 的公式推导。**

**确定其是否满足于项目要求。**

**2018.07.06**

**G熟悉git和GitHub。**