

一、文献信息

作者 Ben Armstrong, Kate Larson

论文题目 Machine Learning to Strengthen Democracy

发表途径 AI for Social Good workshop at NeurIPS (2019), Vancouver, Canada.

发表时间 Sat Dec 14th, 2019

二、问题意义

事实证明，近年来，利用简单多数投票制的民主制度容易受到干扰，并且不能很好地反映选民的利益。虽然许多人同意存在更好的制度，但对于完美的选举制度如何发挥作用还没有明确的想法。这些问题可以通过切换到更强大的投票系统来缓解，但是实现这一目标的一个主要障碍是采用哪一种系统的共识。加拿大总理 Justin Trudeau 选择不进行选举改革，他表示，这样的制度，“尚未形成共识”。基于公正的无党派的投票规则对解决这些问题大有帮助。本文描述了一种机器学习方法来创建一个投票规则，旨在满足理想的选举公平标准集。作者首先提供了一个框架的描述，然后用一个简单的例子说明它能够实现一个强大的结果。

作者开发了一个基于理想的选举标准集来选择获胜候选人的系统。通过使用标准的机器学习技术，系统在许多可能相互排斥的标准之间找到了平衡，这些标准是根据选举的目标确定的。该系统没有外部培训数据，只有选定的标准，并更加代表了选民的偏好。

三、思路方法

本文首先介绍了这个项目的领域和背景，然后描述了系统的功能，并提供了一个训练有素的系统工作的简单例子。重点是开发一种替代传统的规则名为“计分”的投票函数，在这种规则中，每个投票者提交一份选票，将特定的分数分配给每个候选人，得分最高的候选人获胜。在简单多数制下，选票只给一个候选人一分，不给其他候选人分，在博尔达计数选举中，有 m 个候选人的选票时给一个候选人 m 分，另一个候选人 $m - 1$ 分，以此类推。

在本文中，作者开发了一个框架，用于生成新的、可定制的投票规则，目标是允许设计选举的人确定他们认为可以导致公平结果的一组公理(这可能是相互排斥的)，并使用这些公理来为一个神经网络生成训练数据，一旦训练，这个神经网络就充当投票规则。

3.1 选举公理

选举规则的公理定义了基于选举规则和选举结果可能适用或不适用的特定属性。有些公理集不能同时被满足，作者的方法接受任何一组具代表性的公理，并可以找到一个在满足或接近满足尽可能多的公理之间取得平衡的胜出者。

3.2 训练数据

通过成对 (B, c) 代表所选公理集的实际定义，其中 B 包含一组选票， c 是满足特定公理的赢家。如果有不止一个可能的赢家，可以使用多个对 (B, c_1) ， (B, c_2) ，等等进行训练。每个选票包含一个选民对所有 m 个候选人的偏好排序，例如 $b_i = [0, 2, 1]$ 表示选民 i 更喜欢候选人 0，而不是候选人 2，最后一个偏好是候选人 1。为了允许任意数量的投票人，将 B 转换成 $m \times m$ 偏好矩阵 E ，其中 $E_{i,j}$ 包含了更喜欢候选人 i 而不喜欢候选人 j 的投票人的比例

(其中 $E_i, i=0$)。用于训练这个模型的数据由对组成:一组对应于每个希望满足的公理组合的选票, 以及每个选票集的理想获胜候选人。每对重复 $m!$ 次其中在所有可能的顺序中列出一次候选人, 以避免将培训偏向于任何特定的候选人顺序。

3.3 训练过程

首先定义一个函数 $f(B, c)$, 它返回的是给定一组选票和候选人的每个候选人的值。 f 可能对应未满足的公理的数量, 或者用于对某些公理进行加权, 使其比其他公理更重要, 公式:

$c^* = \arg \min_{c \in C} f(B, c)$ 对应于理想的选举获胜者。作者将其描述为一个回归问题, 并训练一个回归神经网络来预测每个候选人的得分/惩罚, 然后选择得分最低的候选人作为获胜者。这也表明了对多个赢家选举的直接扩展: 选择 k 个得分最低的候选人(虽然作者在评估中没有探讨这种可能性, 但可以这样猜想)。

3.4 测试数据

本文使用加拿大选举官方报道提供的选举的数据来评估作者的系统。在每一选区, 居住在该选区的选民通过简单多数投票选出一名代表。为了简单和限制计算需求, 本文将重点限制在 6 个组, 通过使用一些公理来实现框架设计, 并测试经过训练的网络如何准确地选择一个惩罚最小的候选对象。还通过输入来自真实的加拿大选举的数据来衡量系统模拟真实选举的能力。

首先选择一组合适的公理。这可能是一个非常主观的过程, 因此作者将重点限制在一个系统上, 该系统在可能的情况下仅尝试选择孔多塞的获胜者, 或者将对获胜的数量最大化的候选人。作者定义函数 f 来优化一个孔多塞冠军或当他不存在时找到一个“近似孔多塞”冠军:

$$f(B, c) = \begin{cases} 0 & c \text{ is Condorcet winner,} \\ 0.5 & c \text{ wins the most pairwise elections but} \\ & \text{no Condorcet winner exists,} \\ 1 & \text{otherwise} \end{cases}$$

本文使用 $m = 6$ 个候选人(与加拿大选举数据相匹配)构造训练数据 X_{train} , 该数据由三部分组成: $X_{c=p}$ 是一组选票和获胜候选人, 其中有一个孔多塞获胜者, 且也是多数获胜者。 $X_{c \neq p}$ 是其中有孔多塞的获胜者, 而其不是多数的获胜者。 $X_{no c}$ 是一组没有孔多塞获胜者的选票和获胜候选人。这些数据被用来训练一个单一隐含层的回归神经网络。当使用加拿大选举数据进行测试时, 将多数选举结果转换为适当格式的选票。

四、实验结论

| Training Data | Training Accuracy |
|-----------------------------|-------------------|
| X_{train} | 0.76 |
| $X_{c=p}$ | 0.99 |
| $X_{c \neq p}$ | 0.91 |
| $X_{no c}$ | 1.0 |
| $X_{c=p} \cup X_{c \neq p}$ | 0.91 |

Table 1: Training accuracy when training with each subset of training data.

表 1 显示, 在孔多塞 Condorcet 获胜者存在的情况下, 显然能够满足被选择的公理, 其中组合不同的数据可能需要进一步的开发。当训练集表示类似的概念($X_{c=p}$ 和 $X_{c \neq p}$)时, 它们的合并具有合理的准确性。还可以使用每次加拿大选举的结果来测试每个训练集, 从表 2 中可以看到, 系统选择惩罚最小化的候选人比随机猜测要准确得多。且当训练数据只包含多数获胜者是孔多塞获胜者的样本时, 神经网络在准确性上的结果与实际选举中当选候选人的“准确性”非常相似(如表 3 所示)。

| Training Data | Accuracy per Year | | | |
|-----------------------------|-------------------|------|------|------|
| | 2006 | 2008 | 2011 | 2015 |
| X_{train} | 0.59 | 0.70 | 0.69 | 0.54 |
| $X_{c=p}$ | 0.92 | 0.94 | 0.94 | 0.91 |
| $X_{c \neq p}$ | 0.90 | 0.87 | 0.87 | 0.86 |
| $X_{no c}$ | 0.38 | 0.47 | 0.52 | 0.35 |
| $X_{c=p} \cup X_{c \neq p}$ | 0.80 | 0.84 | 0.78 | 0.75 |

Table 2: Testing accuracy using a model trained on artificial data to predict the winner in each district of each federal election.

| True Accuracy per Year | | | |
|------------------------|------|------|------|
| 2006 | 2008 | 2011 | 2015 |
| 0.95 | 0.97 | 0.96 | 0.97 |

Table 3: The frequency with which the true elected candidate in a real district is the candidate that minimizes our penalty function.

(表 2:使用在人工数据上训练的模型来测试准确性,以预测每次联邦选举中每个地区的获胜者。表 3:在一个真正的地区, 当选候选人是使惩罚函数最小化的候选人的概率。)

使用这种制度的好处是, 在小范围内, 会产生一种明确的投票模式或结果, 并且可以被选民普遍接受。

五、 启发思考

像本文这样的系统是对最近集中于投票规则的算法生成的研究领域的一种独特贡献。但目前的方法中用于训练网络的数据结构是一个明显的限制。如果能找到一种更有代表性的数据结构, 使所有选票信息的大小不受选民人数的影响, 或建立一种网络结构, 在不影响任何具体结果的情况下纳入选票, 将加强该系统的灵活性。虽然目前的系统已经显示出很有希望的评估结果, 并可能通过修改培训程序或数据结构加以改进, 但使用该系统的领域有限。由于这个系统, 就像大多数神经网络一样, 是不能完全解释的, 所以只有一个基于统计上的保证, 即它不会产生人们会认为明显“错误”的赢家。因此, 它的使用应该限制在赢家可以造成有限影响, 或者限制了谁可以成为候选人的情况下。

本篇论文把投票机制无论是传统的还是新兴的, 都解释的很详细。根据对一些公理的分析, 结合神经网络学习形成更可靠的投票机制是一种很好的想法。设计了一个惩罚函数, 然后用一个公式代表当函数取得最小时的变量值作为理想的选举获胜者, 然后通过对比经过不同训练集后的准确性与实际情况的概率对比产生结果。目前机器学习已经在很多领域逐步的影响了人们的生活, 包括图片数字识别, 疾病发病率的预测, 儿童阅读障碍分析等等, 还有许多未知领域等待去探索发现。经过课堂上老师的讲解和浏览搜集的论文、课程视频, 我对机器学习产生了浓厚的兴趣, 也想在大作业上尝试分析并运行代码, 帮助自己在机器学习这个模块上有更深的了解和感悟。希望以后有机会无论是研究生还是工作上还可以继续钻研并与实践相结合。