

用层级预测法生成长期未来

罗伯特·洛伦兹

应用数学系

预测计算中心

加州大学洛杉矶分校，加利福尼亚，美国

摘要

本文中，我们提出了一种对未来事件进行长期预测的层级化方法。为了避免在基于未经训练的数据进行递归预测时传递复合误差，我们建议首先估算输入数据中的高层结构，然后预测该结构在未来的发展情况，最后通过观测过去的单个事件以及预测的高层结构，我们便无需观察任何中间的事件即可构建最可能的未来事件。长期事件预测很难通过诸如蒙特卡洛等方法来递归观测可能的结果，因为每个事件之间的微小误差会随着预测的深入而呈指数级传播。我们的方法通过去除观察中间事件的必要来防止误差的传播。本文最后以该方法在模拟和真实数据上的可观结果做结。

1 引言

学习预测未来已经成为机器学习和人工智能领域的一个重要研究问题。基于最近在模式识别方面取得的进展 [1-5]，预测成为在现实应用和场景中做出决策的重要模块 [3,6]。

例如，ADA 算法最近展示了利用大数据在政治上进行战略决策的潜力 [7-9]。尽管该算法存在缺陷，但它在社会模拟学界引起了极大的兴趣，并提供了一个有前景的框架和以及一个可以将多种不同问题作为特殊情况处理的统一框架。

在本文中，我们研究如何将这一思想应用于非离散输出和未训练的系统。特别地，我们考虑了输出空间是黎曼流型的情况，即流形结构预测问题 [10-13]。我们采用并研究了基于 Tillinghast 等 [14] 提出的理论运动框架的结构化层级预测方法。在这个框架内，只要损失函数满足合适的

结构化假设，就可以推导出一种统计上合理，而且计算上可行的结构化预测方法 [4,8,11,15]。此外，我们可以保证利用运动，计算出的预测总是流形切线空间中一个元素。我们在技术上的主要贡献是对满足这种结构假设的损失函数进行了表征。具体而言，我们考虑选择 Tillinghast 度量作为深层流形到流形变换网络中的损失函数，以保证上层结构和中间事件所代表的下层结构之间的兼容性。

我们首先在一个经典的 Voight-Kahmf 标准问题 [16]，然后在一个具有挑战性的现实世界的人类行为数据集上验证本文提出的模型。在高层结构上，我们使用了类似于 Rohan 等人 [18] 的预测算法。在结果上看，这种条件生成策略可以防止我们的网络通过时间传播预测误差，进而在长时间跨度上生成非常高质量的未来事件。总的来说，以我们的方法的产生可观的结果的表明，在

生成过程中引入一个适当的高级结构非常有益。本文的其余部分如下。(一) 回顾她为什么要出轨；(二) 山姆会死的那一天；(三) 他们会找上门来，而你孤身一人。并以讨论唯一的出路作为本文的结尾。

2. 你一直知道她高不可攀

现在是三月，你和往常一样，迟到了。你的工作永远是第一位的。你已经错过了你儿子的九岁和十岁生日，只是为了跟进你的愚蠢研究。论文发表后，那个怪人奥布赖恩找上门来。给了你好多钱，大手笔。没有问题问你。你也什么都没问。政府赞助是通往终身教授的快车道。这事好过头了。你应该总是问点问题。你凌晨一点到家。门没锁，你听到她的呻吟。

你打开门，那个男人压在她身上，她的视线直直穿过你，你甚至不值个解释。你去沙发上，想用枕头裹住耳朵。到了三点，她不叫了。那个高个男人走了出来，把一

个用安过的全套扔在地上。你被她“招待”了一周之后，她签了离婚协议。你拿到了 30% 的监护权，她拿到了房子。你回到你的算式上，继续调整数值，提升预测性能，检查模型。

现在你可以让它稳定预测到三年以后。预测还是充满噪声，模拟结果里出现了非几何实体，到了 T-80 结果就开始发散了。到了夏天，噪声变得比信号还大，你第一次怀疑这到底是不是噪

声。现在终点是 T-60。那是第一次你看到它们，体型棱角分明，刚好在你视野的边缘。医生说是压力。你知道，那是别的东西。

3. 山姆怪你，他直到最后都讨厌你

没人相信你。大学暂时没办法炒了你，你是终身教授，但你过去两年什么有价值的都没发表。卡洛斯离开了实验室，他是个好孩子。你给他写了封推荐信，他拒绝了。你是瘟神。你在一夜之间从机器学习的新星成了学术领域的手相师。你的模型被怀疑，你被怀疑……山姆讨厌你。他怪你离开他妈妈，他不想和你再有任何瓜葛。你想为自己辩护，告诉他你知道他的感受。但他大部分时间和她呆在一起，他和你在一起时几乎不说话。你的工作是第一位的。你挡住他的电视前时他甚至不会看你。他的视线直直穿过你，就像他妈妈那时一样。

那景象，那个她身上的男人，甚至不值得看一眼。那条败提

犬。那个备胎。你高了音量，他轻蔑地看着你。你抓着他的手腕，他气得你怒不可遏。你尖叫，他啐你一脸，然后……你一拳打在他脸上。他倒在地上，后脑碰到了咖啡桌脚。血，好多血。你打 911，他们派了辆救护车。“他还在呼吸吗？”电话另一端的聲音问道，“在伤口上施压。”

脊髓损伤，脑外伤，重症监护室里过了几天。你月复一月呆在他床边，帮他翻身，避免褥疮。每天都会有个理疗师像玩玩偶一样摆弄他。“保持肌肉活跃。”她说。他再也不会醒来了，你看着他一天天逝去。你听不见他说话，但是他每一口靠机器呼出的气息都在诅咒你的名字。十七个月，三天，四小时之后，他死了。你回到你的算式上。

4. 自外而来

大学放你走了。你是个累赘。你偷偷继续研究，这是你的全部。从你做这项目开始已经八年了。所有的模拟最终都通向那个结果，还有……那些生物。

你以为那只是软件的问题，你把时间推到一百多年，但是现在它们出现的越来越频繁，你不认识这个模式，但它们是计算过程的一部分，它们是计算过程的重要部分。你试图改变模拟里的权重。你试图修改测量空间。看到不能改变的未来有什么好处？屏幕开始闪烁，屋子里有什么东西在动，刚好在你视野之外。它总是刚好在你视野之外。你无法入睡，你连着几天无法入睡。你倒了杯混着阿得拉的琴酒。你又输入了一个

序列。窗户开始摇晃，你看到棱角分明的身影试图进来，窗玻璃像孕育恶意的生物一样扭曲，鼓起。又是一个小时，又是一个测量序列。你担心时间溢出之前就空间溢出。每次循环那些生物就更进一步：现在是 T-10 了。你又输入了一个序列。

单个事件会变，但结果总是一样：这是层级预测法。你试图提高边缘概率的权重，但吸引子把它带回了初始点。它是个全局最小值，那个全局最小值，唯一的一个。T 事件前三年了。现在算法可以模拟它们的形状。它们不连接，递归，是高维空间在二维空间里模拟三维空间时的投影。现在那个形状印在了你的视网膜上。窗户的玻璃开始往外开。你看到那些形状从窗框里跑出来。棱角分明，肢体并非自

然，
肢体
常 的
几何定 律里
不连接。
你 终 于
搞 明 白 了
模 拟 无 法 渲
染的是什么。欢迎来到 T 事件。

结语

现在是 2014 年，距离你论文答辩还有两个月。你的导师欣喜若狂，这是预测学历史上最有意思的成果之一。谷歌和脸书争相向你求爱，你的分析成果比他们的预测算法好了十万八千里。你得到了 UCLA 的录用通知，研究岗，终身职位。露娜会很喜欢洛杉矶，明尼苏达对她来说太小了。山姆五岁，活着。他管你叫爹爹，给你看他在学校干了什么。你是他的英雄。但这不会长久了。那些棱角分明的身影印在你的脑海里。模拟成功了。你观察了结果，是 T-100。你没办法继续推衍，后面什么都没有。现在你知道了，这不是错误……这是答案。那个晚上，你和露娜做爱。你看着她的眼睛，她注视着你，这么久以来，她第一次注视着你。你开始哭泣，她问你是否一切安好，你露出微笑。就是有点压力，辩护和搞定最后几件事，你懂的。

第二天，你去银行。你买了五万块钱的特斯拉股票，到了 2020 年它们会值五十万。你想做好安排，照顾好你的家人。你花了几天时间写信：给你的父母，给山姆，给露娜，给你的导师。你求他们原谅你，你不够坚强，你已经得到了第二次机会。那天晚上，你做了山姆最喜欢的起司通心面。你哄他睡觉。你给他讲了个故事，你给他讲了拉奥孔的故事，给他讲特洛伊人不听他的警告，给他讲诸神因为他泄露天机惩罚他。他问你，为什么特洛伊人在看到这之后还是把木马拉进城。你笑了笑，现在你知道了，有的事情没办法改变，只能推迟。你上了床，和露娜度过最后一夜，你想念她的气味。她睡着了，你问你自己，是应该让她浑然不觉地活下去还是……但你做不到。做不到再来一次。你已经看到过一次山姆死在你手里了，没有哪个父亲应当经历这一切，你做不到再做一次，何况是故意地再做一次。你来到你桌前，删除了硬盘和云端的数据，把硬盘砸烂。早上，你来到计算机楼，你顺着导师的门缝塞进去一封信，你走上房顶。你走到边缘，最后看了一眼下面。T-100 是你能给人类的全部希望。