

王博士，您好

我仔细阅读了您的这篇论文：“Learning User-Specific Latent Influence and Susceptibility from Information Cascades”，并对cascade size prediction做了仿真。仿真过程中用到的数据集为train.dat.1、net.1以及source.ret.dat.1(训练和测试均在同一数据集上进行)。仿真结果如图1所示，结果并不好。真实数据在cascade size为2处有一个波峰，而预测结果的波峰在1处。同时，MAPE值为1.27，比您论文中的结果0.140大了一个数量级。

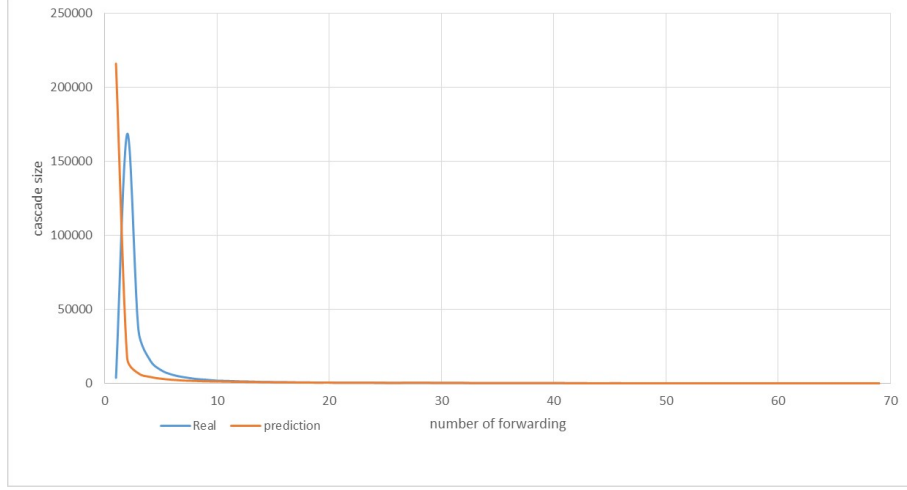


图 1: 预测结果

我有一些疑问，不知您可否帮我解答一下？

- (1) 向量 $I_u$ 和 $S_v$ 是 $n$ 维向量， $n$ 值应该怎么取？
- (2) 在train.dat.1中，某些数据如“1|32335\*-3349|1”中的\*号代表什么含义？
- (3) 在对目标函数

$$\mathcal{L}(C) = - \sum_{v \in V} \sum_{D_{v,i} \in \Gamma(v)}^N (n_{z_{v,i}, D_{v,i}} \log p(z_{v,i} | z_{v,i-1}, D_{v,i}, \delta)) + \gamma_I \|I\|_F^2 + \gamma_S \|S\|_F^2$$

使用投影梯度下降法时，目标函数的初始值大概在 $8.6 \times 10^6$ 左右，经过几百次的迭代下降后(需要数个小时)也只能下降十分之一。目标函数是凸函数吗？投影梯度下降能找到最优解吗？

- (4) 原文公式9中的

$$\frac{\partial \mathcal{L}}{\partial S_v} = -\lambda \sum_{D_{v,i} \in \Gamma(v)} \sum_{u \in D_{v,i}} I_u (n_{z_{v,u}=1, D_{v,u}} \frac{1 - p_{v, D_{v,u}}}{p_{v, D_{v,u}}} - n_{z_{v,u}=0, D_{v,u}}) + \gamma_S S_v$$

中的cascade context为何是 $D_{v,u}$ 而不是 $D_{v,i}$ 呢？

- (5) 能否告知我一些参数的值或范围？比如 $\lambda$ 、 $\gamma_I$ 、 $\gamma_S$ 。