



Curriculum Learning和Self-paced Learning的相关知识及应用

本次分享主要围绕强化学习的方向，向大家介绍，Curriculum Learning和Self-paced Learning的相关知识及应用。

Curriculum Learning和Self-paced Learning的介绍如下：

Curriculum Learning和Self-paced Learning代表了最近提出的学习制度，其受到人类和动物学习过程的启发，这些学习过程逐渐从训练中的简单复杂样本开始。这两种方法具有相似的概念学习范式，但在具体的学习方案上有所不同。

在Curriculum Learning中，课程由先前知识预先确定，并在此之后保持固定。因此，这种方法在很大程度上依赖于先前知识的质量而忽略了关于学习者的反馈。

在Self-paced Learning中，课程是动态决定的，以适应学习者的学习节奏。但是，Self-paced Learning无法处理先前的知识，使其容易过度拟合。

Curriculum Learning

Curriculum Learning有点类似人类学习机制——先学简单的技能，再学困难的。学习有意义的训练数据顺序可以使各种任务受益，即首先选择更容易学习的例子，然后逐渐增加难度。

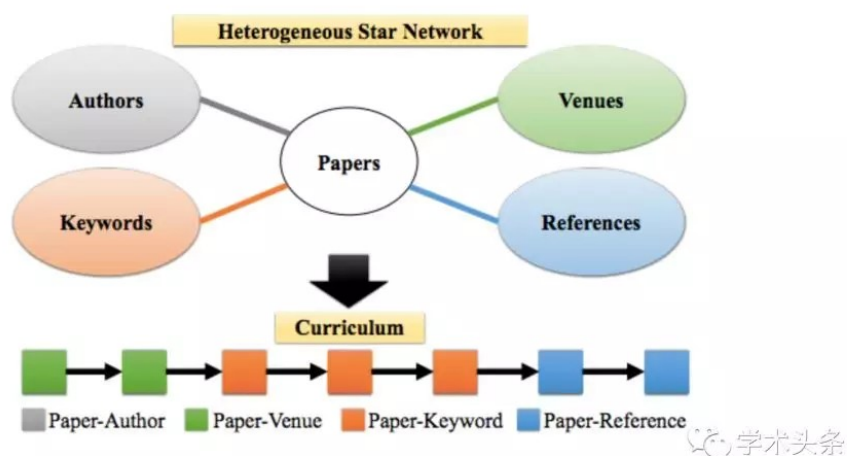
课程学习的概念是2009年 Yoshua Bengio等人在《Curriculum learning》一文中提出来的。其特点包括：

- 提高生成速度和加快收敛速度
- 在非凸的训练准则上找到更好的局部极小值

今天分享的第一篇文章是《Curriculum Learning for Heterogeneous Star Network Embedding via Deep Reinforcement Learning》。

该文主要研究异构星型网络的学习节点表示，该异构星型网络的中心节点类型通过不同类型的边与多属性节点类型相连。

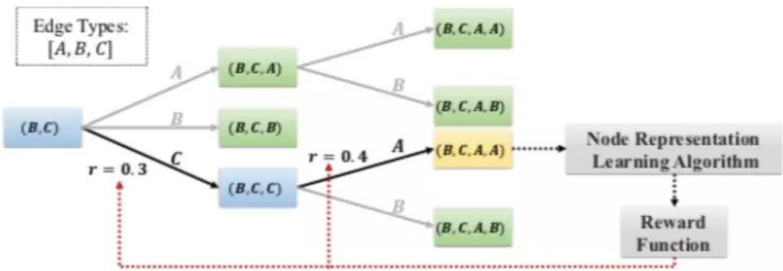
Heterogeneous Star Network



并提出了一个基于深度增强学习的方法，使用LSTM模型来编码状态，并进一步估计每个state-actionpair的期望cumulative reward。该文在深度增强学习的基础上融合了learning和planning的策略。实验证明该方法有效且高效。

规划模块的动作是选择某种边缘类型，在每一步中，状态被定义为到目前为止选择的边缘类型序列。在节点分类任务中，将奖励定义为准确度增益，惩罚每一个动作。目标是采取一系列行动以最大化累积奖励。

Planning Module



Given a state, we simulate a series of actions (the black path) until we reach an unvisited state (the yellow one). Then the rewards are calculated to approximate Q_p .

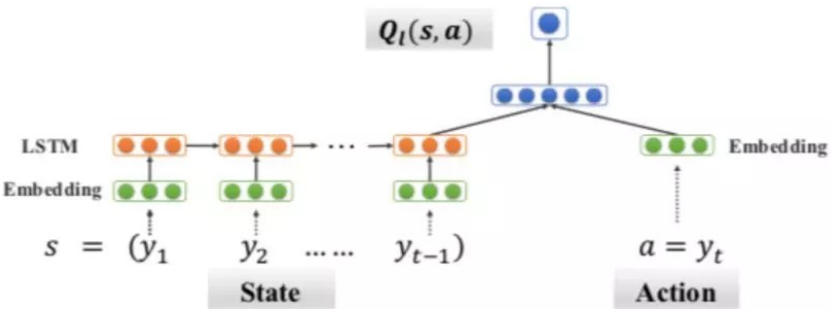
$$a = \arg \max_a \{ \frac{Q_p(s, a)N(s, a)}{N(s, a) + 1} + \frac{Q_l(s, a)}{N(s, a) + 1} + \lambda \sqrt{\frac{\ln N(s)}{N(s, a) + 1}} \}.$$

$$Q_p(s_i, a_i) = Q_p(s_i, a_i) + \alpha [r_i + Q_p(s_{i+1}, a_{i+1}) - Q_p(s_i, a_i)],$$

学术头条

学习模块是通过利用LSTM层对状态进行编码，我们可以有效地捕获不同状态的相关性，以推断新的状态 - 动作对的Q值。

Learning Module



By leveraging LSTM layers to encode states, we can effectively capture the correlations of different states to infer Q values of

new state-action pairs.

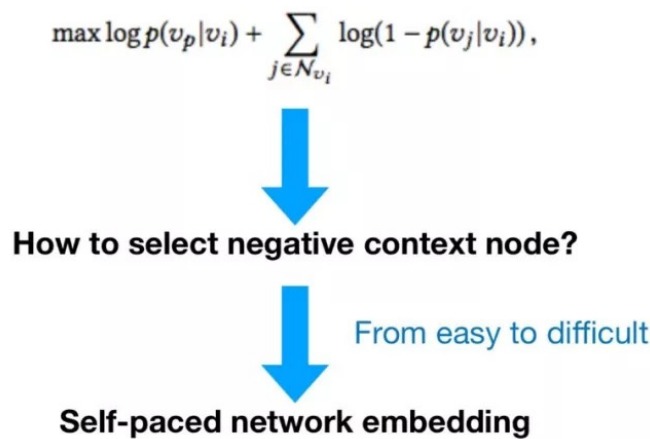
$$w_l = w_l + \alpha [r_i + Q_l(s_{i+1}, a_{i+1}) - Q_l(s_i, a_i)] \nabla_{w_l} Q_l(s_i, a_i),$$

学术头条

Self-paced Learning

今天分享的第二篇文章是《Self-paced network embedding》。

由于传统的抽样分布不能捕获每个节点的真实信息，并且不能反映训练的状态，本文提出了一种新的self-paced network embedding方法来解决此问题。



学术头条

该方法能够根据当前训练状态自适应地捕获每个节点的信息量，并根据其信息量对负上下文节点进行采样。所提出的self-pace采样策略能够随着训练过程的进行，逐步选择困难的负面上下文节点，以学习更好的节点表示。

Self-Paced Negative Sampling

- Always select the more informative nodes at each iteration, yet it fails to utilize all nodes.
- Should feed easy samples at early training stage and gradually provide difficult samples.

$$\max \log p(v_p|v_i) + \sum \log(1 - p(v_j|v_i)) + l(\mu)$$

$$j \sim_{p'_{ij}} \mathcal{N}_{v_i}$$

$$s.t. \quad p'_{ij} = \begin{cases} p_{ij}, & p_{ij} < l(\mu) \\ 0, & otherwise, \end{cases}$$

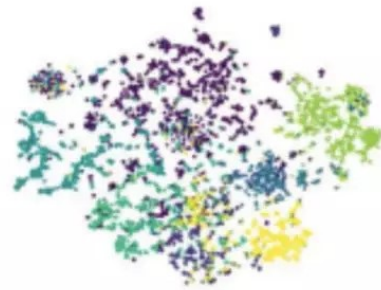
- set the probability of difficult negative context nodes to zero such that the easy ones have **larger** chance to be selected.
- With the training process going on, $l(\mu)$ is **increasing** so that difficult negative context nodes will be included gradually.

此外，为了更好地捕捉节点信息性以学习节点表示，将该文的方法扩展到具有更大发现节点信息能力的生成性对抗网络框架。在基准网络数据集上进行了大量的实验，验证了所提出方法的有效性。

Experiments



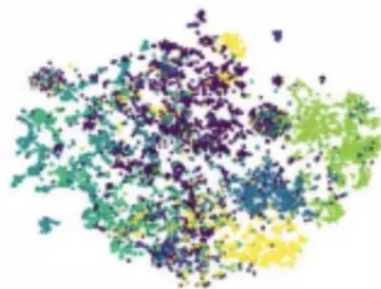
(a) DeepWalk



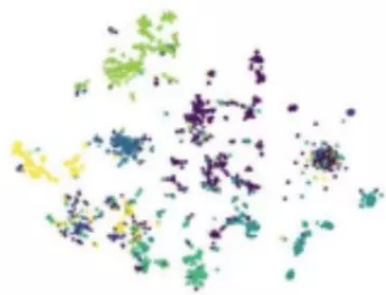
(b) Node2Vec



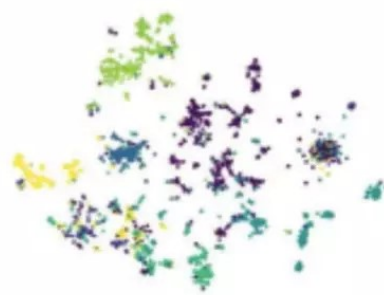
(c) GraRep



(d) LINE



(e) SeedNE



(f) ASeedNE

学术头条



AMiner学术头条

AMiner平台由清华大学计算机系研发，拥有我国完全自主知识产权。系统2006年上线，吸引了全球220个国家/地区800...

<https://www.aminer.cn/>



理论 深度强化学习 Self-paced Learning

❤ 3 💬 🔗

Curriculum Learning LSTM

相关数据

约书亚·本吉奥 • 人物



基准 • 技术



Q值 • 技术



收敛 • 技术



规划 • 技术



展开全部数据 ▾