Personal Recommendation Algorithm

Main Flow

- 个性化召回算法Item2vec背景与物理意义
- Item2vec算法应用主流程
- Item2vec算法依赖model word2vec介绍

背景

- Item2item的推荐方式效果显著
- NN model的特征抽象能力
- 算法论文:ITEM2VEC: NEURAL ITEM EMBEDDING FOR COLLABORATIVE FILTERING

物理意义

- 将用户的行为序列转化成item组成的句子

- 模仿word2vec训练word embedding将item embedding

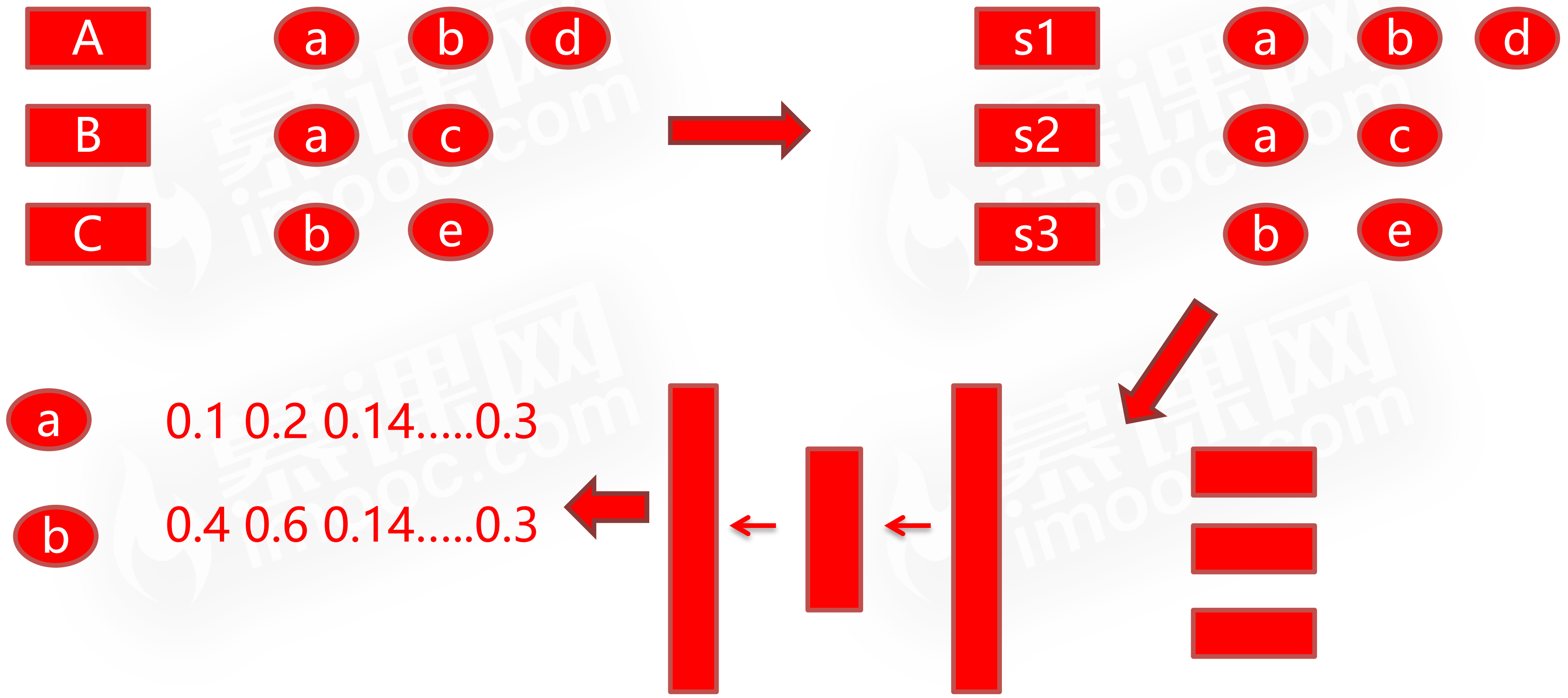
缺陷

- 用户的行为序列时序性缺失
- 用户行为序列中的item强度是无区分性的

Item2vec算法主流程

- 从log中抽取用户行为序列
- 将行为序列当成语料训练word2vec得到item embedding
- 得到item sim关系用于推荐

Example

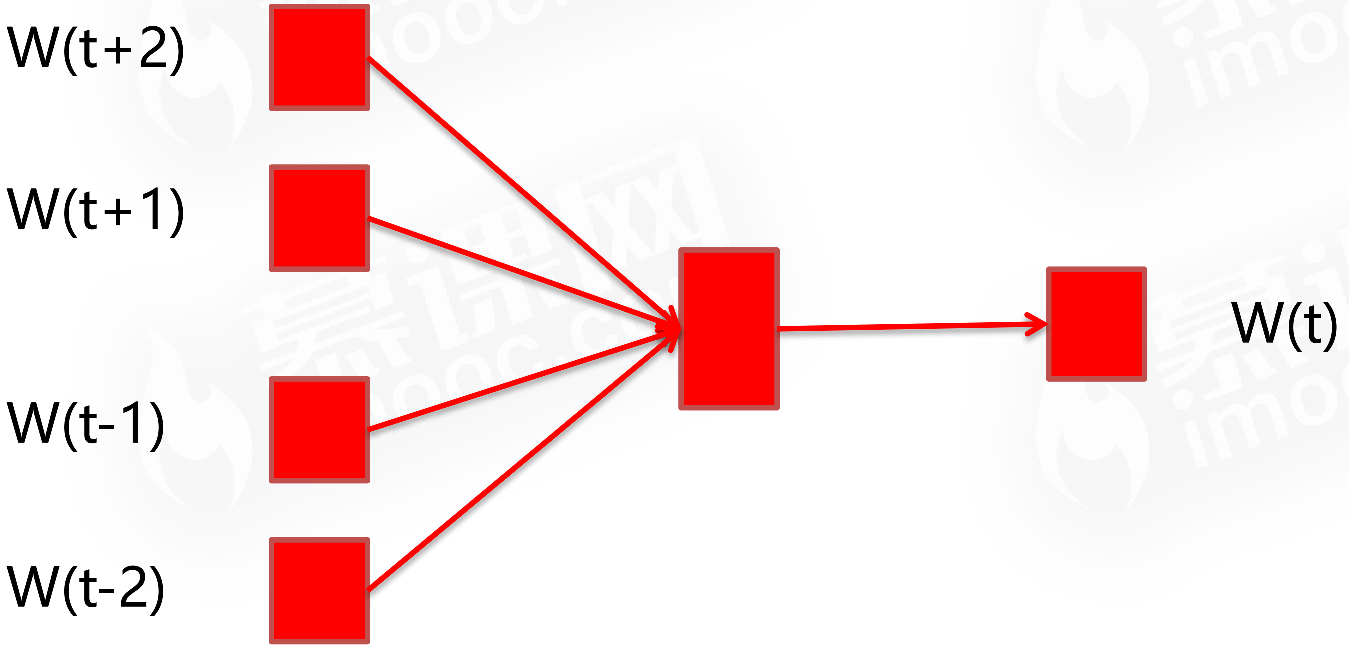


Class Two

Personal Recommendation Algorithm

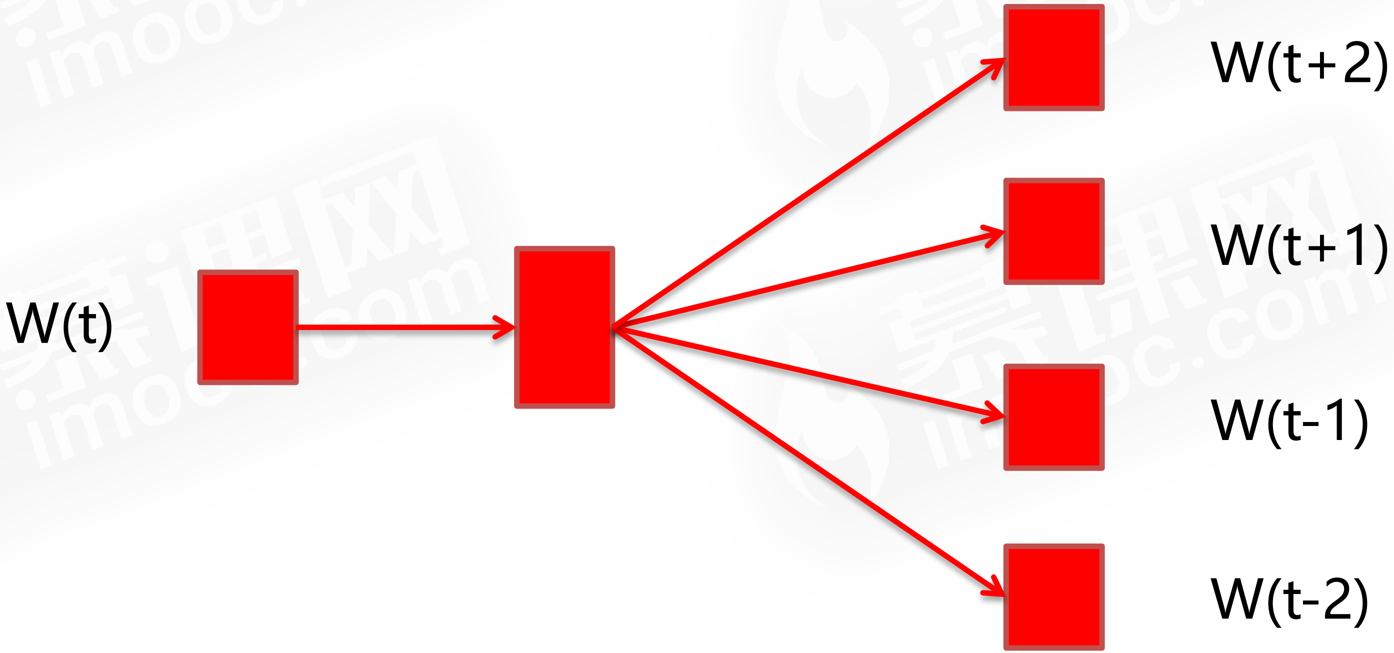
Word2vec Model

- CBOW (continuous bag of words)



Word2vec Model

- Skip Gram



CBOW Word2vec 数学原理

- 问题抽象

$$g(w) = \prod_{u \in w \cup NEG(w)} p\left(u \mid Context(w)\right)$$

$$p\left(u \mid Context(w)\right) = \sigma\left(X_w^T \theta^u\right)^{L^w(u)} \left(1 - \sigma\left(X_w^T \theta^u\right)\right)^{\left(1 - L^w(u)\right)}$$

Loss Function

$$Loss = \log\left(g\left(w\right)\right)$$

$$Loss = \sum\left(L^{w}\left(u\right) * \log\left(\sigma\left(x_{w}^{T}\theta^{u}\right)\right) + \left(1 - L^{w}\left(u\right)\right) * \log\left(1 - \sigma\left(x_{w}^{T}\theta^{u}\right)\right)\right)$$

梯度

$$\frac{\partial Loss}{\partial \theta^{u}}=\left(L^{w}(u)-\delta\left(x_{w}^{T}\theta^{u}\right)\right)x_{w}$$

$$\theta^{u}=\theta^{u}+\alpha*\frac{\partial Loss}{\partial \theta^{u}}$$

$$\frac{\partial Loss}{\partial x_{w}}=\left(L^{w}(u)-\delta\left(x_{w}^{T}\theta^{u}\right)\right)\theta^{u}$$

$$v\left(w_{context}\right)=v\left(w_{context}\right)+\sum_{u\in w\cup NEG(w)}\alpha*\frac{\partial Loss}{\partial x_{w}}$$

CBOW训练主流程

- 选取中心词w以及负采样出NEG(w)
- 分别获得损失函数对于x_w与$theta^u$的梯度
- 更新$theta^u$以及中心词对应的context(w)的每一个词的词向量

Class three

Personal Recommendation Algorithm

Skip Gram Word2vec 数学原理

- 问题抽象

$$G = \prod_{u \in Context(w)} \prod_{z \in u \cup NEG(u)} p(z|w)$$

$$p(z|w) = \left(\delta\left(v(w)^T \theta^z\right)\right)^{L^u(z)} * \left(1 - \delta\left(v(w)^T \theta^z\right)\right)^{1-L^u(z)}$$

Loss Function

$$Loss = \sum_{u \in Context(w)} \sum_{z \in u \cup NEG(u)} L^u(z) * \log\left(\delta\left(v(w)^T \theta^z\right)\right) + \left(1 - L^u(z)\right) * \log\left(1 - \delta\left(v(w)^T \theta^z\right)\right)$$

$$G = \prod_{w^c \in context(w)} \prod_{u \in w \cup NEG(w)} p\left(u \mid w^c\right)$$

$$Loss = \sum_{w^c \in Context(w)} \sum_{u \in w \cup NEG(w)} L^w(u) * \log\left(\delta\left(v(w^c)^T \theta^u\right)\right) + \left(1 - L^w(u)\right) * \log\left(1 - \delta\left(v(w^c)^T \theta^u\right)\right)$$

Skip Gram word2vec 训练主流程

- 对于context(w)中任何一个词w^c,选取词w的正负样本
- 计算Loss对theta以及对w^c的偏导
- 更新w^c对应的词向量

负采样算法

$$len(word) = \frac{(counter(word))^{\alpha}}{\sum_{w \in D} (counter(w))^{\alpha}}$$