

周报——2022.5.29

一、完成工作

这两周主要看了两篇论文——**StyleGAN2**和**Image2StyleGAN**（详细的论文笔记可以看<https://github.com/pang-discussion/gq>），其他完成了两项课程作业，一个是实现了算法作业（通过五种方法完成字符串匹配）另一个是NLP作业（使用Wikipedia和pubmed语料库以及Word2vec和glove方法实现两个词向量模型，具体的代码和预训练权重在<https://github.com/gq-max/NLP>里面）

1. StyleGAN2

这个是一篇**图像生成**的文章，解决了StyleGAN里面存在的两个问题——水滴状伪影和特征滞留问题。严格来说这篇论文是一个实验性的论文，通过问题然后一步步分解模型，找到问题存在的地方，然后通过实验解决这个问题。另外文章还提出了一个PPL损失（因为感知路径长度相比于L2损失来说，更能够反映一张图片的质量问题）以及一个lazy regularization。从这篇论文学到了：

- 感知损失，其中最著名的是LPIPS损失（学习感知图像块相似度），这种感知损失比L1/L2损失相比来说，更能够反映图像的质量，后面可以看看能不能用到对抗样本的防御当中（因为对抗样本添加了部分噪声，图片质量没有原始图片好）
- lazy regularization，这个方法可以加快模型的训练
- stylegan2的潜在变量空间是一个比较好的潜在变量空间，可以将这个潜在空间与词向量的潜在空间一起训练，得到一些多模态东西（比如文本驱动的图像生成和修改）

2. Image2StyleGAN2

这个是**图像操作**的一篇文章，就是把一张图映射到stylegan的潜在空间当中。只不过这个思路生成的时间太长了，后期有一些通过GAN的思想来实时实现图像向潜在空间映射的方法，可以直接使用。从这篇论文主要学到了

- 图像操作的一种有效方法
- 可以直接使用的图像到潜在空间的映射代码

3. word embedding

词嵌入

2022年5月28日 22:19

词嵌入

1. dog apple one (可以用来测试的词)
2. $e_{\text{man}} - e_{\text{woman}} \approx e_{\text{king}} - e_{\text{queen}}$
3. t-sne 可视化、余弦 cosine 相似度 $[-1, 1]$

一. 嵌入矩阵

1. 包含了一个 $\langle \text{unk} \rangle$, 用于标识未出现的词和词频太低词

直接通过 one-hot 编码取出 Embedding Matrix 的特定行得到单词的编码。

目标: 学习一个嵌入矩阵 (10000, 300), 10000 个单词, 300 维。

二. 学习词嵌入

1. 固定的窗口

2. I want a glass of orange juice to go along with my cereal

context
target
context

三. Word 2 Vec

1. Skip-Gram

1. I want a glass of orange juice to go along with my cereal

target
context
target

context c ("orange") \xrightarrow{x} Target t ("juice")

$O_c \rightarrow E \rightarrow e_c \rightarrow O_{\text{softmax}} \rightarrow y$

2. 计算速度: 很难推广到大的词汇表 (bulk softmax 分类)

① 分层 softmax \rightarrow Hierarchical softmax

② 负采样 \rightarrow 如何采样 context c

均匀采样: 每个词会被采样很多次 (x)

四. 负采样

1. softmax 的计算代价很大

2. I want a glass of orange juice to go along with my cereal.

它下一个新的学习任务: 预测两个词 是否是上下文 (例如 orange 和 juice, orange for king)

0 或 1 选取 (orange, juice) $\rightarrow 1$ (正样本)

1

0

① 选取 (orange, juice) $\rightarrow 1$ (正样本)
 context target

② 选择词汇表随机单词, 例如 (orange, king) $\rightarrow 0$ (负样本)
 构造多个负样本 (超参数 k)

③ 构造一个监督学习方法 (分类模型)

k 的选择 (小数据集 5~20, 大数据集 2~5)

$$\text{Softmax: } p(t|c) = \frac{e^{\theta_t^T e_c}}{\sum_{j=1}^{10000} e^{\theta_j^T e_c}}$$

$$p(y=1|c,t) = \sigma(\theta_t^T \cdot e_c)$$

$$O_c \rightarrow E \rightarrow e \begin{cases} \leftarrow \\ \rightarrow \end{cases} \begin{matrix} \uparrow \\ \downarrow \end{matrix} \quad (n \text{ 个负样本, } 1 \text{ 个正样本})$$

2. 选择负样本

① 根据词频采样

$$p(w_i) = \frac{f(w_i)^{\frac{1}{\alpha}}}{\sum_{j=1}^{10000} f(w_j)^{\frac{1}{\alpha}}}$$

五. GloVe: 用于词汇表征的全局变量

1. I want a glass of orange juice to go along with my cereal.

X_{ij} : i 在 j 的上下文中出现的次数 $\rightarrow X$: 对称矩阵 (例如彼此相邻)
 (i,j) 同时出现的次数

$$2. \min_x \sum_{i=1}^{10000} \sum_{j=1}^{10000} f(X_{ij}) (\theta_i^T e_j + b_i - b_j - \log X_{ij})^2$$

$f(X_{ij})$: ① 当 $X_{ij} \rightarrow 0$ 时 $\log X_{ij}$ 无穷大, 因此加罚 $f(X_{ij}) \rightarrow 0 \cdot \log \rightarrow 0$

② 有一些停用词 (this, is, of, a, ...), 降低它们出现概率

$$3. \theta_i \text{ 和 } e_i \text{ 是对称的} \Rightarrow e_w = \frac{e_j + \theta_i}{2}$$

无法确保学习到的维度是可解释的, 特征之间是耦合的 (线性组合)

在训练 Word2vec 的时候, 因为采用的是 softmax 函数, 而词汇表单词的数目基本上都在 10000 左右, 每次迭代计算 softmax 的计算代价太大, 可以采用负采样

二、后面安排

- 看 StyleCLIP 这篇文章
- 看陈志凯师兄和贺勇师兄的毕设论文
- 看一看对抗样本、deepfake 和 CLIP 这三个方法可不可以结合在一起得到文本——图片 (视频) 多模态的对抗样本、deepfake 的攻防

