

基于ACSA-CapNet的方面级情感分析研究

孙阁阁, 游福成

北京印刷学院, 北京

收稿日期: 2022年11月7日; 录用日期: 2022年12月7日; 发布日期: 2022年12月28日

摘要

方面情感分析(Asspect Based Sentiment Analysis, ABSA)是指预测给定方面的情感极性, 已引起广泛关注。在现有的ABSA数据集中, 大多数句子只包含一个方面或具有相同情感极性的多个方面, 这使得ABSA任务退化为句子层面的情感分析。本文中, 我们采用了一个新的大规模多角度多情绪(MAMS)数据集, 其中每个句子至少包含两个具有不同情感极性的不同方面, 但该数据集类别分布极其不平衡, 导致模型准确率受到某个类别的影响。为了解决这一问题, 我们在原数据集的基础上进行了数据增强, 并提出了一个新的MAMS的数据集, 并将其命名为ACSA-V2。此外, 我们在CapsNet模型的基础上, 对压缩函数进行了修改, 使得模型在提升准确率的同时可以更快地收敛。实验结果显示, 相较于江等人的胶囊网络模型, 基于改进的胶囊模型在方面级情感分析任务上准确率提高了约5个点。

关键词

胶囊网络, 方面级, 情感分析, 动态路由

A Study on Aspect-Level Sentiment Analysis Based On ACSA-CapNet

Gege Sun, Fucheng You

Beijing Institute of Graphic Communication, Beijing

Received: Nov. 7th, 2022; accepted: Dec. 7th, 2022; published: Dec. 28th, 2022

Abstract

Aspect Based Sentiment Analysis (ABSA), which refers to the prediction of the sentiment polarity of a given aspect, has attracted a great deal of attention. In existing ABSA datasets, most sentences contain only one aspect or multiple aspects with the same sentiment polarity, which degrades the ABSA task to sentence-level sentiment analysis. In this paper, we use a new large-scale multi-angle

multi-emotional (MAMS) dataset in which each sentence contains at least two different aspects with different sentiment polarity, but the distribution of categories in this dataset is extremely unbalanced, resulting in model accuracy being influenced by a particular category. To address this issue, we augmented the original dataset with data and proposed a new dataset for MAMS, which was named ACSA-V2. In addition, we modified the compression function based on the CapsNet model so that the model could converge faster while improving accuracy. The experimental results show that the improved CapsNet-based model improves accuracy by about 5 points on the aspect-level sentiment analysis task compared to Jiang *et al.*'s CapsNet model.

Keywords

Capsule Networks, Aspect-Level, Sentiment Analysis, Dynamic Routing

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

基于方面的情感分析(ABSA)是一项广泛研究的细粒度语义层面的情感分析任务,旨在识别句子中特定方面的情感极性。它包含两个子任务:基于给定目标的情感分析(Target Based Sentiment Analysis, ATSA)和基于给定方面类别的情感分析(Aspect Category Sentiment Analysis, ACSA) [1]。本文主要研究基于给定方面类别的情感分析,这里的方面类别是提前定义好的。例如,在:“Our waitress was a bit cold but she was still attentive and cakes came quickly.”这句话中,有两个目标的术语“waitress”和“cakes”,分别对应“staff”类别和“food”类别,它们各自的情感极性为:消极情绪和积极情绪。

胶囊网络(capsule network, CapsNet)是近些年新提出的一种网络模型,最先用于图像处理方向。它通过提取向量形式的特征,引入胶囊网络来改善 CNN 和 RNN 模型不能有效编码不同特征的空间位置关系的表达限制。Wang、Sun 和 Han (2018) [2]提出了情感分析中的胶囊网络。他们介绍了基于递归神经网络(RNN)的胶囊模型,每个胶囊不仅能够预测其指定情绪的概率,而且能够重构输入实例表示,胶囊 N 的数量与要建模的情绪类别的数量相同,每个类别对应一个情绪类别。Wang 等人(2019) [3]修改了 RNN 胶囊以适应 ABSA 任务,他们提出了方面级情感胶囊模型(AS capsule),该模型能够以联合方式同时执行方面检测和情感分类。为了解决缺乏方面级标记数据的问题,Chen 和 Qian (2019) [4]提出了一种传输胶囊网络(TransCap)模型,用于将文档级知识传输到方面级情感分类。Du、Sun 和 Wang (2019) [5]提出了胶囊网络的另一种修改,所提出的模型将胶囊网络与注意力机制相结合。Jiang、Chen 和 Xu (2019) [6]为了解决方面级情感分析退化为句子级情感分析提出了 CapNet 网络,在先前的工作上加入了方面词嵌入和残差连接。

我们分析了 Jiang、Chen 和 Xu (2019) [6]的工作,发现其提出 MAMS 中的 ACSA 数据集各个类别分布不平衡导致模型泛化能力弱,且原胶囊网络模型收敛速度较慢。为了解决上述问题,我们提出了一个简单的方面级情感分析胶囊网络模型 ACSA-CapNet。我们的主要工作是:第 1 步采用了同义替换、随机删除等文本增强技术将 ACSA 数据集进行了平衡,并提出了 ACSA-V2 多方面多极性数据集;第 2 步在 Jiang 等人的模型基础上加了 1 层卷积层并修改了原有的挤压函数,使得模型更好地提取不同的特征。实验结果表明,在 ACSA-V2 数据集上,所提出的方法比 Jiang 等人的工作高出约 5 个百分点。

2. ACSA-CapNet 模型

2.1. 整体框架概览

胶囊网络最先开始用在图像上, Jiang 等[6]人第一次将胶囊网络用于细粒度情感分析上, 其模型的泛化能力明显优于卷积神经网络和循环神经网络。

该模型主要由 4 部分组成: 第 1 部分是一个词嵌入层, 通过词嵌入将输入句子 S 转化为向量形式 E , 此外, 将方面嵌入 a 与每个单词嵌入在 $S: E_i^{sa} = [E_i; a]$ 中连接起来, 得到方面感知句子嵌入 E^{sa} ; 第 2 部分为编码层, 使用卷积神经网络的变体双向门控循环单元(BiGRU) [7]和残差网络[8]将体态感知句子表示 E^{sa} 转换为上下文化表示 $H = [h_1, \dots, h_n]$; 第 3 部分为初级胶囊层, 我们通过线性变换和挤压激活获得初级胶囊 P 和方面胶囊 c ; 第 4 部分为类别胶囊层, 基于归一化权重和路由权重, 得到类别胶囊 V , 使用边际损失作为基于方面的情感分类的损失函数, 最后通过 Softmax 函数生成类别胶囊及其对每个类的概率。

本文在 Jiang 等[6]的模型基础上又增加了 1 层卷积层, 用来再次提取不同位置上的特征, 使得特征提取更加地细致; 此外, 还对挤压函数进行了改进, 使得模型收敛速度更快。基于改进胶囊网络的文本分类模型如图 1 所示。

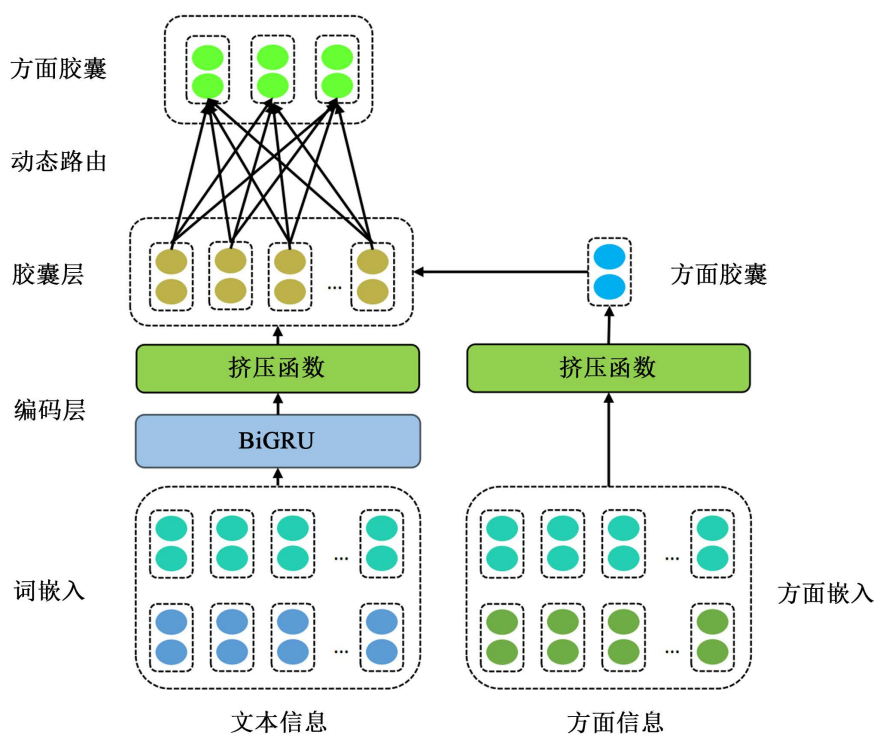


Figure 1. ACSA-CapNet Model Architecture

图 1. ACSA-CapNet 模型架构

2.2. 词嵌入层

在词嵌入层, 我们将句子 S 转换为单词嵌入 E , 对于 ACSA 任务, 由于预期类别嵌入 a 在训练期间被随机初始化和学习。我们通过将方面嵌入 a 与每个单词嵌入在 $S: E_i^{sa} = [E_i; a]$ 中连接起来, 得到方面感知句子嵌入 E^{sa} 。

2.3. 编码层

我们通过双向门控循环单元(BiGRU)和残差连接将方面感知句子表示 E^{sa} 转换为上下文化表示 $H = [h_1, \dots, h_n]$ 。

$$H = BiGRU(E^{sa}) + E^{sa} \quad (1)$$

2.4. 胶囊网络层

在初级胶囊层中, 我们通过线性变换和挤压激活函数获得初级胶囊 $P = [p_1, \dots, p_n]$ 和方面胶囊 c 。

$$P_i = squash(W^p h_i + b^p) \quad (2)$$

$$c = squash(W^a + b^a) \quad (3)$$

这里的 W^p, b^p, W^a, b^a 都是模型的学习参数。原始胶囊网络的挤压函数如等式(4)所示, 本文受沈等人 [9] [10] [11] 的影响, 挤压函数定义见等式(5)所示。

$$squash(s) = \frac{\|s\|^2}{1 + \|s\|^2} \frac{s}{\|s\|} \quad (4)$$

$$squash(s) = \frac{\|s\|^2}{0.5 + \|s\|^2} \frac{s}{\|s\|} \quad (5)$$

由于句子的长度可变, 发送到上层胶囊的初级胶囊的数量因句子而异, 导致训练过程不稳定。长度较长的句子会使挤压激活函数达到饱和, 导致所有类别的高置信度; 而相比之下, 长度十分短的句子将导致所有类别的置信度低。为了缓解这个问题, 我们提出了方面感知归一化, 利用方面胶囊来选择重要的主要胶囊, 并通过以下方式归一化主要胶囊权重 u , 这里的 W^n 是模型的可学习参数。

$$u_i = \frac{\exp(p_i W^n c)}{\sum_{j=1}^n \exp(p_j W^n c)} \quad (6)$$

由于路由的迭代过程, 原始动态路由机制[12]存在训练效率低下的问题。并且没有上层信息用于指导路由过程, 这使得动态路由就像一个自我引导的过程一样工作。

我们设计了一种胶囊引导路由机制, 它利用关于情感类别的先验知识来有效地指导路由过程, 而不是在路由过程中计算主胶囊和类别胶囊之间的耦合系数。具体来说, 我们使用一组情绪胶囊来存储关于情绪类别的先验知识。让 $G \in R^{C \times d}$ 是情感矩阵, 用情感词的平均嵌入来初始化。 C 是情绪类别的数量, d 是情绪嵌入的维度。我们可以通过对情绪矩阵应用挤压激活来获得情绪胶囊 $Z = [z_1, \dots, z_c]$, 并通过计算主胶囊和情绪胶囊之间的相似性来计算路由权重 w :

$$z_i = squash(G_i) \quad (7)$$

$$w_{ij} = \frac{\exp(p_i W^r z_j)}{\sum_{k=1}^n \exp(p_i W^r z_k)} \quad (8)$$

2.5. 类别胶囊层

基于归一化权重和路由权重, 最终类别胶囊 $V = [v_1, \dots, v_c]$ 可以计算为:

$$v_j = squash\left(s \sum_{i=1}^n w_{ij} u_i p_i\right) \quad (9)$$

其中 s 是可学习的缩放参数, 用于将连接权重缩放到合适的水平。我们使用边际损失作为基于方面的情感分类的损失函数:

$$L = \sum_{k=1}^c T_k \max(0, m^+ - \|v_k\|)^2 + \lambda(1 - T_k) \max(0, \|v_k\| - m^-)^2 \quad (10)$$

其中 $T_k = 1$ 当且仅当类别 k 存在时, m^+ 和 m^- 是边际超参数, λ 控制缺失类别的损失。在我们的实验中, m^+ , m^- , λ 分别设置为 0.9, 0.1, 0.6。

3. 实验结果与分析

3.1. 实验数据

本文在 MAMS-ACSA 数据集的基础上采用用随机替换、同义替换以及数据删除等文本数据增强方法对先前数据集进行了类别均衡, 并提出了一个新的数据 ACSA-V2, 并使新的数据集总量与原始数据集大小保持一致(见表 1 所示)。ACSA-V2 数据集中一共定义了八个粗略的方面类别: 食物、服务、员工、价格、环境、菜单、地点和其他。平衡后的数据集在各个类别的占比比较先前而言, food 方面数据集占比降低了约 50%, 而 price、menu 等原先占比较少的方面都相应的提高了约 10%。

Table 1. Statistics of MAMS-ACSA and ACSA-V2 dataset

表 1. MAMS-ACSA 和 ACSA-V2 数据集统计

数据集	类别	数量	类别	数量
MAMS-ACSA	food	2307	service	631
	staff	1383	menu	475
	miscellaneous	954	ambience	324
	place	694	price	322
ACSA-V2	food	1491	service	818
	staff	1068	menu	605
	miscellaneous	1075	ambience	538
	place	901	price	594

3.2. 基线模型

将我们提出的 ACSA-CapNet 模型与 Jiang、Chen 和 Xu 的 CapNet 模型进行比较。为了保证实验的准确性, 我们使用同样的设置对于 ACSA-CapNet 模型和 CapNet 模型。

3.3. 实验细节

实验中, 我们使用由 GloVe (Pennington 等人, 2014) [13] 预先训练的 300 维单词向量来初始化胶囊网络模型的单词嵌入向量。胶囊大小设置为 300。CapsNet 的批次大小分别设置为 64。我们使用 Adam 优化器来训练我们的模型。ACSA-CapsNet 学习率设置为 0.0003。我们运行所有模型 5 次, 每次迭代 20 轮, 并报告测试数据集的平均结果。实验参数设置见表 2。

3.4. 结果与分析

实验结果见表 3。从表 3 中我们得出以下结论: 我们提出的 ACSA-CapNet 模型在 MAMS-ACSA 和

ACSA-V2 数据集都优于 Jiang、Chen 和 Xu (2019)的 CapNet 模型。实验结果证明了我们提出的 ACSA-V2 数据集提高了模型的泛化能力。

Table 2. Experimental parameter setting
表 2. 实验参数设置

参数名称	参数含义	参数值
learning_rate	学习率	0.0003
embed_size	词向量维度	300
dropout	随机失活	0.5
num_layers	模型层数	3
capsule_size	胶囊数	300
batch_size	批训练样本数	64
num_epochs	迭代次数	20

Table 3. Experimental results on MAMS-ACSA and ACSA-v2 for ACSA subtask.
表 3. MAMS-ACSA 和 ACSA-V2 在 ACSA 子任务上的实验结果

模型方法	MAMS-ACSA	ACSA-V2
CapsNet	73.986	74.212
ACSA-CapNet	74.467	74.944
CapsNet+SQ	74.142	75.108
ACSA-CapNet-SQ	74.193	74.532

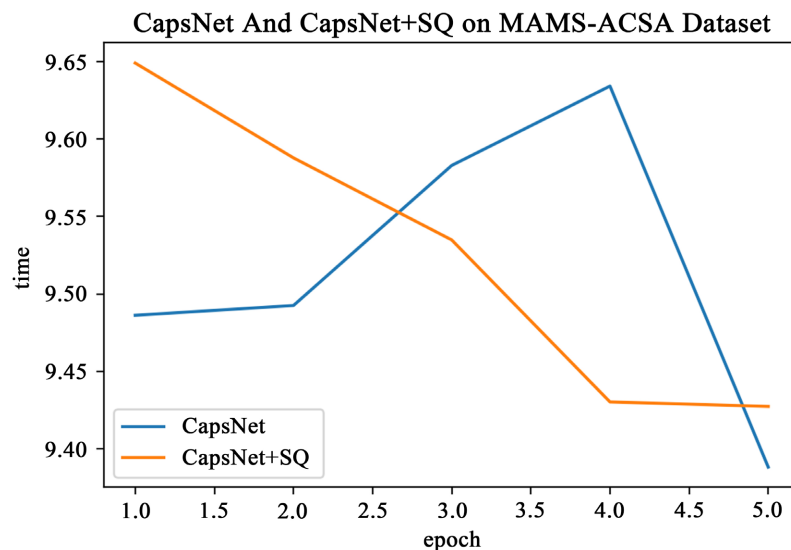


Figure 2. Time consumption of 20 epochs in 5 rounds of CapsNet and CapsNet + SQ
图 2. CapsNet 和 CapsNet + SQ 在 5 轮 20 个 epoch 的耗时情况

3.5. 消融实验

为了分析所提出的 ACSA-CapNet 模型效果的有效性, 我们进行了消融实验, 在 ACSA-CapNet 和 CapsNet 模型通过挤压函数的改变和模型层数的改变, 从而产生 CapsNet + SQ (CapsNet 改变了挤压函数)

和 ACSA-CapNet-SQ (ACSA-CapNet 只增加了模型层数)。从表 3 (最后两行)可以看出, 通过改变挤压函数, Jiang 等人提出的 CapsNet 模型较原先相比, 准确率提高约 2 个百分点。从图 2 可以看出, 当我们改变 CapsNet 模型的挤压函数后, 可以看出耗时情况与没有改变时不相上下, 但准确率较之前得到了提高。从图 3 可以看出, 当对我们提出的 ACSA-CapNet 模型进行挤压函数的消融实验后, 发现模型的收敛速度与准确率都有轻微的下降。

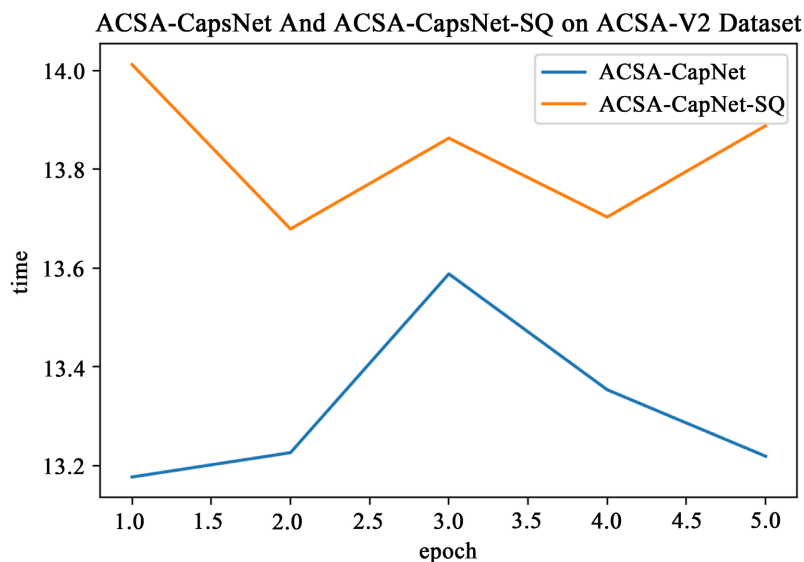


Figure 3. Time consumption of 20 epochs in 5 rounds of ACSA-CapNet and ACSA-CapNet-SQ

图 3. ACSA-CapNet 和 ACSA-CapNet-SQ 在 5 轮 20 个 epoch 的耗时情况

通过消融实验, 虽然我们提出的 ACSA-CapNet-SQ 去掉挤压函数后相较 ACSA-CapNet 模型在 MAMA-ACSA 和 ACSA-V2 两个数据集上准确率都有所下降, 但也侧面证明了压缩函数对模型准确率的影响。

4. 结语

在本文中, 我们提出了 ACSA-V2, 一个用于基于方面的情绪分析的挑战数据集, 其中每个句子包含具有不同情绪极性的多个方面。所提出的 ACSA-V2 数据集不仅可以防止方面级情感分类退化为句子级情感分类, 还提高了模型的泛化能力, 这可能会推动基于方面的情感分析的研究。此外, 我们改进了 Jiang 等人的 CapsNet 胶囊网络的挤压激活函数, 基于此提出了一个新的方面情感分析模型 ACSA-CapNet。实验结果证明, 其性能优于原有模型方法。

虽然本文提出的 ACSA-CapNet 模型可有效地提高方面级文本情感分析的准确性, 但在实验中发现编码层使用的 BiGRU 模型所耗费的时间代价较高, 随着 Transformer 模型的快速发展, 我们下一步计划使用 Transformer 模型代替 BiGRU 模型, 尝试去解决时间代价高的问题。

参考文献

- [1] Tang, D., Qin, B. and Liu, T. (2015) Document Modelling with Gated Recurrent Neural Network for Sentiment Classification. *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2015)*, Lisbon, September 2015, 1422-1432. <https://doi.org/10.18653/v1/D15-1167>

-
- [2] Wang, Y., Sun, A. and Han, J. (2018) Sentiment Analysis by Capsules. *Proceedings of the 2018 World Wide Web Conference*, Lyon, 23-27 April 2018, 1165-1174. <https://doi.org/10.1145/3178876.3186015>
- [3] Wang, Y., Sun, A., Huang, M., Zhu, X., et al. (2019) Aspect-Level Sentiment Analysis Using AS-Capsules. *International World Wide Web Conference Committee*, San Francisco, 13-17 May 2019, 2033-2044. <https://doi.org/10.1145/3308558.3313750>
- [4] Chen, Z. and Qian, T. (2019) Transfer Capsule Network for Aspect Level Sentiment Classification. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, Florence, 28 July-2 August 2019, 547-556. <https://doi.org/10.18653/v1/P19-1052>
- [5] Du, C., Sun, H., et al. (2019) Capsule Network with Interactive Attention for Aspect-Level Sentiment Classification. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*, Hong Kong, 3-7 November 2019, 5488-5498. <https://aclanthology.org/D19-1551> <https://doi.org/10.18653/v1/D19-1551>
- [6] Jiang, Q.N., Chen, L., Xu, R.F., Ao, X. and Yang, M. (2019) A Challenge Dataset and Effective Models for Aspect-Based Sentiment Analysis. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, Hong Kong, 3-7 November 2019, 6280-6285. <https://doi.org/10.18653/v1/D19-1654>
- [7] Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H. and Bengio, Y. (2014) Learning Phrase Representations Using RNN Encoder-Decoder for Statistical Machine Translation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Doha, October 2014, 1724-1734. <https://doi.org/10.3115/v1/D14-1179>
- [8] He, K.M., Zhang, X.Y., Ren, S.Q. and Sun, J. (2016) Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 770-778. <https://doi.org/10.1109/CVPR.2016.90>
- [9] 沈炜域. 基于自注意力与动态路由的文本建模方法[J]. 软件导刊, 2019, 18(1): 56-60+64.
- [10] 应伟志, 于青. 基于 BGRU-CapsNet 的情感分析算法研究[J]. 天津理工大学学报, 2021, 37(5): 7-12.
- [11] 尹春勇, 何苗. 基于改进胶囊网络的文本分类[J]. 计算机应用, 2020, 40(9): 2525-2530.
- [12] Sabour, S., Frosst, N. and Hinton, G.E. (2017) Dynamic Routing between Capsules. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Long Beach, 4-9 December 2017, 3859-3869. <https://dl.acm.org/doi/10.5555/3294996.3295142>
- [13] Pennington, J., Socher, R. and Manning, C.D. (2014) GloVe: Global Vectors for Word Representation. *Proceedings of the Conference on Empirical Methods on Natural Language Processing (EMNLP 2014)*, Doha, October 2014, 1532-1543. <https://doi.org/10.3115/v1/D14-1162>