在人物画像相关研究中，包括性别预测、年龄预测、职业预测、行为分析、立场分析、位置分析，其中立场分析研究比较热。立场分析是用于提取用户对商品、事件、人物等目标主体的观点立场。

**问题描述：**分析Twitter发文数据，判断用户针对“无神论”、“女性运动”、“气候变化”、“堕胎合法化”、“希拉里”、“川普”等目标主体的观点立场。

**数据集：**Twitter数据4870条，具体如下表：

|  |  |
| --- | --- |
| **目标主体** | **数量** |
| 无神论 | 733 |
| 女性运动 | 949 |
| 气候变化 | 564 |
| 堕胎合法化 | 933 |
| 希拉里 | 984 |
| 川普 | 707 |

**方法综述：**相关的方法可以分为深度学习、机器学习方法。其中，机器学习的方法在识别效果上明显不如深度学习方法。目前对深度学习的相关方法进行了总结，具体如下：

文献[1]认为很多方法没有考虑用户全部推文暗含的主观意识，提出使用半监督学习的方法预测用户的主观情感，并将其添加到立场识别的过程中，但效果并不理想，应该是和主观意识的预测准确度不高有关。

文献[2]将层级的注意力机制加入到LSTM中并考虑了多种语言学的注意力机制，实验表明立场识别效果有所提升。

文献[3]使用目标信息注意力机制，将目标信息添加到LSTM模型中，使得模型对目标的立场信息更加敏感，识别效果得到了提升。文献[4]、[5]都与文献[3]相似，文献[4]将目标信息添加到GRU模型中，识别效果同样得到了提升。文献[5]将目标信息添加到GRU-CNN模型中。

文献[6]提出将数据集根据具体情况（例如：目标立场对立）来抽取多个子数据集，分别对各个数据子集进行训练，得到不同的预测模型。最后，对于Twitter发文内容立场检测时，同时考虑多个模型的输出结果，进行投票选出最终的观点立场，提升了立场识别的准确率。

文献[7]在词典粒度上进行了细分，分别训练单词级别和字符级别的CNN模型，并综合考虑这两个模型的分析结果，该方法对立场识别的准确率有较好的提升。

**综上所述**，尽管目前对识别准确率有提升的方法各不相同，但可以分为基于模型的提升和基于数据的提升。**基于模型的提升包括：**注意力机制、层级模型、多模型融合等。目前很多经典的和热门的模型被用于该类问题中，模型创新相对来说难度较大。**基于数据提升的方法包括：**对数据集进行分类提取，从而扩大可用数据集的数量；引入目标信息等外来数据。从基于数据提升的方法和数据集的分布可以看出，每一类的数据量较小，阻碍了模型识别的能力，需要扩大数据集或者引入外来数据辅助模型进行立场识别。

参考文献：

1. Benton A, Dredze M. Using author embeddings to improve Tweet stance classification[C]//Proceedings of the 2018 EMNLP Workshop W-NUT: The 4th Workshop on Noisy User-generated Text. 2018: 184-194.
2. Sun Q, Wang Z, Zhu Q, et al. Stance detection with hierarchical attention network[C]//Proceedings of the 27th International Conference on Computational Linguistics. 2018: 2399-2409.
3. Du J, Xu R, He Y, et al. Stance classification with target-specific neural attention networks[C]. International Joint Conferences on Artificial Intelligence, 2017.
4. Wei P, Mao W, Zeng D. A target-guided neural memory model for stance detection in Twitter[C]//2018 International Joint Conference on Neural Networks (IJCNN). IEEE, 2018: 1-8.
5. Zhou Y, Cristea A I, Shi L. Connecting targets to tweets: Semantic attention-based model for target-specific stance detection[C]//International Conference on Web Information Systems Engineering. Springer, Cham, 2017: 18-32.
6. Wei W, Zhang X, Liu X, et al. pkudblab at semeval-2016 task 6: A specific convolutional neural network system for effective stance detection[C]//Proceedings of the 10th international workshop on semantic evaluation (SemEval-2016). 2016: 384-388.
7. Vijayaraghavan P, Sysoev I, Vosoughi S, et al. Deepstance at semeval-2016 task 6: Detecting stance in tweets using character and word-level cnns[J]. arXiv preprint arXiv:1606.05694, 2016.