采样方法的发展

SMOTE 对少数类随机插值(Chawla et al., 2002) (Lee, 2000)

Ｂｏｒｄｅｒｌｉｎｅ-ＳＭＯＴＥ (Bunkhumpornpat et al., 2009) (于化龙 et al., 2012) (Sanabila et al., 2017)

为少数类样本加入随机噪声(Lee, 2000)

Kubat等人将稀有类和大类交叉分布区域内的样本重新标定成稀有类样本 ,以降低不平衡度(Kubat et al., 1997)

单边采样 去除噪声和冗余及边界样本 (Kubat and Matwin, 1997)

主动欠采样 删去远离决策面的样本点 (杨泽平, 2015)

混合采样 (Raskutti and Kowalczyk, 2004)

变分自编码器 (Kingma and Welling, 2013)

生成式对抗网络 (Goodfellow et al., 2014)

一类学习 (one2class learning) [ 42 ]也被用于处理  
不平衡问题. 当样本数量不平衡时 ,并且当特征空间  
中混杂有大量噪音特征时 ,基于学习单一稀有类样  
本的产生式模型 ,相比于学习两类问题的判别式模  
型具有更好的性能 [ 19 ]

Bunkhumpornpat, C., Sinapiromsaran, K., and Lursinsap, C. (2009). Safe-Level-SMOTE: Safe-Level-Synthetic Minority Over-Sampling TEchnique for Handling the Class Imbalanced Problem. *5476*, 475–482.

Chawla, N.V., Bowyer, K.W., Hall, L.O., and Kegelmeyer, W.P. (2002). SMOTE: synthetic minority over-sampling technique. Journal of Artificial Intelligence Research *16*, 321–357.

Goodfellow, I., Pougetabadie, J., Mirza, M., Xu, B., Wardefarley, D., Ozair, S., Courville, A., and Bengio, Y. (2014). Generative Adversarial Nets. Advances in Neural Information Processing Systems 2672–2680.

Kingma, D.P., and Welling, M. (2013). Auto-Encoding Variational Bayes.

Kubat, M., and Matwin, S. (1997). Addressing the Curse of Imbalanced Training Sets: One-Sided Selection. In International Conference on Machine Learning, pp. 179–186.

Kubat, M., Holte, R., and Matwin, S. (1997). Learning when negative examples abound (Springer Berlin Heidelberg).

Lee, S.S. (2000). Noisy replication in skewed binary classification. Computational Statistics & Data Analysis *34*, 165–191.

Raskutti, B., and Kowalczyk, A. (2004). Extreme re-balancing for SVMs: a case study. Acm Sigkdd Explorations Newsletter *6*, 60–69.

Sanabila, H.R., Kusuma, I., and Jatmiko, W. (2017). Generative oversampling method (GenOMe) for imbalanced data on apnea detection using ECG data. In International Conference on Advanced Computer Science and Information Systems, pp. 572–579.

于化龙, 高尚, 赵靖, and 秦斌 (2012). 基于过采样技术和随机森林的不平衡微阵列数据分类方法研究. 计算机科学 *39*, 190–194.

杨泽平 (2015). 基于神经网络的不平衡数据分类方法研究. 华东理工大学.