EWGAN: Entropy-Based Wasserstein GAN for Imbalanced Learning

Jinfu Ren, Yang Liu, Jiming Liu
Department of Computer Science
Hong Kong Baptist University
Kowloon Tong, Hong Kong
{jinfuren,csygliu,jiming}@comp.hkbu.edu.hk

Abstract

In this paper, we propose a novel oversampling strategy dubbed Entropy-based Wasserstein Generative Adversarial Network (EWGAN) to generate data samples for minority classes in imbalanced learning. First, we construct an entropy-weighted label vector for each class to characterize the data imbalance in different classes. Then we concatenate this entropy-weighted label vector with the original feature vector of each data sample, and feed it into the WGAN model to train the generator. After the generator is trained, we concatenate the entropy-weighted label vector with random noise feature vectors, and feed them into the generator to generate data samples for minority classes. Experimental results on two benchmark datasets show that the samples generated by the proposed oversampling strategy can help to improve the classification performance when the data are highly imbalanced. Furthermore, the proposed strategy outperforms other state-of-the-art oversampling algorithms in terms of the classification accuracy.

Introduction

Oversampling is an effective strategy in sampling methods (Chawla et al. 2002; Douzas and Bacao 2018) for imbalanced learning. It aims to generate data for minority classes to overcome the data imbalance problem. Representative oversampling methods include the simplest data replication, the classical synthetic minority oversampling technique (SMOTE) (Chawla et al. 2002) and its variants (Last, Douzas, and Baçao 2017), etc.

Recently, the generative adversarial network (GAN) (Goodfellow et al. 2014) in deep learning has received much attention as it showed good potential in generating artificial data which resemble the real-world data such as images. Motivated by the success of GAN in unsupervised learning tasks, its idea has been transplanted to supervised learning, including imbalanced learning. Douzas and Bacao (2018) employed the conditional GAN (cGAN) to incorporate label vectors into feature vectors for training, and used the trained generator to generate samples for minority classes. Although this method has demonstrated the state-of-the-art performance in data generation, a shortcoming is that the label vectors used in the method are one-hot vectors, which may not be able to characterize the imbalance between different classes. Moreover, a potential limitation of cGAN is the learning instability, which is caused by the Kullback-Leibler divergence (KL divergence) used in GANs (Arjovsky, Chintala, and Bottou 2017).

To address the aforementioned limitations while inheriting GANs' powerful ability in data generation, we propose a novel oversampling strategy called Entropy-based Wasserstein GAN (EWGAN), which constructs an entropy-weighted label vector for each class to characterize the data imbalance in different classes and trains the generator using a WGAN model (Arjovsky, Chintala, and Bottou 2017), overcoming the learning instability of GAN by using the Wasserstein distance to replace KL divergence.

Entropy-based Wasserstein Generative Adversarial Network

Given the training set $\{\mathbf{x}_1, \cdots, \mathbf{x}_n\}$ with n data samples, where $\mathbf{x}_i \in \mathbb{R}^d$ denotes the i-th sample ($i = 1, \cdots, n$) and d denotes the dimension of the feature vector. The label vector associated with $\mathbf{x}_i$ is denoted as $\mathbf{y}_i$, where $\mathbf{y}_i = [1, 0]^T$ indicates that $\mathbf{x}_i$ belongs to the majority class while $\mathbf{y}_i = [0, 1]^T$ indicates that $\mathbf{x}_i$ belongs to the minority class. The objective of EWGAN is generating samples $\mathbf{x}'_1, \cdots, \mathbf{x}'_p$ for the minority class, by learning the data distribution from the given dataset, where $p = n_{major} - n_{minor}$ denotes the number of samples to be generated, and n_{major} and n_{minor} denote the number of samples in the majority class and that in the minority class, respectively.

To describe the imbalance between different classes, we construct an entropy-weighted label vector for each class. First, we define I_i ($i = 1, \cdots, n$) for each sample to indicates the importance of class $\mathbf{y}_i$. Inspired by the focal loss proposed for object detection (Lin et al. 2017), we introduce the following entropy-based formulation to calculate I_i:

$$I_i = -\frac{(1 - p_i)^2}{p_i} \log(p_i), \tag{1}$$

where p_i denotes the proportion of class $\mathbf{y}_i$ in the dataset. Obviously, small proportion indicates high importance. Accordingly, the entropy-weighted label vector is constructed as $\hat{\mathbf{y}}_i = I_i \cdot \mathbf{y}_i$ ($i = 1, \cdots, n$).

We then concatenate $\mathbf{x}_i$ and $\hat{\mathbf{y}}_i$ as $\mathbf{u}_i = [\mathbf{x}_i^T, \hat{\mathbf{y}}_i^T]^T$, and concatenate a random feature vector $\mathbf{r}'_i$ and $\hat{\mathbf{y}}_i$ as $\mathbf{z}_i =$

Algorithm 1 Entropy-based Wasserstein GAN (EWGAN)

Input: The training dataset $\{(\mathbf{x}_1, \mathbf{y}_1), \cdots, (\mathbf{x}_n, \mathbf{y}_n)\}$.
Output: The generated data $\mathbf{x}'_1, \cdots, \mathbf{x}'_p$ for minority class.

```
for i ← 1, ⋯ , n do
    I_i ← −((1−p_i)^2 / p_i) log(p_i);
    u_i ← [x_i^T, I_i · y_i^T]^T;
    z_i ← [r_i^T, I_i · y_i^T]^T;
end for
for t ← 1, ⋯ , T_outer do
    for t ← 1, ⋯ , T_inner do
        Sample a batch from {u_1, ⋯ , u_n};
        Sample a batch from {z_1, ⋯ , z_n};
        g_w ← ▽_w[(1/m) σ_{i=1}^m f_w(u_i) − (1/m) Σ_{i=1}^m f_w(g_θ(z_i))];
        w ← w + αRMSProp(w, g_w);
        w ← clip(w, −c, c);
    end for
    Sample a batch from {z_1, ⋯ , z_n};
    g_θ ← − ▽_θ (1/m) Σ_{i=1}^m f_w(g_θ(z_i));
    θ ← θ − αRMSProp(θ, g_θ);
end for
for j ← 1, ⋯ , p do
    x'_j ← g_θ([r'_j^T, I_j · y'_j^T]^T);
end for
```

$[\mathbf{r}_i^T, \hat{\mathbf{y}}_i^T]^T$. We feed them into the EWGAN model to train the generator and discriminator. The loss functions for generator and discriminator are given as follows:

$$-\mathbb{E}_{\mathbf{r}\sim P_{\mathbf{r}}}[f_w(g_\theta(\mathbf{z}))], \tag{2}$$

$$\mathbb{E}_{\mathbf{r}\sim P_{\mathbf{r}}}[f_w(g_\theta(\mathbf{z}))] - \mathbb{E}_{\mathbf{x}\sim P_{\mathbf{x}}}[f_w(\mathbf{u})], \tag{3}$$

where $P_{\mathbf{r}}$ and $P_{\mathbf{x}}$ denote the distribution of generated samples and that of the existing samples in a specific class, respectively, $g_\theta(\cdot)$ and $f_w(\cdot)$ denote the generator and discriminator, respectively, and θ and w are the parameters of the generator and discriminator, respectively. Thus, the target of EWGAN is to maximize the following objective function:

$$\arg\max_{\theta,w} \mathbb{E}_{\mathbf{r}\sim P_{\mathbf{r}}}[f_w(g_\theta(\mathbf{z}))] - \mathbb{E}_{\mathbf{x}\sim P_{\mathbf{x}}}[f_w(\mathbf{u})]. \tag{4}$$

Once the model is trained, the generator will be used to generate the data samples for the minority class. The details of EWGAN is described in Algorithm 1.

Experiments

We use two benchmark imbalanced datasets, Vowel0 and Page-blocks0, for performance evaluation (https://sci2s.ugr.es/keel/imbalanced.php). The size, dimension, and imbalance ratio of these datasets are listed in Table 1. To validate the effectiveness of EWGAN, we compare it with three state-of-the-art algorithms, Kmeans-SMOTE (Last, Douzas, and Baçao 2017), WGAN (Arjovsky, Chintala, and Bottou 2017), and cGAN (Douzas and Bacao 2018). The result without oversampling is used as the baseline.

For the Vowel0 dataset, we split it into the training set with 20 positive samples and 828 negative samples and the test set with 70 positive samples and 70 negative samples. For the Page-blocks0 dataset, we split it into the training set with 259 positive samples and 4663 negative samples and the test set with 250 positive samples and 250 negative samples. For the proposed EWGAN, we set $T_{outer} = 40$ and $T_{inner} = 5$. After data generation, the Support Vector Machine (SVM) is utilized for classification. We repeat the experiment for 10 times on randomly selected training/test sets and report the average results and standard deviations.

Table 1: Statistics of the benchmark datasets.

	Vowel0	Page-blocks0
Size	988	5472
Dimension	13	10
Imbalance Ratio	9.98	8.79

Table 2: The classification accuracy of SVM with no oversampling as well as oversampling via Kmeans-SMOTE, WGAN, cGAN, and the proposed EWGAN on Vowel0 and Page-blocks0 datasets. The best performances are highlighted in bold.

	Vowel0	Page-blocks0
No Oversampling	0.5645 ± 0.0001	0.8023 ± 0.0025
Kmeans-SMOTE	0.6393 ± 0.0003	0.7868 ± 0.0014
WGAN	0.5638 ± 0.0007	0.7014 ± 0.0006
cGAN	0.8135 ± 0.0003	0.7567 ± 0.0004
EWGAN	$\mathbf{0.8385 \pm 0.0006}$	$\mathbf{0.8322 \pm 0.0007}$

The performances of all the methods are shown in Table 2. Note that WGAN does not perform well as is unsupervised, and thus the data generated by WGAN cannot take the label information into consideration. With the ability of data generation inherited from WGAN and the incorporation of entropy-weighted label vectors in the learning process, EWGAN outperforms other algorithms on both datasets.

Conclusion

In this paper, we proposed EWGAN, aiming at generating data for the minority class in imbalanced learning. The proposed method achieved good performance on Vowel0 and Page-blocks0 datasets. In our future work, we will extend the proposed model to the multi-class version and validate its performance on more large-scale datasets.

References

Arjovsky, M.; Chintala, S.; and Bottou, L. 2017. Wasserstein GAN. *ArXiv e-prints*.

Chawla, N. V.; Bowyer, K. W.; Hall, L. O.; and Kegelmeyer, W. P. 2002. Smote: Synthetic minority over-sampling technique. *J. Artif. Intell. Res.* 16:321–357.

Douzas, G., and Bacao, F. 2018. Effective data generation for imbalanced learning using conditional generative adversarial networks. *Expert Syst. Appl.* 91:464–471.

Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; and Bengio, Y. 2014. Generative adversarial nets. In *NIPS 27*. 2672–2680.

Last, F.; Douzas, G.; and Baçao, F. 2017. Oversampling for imbalanced learning based on k-means and smote. *CoRR* abs/1711.00837.

Lin, T.; Goyal, P.; Girshick, R.; He, K.; and Dollar, P. 2017. Focal loss for dense object detection. In *Proc. 16th ICCV*, 2999–3007.