

机器学习导论

习题四

学号, 作者姓名, 邮箱

2017 年 4 月 27 日

1 [20pts] Reading Materials on CNN

卷积神经网络(Convolution Neural Network,简称CNN)是一类具有特殊结构的神经网络,在深度学习的发展中具有里程碑式的意义。其中,Hinton于2012年提出的AlexNet可以说是深度神经网络在计算机视觉问题上一次重大的突破。

关于AlexNet的具体技术细节总结在经典文章“ImageNet Classification with Deep Convolutional Neural Networks”, by Alex Krizhevsky, Ilya Sutskever and Geoffrey E. Hinton in NIPS'12, 目前已逾万次引用。在这篇文章中,它提出使用ReLU作为激活函数,并创新性地使用GPU对运算进行加速。请仔细阅读该论文,并回答下列问题(请用1-2句话简要回答每个小问题,中英文均可)。

- (a) [5pts] Describe your understanding of how ReLU helps its success? And, how do the GPUs help out?
- (b) [5pts] Using the average of predictions from several networks help reduce the error rates. Why?
- (c) [5pts] Where is the dropout technique applied? How does it help? And what is the cost of using dropout?
- (d) [5pts] How many parameters are there in AlexNet? Why the dataset size(1.2 million) is important for the success of AlexNet?

关于CNN,推荐阅读一份非常优秀的学习材料,由南京大学计算机系吴建鑫教授¹所编写的讲义Introduction to Convolutional Neural Networks²,本题目为此讲义的Exercise-5,已获得吴建鑫老师授权使用。

Solution. 此处用于写解答(中英文均可)

¹ 吴建鑫教授主页链接为cs.nju.edu.cn/wujx

² 由此链接可访问讲义<https://cs.nju.edu.cn/wujx/paper/CNN.pdf>

2 [20pts] Kernel Functions

(1) 试通过定义证明以下函数都是一个合法的核函数:

(i) [5pts] 多项式核: $\kappa(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i^T \mathbf{x}_j)^d$;

(ii) [10pts] 高斯核: $\kappa(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2})$, 其中 $\sigma > 0$.

(2) [5pts] 试证明 $\kappa(\mathbf{x}_i, \mathbf{x}_j) = \frac{1}{1 + e^{-\mathbf{x}_i^T \mathbf{x}_j}}$ 不是合法的核函数。

Proof. 此处用于写证明(中英文均可)

□

3 [25pts] SVM with Weighted Penalty

考虑标准的SVM优化问题如下(即课本公式(6.35)),

$$\begin{aligned} \min_{\mathbf{w}, b, \xi_i} \quad & \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^m \xi_i \\ \text{s.t.} \quad & y_i(\mathbf{w}^T \mathbf{x}_i + b) \geq 1 - \xi_i \\ & \xi_i \geq 0, i = 1, 2, \dots, m. \end{aligned} \tag{3.1}$$

注意到, 在(3.1)中, 对于正例和负例, 其在目标函数中分类错误的“惩罚”是相同的。在实际场景中, 很多时候正例和负例错分的“惩罚”代价是不同的, 比如考虑癌症诊断, 将一个确实患有癌症的人误分类为健康人, 以及将健康人误分类为患有癌症, 产生的错误影响以及代价不应该认为是等同的。

现在, 我们希望对负例分类错误的样本(即false positive)施加 $k > 0$ 倍于正例中被分错的样本的“惩罚”。对于此类场景下,

(1) [10pts] 请给出相应的SVM优化问题;

(2) [15pts] 请给出相应的对偶问题, 要求详细的推导步骤, 尤其是如KKT条件等。

Solution. 此处用于写解答(中英文均可)

4 [35pts] SVM in Practice - LIBSVM

支持向量机(Support Vector Machine, 简称SVM)是在工程和科研都非常常用的分类学习算法。有非常成熟的软件包实现了不同形式SVM的高效求解, 这里比较著名且常用的如LIBSVM³。

(1) [20pts] 调用库进行SVM的训练, 但是用你自己编写的预测函数作出预测。

(2) [10pts] 借助我们提供的可视化代码, 简要了解绘图工具的使用, 通过可视化增进对SVM各项参数的理解。详细编程题指南请参见链接: http://lamda.nju.edu.cn/ml2017/PS4/ML4_programming.html。

(3) [5pts] 在完成上述实践任务之后, 你对SVM及核函数技巧有什么新的认识吗? 请简要谈谈。

Solution. 此处用于写解答(中英文均可)

³LIBSVM主页可参见链接: <https://www.csie.ntu.edu.tw/~cjlin/libsvm/>