

Assignment #4 - SVM

- SVM (Support Vector Machines) for Classification (台大 [林智仁 教授](#))
- SVM工具套件 - LIBSVM
 - [LIBSVM](#), a support vector machines (SVM) library for classification
 - Download LIBSVM [zip file](#)
 - you need to install [python](#) and [gnuplot](#) if you use easy.py and grid.p you should modify "path" to these two package.
 - \$ python easy.py train.1 test.1
- 套件下載 & 函式說明

<http://dns2.asia.edu.tw/~jdwang/project/svm.htm>

你可能會遇到的專有名詞或參數設定...

- Kernal:

- linear: $K(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i^T \mathbf{x}_j$.
- polynomial: $K(\mathbf{x}_i, \mathbf{x}_j) = (\gamma \mathbf{x}_i^T \mathbf{x}_j + r)^d, \gamma > 0$.
- radial basis function (RBF): $K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2), \gamma > 0$.
- sigmoid: $K(\mathbf{x}_i, \mathbf{x}_j) = \tanh(\gamma \mathbf{x}_i^T \mathbf{x}_j + r)$.

RBF Kernel

- In general, the RBF kernel is a reasonable first choice.
- RBF kernel nonlinearly maps samples into a higher dimensional space so it, can handle the case when the relation between class labels and attributes is nonlinear.
- Furthermore, the **linear kernel** is a special case of RBF with some parameters ($C; \gamma$). In addition, **the sigmoid kernel** behaves like RBF for certain parameters.
- The second reason is the number of hyperparameters which in influences the complexity of model selection. The polynomial kernel has more hyperparameters than the RBF kernel.

Cross-validation and Grid-search

- There are two parameters for an RBF kernel: C and γ . It is not known beforehand.
- Model selection (parameter search) must be done.
- The goal is to identify good $(C; \gamma)$ by cross-validation and grid-search.
- Trying exponentially growing sequences of C and γ is a practical method to identify good parameters
 - For example, $C = 2^{-5}, 2^{-3}, \dots, 2^{15}; \gamma = 2^{-15}, 2^{-13}, \dots, 2^3$.

Scaling

- Linearly scaling each attribute to the range $[-1; +1]$ or $[0; 1]$.
- Of course we have to use the same method to scale both training and testing data.
- For example, suppose that we scaled the attribute of training data from $[-10; +10]$ to $[-1; +1]$. If the attribute of testing data lies in the range $[-11; +8]$, we must scale the testing data to $[-1.1; +0.8]$.
- See Appendix B in the following file for some real examples.
 - **A Practical Guide to Support Vector Classification**
 - <https://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf>

Unbalanced Data Handling

- Yes, there is a -w options. For example, if you use
 `> svm-train -s 0 -c 10 -w1 1 -w-1 5 data_file`
- The penalty for class "-1" is larger.
- Note that this -w option is for C-SVC only.
- **A Practical Guide to Support Vector Classification**
 - <https://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf>

執行步驟

- Many beginners use the following procedure now:
 1. Transform data to the format of an SVM package
 2. Randomly try a few kernels and parameters
 3. Test
- Recommend that beginners try the following procedure:
 1. Transform data to the format of an SVM package
 2. Conduct simple scaling on the data
 3. Consider the RBF kernel
 4. Use cross-validation to find the best parameter C and γ
 5. Use the best parameter C and γ to train the whole training set
 6. Test

作業要求

- RBF SVM

1. Use the linear model $y = 2x + \varepsilon$ with zero-mean Gaussian noise $\varepsilon \sim N(0, 1)$ to generate 500 data points with (equal spacing) $x \in [-100, 100]$.
2. 訓練一個RBF SVM
3. 利用5-fold cross validation，至少找三組C and γ 參數，挑一組最好的
4. 用上題找到的參數，比較scaling方法、或不用scaling，影響準確率的差異為何。

- 繳交期限: 4/25 11:59pm

- 繳交程式碼與報告，如同之前的規定一樣