Electronic Supplementary Material (ESI) for Analytical Methods.

Supplementary Materials

Ensemble-based support vector machine classifiers as an efficient tool for quality assessment of beef fillets from electronic nose data

Fady Mohareb[a, *], Olga Papadopoulou[b], Efstathios Panagou[b], George-John Nychas[b], Conrad Bessant[c]

S1. Naïve Bayes

Let $P(s_j)$ be the probability that classifier D_j predicts sample x to in class $s_j \in \Omega$. The conditional independence can therefore presented as follows:

$$P(s \mid \omega_k) = P(s_1, s_2, \dots, s_T \mid \omega_k) = \prod_{i=1}^{T} P(s_i|\omega_k)$$

The posterior probability can be obtained as follows:

$$P(\omega_k \mid s) = \frac{P(\omega_k)P(s|\omega_k)}{P(s)} = \frac{P(\omega_k)\prod_{i=1}^{T} P(s_i|\omega_k)}{P(s)}$$

Where $k = 1, \dots, C$.

P(s) is constant for each class, and can be therefore ignored. The support for class ω_k can be calculated as follows:

$$\mu_k(x) \propto P(\omega_k)\prod_{i=1}^{T} P(s_i|\omega_k)$$

For each classifier D_i, a $c \times c$ confusion matrix CM^i is generated based on labeling the training subset. The $(k,s)^{th}$ element of this matrix, $cm^i_{k,s}$ is the number of element classified as ω_s while their true class is ω_k. N_s is the total number of elements of Z from class ω_s. This can be represented as follows:

$$\mu_k(x) \propto \frac{1}{N_k^{T-1}}\prod_{i=1}^{T} cm^i_{k,s_i}$$

Where ${cm^i_{k,s_i}}/{N_k}$ is an estimate of the probability $P(\omega_k \mid s)$, and ${N_k}/{N}$ is an estimate of the prior probability for class ω_s.

The previous formula suffers from a drawback, that it when cm_{k,s_i}^{i} equals zero, it will automatically set $\mu_k(x)$ as zero, regardless to the rest of the estimates. That is why Titterington et al. (Titterington et al., 1981) proposed a modification to the estimate to overcome the null value for $\mu_k(x)$ as follows:

$$P(s|\omega_k) \propto \left\{ \prod_{i=1}^{T} \frac{cm_{k,s_i}^{i} + {}^{1}\!/_{c}}{N_k + 1} \right\}^{B}$$

Where $N_k$is the number of elements belonging to the class ω_k, and *B* is a constant (Kuncheva, 2004).

S2. Boosting using Adaboost.M1

- Let *S* be the training subset of *N* number of samples S: $S = [(x_i, y_i), i = 1, \ldots, N$ where x_i is the i^{th} sample, and y_i is the i^{th} label for sample x_i where $y_i \in \Omega, \Omega = \{\omega_1, \ldots, \omega_C\}$.
- Let *D* be the classifiers ensemble $A = \{h_1, \ldots, h_T\}$, Where *T* is the number of classifiers to train (*i.e.* number of iterations). A grid search is performed to identify the best classification parameters using the parameter ranges $C = [1,2,3, \ldots, 30]$ and $\gamma = [0.1, 0.2, 0.3, \ldots, 5]$. For each cycle within the grid search, a bootstrapped subset of S, S_t is developed using resampling with replacement. S_t is then split into $S_t training_{c,g}$ and $S_t testing_{c,g}$ .If $\varepsilon_t > 0.5$, h_t h_t, otherwise h_t to the ensemble.
- The classification error for *S* is initialised as follows:

$$D_1(i) = \frac{1}{N}. i = 1, \ldots, N;$$

- The error of f h_t is calculated such as $\varepsilon_t = \sum_{i:h_t(x_i) \neq y_i} D_t(i).$ The Adaboost Weight for h_t is then calculated as follows

$$\rho = \ln\left(\frac{1 - \varepsilon_t}{\varepsilon_t}\right)\frac{1}{2}$$

- The classification error is calculated as follows:

$$D_t : D_{t+1}(i) = D_t(i) * e^{-\rho c}$$

Where: $c = 1 \; if \; h_t(x_i) = y_i; c = -1 \; if \; h_t(x_i) \neq y_i$;

- Test the ensemble classification using weighted majority voting by calculating the total votes received by each class as follows:

$$V_j = \sum_{t:h_t(x) = \omega_j} \rho_t, j = 1, \ldots, C .$$

- Choose the class that receives the highest total vote as the final classification.

References:

Kuncheva, Ludmila I. (2004). *Combining pattern classifiers : methods and algorithms*. Hoboken, N.J. ; [Chichester]: John Wiley ;.

Titterington, D. M., Murray, G. D., Murray, L. S., Spiegelhalter, D. J., Skene, A. M., Habbema, J. D. F., & Gelpke, G. J. (1981). Comparison of Discrimination Techniques Applied to a Complex Data Set of Head Injured Patients. *Journal of the Royal Statistical Society. Series A (General), 144*(2). doi: citeulike-article-id:9414776