# Machine learning algorithms comparing experts vs nurses 2-class (agitated night – yes/no) classification problem

| **Evaluator** | **Experts** | | | **Nurses** | | |
| --- | --- | --- | --- | --- | --- | --- |
| **Method** | **Validation accuracy (%)** | **Tuning parameters** | **Options** | **Validation accuracy (%)** | **Tuning parameters** | **Options** |
| Naïve Bayes | 91.5 |  | Gaussian distribution for predictors | 8.2 |  | Gaussian distribution for predictors |
| *k*-Nearest Neighbour | 94.9 | *k* = 11 | Standardise predictors, Minkowski (cubic) distance (no weighting) | 6.3 | *k* = 11 | Standardise predictors, Euclidean distance (no weighting) |
| Decision tree | 94.9 | Maximum deviance reduction criterion, maximum number of splits at each node = 3 | Surrogate decision splits = off, prune = on, merge leaves = on, minimum leaf size = 1, minimum parent size = 2, number of variables to sample = all |  |  |  |
| Support Vector Machine (SMO) | 94.9 | Kernel = linear, Cost = 7.37 | Standardise predictors | 6.8 | Kernel = linear | Standardise predictors |
| Support Vector Machine (SMO) | 95.2 | Kernel = quadratic, Cost = 1.06 | Standardise predictors |  |  |  |
| Support Vector Machine (SMO) | 94.6 | Kernel = cubic, Cost = 10.53 | Standardise predictors |  |  |  |
| Support Vector Machine (SMO) | 95.2 | Kernel = Gaussian, kernel scale = 6, Cost = 4.22 | Standardise predictors |  |  |  |
| Support Vector Machine (LibSVM) |  |  |  |  |  |  |
| Neural network |  |  |  |  |  |  |
| Logistic regression |  |  |  |  |  |  |
| Discriminant analysis | 5.7 | Type = linear |  | 7.7 | Type = linear |  |
| ***Ensemble methods*** |  |  |  |  |  |  |
| AdaBoost-M1 | 94.6 | Learning cycles = 9 | Method = tree, learning rate = 0.1 | 8.2 | Learning cycles = 15 | Method = tree |
| GentleBoost | 94.0 | Learning cycles = 1 | Method = tree, learning rate = 0.1 | 6.8 | Learning cycles = 6 | Method = tree |
| RobustBoost | 94.0 | Learning cycles = 1 | Method = tree | 6.8 | Learning cycles = 3 | Method = tree |
| LPBoost | 91.2 | Learning cycles = 9 | Method = tree | 11.1 | Learning cycles = 23 | Method = tree |
| TotalBoost | 91.2 | Learning cycles = 45 | Method = tree | 25.8 | Learning cycles = 46 | Method = tree |
| RUSBoost | 92.0 | Learning cycles = 83 | Method = tree, learning rate = 0.1 | 26.7 | Learning cycles = 99 | Method = tree |
| LogitBoost | 94.3 | Learning cycles = 1 | Method = tree, learning rate = 0.1 |  | Learning cycles = | Method = tree, learning rate = 0.1 |
| Random subspace | 92.9 | Learning cycles = 7, Method = discriminant | Number of predictors to sample = 1 | 6.8 | Learning cycles = 5, Method = discriminant |  |
| Bagging | 94.9 | Learning cycles = 2 | Method = tree | 7.7 | Learning cycles = 3 | Method = tree |

Remarks:

10-fold cross validation;

SMO = Sequential Minimal Optimisation;

Decision tree was used for different ensemble algorithms as a weak learner, except for ‘subspace’ algorithm.

Comments:

* Overall, support vector machine (Gaussian or quadratic kernel) was found to be the best performing method in differentiating an agitated from a non-agitated night. The ensemble methods did not improve classification from the weak learner, *i.e.*, classification tree. Given only 2 predictors in the model (*i.e.*, frequency and duration), expectedly the number of learning cycles required for a bagging classifier to attain optimum performance were very few.
* Experts’ classification performed better than (or equal to) nurses’ classification except for naïve Bayes, which was not a very good classifier among the different methods applied.