BRAINSTORMING
Agent based Meta-learning Approach

Dariusz Plewczynski
ICM, Interdisciplinary Centre for Mathematical and Computational Modelling, University of Warsaw
Pawinskiego 5a Street, 02-106 Warsaw, Poland

Keywords: Machine learning, Ensemble methods, Agent based modelling, Meta-learning, Brainstorming.

Abstract: Brainstorming meta-learning approach is performed by a set of cognitive agents (CA), each implementing different machine learning (ML) algorithm, and/or trained using diverse subsets of available features describing input examples. The goal of the meta-learning procedure is providing a general and flexible classification meta-model for a given training data. In the first phase all agents, when trained using different features describing training objects, construct the ensemble of classification models independently. In the second step all solutions are gathered and the consensus is built between them by optimizing the voting weights for all agents. No early solution, given even by a generally low performing agent, is not discarded until the late phase of prediction, when comparing different learning models draws the final conclusion. The final phase, i.e. brainstorming tries to balance the generality of solution and the overall cognitive performance of all CAs. The classification meta-model is than used for predictions of the classification membership for given testing examples. The method was recently used in several ML applications in bioinformatics and chemoinformatics by the author.

1 INTRODUCTION

Although quite a few machine learning algorithms have been developed, such as fuzzy or crisp clustering, support vector machine (SVM), artificial neural network (ANN), or K-nearest neighbor (KNN), and many others classifiers, the way they operate the identification is basically individual. Yet, the proper approach usually takes into account the opinions from several experts rather than rely on only one when they are making critical decisions. Likewise, a sophisticated identifier should be trained by several different modes. This is the core idea of brainstorming, i.e. the consensus meta-learning algorithm that will be described in this manuscript, which applications were described previously for selected types of machine learning methods such as clustering (Can, T. et al., 2004), (Han, X., 2007), (Schulze-Kremer and King, 1992), (Vernikos and Parkhill, 2008), support vector machine (Arimoto et al., 2005), (Bhasin and Raghava, 2004), (Briem and Gunther, 2005), (Burton, J. et al., 2009), (Yao, X. Q. et al., 2008), (Abrusan, G. et al., 2009), (Hwang, S. et al., 2007), artificial neural networks (Yao and She, 2008), (Garg, P. et al., 2009), (Miller and Blom., 2009), or K-nearest neighbor (Arimoto et al., 2005), (Briem and Gunther, 2005), (Garg, P. et al., 2009), (Bindewald and Shapiro, 2006), including our own findings (Plewczynski and Ginalski, 2009), (Plewczynski, D. et al., 2007), (Plewczynski, D. et al., 2009a), (Plewczynski, D. et al., 2005), (Plewczynski, D. et al. 2009b), (Plewczynski, D. et al., 2006), Bagging and boosting are previously known meta-learning techniques had a wide array of applications as recapitulated in various manuscripts (Bruce, C. L. et al., 2007), (Islam, M. M. et al., 2008), (Plewczynski, D. et al., 2009c), (Schwenk and Bengio, 2000), (Serpen, G. et al., 2008), (Shrestha and Solomatine, 2006), (Peng, Y., 2006), (Wang, C. W., 2006), (Yang, J. Y. et al., 2008), The meta-learning procedure is here implemented using Agent Based Modelling framework developed recently for various applications in Life Sciences (Devillers, J. et al.), (Moore, D. et al., 2009), (Farmer and Foley,2009), (Gu and Novak, 2009), (Pogson, M. et al., 2008), (Sun, T. et al., 2008), (Bryson, J. J. et al., 2007), (Pogson, M. et al., 2006), (Walker, D. C. et al., 2004).

Plewczynski D..
BRAINSTORMING - Agent based Meta-learning Approach.
DOI: 10.5220/0003310104870490
In *Proceedings of the 3rd International Conference on Agents and Artificial Intelligence* (ICAART-2011), pages 487-490
ISBN: 978-989-8425-41-6

2 CONSENSUS OF MACHINE LEARNING METHODS

The model of meta-learning within Agent Based Modelling (ABM) framework is based on several assumptions:

1. Binary Logic

I assume the binary logic of individual intelligent, or so called cognitive agents CAs, i.e. we deal with *N* different software agents. For the single prediction, each algorithm gives one of two opposite decisions ("YES" or "NO") described here by the variable $\sigma_j = \pm 1$. Typically ML algorithms, such as support vector machines, decision trees, trend vectors, artificial neural networks, random forest, predict two classes for incoming data, based on trained models. Therefore the prediction of all ML algorithms addresses the same question: is a query item in the class of positives ("YES"), or it is not ("NO").

2. Strength Parameters

Each CA is characterized typically by two parameters: $p_j = f(precision, j)$ and $s_j = f(recall, j)$ that describe the quality of predictions for individual algorithm implemented by an agent (described by *j* index). It depends on a training dataset; the values of those parameters will be different for each of training session, or cognitive task. Therefore the parameters should be averaged over different cognitive tasks in order to make them data-independent. The quality of brainstorming approach depends on mean values $p = \Sigma \frac{p_j}{N}$ and $s = \Sigma \frac{s_j}{N}$ calculated over used learners.

3. Probability of Success

The weighted majority-minority balance in the system is given by the equation:

$$m = \frac{\Sigma_j \frac{\left(s_j + p_j\right)\sigma_j}{N(s+p)} + 1}{2}. \quad (1)$$

The normalized and nonnegative value of *m* describes the probability for the correct prediction, i.e. we assume here the modified or weighted vote rule. Each learner votes for the final prediction outcome, all votes are gathered, and the relative probability of correct answer is calculated, as given by the set of individual learners.

4. Brainstorming: The Procedure of Meta-learning

The global preference toward selected solution in brainstorming method is described as the global order parameter that is calculated using all used CAs. Each cognitive network node (*learner*, or in other words *intelligent agent*) performs its own and independent training on available input data (both the training and testing datasets are identical for all learners). In the prediction step, a query of testing items is analyzed independently by each agent, which predicts the query item classification (positive or negative). Then, all predictions done by a set of learners are gathered and integrated into the single prediction via majority rule. This view of the consensus as between various machine learning algorithms is especially useful for artificial intelligence, or robotic applications, where adaptive behavior given by the integration of results from a set of ML methods.

3 CONCLUSIONS

Generally there are two competing philosophies in supervised learning, where goal is to minimize the probability of model errors on future data. A single model approach tries to build a single good model: either not using Occam's razor principle (Minimax Probability Machine, trees, Neural Networks, Nearest Neighbor, Radial Basis Functions) or those based on Occam's razor models that select the best model as the simplest one (Support Vector Machines, Bayesian Methods, other kernel based methods such as Kernel Matching Pursuit). An ensemble of models states that a good single model is difficult to compute, so it tries to build many of those and combine them. Combining many uncorrelated models produces better predictors as was observed in models that don't use randomness or use directed randomness (Boosting, Specific cost function, Gradient Boosting, a boosting algorithm derivative for any cost function), or in models that incorporate randomness (Bagging, Bootstrap Sample: Uniform random sampling with replacement, Stochastic Gradient Boosting, Random Forests, or by inputs randomizations for splitting at tree nodes).

Meta-learning approach trains an ensemble of machine learning algorithms on the whole or different subset of all available training examples. The consensus gathers all solutions and tries to balance between them in order to maximize the prediction performance. No early solution, even provided by a generally low performing module, is not discarded until the late phase of prediction, when comparing different machine learning classifiers draws the final conclusion. This final phase is focusing on balance the generality of solution and the overall performance of trained model. Early results shows, that brainstorming approach reaches higher performance than any single method used in consensus. This confirms reported results of other meta-learning approaches based on different versions of single machine learning algorithm or those that use a set of different ML (Plewczynski, D., 2009), (Plewczynski, D., 1998), (Plewczynski, D., 2010).

ACKNOWLEDGEMENTS

This work was supported by Polish Ministry of Education and Science (N301 159735, N518 409238 and others).

REFERENCES

Can, T., et al., *Automated protein classification using consensus decision.* Proc IEEE Comput Syst Bioinform Conf, 2004: p. 224-35.

Han, X., *Cancer molecular pattern discovery by subspace consensus kernel classification.* Comput Syst Bioinformatics Conf, 2007. 6: p. 55-65.

Schulze-Kremer, S. and R.D. King, *IPSA-Inductive Protein Structure Analysis.* Protein Eng, 1992. 5(5): p. 377-90.

Vernikos, G. S. and J. Parkhill, *Resolving the structural features of genomic islands: a machine learning approach.* Genome Res, 2008. 18(2): p. 331-42.

Arimoto, R., M. A. Prasad, and E.M. Gifford, *Development of CYP3A4 inhibition models: comparisons of machine-learning techniques and molecular descriptors.* J Biomol Screen, 2005. 10(3): p. 197-205.

Bhasin, M. and G. P. Raghava, *Prediction of CTL epitopes using QM, SVM and ANN techniques.* Vaccine, 2004. 22(23-24): p. 3195-204.

Briem, H. and J. Gunther, *Classifying "kinase inhibitor-likeness" by using machine-learning methods.* Chembiochem, 2005. 6(3): p. 558-66.

Burton, J., et al., *Virtual screening for cytochromes p450: successes of machine learning filters.* Comb Chem High Throughput Screen, 2009. 12(4): p. 369-82.

Yao, X. Q., H. Zhu, and Z. S. She, *A dynamic Bayesian network approach to protein secondary structure prediction.* BMC Bioinformatics, 2008. 9: p. 49.

Abrusan, G., et al., *TEclass--a tool for automated classification of unknown eukaryotic transposable elements.* Bioinformatics, 2009. 25(10): p. 1329-30.

Hwang, S., Z. Gou, and I. B. Kuznetsov, *DP-Bind: a web server for sequence-based prediction of DNA-binding residues in DNA-binding proteins.* Bioinformatics, 2007. 23(5): p. 634-6.

Garg, P., et al., *SubCellProt: predicting protein subcellular localization using machine learning approaches.* In Silico Biol, 2009. 9(1-2): p. 35-44.

Miller, M. L. and N. Blom, *Kinase-specific prediction of protein phosphorylation sites.* Methods Mol Biol, 2009. 527: p. 299-310, x.

Bindewald, E. and B. A. Shapiro, *RNA secondary structure prediction from sequence alignments using a network of k-nearest neighbor classifiers.* RNA, 2006. 12(3): p. 342-52.

Plewczynski, D. and K. Ginalski, *The interactome: predicting the protein-protein interactions in cells.* Cell Mol Biol Lett, 2009. 14(1): p. 1-22.

Plewczynski, D., et al., *The RPSP: Web server for prediction of signal peptides.* Polymer, 2007. 48: p. 5493-5496.

Plewczynski, D., A. H. Spieser, and U. Koch, *Performance of Machine Learning Methods for Ligand-Based Virtual Screening.* Combinatorial Chemistry & High Throughput Screening, 2009.

Plewczynski, D., et al., *AutoMotif server: prediction of single residue post-translational modifications in proteins.* Bioinformatics, 2005. 21(10): p. 2525-7.

Plewczynski, D., et al., *Virtual High Throughput Screening Using Combined Random Forest and Flexible Docking.* Combinatorial Chemistry & High Throughput Screening, 2009.

Plewczynski, D., S. A. H. Spieser, and U. Koch, *Assessing different classification methods for virtual screening.* Journal of Chemical Information and Modeling, 2006. 46(3): p. 1098-1106.

Bruce, C. L., et al., *Contemporary QSAR classifiers compared.* J Chem Inf Model, 2007. 47(1): p. 219-27.

Islam, M. M., et al., *Bagging and boosting negatively correlated neural networks.* IEEE Trans Syst Man Cybern B Cybern, 2008. 38(3): p. 771-84.

Plewczynski, D., S. A. Spieser, and U. Koch, *Performance of machine learning methods for ligand-based virtual screening.* Comb Chem High Throughput Screen, 2009. 12(4): p. 358-68.

Schwenk, H. and Y. Bengio, *Boosting neural networks.* Neural Comput, 2000. 12(8): p. 1869-87.

Serpen, G., D. K. Tekkedil, and M. Orra, *A knowledge-based artificial neural network classifier for pulmonary embolism diagnosis.* Comput Biol Med, 2008. 38(2): p. 204-20.

Shrestha, D. L. and D. P. Solomatine, *Experiments with AdaBoost.RT, an improved boosting scheme for regression.* Neural Comput, 2006. 18(7): p. 1678-710.

27. Peng, Y., *A novel ensemble machine learning for robust microarray data classification.* Comput Biol Med, 2006. 36(6): p. 553-73.

Wang, C. W., *New ensemble machine learning method for classification and prediction on gene expression data.* Conf Proc IEEE Eng Med Biol Soc, 2006. 1: p. 3478-81.

Yang, J. Y., et al., *A hybrid machine learning-based method for classifying the Cushing's Syndrome with* *comorbid adrenocortical lesions.* BMC Genomics, 2008. 9 Suppl 1: p. S23.

Devillers, J., et al., *Internet resources for agent-based modelling.* SAR QSAR Environ Res. 21(3-4): p. 337-50.

Moore, D., et al., *Extending drug ethno-epidemiology using agent-based modelling.* Addiction, 2009. 104(12): p. 1991-7.

Farmer, J. D. and D. Foley, *The economy needs agent-based modelling.* Nature, 2009. 460(7256): p. 685-6.

Gu, W. and R.J. Novak, *Agent-based modelling of mosquito foraging behaviour for malaria control.* Trans R Soc Trop Med Hyg, 2009. 103(11): p. 1105-12.

Pogson, M., et al., *Introducing spatial information into predictive NF-kappaB modelling--an agent-based approach.* PLoS ONE, 2008. 3(6): p. e2367.

35. Sun, T., et al., *Agent based modelling helps in understanding the rules by which fibroblasts support keratinocyte colony formation.* PLoS ONE, 2008. 3(5): p. e2129.

Bryson, J. J., Y. Ando, and H. Lehmann, *Agent-based modelling as scientific method: a case study analysing primate social behaviour.* Philos Trans R Soc Lond B Biol Sci, 2007. 362(1485): p. 1685-98.

Pogson, M., et al., *Formal agent-based modelling of intracellular chemical interactions.* Biosystems, 2006. 85(1): p. 37-45.

Walker, D. C., et al., *The epitheliome: agent-based modelling of the social behaviour of cells.* Biosystems, 2004. 76(1-3): p. 89-100.

Plewczynski, D., *Mean-field theory of meta-learning.* Journal of Statistical Mechanics-Theory and Experiment, 2009: p. -.

Plewczynski, D., *Landau theory of social clustering.* Physica A, 1998. 261(3-4): p. 608-617.

Plewczynski, D., *Brainstorming: weighted voting prediction of inhibitors for protein targets.* J Mol Model, 2010.