Project Part 2

Group 6: Janette Rounds, David Rice, Mitch Vander Linden

September 30, 2015

In this paper we discuss three Machine Learning algorithms, a classifier that uses Bayesian Optimization, an algorithm for automatically scheduling drip irrigation and an algorithm for identifying orchards from images. Scheduling drip irrigation and identifcation of fruit trees are both problems in the domain of Precision Agriculture. The other problem, creating a powerful Bayesian classifier with less stringent assumptions, is well known in the field of Machine Learning.

1 Bayesian Optimization Classifier

In the field of Machine Learning, there are many models based off of Bayes Theorem:

$$P(x|y) = \frac{P(y|x)P(x)}{P(y)}$$

where x and y are events, and $P(x)$ is the probability of x [1]. In short, Bayes Theorem states that the probability of an event x, given that we know event y happened, is proportional to the probability of event y given event x times the probability of x. Bayes Theorem allows us to use prior information about a domain or a model in order to assess what can be thought of as a "degree of belief".

There are two Bayesian methods that we are concerned with in Liu et al.[2, Liu 2013], Bayesian Optimization and Naive Bayes Classification (NB). NB is a powerful Machine Learning method that allows us to assign a category (or a label) to an observation based on a training data set [1]. NB is a powerful model, but it also has some very strong assumptions, primarily that all attributes (variables that help us figure out which label to assign) are independent of one another. There are countless domains and problems where this assumption does not hold.

Our other Bayesian method of interest is Bayesian Optimization (BO). BO involves approximating functions, especially functions that are computationally expensive, where derivatives do not exist or are unknown, or where there are only a few observations available [3]. In Liu et al.[2, Liu 2013], the authors show that, through an algorithm that converts the BO model into a classifier, we can create a Bayesian Classifier without the strong independence assumption of Naive Bayes Classifiers.

2 Tree Counting Algorithm

Tree identification from images is ordinarily a process that contains plenty of noise. However, much academic insight and effort has been put in to the challenge of image processing, geospatial data integration, handling optical sensor input, and understanding reflectance data.

Although much work has been done to analyze data from forest scans, a unique challenge has presented itself when it comes to man made nurseries consisting of fruit trees[4, Santoro 2013]. The trees tend to be spaced further apart and the crowns of the trees spread more. These tendencies mean that identification of trees can be confounded by reflections from man-made structures, soil conditions and even things like shadows beneath the tree. The authors show that their four-part algorithm (asymmetrical smoothing filter, local minimum filter, mask layer, and spatial aggregation operator) can identify tree position quickly and consistently.

3   Automated Scheduling Algorithm

The question of efficient crop irrigation is one of great importance in agricultural pursuits of all kinds. On the one hand, the results of under-watering are quite apparent. In contrast, over-watering can also have detrimental effects, and in many climates water is a limited resource that cannot be wasted.

This algorithm aims to solve the problem of automated scheduling of drip irrigation in tree crops by efficiently streamline seven distinct but related tasks: estimation of irrigation needs, adaptation to the irrigation setup, execution of the schedule, plant and soil monitoring, automated interpretation of data, reaction to unforeseen events, and continuous tuning of the model[5, Casadesus 2012]. It is this final constraint that requires the algorithm to implement machine learning - by constantly fine-tuning the irrigation process based on recent data, it attempts to construct a solution that is best-fit to the current conditions. [5, Casadesus 2012]

References

[1] D. Koller and N. Friedman, *Probabilistic graphical models: principles and techniques*, ch. 3, The Bayesian Network Representation, pp. 45–102. MIT press, 2009.

[2] S. Liu, M. Zhu, and Y. Yang, "A bayesian classifier learning algorithm based on optimization model," *Mathematical Problems in Engineering*, vol. 2013, 2013.

[3] E. Brochu, V. M. Cora, and N. De Freitas, "A tutorial on bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning," *arXiv preprint arXiv:1012.2599*, 2010.

[4] F. Santoro, E. Tarantino, B. Figorito, S. Gualano, and A. M. D'Onghia, "A tree counting algorithm for precision agriculture tasks," *International Journal of Digital Earth*, vol. 6, no. 1, pp. 94–102, 2013.

[5] J. Casadesús, M. Mata, J. Marsal, and J. Girona, "A general algorithm for automated scheduling of drip irrigation in tree crops," *Computers and electronics in agriculture*, vol. 83, pp. 11–20, 2012.