**Lesson 05 Discussion**

Support Vector Machine (SVM) and logistic regression models are closely related. Comment on at least two of the following:

* How can you describe the differences in loss functions of these two methods? How might this matter in practice.
* Both methods can be extended to nonlinear problems. How might the differences in these methods have advantages and disadvantages in practice?
* Which algorithm produces a probabilistic output? How might a probabilistic output be useful in practice?
* Describe a use case where one or the other algorithm might be preferred. Explain what the advantages of that algorithm over the other are.

To receive credit post you thought and then comment on at least one other post.

One of the main differences between Logistic Regression and Support Vector Machines is by their level of sensitivity to outliers. Support Vector Machines are unsensible to outliers as the vectors (data points) that provide the highest value for the model are those in the proximity of the separation hyperplane (class boundary region) compared to the Logistic Regression where outliers could be an issue as all the data points are considered in the calculation of the Maximum Likelihood which is the way the mode is constructed.

Even though given their similarities both models can yield very similar results, Support Vector Machines then to be a little bit better when there is a clear separation between classes compared to Logistic Regression. In contrast when there is more overlapping among the different classes, Logistic Regression is a usually selected.

Same here, it did not occur to me before that by using K Fold Cross Validation one can get an indirect probability, by what I imagine to be a voting system that based on the number of K Folds will give a probability estimate. I am guessing that potentially that probability could be develop a little bit further by means of bootstrapping.