Purpose of this report:

I have prepared this report about my experience with Parkinsons data. I am submitting this report so that tomorrow on exam it will be easier for you to understand me. This report may help me to reduce language problem between you and me.

What? I have worked!

1. I have implemented four algorithms to classify and predict using Parkinsons data. The algorithms are following:
   a. Binary logistic regression to solve multi classification problem(one vs rest)
   b. Binary logistic regression to solve multi classification problem(All pairs)
   c. Binary Support Vector machine to solve multi classification problem(one vs rest)
   d. Binary Support Vector machine to solve multi classification problem(All pairs)
2. After that I have measured accuracy of these methods on Parkinsons dataset using cross validation.

What contains this report?

1. At first Binary logistic regression
2. Then multiclass logistic regression
3. Then way how can we use binary logistic regression to solve multiclass problem
4. At last I will show what result I got using those methods on Parkinsons dataset
5. Finally, I will describe the trends of data: i.e how the results differs from each other.

1. Binary Logistic regression:

This algorithm gives a probabilistic output, that whether a data instance (x) is likely to be in a class of two classes. Set of class K={1,0}

Model:

$$\mathbf{P}(y=1\,|\,\boldsymbol{x}) = \frac{1}{1+\exp\big(-r(\boldsymbol{x})\big)} \;,$$

where

$$r(\boldsymbol{x}) \;=\; \beta_0 + \boldsymbol{w}^\top \boldsymbol{x} = \beta_0 + \sum_{i=1}^{m} w_i \cdot x_i$$

Where w = vector regression coefficient.

Beta= constant bias

y is the assigned class for a training data in training dataset.

2. Multi class logistic regression:

This algorithm gives a probabilistic output, that whether a data instance (x) is likely to be in a class of n classes. Here set of class K={k1, k2, k3…..kn}

$$\begin{aligned}\mathbf{P}(y = c_k \,|\, \boldsymbol{x}) &= \mathbf{P}(y \leq c_k \,|\, \boldsymbol{x}) - \mathbf{P}(y \leq c_{k-1} \,|\, \boldsymbol{x}) \\ &= \pi_k(\boldsymbol{x}) - \pi_{k-1}(\boldsymbol{x})\end{aligned}$$

$$\begin{aligned}\pi_k(\boldsymbol{x}) = \pi_k(\boldsymbol{x} \,|\, \boldsymbol{\beta}, \boldsymbol{w}) = \mathbf{P}(y \leq c_k \,|\, \boldsymbol{x}) \\ = \frac{\exp(\beta_k)\exp(\boldsymbol{w}^\top \boldsymbol{x})}{1 + \exp(\beta_k)\exp(\boldsymbol{w}^\top \boldsymbol{x})}\end{aligned}$$

Where w = vector regression coefficient.

Beta= constant bias

y is the assigned class for a training data in training dataset.

3. Then way how can we use binary logistic regression to solve multiclass problem:

In practice it is very much convenient to use binary Logistic regression than multi class logistic regression. Now I will describe how we can we use binary logistic regression for Multi class problem. I will discuss one method named One versus Rest.

One vs Rest: let we have a set class

$$\mathcal{Y} = \{c_1, c_2, \ldots, c_K\}$$

The idea is to learn K models

$$h_1, h_2, \ldots, h_K$$

The task of the model hk is to distinguish instances belonging to class from instances that do not belong to this class. Correspondingly, if the original training data is given by

$$\mathcal{D} = \big\{(\boldsymbol{x}_1, y_1), (\boldsymbol{x}_2, y_2), \ldots, (\boldsymbol{x}_N, y_N)\big\} \subset (\mathcal{X} \times \mathcal{Y})$$

The model hk is trained on the transformed data

$$\mathcal{D}_k = \big\{(\boldsymbol{x}_1, y_1^{(k)}), (\boldsymbol{x}_2, y_2^{(k)}), \ldots, (\boldsymbol{x}_N, y_N^{(k)})\big\} \subset (\mathcal{X} \times \{-1, +1\}) \ ,$$

where

$$y_n^{(k)} = \begin{cases} -1 & \text{if } y_n \neq c_k \\ +1 & \text{if } y_n = c_k \end{cases} \ .$$

In practice we use a base learner (model) as hk and train it K times (once for each class)

Now, suppose that all models have been trained, and that a new query instance x is submitted for classification. Ideally, there is one k such that

hk(x) = +1 while hi(x) = -1 for all i not equals to k; in that case, the predicted class would be ck.

Therefore, instead of producing mere binary classifications, one typically trains scoring classifiers. Given predictions of that kind, a natural classification rule is to choose the class ck such that

$$h_k(\boldsymbol{x}) = \max_{1 \le i \le K} h_i(\boldsymbol{x}) \ .$$

And finally class k of hk(x) will be the predicted class for data instance x

4. <u>Result I got classifying the Parkinson dataset:</u>

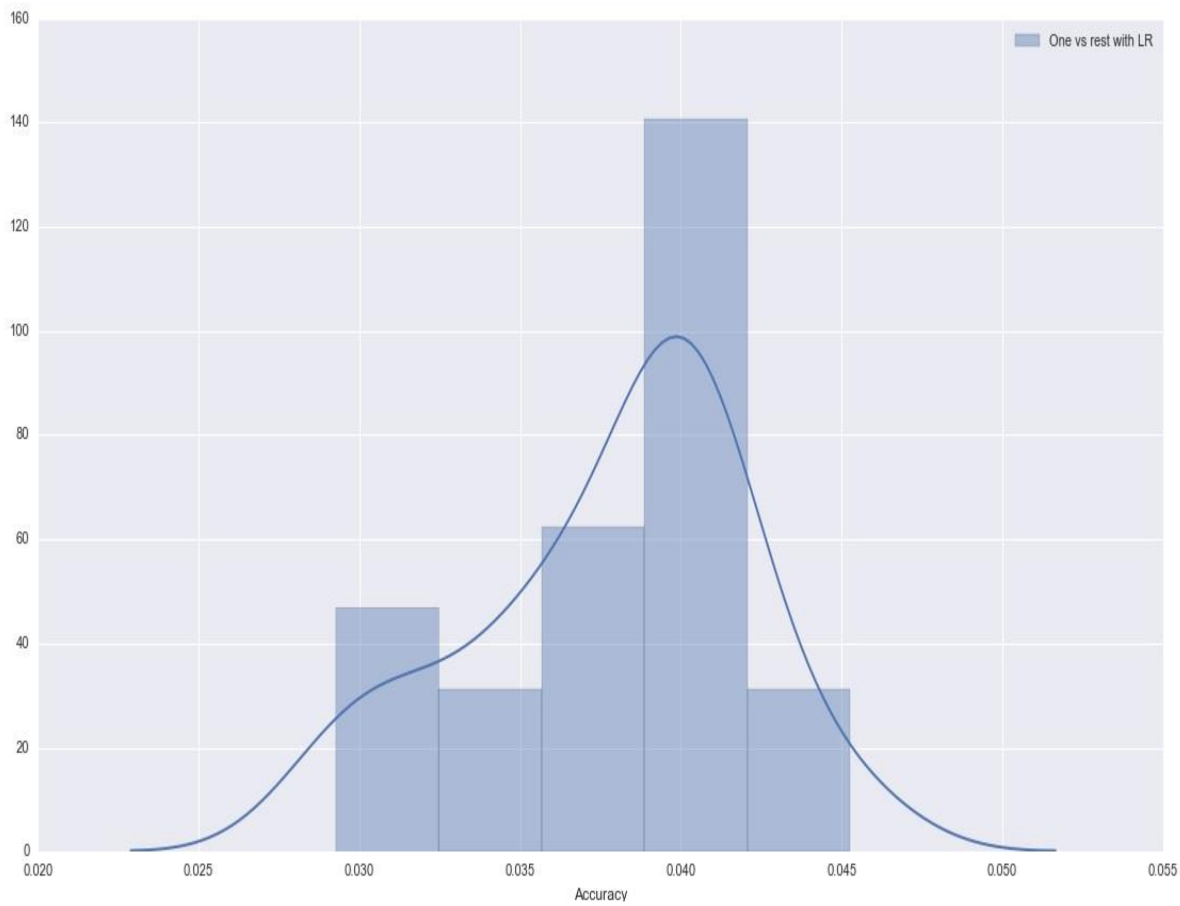
I have predicted total_UPDRS in Parkinson data by using the last 16 column of each observation.

At first I implemented **Binary logistic regression for Multiclass problem using One vs Rest fashion**:

Cross Validation{folding =20}

Mean=3.6, Standard deviation =0.4

And the distribution of accuracy found from 20 foldings is following:

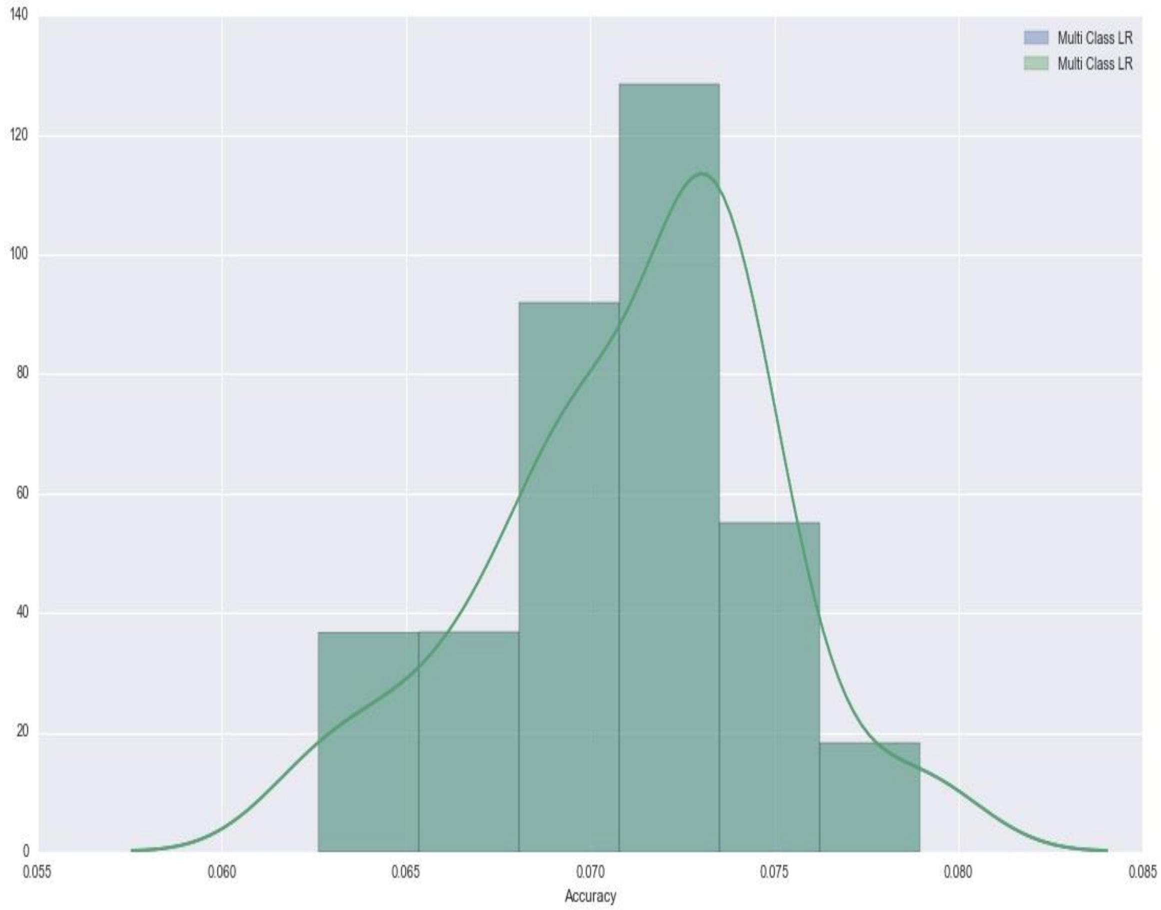


Then I have implemented **Binary logistic regression for Multiclass problem using All pairs** fashion:

Cross Validation{folding =20}

Mean=7.1, SD=0.4

And the distribution of accuracy found from 20 foldings is following:

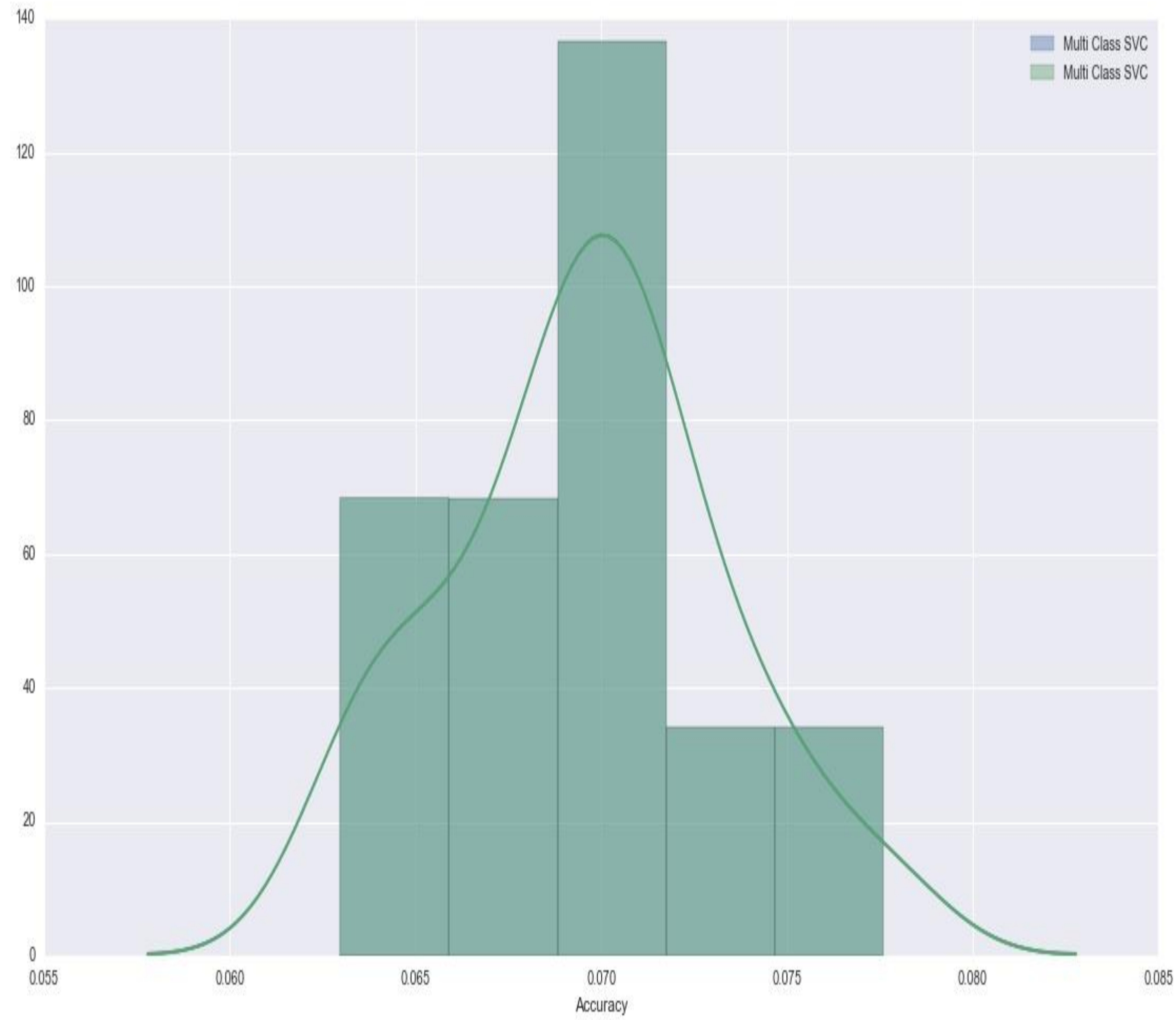


Then I have implemented **Binary Support vector Machine for Multiclass problem using All pairs** fashion:

Cross Validation{folding =20}

Mean=7.0, SD=0.4

And the distribution of accuracy found from 20 foldings is following:

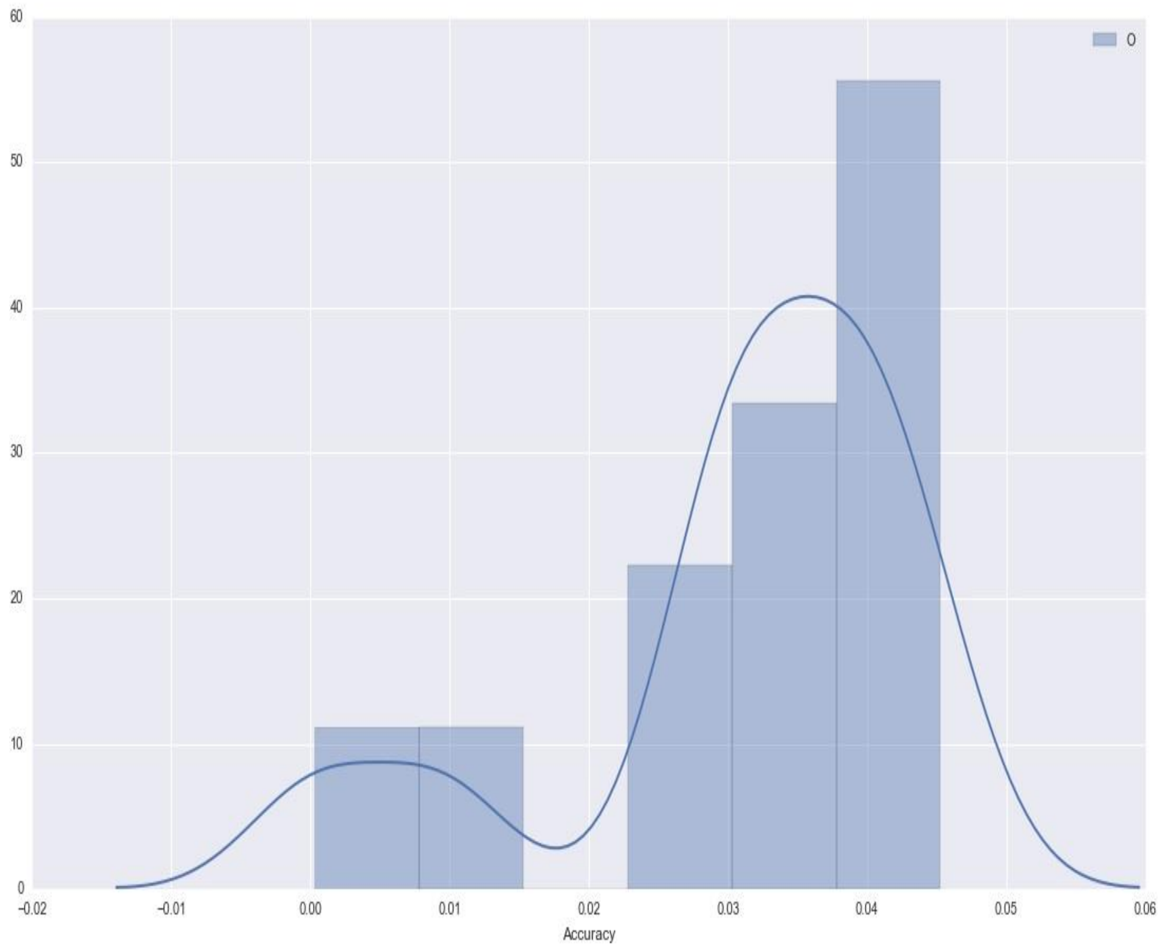


Then I have implemented **Binary Support vector Machine for Multiclass problem using One vs Rest** fashion:

Cross Validation{folding =12}

Mean=3.1, SD=1.3

And the distribution of accuracy found from 12 foldings is following:



5. <u>Result analysis:</u>

The graphs above shows that:

Binary Logistic regression for Multiclass problem using One vs Rest,

Binary Logistic regression for Multiclass problem using All pairs,

Binary Support vector Machine for Multiclass problem using All pairs

all of these methods have the same standard deviation=0.4 which indicates that all these methods perform about the same as classifiers.

But Binary Support vector Machine for Multiclass problem using one vs rest have SD=1.3 which to high and indicates that classification performance is too low here. It may because of folding number. In this case folding number=12 where in other cases it was 20.