Part II. Understand the Dataset:

Mushroom Dataset:

- Relevant Paper: **Extraction of crisp logical rules using constrained backpropagation networks**
- Author: Duch W, Adamczak R, Grabczewski K

Abstract:

In this paper, the two methods Iris and Mushroom classification are compared to create a network with minimum number of connections that ultimately cause small number of crisp logical rules that gives better accuracy than the original nueral classifiers.

Mushroom Dataset Information:

This dataset is obtained from the UCI machine learning repository. The Mushroom dataset consists of 8124 instances which has 22 discrete attributes. Amongst these half of the instances represents the edible mushrooms and the rest represents the non-edible/poisonous mushrooms.

Summary:

Here, there are two classification methods used:

1. MLP2LN – Multi-layered Perceptrons to a Logical Network
   This classifier does mapping from input to the output space. Here input and output can be vectors.
2. SLF – Structural Learning with Forgetting
   It uses the Laplace-type regularization term.
   $E(W) = E0(W) + \lambda \sum_{i,j} |Wij|$

A single neuron can learn all the given training samples but the network gets tough to interpret. Thus, both the methods are applied, MLP2LN as well as SLF.

This approach results in 99.14% accuracy with 48 errors.

The following single approach is used:

edible if odor=(almond$\vee$anise$\vee$none)$\wedge$ spore-print-color= $\neg$green

This means that the mushroom is edible with the above feature combination of odor and spore print color i.e. if its odor is one amongst almond or anise or none and if the spore-print-color is not green.

This approach has two features odor and spore-print-color and their respective four antecedents.

SLF alone arrive at perfect accuracy by using weaker regularization parameters for edible mushrooms.

MLP2LN arrived at the following disjunctive rule for poisonous mushrooms:

- Rule-1: odor=$\neg$(almond$\vee$anise$\vee$none)
  This means that odor is not almond or anise or none which gives 120 errors and 98.52% accuracy on the whole dataset.

- Rule-2: spore-print-color=green
  This means that the only attribute spore-print-color is green which gives 48 errors and 99.41% accuracy.

- Rule-3: odor=none$\wedge$stalk-surface-below-ring=scaly$\wedge$ (stalk-color-above-ring=$\neg$brown)
  This rule with 3 attributes odor, scaly and ring gives only 8 errors and 99.90% accuracy.

- Rule-4: habitat=leaves$\wedge$cap-color=white
  This rule with 2 attributes gives perfect classification and accuracy.

Same results are obtained by training on random 10% of the database as well as on the whole dataset.

The rule extraction from neural network gets natural flexibility which is not common in other approaches. One important concept here is the stability of rules. The number of rules increases if the number of linguistic variables increases due to which logical approximation may become accurate. The important factor for stability is the choice of linguistic variables. Thus, the linguistic variables are optimized using extracted rules. Crisp logical rules that are found by analyzing the trained network nodes performs better than the other methods of rule extraction.