

Problem Statement: PARKINSON'S DISEASE DETECTION

Description:- Parkinson's disease (PD) is a long-term degenerative disorder of the central nervous system that mainly affects the motor system. The symptoms usually emerge slowly. Early in the disease, the most obvious symptoms are shaking, rigidity, slowness of movement, and difficulty with walking. Thinking and behavioral problems may also occur. Dementia becomes common in the advanced stages of the disease. Depression and anxiety and other symptoms include sensory, sleep, and emotional problems. The main motor symptoms are collectively called "parkinsonism", or a "parkinsonian syndrome". The cause of the disease are genetics and environmental factors. Parkinson's disease is the most prevalent neurodegenerative disorder affecting more than 10 million people worldwide. There is no single test which can be administered for diagnosing Parkinson's disease. Because of these difficulties, to investigate a machine learning approach to accurately diagnose Parkinson's, using a given dataset. To prevent this problem in medical sectors, have to predict the disease affected or not by finding accuracy calculation using machine learning techniques. The aim is to investigate machine learning based techniques for Parkinson disease by prediction results in best accuracy with finding classification report. The analysis of dataset by Support Vector Machine. It is a new generation learning system based on recent advances in statistical learning theory. It is an algorithm for both linear and nonlinear data. It transforms the original data in a higher dimension, from where it can find a hyper plane for separation of the data using essential training tuples called support vector. A SVM constructs a hyper plane or set of hyper planes in a high or infinite-dimensional space, which can be used for classification, regression, or other tasks. Intuitively, a good separation is achieved by the hyper plane that has largest distance to the nearest training data point of any class so called functional margin, since in general the larger the margin the lower the generalization error of the classifier.

Attribute Information:

Given is the attribute name, attribute type, measurement unit, and a brief description. The number of rings is the value to predict: either as a continuous value or as a classification problem.

Name / Data Type / Measurement Unit / Description

- 1- Sex / nominal / -- / M, F, and I (infant)
- 2- Length / continuous / mm / Longest shell measurement
- 3- Diameter / continuous / mm / perpendicular to length
- 4- Height / continuous / mm / with meat in shell
- 5- Whole weight / continuous / grams / whole abalone
- 6- Shucked weight / continuous / grams / weight of meat
- 7- Viscera weight / continuous / grams / gut weight (after bleeding)
- 8- Shell weight / continuous / grams / after being dried
- 9- Rings / integer / -- / +1.5 gives the age in years

Building a Regression Model

1. Download the dataset: Dataset
2. Load the dataset into the tool.
3. Perform Below Visualizations.
 - Univariate Analysis
 - Bi-Variate Analysis
 - Multi-Variate Analysis
4. Perform descriptive statistics on the dataset.
5. Check for Missing values and deal with them.
6. Find the outliers and replace them outliers
7. Check for Categorical columns and perform encoding.
8. Split the data into dependent and independent variables.
9. Scale the independent variables
10. Split the data into training and testing
11. Build the Model
12. Train the Model
13. Test the Model

14. Measure the performance using Metrics.