## Comparison of Various Supervised Learning Models:

We have experimented with 6 different supervised learning approaches until now in the two domains. The performance of these algorithms has been compared with various aspects of performance like Accuracy, complexity, precision, recall, confusion matrix, training speed, prediction speed, performance when number of observations are small and noise handling.

**Accuracy:**

Accuracy is a measure of how often the classifier correctly predicted the label.

**Complexity:**

Complexity is a measure of how long the algorithm takes to run and classify data.

**Precision:**

Precision is the ratio of the number of True Positives to the number of predicted positives.

**Recall:**

Recall is the ratio of the number of True Positives to the number of actual positives in the dataset.

**Confusion Matrix**:

Confusion Matrix is a matrix that is used to describe the performance of a classification model.

**Training Speed:**

Training Speed is a measure of how fast or slow the algorithm takes to train.

**Prediction Speed:**

Prediction Speed is a measure of how fast or slow the algorithm takes to predict the label.

**Performance with small number of observations:**

This is used to measure how the classifier performs when the number of training samples is small.

**Handles noise:**

This is used to measure if high noise in the dataset is handled by the classifier or not.

The comparison of results is as follows:

**Digit Recognition Dataset:**

| | **Decision Tree** | **Artificial Neural Networks** | **KNN** | **Ada Boosting** | **SVM** | **Naïve Bayes** |
|---|---|---|---|---|---|---|
| **Accuracy** | 85% | 95.77% | 98.21% | 96.38% | 97.88% | 89.14% |
| **Complexity** | O(m*n*log(n)) | $O(n*m*h^{k}*o*i)$ | O(nd+nk) | O(mnp) | $O(m*n^{2})$ to $O(m*n^{3})$ | O(n) |
| **Precision** | 0.84 | 0.94 | 0.97 | 0.95 | 0.97 | 0.89 |
| **Recall** | 0.77 | 0.90 | 0.91 | 0.90 | 0.92 | 0.85 |
| **F Score** | 0.81 | 0.92 | 0.95 | 0.93 | 0.95 | 0.87 |
| **Training Speed** | Fast | Slow | Fast | Slow | Medium | Fast |
| **Prediction Speed** | Fast | Fast | Slow | Fast | Fast | Fast |
| **Performance with small no. of observations** | Performs poorly | Performs poorly | Performs poorly | Performs poorly | Performs poorly | Performs well |
| **Handles high Noise** | No | Yes | No | No | Yes | Yes |

m – number of features
n – number of samples
d - distance to one example
k – number of hidden layers (ANN); k – 'k' neighbors (KNN)
p – number of rounds
h – total number of neurons
o- number of output neurons
i – number of iterations