**Literature Survey – Classification Algorithms**

**Classification:** In machine learning and statistics, classification is the problem of identifying to which a set of categories a new observation belongs, on the basis of a training set of data containing observations whose category membership is known. An example would be assigning a given email into “spam” or “not-spam” categories.

**Classification Algorithms:**

* **Naïve Bayes Classifier:** In Machine Learning, naïve Bayes classifiers are a family of simple probabilistic classifiers based on applying Bayes’ theorem with string independence assumptions between the features. Naïve Bayes classifiers are highly scalable, requiring a number of parameters linear in the number of variables in a learning problem. Maximum-likelihood training can be done by evaluating a closed-form expressing, which takes linear time, rather than by expensive iterative approximation as used for many other types of classifiers.
* **Support Vector Machine:** In Machine Learning, support vector machines(SVMs) are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. Given a set of training examples, each marked as belonging to one or the other of two categories, an SCM training algorithm builds a model that assigns new examples to one category or the other, making it a non-probabilistic binary linear classifier. An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall.
* **K –** **Nearest Neighbors:** In pattern recognition, the k-nearest neighbor’s algorithm(k-NN) is a non-parametric method used for classification and regressing. The input consists of the k closest training examples in the feature space. In classification, the output is class membership. An object is classified by a majority vote of its neighbors, with the object being assigned to the class most common among its k nearest neighbors (k is a positive integer).
* **Random Forests:** Random Forests are an ensemble learning method for classification, regression and other tasks, that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes of the individual trees. Random decision forests correct for decision trees habit of overfitting to their training set.
* **Convolutional Neural Networks:** In Machine Learning, a convolutional neural network(CNN or ConvNet) is a class of deep, feed-forward artificial neural networks, that has successfully been applied to analyzing visual imagery. CNNs use a variation of multilayer perceptrons designed to required minimal processing. They are also known as shift invariant or space invariant artificial neural networks.