Rish, Irina [34] performed a study on naïve bayes classifier and analyzed the impact of the distribution entropy on the classification error showing that low entropy feature distributions yield good performance of naïve Bayes. This classifier has also proved to be effective in many practical applications, including text classification, medical diagnosis, and systems performance management.

Daniela Xhemali et. al. [35] designed an Enhanced naïve Bayes classifier model to classify data and compared the results with naïve Bayes classifier and decision tree classifier. This research showed that the Enhanced Naïve Bayes classifier performance was far superior compared to traditional naïve Bayes classifier and the classification accuracy was 95.20%.

The concept of using radial basis functions for learning in adaptive layered networks was first discussed by Broomhead and Lowe in 1988 [36]. They explained that radial basis functions can be applied to general multi-layered perceptron for which the output units have an invertible non linearity. The input-hidden weights adjust slowly because of the non-linear optimization but the hidden-output weights adjust very rapidly due to the linear optimization and hence will always remain in a global minimum of error surface over the hidden-output weights which are controlled by the input-hidden weights.

Dong Xiao Ni [37] developed a neural network model for Optical Character Recognition (OCR) application. Several neural software packages have been developed to study the problem with different neural architectures and learning rules.

Guang-Bin Huang [38] designed a new sequential learning algorithm called Generalized Growing And Pruning algorithm for RBF (GGAP-RBF) which is based on the concept of linking the learning accuracy with the significance of the nearest neuron. This algorithm was found to produce excellent results and outperformed several other
sequential algorithms in terms of learning speed, network size and generalization.

M. Bortman and M. Aladjem [39] constructed a new network classifier for Resource Allocation Network (RAN). In this work a modified form of GGAP-RBF algorithm is proposed and found to outperform the original GGAP in terms of lower prediction error and the trained network exhibited reduced complexity.

A Hybrid Forward neural network Algorithm (HFA) for RBF was studied by Jian-Xun Peng et. al.[40]. In the proposed HFA, the network is constructed within a unified analytic framework which improved network performance and reduced the memory usage. The parameter optimization method used is found to speed up the convergence of the network search.

Tianping Chen and Hong Chen [41] studied the problem of function approximation using radial basis function neural networks and a suitable necessary and sufficient condition for a continuous function to be qualified as an activation function has been designed. The results show excellent improvement in function approximation methods.

The identification of system parameters with uncertain data is termed as blind equalization and Nan Xie and Henry Leung [42] used Radial Basis Function Neural networks to solve the problem of improving least square errors to reduce the impact of noisy data. The convergence of the improved least square method is studied and a new asymptotic mean square error is proposed for RBF identification method.

Radial basis networks are aimed at an approximation of nonlinear mappings from $\mathbb{R}^n$ to $\mathbb{R}$. Witold Pedrycz [43] has presented a method to use fuzzy clustering techniques to preserve the homogeneity of the patterns in the input space. The research claims that the incorporation of fuzzy clustering has resulted in faster learning in Radial Basis Function neural network.
Nicolaos B. Karayiannis [44] has devised a method of constructing reformulated radial basis function neural networks to the selection of admissible generator functions. This research introduced a new family of functions called cosine radial basis functions. These functions when used for pattern classification outperformed other feed forward neural networks considerably.

In the paper [45] Francesco Lampariello and Marco Sciandrone, formulated a new method to train radial basis function efficiently. In this work the problem is formulated based on the system of nonlinear inequalities, and a suitable error function is determined. This training method has produced excelled result in terms of computational time saving.

Meng Joo Er et. al. [46] used RBF classifier for face recognition problem. In this work, a new hybrid learning algorithm is introduced in which the training is done with drastically reduced data set in terms of gradient paradigm. The simulation results produced excellent performance both in terms of reduced classification error rates and improved learning efficiency.

Sunan Huang and Kok Kiong Tan [47] used RBF neural networks for the study of fault detection. The RBF is trained to study the characteristics of a fault mode and is made to identify any failure mode by comparing it with any known failure mode.

A supervised learning algorithm for reformulated radial basis function neural network was formulated by Nicolaos B. Karayiannis [48] for gradient descent learning with improved convergence. This method is found to perform better compared to traditional gradient descent learning methods used in RBF neural networks.

Meng Joo Er. [49] has presented a high-speed face recognition method combining discrete cosine transforms, Fisher’s linear discriminant and RBF neural classifier. The proposed method is found to be robust against uniform brightness in
variations of images. This method also helps in the optimum fixation of hidden units in RBF neural networks.

2.4.2 Unsupervised Learning Algorithms

When the learning is based only upon the input data and is independent of the desired output data no error is calculated to train a network. This type of learning is called unsupervised learning and the input data are called unlabelled data. The leaning in such neural networks takes place based on the outcome of the output neuron competition and hence these are termed as competitive networks.

Clustering algorithms are the widely used unsupervised learning algorithms whose function is to group data based on the similarities found in the data. The main characteristic of clustering algorithms is that the effectiveness of the method depends on the definition of the distance measure.

![Unsupervised Classification Diagram](image)

Figure 2.3 Unsupervised Learning Procedures
The clustering algorithms are categorized into four categories as follows:

- Exclusive Clustering
- Overlapping Clustering
- Hierarchical Clustering
- Probabilistic Clustering

K-means clustering algorithm is a widely used unsupervised learning algorithm for purposes of classification. The K-means clustering algorithm proposed by Hartigan, John A., and Manchek A. Wong [50] easily identifies the similarities in the given data set and produces a fixed number of cluster centers. Since the initial allocation of the cluster center may not be a good model for the probability distribution of the input data, this process of fixing the centers is repeated where each cluster center is moved to the mean position of all the training patterns in a region.

Wagstaff, Kiri, et al [51] conceived a new constrained k-means clustering algorithm in which the authors define two types of constraints namely must-link and cannot-link constraints for the instances to be placed in a cluster. This method has showed improved accuracy for some data sets.

An efficient k-means clustering algorithm has been developed by Kanungo, Tapas, et al. [52] based on a filtering method where the multidimensional data points are stored in a kid-tree. Through this algorithm the authors have proved the more separated the clusters, the faster the algorithm runs.

A generalized variant of K-means clustering which is otherwise called as Adaptive K-means clustering [53] is used to change the learning rate dynamically based on the quality of current clustering. This method is suitable to situations where the source of the pattern generated is unstable.
Hierarchical clustering is a variation of linear k-means. In hierarchical clustering, clusters are organized into a tree-like structure. The results of hierarchical clustering are represented in the form of dendrogram. Davidson and Ravi [54] introduced an instance and cluster-level constraints for improving the performance of hierarchical agglomerative clustering algorithm.

In [55] Mark J Embrechts describe a new variant of agglomerative hierarchical algorithm using a method called cluster seeding for speeding up the clustering process. In this paper they also discuss several cluster validation indices that can be used to evaluate the quality of clusters formed.

2.5 Feature Selection Techniques

Feature Selection is a technique used to cut down the number of features before being applied to neural network classifier. Features that are irrelevant to the task under study will have a negative effect on classification and may result in overgeneralization or poor classification. Moreover, the computational complexity of a classification algorithm may suffer from the curse of dimensionality caused by several features. Hence to improve speed and accuracy of classification, feature selection process plays a vital role. The use and performance of feature selection techniques used in classification problems in view of the literature is given in the following section.

Leray et. al. [56] discusses the different approaches used for feature selection using neural networks for classification. They also propose a series of rules that could be used to evaluate the feature selection methods at a reasonable extra computational cost so that the neural network can be retrained taking into account some correlation between the extracted features.

S. Ledesma et. al. [57] devised a hybrid feature selection method using Genetic Algorithms with Neural networks to assist the search process and thereby retrieving
the most relevant features for classification. Their experimental shows that using Genetic Algorithm for feature selection has resulted in very few hidden units in the neural network.

Huang, Li-li, et al. [58] describe a new robust feature selection method for multi-class learning problem. In the proposed method feature correlation and feature selection are done simultaneously which is tested to produce faster convergence. The feature selection process is also found to be more effective and the experimental results show that the proposed method outperforms many of the existing feature selection methods.

Kira and Rendell [59] describe a new algorithm for feature selection called RELIEF which assigns a relevance weight to each feature based on instance learning method. Features are ranked by their weights which distinguish each feature among the class values. An attribute’s weight is determined by finding how well its values distinguish its membership into a class. An attribute is assigned a high weight if it differentiates between instances from different classes and is given the same weight if it belongs to the same class.

In [60] Scherf and Brauer describe a similar instance based approach for assigning feature weights called EUBAFES. In this method, a gradient descent based approach is followed to assign weights to reinforce similarities between instances of the same class and at the same time the similarities between the instances of different class is decreased.

2.6 Neuro-Fuzzy Classifiers

In artificial neural networks, neuro-fuzzy system is the hybridization of fuzzy logic with neural networks. Neuro-fuzzy system was first proposed by J. S. R. Jang. The hybridization of neuro-fuzzy systems result in hybrid intelligent systems which
combines the human like reasoning facility in fuzzy logic with the connectionist structures of neural networks.

Nauck et. al.[61] has developed a new learning method for fuzzy classification rules. The proposed new learning algorithm is a simple heuristics that derives fuzzy rules from the training set quickly and tunes them according to the parameters of membership functions. This approach is based on the NEFCLASS, a neuro-fuzzy model used for pattern classification.

A new adaptive neural network with fuzzy logic for classification is discussed in [62]. In this paper, the membership functions defined for each feature and parameterized t-norms are used along with back propagation. This method uses a supervised learning algorithm using a gradient descent function to update the parameters is used to train the adaptive neural network and the classifier performance is studied for two classification problems: two-spiral classification and Iris classification. The classifier performs well for the Iris classification problem.

Angelov, Plamen, Edwin Lughofer, and Xiaowei Zhou discuss two novel approaches for online evolving fuzzy classifiers [63]. The proposed Fuzzy Rule Based (FRB) classifier adapts its parameters and structure whenever a new instance is added. A special emphasis is placed on methods for improving accuracy and robustness. The method is evaluated with handwritten image recognition and image classification problems and found to work well in online and interactive mode.

An adaptive neuro-fuzzy system to extract fuzzy if then rules from input and output data used for training was studied by Cho, Kwang Bo and Bo Hyeun Wang [64]. The proposed method uses radial basis function based Adaptive neuro Fuzzy System (AFS) and uses Gaussian functions as membership functions to extract the fuzzy rules. The new adaptive system is implemented to study system identification and time series
prediction problems to evaluate the validity and effectiveness of the system using RBF based AFS.

Sarimveis, Haralambos, et al.[65] developed nonlinear dynamic training methodology to train radial basis function neural network based on a fuzzy partition of the input space combined with self-organized and supervised learning. The experimental results show that the methodology is faster and produces more accurate model to train RBF networks and the proposed fuzzy partition is able to determine the proper network structure easily.

In paper [66] the authors present a Fuzzy Hybrid Learning Algorithm (FHLA) for the radial basis function neural network (RBFNN). In this method, the parameters of the hidden units are determined using advanced fuzzy clustering method. The FHLA combines the gradient method and the linear least-squared method for adjusting the RBF parameters and connection weights. The proposed method is tested with face recognition system and was found to provide a faster convergence in the training phase, resulted in less number of hidden neurons. Comparative results show that FHLA yields excellent recognition rate in human face recognition problems.

Sarimveis et. al. [67] developed a new classification algorithm using RBF neural networks that is used exclusively to classify products during manufacturing process. The proposed neural model uses fuzzy means algorithm to classify the input data and is found to be capable of acting as an artificial sensor to classify products based on their quality in real time. The method produced successful results and outperforms a number of other feed forward neural network (FNN) classifiers.

Maglogiannis, Ilias, et al.[68] investigated the potential of using RBF neural network for the classification of biological microscopic images showing lung tissue sections with idiopathic pulmonary fibrosis. The proposed technique uses fuzzy means
clustering algorithm to identify the structure and parameters of the RBF classifier. This technique was compared with the results obtained using Support Vector Machine (SVM) classifier. The comparative result shows that RBF classifier outperforms SVM classification outcome.

A general methodology for constructing fuzzy-neural control model using B-spline membership function is proposed in [69]. By using the local control property of a B-spline membership function (BMF), the local tuning of BMF is possible and this results in the reduced number of iterations needed for training the fuzzy-neuro control model. The effectiveness of the method is studied by generating a fuzzy-neural model to control a car and is found to perform well.

Mitra, Sushmita, and Jayanta Basak [70] designed a fuzzy radial basis function neural network by integrating the principles of radial basis function and fuzzy c-means algorithm. The architecture of the model is modified to incorporate fuzzy concepts to handle both linguistic and numeric inputs and to provide a soft decision output. The effectiveness of the model is studied using a speech recognition problem.

The research paper [71] titled, “survey of neural transfer functions”, elaborates the taxonomy of several activation and output functions. A detailed description of the various activation functions that can be used to implement RBF neural classifiers is also clearly given in this paper.

2.7 Summary

In this chapter a review of various classifiers used for classification problems has been done in view of the literature available. The review shows that multilayer feed forward neural networks are the widely used neural networks for classification task. Compared to multilayer perceptrons, radial basis function neural networks perform well because of their insensitiveness to local minima. RBF neural network does not suffer
from the local minima condition because the only parameter adjusted to train the networks is the linear mapping of weights from hidden layer to output layer which results in a single minimum. The linear mapping ensures that the error surface is quadratic. Many variations of RBFNN like hierarchical RBF neural networks, fuzzy RBF neural networks, and median RBF neural networks have been proposed to make them suitable for the varied problems to which they have been applied.

With an appropriate feature selection and data reduction methods, the RBF neural network can be trained faster and their convergence rate is very quick. The classification performance of the RBF neural classifier can be tuned fine by choosing a proper method to fix the parameters like center and spread used in the learning algorithm. The literature study also reveals that using clustering methods to fix the center and spread results in better classification. To work with uncertain and noisy data, fuzzy logic is proposed in many of the existing work. Taking into consideration all the suggestions given in various papers, a new fuzzy RBF classifier has been modelled and implemented in this research thesis.