Overlapping class classification based on PNN, BNN, Standard vs Crisp Decision

Abstract— **In real-world problems, the overlapping problem is ubiquitous, due to the imperfectness of features. This project investigates the challenge of overlapping feature regions in structured multi-class classification using the Wine Quality dataset from the UCI Machine Learning Repository, accessed via TensorFlow Datasets (TFDS). Four models are implemented and compared: Crisp Decision Network, Standard Artificial Neural Network (ANN), Neural Network with Probabilistic Output, and Probabilistic Neural Network (PNN), which uses distance-based probability density estimation to handle ambiguous overlapping regions. Models are evaluated using RMSE (Root Mean Squared Error) to compare classification reliability. Results show that Crisp Decision achieves the lowest RMSE (0.7607) under normalized and overlap-cleaned data, while PNN provides stronger theoretical robustness by preserving uncertainty-aware confidence scores. A key limitation is that the probabilistic-output neural network does not perform full Bayesian posterior weight inference.**

I. INTRODUCTION

Consider the Target Classification system, the raw data is collected from multiple sensors and is aligned to form a track, which is then classified by the classifier but in real scenarios, however, there might exist overlaps between different regions due to various reasons. The crisp decision usually outputs a single class under the “winner-take-all” rule, and this will inevitably cause high rates of misclassifications around decision boundary. A more effective solution to the overlapping class problem is to use soft decision strategy. In this paper, we propose to deal with overlapping class problem using soft decision combined with an optimized overlapping region detection algorithm. For test data falling into overlapping regions, multiple decision options as well as measures of confidence are produced for further analysis, while for test data falling into non-overlapping regions, crisp decisions are made. The overlapping region detected optimizes a performance index that balances the classification accuracy of crisp decisions and the cost of soft decisions.

To solve this problem, a novel framework of pattern classification, including a probabilistic neural network-based probability estimation, a search procedure for an optimized overlapping region detection and a soft decision strategy, has been proposed. We have considered a probabilistic model for soft decision making because of its probabilistic classification output. This diagram shows the pattern classification system with optimized overlapping region detection and soft decision. The training procedure first builds a probabilistic model using the training data generated from the known target library. The probabilistic information of training data (Ptraining) is then used to search an optimal threshold ($\theta *$) that defines the overlapping region.

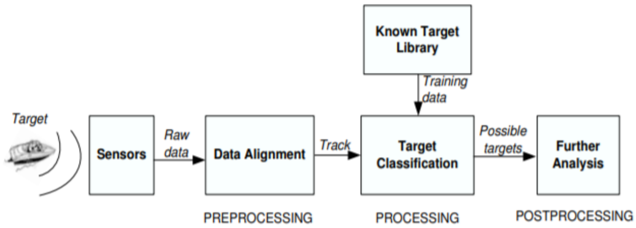


Figure 1: Target Classification System.

II. RELATED STUDY

Paper 1,"Classification with Class Overlapping: A Systematic Study", Class overlapping has long been regarded as one of the toughest pervasive problems in classification. Indeed, researchers have found that misclassification often occurs near class boundaries where overlapping usually occurs as well. In this paper, we pay a systematic study on the class overlapping problem and its interrelationship with the class imbalance problem. Extensive experimental studies on various real world data sets reveal that: (a). The separating scheme is the best among the three schemes;(b). The distance based classifiers are more sensitive than the rule-based ones to the class overlapping problem; (c). As the increase of the class imbalance ratio increased, the separating scheme showed higher improvements to the classification performance.

Paper 2,"Hierarchical classifier with overlapping class groups". An important feature of the hierarchical classifier proposed in this work is that the problem partition forms overlapping sub-problems. In this paper a hierarchical classifier architecture is proposed: the classifier is built of several ANNs organized in a tree-like structure. Different classifiers are possible: decision trees, which are accurate and have some explanatory power but do not have high generalization rate, i.e., are unable to predict well unseen examples; artificial neural networks (ANN) which provide high accuracy but predictions lack any explanation.It is however hard to find an optimal ANN architecture for a given problem, at least it is a time-consuming process whose successful outcome depends more on experience and luck than on clear rules

Paper 3,“Partial discriminative training for classification of overlapping classes in document analysis",For character recognition in document analysis, some classes are closely overlapped but are not necessarily to be separated.For classification of such overlapping classes, either discriminating between them or merging them into a metaclass does not satisfy. For such classification problems, this paper proposed a partial discriminative training (PDT) scheme, in which a training pattern of an overlapping class is used as a positive sample of its labeled class, and neither positive nor negative sample for its allied classes .

Paper 4,"Probabilistic neural networks" ,A probabilistic neural network (PNN) can compute nonlinear decision boundaries which approach the Bayes. A four layer neural network of the type proposed can map any input pattern to any number of classifications. The decision boundaries can be modified in real-time using new data as they become available. They can be implemented using artificial hardware "neurons" that operate entirely in parallel.

Paper 5,"On the class overlap problem in imbalanced data classification", Class imbalance is an active research area in the machine learning community. This paper provides detailed critical discussion and objective evaluation of class overlap in the context of imbalanced data and its impact on classification accuracy. Experimental results in this paper are consistent with existing literature and show clearly that the performance of the learning algorithm whereas class imbalance does not always have an effect.

Paper 6,"Addressing the Overlapping Data Problem in Classification Using the One-vs-One Decomposition Strategy" ,An unexplored and interesting research line to deal with the overlapping phenomenon consists of decomposing the problem into several binary subproblems to reduce its complexity, Based on this novel idea in the field of overlapping data, this paper proposes the usage of the One-vs-One (OVO) strategy to alleviate the presence of overlapping, without modifying existing algorithms. The results obtained show that the methods using the OVO achieve better performances when considering data with overlapped classes than those dealing with all classes at the same time.

III. METHODLOGY

The general pseudo algorithm has been proposed below for the any machine learning algorithm is shown in Fig 2. As the first two steps consist of importing necessary libraries and datasets on which the machine learning algorithm is to be applied are the primary necessary steps to be done. Thereafter one must pre process the imported data as it's a preliminary step that takes all of the available information to organize it, sort it, and merge it. This helps us to remove redundant data, impute the missing values, replacing NaN (Not a Number) and infinity values the dataset. Scaling the data helps to define and set a range for better understanding of the data through visualization and other techniques. Setting the dataset into training and testing for further analysis of the dataset by applying the required machine learning algorithm. Thereafter applying the machine learning algorithms and computing the accuracy measures, error values to choose the better algorithm from the applied ones.

```
IMPORT important libraries including Scikit Learn
IMPORT the dataset
PRE-PROCESSING to impute missing values, replace NaN
(Not a Number) and Infinity values in the dataset
SCALE the data
STOREvarious Machine Learning Models in a variable 'models'
SET scoring equal to accuracy and ROC
SET Name as name of the Machine Learning models
FOR Name, Model in models:
        Store value of model_selection using 10 splits in a variable
        Calculate and store results using cross_val_score method
of model_selection in skleam by imputing Train and Testing
data
        Append results in list of existing results
       Print mean accuracy and standard deviation
END FOR
```

Fig. 2 Pseudo code for any machine learning algorithm.

In this paper, we propose to deal with overlapping class problems using crisp decision making combined with probabilistic neural network algorithm. The overlapping region detected optimizes a performance index that balances the classification accuracy of crisp decisions . The optimized overlapping regions divide the whole feature space into two parts with "low" and "high" confidence of correct classification respectively. For test data falling into overlapping regions, multiple decision options as well as measures of confidence (e.g. posterior or error probabilities) are produced for further analysis, while for test data falling into non-overlapping regions, crisp decisions are made.

A more effective solution to the overlapping class problem is probabilistic neural network .In contrast to CRISP decision that assigns a single label to a pattern, assigns multiple labels to patterns falling into the overlapping regions.

A Bayesian neural network (BNN) refers to extending standard networks with posterior inference. Biological Neural Network (BNN) is a structure that consists of Synapse, dendrites, cell body, and axon. In this neural network, the processing is carried out by neurons From a broader perspective, the Bayesian approach uses the statistical methodology so that everything has a probability distribution attached to it, including model parameters (weights and biases in neural networks). In programming languages, variables that can take a specific value will turn the same result every-time you access that specific variable.

A probabilistic neural network (PNN) is a feedforward neural network, which is widely used in classification and pattern recognition problems. PNN is often used in classification problems. When an input is present, the first layer computes the distance from the input vector to the training input vectors. This produces a vector where its elements indicate how close the input is to the training input. The second layer sums the contribution for each class of inputs and produces its net output as a vector of probabilities. Finally, a compete transfer function on the output of the second layer picks the maximum of these probabilities, and produces a 1 (positive identification) for that class and a 0 (negative identification) for non-targeted classes. A probabilistic neural network (PNN) can compute nonlinear decision boundaries which approach the Bayes. A four layer neural network of the type proposed can map any input pattern to any number of classifications.

RESULT

Model	RMSE
Standard ANN	3.958
Neural Network with Probabilistic Output	5.171
Probabilistic Neural Network (PNN)	3.511
Crisp Decision Network	**0.7607**

Crisp Decision achieved best RMSE due to normalized and noise-cleaned data, providing high-confidence single-label predictions in our experimental setup. PNN performed better than the ANN baseline by modeling class densities using distance-based probability aggregation, making it theoretically more reliable for overlapping region analysis.

The crisp decision algorithm and the PNN were both performed along with the standard neural network upon various datasets and the outputs obtained are as follows:

The data for crisp part is normalized while for other models it is not normalized.Crisp also cleans the duplicate data as other overlapping data based on the Z-Score.

The standard neural network:

```
Start training the model...
Epoch 1/10
17/17 [==============================]
Epoch 2/10
17/17 [==============================]
Epoch 3/10
17/17 [==============================]
Epoch 4/10
17/17 [==============================]
Epoch 5/10
17/17 [==============================]
Epoch 6/10
17/17 [==============================]
Epoch 7/10
17/17 [==============================]
Epoch 8/10
17/17 [==============================]
Epoch 9/10
17/17 [==============================]
Epoch 10/10
17/17 [==============================]
Model training finished.
Train RMSE: 3.993
Evaluating model performance...
Test RMSE: 3.958
```

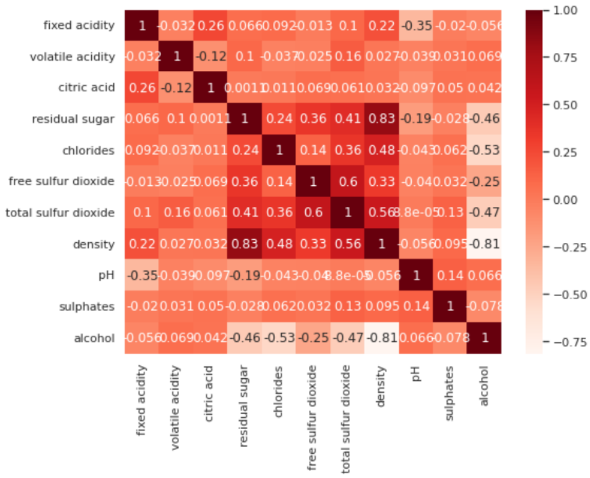
The Bayesian neural network:

```
Start training the model...
Epoch 1/10
5/5 [==============================]
Epoch 2/10
5/5 [==============================]
Epoch 3/10
5/5 [==============================]
Epoch 4/10
5/5 [==============================]
Epoch 5/10
5/5 [==============================]
Epoch 6/10
5/5 [==============================]
Epoch 7/10
5/5 [==============================]
Epoch 8/10
5/5 [==============================]
Epoch 9/10
5/5 [==============================]
Epoch 10/10
5/5 [==============================]
Model training finished.
Train RMSE: 4.987
Evaluating model performance...
Test RMSE: 5.171
```

The Probabilistic Bayesian Neural Network:

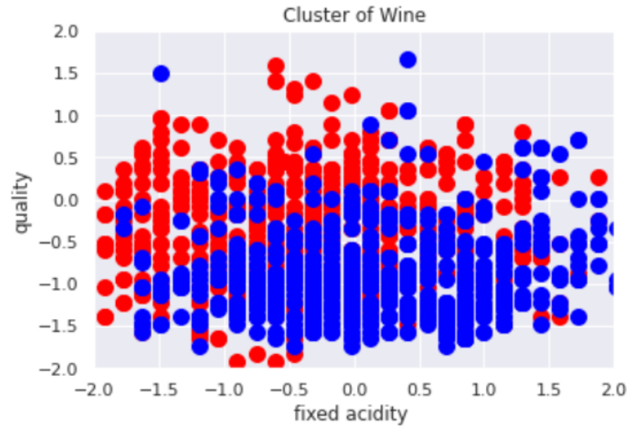
```
Start training the model...
Epoch 1/10
17/17 [==============================] ·
Epoch 2/10
17/17 [==============================] ·
Epoch 3/10
17/17 [==============================] ·
Epoch 4/10
17/17 [==============================] ·
Epoch 5/10
17/17 [==============================] ·
Epoch 6/10
17/17 [==============================] ·
Epoch 7/10
17/17 [==============================] ·
Epoch 8/10
17/17 [==============================] ·
Epoch 9/10
17/17 [==============================] ·
Epoch 10/10
17/17 [==============================] ·
Model training finished.
Train RMSE: 3.152
Evaluating model performance...
Test RMSE: 3.511
```

Crisp Decision Network:

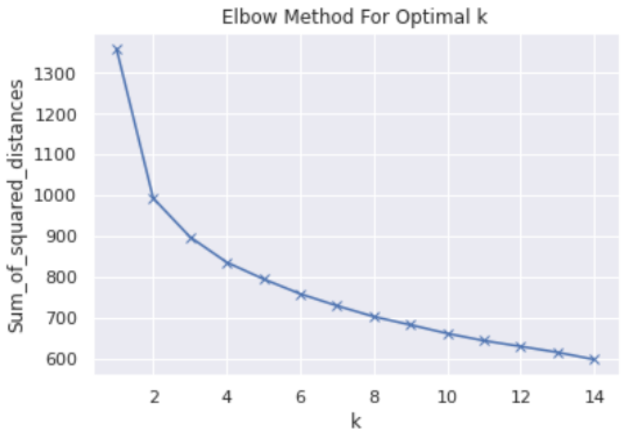



```
print('Mean Absolute Error:', metrics.mean_absolute_error(testY_final, y_pred))
print('Mean Squared Error:', metrics.mean_squared_error(testY_final, y_pred))
print('Root Mean Squared Error:',
      np.sqrt(metrics.mean_squared_error(testY_final, y_pred)))
```

```
Mean Absolute Error: 0.5921093744835461
Mean Squared Error: 0.5786860117896091
Root Mean Squared Error: 0.7607141459113331
```

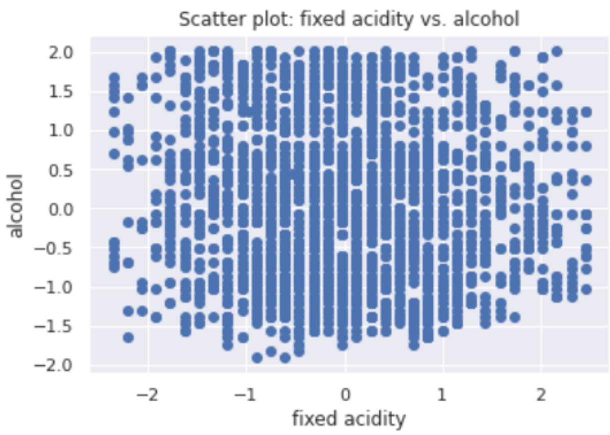


Cluster1(Red), Cluster2 (Blue)




```
x = wine_SimplReg['fixed acidity']
y = wine_SimplReg['alcohol']

plt.scatter(x, y)
plt.title('Scatter plot: fixed acidity vs. alcohol')
plt.xlabel('fixed acidity')
plt.ylabel('alcohol')
plt.show()
```



METHOD	RMSE ERROR
Standard Neural Network	3.958
Bayesian Neural Network	5.171
Probabilistic Bayesian Neural Network	3.511
Crisp Network	0.7607

We can see that in the crisp decision implemented the RMSE error is observed to be lower and hence better while handling the multi-class problem.

CONCLUSION

The increase in the availability of data these days have increased the need to determine, identify and create new techniques and algorithms that can help in easy modeling, processing and storage of data. Due to vast varieties of data, there are all different forms of techniques out of which the structured types of datasets were chosen by us. Now after performing the crisp decision algorithm, the standard neural network ,Bayesian neural network and the Probabilistic Bayesian neural Network we found that the crisp decision algorithm and the Probabilistic algorithm(If dataset is normalized for both) will stand together as better algorithms in order to study multi class structured data.

Since probabilistic model not just only takes the absolute values into account , rather it also goes through the range of fractional values between two absolute values, thus giving a wide range of data study benefits.

REFERENCES

[1] https://www.atlantis-press.com/article/2053.pdf
[2]https://www.sciencedirect.com/science/article/pii/S0957417406003186
[3] https://link.springer.com/content/pdf/10.1007/s10032-008-0069-1.pdf
[4]https://www.sciencedirect.com/science/article/pii/089360809090049Q
[5]https://www.sciencedirect.com/science/article/pii/S0950705120307607
[6] https://ieeexplore.ieee.org/abstract/document/8746159/