Chapter 7
Performance Study on Real-valued Classification Problems

As mentioned in Chapter 5, the orthogonal decision boundaries of fully complex-valued neural networks help them to perform classification tasks efficiently. Therefore, in this chapter, we study the classification performance of FC-MLP and IC-MLP described in Chapter 2, FC-RBF and Mc-FCRBF explained in Chapter 3, FCRN and CC-ELM described in the chapters 5 and 6 respectively. First, the study is conducted on a set of benchmark real-valued classification problems from the UCI machine learning repository [1] and then, using a practical acoustic emission signal classification problem for health monitoring [2].

7.1 Descriptions of Real-valued Benchmark Classification Problems

We consider a set of real-valued benchmark problems (both multi-category/binary classification problems) from the UCI machine learning repository [1]. Based on a wide range of Imbalance Factors (I. F.) (as defined in [3]) of the data set, three multi-category and four binary data sets are chosen for this study. To recap, the imbalance factor is defined as

\[
(I. \ F.) = 1 - \frac{C}{N} \min_{j=1,\ldots,C} N_j
\]

(7.1)

where \(N_j\) is the total number of samples belonging to the class \(j\).

The detailed description of these data sets including the number of classes, the number of input features, the number of samples in the training/testing and the imbalance factor are presented in Table 7.1. From the table, one can see that the problems chosen for this study have both balanced and unbalanced data sets and also that the imbalance factors of the data sets vary over a wide range.
Table 7.1 Description of benchmark data sets selected from [1] for performance study

<table>
<thead>
<tr>
<th>Type of data set</th>
<th>Problem</th>
<th>No. of features</th>
<th>No. of classes</th>
<th>No. of samples</th>
<th>I. F.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Training</td>
<td>Testing</td>
<td>Training</td>
<td>Testing</td>
</tr>
<tr>
<td>Multi-Category</td>
<td>Image Segmentation (IS)</td>
<td>19</td>
<td>7</td>
<td>210</td>
<td>2100</td>
</tr>
<tr>
<td></td>
<td>Vehicle Classification (VC)</td>
<td>18</td>
<td>4</td>
<td>424</td>
<td>422</td>
</tr>
<tr>
<td></td>
<td>Glass Identification (GI)</td>
<td>9</td>
<td>7</td>
<td>109</td>
<td>105</td>
</tr>
<tr>
<td>Binary</td>
<td>Liver Disorder</td>
<td>6</td>
<td>2</td>
<td>200</td>
<td>145</td>
</tr>
<tr>
<td></td>
<td>PLIMA Data</td>
<td>8</td>
<td>2</td>
<td>400</td>
<td>368</td>
</tr>
<tr>
<td></td>
<td>Breast Cancer</td>
<td>9</td>
<td>2</td>
<td>300</td>
<td>383</td>
</tr>
<tr>
<td></td>
<td>Ionosphere</td>
<td>34</td>
<td>2</td>
<td>100</td>
<td>251</td>
</tr>
</tbody>
</table>

7.2 Performance Study

First we present the performance study results on three real-valued multi-category benchmark classification problems. Next, we consider four binary benchmark classification problems.

7.2.1 Performance Measures

The classification/confusion matrix \( Q \) is used to obtain the statistical measures for both the class-level and global performance of the various classifiers. Class-level performance is measured by the percentage classification \( \eta_j \) which is defined as:

\[
\eta_j = \frac{q_{jj}}{N_j} \times 100\% \quad (7.2)
\]

where \( q_{jj} \) is the total number of correctly classified samples in the class \( c_j \).

The global measures used in the evaluation are the average per-class classification accuracy \( \eta_a \) and the over-all classification accuracy \( \eta_o \) defined as:

\[
\eta_a = \frac{1}{C} \sum_{j=1}^{C} \eta_j
\]

\[
\eta_o = \frac{\sum_{j=1}^{C} q_{jj}}{\sum_{j=1}^{C} N_j} \times 100\% \quad (7.3)
\]