**Motivation & Introduction**

Classifier: \( f(X) \rightarrow Y \)
- \( X \): features
- \( Y \): final label
- Combination of all feature all at once not always optimal
- We determine “Feature Clusters” in the data space (as opposed to data clusters in feature space)

**Algorithm**

```
\text{Input}: \text{Features } \{x_1, x_2, \ldots, x_n\}, \text{ Accuracy Threshold } \epsilon, \text{ Number of Clusters } K
\text{Output}: \text{Feature Clusters } \{X_1, X_2, \ldots, X_K\}
\begin{align*}
\text{begin} & \\
\text{for } i = 1, 2, \ldots, K \text{ do} & \\
X_i & = \{x_i\} \text{; } Y_i = f(x_i) \text{; } \text{score} = 0 \text{; } \text{size} = 0 \\
\text{end} & \\
\text{for } x \in X_i \text{ do} & \\
Y(x) & = f(D(x)) \\
\text{end} & \\
\text{end} & \\
\text{Output } & \{X_1, X_2, \ldots, X_K\}
\end{align*}
```

- Forward feature selection with different initial “seeds”
- Feature clusters optimize accuracy
- Last feature cluster collection of all unselected features
- Feature clusters span entire feature space

**Classifier fusion schemes**

- Simple plurality
- Entropy based selection
- Log-likelihood summation
- Weighted log-likelihood summation
  - Weights determined by gradient ascent on accuracy function
- Entropy based weighted log-likelihood summation
- Confidence region based log-likelihood summation

**Description**

Higher Dimensional space broken down into lower dimensions

**Results**

<table>
<thead>
<tr>
<th>Database</th>
<th>Dataset</th>
<th>No. of Points</th>
<th>No. of Features</th>
<th>No. of Classes</th>
<th>Baseline Accuracies</th>
<th>Fusion Accuracies</th>
<th>Weighted Log Likelihood Summation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wine</td>
<td>Q1</td>
<td>238</td>
<td>8</td>
<td>3</td>
<td>71.70</td>
<td>70.56</td>
<td>70.79</td>
</tr>
<tr>
<td>Iris</td>
<td>Q1</td>
<td>150</td>
<td>8</td>
<td>3</td>
<td>88.55</td>
<td>88.24</td>
<td>88.40</td>
</tr>
<tr>
<td>KDD</td>
<td>Q1</td>
<td>224</td>
<td>36</td>
<td>2</td>
<td>91.65</td>
<td>91.63</td>
<td>91.65</td>
</tr>
<tr>
<td>HAMACHF</td>
<td>Q1</td>
<td>681</td>
<td>11</td>
<td>3</td>
<td>88.33</td>
<td>88.30</td>
<td>88.30</td>
</tr>
</tbody>
</table>

**Conclusion and Future work**

- Performance of fusion scheme depends on the database
- Feature clusters have better classification capability as compared to all the features taken together

**Future work**

- More feature clustering schemes (based on correlation, mutual information)
- Finding fusion schemes given the database characteristics.