Interpretable Rules in Relaxed Logical Form

Bishwamittra Ghosh
ML algorithms continue to permeate critical application domains
  ▶ medicine
  ▶ legal
  ▶ transportation
  ▶ ...

It becomes increasingly important to
  ▶ understand ML decisions
  ▶ interact with ML solutions

Interpretability has become a central thread in ML research
ML predictions in the form of **rules** are arguably more interpretable.

- Decision lists
- Decision trees
- Decision rules (CNF/DNF)
A CNF (Conjunctive Normal Form) formula is a conjunction of clauses where each clause is a disjunction of literals.

A DNF (Disjunctive Normal Form) formula is a disjunction of clauses where each clause is a conjunction of literals.

Example

- CNF: \((a \lor b \lor c) \land (d \lor e)\)
- DNF: \((a \land b \land c) \lor (d \land e)\)
Example of CNF classification rules

A sample is Iris Versicolor if
(sepal length > 6.3 OR sepal width > 3 OR petal width ≤ 1.5 )
AND
(sepal width ≤ 2.7 OR petal length > 4 OR petal width > 1.2)
AND
(petal length ≤ 5)
Key Contribution

- generalize the widely popular CNF rules
- introduce relaxed-CNF rules
Definition of Relaxed-CNF formula

- Relaxed-CNF formula has two extra parameters $\eta_l$ and $\eta_c$
- A clause is satisfied if at least $\eta_l$ literals are satisfied
- A formula is satisfied if at least $\eta_c$ clauses are satisfied

more restriction on literals, less restriction on clauses
Relaxed-CNF rule for Breast Cancer Prediction

Tumor is diagnosed as malignant if,

\[
[( \text{smoothness} \geq 0.089 + \text{standard error of area} \geq 53.78
+ \text{largest radius} \geq 18.225) \geq 2]
\]

\[
+ [(98.76 \leq \text{perimeter} < 114.8 + \text{largest smoothness} \geq 0.136 + 105.95 \leq \text{largest perimeter} < 117.45) \geq 2] \geq 1
\]
Benefit of Relaxed-CNF

- Relaxed-CNF is more succinct than CNF
- Relaxed-CNF has similar interpretability/expressiveness as CNF
- Smaller relaxed-CNF rules reach the same level of accuracy compared to plain CNF/DNF rules and decision lists
**IRR: Interpretable Rules in Relaxed Form**

- We formulate an Integer Linear Program (ILP) for learning relaxed rules
- We incorporate incremental learning in ILP formulation to achieve scalability
## Accuracy of relaxed-CNF rules and other classifiers

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Size</th>
<th>Features</th>
<th>NN</th>
<th>SVC</th>
<th>RF</th>
<th>RIPPER</th>
<th>BRS</th>
<th>IMLI</th>
<th>IRR</th>
<th>inc-IRR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heart</td>
<td>303</td>
<td>31</td>
<td>83.6</td>
<td>85.48</td>
<td>83.87</td>
<td>81.59</td>
<td>80.65</td>
<td>80.65</td>
<td><strong>86.65</strong></td>
<td>86.44</td>
</tr>
<tr>
<td>WDBC</td>
<td>569</td>
<td>88</td>
<td>96.49</td>
<td>98.23</td>
<td>96.49</td>
<td>96.49</td>
<td>97.35</td>
<td>96.46</td>
<td>97.34</td>
<td>96.49</td>
</tr>
<tr>
<td>ILPD</td>
<td>583</td>
<td>14</td>
<td>71.56</td>
<td>71.19</td>
<td>71.19</td>
<td>72.41</td>
<td>66.67</td>
<td>71.31</td>
<td>69.57</td>
<td><strong>74.14</strong></td>
</tr>
<tr>
<td>Pima</td>
<td>768</td>
<td>30</td>
<td>79.22</td>
<td>77.13</td>
<td>78.57</td>
<td>77.27</td>
<td>77.92</td>
<td>74.51</td>
<td><strong>78.57</strong></td>
<td>77.27</td>
</tr>
<tr>
<td>Tic Tac Toe</td>
<td>958</td>
<td>27</td>
<td>87.5</td>
<td>98.44</td>
<td>99.47</td>
<td>98.44</td>
<td><strong>100</strong></td>
<td>82.72</td>
<td>84.37</td>
<td>84.46</td>
</tr>
<tr>
<td>Titanic</td>
<td>1309</td>
<td>26</td>
<td>77.1</td>
<td>78.54</td>
<td>79.01</td>
<td>78.63</td>
<td>77.78</td>
<td>79.01</td>
<td><strong>81.22</strong></td>
<td>78.63</td>
</tr>
<tr>
<td>Tom’s HW</td>
<td>28179</td>
<td>910</td>
<td>—</td>
<td>97.6</td>
<td>97.46</td>
<td>97.6</td>
<td>—</td>
<td>96.01</td>
<td>97.34</td>
<td>96.52</td>
</tr>
<tr>
<td>Credit</td>
<td>30000</td>
<td>110</td>
<td>80.69</td>
<td>82.17</td>
<td>82.12</td>
<td>82.13</td>
<td>—</td>
<td>81.75</td>
<td><strong>82.15</strong></td>
<td>81.94</td>
</tr>
<tr>
<td>Adult</td>
<td>32561</td>
<td>144</td>
<td>84.72</td>
<td>87.19</td>
<td>86.98</td>
<td>84.89</td>
<td>—</td>
<td>83.63</td>
<td><strong>85.23</strong></td>
<td>83.14</td>
</tr>
<tr>
<td>Twitter</td>
<td>49999</td>
<td>1511</td>
<td>—</td>
<td>—</td>
<td>96.48</td>
<td><strong>96.14</strong></td>
<td>—</td>
<td>94.57</td>
<td>95.44</td>
<td>93.22</td>
</tr>
</tbody>
</table>
## Rule-size of different interpretable models

<table>
<thead>
<tr>
<th>Dataset</th>
<th>RIPPER</th>
<th>BRS</th>
<th>IMLI</th>
<th>inc-IRR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heart</td>
<td>7</td>
<td>35.5</td>
<td>14</td>
<td>19.5</td>
</tr>
<tr>
<td>WDBC</td>
<td>7</td>
<td>18</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>ILPD</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Pima</td>
<td>8</td>
<td>8</td>
<td>15</td>
<td>21.5</td>
</tr>
<tr>
<td>Tic Tac Toe</td>
<td>25</td>
<td>24</td>
<td>11.5</td>
<td>12</td>
</tr>
<tr>
<td>Titanic</td>
<td>5</td>
<td>2</td>
<td>7</td>
<td>12.5</td>
</tr>
<tr>
<td>Tom’s HW</td>
<td>16.5</td>
<td>—</td>
<td>32</td>
<td>5.5</td>
</tr>
<tr>
<td>Credit</td>
<td>33</td>
<td>—</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td>Adult</td>
<td>106</td>
<td>—</td>
<td>35.5</td>
<td>13</td>
</tr>
<tr>
<td>Twitter</td>
<td>56</td>
<td>—</td>
<td>67.5</td>
<td>7</td>
</tr>
</tbody>
</table>
Effect of threshold parameter

- Rule Size
  - threshold, $\eta_l$
  - 7.5
  - 10.0
  - 12.5
  - 15.0
  - 17.5

- Test Acc %
  - threshold, $\eta_l$
  - 78
  - 79
  - 80
  - 81
Effect of data-fidelity parameter

![Graph showing the effect of data-fidelity parameter on rule size and test accuracy.](image)

- **Rule Size** increases as the fidelity, \( \lambda \), increases.
- **Test Acc %** also increases with increasing fidelity, \( \lambda \).
Effect of partitioning

- Rule Size
  - #partition, $\tau$
  - 1 4 8 16 32

- Time (s)
  - #partition, $\tau$
  - 0 500 1000 1500

- Test Acc %
  - #partition, $\tau$
  - 70.0 72.5 75.0 77.5 80.0
Conclusion

- Relaxed-CNF rules allow increased flexibility to fit data
- The size of relaxed-CNF rule is less for larger datasets, indicating higher interpretability
- Relaxed-CNF rule can be applied to various applications, for example checklists