Incremental Approach to Interpretable Classification Rule Learning

Bishwamittra Ghosh and Kuldeep S. Meel
School of Computing, National University of Singapore

Interpretable Machine Learning

- The wide adoption of machine learning in the critical domains has propelled the need for interpretable techniques
- Interpretable machine learning model provides end users the reasoning behind decision-making
- We propose an incremental approach for learning interpretable classification rules

Examples of Interpretable Rules

- **Credit Default Dataset**
  A client payment defaults if
  Repayment status in September: payment delay $>$ 1 month OR
  Repayment status in August: payment delay $>$ 2 months OR
  Repayment status in June: payment delay $>$ 2 months OR
  Education type = other
- **Pima Indian Diabetes Dataset**
  A person is tested positive for diabetes if
  Plasma glucose concentration $>$ 125 AND
  Triceps skin fold thickness $\leq$ 35 mm AND
  Diabetes pedigree function $>$ 0.259 AND
  Age $>$ 25 years

Learning Interpretable Classification Rules

- In rule-based classifiers, sparsity refers to the interpretability of the rule, i.e., a sparser rule is more interpretable [1]
- Consider decision variables:
  - $b_i^j = \mathbb{1}\{j\text{-th feature is selected in } i\text{-th clause}\}$
  - $\eta_q = \mathbb{1}\{\text{sample } q \text{ is misclassified}\}$
- To learn an interpretable classification rule, the objective function is:

$$\min \sum_{i,j} b_i^j + \lambda \sum_q \eta_q$$

- Constraints: positive labeled samples satisfy the rule, and negative labeled samples do not satisfy the rule
- In MaxSAT, the objective function is encoded as soft clauses and the constraints are encoded as hard clauses

Analysis

- To generate a k-clause CNF rule for a dataset of n samples over m boolean features, the number of clauses of the MaxSAT instance is $\mathcal{O}(n \cdot m \cdot k)$
- Suffers from poor scalability when dataset is large

References

[1] Dmitry Malioutov and Kuldeep S Meel. MLIC: A MaxSAT-based framework for learning interpretable classification rules. In *Proc. of CP*, 2018.

[2] Bishwamittra Ghosh and Kuldeep S. Meel. IMLI: An incremental framework for MaxSAT-based learning of interpretable classification rules. In *Proc. of AIES*, 2019.

An Incremental Rule-learning Approach

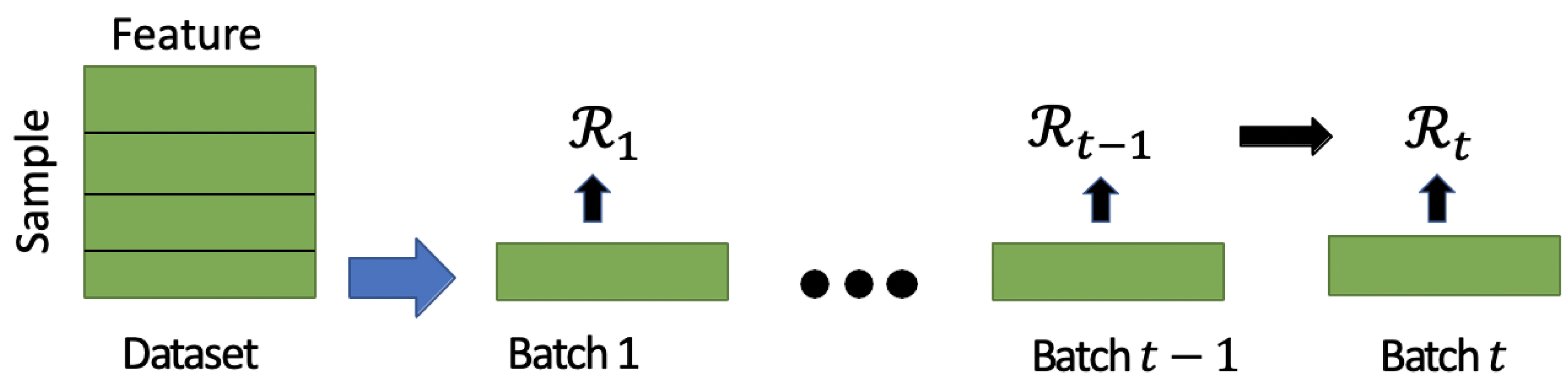
- We attribute large formula size of the MaxSAT instance for the poor scalability
- We propose a mini-batch incremental learning framework with the following objective function

$$\min \sum_{i,j} b_i^j \cdot I(b_i^j) + \lambda \sum_q \eta_q.$$

where indicator function $I(\cdot)$ is defined as follows.

$$I(b_i^j) = \begin{cases} -1 & \text{if } b_i^j = 1 \text{ in the } (t-1)\text{-th batch } (t \neq 1) \\ 1 & \text{otherwise} \end{cases}$$

Solution Technique



- Divide the training data into a fixed number of batches p
- The MaxSAT instance constructed for the t-th batch depends on the training data in the $(t-1)$-th batch and the rule learned in the $(t-1)$-th batch
- Construct soft unit clauses to encode the learned assignment of decision variables in the previous batch

Key Contribution

IMLI makes p queries to the MaxSAT solver with each query of the formula size $\mathcal{O}(\frac{n}{p} \cdot m \cdot k)$

Experimental Results

Dataset	Size n	Features m	LR	SVC	RIPPER	IMLI [2]
PIMA	768	134	75.32 (0.3)	75.32 (0.37)	75.32 (2.58)	73.38 (0.74)
Credit-default	30000	334	80.81 (6.87)	80.69 (847.93)	80.97 (20.37)	79.41 (32.58)
Twitter	49999	1050	95.67 (3.99)	Timeout	95.56 (98.21)	94.69 (59.67)

Table 1: Each cell in the last 5 columns refers to test accuracy (%) and training time (s). IMLI exhibits better training time by costing a little bit of accuracy.

Dataset	RIPPER	IMLI
PIMA	8.25	3.5
Twitter	21.6	6
Credit	14.25	3

Table 2: Average size of the rules of different rule-based models. IMLI generates shorter rules compared to other rule-based models.

Conclusion

- IMLI achieves up to three orders of magnitude improvement in training time by sacrificing a bit of accuracy
- The generated rules appear to be more interpretable