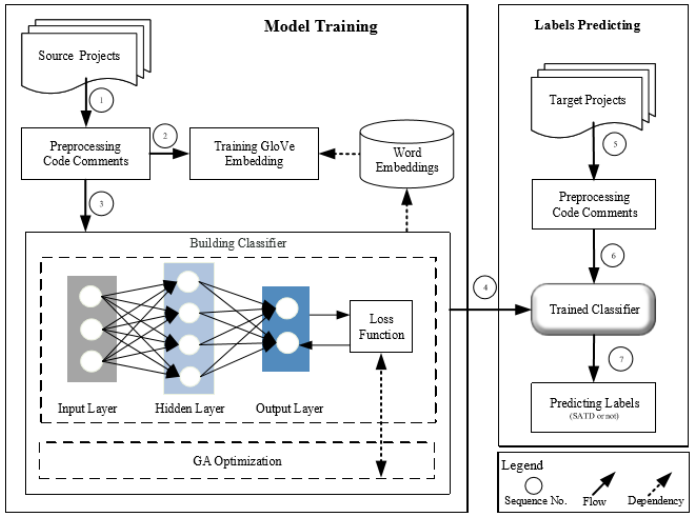
Using BiLSTM with attention mechanism to automatically detect self-admitted technical debt

Dongjin YU, Lin WANG, Xin CHEN, Jie CHEN

Frontiers of Computer Science, DOI: 10.1007/s11704-020-9281-z

Problems & Ideas

- **Problems of detecting self-admitted technical debt**
  - Low detection accuracy
  - Imbalanced positive/negative sample numbers
  - Sharply varied comment lengths



The overall framework of our approach

- **Ideas: Automatically identify self-admitted technical debt**
  - Building a classifier to automatically identify self-admitted technical debt.
  - Adopting a balanced cross entropy loss function to overcome the class imbalance problem.
  - Applying genetic algorithms to adaptively learn the balancing factor of the loss function for improving the expansibility of our approach .

Main Contributions

- We propose a new approach based on BiLSTM with the attention mechanism to automatically identify self-admitted technical debt.
- We conduct extensive experiments to evaluate the performance of the proposed approach and compare our approach with baseline methods on a public dataset. The results show that our approach outperforms these methods.

F1-score of all methods on the 10 projects

Project	Ours	Patterns	NLP	Text Mining
Ant	**61.22%**	14.08%	52.10%	47.96%
JMeter	**84.81%**	8.28%	79.00%	81.69%
ArgoUML	**87.54%**	4.40%	86.80%	83.73%
Columba	**91.21%**	10.53%	85.90%	84.36%
EMF	**60.62%**	8.89%	44.70%	50.95%
Hibernate	**82.83%**	12.23%	80.00%	79.01%
JEdit	**55.87%**	26.71%	51.70%	43.93%
JFreeChart	**73.14%**	8.70%	67.00%	69.46%
JRuby	**87.36%**	8.69%	84.00%	83.98%
SQuirrel	**73.98%**	7.44%	62.90%	67.15%
Average	**75.86%**	10.99%	69.41%	69.22%