OpenEnsemble System and Its Reputation Mechanism

Jinhan Kim

11th March 2011

Part I

Traditional ensemble learning

- Fixed set of base classifiers
- Fixed set of original data set
- Simple aggregation mechanism
  - Focus on more diverse and accurate base classifiers
- Classification
- Eg) Boosting, Bagging and Random Forest

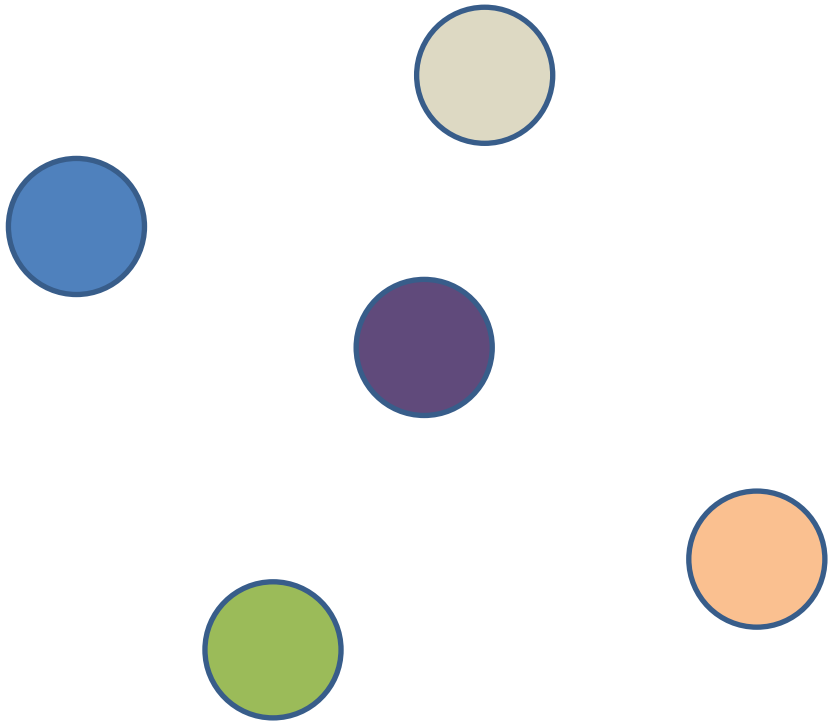
Environment changed

- We live in the open world
  - Unbounded set of classifiers - simple sensors or even humans
  - Someone appears and someone disappears.
- Data is changing over time.

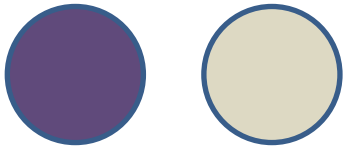
Service Choreography System

- Under the changed environment
- OpenKnowledge project
- We can program interaction models (workflows) to be aggregation methods.
  - IM0: accuracy (general interaction model)
  - IM1: F-measure
  - IM2: AUC
  - IM3: quick classification
  - IM4: microRNA classification
  - IM5 …

Contact Problem

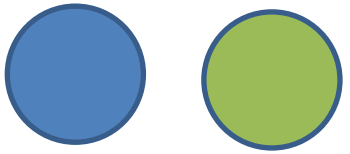


Attended to IM



Not good classification performance

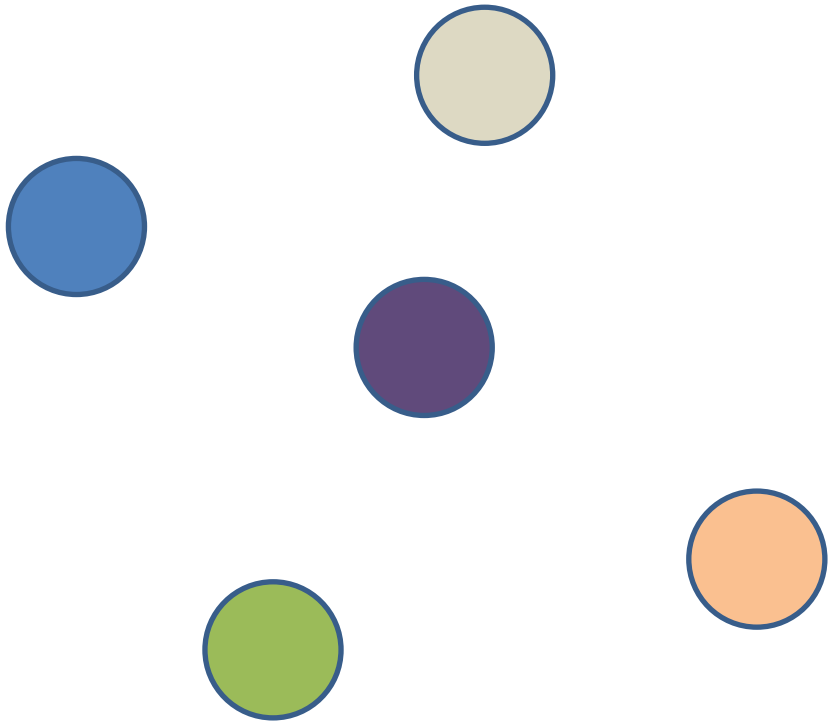
Good Ones



Good classification performance

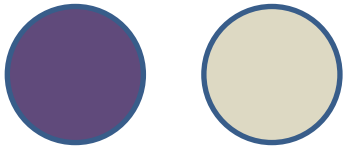
- When apply an IM, we cannot expect a good classification results from classifiers that attend to the IM.

The missing link – Reputation mechanism



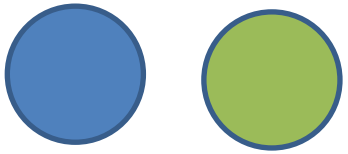
R

Attended to IM



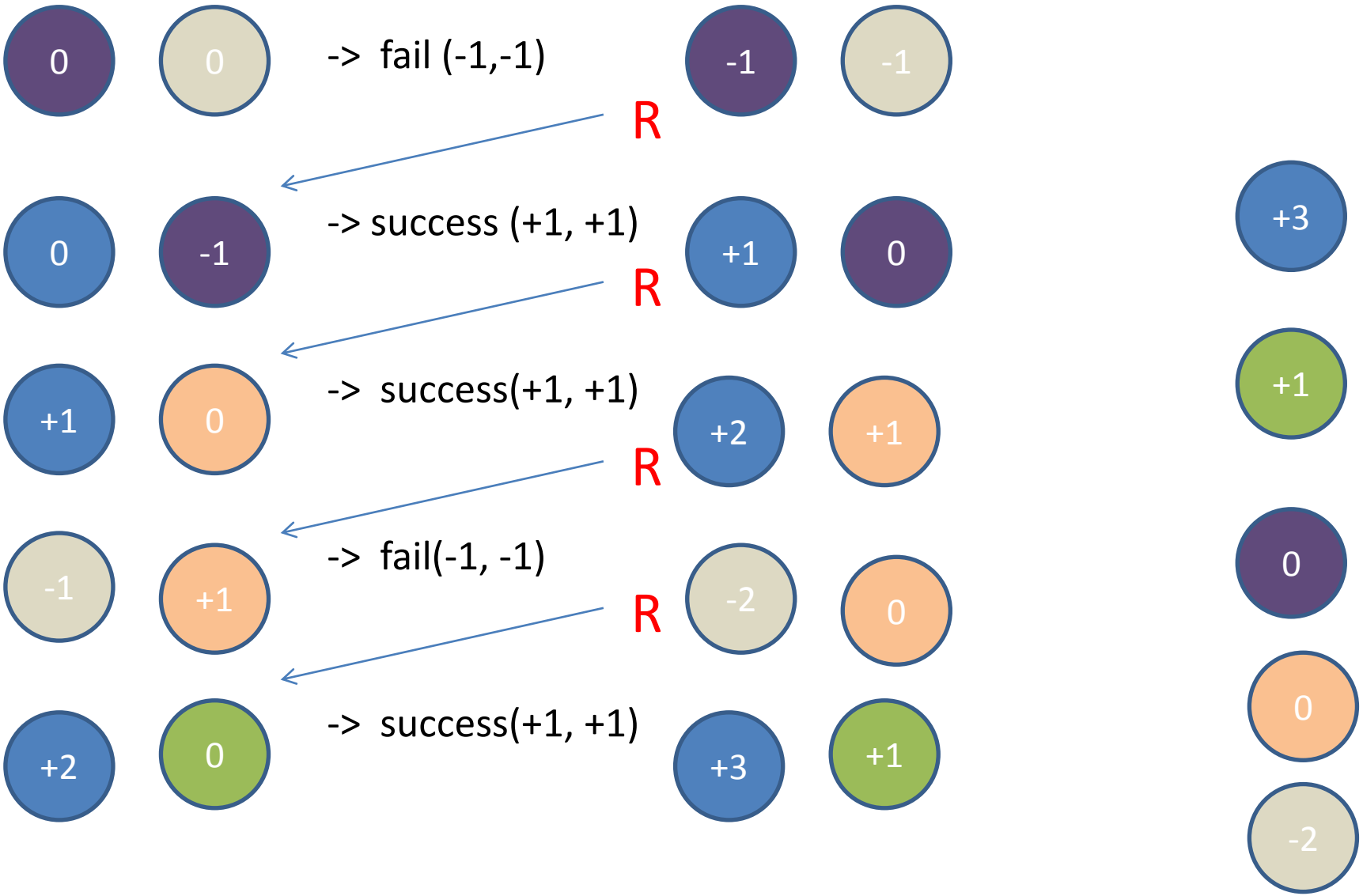
Not good classification performance

Good Ones



Good classification performance

Higher score and avaiable



Research hypothesis

- Traditional ensemble learning > OpenEnsemble system without the reputation mechanism (OE_null)
- Traditional ensemble learning == (<) OpenEnsemble system with the reputation mechanism (OE_R)
- We can define a general and automatic reputation mechanism for ensemble classification.

Reputation mechanism should statisfy

- {OE}_R's performances approaches TE
  - IM_0, IM_f, $IM_{U,..}$
- {OE}_R approahces TE in practical time
  - IM_q
- {OE}_R applicable to realistic problems
  - IM_p, IMi
- R needs minimal or no parameterisation
  - R

Part II

Experiments

- Interaction models: LCC
- Classifiers: Weka machine learning framework

Table 1. Datasets. The sizes of examples are for one fold from total five folds of a dataset.

Dataset	Examples	Training	Query	Test	Attr.	Ratio of Class
spambase	4601	2761	920	920	58	2788:1813
sick	3772	2264	754	754	30	3541:231
kr-vs-kp	3196	1918	639	639	37	1669:1527
credit-g	1000	600	200	200	21	700:300
tic-tac-toe	958	576	191	191	10	332:626
diabetes	768	468	153	153	9	500:268

- Ensemble pool size: 50 base classifiers
- Ensemble size: 3
- Accuracy, F-measure and AUC
- IM0 (general for accuracy)

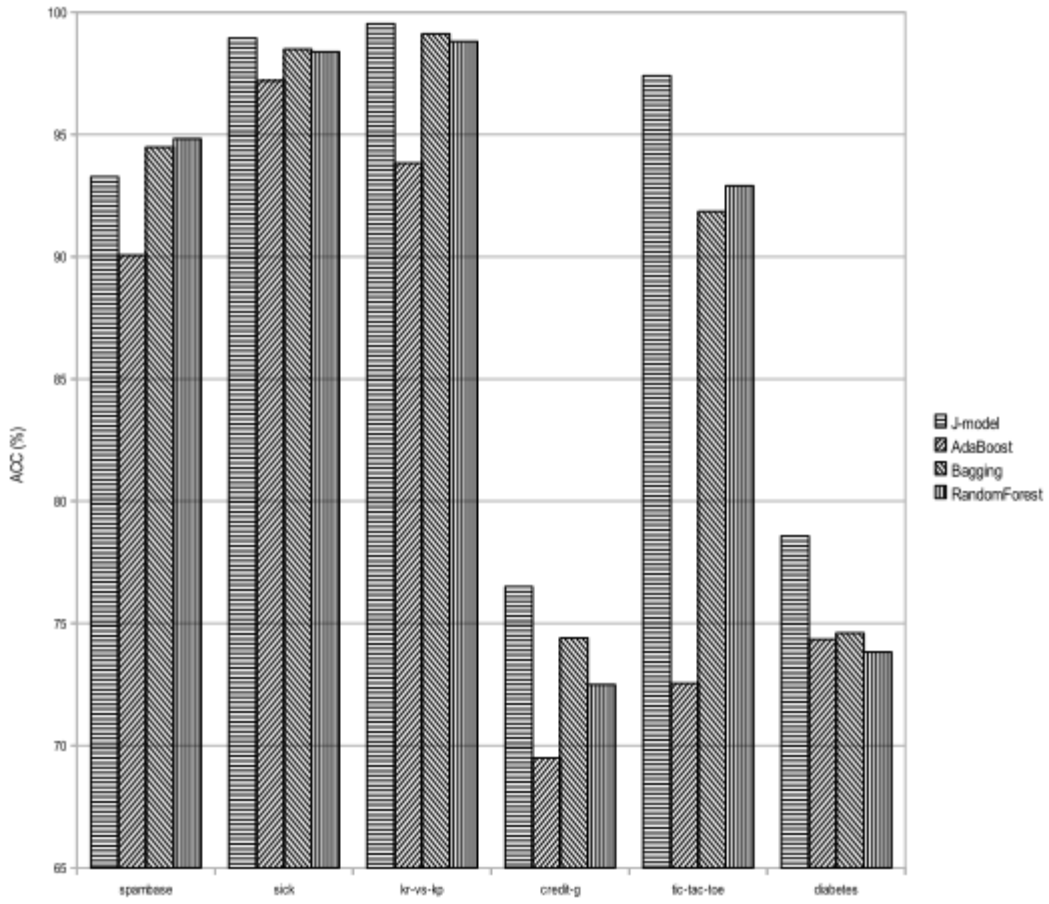


Figure 9. Accuracy comparison

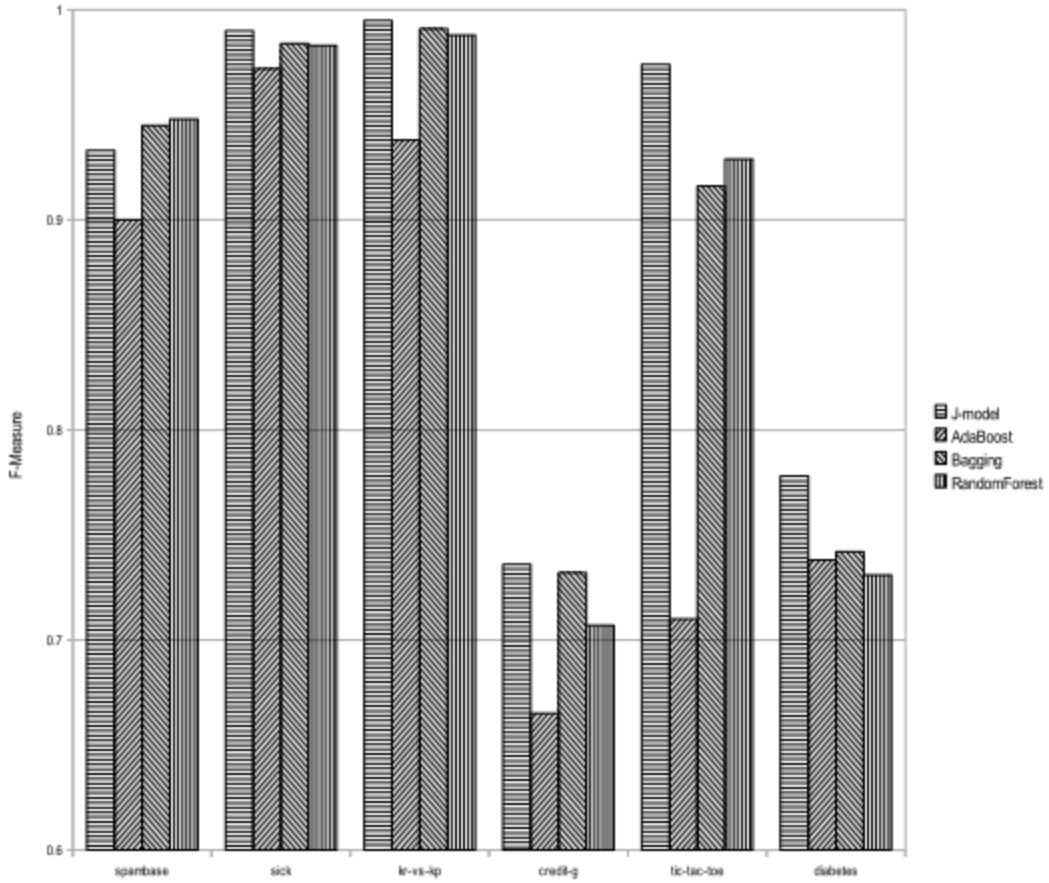


Figure 10. F-measure comparison

Comparison Using ROC Area

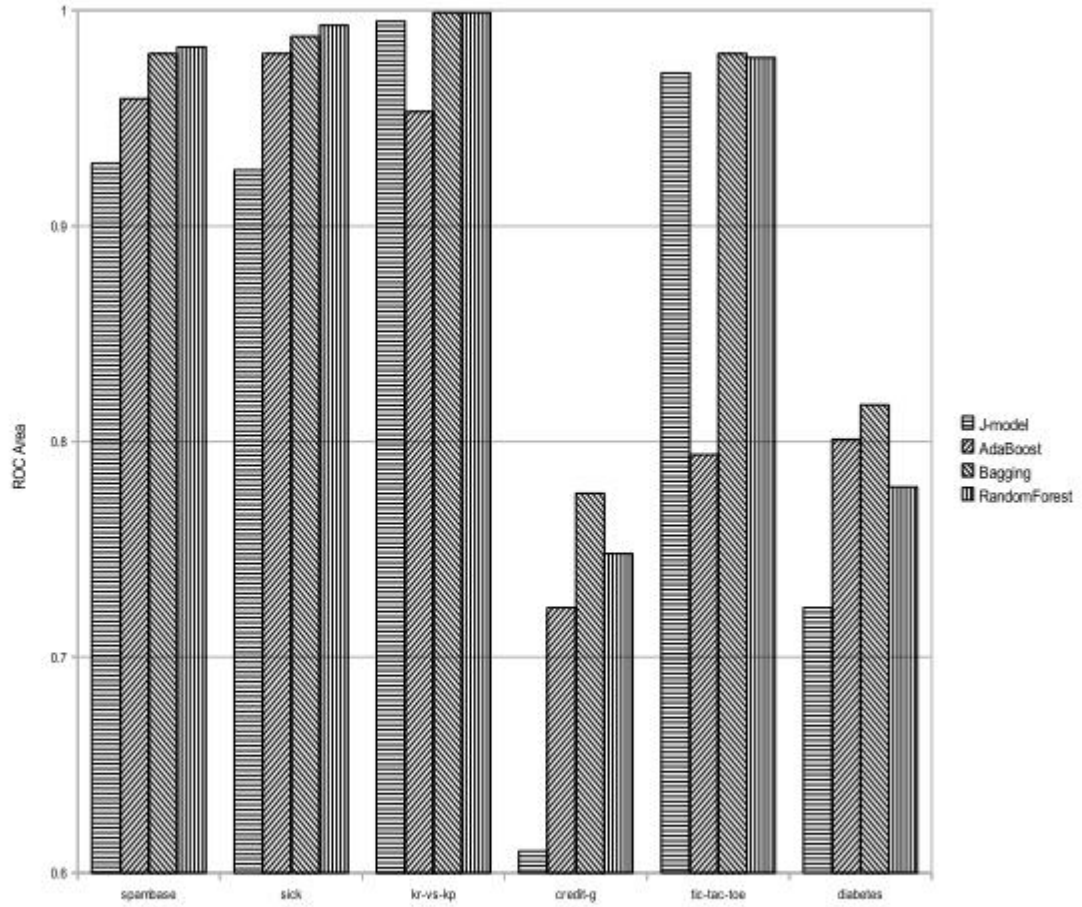


Figure 11. Accuracy comparison

Thank you, DIR group