Analysis of Instance Hardness in Machine Learning Using Item Response Theory

Ricardo B. C. Prudêncio
– Centro de Informática, Recife, Brazil

José Hernández-Orallo and
Adolfo Martínez-Usó
-DSIC, Universitat Politècnica de València, Spain

Content

- Instance hardness
- Item Response Theory
- IRT for Instance-Wise Evaluation
- Case Study
- Conclusions

Instance Hardness

- Hard instances are those for which learning models have a low probability of predicting the correct class label
- Analysis of instance hardness can portray valuable insights about learning algorithms
- Connected to different topics like instance selection, noise reduction, active learning,...

Instance Hardness

- Instance hardness as the average behaviour of a pool of classifiers (Smith et al. 2014)
  - E.g., Average error

- Important information about instance difficulty can be missed
  - E.g., Which classifiers produced correct answers for that instance?

Item Response Theory (IRT)

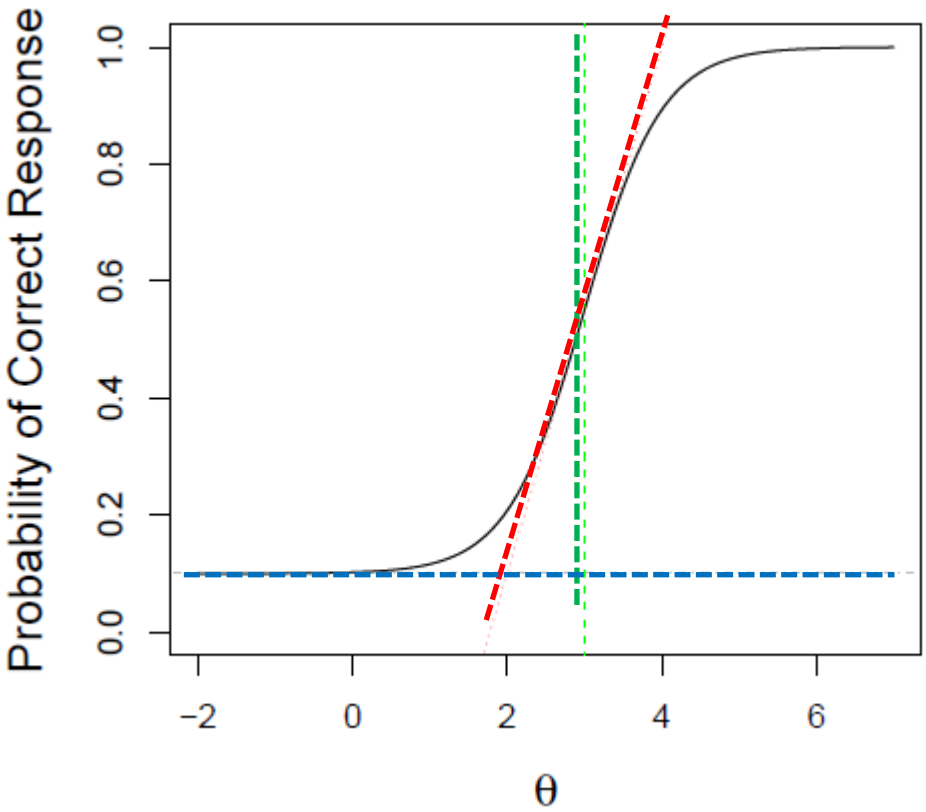
- Relates probability of correct response of items and latent skills of the respondents

- Educational testing and psychometric evaluation
  - E.g., Student's ability vs question's difficulty

- Ability = level of hard items a respondent is able to solve

Item Characteristic Curve – 3P IRT Model

$$P(U_{ij}=1|\theta_j)=c_i+\frac{1-c_i}{1+\exp(-a_i(\theta_j-b_i))}$$



Respondent ability θ_j

Item's parameters

b_i (**Difficulty or location**)

a_i (**Discrimination**)

c_i (**Guessing**)

Item Response Theory (IRT)

- Dual relationship:
  - Difficult items are those ones only solved by the most skilled respondents
  - Skilled respondents in turn are those ones who solve the hard questions

- Both item's parameters and abilities are usually unknown

Item Response Theory (IRT)

- Birnbaum's estimation method
  - Step (1) Start with some initial values for abilities θ_j and estimate the item's parameters;
  - Step (2) Adopt the estimated parameters in the previous step as known values and estimate the abilities θ_j

OBS.: Hence, unknown abilities and items' parameters are estimated simultaneously

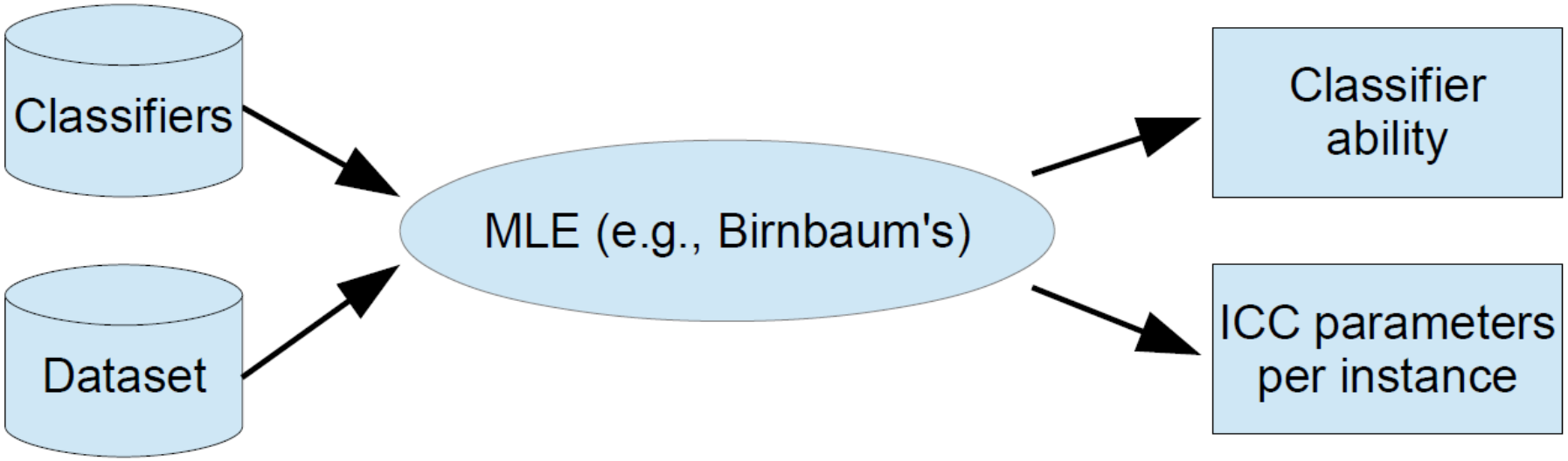
IRT for Instance-Wise Evaluation

- IRT models to evaluate classifiers
  - Items → instances
  - Respondents → Classifiers

  - Items' parameters → Instance hardness
  - Ability → Classifier performance

- ICC plots the probability of success on the instance, given the ability of the classifier

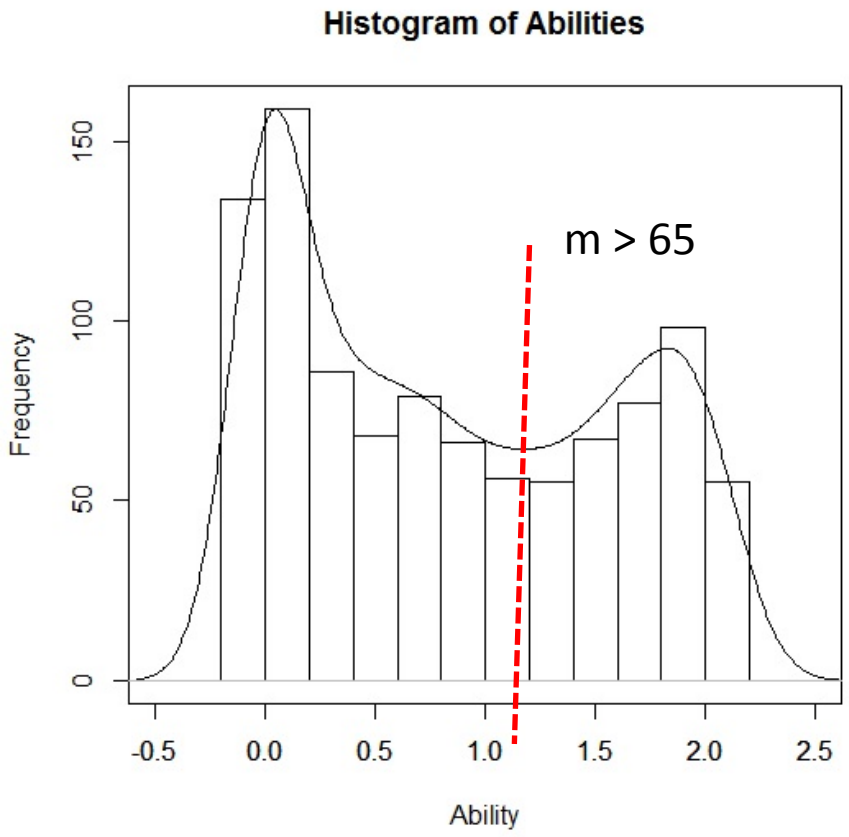
IRT for Instance-Wise Evaluation

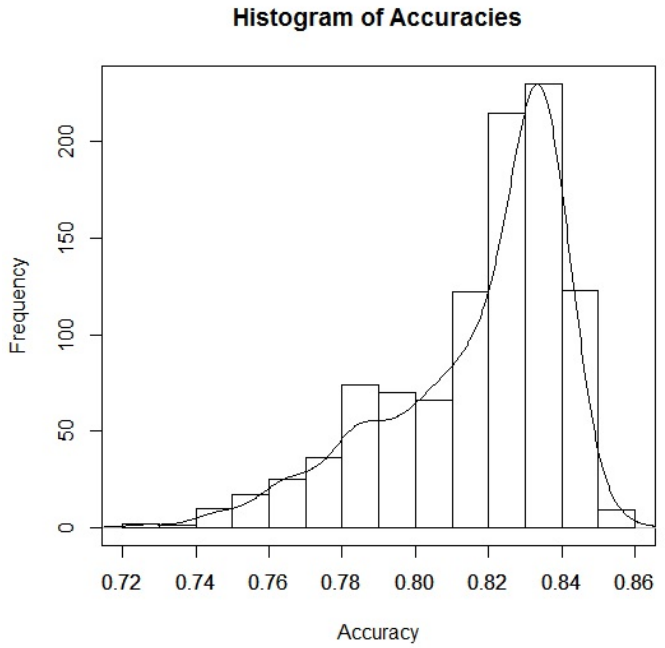


Case Study

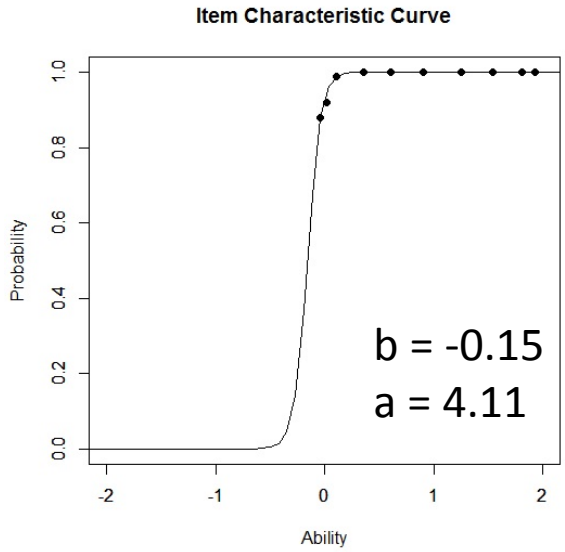
- Random Forests with different number of trees
  - $1 + 2^m$, m = 1,..., 10
  - 100 runs
- Heart dataset
  - 270 instances
- 270 IRT models and 1000 classifiers
  - 2P model (no guessing parameter)

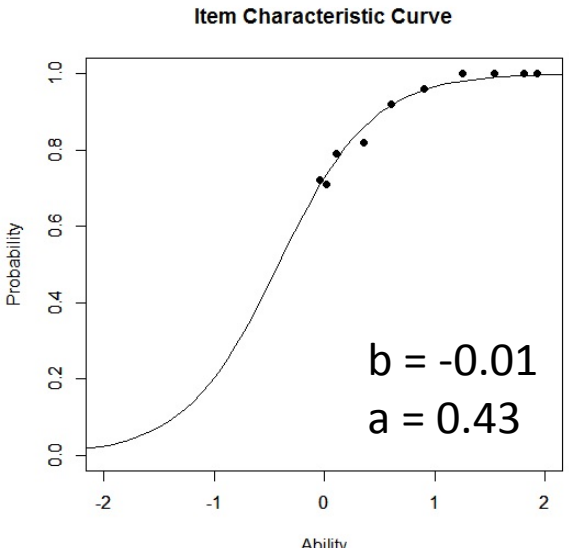
Classifier Ability

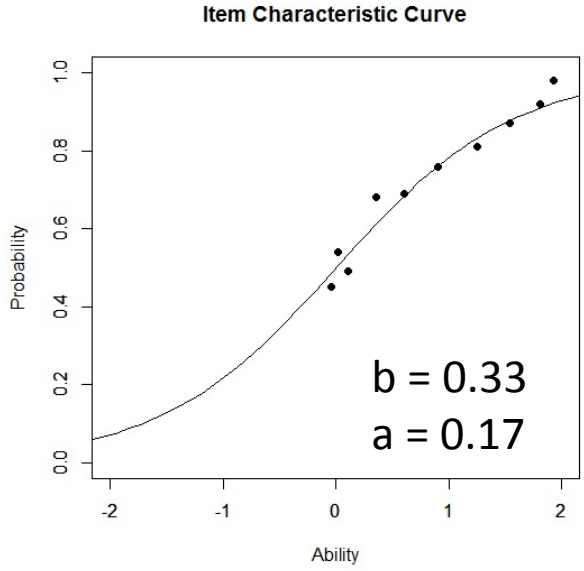


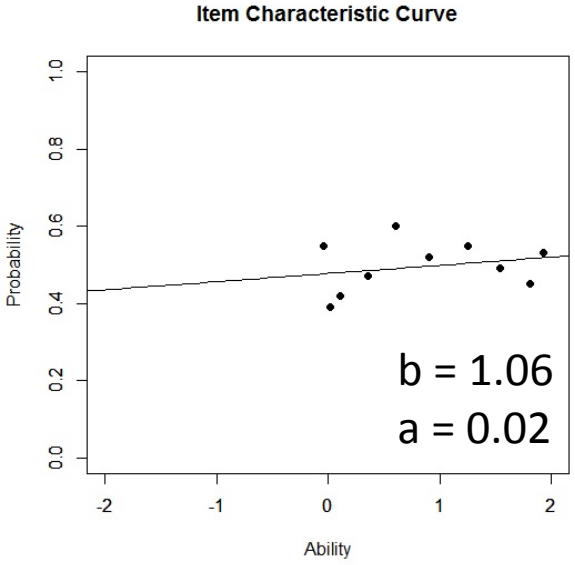


Item Characteristic Curves - Examples

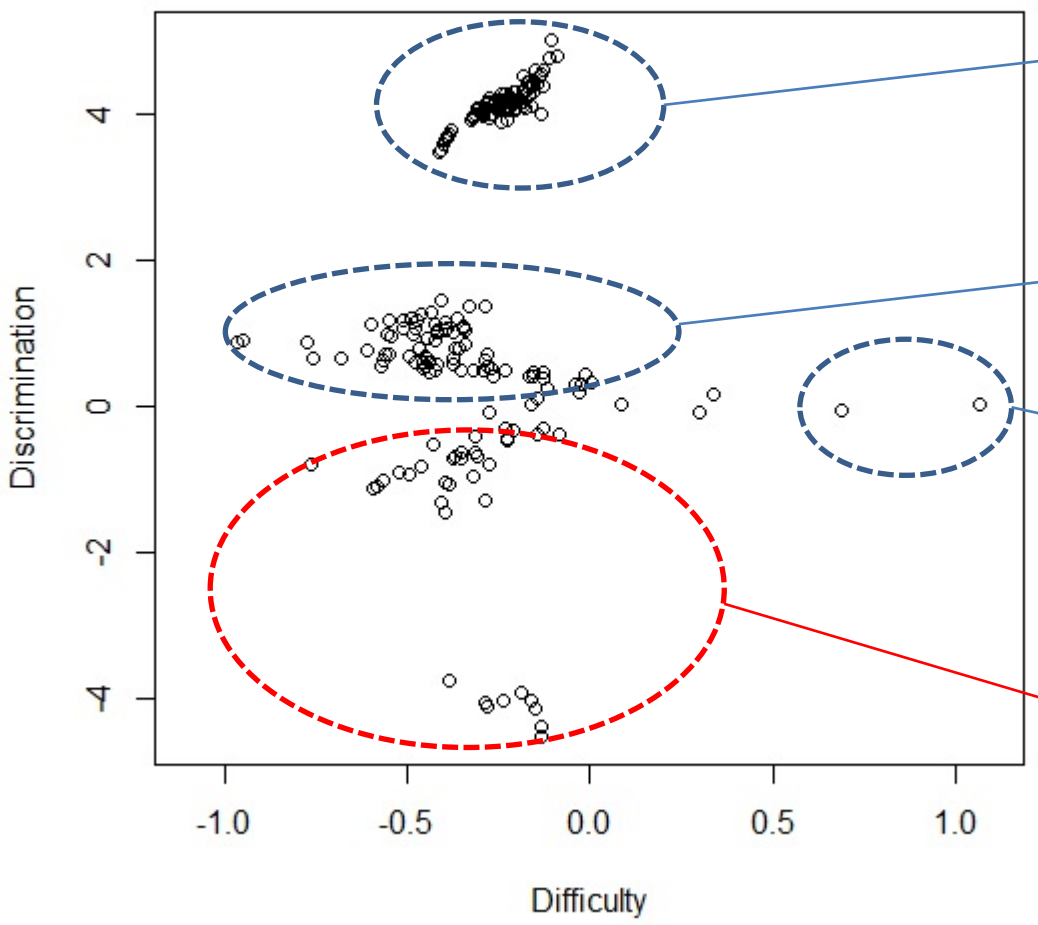








Difficulty vs Discrimination



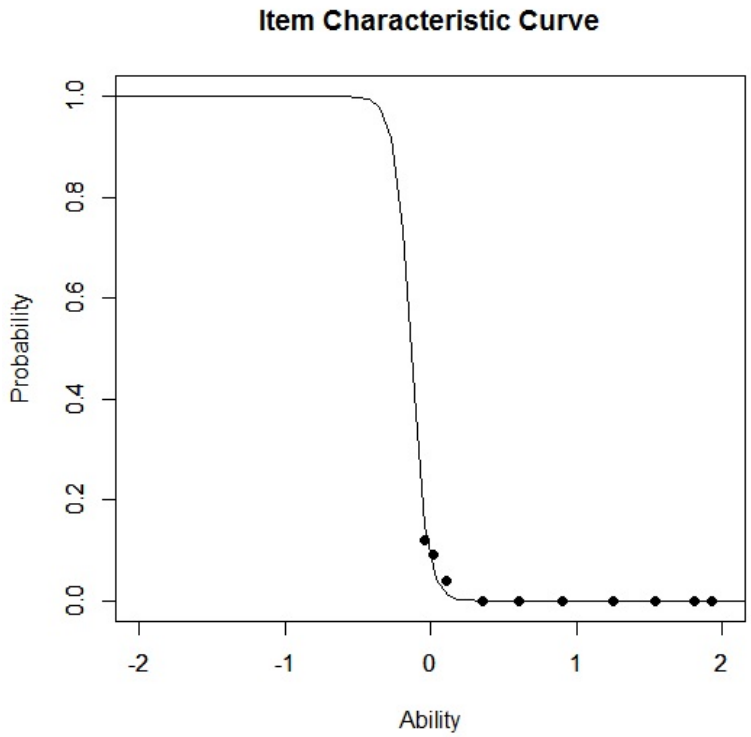
Not solved using few trees

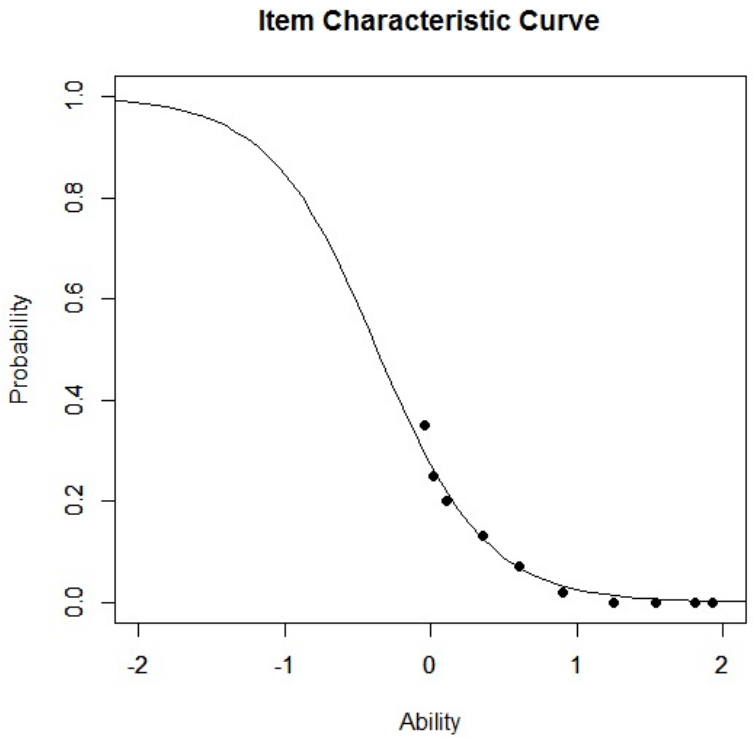
Different levels of difficulty

Very difficulty but low discriminating

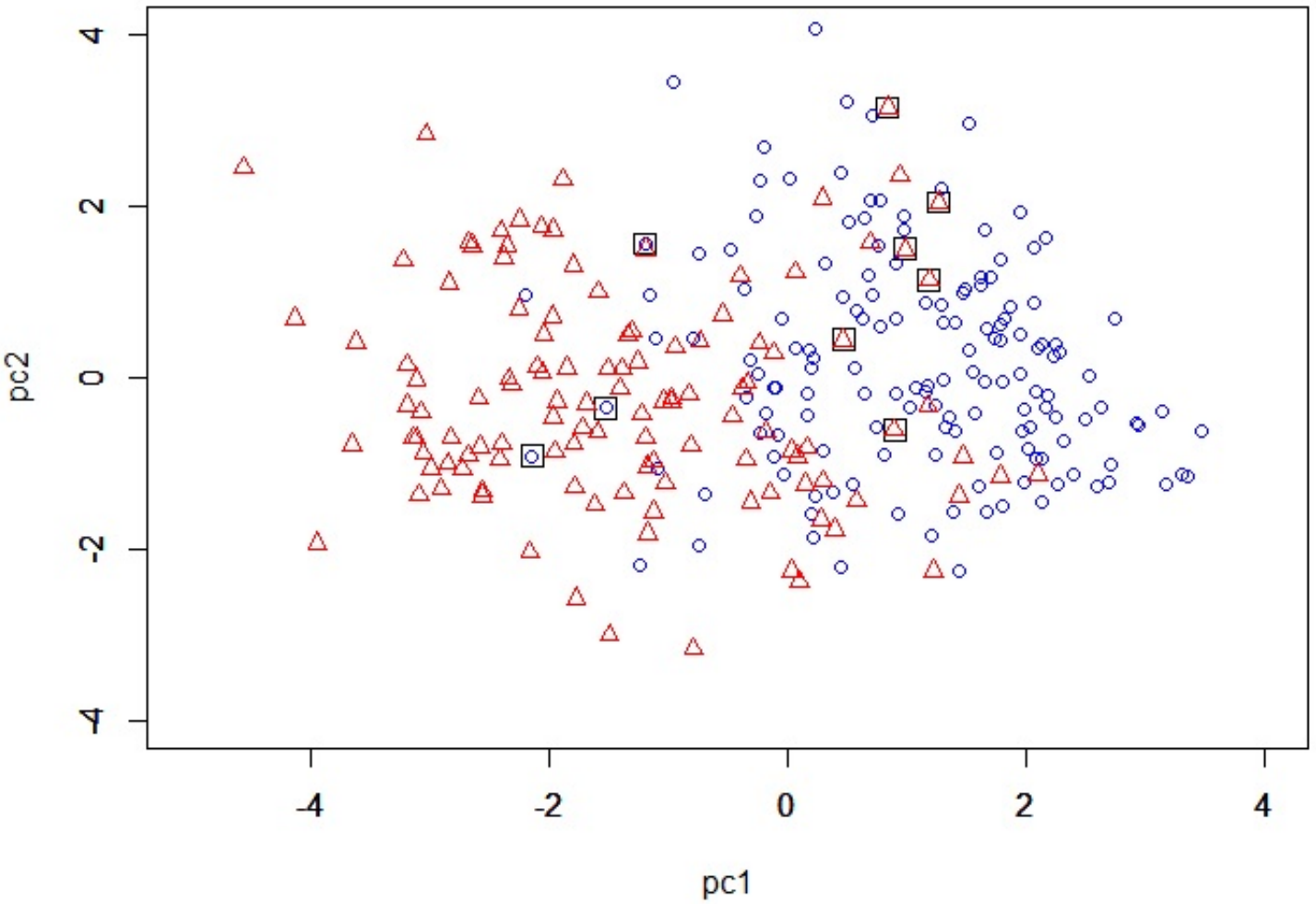
Negative slopes, i.e., easier for the worst classifiers!!!

Item Characteristic Curves with Negative Slopes - Examples





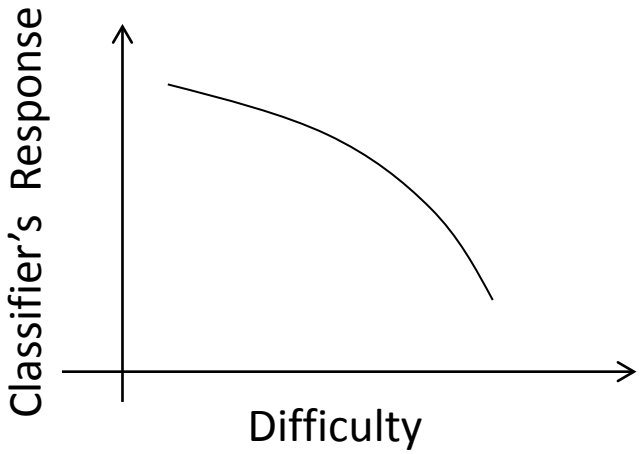
Visualization of the Heart dataset (2 PCA)



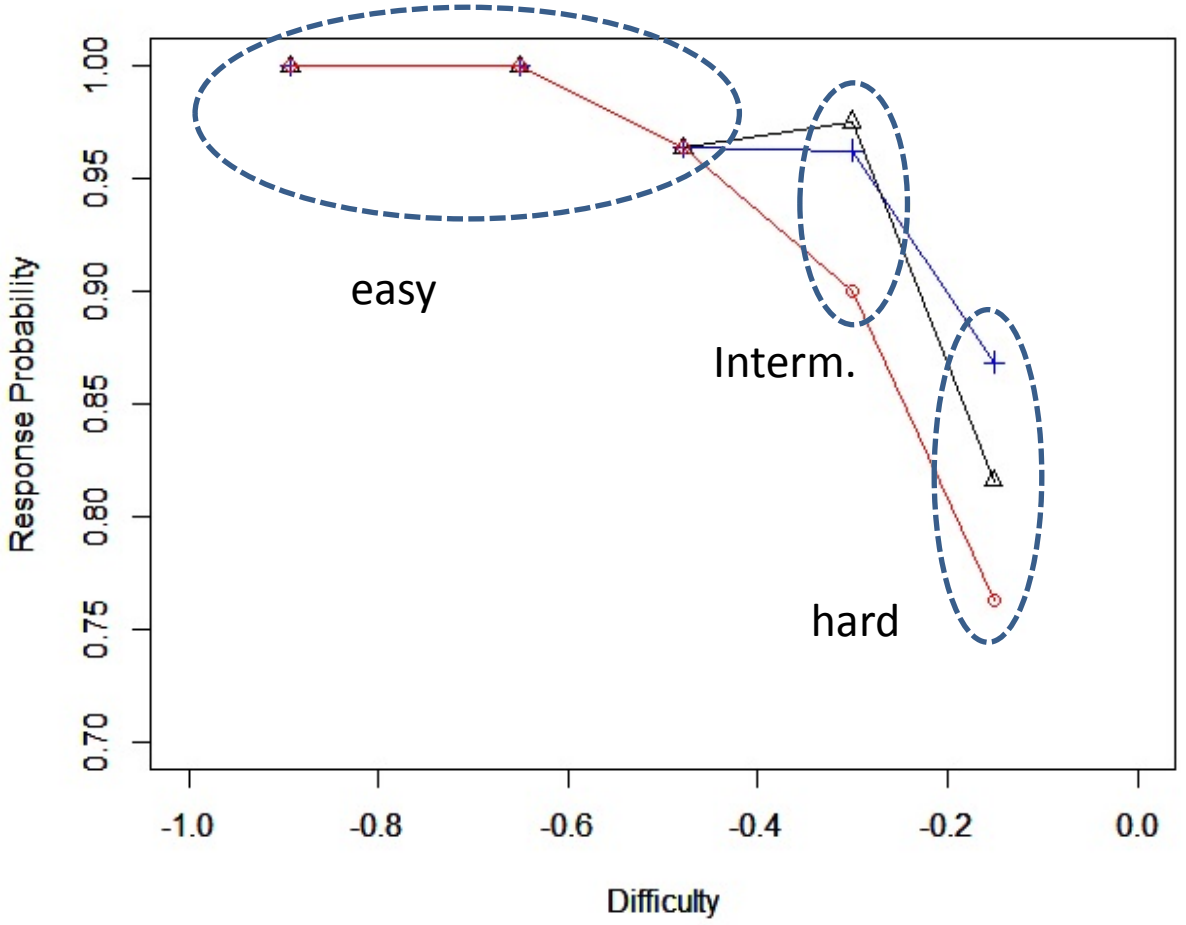
Squares indicates the instances with the top 9 negative slopes

Model Characteristic Curve (MCC)

- Once the hardness is estimated, we can define characteristic curves for models
- MCC plots the response probability of a classifier along the instance difficulty



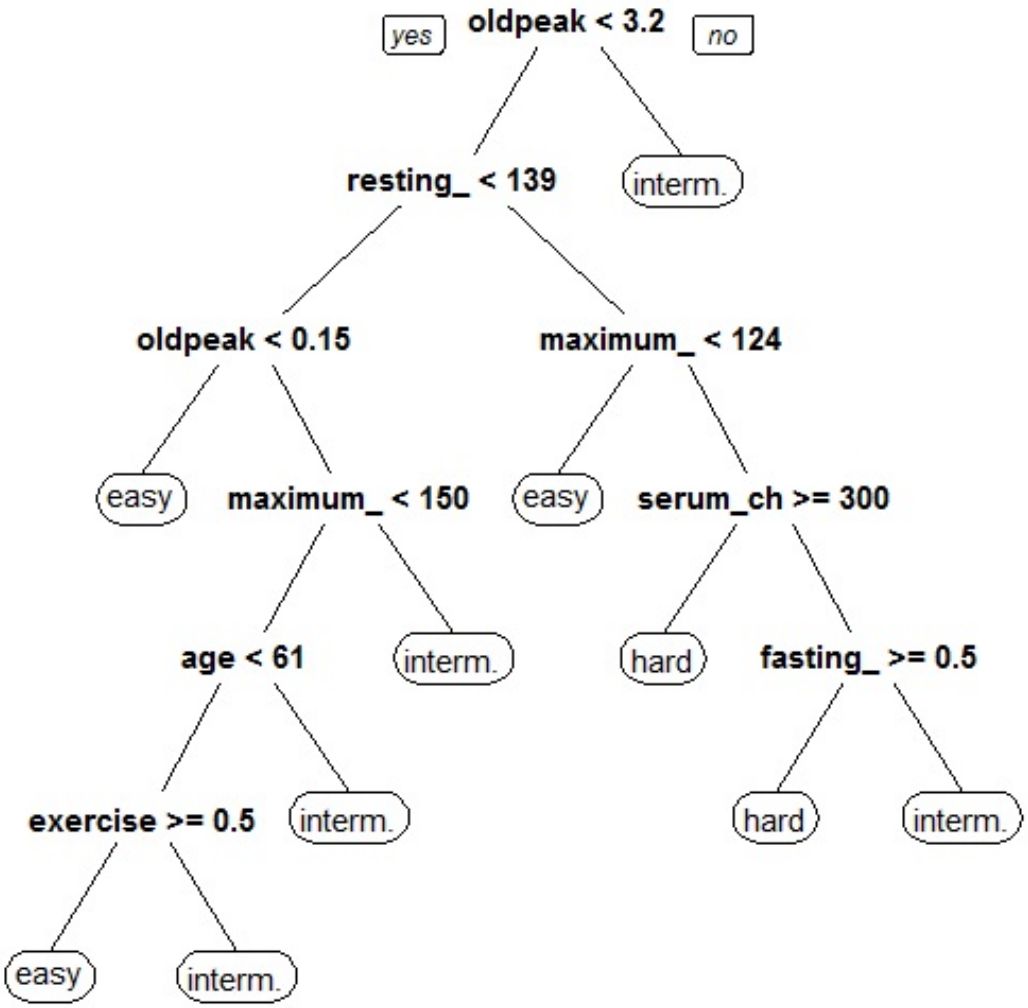
Model Characteristic Curve



+ Naive Bayes

△ Logistic Regression

○ Decision Trees


yes
oldpeak < 3.2
no
resting_ < 139
interm.
oldpeak < 0.15
maximum_ < 124
easy
maximum_ < 150
easy
serum_ch >= 300
age < 61
interm.
hard
fasting_ >= 0.5
exercise >= 0.5
interm.
hard
interm.
easy
interm.

Conclusion

- A completely different way of analysing classifiers at instance level

- Possibilities and future issues:
  - Computational issues, alternative IRT models, which classifiers to use,...
  - ...
  - IRT models for learning, instance selection and weighting, ensemble and classifier selection,..

Analysis of Instance Hardness in Machine Learning Using Item Response Theory

(Easy) Questions ???