

The alternating decision tree learning algorithm

Yoav Freund

AT&T

180 Park Avenue
Florham Park, USA

Llew Mason

Department of Systems Engineering
Australian National University
Canberra, Australia

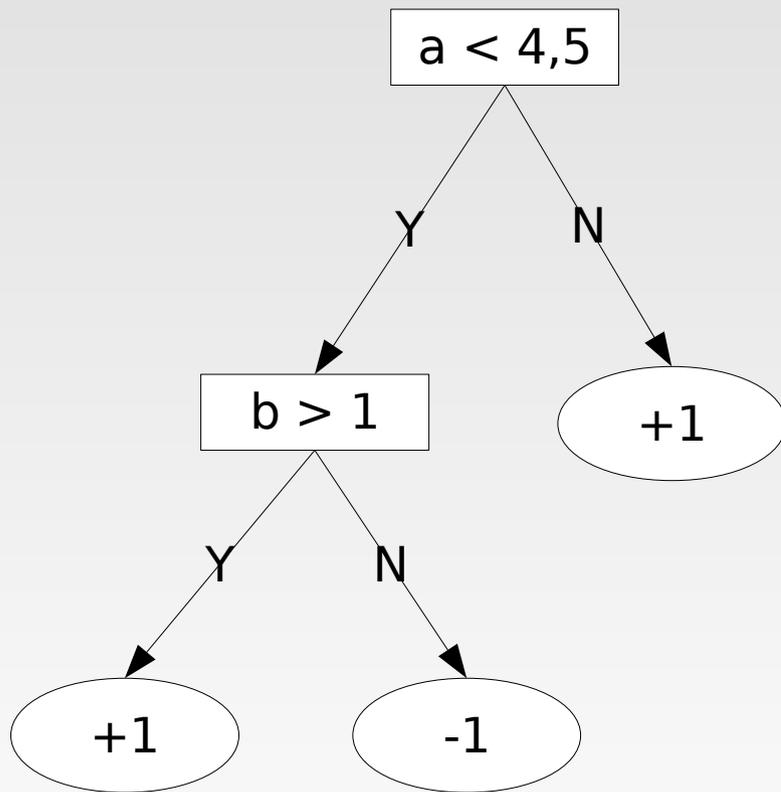
Présenté par Laurent Sieczkowski, M2R Université Paris-Sud

Plan

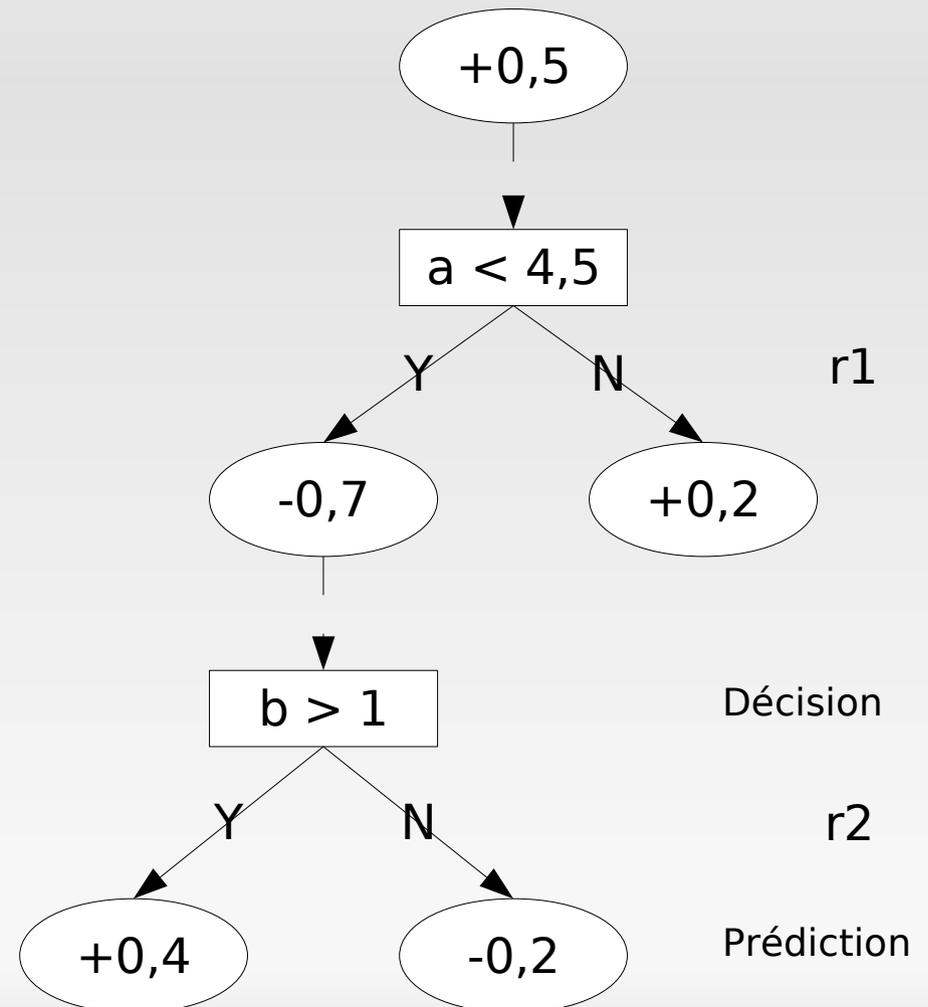
- ♦ Arbre alterné de décision
- ♦ Boosting des arbres alternées de décision
- ♦ Expériences et résultats

Arbre alterné de décision

Arbre classique



Arbre alterné



Arbre alterné de décision

- Règle de classification:
 $\text{sign}(0,5 + r_1(a,b) + r_2(a,b))$
- Généralisation:
 - arbres de décisions
 - “stumps” votés

Plan

- ♦ Arbre alterné de décision
- ♦ Boosting des arbres alternées de décision
- ♦ Expériences et résultats

Boosting des Arbres alternées de décision

- Application simple d'algorithme de boosting
 - Somme de règles
- Utilisation Adaboost modifié
- Validation croisée (choix d'arrêt)

Boosting des Arbres alternées de décision

- Algorithme:

- Trouver la meilleur prédiction constante

- Minimiser Z_t

$$Z_t(c1, c2) = 2(\sqrt{(W+(c1 \wedge c2)W-(c1 \wedge c2))} + \sqrt{(W+(c1 \wedge \neg c2)W-(c1 \wedge \neg c2))}) + W(\neg c2)$$

- ajout des règles $c1$ et $c2$

- calcul et ajout des valeurs de prédiction

- Précondition

- ajout de règles $c1 \wedge c2$ et $c1 \wedge \neg c2$

- Mise à jour des poids

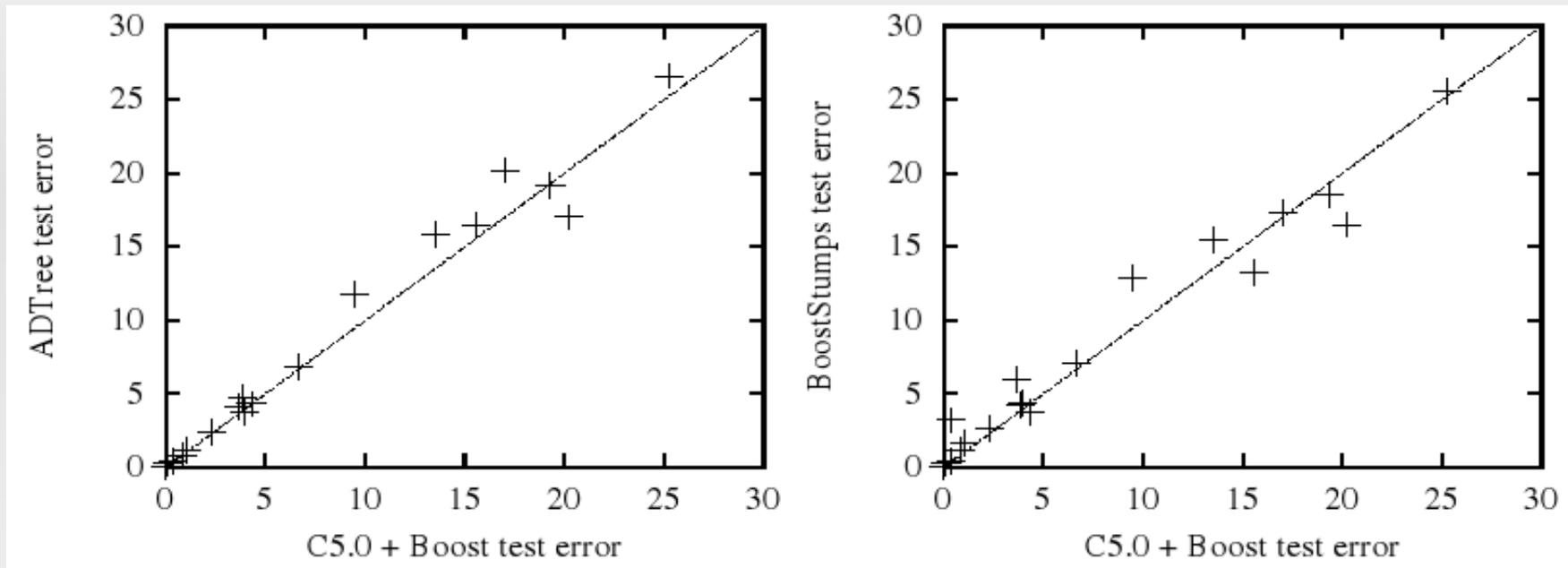
Plan

- ♦ Arbre alterné de décision
- ♦ Boosting des arbres alternées de décision
- ♦ Expériences et résultats

Expériences et résultats

- Comparaison:
 - C5,0 (avec ou sans boosting)
 - StumpBoost
 - ADTree
- Arbre plus facilement interprétable
 - faible taille
 - décision compréhensible
- Sujet au sur-apprentissage
 - arrêt plus rapide

Expériences et résultats



- Performance très correcte
- Prend les avantages de C5,0 et StumpBoost
- Représente aussi bien une arbre de décision qu'un “stump” voté