Boosting	 Ensemble Methods Weak and Strong Learners Useful Weak Learners
2000000	2 Boosting Algorithms
	 Bagging Boosting AdaBoost

Örjan Ekeberg	Brain Modeling and Machine Learning	Örjan Ek	eberg	Brain Modeling and Machine Learning
Ensemble Methods	Weak and Strong Learners	Ensemble Me	hods	Weak and Strong Learners
Boosting Algorithms	Useful Weak Learners	Boosting Algori	thms	Useful Weak Learners

1 Ensemble Methods

- Weak and Strong Learners
- Useful Weak Learners

2 Boosting Algorithms

- Bagging
- Boosting
- AdaBoost

Ensemble Method

Combining hypotheses from several learners



Terminology

Weak Learner

Learning algorithm capable of always producing hypotheses that perform *better than chance*.

Strong Learner

Learning algorithm capable of producing hypotheses that perform *arbitrarily well*.

A *strong learner* can always be constructed by combining multiple instances of any *weak learner*.

• Instead of inventing very good learning algorithms we can use multiple simple ones

Örjan Ekeberg	Brain Modeling and Machine Learning	Örjan Ekeberg	Brain Modeling and Machine Learning
Ensemble Methods Boosting Algorithms	Weak and Strong Learners Useful Weak Learners	Ensemble Methods Boosting Algorithms	Bagging Boosting AdaBoost
Revival of simple learning algorith • Single layer perceptrons • Limited height trees ("Stump • Naïve Bayesian classifiers	ıms ping'')	 Ensemble Methods Weak and Strong Learners Useful Weak Learners Boosting Algorithms Bagging Boosting AdaBoost 	

Bagging — Bootstrap Aggregating

- Given a set of training examples D
- Form new sample sets D_i by randomly sampling D

Ensemble Methods Boosting Algorithms

- Use any weak classifier to get hypotheses: $D_i \rightarrow h_i$
- Create an aggregated classifier h^* which delegates to all h_i and returns the majority vote

Bagging

Boosting AdaBoost

Boosting

- Evaluate each learner on the training data
- Force next learner to concentrate on hard examples
- Weighted majority vote, based on performance

Örjan Ekeberg Brain Modeling and Machine Learning	Örjan Ekeberg Brain Modeling and Machine Learning				
Bagging	Barging				
Ensemble Methods Boosting Algorithms AdaBoost	Ensemble Methods Boosting Boosting Algorithms AdaBoost				
	1 Initialize weights $w_i = \frac{1}{N} \forall \text{ examples } i$				
	(2) Repeat until $t = T \lor \epsilon_t \ge 0.5$				
	• Train weak classifier: $\{D, \vec{w}\} \rightarrow h_t$				
AdaBoost — Adaptive Boosting	2 Evaluate the performance of h_t :				
Most common variant of boosting	$\sum 1 \cdot 1 - \epsilon_t$				
Assign a weight we to each training example	$\epsilon_t \leftarrow \sum_{i \in I : i \in I} w_i \text{set } \alpha_t \leftarrow \frac{1}{2} \log \frac{1}{\epsilon_t}$				
• Assign a weight with the case mene ettention to high weight	$\iota:h(x_i) \neq t_i$				
• Use a weak learner which pays more attention to high-weight	Indate weights:				
examples					
 Increase w_i for examples which are incorrectly classified 	$w_i \leftarrow \begin{cases} w_i e^{-\alpha_t} & \text{for correctly classified points} \end{cases}$				
	$w_i e^{\alpha_t}$ for incorrectly classified points				
	wightharpoonup wigh				
	3 Final classifier: $h^*(x) \equiv \operatorname{sign}\left(\sum_{t=1}^T \alpha_t h_t(x)\right)$				



- Adding more learners reduces training error
- Risk of overlearning
- In practice, this does not happen!
- Recent theoretical finding: AdaBoost tends to maximize margins

Örjan Ekeberg Brain Modeling and Machine Learning