

A matrix corresponding to boosting

	ex. 1	ex. 2	ex. 3	ex. 4	...
base rule 1	1	0	1	0	...
base rule 2	0	0	0	1	...
base rule 3	1	0	0	1	...

- ▶ 0 mistake, 1 correct.
- ▶ A weak learning algorithm: can find a base rule whose weighted error is smaller than $1/2 - \gamma$ or any distribution over the examples.
- ▶ There is a distribution over the base rules such that for any example the expected error is smaller $1/2 - \gamma$.
- ▶ Implies that the majority vote wrt this distribution over base rules is correct on **all** examples.
- ▶ Moreover - the weight of the majority is at least $1/2 + \gamma$, the minority is at most $1/2 - \gamma$.