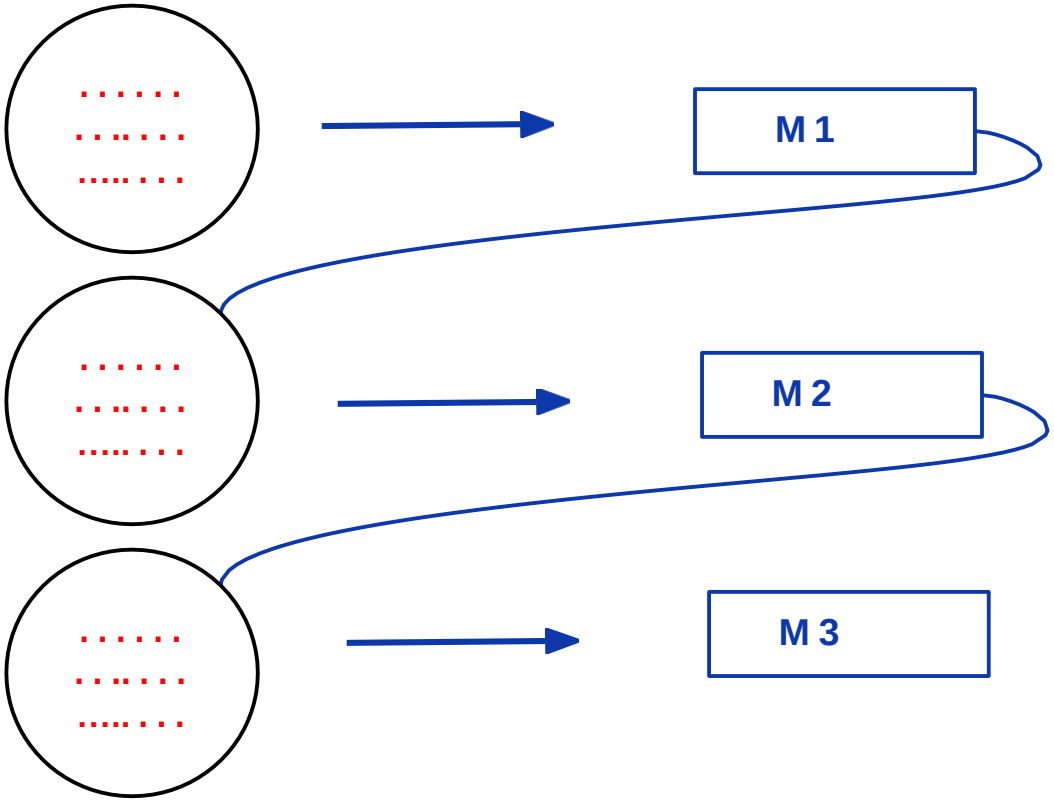
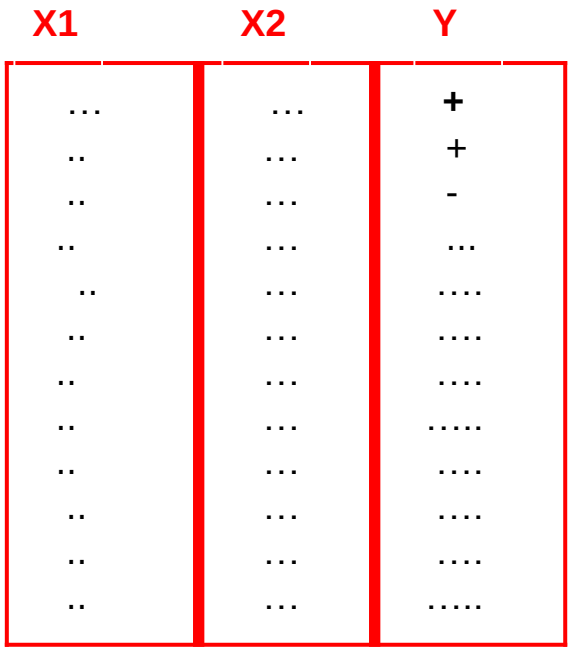
Boosting



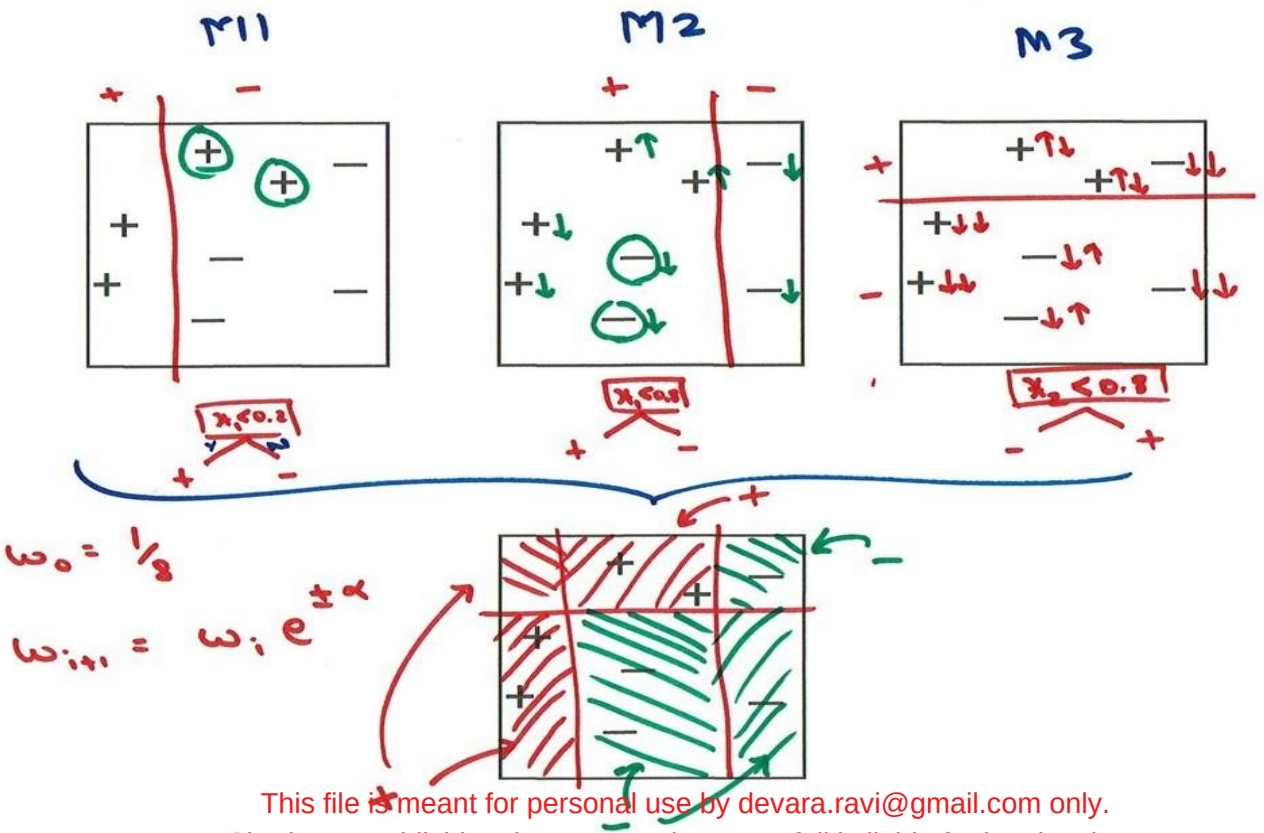
Boosting Methods

- **AdaBoosting** (Adaptive Boosting)
  - In AdaBoost, the successive learners are created with a focus on the ill fitted data of the previous learner
  - Each successive learner focuses more and more on the harder to fit data i.e. their residuals in the previous tree
- **Gradient Boosting** (GBM)
  - Each learner is fit on a modified version of original data. Original data is replaced with the x values and residuals from previous learner
  - By fitting new models to the residuals, the overall learner gradually improves in areas where residuals are initially high
- **XG Boost (Extreme Gradient Boosting)** (XGM)
  - Upgraded implementation of Gradient Boosting. Developed for high computational speed, scalability, and better performance.
  - Parallel Implementation, Cross-Validation, Cache Optimization, Distributed Computation

AdaBoost

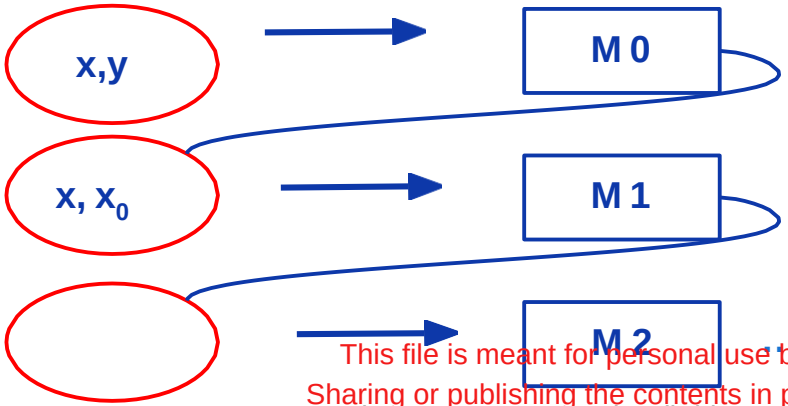


AdaBoost



Gradient Boosting

X			y	y_0	$y - y_0$	h
			50	40	10	8
			92	100	-8	-8
			60	80	-20	-10
			64	50	14	12
			.....	....	.....	.....

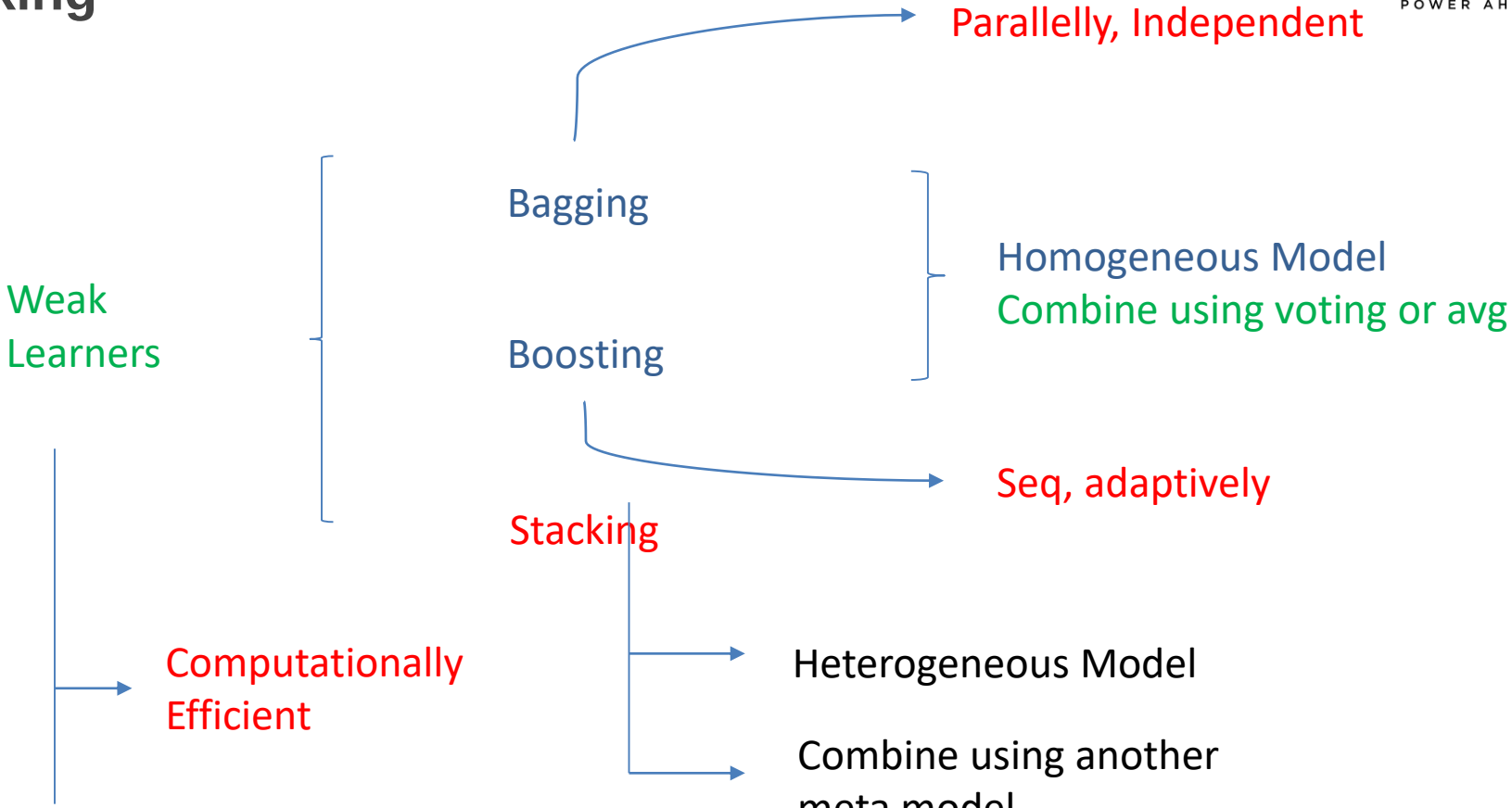


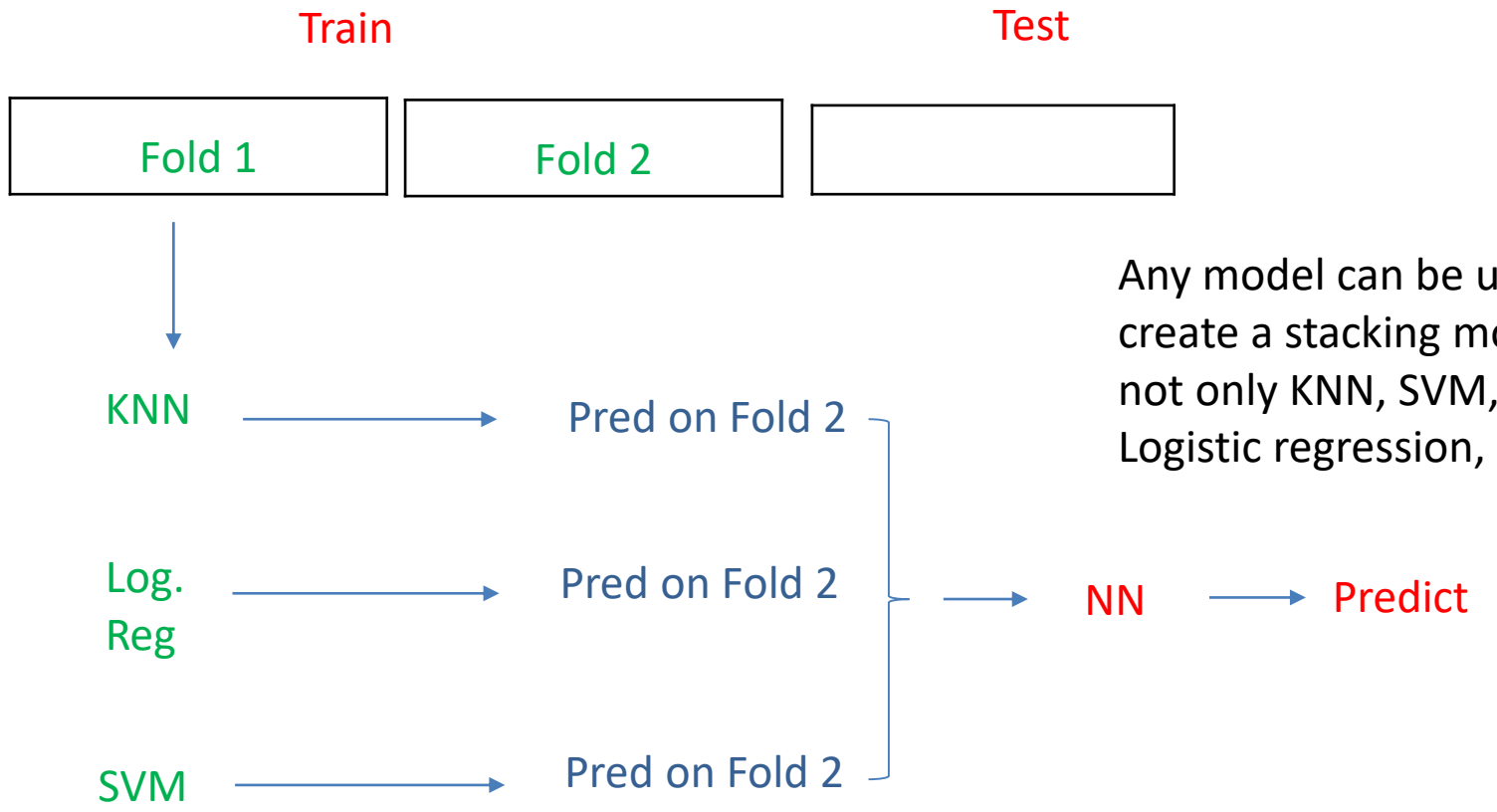
$y_0 = h_0(x)$

$y = h_0(x) + \alpha_1 h_1(x)$

$y = h_0(x) + \alpha_1 h_1(x) + \alpha_2 h_2(x) + \ldots\ldots\ldots$

Stacking





Any model can be used to create a stacking model not only KNN, SVM, Logistic regression, or NN