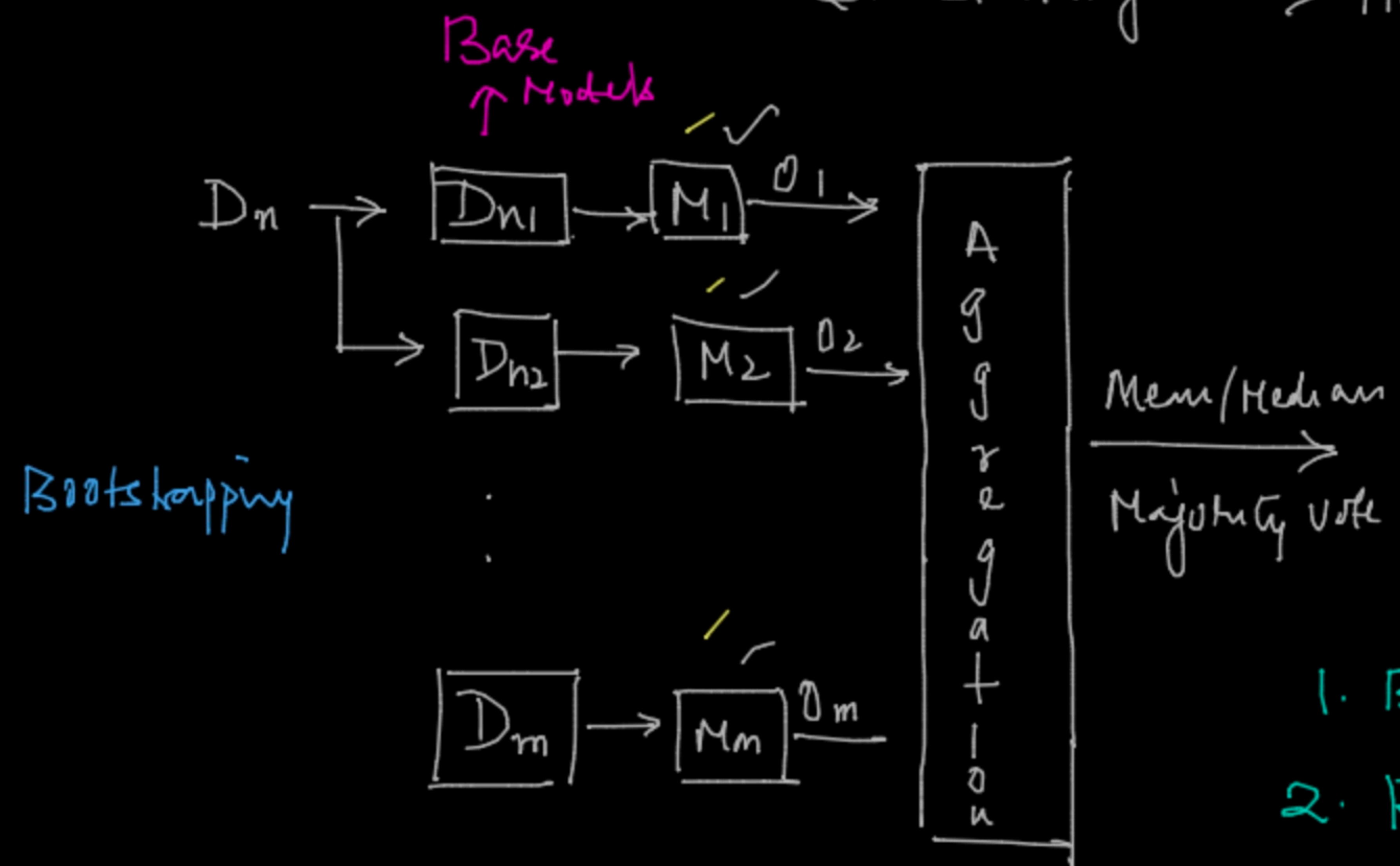


- 1. Bagging \rightarrow Bootstrapped Aggregation } Homogeneous
- 2. Boosting $\checkmark \rightarrow$ AdaBoost
- 3. Stacking \rightarrow Heterogeneous Ensemble



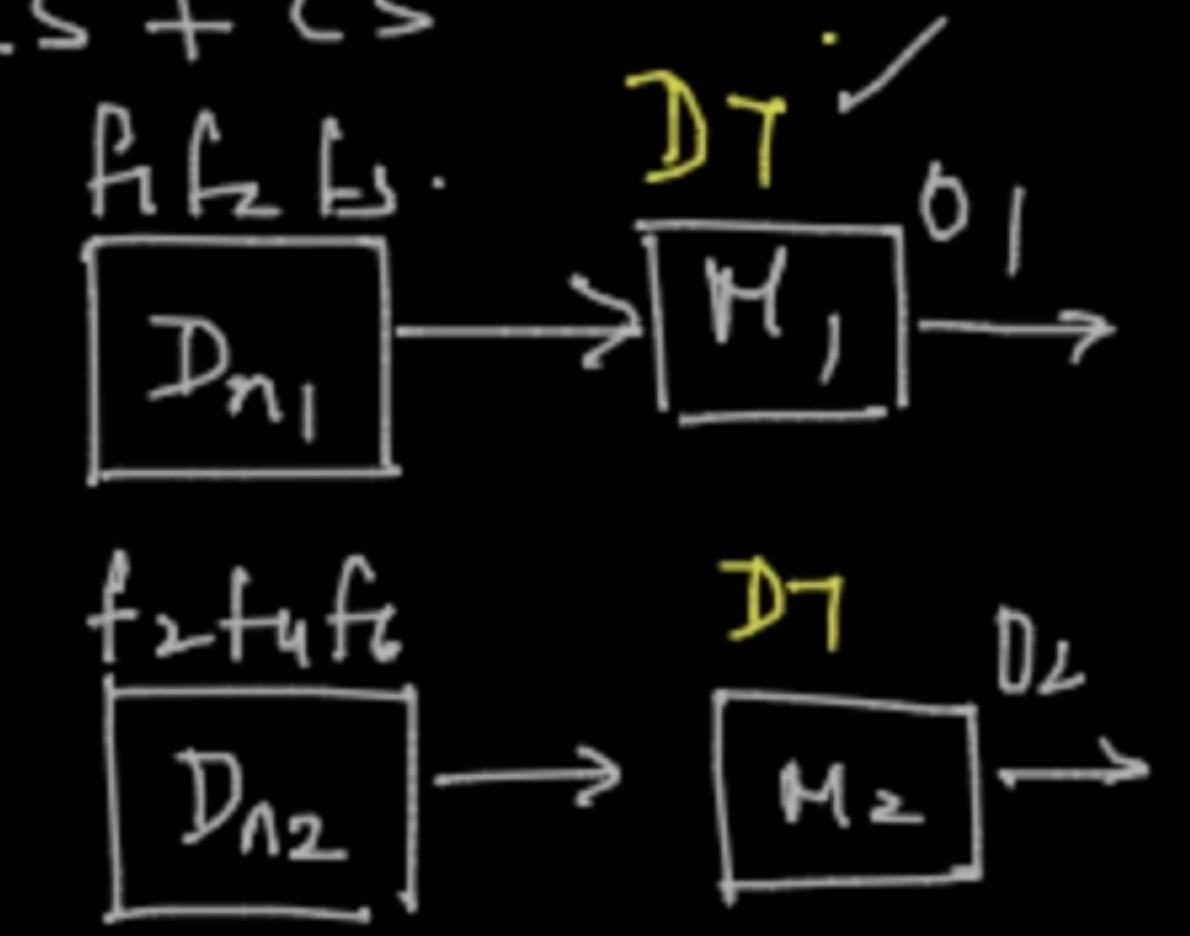
Bagging reduces Variance

Base Models — High Variance + Low Bias
 \downarrow
DT Built to large depth

Parallelize the Execution \rightarrow Faster

- 1. Bagging classifier / Bagging Regressor — Sklearn
- 2. Random Forest \rightarrow 1. All BM's are DT's

2. RS + CS



All models have equal Weightage

2. Boosting:

- only the first Base Model will predict y
- All Subsequent models will predict the errors from the previous models

ADA BOOST

75%

90%

87%

↑

✓

↑

$$\hat{y} = \alpha_0 f_0(x) + \alpha_1 f_1(x) + \alpha_2 f_2(x) + \dots$$

↓

	x_1	x_2	x_3	y	Wt	$f_0(x) = \hat{y}$
1				1	0.2	1 → 0.175 ↓
2				1	0.2	0 → 0.25 ↑
3				0	0.2	0 → 0.175 ↓
4				0	0.2	1 → 0.25 ↑
5				1	0.2	1 → 0.175 ↓
					1.0	1.0

2 Weights

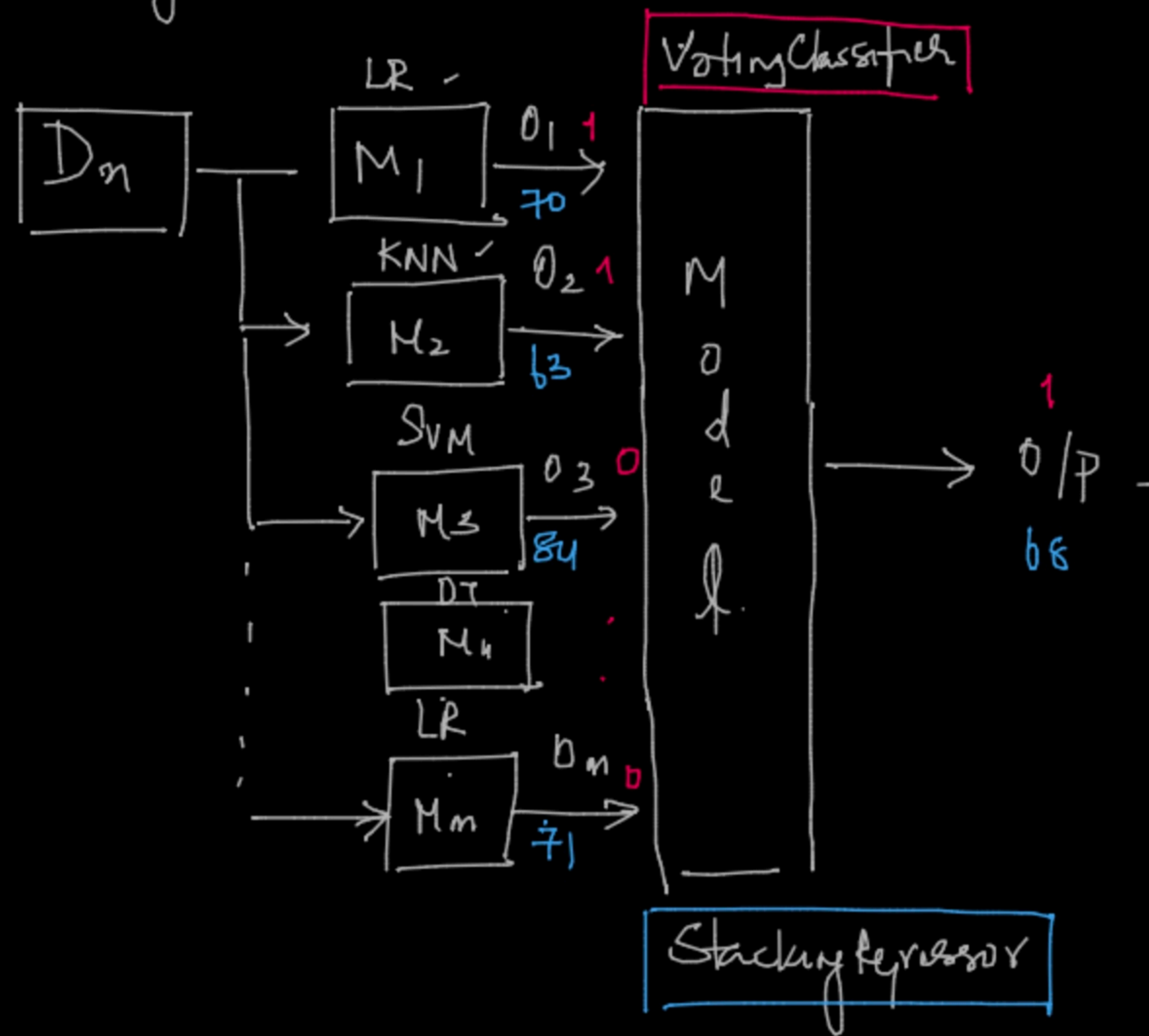
- Every model is given a weightage Based on its performance -
- Every datapoint is given a weightage -

XGBoost
 LightGBM
 CatBoost

} Optimizations on Boosting technique

Talking:

3. Stacking → Classification Model.



Bagging Model → Any model.

Random Forest → DT

AdaBoost → DT.

Stacking → we can choose the base models

