Boosting (form of ensemble)
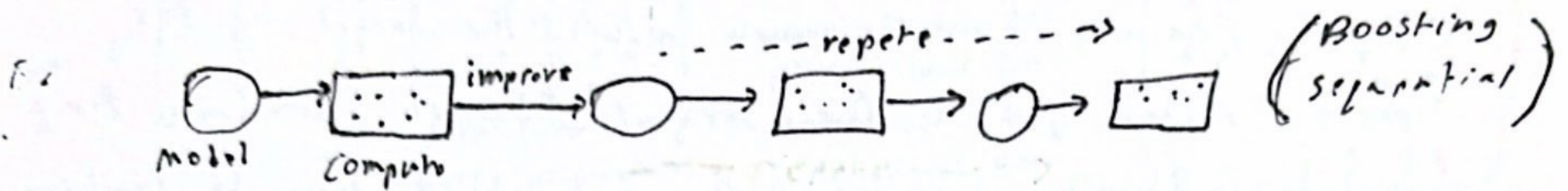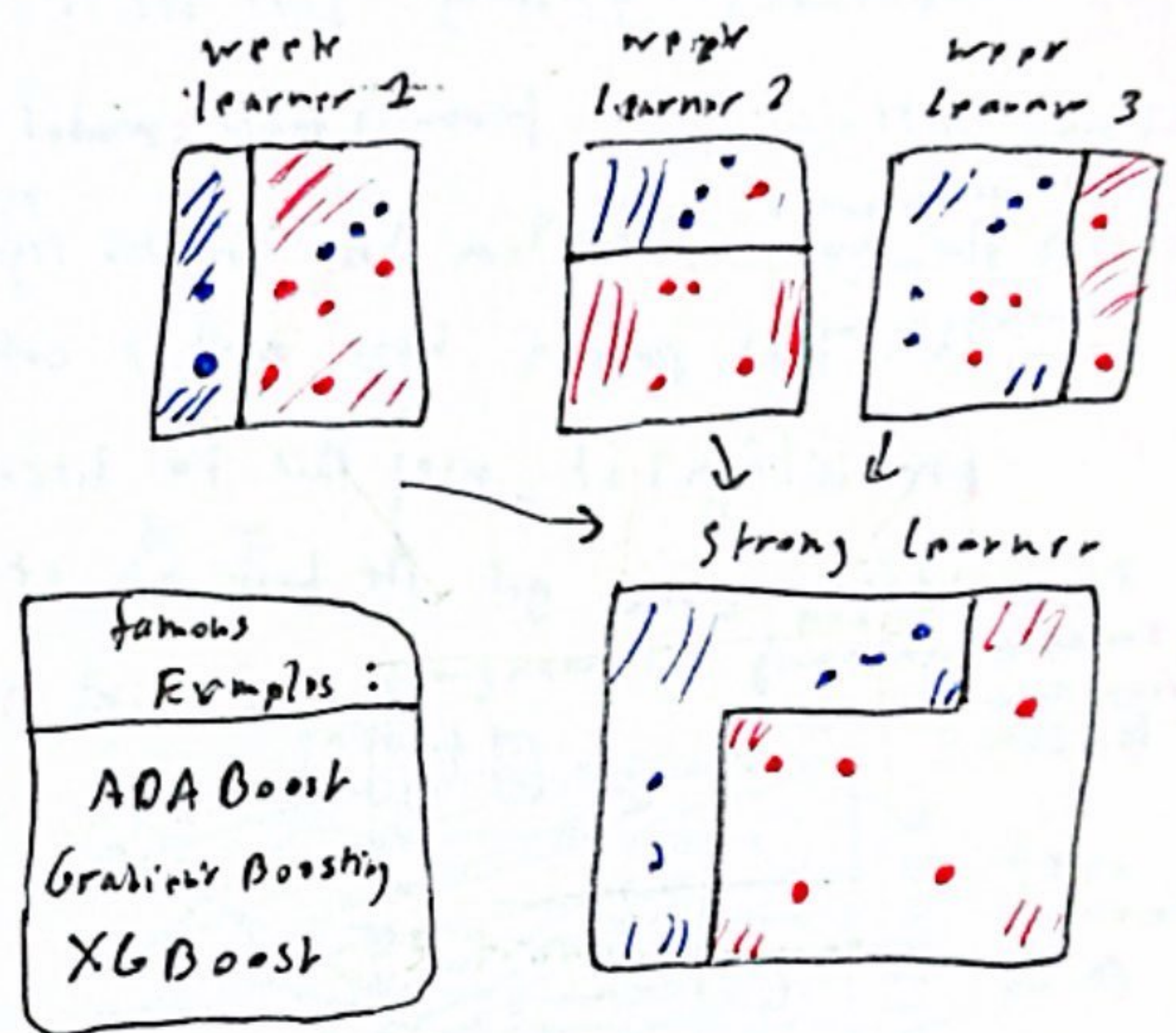
- another type of ensembel algorith popular for decision tree is Boosting where instead of running in parallel Like Bagging we train models in sequence (1 by 1) where each model focuses on fixing the errors made by the last one

Ex:
Model   Computer   improve   ----- repete ----->   (Boosting sequential)

- we combine a series of week models in sequence thus becoming a Strong model as each model becomes better then last by fixing in errors

So Boosted trees are better then Random forrests in general But more prone to overfiting plus Slower to train then Bagging as its not parranell

week learner 1   week learner 2   week learner 3

↓   ↓

Strong learner

famous
Examples :
ADA Boost
Gradient Boosting
XGBoost

Voting (for of ensemble) (and averaging)

- voting is a type of ensemble for Classification here take different models often of different type and combine predictions through voting there are 2 main types hard and soft
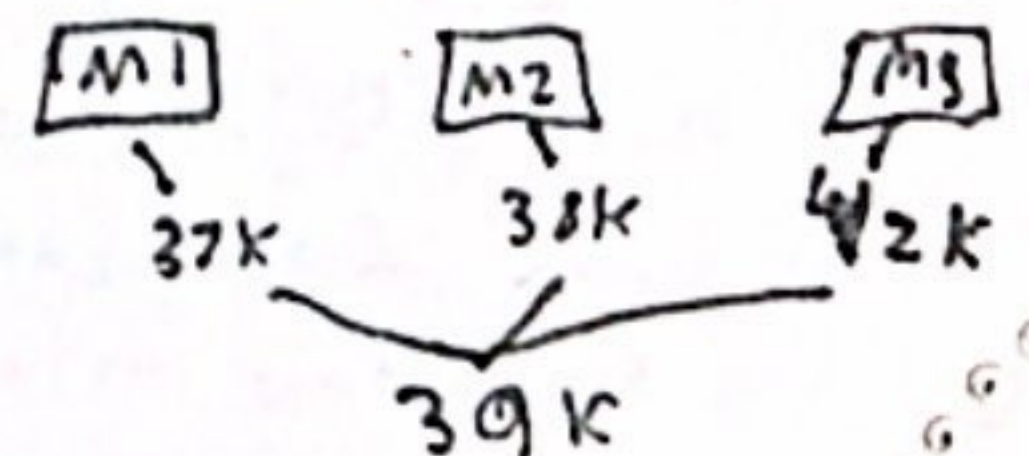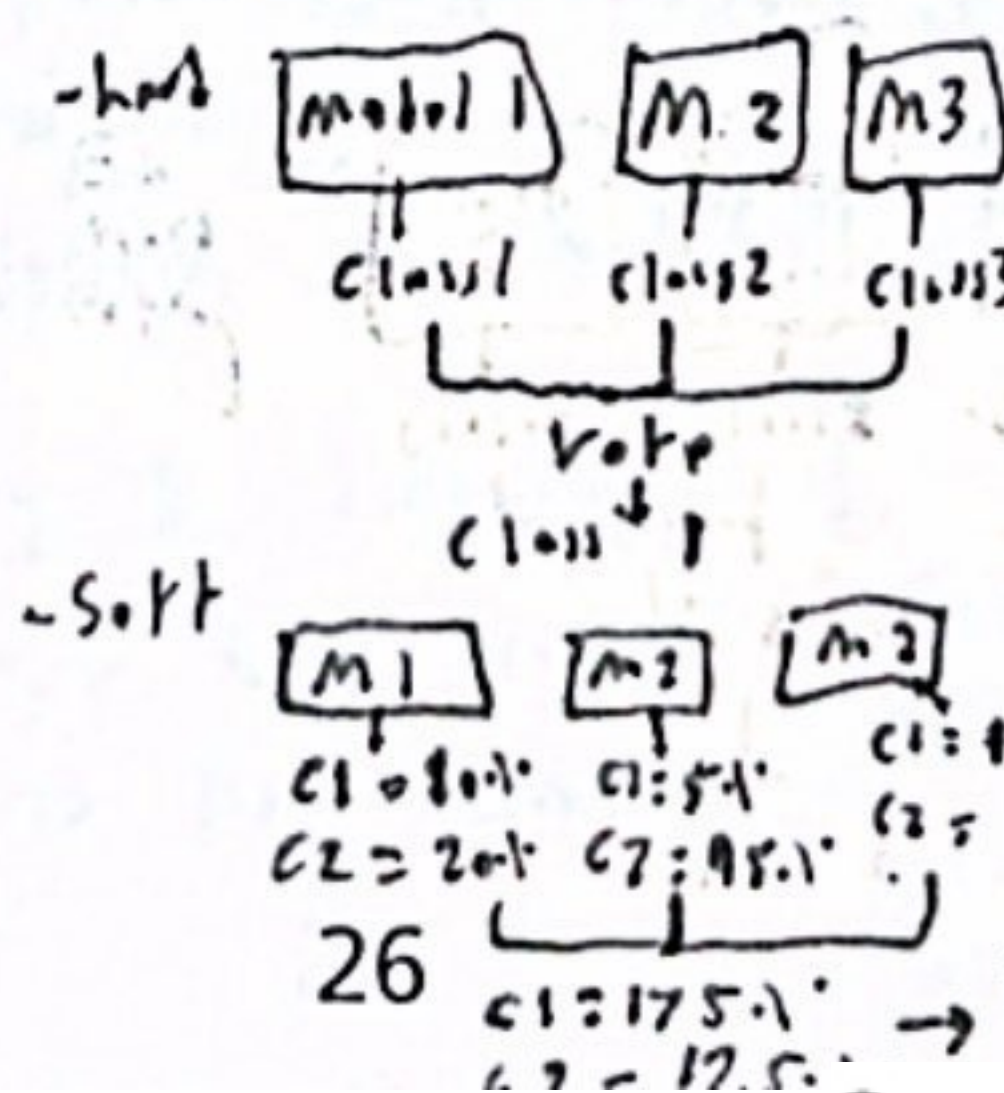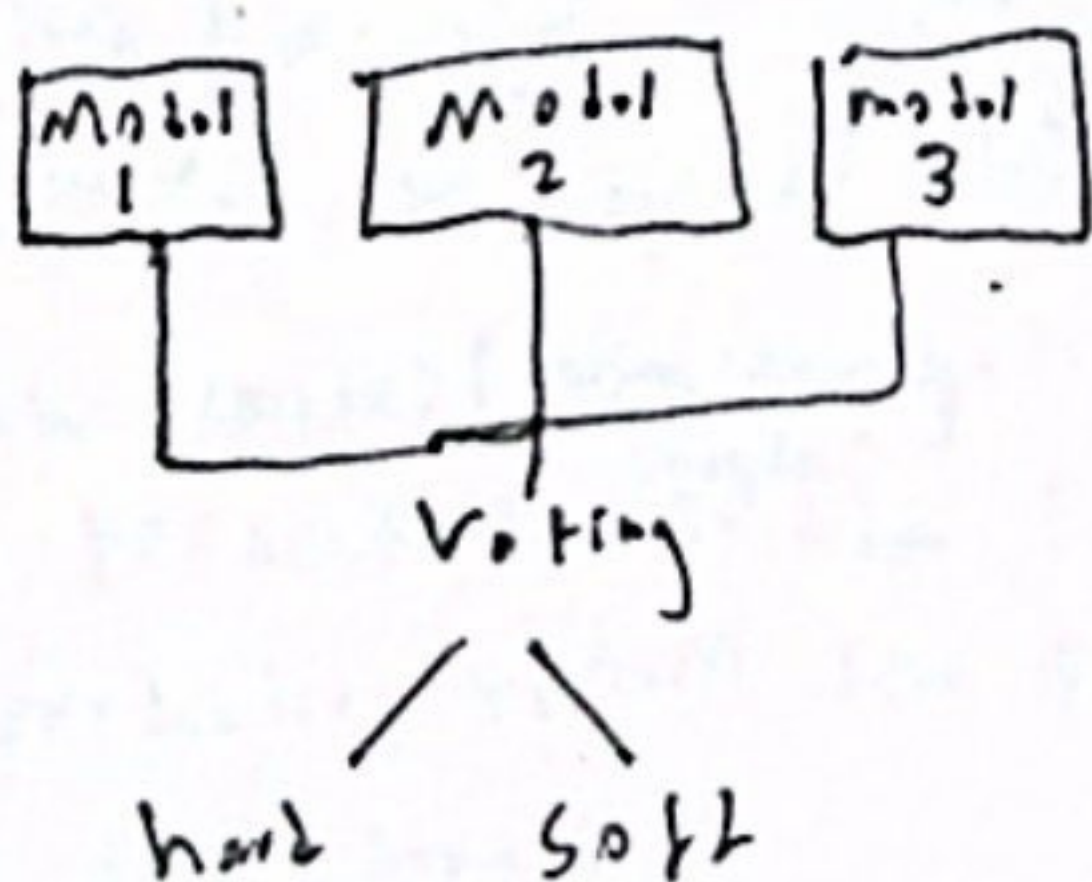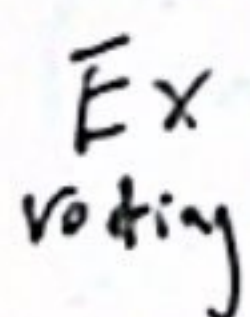
- hard voting is when models predict the class and the class with majority of votes is Choosen as final output

- Soft voting is when models output probabilities and classes probabilities are summed up and highest probability wins and is chossen

- hard vs soft depends on if classifiers give well defined prob Ex: NN, Niev Bayes & Logistic R

- Averaging is used for Regression where we avg outputs of models

EX
voting

| Model 1 | Model 2 | model 3 |

Voting
/ \
hard   Soft

-hard | Model 1 | M 2 | M3 |

class1   class2   class3

Vote
Class 1

-Soft | M1 | M2 | M3 |
C1 = 10% C1:54% C1: 40%
C2 = 20% C2:95.1% C2 = 10%

26   C1:175.1% → vote 1

| M1 | M2 | M3 |
37k   38k   42k

39K

-Ex of averaging