

Low bias High variance

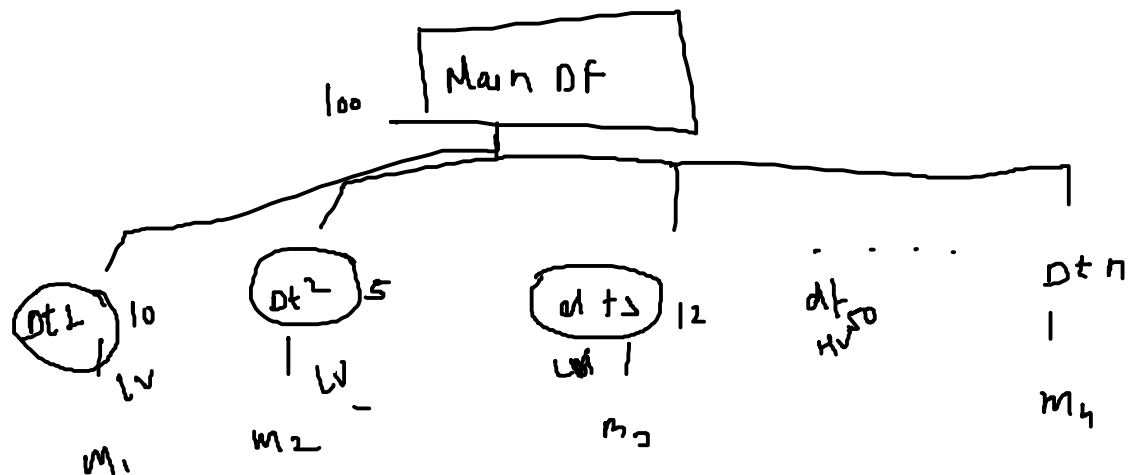
decision tree

low bias \rightarrow training data $\gamma\gamma$ work well

high variance \rightarrow train & test $\gamma\gamma$ Variance $\gamma\gamma$ does not work well.

0 — dt LB HV
0 — dt LB HV
0 — dt LB HV
0 — dt LB HV

Bagging (Bootstrapping + Aggregate)



Sampling :- Randomly.

Main df

name	Age	wt	BMI	df
A1	28	68	24	normal
A2	30	52	19	under wt
A3	25	80	28	over wt
A4	31	75	28	normal
A5	29	76	19	under wt
A6	27	80	30	over wt

400
Subset (1)
A1
A2
A3
A1

with replacement
without replacement

MS	SI	-	-	and or
AS	29	76	17	overw.
AG	27	80	30	

IPL

team

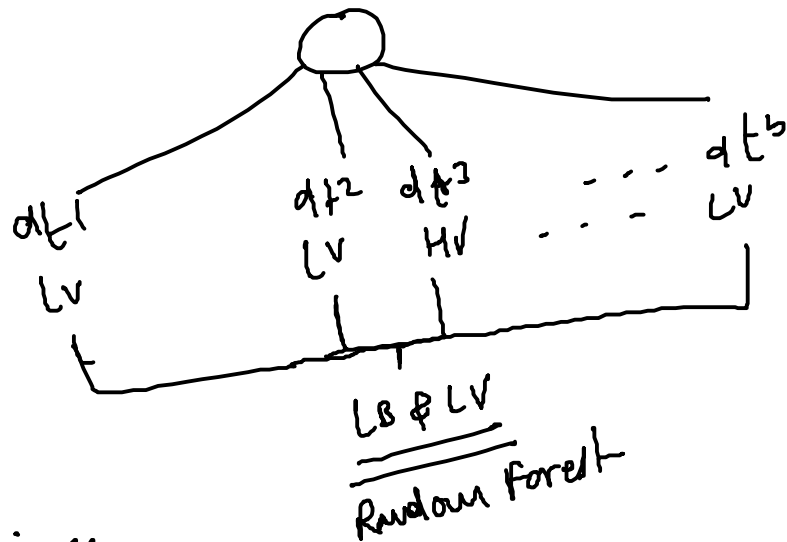
CSK
RCB
MI

Sachin, Chetnar, Yuvraj
Kohli, Singh
Rohit, Sharma, Hardik.

Zodiga

Kohli
Jodiga
Rohit

out of bag data \Rightarrow which were never picked by any subset.



Bias & variance

Random forest

out of bag \Rightarrow usually 20 to 30 %

Blog \Rightarrow data
 \downarrow
Articles \rightarrow
 \downarrow
research paper \rightarrow

Boasting

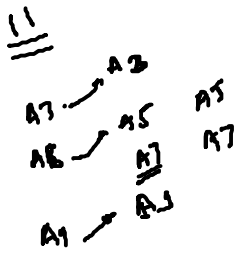
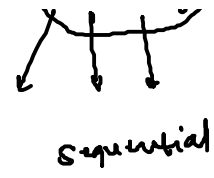
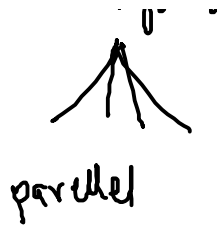
Boasting

Mai

T1 T2 T3

A2

as



hi