

25 students in class
60% accuracy model

4a) 3 models & build majority vote classifier C_3 , what would be accuracy of new classifier?

$$P = \binom{n}{k} p^k (1-p)^{n-k}$$

$$\begin{aligned} n &= 3 \\ k &= 2 \\ p &= 0.6 \end{aligned}$$

need to consider majority vote at least 2 of 3

$$P = \binom{3}{2} (0.6)^2 (1-0.6)^{3-2} + \binom{3}{3} (0.6)^3$$

$$\frac{\cancel{3} \cdot \cancel{2} \cdot 1}{\cancel{(2 \cdot 1)} \cdot 1} (0.36) (0.4) + \frac{\cancel{3} \cdot \cancel{2} \cdot \cancel{1}}{\cancel{(3 \cdot 2 \cdot 1)} \cdot 1} (0.216)$$

$$(3) (0.36) (0.4) + 0.216$$

$$0.432 + 0.216$$

$$0.648 \rightarrow 64.8\%$$

4b) 5 models & build majority vote classifier C_5 , what would be accuracy of new classifier

$$\begin{aligned} n &= 5 \\ k &= 3 \\ p &= 0.6 \end{aligned}$$

need to consider majority vote at least 3 of 5

$$P = \binom{5}{3} (0.6)^3 (1-0.6)^{5-3} + \binom{5}{4} (0.6)^4 (1-0.6)^{5-4} + \binom{5}{5} (0.6)^5$$

$$(\cancel{5} \cdot \cancel{4} \cdot \cancel{3} \cdot \cancel{2} \cdot 1) / (\cancel{3 \cdot 2 \cdot 1} \cdot 1) (0.216) (0.4) + \frac{\cancel{5} \cdot \cancel{4} \cdot \cancel{3} \cdot \cancel{2} \cdot 1}{\cancel{4 \cdot 3 \cdot 2 \cdot 1} \cdot 1} (0.1296) (0.4) + \frac{\cancel{5} \cdot \cancel{4} \cdot \cancel{3} \cdot \cancel{2} \cdot \cancel{1}}{\cancel{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} \cdot 1} (0.07776)$$

$$0.3456 + 0.2592 + 0.07776$$

$$= 0.68256 \rightarrow 68.3\%$$

4c) Provided a program to calculate this. The output yielded was 0.8462 or 84.6%

4d) The calculations in part c assume that all models are independent and have identical accuracy rates across diverse conditions. This does not hold in reality for the following reasons:

- 1) Models can be correlated, if they are trained on the same dataset or share similar features, algorithms or hyperparameters
- 2) The performance of models can vary significantly across different segments of data or under different conditions, which is not captured in a single accuracy metric
- 3) The assumption of a constant accuracy rate such as that of 60% across all possible inputs is overly simplistic as real world data often exhibits more complexity & variability

4e) Again I used the program from part c to calculate this. The output yielded was 0.3063 or 30.6%