Bayes Optimal Classifier

What if we have following hypotheses classifying a new instance as follows:

$h_1 \rightarrow$ classifies x as (+) with 0.4 → MAP!

$h_2 \rightarrow$ (−) with 0.3 → summing up (−) seems more probable

$h_3 \rightarrow$ (−) with 0.3 ↗

* If new example can only take on any value v_j from V.

then $P(v_j|D)$ - correct classification- is

$$P(v_j|D) = \sum P(v_j|h_i)\,P(h_i|D)$$

$P(v_j|D)$ ↓ look for the maximum

$P(v_j|h_i)$ → how hypothesis classifies instance in each label & their probabilities

$P(h_i|D)$ → prob of hypothesis given data

$$\text{argmax} \sum P(v_j|D)$$

$$= \text{argmax} \sum P(v_j|h_i)P(h_i|D)$$

↳ Bayes optimal classification

ex//

$P(h_1|D)=0.4$ $P(-|h_1)=0$ $P(+|h_1)=1$

$P(h_2|D)=0.3$ $P(-|h_2)=1$ $P(+|h_2)=0$

$P(h_3|D)=0.3$ $P(-|h_2)=1$ $P(+|h_3)=0$

$\sum_h P(+|h_i)P(h_i|D) = 0.4$

$\sum_h P(-|h_i)\,P(h_i|D) = 0.6$ → classified as (−)!

* If we're learning boolean concepts we can take weighted vote among all hypotheses in version space with weights = posterior probabilities.