FEATURES = {NH, SME, SPT, SMO}

$|D| = 8 \Rightarrow 3+, 5-$

$NH^+ = 2+, 3-$ $NH^- = 1+, 2-$
$SME^+ = 1+, 2-$ $SME^- = 2+, 3-$
$SPT^+ = 1+, 2-$ $SPT^- = 2+, 3-$
$SMO^+ = 1+, 3-$ $SMO^- = 2+, 2-$

$$IG(F) = E(D) - \sum_{v \in Values(F)} \frac{|D_v|}{|D|} \cdot E(F_v)$$

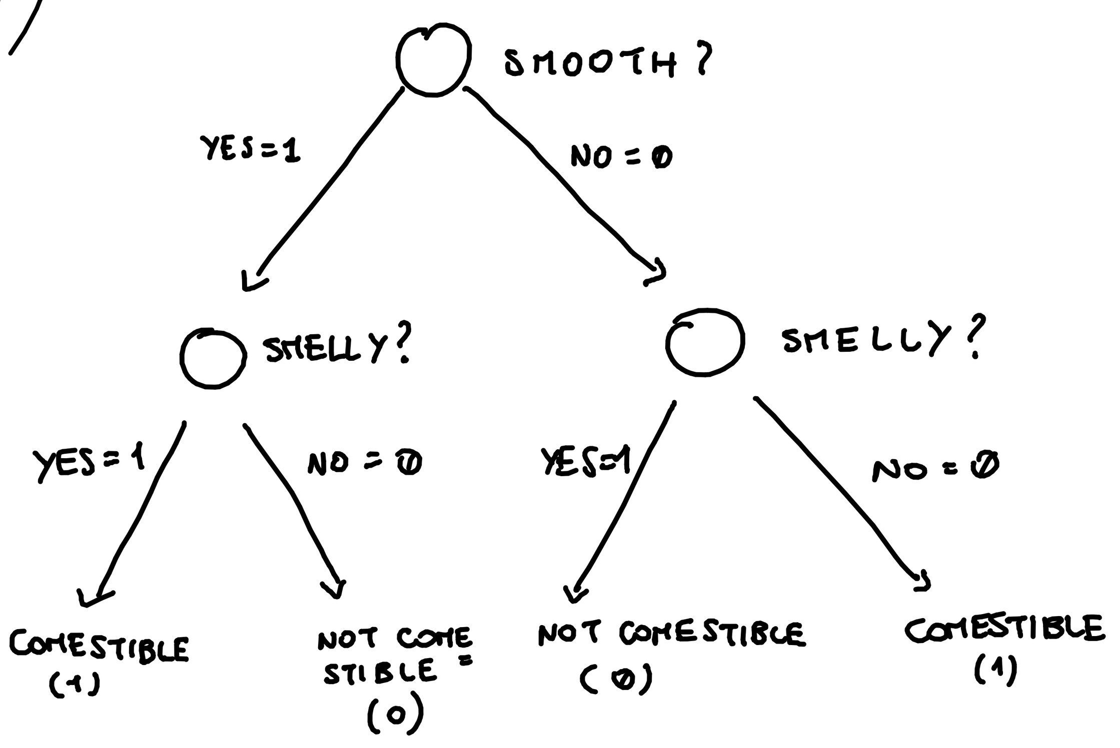
BINARY ENTR. IN THIS CASE

$$E(D) = \frac{3}{8}\log\frac{8}{3} + \frac{5}{8}\log\frac{8}{5} = 0{,}53 + 0{,}42 = 0{,}95$$

$$IG(NH) = 0{,}95 - \left(\frac{1}{8}\left(5\left(\frac{2}{5}\log\frac{5}{2} + \frac{3}{5}\log\frac{5}{3}\right) + 3\left(\frac{1}{3}\log 3 + \frac{2}{3}\log\frac{3}{2}\right)\right)\right)$$
$$= 0{,}95 - \left(\frac{1}{8}\left(5(0{,}529 + 0{,}442) + 3(0{,}91)\right)\right)$$
$$= 0{,}95 - 0{,}948 = 0{,}002$$

$IG(SME) = IG(SPT) =$ // SETS HAVE SAME SPLIT OF +/- $\Rightarrow$ SAME ENTROPY

$$IG(SMO) = 0{,}95 - \left(\frac{1}{8}\left(4\left(\frac{1}{4}\log 4 + \frac{3}{4}\log\frac{4}{3}\right) + 4\left(2\left(\frac{2}{4}\log\frac{4}{2}\right)\right)\right)\right)$$
$$= 0{,}95 - \frac{1}{8}\left(4(0{,}5 + 0{,}311) + 8(0{,}5)\right)$$
$$= 0{,}95 - 0{,}9055 = \boxed{0{,}004}$$




---

FEATURES = {SPT, SME, NH} // NOW WE CALCULATE METRICS OVER THE $SMO^+$ / $SMO^-$ SUBSETS!

class (u) = 1
class (v) = 1
class (w) = 0 :(

SMOOTH = 1 $|D| = 4$

$$E(D) = \frac{1}{4}\log 3 + \frac{3}{4}\log\frac{4}{3} = 0{,}70$$

$NH^+ = 0+, 2-$ $NH^- = 1+, 1-$
$SME^+ = 1+, 0-$ $SME^- = 0+, 3-$
$SPT^+ = 0+, 1-$ $SPT^- = 1+, 2-$

$$IG(NH) = 0{,}7 - \left(\frac{1}{4}\left(2\left(\frac{2}{2}\log\frac{2}{2}\right) + 2\left(\frac{1}{2}\log 2 + \frac{1}{2}\log 2\right)\right)\right)$$
$$= 0{,}7 - \left(\frac{1}{4}(0 + 2)\right) =$$
$$= 0{,}7 - 0{,}5 = 0{,}2$$

$$IG(SME) = 0{,}7 - \left(\frac{1}{4}\left(1(1\log 1) + 3\left(\frac{3}{3}\log\frac{3}{3}\right)\right)\right)$$
$$= 0{,}7 - 0 = \boxed{0{,}7}$$

$$IG(SPT) = 0{,}7 - \left(\frac{1}{4}\left(1(1\log 1) + 3\left(\frac{1}{3}\log\frac{3}{1} + \frac{2}{3}\log\frac{3}{2}\right)\right)\right)$$
$$= 0{,}7 - \left(\frac{1}{4}\left(3(0{,}91)\right)\right) =$$
$$= 0{,}7 - 0{,}6825 = 0{,}0175$$

SINCE ALL INST. THAT ARE SMO = 1 ∧ SME = 1 OR SMO = 1 ∧ SME = 0 HAVE THE SAME CLASS, WE CAN OUTPUT A DECISION NOW

---

SMOOTH = 0 $|D| = 4$

$$E(D) = \left(\frac{2}{4}\log\frac{4}{2}\right)2 = 1$$

$NH^+ = 2+, 1-$ $NH^- = 0+, 1-$
$SPT^+ = 1+, 1-$ $SPT^- = 1+, 1-$
$SME^+ = 0+, 2-$ $SME^+ = 2+, 0-$

$$IG(NH) = 1 - \left(\frac{1}{4}\left(3\left(\frac{2}{3}\log\frac{3}{2} + \frac{1}{3}\log 3\right) + \log 1\right)\right)$$
$$1 - \left(\frac{1}{4}\left(3(0{,}91)\right)\right) =$$
$$1 - 0{,}6825 = 0{,}3175$$

$$IG(SPT) = 1 - \left(\frac{1}{4}\left(2\left(\frac{1}{2}\log\frac{2}{1} + \frac{1}{2}\log\frac{2}{1}\right) + 2(//)\right)\right)$$
$$= 1 - \frac{1}{4}(4) = 0$$ // AS EXPECTED, THE "EVEN" SPLIT TELLS US NOTHING

$$IG(SME) = 1 - \frac{1}{4}\left(2\left(\frac{2}{2}\log\frac{2}{2}\right) + 2(//)\right)$$
$$= 1 - \frac{1}{4}(4(0)) = \boxed{1}$$

WE CAN NOW STOP FOR THE SAME REASONING AS BEFORE. CONTINUING IS POSSIBLE, BUT THE DEEPER WE GO, THE MORE WE RISK OVERFITTING