	Loves Popcorn	Loves Soda	Age	Loves Troll 2
①	Yes	Yes	7	No
②	Yes	No	12	No
③	No	Yes	18	Yes
④	No	Yes	35	Yes
⑤	Yes	Yes	38	Yes
⑥	Yes	No	50	No
⑦	No	No	83	No

Loves Popcorn

Yes / No

①⁻, ②⁻   ③⁺, ④⁺, ⑦⁻
⑤⁺, ⑥⁻
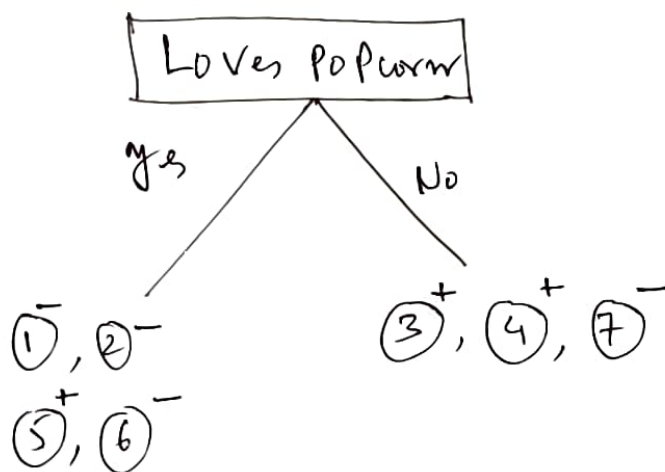
$$G(\text{Loves Popcorn}) = \frac{4}{7}\left\{1-\left(\frac{1}{4}\right)^2-\left(\frac{3}{4}\right)^2\right\}$$
$$+ \frac{3}{7}\left\{1-\left(\frac{2}{3}\right)^2-\left(\frac{1}{3}\right)^2\right\}$$

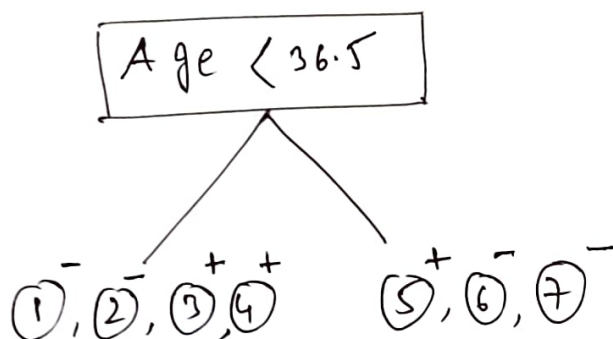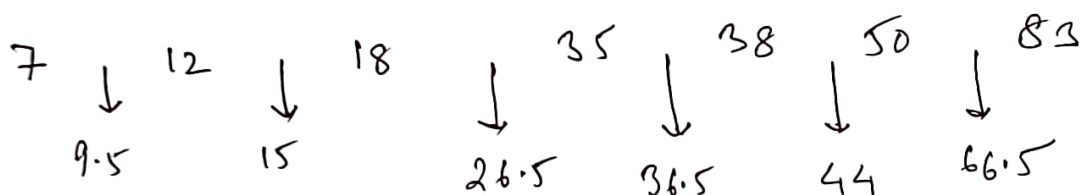$$= 0.405$$

Loves soda

Yes / No

①̄, ③⁺, ④⁺   ②̄, ⑥̄, ⑦̄

⑤⁺

$$G\,(\text{Loves soda})$$

$$= \frac{4}{7}\left\{1-\left(\frac{3}{4}\right)^r-\left(\frac{1}{4}\right)^r\right\} + \frac{3}{7}\left\{\left(\frac{0}{3}\right)^r+\left(\frac{3}{3}\right)^r\right\}$$

$$= 0.214$$

Age
$$7 \downarrow \quad 12 \downarrow \quad 18 \downarrow \quad 35 \downarrow \quad 38 \downarrow \quad 50 \downarrow \quad 83 \downarrow$$
$$9.5 \qquad 15 \qquad\qquad 26.5 \qquad 36.5 \qquad 44 \qquad 66.5$$

Age < 36.5

①̄, ②̄, ③⁺, ④⁺   ⑤⁺, ⑥̄, ⑦̄

$$G\,(\text{Age} < 36.5)$$

$$= \frac{4}{7}\left\{1-\left(\frac{2}{4}\right)^r-\left(\frac{2}{4}\right)^r\right\} + \frac{3}{7}\left\{1-\left(\frac{1}{3}\right)^r-\left(\frac{2}{3}\right)^r\right\}$$

$$= 0.476$$

Age < 9.5 Age < 15 Age < 26.5 Age < 36.5
0.429   0.343   0.476   0.476

Age < 44 Age < 66.5
0.343   0.429

Age < 15 / Age < 44 $\longrightarrow$ Winner

Now

G (Loves Soda) $= 0.214$

G (Loves Popcorn) $= 0.405$

G (Age < 15) $= 0.343$

Loves Soda

Yes / No

$\textcircled{1}^{-}, \textcircled{3}^{+}, \textcircled{4}^{+}$
$\textcircled{5}^{+}$

$\textcircled{2}^{-}, \textcircled{6}^{-}, \textcircled{7}^{-}$

Loves Soda

$\textcircled{1}^{-}, \textcircled{3}^{+}$
$\textcircled{4}^{+}, \textcircled{5}^{+}$

NO
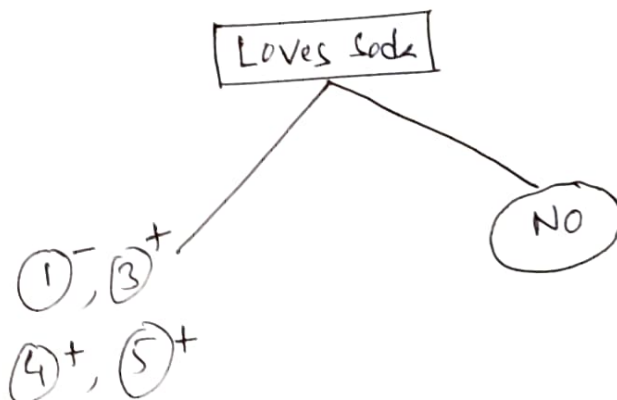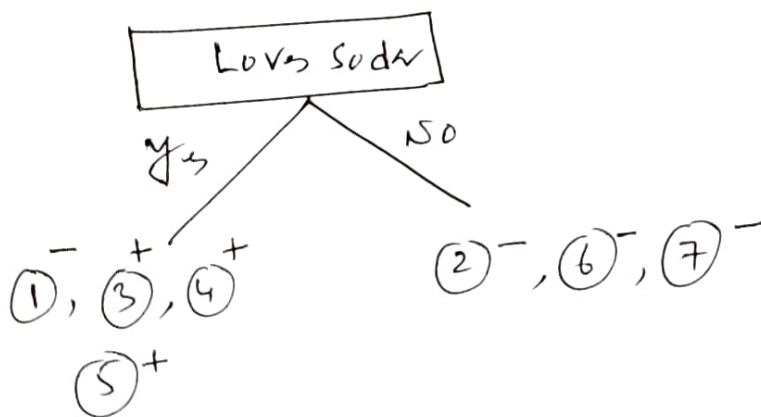
Loves Popcorn	Loves Soda	Age	Loves Troll 2
① Yes	Yes	7	No
③ No	Yes	18	Yes
④ No	Yes	35	Yes
⑤ Yes	Yes	38	Yes

~~Loves Popcorn~~

```
        Loves Popcorn
       Yes /        \ No
     ⁻ ①, ⑤ ⁺      ⁺③ ④⁺
```
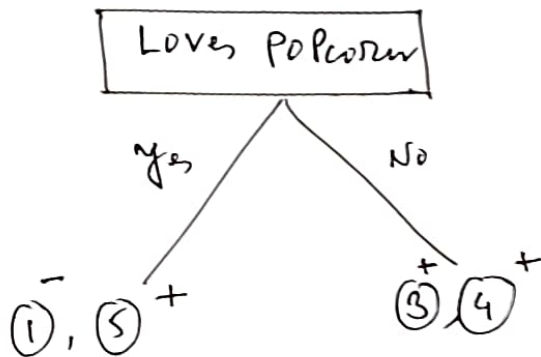
$G(\text{Loves Popcorn}) = 0.25$

Age
_____

```
  7      18      35      38
   ↓       ↓       ↓
  12·5   26·5    36·5
```

$\circledast$ $(Age < 12.5)$

$① ^-$ 　　　 $③^+ , ④^+ , ⑤^+$

$G\left(Age < 12.5\right) = \dfrac{1}{4}\left\{1 - \left(\frac{0}{1}\right)^{v} - \left(\frac{1}{1}\right)^{v}\right\}$

$+ \dfrac{3}{4}\left\{1 - \left(\frac{3}{3}\right)^{v} - \left(\frac{0}{3}\right)^{v}\right\}$

$= 0$
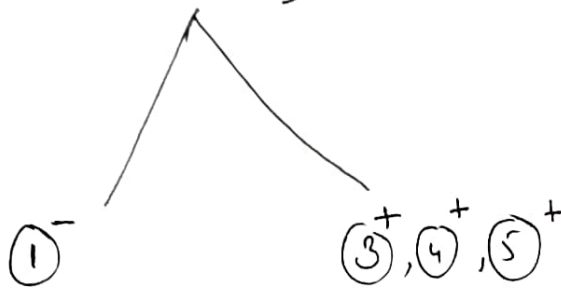
Calculate $G\left(Age < 26.5\right)$ and $G\left(Age < 36.5\right)$

$G\left(Age < 12.5\right) = 0$ 　is the minimum

$G\left(Loves\ Po\ Pcorn\right) = 0.25$

$G\left(Age < 12.5\right) = 0$

$G\left(Age < 12.5\right)$ is the winner

| Loves Soda |
yes / No

| Age < 12.5 |
yes / No

No 　　 yes
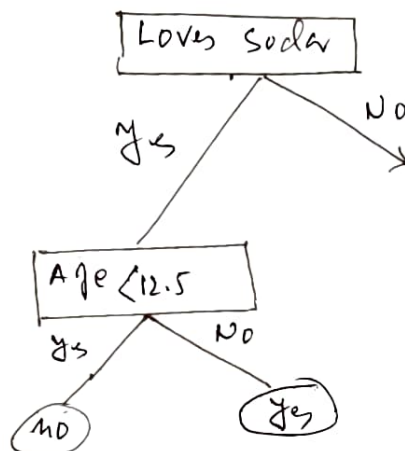
To avoid over fitting we may put a restriction on the number of examples in a ~~leaf~~ node. If a ~~leaf~~ node contains, i.e., 10 or less examples, this ~~leaf~~ node should not be extended further and converted to a leaf. The majority class in the leaf will determine the label of the leaf.

The cross validation approach may be used to find out the appropriate number of examples.

The other more effective approach for avoiding over fitting is tree pruning.