

Part B

Question 1

Yes, if and only if each cluster contains no more the 1 different point.

Question 2

2.1

The search engine will be get an accuracy of 1, if and only if, given q , every document in C will be labeled as 0.

2.2

Let $C_q = [x_1, x_2, \dots, x_n] \mid s.t. x_i \in \{0, 1\}$

The true positives are defined as the amount of labeled relevant files that are actually relevant.

\therefore We label all files as non-relevant, TP and FP will always be equal to zero, but the FN isn't always zero.

$$\begin{aligned} Recall &= \frac{TP}{TP + FN} = \frac{0}{0 + \text{Non-Zero}} = 0 \\ Precision &= \frac{TP}{TP + FP} = \frac{0}{0 + 0} := 1 \end{aligned} \tag{1}$$

\therefore Recall = 0, Precision = 1 (by definition).