

## SEARCH ENGINE

ปัจจุบันระบบการค้นหาเอกสารหรือเว็บไซต์ที่เกี่ยวข้องมีความสำคัญต่อสังคมยุคปัจจุบันเป็นอย่างมาก ระบบการค้นหาที่ดีนั้น นอกจากจะต้องมีความถูกต้องแล้ว ยังจะต้องสามารถเรียงลำดับการนำเสนอของเอกสารที่เกี่ยวข้อง เพื่อให้ผู้ใช้งานได้คัดเลือกเอกสารที่เกี่ยวข้องได้สะดวกที่สุด เช่น เว็บไซต์ google.com ที่พยายามจะหาเว็บไซต์ที่ดีที่สุดมาลิบเว็บไซต์เพื่อแสดงผลในหน้าแรก แม้ว่าจะมีเว็บไซต์ที่เกี่ยวข้องเป็นหมื่นเป็นล้านเว็บไซต์

สำหรับโจทย์ระบบค้นหาที่เราจะทำนั้น จะกำหนด **คะแนนความเกี่ยวข้อง** ของเอกสาร และ **นำเสนอเอกสารที่มีคะแนนความเกี่ยวข้องมากที่สุด** ดังต่อไปนี้

$$\text{คะแนนความเกี่ยวข้อง} = \text{คะแนนความถี่ของคำ} \times \text{คะแนนความจำเพาะของคำ}$$

โดย

$$\text{คะแนนความถี่ของคำ} = \frac{\text{จำนวนครั้งที่คำที่ค้นหาปรากฏในเอกสาร}}{\text{จำนวนคำทั้งหมดในเอกสาร}}$$

$$\text{คะแนนความจำเพาะของคำ} = \frac{1}{\text{จำนวนคำที่ไม่ซ้ำกันทั้งหมดในเอกสาร}}$$

เช่น ถ้าเราต้องการค้นหาคำว่า **APPLE** และมีเอกสารทั้งหมดสองเอกสาร

เอกสาร **PPAP**

คำในเอกสาร I HAVE A PEN I HAVE AN **APPLE** AH PINEAPPLE PEN

เอกสาร **MYAPPLE**

คำในเอกสาร **APPLE** WATCH MACBOOK AIR IPOD **APPLE** ORANGE

เอกสาร **PPAP** มี

- **APPLE** 1 คำ
- คำทั้งหมด 11 คำ
- คำที่ไม่ซ้ำ 8 คำ (I HAVE A PEN AN **APPLE** AH PINEAPPLE)
- ดังนั้น **คะแนนความเกี่ยวข้อง** ของ **PPAP** คือ  $\frac{1}{11} \times \frac{1}{8} = 0.0113636$   
(ข้อสังเกต คำว่า PINEAPPLE มี **APPLE** อยู่ แต่ไม่นับเป็นคำว่า **APPLE**)

เอกสาร **MYAPPLE** มี

- **APPLE** 2 คำ
- คำทั้งหมด 7 คำ
- คำที่ไม่ซ้ำ 6 คำ (APPLE WATCH MACBOOK AIR IPOD ORANGE)
- ดังนั้น **คะแนนความเกี่ยวข้อง** ของ **MYAPPLE** คือ  $\frac{2}{7} \times \frac{1}{6} = 0.047619$

เลือกค่าสูงสุด

**คะแนนความเกี่ยวข้อง** ของ **MYAPPLE** มีค่าสูงสุด จึงแสดงคำว่า **MYAPPLE** ให้ผู้ใช้

หมายเหตุ ในกรณีทดสอบจะไม่มีกรณีที่คะแนนสูงสุดเท่ากัน

## ข้อมูลนำเข้า

แบ่งเป็นสองส่วน

### 1. ส่วนเอกสาร

เริ่มด้วยตัวเลขจำนวนเอกสารที่จะใส่เข้าระบบ หนึ่งบรรทัด

ตามด้วยชื่อเอกสาร (หนึ่งบรรทัด) และข้อมูลในเอกสาร (หนึ่งบรรทัด) จนครบจำนวนเอกสารที่ระบุ

### 2. ส่วนค้นหา

ผู้ใช้จะใส่คำที่ต้องการค้นหาทีละบรรทัด ซึ่งโปรแกรมจะต้องแสดงผลการค้นหาทันที และรอรับคำค้นหาคำถัดไป หาก

ผู้ใช้ใส่คำว่า **-1** ให้จบโปรแกรม

(ตัวอักษรภาษาอังกฤษของข้อมูลนำเข้าจะเป็นตัวพิมพ์ใหญ่ทั้งหมด)

## ข้อมูลส่งออก

ชื่อเอกสารที่มีคะแนนสูงสุด หากคะแนนความเกี่ยวข้องที่สูงที่สุดคือ 0 ให้แสดง **NOT FOUND**

## ตัวอย่าง Input Output

### Input (จากแป้นพิมพ์)

3  
PPAP  
I HAVE A PEN I HAVE AN APPLE AH PINEAPPLE PEN  
MYAPPLE  
APPLE WATCH MACBOOK AIR IPOD APPLE ORANGE  
GOOGLE  
GOOGLE MY WATCH AND GET MACBOOK AIR AND FREE I PHONE WATCH YOUTUBE  
PEN  
GOOGLE  
APPLE  
I  
AKB  
-1

### Output (ทางจอภาพ)

PPAP  
GOOGLE  
MYAPPLE  
PPAP  
NOT FOUND