

Modelo parcial BD 2:

1. usuarios (uid, nombre)
 fotos (fid, dueño, tamaño)
 comentarios (cid, autor, foto, texto)

• Revisar:

- APTME Aula
- clase 2 procl. low 2021

$\Pi_{fid, uid} (fotos \bowtie \sigma_{nombre = 'Luis Pérez'} (usuarios))$

- Id prim. en nombre para usuario.
- Usuario con 1000 tuplas
- bloque 4096 B
- 10 000 nombres dif. usuario
- Mismo nombre no repite mas de 15 veces.

Desarrollo:

Cant. Claves Busqueda:

$$tanBlaq. = N \cdot tanCB + (N+1) \cdot tanPuntero$$

$$4096 \quad \downarrow \quad 50$$

cantCB ↓

chars nombre
(Es arbitrario)

$$6 \text{ u } 8$$



cent bits direccion?
(Arbitrario)

Profe dice que da igual el valor mientras sea razonable.

Despejo N:

$$4090 = 56N \Rightarrow N = 4090/56 = \underline{73} \rightarrow \text{cantCB}$$



redondeo para abajo

$$f_{char} = 10 \text{ B}$$

Altura árbol:

$$\text{alt Arbol} = \left\lceil \log_{\left\lfloor \frac{M}{2} \right\rfloor} K \right\rceil = \left\lceil \log_{\left\lfloor \frac{75}{2} \right\rfloor} 10000 \right\rceil = \left\lceil \log_{37} 10000 \right\rceil$$

\downarrow
 Cant Nombres

$$= \left\lceil \log_{10} 1000 \right\rceil = 2,55 = 3$$

Operador Físico:

op Físico = sel. para igualdad en no clave candidato, con IndPrim

h: transf Bloques para llegar primer registro con información = alt Arbol

b: transf Bloq. que contienen registros con valor buscado.

↳ calculamos costo superior

6 bits alcanza para direccionar

$$\text{cant Tuplas Bloq.} : 4096 = N \cdot (\text{tam Nombre} + \text{tam bid} + \text{tam Puntero})$$

$$= N \cdot (50 + 6 + 6)$$

$$= N \cdot 62$$

$$N = \frac{4096}{62} = 66$$

↑
ver cuanto espacio
se puede direccionar.
Siempre es el mismo para
toda la tabla.

⇒ b = 2 en peor caso. Ya que como mucho los nombres se

$$\text{Costo} = h + b = 3 + 2 = 5$$

guardan en 2 bloques, ya que solo
hay como mucho 15 con mismo nombre
Y puede que no todos nombres estén
mismo bloque.

2.3. - Res. ~~ord~~ ord por relevancia

- Modelo espacio vec. - Docs en HDD
- Ind. invertido
- Consultas bool



j : ID doc.
 f : frecuencia o $\frac{TF \cdot IDF}{\text{Max relevancia}}$

Relevancia: $\sum_{j \in Q} TF_{j,i} \cdot IDF_j$

b) Listar Parar

Relevancia de docs con similitud vectores

- Colección en HDD
- Modelo esp. Vec
- Cons. bool \rightarrow para cada termino en que docs estan

Parar proc. una consulta?

1. Encontrar vocab consulta
2. Buscar terminos consultas en ind. invertido $II(i)$, termino de \emptyset
3. Calcular $S(i) = \{\text{docs con termino } i\}$
4. Evaluar expresion de consulta usando operadores asociados a conectivos booleanos
5. Calcular vector consulta Q
6. Calc. relevancia docs usando similitud. \rightarrow formula coseno de vectores
7. Ordenar docs por relevancia

Ind. Inv
 \uparrow

3. Identificar tipos contexto

Mientras tomaba mate en la plaza principal de mi pueblo

(plaza → principal) → mi pueblo

escuchando el sonido de las campanas de la iglesia

recorde las protestas globales por el cambio climático.

Pense en nuestros ancestros, que vivían en armonía con la

naturaleza, y en como nuestra sociedad actual ha perdido

esa conexión. me pregunté si volveríamos a esos tiempos

o si seguiremos repitiendo los mismos errores del pasado.

Contextos:

- Local → Contexto oración u oraciones cercanas

(plaza → principal) → mi pueblo

sonido → campanas, cambio → climático, conexión → naturaleza

- Global: vivir en armonía con la naturaleza.

- situacional: plaza principal de mi pueblo, ~~estaba~~ escuchando campanas iglesia.

tomaba mate (cultural), protestas globales, si volveremos a esos tiempos (emocional)

Si no son muy claras mejor no ponerlas

- Histórico: conocimiento acumulado sobre un tema.
 - ^{incertidumbre} Vivien en armonía con la naturaleza.
 - Errores del pasado.
 - Cambio climático (conocimiento acumulado)

(puedo faltar alguno más)

- Cultural: intersección a situacional? → cambio foco.
- Tomeba mate → ~~mapa~~ Sudamérica.
- Campanas de teleria → país cristiano?

4) Conecta NLP con lo que hacen los Transformers

a) Reconocimiento intenciones: NO explícitamente

- Forma parte identificación contexto. → ~~de~~ ~~de~~
- Participa fuertemente capa autoatención.

b) Determinación contenido: No explícitamente

- Sin embargo, se determina una palabra por vez del contenido, teniendo en cuenta todas las anteriores de la secuencia, input y contexto.
- La capa lineal participa en esto.

c) (~~es~~ similar a anterior)

d) Post procesamiento: No hay capa en org.

- A veces se ~~se~~ agrega en chatbots

centNodos/Arbol: punteros tuplas hojas