

:1 דף

ישל 0 bit אבסורד & Jaccard similarity של (a)

: פונקציה (מרחב המרחב) min-hash - ה

$$\Pr[h_T(c_1) = h_T(c_2)] = \text{sim}(c_1, c_2) = 0$$

פונקציה minhashing-ה של, $h_T(c_1) \neq h_T(c_2)$: פס

: Jaccard similarity-ה מרחב הקבוצות

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 1 & 0 \end{bmatrix}$$

: מרחב הקבוצות

$$h_1(x) = x$$

(b) המרחב הקבוצות

$$h_2(x) = [x+1]_3$$

: מרחב הקבוצות

$$h_3(x) = [x+2]_3$$

: 3 מרחב הקבוצות

$$h_1(r_3) = 1 \ 0$$

$$h_2(r_3) = 2 \ 1$$

$$h_3(r_3) = 0 \ 2$$



: 1 מרחב הקבוצות

$$1 \quad 1$$

$$2 \quad 2$$

$$0 \quad 0$$

: 2 מרחב הקבוצות

$$h_1(r_2) = 1 \ 0$$

$$h_2(r_2) = 2 \ 1$$

$$h_3(r_2) = 0 \ 2$$



: 1 מרחב הקבוצות

$$1 \quad 1$$

$$2 \quad 2$$

$$0 \quad 0$$

: 1 מרחב הקבוצות

$$h_1(r_1) = 1 \ 1$$

$$h_2(r_1) = 2 \ 2$$

$$h_3(r_1) = 0 \ 0$$



: 2 מרחב הקבוצות

$$1 \quad 1$$

$$2 \quad 2$$

$$0 \quad 0$$

$$\text{Jaccard} = \frac{1}{3}$$

: פס

: sim = 1

: פס

$$1 \quad 1$$

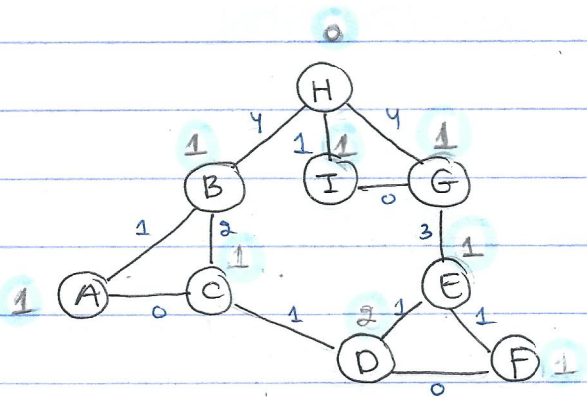
$$2 \quad 2$$

$$0 \quad 0$$

: מרחב הקבוצות

: פס של המרחב הקבוצות של המרחב הקבוצות של המרחב הקבוצות

: Jaccard similarity של המרחב הקבוצות

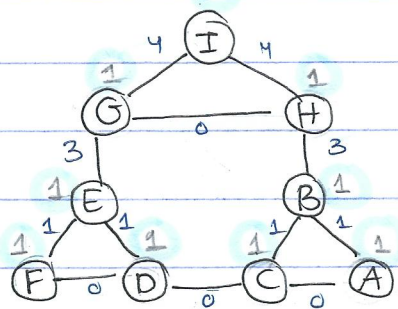


2 דיליט

shortest paths - ה' ע'נ .1 (2)

פיראמל, H זיכרונ

(סימא) - א' ב' ג'

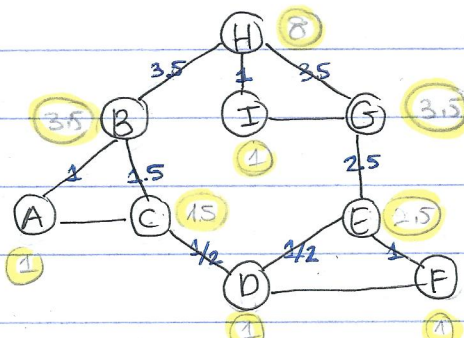
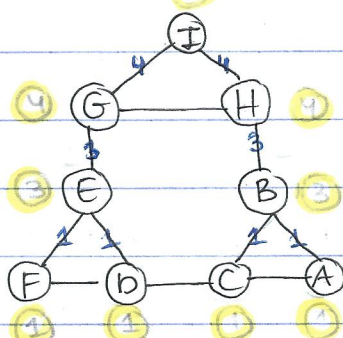


shortest paths - ה' ע'נ .2

פיראמל, I זיכרונ

(סימא) - א' ב' ג'

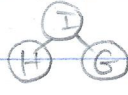
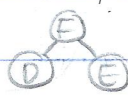
ה' ע'נ (פיראמל) - א' ב' ג' (2) (3) (4) (5) (6) (7) (8) (9) (10) (11) (12) (13) (14) (15) (16) (17) (18) (19) (20) (21) (22) (23) (24) (25) (26) (27) (28) (29) (30) (31) (32) (33) (34) (35) (36) (37) (38) (39) (40) (41) (42) (43) (44) (45) (46) (47) (48) (49) (50) (51) (52) (53) (54) (55) (56) (57) (58) (59) (60) (61) (62) (63) (64) (65) (66) (67) (68) (69) (70) (71) (72) (73) (74) (75) (76) (77) (78) (79) (80) (81) (82) (83) (84) (85) (86) (87) (88) (89) (90) (91) (92) (93) (94) (95) (96) (97) (98) (99) (100)



$$\{(A,B), (A,C), (E,D), (E,F), (G,I), (I,H)\} = \frac{8}{2}, \{(B,C), (H,G), (D,E)\} = \frac{13}{2}, \{(B,H), (C,D), (G,E)\} = \frac{19}{2}$$

threshold = 6 (c)

$\{A,B,C\}, \{D,E,F\}, \{G,H,I\}$



ה' ע'נ (פיראמל) - א' ב' ג' (2) (3) (4) (5) (6) (7) (8) (9) (10) (11) (12) (13) (14) (15) (16) (17) (18) (19) (20) (21) (22) (23) (24) (25) (26) (27) (28) (29) (30) (31) (32) (33) (34) (35) (36) (37) (38) (39) (40) (41) (42) (43) (44) (45) (46) (47) (48) (49) (50) (51) (52) (53) (54) (55) (56) (57) (58) (59) (60) (61) (62) (63) (64) (65) (66) (67) (68) (69) (70) (71) (72) (73) (74) (75) (76) (77) (78) (79) (80) (81) (82) (83) (84) (85) (86) (87) (88) (89) (90) (91) (92) (93) (94) (95) (96) (97) (98) (99) (100)

$$\frac{5 - \sqrt{13}}{2} \approx 0.697$$

ה' ע'נ (פיראמל) - א' ב' ג' (2) (3) (4) (5) (6) (7) (8) (9) (10) (11) (12) (13) (14) (15) (16) (17) (18) (19) (20) (21) (22) (23) (24) (25) (26) (27) (28) (29) (30) (31) (32) (33) (34) (35) (36) (37) (38) (39) (40) (41) (42) (43) (44) (45) (46) (47) (48) (49) (50) (51) (52) (53) (54) (55) (56) (57) (58) (59) (60) (61) (62) (63) (64) (65) (66) (67) (68) (69) (70) (71) (72) (73) (74) (75) (76) (77) (78) (79) (80) (81) (82) (83) (84) (85) (86) (87) (88) (89) (90) (91) (92) (93) (94) (95) (96) (97) (98) (99) (100)

$$[-0.58 \quad -0.38 \quad -0.38 \quad 0.08 \quad 0.29 \quad 0.29 \quad 0.29 \quad 0.08 \quad 0.29]$$

ה' ע'נ (פיראמל) - א' ב' ג' (2) (3) (4) (5) (6) (7) (8) (9) (10) (11) (12) (13) (14) (15) (16) (17) (18) (19) (20) (21) (22) (23) (24) (25) (26) (27) (28) (29) (30) (31) (32) (33) (34) (35) (36) (37) (38) (39) (40) (41) (42) (43) (44) (45) (46) (47) (48) (49) (50) (51) (52) (53) (54) (55) (56) (57) (58) (59) (60) (61) (62) (63) (64) (65) (66) (67) (68) (69) (70) (71) (72) (73) (74) (75) (76) (77) (78) (79) (80) (81) (82) (83) (84) (85) (86) (87) (88) (89) (90) (91) (92) (93) (94) (95) (96) (97) (98) (99) (100)

$\{A,B,C\}, \{D,E,F,G,H,I\}$

$x_{j-1}, x_{j+1} \in C$ and $x_j \notin C$: e.g. x_j

\tilde{X} 'ngg' c (k' center \rightarrow sb' b'ro) : center (an) xj p/c (*)

cluster - i will ω_i is center μ_i x_{j-1}, x_{j+1} "pre par" $k \in$

1. $p \cdot p \wedge N$ x_{j-1} de center $\rightarrow \tilde{x} \in p \wedge N$ $x_j < \tilde{x} : e$ 3.2.2 (11)

$$|x_j - x_{j-1}| \leq |\tilde{x} - x_{j-1}| \leq \delta \quad |x_{j-1}| \leq x_j \leq \tilde{x} \quad \text{p.s.} \quad |\tilde{x} - x_{j-1}| \leq |x_j - x_{j-1}|$$

המחיר

x_i - is the center of the i -th PLC ; PLC (*)

1. $P \cap N \neq \emptyset$, $\hat{X} = P \cap N$, X_{i-1}, X_{i+1} (b) Center: $-\infty \leq x \leq \infty$

∴ $P_{\text{PAN}} \text{ PSHC} \cdot |x_{j-1} - x'| > |x_j - \tilde{X}|$ (ii)', $|x_j - x'| < |x_j - \tilde{X}|$ (i)

$$|x_j - \tilde{x}| = |x_j - x_{j-1}| + |x_{j-1} - \tilde{x}| \underset{(ii)}{<} |x_j - x_{j-1}| + |x_{j-1} - x| = |x_j - x|$$

(ii) - δ נבחר כך ש $|x_j - \tilde{x}| < |x_j - x|$ וכל הבהרה

← וכל מקום קיבלו מידה צדקה, עם החזקה האבסולוטי של מקבל

• IIR yop $\epsilon_{3 \times N}$ cluster bed

የጥገናው ዓላማ (d)

(*) נ"ל כי $x_1 \leq \dots \leq x_n$ וכן \bar{x} הוא ממוצע x_1, \dots, x_n ולכן $\bar{x} \in [x_1, x_n]$.

← נתונה מטריצה M בודות $n \times K$ ובה n שורות ו- K עמודות, \hat{c}_k cluster-ה k ו- μ_k ממוצע cluster-ה k של M .

Centroid = $(\sum_{i=j}^n x_i / (n-j))$; P_1, P_2, \dots, P_n ; $j \in \{1, \dots, n\}$; $k=1$; $k=2$; $k=3$; $k=4$; $k=5$; $k=6$; $k=7$; $k=8$; $k=9$; $k=10$; $k=11$; $k=12$; $k=13$; $k=14$; $k=15$; $k=16$; $k=17$; $k=18$; $k=19$; $k=20$; $k=21$; $k=22$; $k=23$; $k=24$; $k=25$; $k=26$; $k=27$; $k=28$; $k=29$; $k=30$; $k=31$; $k=32$; $k=33$; $k=34$; $k=35$; $k=36$; $k=37$; $k=38$; $k=39$; $k=40$; $k=41$; $k=42$; $k=43$; $k=44$; $k=45$; $k=46$; $k=47$; $k=48$; $k=49$; $k=50$; $k=51$; $k=52$; $k=53$; $k=54$; $k=55$; $k=56$; $k=57$; $k=58$; $k=59$; $k=60$; $k=61$; $k=62$; $k=63$; $k=64$; $k=65$; $k=66$; $k=67$; $k=68$; $k=69$; $k=70$; $k=71$; $k=72$; $k=73$; $k=74$; $k=75$; $k=76$; $k=77$; $k=78$; $k=79$; $k=80$; $k=81$; $k=82$; $k=83$; $k=84$; $k=85$; $k=86$; $k=87$; $k=88$; $k=89$; $k=90$; $k=91$; $k=92$; $k=93$; $k=94$; $k=95$; $k=96$; $k=97$; $k=98$; $k=99$; $k=100$; $k=101$; $k=102$; $k=103$; $k=104$; $k=105$; $k=106$; $k=107$; $k=108$; $k=109$; $k=110$; $k=111$; $k=112$; $k=113$; $k=114$; $k=115$; $k=116$; $k=117$; $k=118$; $k=119$; $k=120$; $k=121$; $k=122$; $k=123$; $k=124$; $k=125$; $k=126$; $k=127$; $k=128$; $k=129$; $k=130$; $k=131$; $k=132$; $k=133$; $k=134$; $k=135$; $k=136$; $k=137$; $k=138$; $k=139$; $k=140$; $k=141$; $k=142$; $k=143$; $k=144$; $k=145$; $k=146$; $k=147$; $k=148$; $k=149$; $k=150$; $k=151$; $k=152$; $k=153$; $k=154$; $k=155$; $k=156$; $k=157$; $k=158$; $k=159$; $k=160$; $k=161$; $k=162$; $k=163$; $k=164$; $k=165$; $k=166$; $k=167$; $k=168$; $k=169$; $k=170$; $k=171$; $k=172$; $k=173$; $k=174$; $k=175$; $k=176$; $k=177$; $k=178$; $k=179$; $k=180$; $k=181$; $k=182$; $k=183$; $k=184$; $k=185$; $k=186$; $k=187$; $k=188$; $k=189$; $k=190$; $k=191$; $k=192$; $k=193$; $k=194$; $k=195$; $k=196$; $k=197$; $k=198$; $k=199$; $k=200$; $k=201$; $k=202$; $k=203$; $k=204$; $k=205$; $k=206$; $k=207$; $k=208$; $k=209$; $k=210$; $k=211$; $k=212$; $k=213$; $k=214$; $k=215$; $k=216$; $k=217$; $k=218$; $k=219$; $k=220$; $k=221$; $k=222$; $k=223$; $k=224$; $k=225$; $k=226$; $k=227$; $k=228$; $k=229$; $k=230$; $k=231$; $k=232$; $k=233$; $k=234$; $k=235$; $k=236$; $k=237$; $k=238$; $k=239$; $k=240$; $k=241$; $k=242$; $k=243$; $k=244$; $k=245$; $k=246$; $k=247$; $k=248$; $k=249$; $k=250$; $k=251$; $k=252$; $k=253$; $k=254$; $k=255$; $k=256$; $k=257$; $k=258$; $k=259$; $k=260$; $k=261$; $k=262$; $k=263$; $k=264$; $k=265$; $k=266$; $k=267$; $k=268$; $k=269$; $k=270$; $k=271$; $k=272$; $k=273$; $k=274$; $k=275$; $k=276$; $k=277$; $k=278$; $k=279$; $k=280$; $k=281$; $k=282$; $k=283$; $k=284$; $k=285$; $k=286$; $k=287$; $k=288$; $k=289$; $k=290$; $k=291$; $k=292$; $k=293$; $k=294$; $k=295$; $k=296$; $k=297$; $k=298$; $k=299$; $k=300$; $k=301$; $k=302$; $k=303$; $k=304$; $k=305$; $k=306$; $k=307$; $k=308$; $k=309$; $k=310$; $k=311$; $k=312$; $k=313$; $k=314$; $k=315$; $k=316$; $k=317$; $k=318$; $k=319$; $k=320$; $k=321$; $k=322$; $k=323$; $k=324$; $k=325$; $k=326$; $k=327$; $k=328$; $k=329$; $k=330$; $k=331$; $k=332$; $k=333$; $k=334$; $k=335$; $k=336$; $k=337$; $k=338$; $k=339$; $k=340$; $k=341$; $k=342$; $k=343$; $k=344$; $k=345$; $k=346$; $k=347$; $k=348$; $k=349$; $k=350$; $k=351$; $k=352$; $k=353$; $k=354$; $k=355$; $k=356$; $k=357$; $k=358$; $k=359$; $k=360$; $k=361$; $k=362$; $k=363$; $k=364$; $k=365$; $k=366$; $k=367$; $k=368$; $k=369$; $k=370$; $k=371$; $k=372$; $k=373$; $k=374$; $k=375$; $k=376$; $k=$

$$\text{pay}(k, x_1, x_n) = \sum_{j=1}^n \min_{i \in [k]} \|x_j - c_i\|^2$$

$O(n) \rightarrow \text{pay}(1, x_j, x_n) \text{ w/c prg31: p51. (Centroid} \rightarrow \text{map p1}$

(ii) פונק' הקורסיב: נבחן אם אנו יוצאים אל החלוקה הניסית מה הטל

אופן הקצאה של 4 מרכזים לבד, clusters $(K-1) = 5$, $[x_1, x_m]$

איבראןאן בן יפ. $[X_M, X_n]$: פ'לען יפן נא 30% cluster - 8

לפניי M נהנו אל פאספ'א . ps :

$$\min \{ \text{pay}(k, x_1, x_n) \} = \min_{1 \leq M \leq n-1} \{ \min \{ \text{pay}(k-1, x_1, x_M) \} + \text{pay}(1, x_M, x_n) \}$$

(*) קבוצת היתן נמצאת אף הטבלה ולקוק אתר קטנים ה- M שנה לבחור

PN 600 - m

21/02/2021

4

(iii) $O(n) - P, K=1$ מספר הסיבוכים $O(n)$

$O(k \cdot \underline{n^3})$: $k \neq 1$ ופחות $O(n^2)$: $k=1$ ויותר

[illegible]

$\text{pay}(k-1, x_1, x_M)$ כר מ' קטן מ' ולק' נ'ש ר- $\phi(M)$ sk1

משפט 1.1.1. \mathcal{A} הוא אלגברה של σ -אלגברות. \mathcal{A} הוא אלגברה של σ -אלגברות.

"קרי $O(n)$ משמעותו קטטה, וכך נקטן את n הריבוי

$$e, 3, 2, O(n^2) + O((kn)n) = O(kn^2) : f$$



: p_n & p_{n+1}