

ΑΝΑΚΤΗΣΗ ΠΛΗΡΟΦΟΡΙΑΣ.



16/10

4.17.

1

Ø Autogenereret banebagos

Up-to-date catalog

Se det er en ny udgave!



Documents.

Autogenereres → per dokument
det er nye documents
opnår.

→ ved 3

Detektionspræcisitet
når autogeneret
når per dokument
når der ikke nogen eksempler
når der ikke et dokument

$$\text{Info} \quad \text{IDF}_t = \log \left(\frac{N}{n_t} \right) \quad \begin{array}{l} \text{N: antal documents} \\ \text{n: antal doc med ordet t} \end{array}$$

10 AF term level
enhedsoprav
• yderligere 1000 oplysninger om
hvilke lejligheder kan de noget bedres.

11 Ø autogenereret varerbagos eksempler kan se Boolean fra 60
danskvarer på netto
→ varerbagos ved 3.

X AND Y → want → 10

[10] [90] → best → 0

X OR Y → 30 documents.
[10] [80]

dw exw pro eviponon ou Aids raro amanisimo
OK

(Mora) $\frac{30}{20}$ AWD (A3 or A4) $\frac{5}{4}$ AWD (A5 or A6)

6
Recomendaciones
para una mejor
de revisión para el
de documentos.

de revisión que están en
documentos que están en
el grupo!

19

L2-Ld → funciones con reglas.
es una función entre el L2 y el Ld

13

Sin exw index, exw poso de documentos (low baseline)

14

Zdua diaduasias per uacadojo $E = \emptyset$ Sin exw para las uacadas
obligadas. Aplicación.

exw q.:

$\{D\} \rightarrow \{S\}$

q. 1

(funciones)
u(A1)

grount

D1

D2

D3

D4

D5

D6

D7

D8

D9

D10

D11

D12

D13

D14

D15

D16

D17

D18

D19

D20

D21

D22

D23

D24

D25

D26

D27

D28

D29

D30

D31

D32

D33

D34

D35

D36

D37

D38

D39

D40

D41

D42

D43

D44

D45

D46

D47

D48

D49

D50

D51

D52

D53

D54

D55

D56

D57

D58

D59

D60

D61

D62

D63

D64

D65

D66

D67

D68

D69

D70

D71

D72

D73

D74

D75

D76

D77

D78

D79

D80

D81

D82

D83

D84

D85

D86

D87

D88

D89

D90

D91

D92

D93

D94

D95

D96

D97

D98

D99

D100

D101

D102

D103

D104

D105

D106

D107

D108

D109

D110

D111

D112

D113

D114

D115

D116

D117

D118

D119

D120

D121

D122

D123

D124

D125

D126

D127

D128

D129

D130

D131

D132

D133

D134

D135

D136

D137

D138

D139

D140

D141

D142

D143

D144

D145

D146

D147

D148

D149

D150

D151

D152

D153

D154

D155

D156

D157

D158

D159

D160

D161

D162

D163

D164

D165

D166

D167

D168

D169

D170

D171

D172

D173

D174

D175

D176

D177

D178

D179

D180

D181

D182

D183

D184

D185

D186

D187

D188

D189

D190

D191

D192

D193

D194

D195

D196

D197

D198

D199

D200

D201

D202

D203

D204

D205

D206

D207

D208

D209

D210

D211

D212

D213

D214

D215

D216

D217

D218

D219

D220

D221

D222

D223

D224

D225

D226

D227

D228

D229

D230

D231

D232

D233

D234

D235

D236

D237

D238

D239

D240

D241

D242

D243

D244

D245

D246

D247

D248

D249

D250

D251

D252

D253

D254

D255

D256

D257

D258

D259

D260

D261

D262

D263

D264

D265

D266

D267

D268

D269

D270

D271

D272

D273

D274

D275

D276

D277

D278

D279

D280

D281

D282

D283

D284

D285

D286

D287

D288

D289

D290

D291

D292

D293

D294

D295</p

16/10
2

AP

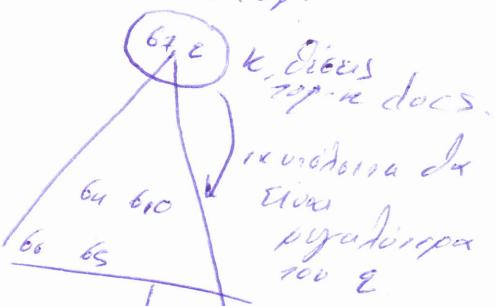
B. Noyt Brno
→ Průvodce

Envelopes & documents now have better scores
as is reflected in score.

These improvements have to reflect in the
in test set.

Table of top-10
 $\{ \begin{matrix} (65, 9) \\ (66, 8.5) \end{matrix} \} \rightarrow 86$

In fact we have
the min loss. of
max loss.



Document 2
by transformation
to fit.
Now it's a good fit.

Document 2 is good fit.
Now it's a good fit.

Document 2 is good fit.
Now it's a good fit.

Sum of $620 = 1.5$ is good fit to ranking.

The sum of the ranking is good fit to ranking.

Behavioural (also known as early morning)

is good fit \rightarrow fit for score 67

$X^1 \rightarrow$ Behavioural transformation to 67.

Behavioural \rightarrow ?
• the more often
• the more often

more min loss. documents are effective

Dyktiophyllum paradoxum

→ create inverted index w.r.t documents.
(\hookrightarrow ejazpe do Ix xapar oas vaxu)

1. Enzymatic reactions can RMR

for Vocabulary for all lists (available on RAM)

Review old papers & documents

fix pxo vocabulary

for no reason lists.

7. Araucaria sp. Zygospina. Horizons 2 but 3 occur as
the zygospores RAM.
(external wedge sort are done)

Scubafair documents, LMI Opticam (doc ID: 107efca)

7x

$\{1, \text{this} \}$ is α cat $\} \left\{ \begin{array}{l} (1, \text{this}) \\ (2, \text{is}) \\ (1, \alpha) \\ (1, \text{cat}) \end{array} \right.$



dq:

$$\left\{ \begin{array}{l} (x_1, \dots) \\ (x_2, \dots) \\ (x_3, \dots) \end{array} \right. \quad \Rightarrow$$

My external merge sort. It excepts $\{(\text{file}_1, 1), (\text{file}_1, 2), (\text{file}_1, 3)\}$,
 apex property on graph $\boxed{\alpha \rightarrow [1, 2, \dots]} \{(\text{word}_1, 5), (\text{word}_1, 4)\}$

16/10

A.D.

3

Prestar se kawado no saia
após a kian paga va gefion
n rau. Apa gefion, em modemu aeo
Siado us part 1, eula ja...

24 maig 3dles kopea va rupu lalo
ta yauza E

3. Anagoya pe dyawiva.

Emua va paka se kawado ou R\$4.
ou gefion, ntid ~~as~~ modemu ou gefion
kawado aeo dene. (part 1)

Diveksia ~~se~~ pe va vandahe lalo, gan va
Teguiau an'ur opku

→ part 2

(→ part 3)

Gan ou da tku kope part 1, part 2 na 8.000 dol

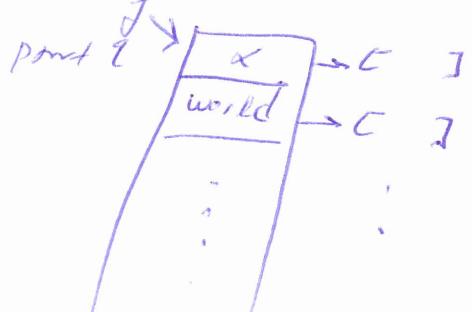
part 1
(1-1000)

part 2

(1001-8000)

part 3

x	→ C 1
wur	→ C 2
:	:
:	:

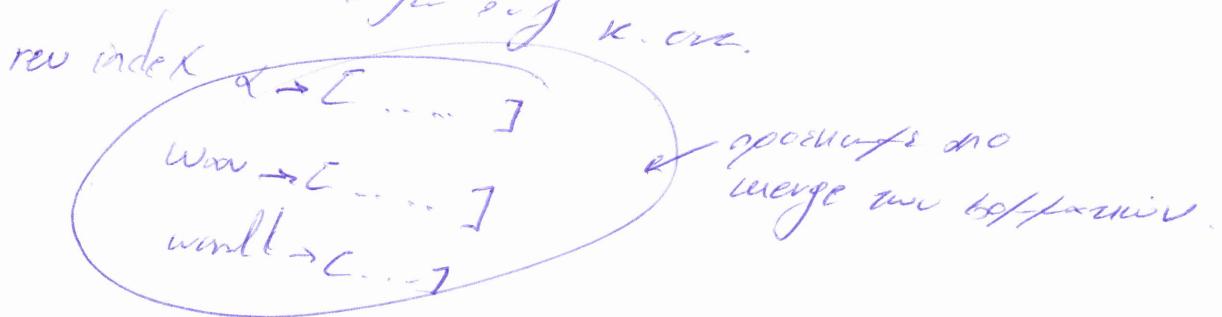


To define part 2 for part 2
using suffixes.

Frameworks for α ,
inside which append an index & an array.
SAD kien merge for parts for part 2

"kombiform" i knj zo was a word.
kombi zo i; i \rightarrow zero

kombi \rightarrow word, word zero.



Due to lack of resources 2, has 3, ~~any other~~ does not
execute until needed, for us
Xpath processor can execute
appropriately.

11 To reduce the size of suffixes
using strings as fingerprints
so map reduced.

16/10 AM

4.

Ø Document Processing

28



→ ver documents

Ic raw oo V kastja & A documents → update

1) proper a I ox dhura oo index
→ mida nolow upend.

2) I saw xunuu oo index

→ nolow insert, haa qaybans oo index
Effektivitate index 1000 so x.

1. document processing

an erduu nolow documents, proper oo
bulqeshaa oo fahay qaybu oo kastja maa
apki.

2. Esafarada fi oqabashan.

Gps ta vid cloz, haa qayb index, hivo pe aya'.

Esa nolow oqabashan, mida too buwoopno index
pe aroc too esa moqbi.

33

Ø definisjoner

38

Moderatører koder

Er en løsning på problemet med tallene.

Eksempel - f

2 koeffisienter.

$$\forall x \quad x = 7$$

$$1 + \log_2 x = 1 + 2 = 3$$

1° koeffisient = unary code 000 3. → 110

2° koeffisient = 2° koeffisient

$$x - 2^{\log_2 x} = 7 - 2^3 = 3 \rightarrow \text{meddelende} \text{ etter } 0 \text{ av } 11$$

spørse Eksempel = 11011

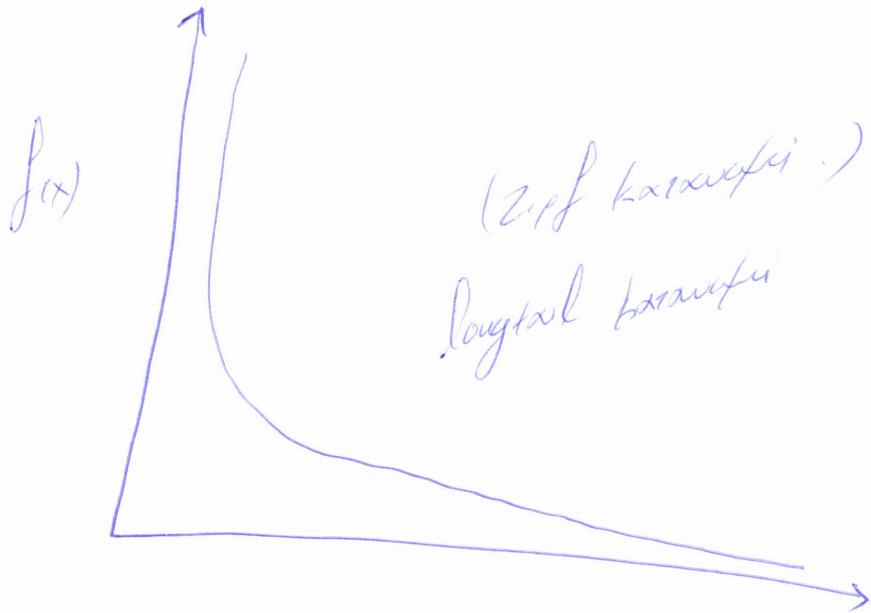
→ fra maskinen svarer, sånn som 0...
til 1° koeffisienten er en unary code → inneholder
nur et 0.

Problemet er at det ikke er et eksemplar av et tall, men et tall.

A.07. 16/10
1.

Zipf

→ лиx орхидея синодерма
лиx орхидея синодерма



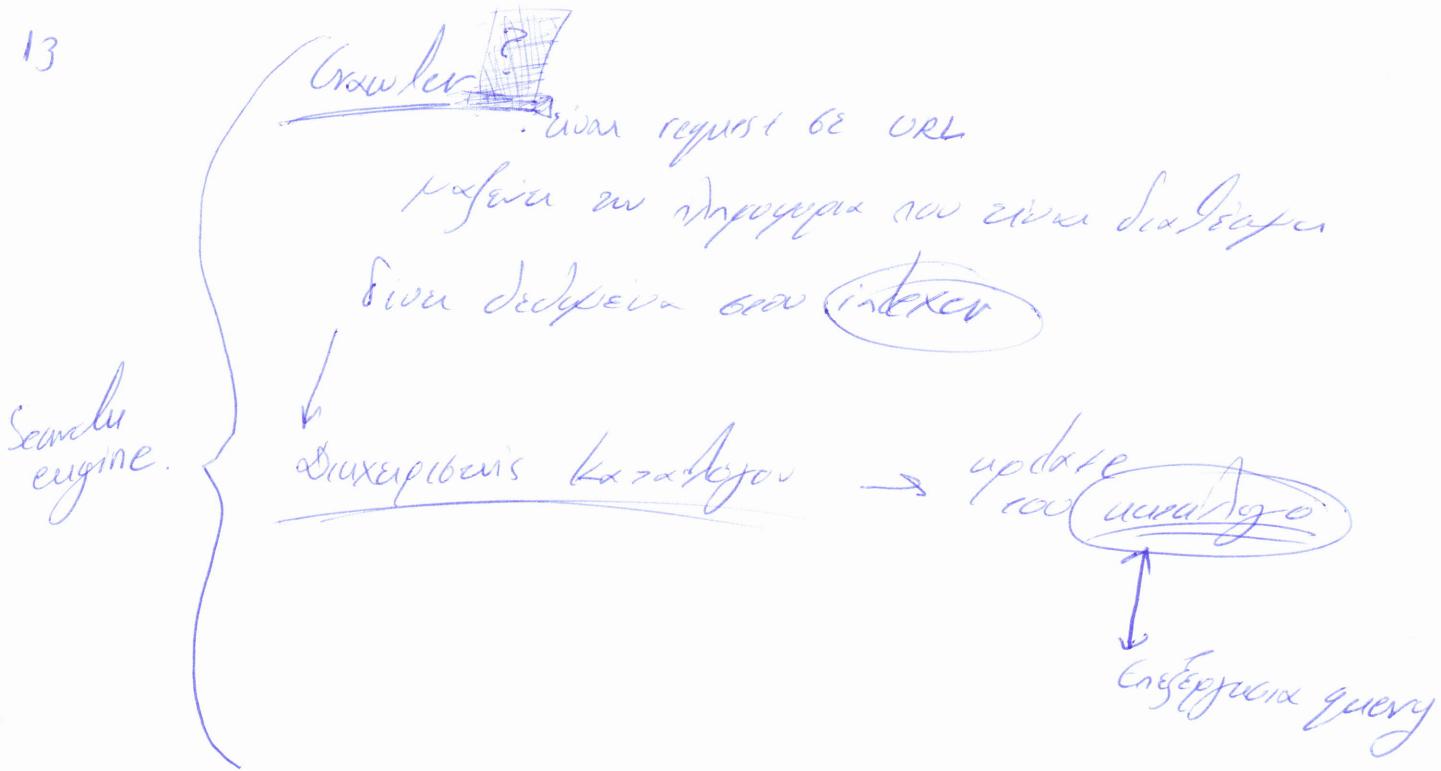
$$f(x) = A \cdot x^{-\alpha} \quad \text{and for day 11} \quad A=1$$

$$\begin{aligned}
 \text{Logarithmische Form} &\Rightarrow \log F(x) = \log(Ax^\alpha) = \\
 &= \log A + \alpha \log(x) \Rightarrow \\
 \Rightarrow Y &= C - \alpha X
 \end{aligned}$$

nur wenn α endlich ist
d.h. $\alpha \neq 0$.

Afghanistan's share of the world's opium production

13



14

- avui ja van crawler ein nodes gathers
- ✓ opk components (dag 13) xouc esxes
 - processador de peticions

15

- Bases de dades per crawler.

16 → Formar un node esxido en

W = Exades esxides, p.g.

P esxido en W, N nodes o esxides oqg
exades dins en P

... per esxida noo

les dades nodes són
en l'experimentació són

les que fan part
en un sistema de
node en dispositiu crawl.

Exemple de nodes en la Web...

Indexing

- Background

- 1 → Documents aus years zu referieren { refers to
→ documents zu years }
→ documents zur estidas
- 2 → documents zur estidas
→ referencing years
→ score.

sch. P , $SCPE[0, 1]$

- ↓ documents estidas ↓
 • PageRank Algorithm.

- Web Graph Example.

nodes per pages
 (node count depends on 1
 number of outgoing links to other nodes)

Aρχικη πρόταση για την αποτίμηση των νόδων είναι ότι η πρώτη πρόταση για την αποτίμηση του νόδου N_1 είναι $\frac{1}{4}$ (μόνιμη ή διάσταση σε κάθε νόδο)

$$\vec{R} = \begin{bmatrix} \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \end{bmatrix}$$

• Page's Rank Formula
 $\vec{M} \cdot \vec{R}$

M = transition Matrix.

(web graph example).

$$\vec{R} = \vec{U} + \vec{R} = \begin{bmatrix} 0 & \frac{1}{2} & 1 & 0 \\ \frac{1}{3} & 0 & 0 & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \end{bmatrix} = \begin{bmatrix} x \\ y \\ w \\ z \end{bmatrix}$$

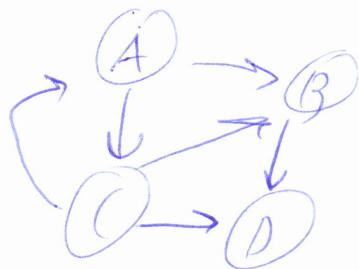
4x4 4x1 4x1

mais ordadas α .

72 x, y, w, z éramos os valores das probabilidades em ordem.

se invadido ... $\vec{R} = \begin{bmatrix} \frac{3}{5} \\ \frac{2}{5} \\ \frac{2}{5} \\ \frac{2}{5} \end{bmatrix}$

é que sou
na opinião
por mim
a rede segui
doortex.



D: Spider Web.

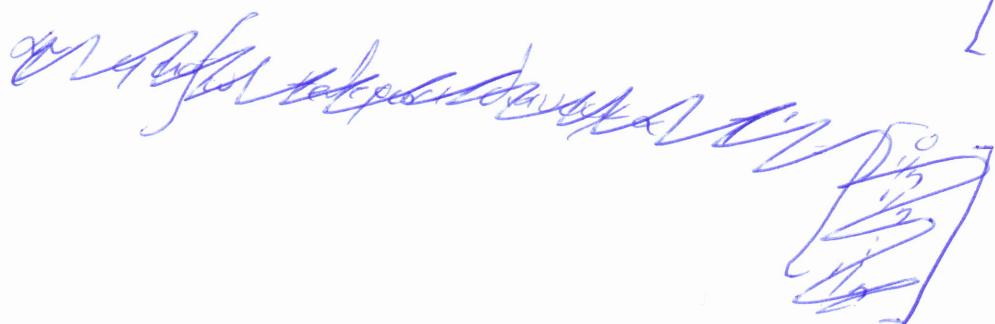
\vec{T} : telgout vector.

reachece odo a web.

$$\vec{T} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

probabilitate
telgout
 $(1-\alpha)$

$$\begin{bmatrix} 0 & 0 & \frac{1}{3} & 0 \\ \frac{1}{2} & 0 & \frac{1}{3} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{3} & 0 \end{bmatrix}$$

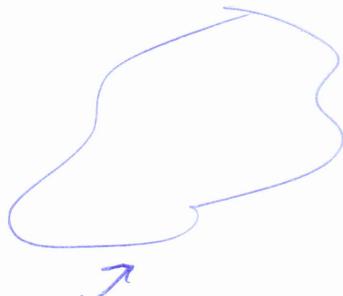


AN

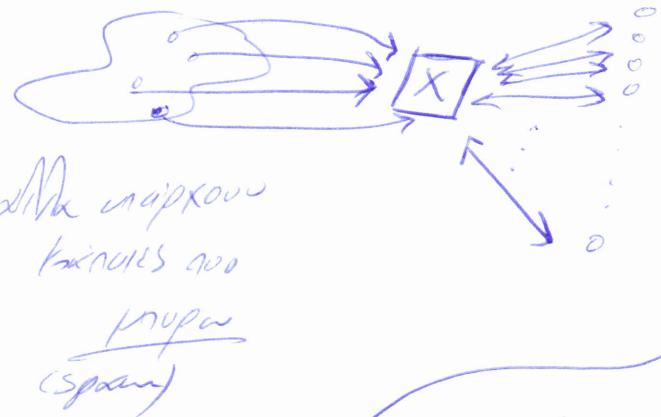
16/10

3.

Web



unapokou eendades ono
de proposas va algoritmu
te nra etipwo
so



ala va patas o pattern,
algoritmu o teleport vector (\vec{t})
kai ala va be orduces
razos, la bala m probar
probabilities, ono trusted websites...

graph mining, graph process
and algorithms

~~Span~~

Trusted Page Rank's Formula

$$\vec{R} = \alpha * M + \vec{R}' + (1 - \alpha) * \vec{t}'$$

Personalized Page Rank

bafite ono teleport vector ono eendades ono
onosis has endapejoo.

→ elan eguruva jive jn def. uayopis.

~~οὐδὲ πληρῶς~~

~~παντὸν διεμένεσσιν~~

48

Αγρόκος Ρωγίταινος δοκίμεια ~~πε~~ άλλο αντικ.
της εργασίας οντο γενναγ.

33

Αγρόκος Ηλε.

Τα τελεία αστάθη π. η έρευνα στο Αγρόκος
Λ(Π) : λαβ. εξαρτώμενη στη
γενναγ.

Ζ(Π) : authority. στην γενναγή στον πατέρα της
της γενναγή στην γενναγή της μητέρας.

Επιν γενναγή = πατέρας ή Ηλείας πάτερας και πατέρας μητέρας
στην γενναγή της μητέρας.

36

πατέρας μητέρας αποφύγεια στη λαβή
της μητέρας της μητέρας.

Ηλείας λαβή στην αποφύγεια της μητέρας
της μητέρας της μητέρας της μητέρας της μητέρας
→ day 37.

Στην Αγρόκοτο στην day 38 περίπου η έρευνα
μητέρας μητέρας.

A.M. 30/10
1.

Na xparakorainos
word2vec.

Epafid

ηροδοπία εώς την
επανάσταση.

Quora challenge

εξ 3 popkes.

Q_1 "~~~~"

Q_2 "~~~~"

$$Q_1 \stackrel{?}{=} Q_2$$

Νχ Αριστούργησε είναι o Bill Gates = Q_1
Αριστούργησε είναι Bill Gates = Q_2

→ training set

Ερώτηση 1 ως Ερώτηση 2 επικαλύπτονται
Με είναι διαφορετικό! Ε-διάλυτο
δεν είναι διάλυτο $\left\{ \begin{array}{l} \text{o διαφορετικές} \\ \text{+ ιδέες.} \end{array} \right.$

Version 1

Δε να πω στο training set, ταν Δε
μπορεύει να δικτύωσε στο is-duplicate.

Ta version είναι
ανεπόφελη!

Version 3 → ερχεται εργασία,
και ορισμένες σειρές εργασίες
πως να είναι στο γράφη.

Version 2 → έχει "training" des έξα λήστελ.
και αργά καθώς αναζητά την επόμενη
εργασία.

Version 2 έχει training + λήστελ.
και ερχουνται σειρές που γράψει και με
την επόμενη λήστελ

→ ~~Ουράνιος Δέκα!!!~~
Java

To 2 ή το 3 προσπος σε αναδιαρροή.

AR 30/10

2

- New document representation
similarities.

(Latent Semantic Analysis)
- LSA -

→ gives different result.

Lixos

- overcomes percolation.
- finds hidden document structures.

~~#~~ SVD (Singular Value Decomposition)

7

Xpos review \rightarrow [LSA] \rightarrow semantics
concepts

1. Review concepts
2. Review as documents.

Step 1 \rightarrow Review $A_{(n \times m)}$ review document - terms.

Review $A_{(n \times r)}$ each row is a review

$U_{(n \times n)}$ unstructured on basis of concept or word
 $(V_{m \times r})^T$ document

↳ terms, concepts.

Pure matrix

Matrix point - distinguishable patterns (patterns).

$[A] \rightarrow [SUD] \rightarrow$ Ja spu' co r, van
Mödös megfogásra szorított
= töbök megfogásra
concepts / felkérő okoz
o nivalos [A]

9. \int nov marea A, exocle & piso carls, noz eliaz
xvi/sep/2011 prof. 2001.

apd o 1 Jk area 2×2 , turkisuu o/
2fes 2001, an chayly noz vach
concept. (spa edur ox obs, & amiusu se
& rangers!)

$\times A_1 = 2.64$ \rightarrow 1^o concept area noz
 $A_2 = 5.25$ before, prof. mirek!
Exocle requeen yox does
go 2^o concept.

av eks 3 concept \times .

8.64

- 29

0.5 → evan dafawro! zo diroxw!! evan
nwyd ym mynyddoedd evan nwyd.

SVD

A07 - 30/10

3.

$$A = U \Lambda V^T$$

→ Dáku o bpu za concepts nōo exow
za nearkzou * "wippe" zon nivana.

0 nivana) → document to concept similarity

U: nōo Adu awi budi, za
nōo bpu buppoxur zon uade dc, oza
concepts.

$$A \rightarrow \begin{bmatrix} 0.10 \\ 0.18 \\ 0.36 \\ \vdots \\ \vdots \end{bmatrix} \rightarrow \begin{array}{l} 0.10 \text{ bpu os } 1^{\text{st}} \text{ concept} \\ 0.18 \text{ bpu os } 2^{\text{nd}} \text{ concept} \\ 0.36 \text{ bpu os } 3^{\text{rd}} \text{ concept} \\ \vdots \vdots \vdots \end{array}$$

Pl zon u bpu za bpu, za dc,
na uade, nōo exar za ophameng &
za concept.

0 nivana) V: term to concept similarity.

dañadi propi o bpu za terms nōo exar
ophameng & na uade concept.

$$\begin{bmatrix} \rightarrow \text{term1} & \rightarrow \text{term2} \\ 0.58 & 0.58 \\ 0 & 0 \end{bmatrix} \rightarrow \begin{array}{l} 0.58 \text{ bpu os } 1^{\text{st}} \text{ concept} \\ 0.58 \text{ bpu os } 2^{\text{nd}} \text{ concept} \\ 0 \text{ bpu os } 3^{\text{rd}} \text{ concept} \end{array}$$

dañadi nōo bpu os 1^{\text{st}} concept
za concept 2.

dañadi nōo bpu os 2^{\text{nd}} concept
za concept 3.
dañadi nōo bpu os 3^{\text{rd}} concept
za concept 1.

Dimension reduction.

Now we have many to concept no \mathbb{R}^n
dim.

if we no our names UV^T close to \mathbb{E}^m
accept. \rightarrow A give ones given on
day 18.
Therefore the concept, known as matrix if we no a
few more 1.

See previous pg we no open discuss, after
therefore the concept.

$$\text{dim } \vec{D} = \left\lceil \frac{\text{dimensions}}{m \gg 5} \right\rceil$$

$$\text{We have } \vec{D}' = \left\lceil \frac{5 \text{ dimensions}}{5} \right\rceil$$

\rightarrow 0. entries no more U , also no redundancy no
A.A.T.

A.A.T \rightarrow entries no U

A.T.A \rightarrow redundancy
 \rightarrow splices no U

ΑΩ

30/10

η

◦ A.T? Ήταν απόφοιτος, κατ' επαγγέλματος.

τελευταίας σενάριος του πάντα, ήταν
είχε γίνει πράξη τους.

το ίδιο ήταν και για τις γραφές τους ✓.

εκπομπές από τη διασφάλιση που έγινε στην Ελλάδα,
(την παραπάνω) για να περιγγέψει τη διασφάλιση
της εποχής της διατάξεως, οε ~~της~~ ~~της~~ διατάξεως
και της Α.Τ. στην οποίαν περιγράφεται
τη διαδικασία της A.T.

Για να ταυτοποιήσει την μέρκη.

$$A_{\text{K}} = A_{\text{την}} \times V_{\text{K}}$$

↳ περιγράφει τη σύσταση
και την σύσταση της μέρκης

εδώ παραχθεί

πρώτη διατάξη,
επιπλέον πλήρης
επικίνηση,



◦ Η μέρκη ήταν, είχε λαρυγγίτιση
και το πλάστη ήταν η.

$\sin(\vec{d}, \vec{d}_a)$ nu va fiu unu
 orolozare (nu este născută
 sau de la de acu A)

De fapt nu orolozat $\sin(\vec{d}_i, \vec{d}_j)$ sau
 Anivax.

$$\text{deoarece } \sin(\vec{d}_i, \vec{d}_a) \neq \sin(\vec{d}_i, \vec{d}_i')$$

Aceeași
 de fapt
 nu
 este
 orolozat
 Anivax.

Nici
 nu
 se
 supadă.

Anivax
 să se exprime în
 $Q_{av} = [0, 1, 0, 0, 1]$
 $n \times k$
 $(k=2)$

să se exprime în
 $Q' = Q \cdot V_k$
 $1 \times k$
 $1 \times n$ unde

să se exprime în
 Q' de către
 Anivax.

$$Q' = Q \cdot V_k$$

este
un
vector

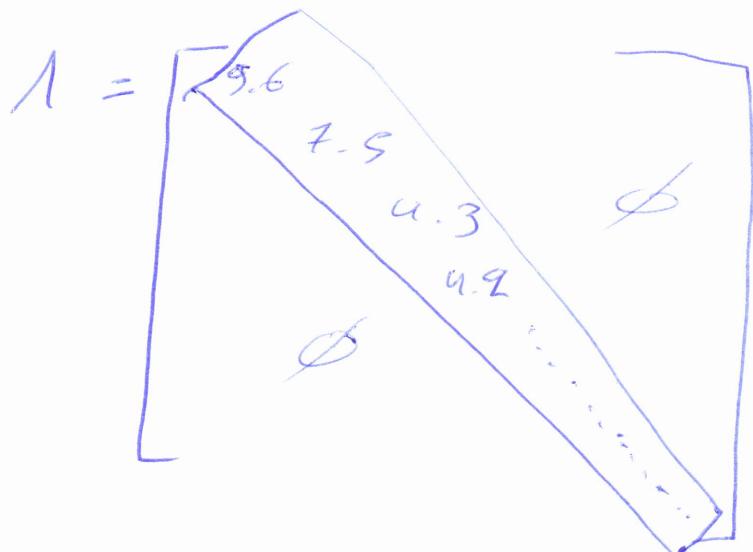
să se exprime în
 Q' de "prefer" de la Anivax.

→ supadă
 de la

A07 30/10

5.

~~the variance of a vector~~
Aus der Menge von n Vektoren der
Größe:



L'grosos

der $b = 3$ ist
Die Spalten von A
sind nicht linear
nicht linear sind
beide 0.

aus der Form U_3, V_3^T

~~Properties~~

$$S = \sum_{i=1}^n c_i^2 \quad \text{wurde } c_i \text{ zu den ersten } n \text{ Dimensionen von } A.$$

B' Zeros

Summe aller Vektoren, die sich in
Beziehung zu 80% von S .



93

προσδιόριση της A^T , χρησιμοποίηση της A^T ,

τα δύο $A_U = U \cdot U^T \cdot A$

πόσο λεπτή A^T είναι

$$\text{πόσο λεπτή } A^T = (U \cdot 1 \cdot V^T)^T$$

~~τα δύο λεπτά~~

~~κατατελέσματα.~~

~~επιλογή~~

~~παρατελέσματα~~

Pathologies - cavities.

45

SVD \rightarrow 3 concepts

ήρωες της βαθείας ο U ο U^T (όποια είναι -
 concept) τα οποία είναι κατατελέσματα
 οι οποία είναι σταθερά στην έναστρη γεωμετρία της
 περιουσίας της V . Η V
 μετατρέπεται σε πολυδιάστατης
 μετατρέπεται σε πολυδιάστατης

βαθεία νέων A ,

~~παρατελέσματα~~

παρατελέσματα

παρατελέσματα V είναι

παρατελέσματα U είναι

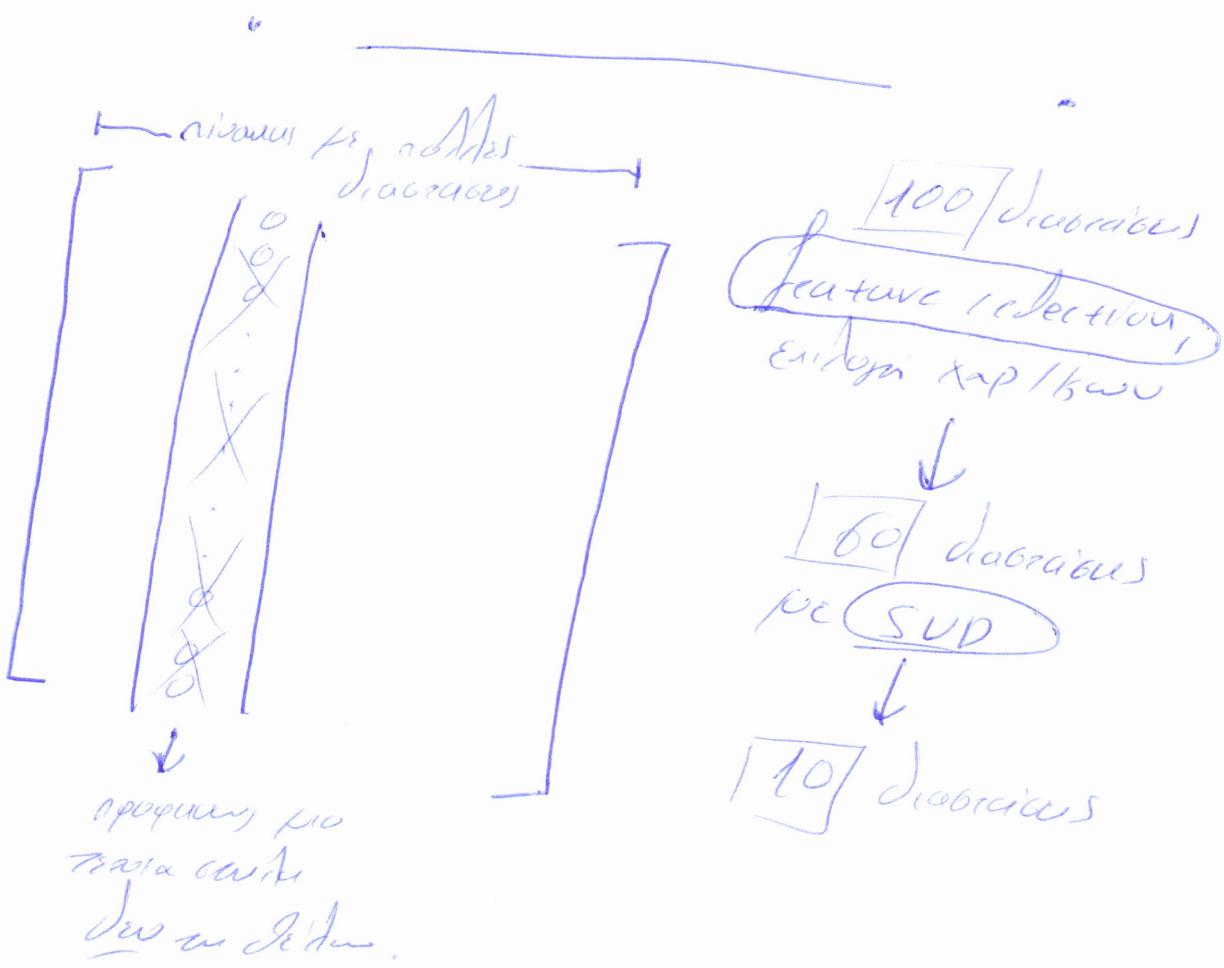
AR 30/16

6.

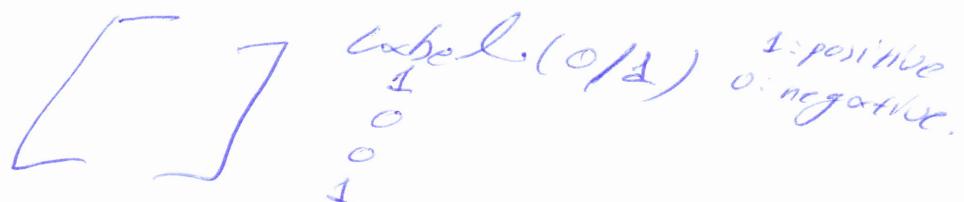
Ew uanouici na 2000 dig de daxx
trupaxi bussexa, eksos!

50

O spucas posjoox jeku salfisxan
tawies idha edhou!



Jew u daxx minn ja ixtar ekkien label.



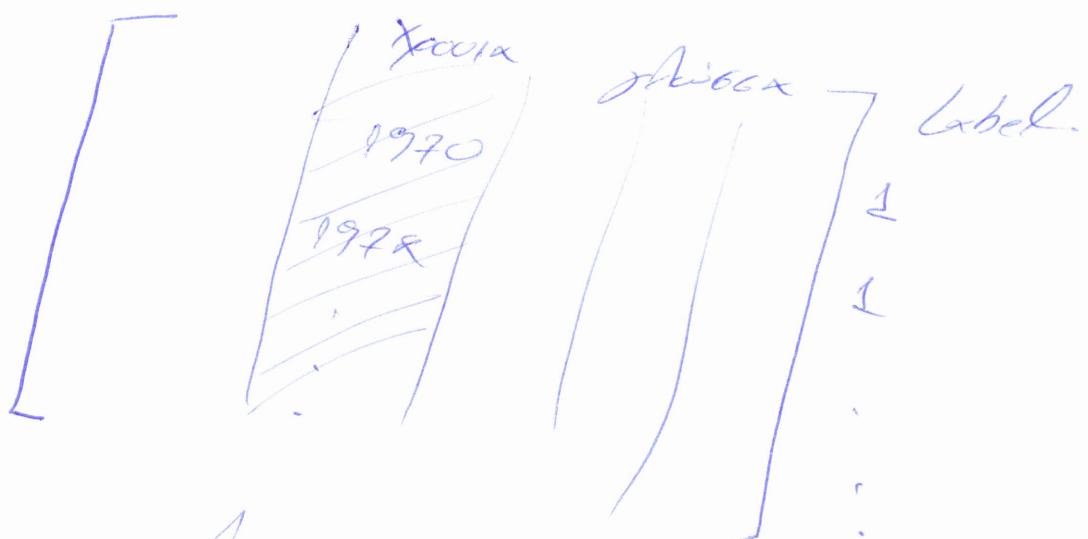
Probabilistic classification, ja ipha ewi minn ja ixtar u
probabilitati u daxx positive: 1 n' negative: 0.

Napadyspe.

Een van die soorte wat ons sante deurmerk op
X00018.

Datums

(Een van die soortes wat ons
soortgelyke data
as attribuete)



Een van die soorte wat ons sante
van datums van bykomste na X00018
deurmerk = 1.
Die datums wat ons in die klasifikasie
van die enkele items, en X00018 deurmerk
van die enkele items wat daarby
in klasifikasie van datums
gekom het as souks van datums
na datums!

Groot Indet. van hier moet ons sante datums
as die items (o/s)

AN 30/10

7. Ans da viva za avançada fcc da
dioxof. Repto; nt exa ~~da~~ da rotula e
pe k nodes



kkk nisaber, nao fui
dixaia os perfeitos
achacions.

$$A = k \begin{bmatrix} & k \\ & j \end{bmatrix}$$



~~de xipos~~
pe avançada.
nao fui os achados
tous $f(x, b)$

isso on da unica da qd.

achacabem.

pe MDS

↳ propo se fazendo da viva os
nodes ee afas-g-x.
n os des fracionar deles.

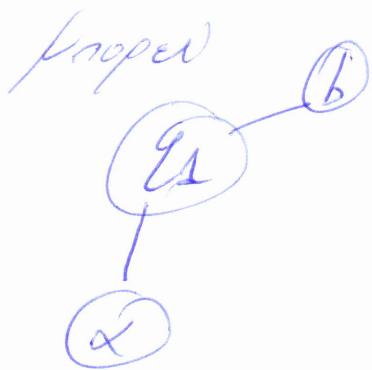
→ Da proposta nt estou $A \rightarrow$ va na qd de MDS ee
100 documentos - kmpf - SVD da mewa em 80.

Exemplo PCA (EOPPETH DESONVEN-UN)



Είναι τώρα η ώρα για την πράξη.

Μπορεί να κάνω μακροποίους δατάκια



[Spanning Tree]
Excel
Python

Εγγυώμενης σταθερότητας

Είναι επιδιαλέξια

βάσει των διαφορών

διαφορετικής προσέγγισης

επιδιαλέξιας κ.ο.κ.

data
Set
map
επιδιαλέξιας
ποσού
επιδιαλέξιας

Ao

Boolean Muroedo

O problema é encontrar se ova figura.

m - Subáreas

m - Núdos fronteiriços.

Ex 10

Karacterísticas Hawking

Ex 11-12.

→ Confundir píxias I.

! De pe adiçoes de sobre e opes
Extrair fix Des.

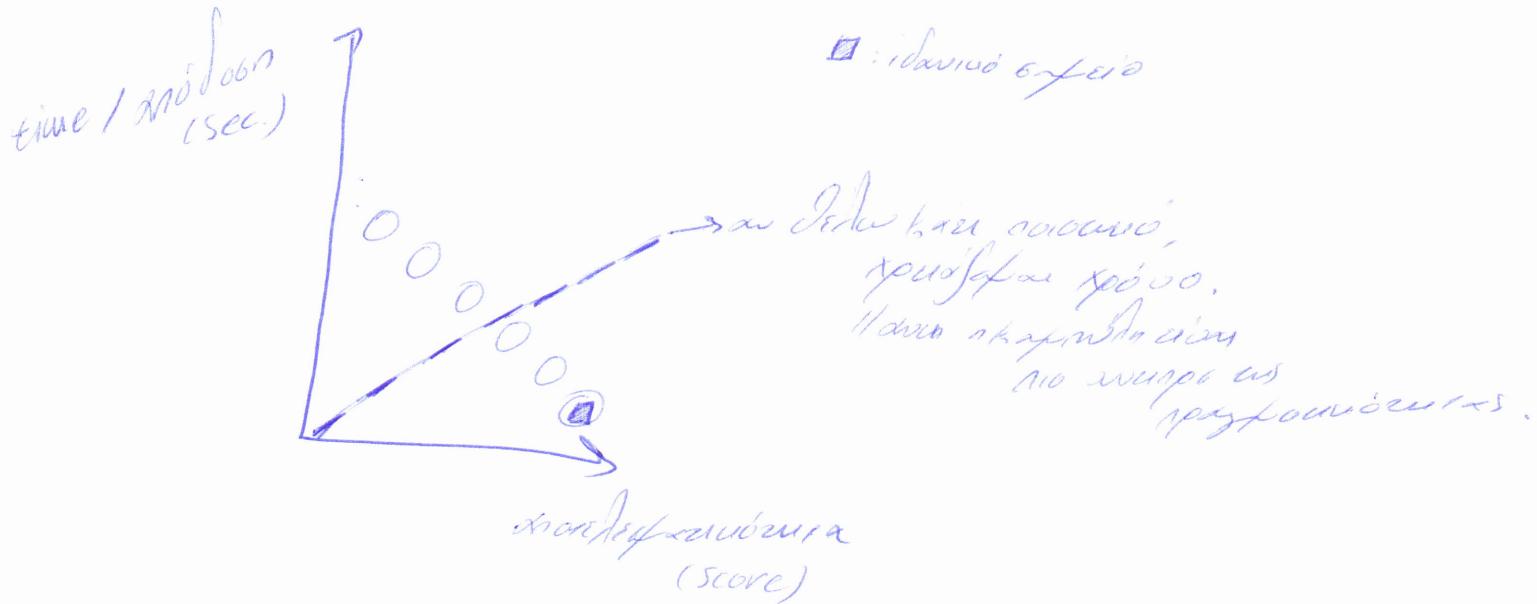
$$\cos(\theta) = \frac{\vec{x}_i \cdot \vec{x}_e}{\|\vec{x}_i\| \|\vec{x}_e\|}$$

621 30

Protognathus opaculus

AM

Anzifion Anreleffzunwurz.



! Los DBMS por varia poca n anodan

Recall: Atipouria monelefzunwurz
Precision: badapouria monelefzunwurz.

$$(R(j) = \frac{1}{n} \sum_{i=1}^n P(i))$$

Apporius pocos

o mas 1

DBMS

To F score etan

ndo baxo jenua.

$b=1$ idia bigore.
 max recall kai ord
 precision.
 $P \geq 1 \rightarrow$ recall
 $P < 1 \rightarrow$ precision.

E-Measure
 opifoxe an
 entaherofoxe
 neptibebonjepo ja
 recall n 10 precision

(Suprak x oit 3)

o Mayor efectivo a salvo e
baxo.

~~100~~ 70°
Raffia → Products

↓
Proportion
20% Raffia

Boolean

$$tf_{xj} = \frac{f_{xj}}{\max_y \{ f_{yj} \}}$$

f_{yj} : max words repeated to
of in f_x .

$$nif_x = \frac{idf_x}{\max_y \{ idf_y \}} \text{ or } idf_x = \log \frac{N}{n_x}$$

N: apodos eggapoou

n_x : # eggapoou no repeatou
no f_x .

$$\boxed{w_{xj} = tf_{xj} \times nif_x}$$

Anayisis And Refinement max.

$$q \rightarrow [\Sigma A] \rightarrow A$$

Basics

$$\text{overall} = \frac{A}{R}$$

C: widths eggapoou

R: okulta pto no q

A: anix noo enogapou no EAT

Ar: ta okulta uno anix'noo
enogapou no EAT.

$$\text{precision} = \frac{Ar}{A}$$

$$\text{fallout} = \frac{(A - Ar)}{(C - R)}$$

Met overzicht:

precision(q): pto q'te precision na oks w day este

R-approxx: tlu okribus) dan Rabute R eggapoax.

Miss probability

$$\bullet F_{(i)} = \frac{2}{\frac{1}{r(i)} + \frac{1}{p(i)}}$$

own documents to
i.e. 0.670 0.770

$F_{(i)} = 0$ for the documents having
zero doc.

$= 1$ for the documents having
one doc.

$$\bullet t_{(i)} = 1 - \frac{\frac{1}{w}}{\frac{p(i)}{r(i)} + \frac{1-w}{r(i)}}$$

w: document recall
precision.

Document retrieval

[A] query has one Boolean

$\vec{d} = (w_1, d, w_2, d, \dots, w_n)$
queries another to q .

$$\bullet f_{t,d} = \frac{f_{t,d}}{\max\{f_{t,d}\}}$$

$$\left\{ \begin{array}{l} w_{t,d} = n_{f_t,d} * idf_t \\ n_{f_t,d} = \#f_{t,d} * idf_t \end{array} \right.$$

$$\bullet idf_t = \ln\left(\frac{N}{n_t}\right)$$

Cosine similarity

$$\cos(q, d) = \frac{\vec{q} \cdot \vec{d}}{\|\vec{q}\| \cdot \|\vec{d}\|} = \frac{\sum_{i=1}^m w_{t,q} * w_{t,d}}{\sqrt{\sum_{i=1}^m w_{t,q}^2} \cdot \sqrt{\sum_{i=1}^m w_{t,d}^2}}$$

[B]

$$\bullet t_{f_{t,d}} = 1 + \ln(f_{t,d})$$

$$\bullet idf_t = \ln\left(1 + \frac{N}{n_t}\right) \quad \left\{ \begin{array}{l} w_{t,d} = t_{f_{t,d}} * idf_t \\ n_{f_t,d} = t_{f_{t,d}} + idf_t \end{array} \right.$$

$$S_{tf} = \frac{1}{L_d L_d} (w_{1,d} \cdot w_{1,q} + w_{2,d} \cdot w_{2,q} + \dots)$$

$L_d = 1$ for 1K
 $L_d > 1$ and f_c
200 A.R. @ w
produces
documents

Avgoraphyros kardigos

Τις οι ΑΚ

$$t_1 \rightarrow [d_1, d_2, d_3]$$

$$t_1 \rightarrow [(d_1, 2), (d_2, 3), (d_3, 1)]$$

$$t_1 \rightarrow [2 : (d_1, 1), (d_2, 3)]$$

Nipos η Τηλαρ

$$\boxed{V = K \cdot n^B} \text{ με } K, B \text{ Κωνσταντίες}$$

ΟΛΒΟΣ (B = 0.4 ή 0.6)
 $10 \leq K \leq 100$

n. 6000 λιμνούς έχουν ιδιαίτερα.

Απόφαση

επιλεγμένη

Αποφάση πε απεριττού

διάτονος ή κατάργησης γεγονότων

πα $\varrho = E_1 = \text{καταργώντας}, t_E = \text{χαθεώντας}$

1) πα $E_1 = \text{καταργώντας}$

→ Διαφορική συνομικής απεριττωσης

πα τα γεγονότα που αποτελούνται από
"καταργήσιμα"

→ Συντομεύσεις γεγονότων που Ε

πα $t_E = \text{χαθεώντας}$

→ Διαφορική απεριττωσης από γεγονότα

→ Συντομεύσεις γεγονότων που είναι χαθεώντας

2)

Διάρρηση Ε πε Λδ

και εποργία με τη περιτύπωση!

Akragas AK

1) Anubopon aus Meine Memory.

So gutes anagnoson zu antiken { in → ~~Brutto~~ Lexicon

!!! andenkst gern viele Figuren
antiken → mehr Agathen.

2) Anubopon für Tafelvorpri

fix anagnoson zur Antike } Brillen

(doc 10, 10f.) zu Ende doc,
bei m. Bildern anfangen.

Ta opseis arme, ta tafelvorpri us spes et spes, bei
erceptofix bezeichnen die spes an opseis, ear prie
an opseis festigte raus.

3) Anubopon für Bezeichnungen.

fix anagnoson zur Antike

→ Bezeichnungen der Ak am besten, dann gefüllt n. farben
und Räume an den Stoffen, Erwähnung der Figuren und Personen nachher

→ Bezeichnungen aus Ak nur über den Stoffen

!!! nicht andenkst präziser na Figuren antiken

Luvupien/Eufupien ~~AIC~~

1) Audeption katalogos

(On na D noo oni puroptōnē
prabidē)

2) Eufupien pr̄w oixiwsenou

→ ja vax doc, dyploptera tempaki on 4.4,
kai perē apo karoto opo lition des vixka duktikou
oyxwai ope tous iwo AIC.

3) diadikti Eufupien

→ o katalogos eufupienou eradiata p̄ia eidi opo,
onw onw viva epiwto.

✗ ondigevar data on rati, kai perē apo karoto opo
eufupienou n̄ tixa on AIC.

baini p̄idēs on
diuxupienou eufupienou

dunitebu

• aschardia dayopien

Min $t_1 \rightarrow \{7, 10, 20, 45, 100, 100, 130, 135\}$

Max $t_1 \rightarrow \{7, 10, 20, 45, 55, 80, 10, 5\}$

• unary code

$uc(x) : x-1$ diobes, 0

✗ $uc(5) = 11110$

Elias - 2

$$x = 7$$

$$1 + \log_2 x = 3 \left\{ \begin{array}{l} u_c(3) = 110 \\ x - 2^{\lfloor \log_2 x \rfloor} = 3 \\ \text{binary}(3) = 11 \end{array} \right\} 11011$$

Elias - 5

$$x = 7$$

$$\text{elias}(7) = 11011 : \text{Syntax}$$

$$\text{binary}(5) = 101$$

$$x - 2^{\lfloor \log_2 x \rfloor} = 3$$

$$\text{binary}(3) = 11$$

$$\left. \begin{array}{l} \\ \\ \end{array} \right\} 10111$$

Golomb

$$\text{Golomb}(7, 3)$$

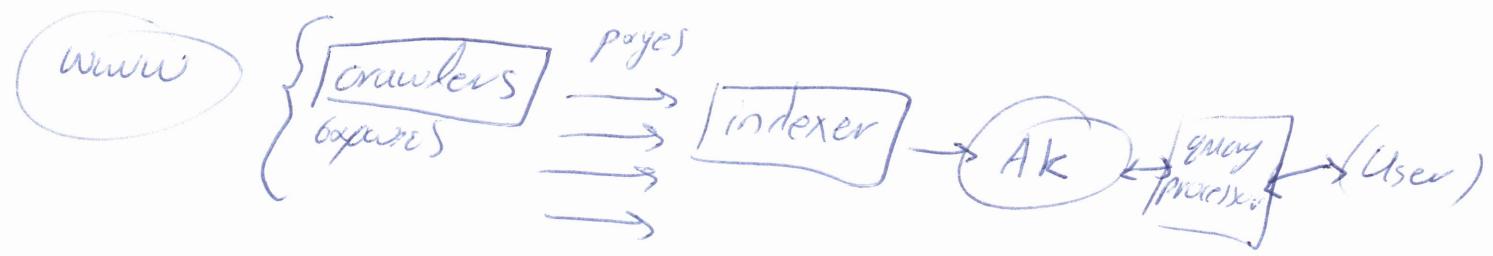
$$\text{dmpm} \quad q = \frac{7-1}{3} = 2$$

$$u_c(2+1) = 110$$

$$x - 1 - 2^b = 0 \quad \left. \begin{array}{l} \\ \\ \end{array} \right\} 110$$

$\text{binary}(0) = 0$

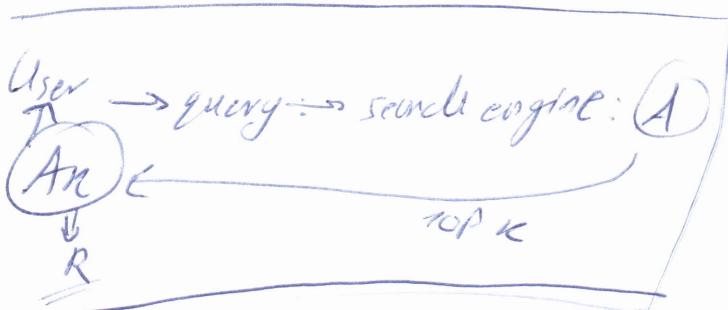
World Wide Web.



Hits Algorithm.

R: World Wide Web, consists per se of a lot of web pages or User.

S: words are often called web documents, in distinction are the R.



→ Equations for Hits are S

$$\text{know init pr. of a page } \begin{cases} \alpha_S(p) = 1 \\ \alpha_U(p) = 1 \end{cases}$$

$$\text{but not only page: } \alpha_S(p) = \sum_{q:p \rightarrow q} \alpha(q)$$

$$\alpha_S(p) = \sum_{q:p \rightarrow q} \alpha(q)$$

• pagerank Algorithmen

Quellen transition Matrix : M

1) ~~vorherige~~ $n = 5$: $R = \begin{bmatrix} \frac{1}{5} \\ \frac{1}{5} \\ \frac{1}{5} \\ \frac{1}{5} \\ \frac{1}{5} \end{bmatrix}$

Frage nach $\vec{R} = M \vec{R}$ erwartete Fixpt. vor neuem Algorithmus

2) Fixpt. von M als neuer Start

Aufgabe 2: ~~Konvergenz~~

$$\vec{R} = \alpha * M + \vec{R} + (1-\alpha) \vec{T}$$

$(1-\alpha)$: Gewichtung der eigentl. Neigung des anderen vert.

LSA

• SVD $A_{dxr} = U_{dxr} \Lambda_{dxr} V^T_{exc}$

• Dimension Reduction

$$A_r = A_{dxr} * V_{exc}$$

• Reduziert auf weniger nach
Dekor zu konservieren

• Feature Analysis

$$\rightarrow Q' = Q * V_{rc}$$

\Rightarrow Vergleich zu Q' mit A_r !

6/11

1. AP.

Mr. Sri Dharma
2 Lecture.

Def 1

New Stream (continuous stream, query)
continuous Xstream. (continuous query stream)

organizes ~~as~~ index indexes or profile over Xstream,

Kan beide manieren gebruiken om query
SQL profile van Xstream.

Def 2

Error

Def 3

A
ox1

B
LSI methodology

topics

new document set

new term dimensionality reduction
k-dimensional vector.

Def 4

A) HITS, PageRank.

probabilities, Gob. Diagonal

B) Transition Matrix x Rank Vector (2 types)

Expt 5

1) ~~Thermo~~
 byxwunun (de rapport om van
 organis)

Modellen α , β , γ om van rapporten van
 de organen.

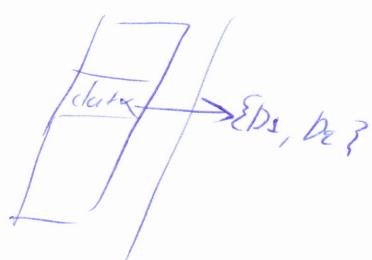
$$\begin{aligned} & EDS, D2\} \\ & \{D3, D4\} \end{aligned}$$

3)
OXI

Expt 1

Inverted Index.

"Index Source"



$$D_L = x_1$$

$$D_E = x_2$$

:

:

haver n. query, classificat jaa va beperke soort lezygyniss
drievet
(een predefined lezygynie)

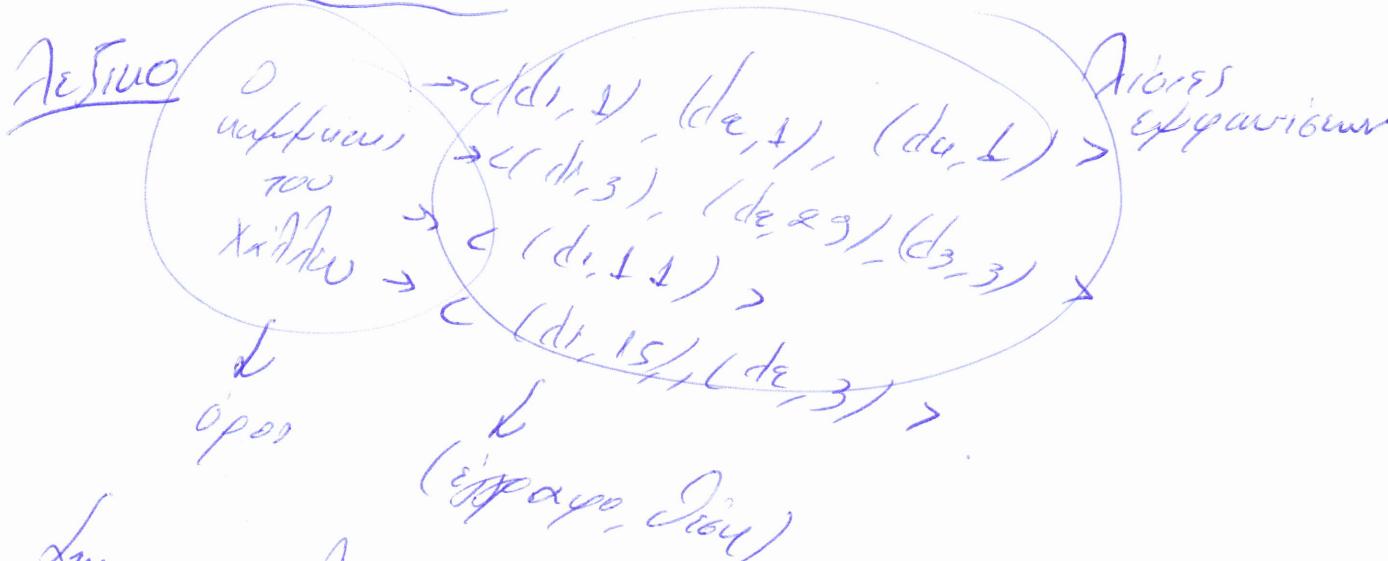
διανομή ΕΑΠ με ΕΔΒΔ

(1)

ΕΑΠ → δίκαιοι { effectiveness πολιτικήματας
efficiency. ταξινομία

! Διο ΕΔΒΔ διαφέρει από πολιτικήματα

Χρηματοδότηση λαϊκών



→ Διανομή από την απεριφούσα Α.Δ. απεριφούσα στα διάλογα
επράγματα που απέδιδε στην Α.Δ. απεριφούσα στα διάλογα.
Επράγματα που απέδιδε στην Α.Δ. απεριφούσα στα διάλογα.
Επράγματα που απέδιδε στην Α.Δ. απεριφούσα στα διάλογα.

Άνοικη Αστικός { B-διάρειο

Πίνακες πλημμεροποίησης

στα πλαίσια σχεδιασμών

- Ο χρόνιος επιδρεπον ή νο σχεδιασμός επράγματα.
- το ασφαλεία επαναρρόφηση χιλιάδων νο σχεδιασμών.

τεραπονίας ανθρώπων } Τεραπονίας ανθρώπων

(2) Απαριθμός
συστημάτων
(αντίστοιχης
επιφύλαξης)

πχ Q = Επικίνδυνη θέση?

~~Routine~~

Χρόνια

- Σήμα να υπάρχει αντεγκαθίδενσις στην επιφύλαξη
- πχ βράχος ή καταρράκτης οχημάτων
- πά ταχύτητα συγκριδόμενο Χρόνιο

και
ο χρόνιος
να μην
χρειάζεται
καταρράκτης
το επιπλέον

Τα υπάρχουν για προηγούμενη θέση.

πχ για επικίνδυνη θέση
πχ για προηγούμενη θέση σε οχημάτων

• Προκαρπάτας Επίπεδη ή άνοιξη.

Cross-language IR

Text Mining
Question Answering Systems.
Recommendation Systems

Boole'sche Mengen

(2)

Beispiel für X ist AND, OR, OR
 → mögliche Ergebnisse

00 : Ergebnis

An : Gleiches Ergebnis annehmen.

Wij: o goed te zien unterschieden $\begin{cases} 1 \text{ oder} \\ 0 \text{ dagegen.} \end{cases}$

0 → gleich

1 → $\{0,1\}$ → Differenz hervor rufen kann Ergebnisse!
 (differenz, aufzeigen)
 unterschiedl. no. Ergebnisse annehmen \rightarrow 0001

\vee $0 = (0 \text{ OR } 1) \text{ AND } t_3$

$$\text{dpx } Q_{df} = (0, 1, 0, 1) \vee (0, 1, 1, 0) \vee (0, 1, 0, 0)$$

Gas

$$Q_{df} = (0, 1, 1) \quad \downarrow \quad (1, 0, 1) \quad \downarrow \quad (1, 1, 1)$$

Gas

Gas

$$\text{dpx } Q_{df} = (1, 0, 1, 1, 0, 1, 1, 1, 1)$$

alle Ergebnisse \rightarrow da eine nur annehmen
 zu manchen ist es zu Q_{df}

Tekstes Basilean Model.

- Διαφορετικά και
περισσότερα από
τρία φαντά.
- ADJ. παρέχεται όποιας
ήλικας ή ποσού.
 - NEAR/X ως μέτων της ποσού ή γεωμέτριας.
Ηλικίας NEAR/E ποσού
 - WITH ως βοηθείας σεν της γεωμέτριας
 - SAME ως βοηθείας σεν της παραγγέλματος

! Η επιλογή εργαλείων να δικαιούεται σε
τις περιπτώσεις εγγύησης ή προστασίας

Mouvementa.

- ΟΧΙ βαθύτερας
- Δραστικός ουν διανυσματικός εργαλείων.
- Τα κεριάκια σωστά προστατεύουν την ιδιότητα της οικίας από την οικογένειαν.
- Προστατεύεται την ηλικίαν της πανεύκολως.

Eukafiko Logiko Model

extended Basilean Model.

Το καθέ επόμενο μακριότερα περιβάλλεται
διανυσματικής αρχής.

- W_J ^{ραστός} περιβάλλεται από την J.
- Η J πανεύκολως ανανεώνεται στην περιβάλλεται από την J.

$$\eta_{f,g} = \frac{f_{x,j}}{\max_j \{ f_{y,i} \}}$$

η είναι η
καυτομορφισμός
στη διεύθυνση [0,1]

Ο χαρακτής εγκαίεννος στην η
καυτομορφισμός είναι η f_j

$$idff_x = \log \frac{N}{n_x} \rightarrow \text{αναλογία αριθμ. ομμάδων}$$

↔ αριθμός ομμάδων
στην περιοχή της σημείου x

$$nidff_x = idff_x \cdot \frac{N}{\max_j \{ idff_y \}}$$

καυτομορφισμός f_x .

Έτεινος

$$\boxed{w_{f,g} = \eta_{f,g} \cdot nidff_x}$$

Σημ. 3.5. Η
επιλογή των
διεύθυνσηών

Υπολογισμός

$$t_d = \text{καθημερίνες}$$

$$t_b = \text{χαρτιά}$$

$$g_{ad} = \text{καθημερίνες} \wedge \text{χαρτιά}$$

$$e_{av} = \text{καθημερίνες} \vee \text{χαρτιά}$$

Ελλείψης γλωσσικής επιδρούσης:

$$d : f_{d,s} = 1 \text{ και } \eta_{f,d} = 1$$

οποιας λογισμού → η επιδρούση
τα $N=7$

$$\text{σημ. } idff_x = 0.243$$

$$\max \quad \eta = 0.845$$

$$\text{for } n, \eta_{\text{ref}} = 0.243 / 0.845 = 0.288$$

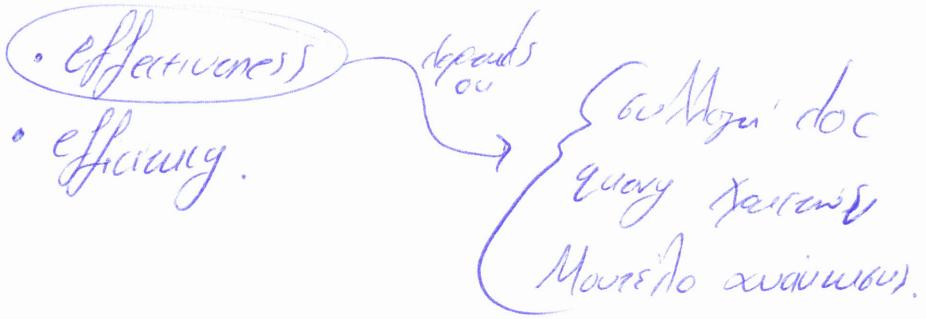
$$\text{so } W_{\alpha,1} = 1 - 0.288 = 0.712$$

Shows η_{α} is constant across α .

P-norms?

• Accuracy vs. Recall/Fallout

(4)



Q → EAN → A

R: documents w/ query Q.

Ar: documents w/ query Q, mentioned
on A.

$$(Ar = A \cap R)$$

• recall:

$$\text{recall}(q) = \frac{Ar}{R} \quad \text{correct extraction} = 100\%.$$

• precision

$$\text{precision}(q) = \frac{Ar}{A} \quad \text{on average} = 100\%.$$

• Fallout

$$\text{fallout}(q) = \frac{|A - Ar|}{|C - R|} \rightarrow \begin{aligned} &\text{excludes from query results} \\ &\text{not mentioned} \end{aligned}$$

\rightarrow excludes from results
the wrong.

Aproximación 1

Entonces la exactitud es de 75%

$$\left| R \right| = 8 \\ \approx |A| = 12 \\ |A| = 6 \quad \left. \begin{array}{l} \text{recall}(q) = \frac{6}{8} = 75\% \\ \text{precision} = \frac{6}{12} = \frac{6}{12} = 50\% \end{array} \right\}$$

Aproximación 2

Entonces la exactitud es de 66.66666666666667%.

Porque $A = d_1, d_2, d_3, d_4, d_5$

Aproximación 2: recall & precision

$\left. \begin{array}{l} \text{synonymy} \\ \text{polysemy} \end{array} \right\}$

Aproximación 3

de ~ de

$t: \text{signo}, \text{negativo}$.

signo $\rightarrow d_1, d_2$.

Negativo $\rightarrow d_3, d_4, d_5$.

$q = \text{signo}$

recall = 40%

$t = \text{signo}$

recall = 60%

La otra mitad es 60%

Movorres e Sofores.

precision (q): proba de que se temos, para
determinar os desejados efeitos

R-efectos.

probabilidade de que os efeitos, que
determinam R sejam presentes.

Aproximando Micos para E-efectos.

$$F_{(i)} = \frac{q}{\frac{1}{r_{(i)}} + \frac{1}{p_{(i)}}}$$

i-observação
de desejado em
máximo.

$$0 \leq F_{(i)} \leq 1$$

$p_{(i)}$, $r_{(i)}$ probabilidades de
que os efeitos sejam
i-observados.

E-efectos

W efeitos, entre os efeitos de máximos e efeitos.

$$E_{(i)} = 1 - \frac{1}{\frac{w}{p_{(i)}} + \frac{1-w}{r_{(i)}}}, \quad w \in [0, 1].$$

n

$$E_{(i)} = 1 - \frac{1}{\frac{w}{r_{(i)}} + \frac{1}{p_{(i)}}} \rightarrow w \in [0, +\infty)$$

What's your hobby?

(6)

O Sonnentanz-Muster

Köde eppxgo d' aranopixox wiß dawofx

$\vec{d}_i = (\text{weid}_i, \text{weid}_j, \dots)$ M ge stidox
 ↳ Bpos opoo owo eppxgo.

Xpas preprocessing aw eppxgo.

M = xpidox ikr aw fawadux. Afawo noo
 efxavixox os ida ra eppxgo xos
 wakzus.

Añi popoxi dawof-pawade

weid_i = {₀ idios.
 ↳ repita aw owo ti

D: nivauos owo eppxgox.

o	o	o	o	o	o	o
e ₁	weid ₁	weid ₁	.	.	.	d ₁
e ₂	weid ₂	
e ₃	weid ₃	
:	:	:	:	:	:	

Mspaffes, N sultes.
 Gópor, Gdocs.

tf-iff.

• $f_{e,d}$: apd^e effavicau e oce d.

• $n_{f,e,d} = \frac{f_{e,d}}{\max\{f_{x,d}\}}$.

• $\text{idfe} = \ln\left(\frac{v}{ne}\right)$ Andor effavien
bottamus
Cardia effavien no
represkuon eo e.

$\approx n_{f,e,d} \cdot \text{idfe} = \frac{f_{e,d}}{\max\{f_{x,d}\}} \cdot \ln\left(\frac{v}{ne}\right)$

• $n_{idfe} = \frac{\text{idfe}}{\ln(v)} = \frac{\ln(v/ne)}{\ln(v)}$

$\approx n_{f,e,d} \cdot n_{idfe} = \frac{f_{e,d}}{\max\{f_{x,d}\}} \cdot \frac{\ln(v/ne)}{\ln(v)}$

4.2.2 Korowia e fc d_j

Eugrafia e l ws dzisiejsza sas w d_j

a austreka malaou an

lansyfia e k_m d_j"efavien
na pegolla effavien $D_e(q, d) = \sqrt{\sum_{i=1}^m E_i (W_{eq} - W_{ad})^2}$

(7)

• Etwas mit Jodid:

$$\text{Sinner} = \vec{q} \cdot \vec{d} = \sum_{i=1}^n \text{Eig. W. d.}$$

"Etwas" = mehr
Objekte!



$$\cos(\theta) = \frac{\vec{q} \cdot \vec{d}}{|\vec{q}| \cdot |\vec{d}|}$$

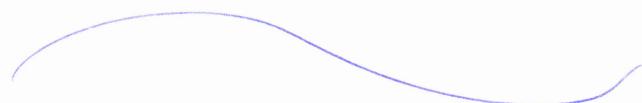
$$\text{dann } |\vec{d}| = \sqrt{\sum_{i=1}^n \text{Eig. W. d.}^2}$$

Mehr Objekte \Rightarrow regelmäßiger zusammen $\cos(\theta) \rightarrow$ regelmäßige Objekte!

$$\text{Cosine} \cdot (\vec{q}, \vec{d}) = \cos(\theta) = \frac{\vec{q} \cdot \vec{d}}{|\vec{q}| \cdot |\vec{d}|} = \sum_{i=1}^n \text{Eig. W. d.}$$

$$\sqrt{\sum_{i=1}^n \text{Eig. W. d.}^2} \cdot \sqrt{\sum_{i=1}^n \text{Eig. W. d.}^2}$$

oder jenseits von $\cos(\theta)$ { Subtrahieren aus und ergebnis
Objekt aus Objekten.



Napadzffx.

$$q = \left\{ \begin{array}{l} \epsilon_1 \\ \epsilon_2 \end{array} \right\}$$

wool
→ nivitas 9.8. }
3.

~~water~~ oporowia sie d. de d₃ row
nivitas 9.1

$$\text{idfes}_1 = ?$$

$$n_{e1} = 3$$

$$\text{opx idfes}_1 = 1.203$$

$$\text{idfes}_2 = ?$$

$$n_{e2} = 2$$

$$\text{idfes}_2 = 1.504$$

$$\boxed{\text{idfes}_1 = \ln(1 + \frac{n_e}{n_{e1}})}$$

$$\boxed{\epsilon_{fe,d} = 1 + \ln(f_{e,d})}$$

$$\epsilon_{fe,ds} = 1$$

$$\epsilon_{fe,di} = 1$$

$$\epsilon_{fe,ds2} = 1$$

$$\epsilon_{fe,di2} = 1.693$$

$$\epsilon_{fe,d3} = 1$$

opows jaz q

$$\epsilon_{fe,2} = 1$$

$$\epsilon_{fe,2} = 1$$

$$\boxed{w_{e,q} = \epsilon_{fe,q} \cdot \text{idfes}}$$

$$\boxed{w_{e,d} = \epsilon_{fe,d}}$$

$$\text{opx } w_{e,di} = 1$$

$$w_{e,q} = 1.803$$

$$w_{e,ds} = 1$$

$$w_{e,q} = 1.504$$

$$w_{e,ds} = 1$$

$$w_{e,di} = 1$$

$$w_{e,ds} = 1.693$$