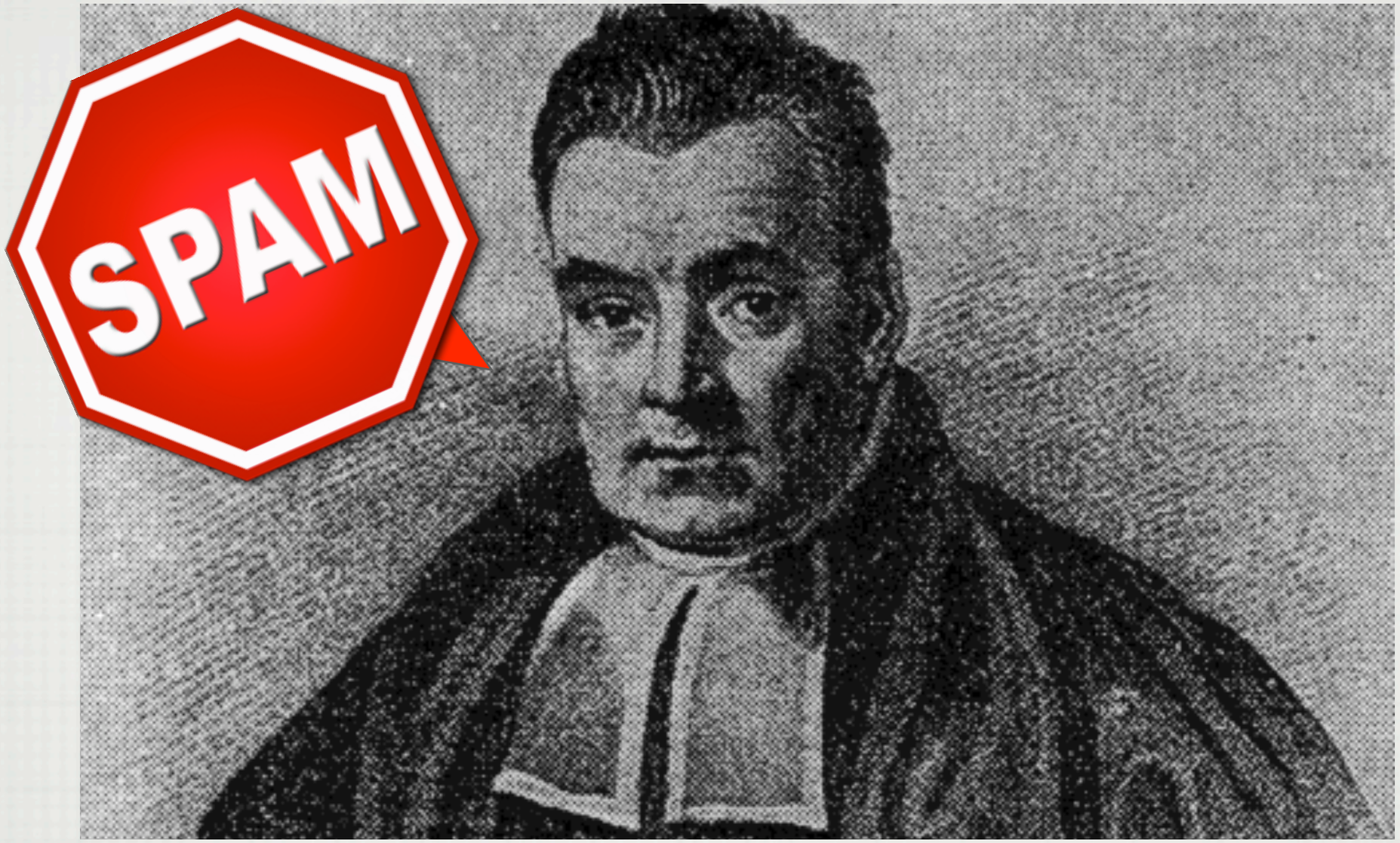
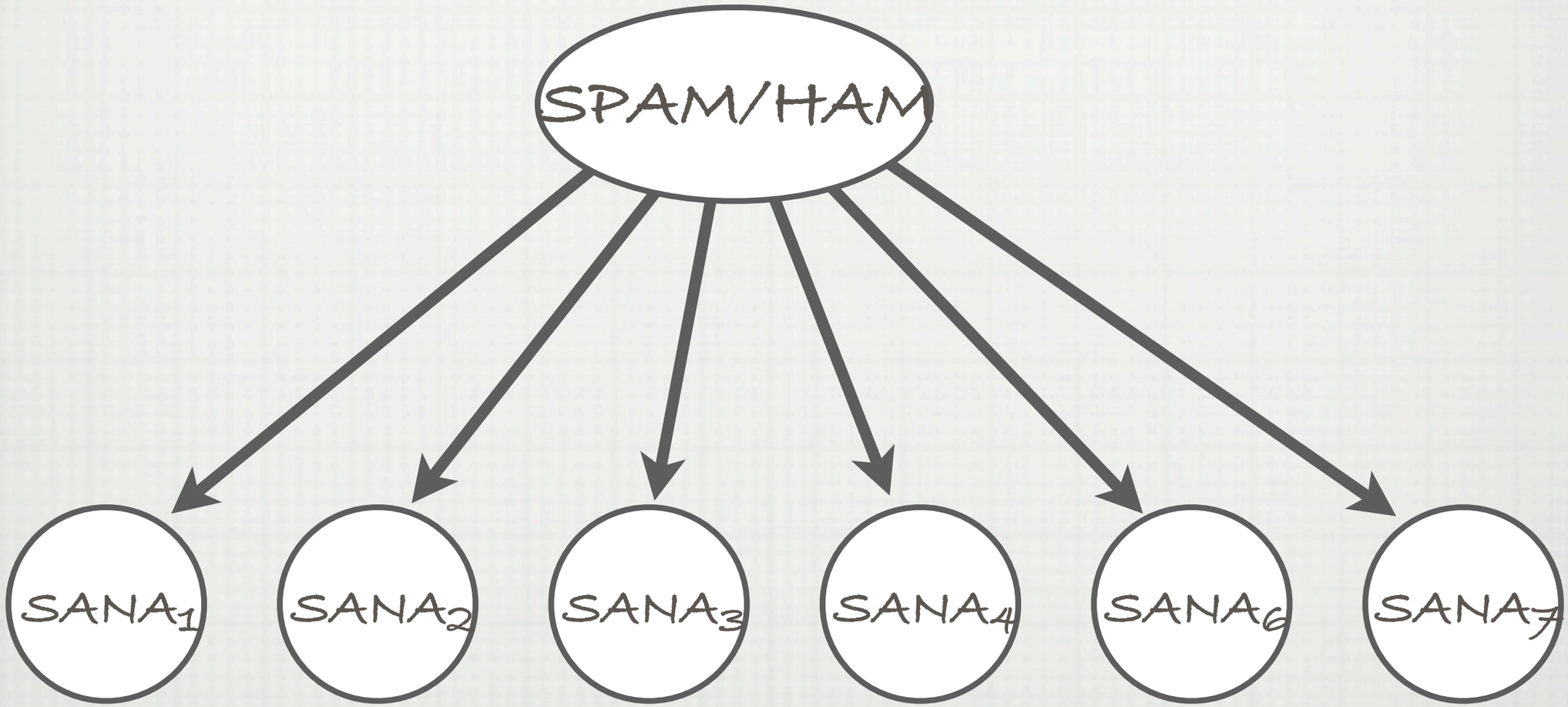


JOHDATUS TEKOÄLYYN

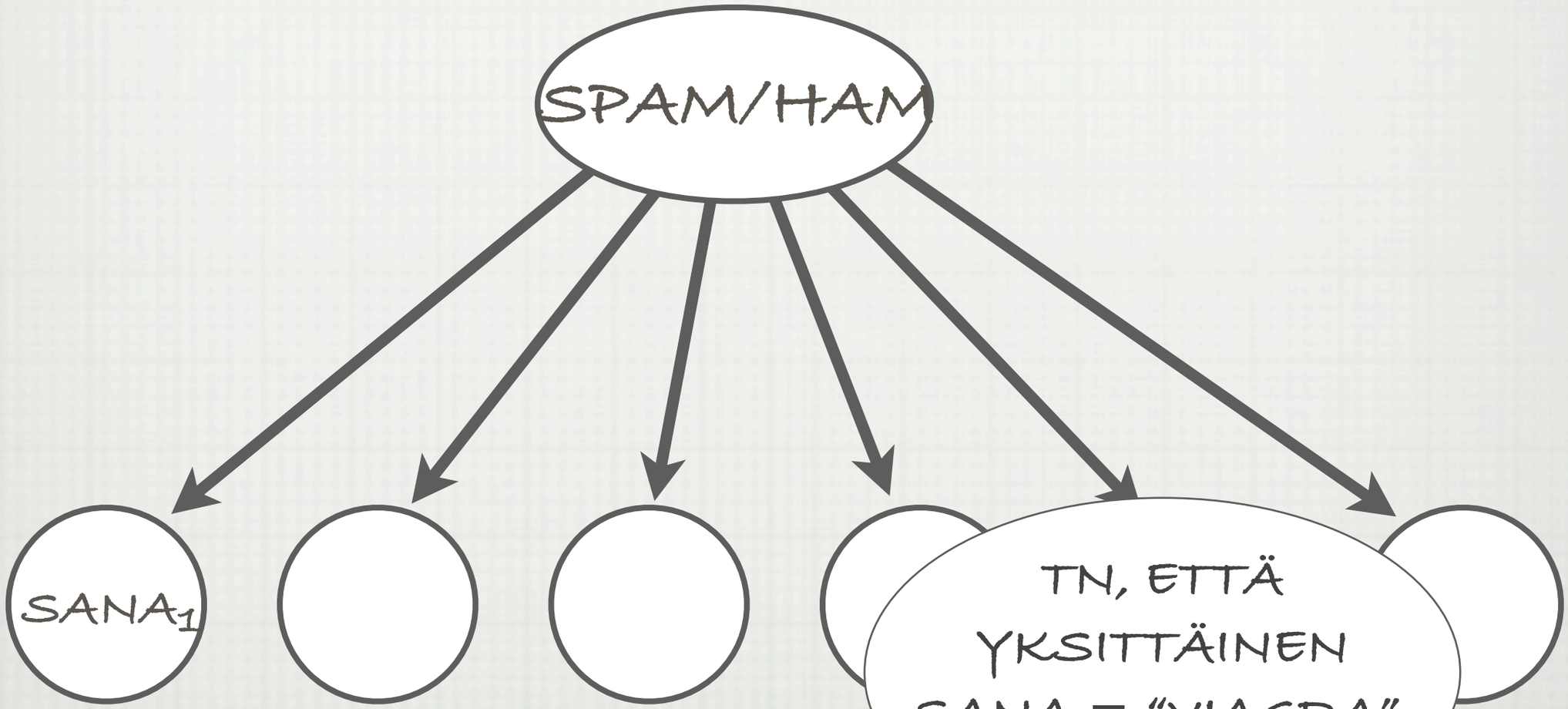
TEEMU ROOS

HELSINGIN YLIOPISTO

NAIVI BAYES



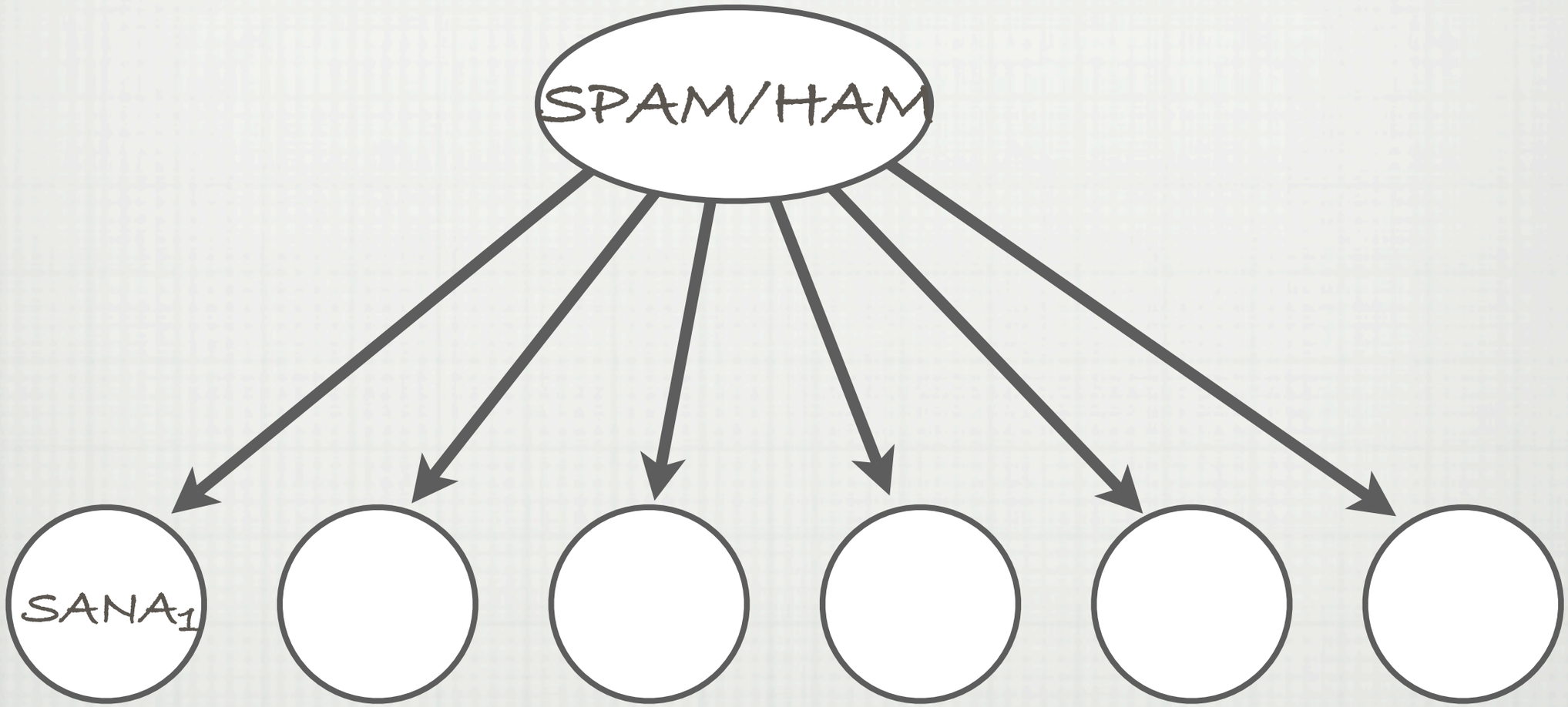
NAIVI BAYES



$P(SANA_i = VIAGRA \mid HAM) = 0.0001$

$P(SANA_i = VIAGRA \mid SPAM) = 0.002$

NAIVI BAYES



$P(SANA_i = \$ \mid HAM) = 0.0002$

$P(SANA_i = \$ \mid SPAM) = 0.005$

NAIVI BAYES

```
FROM: "MARGARETTA NITA" <MARGUERITESEBRINA@WMLE.COM>
SUBJECT: SPECIAL OFFER : VIAGRA ON SALE AT $1.38 !!!
X-BOGOSITY: YES, TESTS=BOGOFILTER, SPAMICITY=0.99993752,
VERSION=2011-08-29
DATE: MON, 26 SEP 2011 21:52:26 +0300
X-CLASSIFICATION: JUNK - AD HOC SPAM DETECTED (CODE = 73)
```

SPECIAL OFFER : VIAGRA ON SALE AT $1.38 !!!

COMPARE THE BEST ONLINE PHARMACIES TO **BUY VIAGRA**. ORDER **VIAGRA** ONLINE WITH HUGE **DISCOUNT**.
MULTIPLE BENEFITS INCLUDE **FREE** SHIPPING, REORDER **DISCOUNTS**, BONUS PILLS
HTTP://RXPHARMACYCVS.RU

JOHDATUS TOD.NÄK.LASKENTAAN

1. $P(A,B,C) = P(A)\,P(B|A)\,P(C|A,B)$ // KETJUSÄÄNTÖ
2. $P(A) = P(A,B) + P(A,\neg B)$ // "MARGINALISOINTI"
3. $P(A|B) = P(A,B) / P(B)$ // EHDOLLINEN TN.

$P(A)\,P(B|A) = P(B)\,P(A|B)$??? (KS. SÄÄNTÖ 1)

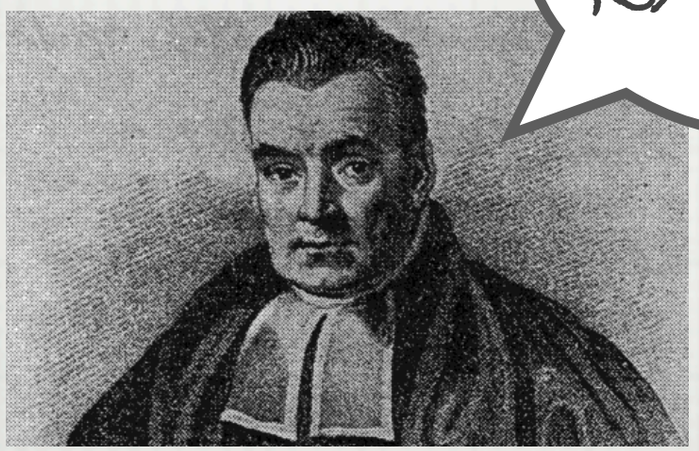
4. $P(B|A) = P(B)\,P(A|B) / P(A)$ // BAYESIN KAAVA
5. $A \perp B \Rightarrow P(A|B) = P(A)$ // RIIPPUMATTOMUUS

NAIVI BAYES

PÄÄTTELY:

1. $P(SPAM) = 0.5$
2. $P(SPAM \mid SANA_1 = VIAGRA) = \frac{P(SPAM)\, P(SANA_1 = VIAGRA \mid SPAM)}{P(SANA_1 = VIAGRA)}$

BAYESIN KAAVA!



NAIVI BAYES

PÄÄTTELY:

1. $P(SPAM) = 0.5$
2. $P(SPAM \mid SANA_1 = VIAGRA) = \frac{P(SPAM)\, P(SANA_1 = VIAGRA \mid SPAM)}{P(SANA_1 = VIAGRA)}$
3. $P(SPAM \mid SANA_1 = VIAGRA, SANA_2 = IS) = \frac{P(SPAM)\, P(SANA_1 = VIAGRA, SANA_2 = IS \mid SPAM)}{P(SANA_1 = VIAGRA, SANA_2 = IS)}$
4. $P(SPAM \mid SANA_1 = VIAGRA, SANA_2 = IS, SANA_3 = ALGORITHM) = \frac{P(SPAM)\, P(SANA_1 = VIAGRA, SANA_2 = IS, SANA_3 = ALGORITHM \mid SPAM)}{P(SANA_1 = VIAGRA, SANA_2 = IS, SANA_3 = ALGORITHM)}$

NAIVI BAYES

PÄÄTTELY:

1. $P(SPAM) = 0.5$
2. $P(SPAM \mid SANA_1 = VIAGRA) = \frac{P(SPAM)\, P(SANA_1 = VIAGRA \mid SPAM)}{P(SANA_1 = VIAGRA)}$
3. $P(SPAM \mid SANA_1 = \ldots$

$P(SPAM)\, P(\ldots$

$= \frac{\ldots}{P(SAN\ldots}$

$\ldots PAM)$

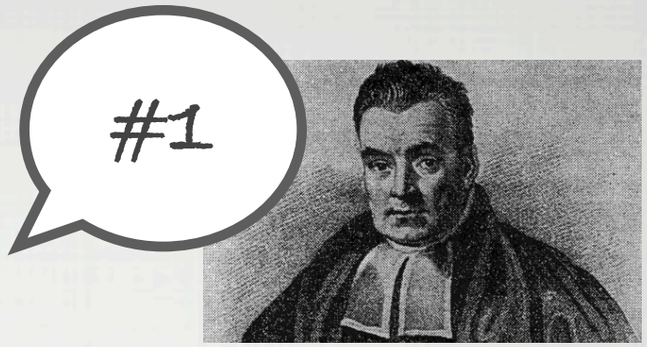
4. $P(SPAM \mid SANA_1 = VIAGRA, SANA_2 = IS, \ldots$

$P(SPAM)\, P(SANA_1 = VIAGRA, SANA_2 = IS, \ldots$

$= \frac{\ldots}{P(SANA_1 = VIAGRA, SANA_2 = IS, \ldots}$

PARI TÄRKEÄÄ JUTTUA TÄSTÄ...

NAIVI BAYES



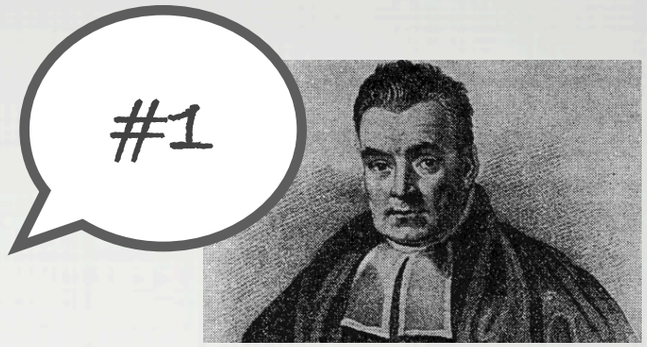
PÄÄTTELY:

4. $P(SPAM \mid SANA_1=VIAGRA, SANA_2=IS, SANA_3=ALGORITHM)$

$$= \frac{P(SPAM)\, P(SANA_1=VIAGRA, SANA_2=IS, SANA_3=ALGORITHM \mid SPAM)}{P(SANA_1=VIAGRA, SANA_2=IS, SANA_3=ALGORITHM)}$$

$$P(SANA_1=VIAGRA, SANA_2=IS, SANA_3=ALGORITHM)$$
$$= P(SPAM, SANA_1=VIAGRA, SANA_2=IS, SANA_3=ALGORITHM)$$
$$+ P(\neg SPAM, SANA_1=VIAGRA, SANA_2=IS, SANA_3=ALGORITHM)$$

NAIVI BAYES



PÄÄTTELY:

4. $P(SPAM \mid SANA_1 = VIAGRA, SANA_2 = IS, SANA_3 = ALGORITHM)$

$$= \frac{P(SPAM)\, P(SANA_1 = VIAGRA, SANA_2 = IS, SANA_3 = ALGORITHM \mid SPAM)}{P(SANA_1 = VIAGRA, SANA_2 = IS, SANA_3 = ALGORITHM)}$$

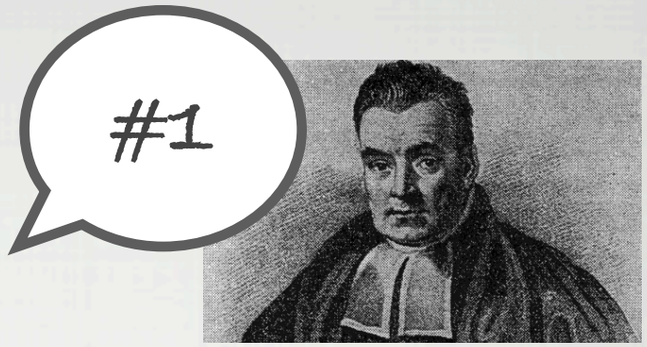
$P(EVIDENSSI)$

$= P(SPAM, EVIDENSSI)$ // MARGINALISOINTI

$+ P(\neg SPAM, EVIDENSSI)$

$$P(SPAM \mid EVIDENSSI) = \frac{P(SPAM)P(EVIDENSSI \mid SPAM)}{P(SPAM, EVIDENSSI) + P(\neg SPAM, EVIDENSSI)}$$

NAIVI BAYES



PÄÄTTELY:

4. $P(SPAM \mid SANA_1 = VIAGRA, SANA_2 = IS, SANA_3 = ALGORITHM)$

$$= \frac{P(SPAM)\,P(SANA_1 = VIAGRA, SANA_2 = IS, SANA_3 = ALGORITHM \mid SPAM)}{P(SANA_1 = VIAGRA, SANA_2 = IS, SANA_3 = ALGORITHM)}$$

$P(EVIDENSSI)$

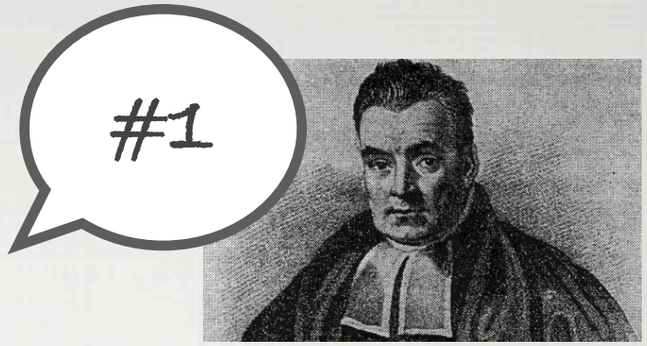
$= P(SPAM, EVIDENSSI)$

$+ P(\neg SPAM, EVIDENSSI)$

$$P(SPAM \mid EVIDENSSI) = \frac{P(SPAM)P(EVIDENSSI \mid SPAM)}{\cancel{P(SPAM, EVIDENSSI) + P(\neg SPAM, EVIDENSSI)}}$$

$$P(\neg SPAM \mid EVIDENSSI) = \frac{P(\neg SPAM)P(EVIDENSSI \mid \neg SPAM)}{\cancel{P(SPAM, EVIDENSSI) + P(\neg SPAM, EVIDENSSI)}}$$

NAIVI BAYES



PÄÄTTELY:

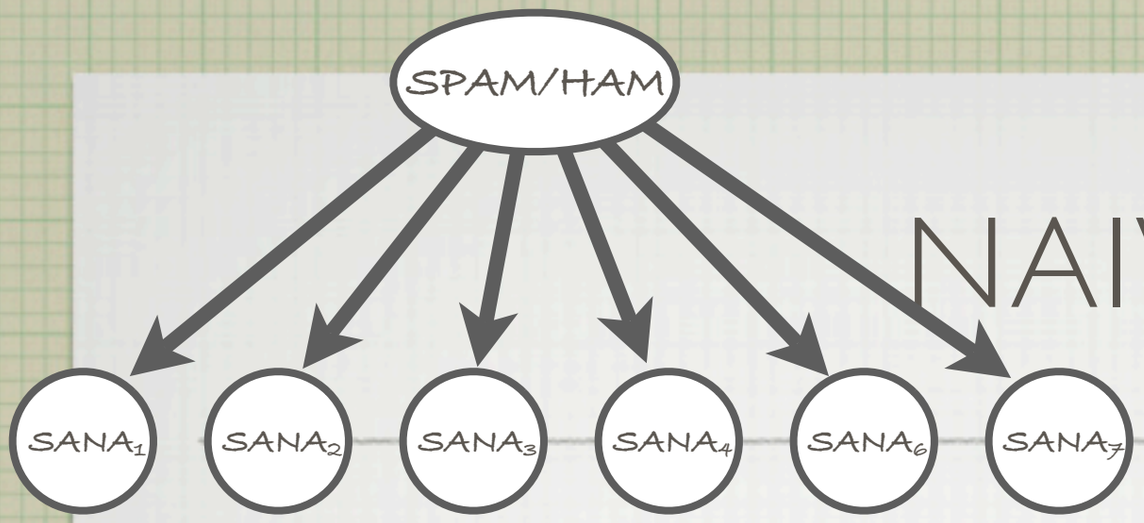
4. $P(SPAM \mid SANA_1 = VIAGRA, SANA_2 = IS, SANA_3 = ALGORITHM)$

$$= \frac{P(SPAM)\,P(SANA_1 = VIAGRA, SANA_2 = IS, SANA_3 = ALGORITHM \mid SPAM)}{P(SANA_1 = VIAGRA, SANA_2 = IS, SANA_3 = ALGORITHM)}$$

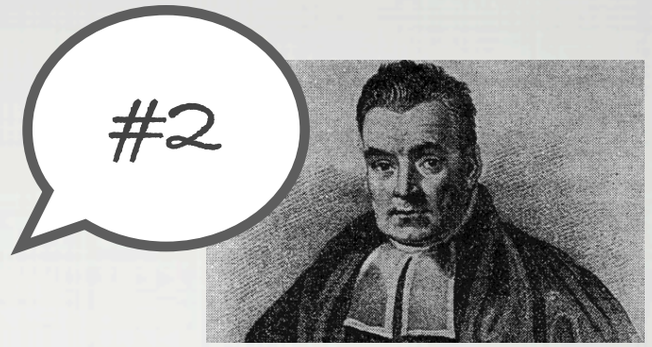
$$P(EVIDENSSI) = P(SPAM, EVIDENSSI) + P(\neg SPAM, EVIDENSSI)$$

$$\frac{P(S \mid E)}{P(\neg S \mid E)} = \frac{P(S)P(E \mid S)}{P(\neg S)P(E \mid \neg S)}$$

$$\frac{P(SPAM \mid EVIDENSSI)}{P(\neg SPAM \mid EVIDENSSI)} = \frac{P(SPAM)P(EVIDENSSI \mid SPAM)}{P(\neg SPAM)P(EVIDENSSI \mid \neg SPAM)}$$



NAIVI BAYES



PÄÄTTELY:

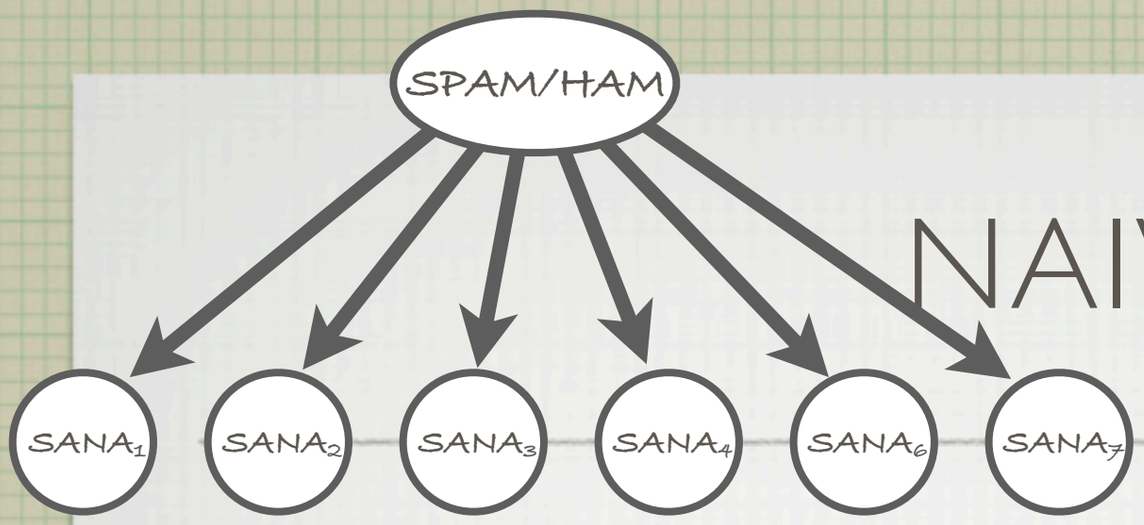
$P(SANA_1 = VIAGRA, SANA_2 = IS, SANA_3 = ALGORITHM \mid SPAM)$

$= P(SANA_1 = VIAGRA \mid SPAM)$ // KETJUSÄÄNTÖ

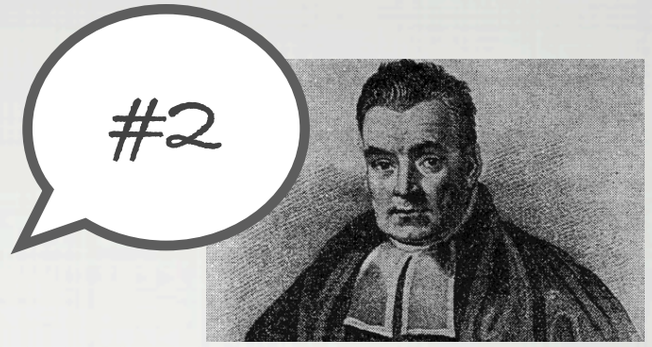
$P(SANA_2 = IS \mid \cancel{SANA_1 = VIAGRA}, SPAM)$

$P(SANA_3 = ALGORITHM \mid \cancel{SANA_1 = VIAGRA, SANA_2 = IS}, SPAM)$

// RIIPPUMATTOMUUS



NAIVI BAYES



PÄÄTTELY:

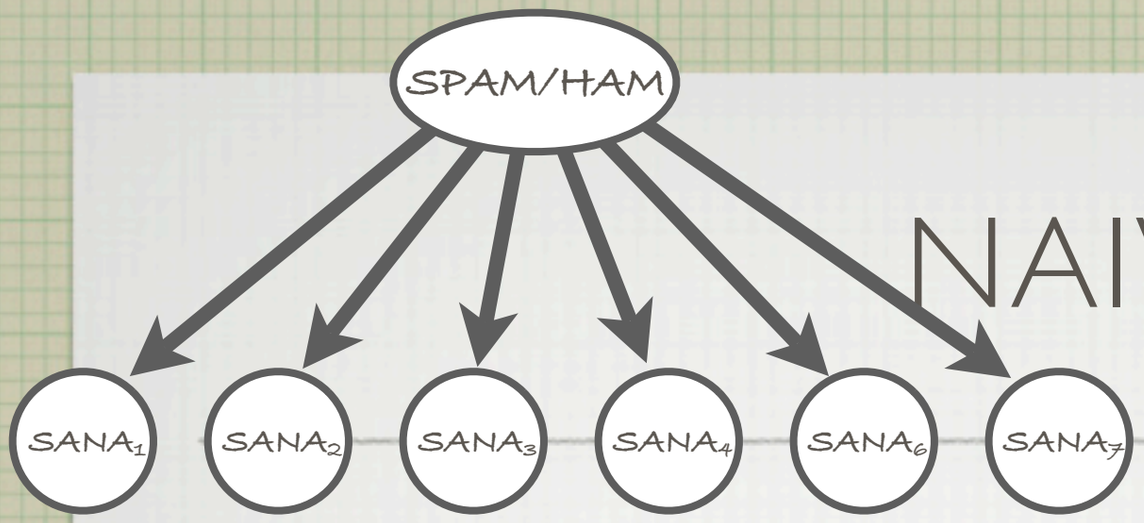
$P(SANA_1 = VIAGRA, SANA_2 = IS, SANA_3 = ALGORITHM \mid SPAM)$

$= P(SANA_1 = VIAGRA \mid SPAM)$
$P(SANA_2 = IS \mid SPAM)$
$P(SANA_3 = ALGORITHM \mid SPAM)$

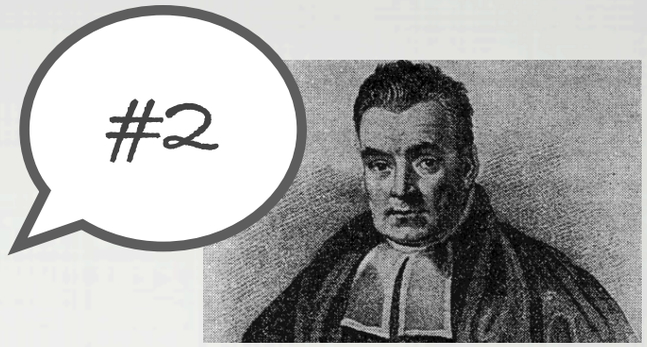
$$\frac{P(S \mid E)}{P(\neg S \mid E)} = \frac{P(S)P(E \mid S)}{P(\neg S)P(E \mid \neg S)}$$

$P(SANA_1 = VIAGRA, SANA_2 = IS, SANA_3 = ALGORITHM \mid \neg SPAM)$

$= P(SANA_1 = VIAGRA \mid \neg SPAM)$ // KETJUSÄÄNTÖ
$P(SANA_2 = IS \mid \neg SPAM)$
$P(SANA_3 = ALGORITHM \mid \neg SPAM)$



NAIVI BAYES



PÄÄTTELY:

$P(SPAM|EVIDENSSI) / P(\neg SPAM|EVIDENSSI)$

$= P(SPAM) / P(\neg SPAM)$

$P(SANA_1=VIAGRA|SPAM) / P(SANA_1=VIAGRA|\neg SPAM)$ >1

$P(SANA_2=IS|SPAM) / P(SANA_2=IS|\neg SPAM)$ $=1$

$P(SANA_3=ALGORITHM|SPAM) / P(SANA_3=ALGORITHM|\neg SPAM)$ <1

...

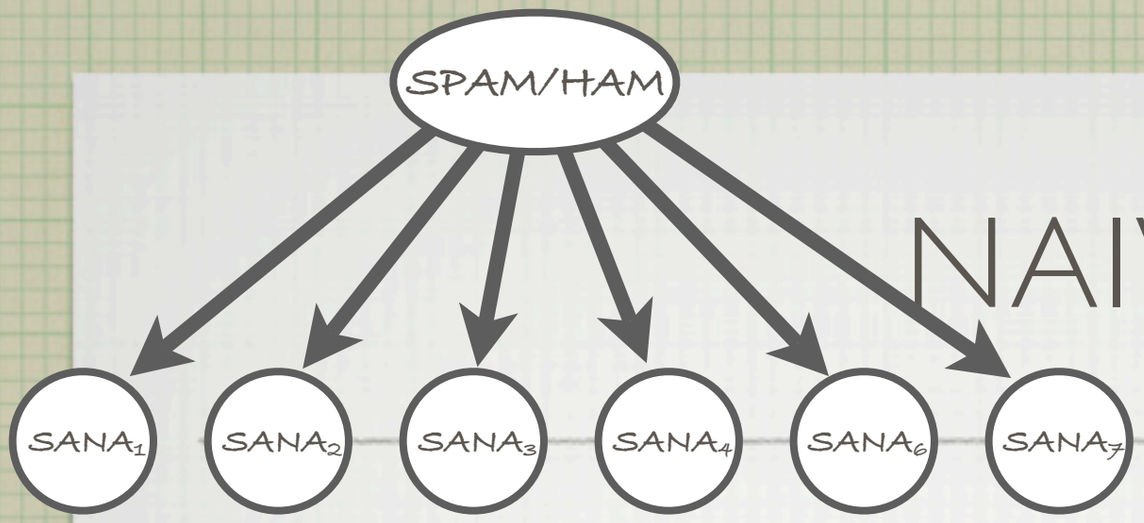
...

...

OSAMÄÄRÄ MÄÄRÄÄ!

$> 1 \Rightarrow$ "SPAM"

$< 1 \Rightarrow$ "HAM"



NAIVI BAYES

YHTEENVETO TOISTAISEKSI:

- TARVITAAN:
  - "PRIORIJAKAUMA" $P(SPAM) = 0.\_\_\_$
  - "LUOKKAEHDOLLISET" JAKAUMAT

$P(SANA_i=VIAGRA|SPAM)=0.\_\_\_$ $P(SANA_i=VIAGRA|\neg SPAM)=0.\_\_\_$

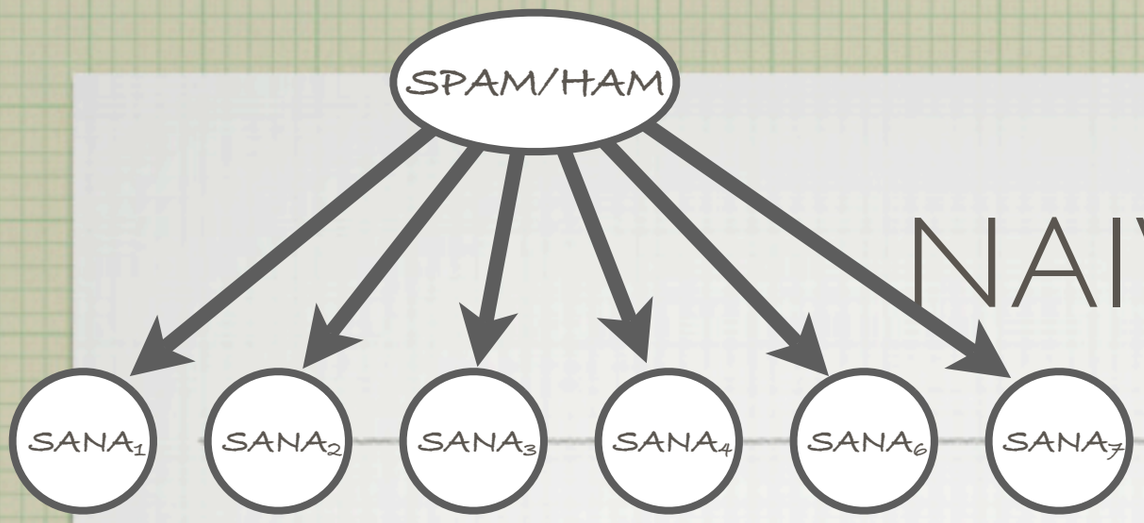
$P(SANA_i=IS|SPAM) = 0.\_\_\_\_$ $P(SANA_i=IS|\neg SPAM)=0.\_\_\_$

$P(SANA_i=ALGORITHM|SPAM) = 0.\_\_\_$ $P(SANA_i=ALG.|\neg SPAM)=0.\_\_\_$

- OLETETAAN ETTÄ $P(SANA_i|SANA_j,SPAM) = P(SANA_i|SPAM)$
  (EHDOLLINEN RIIPPUMATTOMUUS)

- OLENNAISTA ON OSAMÄÄRÄ $\dfrac{P(SANA_i=VIAGRA|SPAM)}{P(SANA_i=VIAGRA|\neg SPAM)}$

  (OTETAAN NÄIDEN TULO)



NAIVI BAYES

PSEUDOKOODINA:

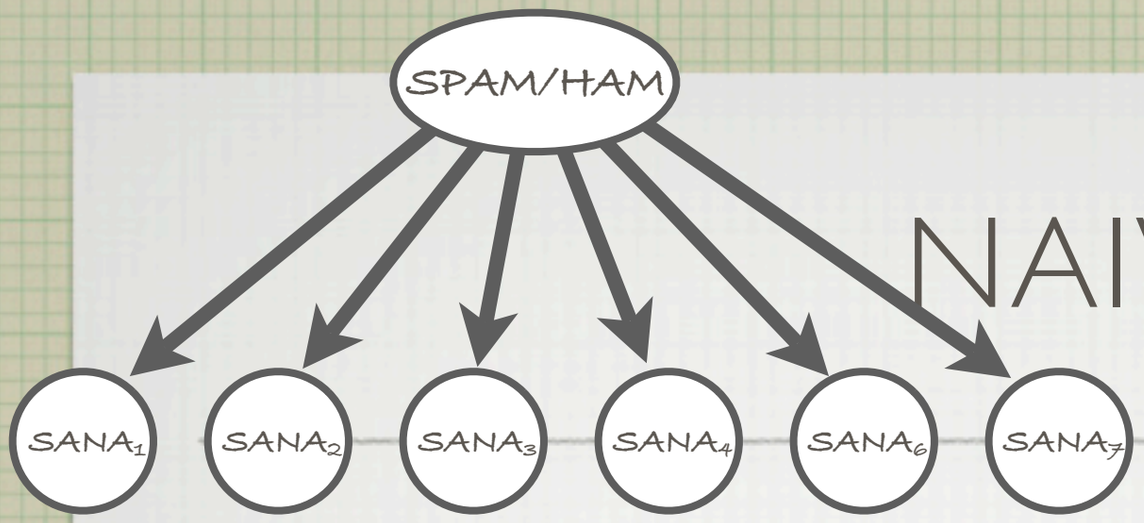
SPAMICITY(Viesti, P):

Odds = P.Spam / P.noSpam // P.Spam + P.noSpam = 1
for each Sana in Viesti
Odds = Odds * P.Sana_Spam(Sana) / P.Sana_noSpam(Sana)
return(Odds)

PÄÄTTELY:

$P(SPAM|EVIDENSSI)/P(\neg SPAM|EVIDENSSI)$

$= P(SPAM) / P(\neg SPAM)$
$P(SANA_1=VIAGRA|SPAM) / P(SANA_1=VIAGRA|$
$P(SANA_2=IS|SPAM) / P(SANA_2=IS|\neg SPAM)$
$P(SANA_3=ALGORITHM|SPAM) / P(SANA_3=ALG$



NAIVI BAYES

PSEUDOKOODINA:

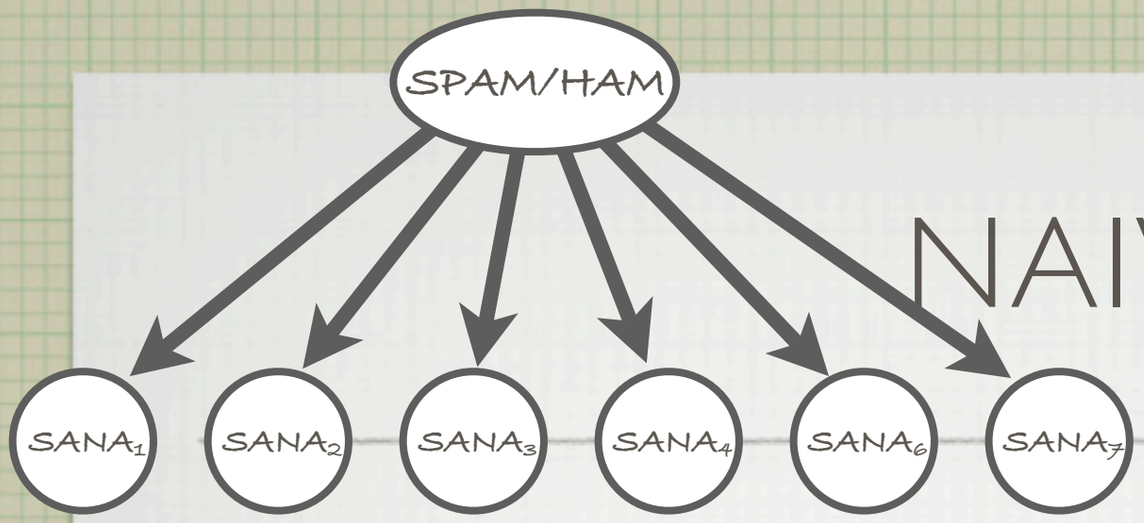
SPAMICITY(Viesti, P):

Odds = P.Spam / P.noSpam // P.Spam + P.noSpam = 1
for each Sana in Viesti
  Odds = Odds * P.Sana_Spam(Sana) /P.Sana_noSpam(Sana)
return(Odds)

PÄÄTTELY:

P(SPAM|EVIDENSSI)/P(¬SPAM|EVIDENSSI)

= P(SPAM) / P(¬SPAM)
P(SANA1=VIAGRA|SPAM) / P(SANA1=VIAGRA|
P(SANA2=IS|SPAM) / P(SANA2=IS|¬SPAM)
P(SANA3=ALGORITHM|SPAM) / P(SANA3=ALG



NAIVI BAYES

PSEUDOKOODINA:

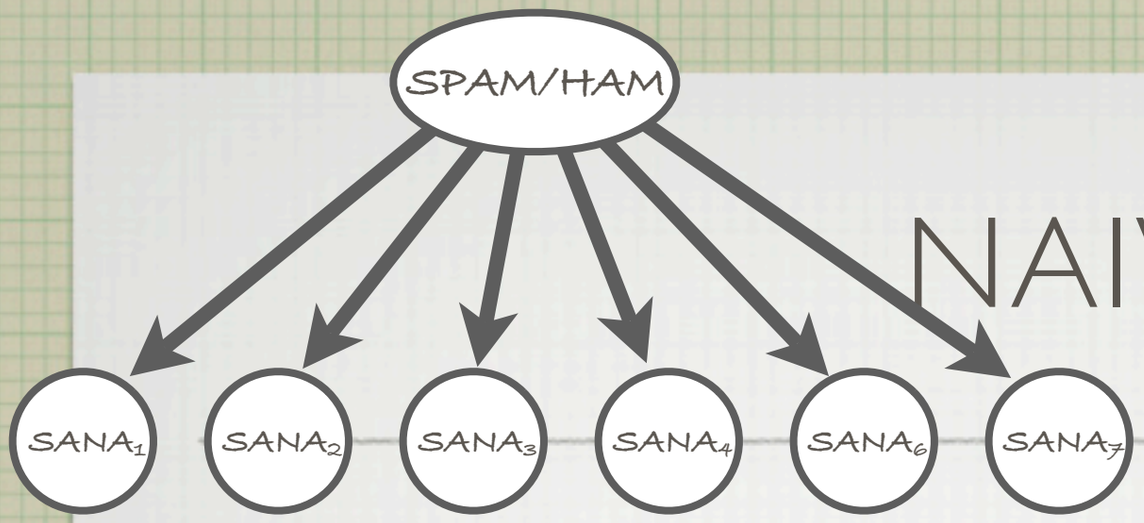
SPAMICITY(Viesti, P):

```
Odds = P.Spam / P.noSpam    // P.Spam + P.noSpam = 1
for each Sana in Viesti
   Odds = Odds * P.Sana_Spam(Sana) /P.Sana_noSpam(Sana)
return(Odds)
```

JOS SPAMICITY(Viesti, P) >1, LUOKITTELE VIESTI SPAMIKSI

JOS SPAMICITY(Viesti, P) <1, LUOKITTELE VIESTI HAMIKSI

JOS SPAMICITY(Viesti, P) =1, EN TIEDÄ



NAIVI BAYES

PSEUDOKOODINA:

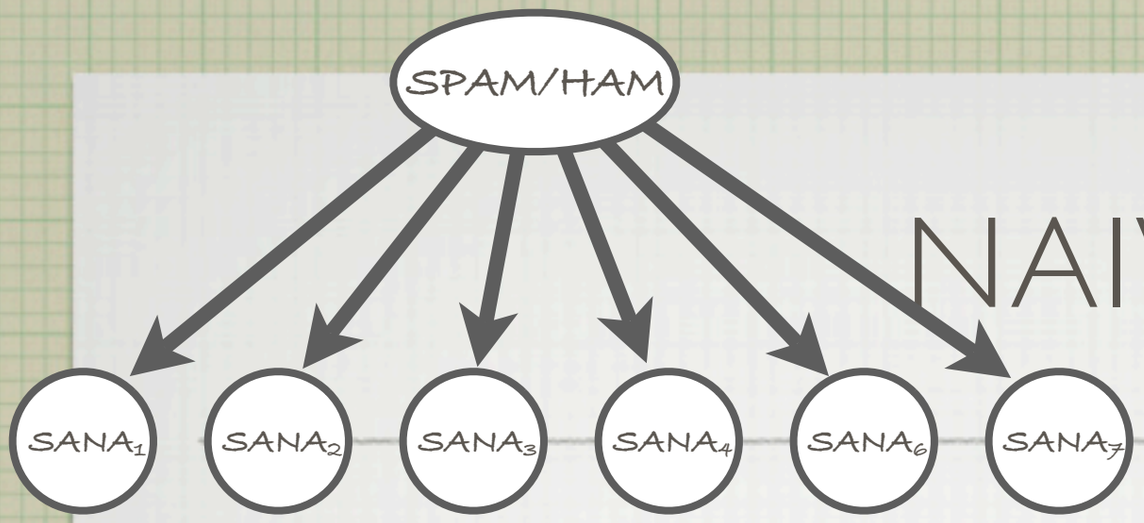
SPAMICITY(Viesti, P):

```
Odds = P.Spam / P.noSpam    // P.Spam + P.noSpam = 1
for each Sana in Viesti
   Odds = Odds * P.Sana_Spam(Sana) /P.Sana_noSpam(Sana)
return(Odds)
```

JOS SPAMICITY(Viesti, P) $>1+\alpha$, LUOKITTELE VIESTI SPAMIKSI

JOS SPAMICITY(Viesti, P) $<1-\beta$, LUOKITTELE VIESTI HAMIKSI

MUUTEN, EN TIEDÄ



NAIVI BAYES

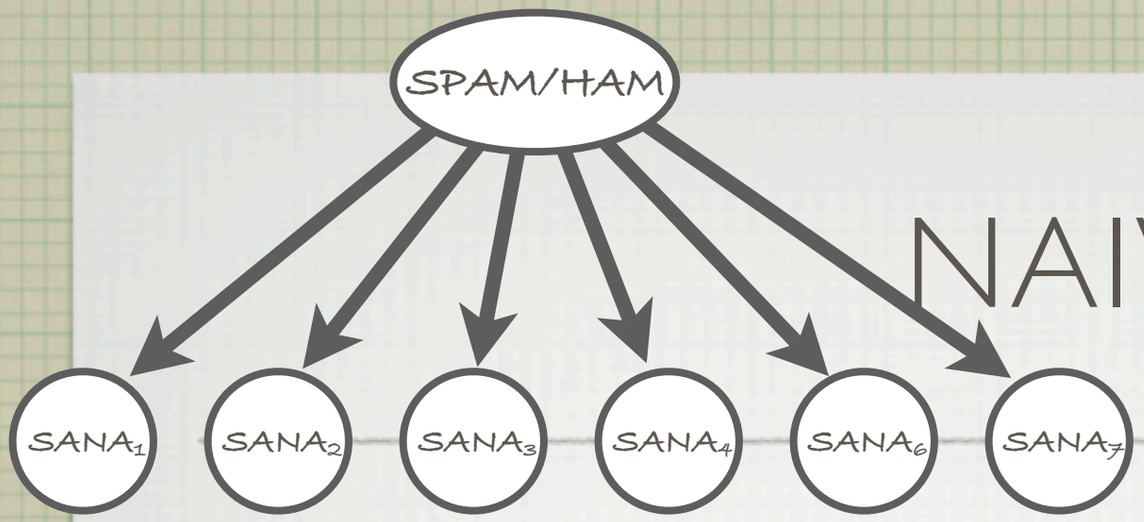
PSEUDOKOODINA:

LOG(A*B) = LOG(A) + LOG(B)

SPAMICITY(Viesti, P):

Odds = P.Spam / P.noSpam // P.Spam + P.noSpam = 1
for each Sana in Viesti
  Odds = Odds * P.Sana_Spam(Sana) / P.Sana_noSpam(Sana)
return(Odds)

KÄYTÄNNÖN ONGELMA: ALI- JA YLIVUODOT
Odds ARVOSTA TULEE HELPOSTI LIIAN PIENI (LÄHELLÄ NOLLAA) TAI LIIAN SUURI.
RATKAISU:
KÄYTÄ log(Odds)



NAIVI BAYES

PSEUDOKOODINA:

LOG(A*B) = LOG(A) + LOG(B)

SPAMICITY(Viesti, P):

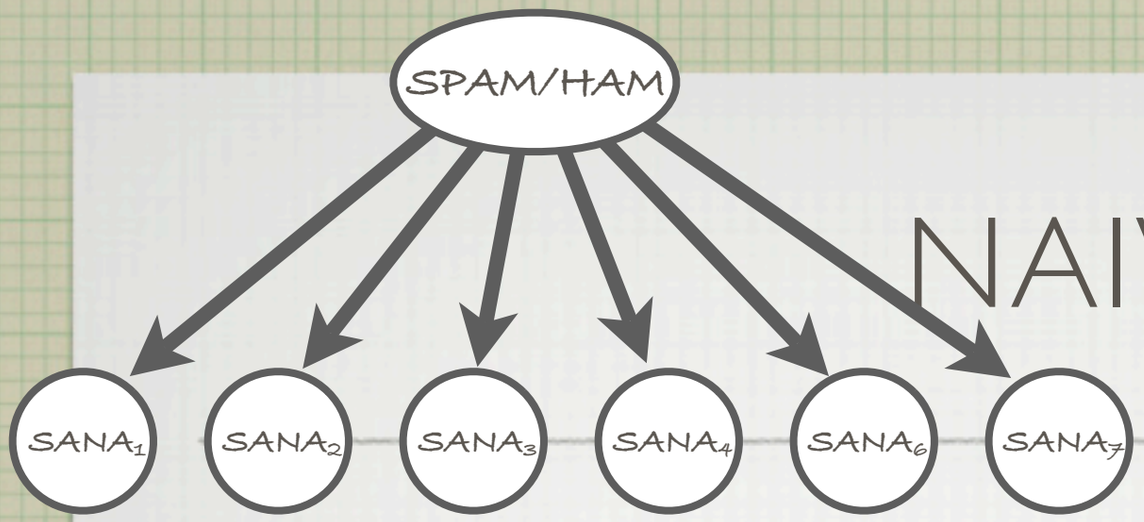
logOdds = log(P.Spam / P.noSpam) // P.Spam + P.noSpam = 1
for each Sana in Viesti
  logOdds = logOdds + log(P.Sana_Spam(Sana) / P.Sana_noSpam(Sana))
return(exp(logOdds))

KÄYTÄNNÖN ONGELMA: ALI- JA YLIVUODOT

Odds ARVOSTA TULEE HELPOSTI LIIAN PIENI (LÄHELLÄ NOLLAA) TAI LIIAN SUURI.

RATKAISU:

KÄYTÄ log(Odds)

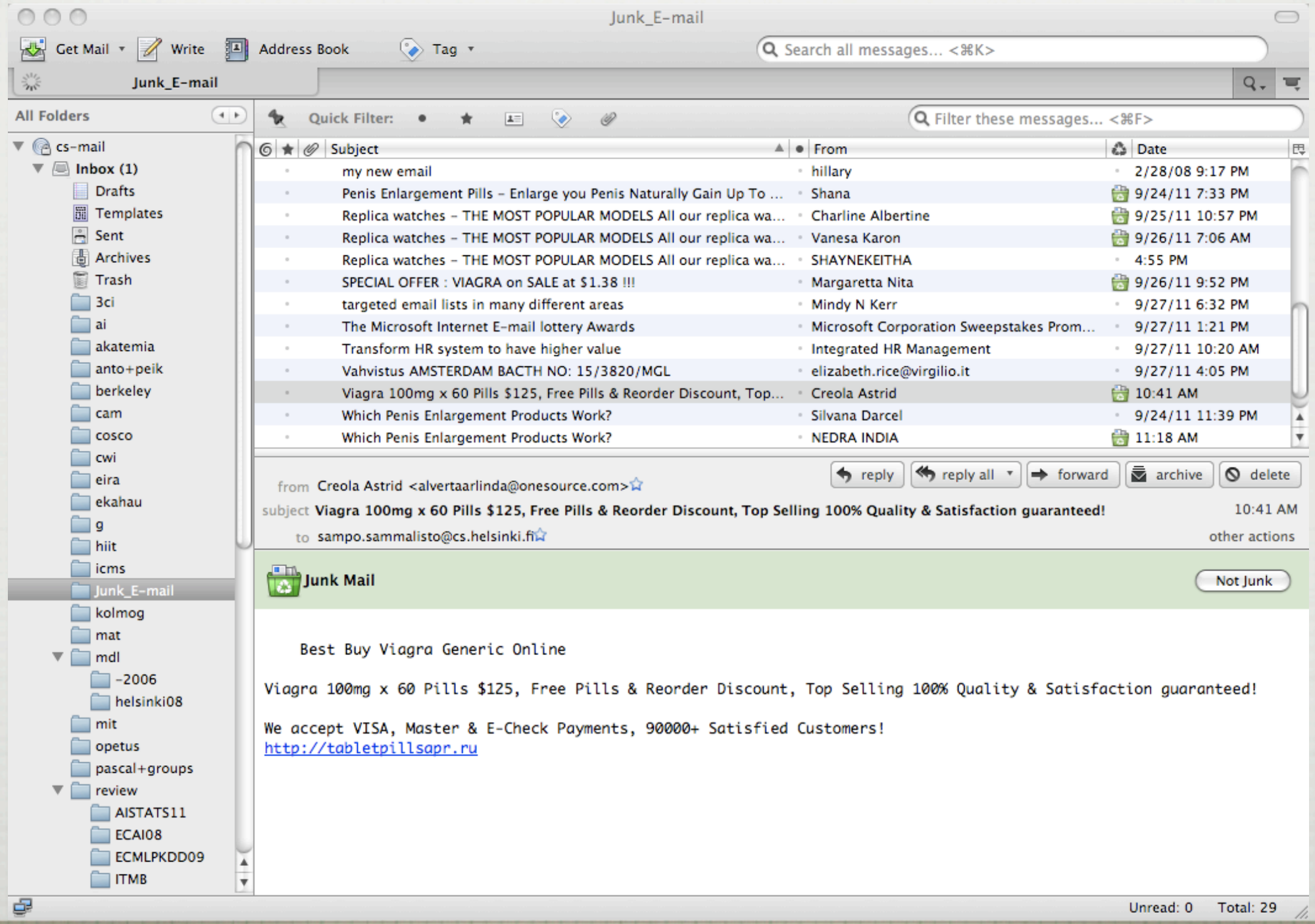


NAIVI BAYES

PARAMETRIEN OPPIMISESTA

- VAIKEA KEKSIÄ PÄÄSTÄ EHDOLLISIA TN:IÄ.
- HUONOT ARVOT HUONONTAVAT FILTTERIN TOIMINTAA
- PAREMPI RATKAISU:
  1. KERÄÄ ISO KASA SPAM-VIESTEJÄ
  2. KERÄÄ ISO KASA HAM-VIESTEJÄ
  3. ARVIOI $P(SANA_i = \_\_\_\_) = 0.\_\_\_\_$ LASKEMALLA DATASTA
     (VRT. LASKUHARJ. 2.3)
- VAROTTAVA NOLLATODENNÄKÖISYYKSIÄ!
  (JOS KASA EI TARPEEKSI ISO, JOTKUT SANAT EIVÄT VAIN SATU ESIINTYMÄÄN SIINÄ.)

ESIM.


Junk_E-mail
Get Mail
Write
Address Book
Tag
Search all messages... <⌘K>
Junk_E-mail
All Folders
Quick Filter:
Filter these messages... <⌘F>
cs-mail
Inbox (1)
Drafts
Templates
Sent
Archives
Trash
3ci
ai
akatemia
anto+peik
berkeley
cam
cosco
cwi
eira
ekahau
g
hiit
icms
Junk_E-mail
kolmog
mat
mdl
-2006
helsinki08
mit
opetus
pascal+groups
review
AISTATS11
ECAI08
ECMLPKDD09
ITMB
Subject
From
Date
my new email
hillary
2/28/08 9:17 PM
Penis Enlargement Pills – Enlarge you Penis Naturally Gain Up To ...
Shana
9/24/11 7:33 PM
Replica watches – THE MOST POPULAR MODELS All our replica wa...
Charline Albertine
9/25/11 10:57 PM
Replica watches – THE MOST POPULAR MODELS All our replica wa...
Vanesa Karon
9/26/11 7:06 AM
Replica watches – THE MOST POPULAR MODELS All our replica wa...
SHAYNEKEITHA
4:55 PM
SPECIAL OFFER : VIAGRA on SALE at $1.38 !!!
Margaretta Nita
9/26/11 9:52 PM
targeted email lists in many different areas
Mindy N Kerr
9/27/11 6:32 PM
The Microsoft Internet E-mail lottery Awards
Microsoft Corporation Sweepstakes Prom...
9/27/11 1:21 PM
Transform HR system to have higher value
Integrated HR Management
9/27/11 10:20 AM
Vahvistus AMSTERDAM BACTH NO: 15/3820/MGL
elizabeth.rice@virgilio.it
9/27/11 4:05 PM
Viagra 100mg x 60 Pills $125, Free Pills & Reorder Discount, Top...
Creola Astrid
10:41 AM
Which Penis Enlargement Products Work?
Silvana Darcel
9/24/11 11:39 PM
Which Penis Enlargement Products Work?
NEDRA INDIA
11:18 AM
reply
reply all
forward
archive
delete
from Creola Astrid <alvertaarlinda@onesource.com>
subject Viagra 100mg x 60 Pills $125, Free Pills & Reorder Discount, Top Selling 100% Quality & Satisfaction guaranteed!
10:41 AM
to sampo.sammalisto@cs.helsinki.fi
other actions
Junk Mail
Not Junk
Best Buy Viagra Generic Online
Viagra 100mg x 60 Pills $125, Free Pills & Reorder Discount, Top Selling 100% Quality & Satisfaction guaranteed!
We accept VISA, Master & E-Check Payments, 90000+ Satisfied Customers!
http://tabletpillsapr.ru
Unread: 0
Total: 29

ESIM.

ESIM.

SPAM			HAM		
1	MONEY	0.04 %	21	ALGORITHM	0.01 %
...			...		
5	VIAGRA	0.21 %	62	MONEY	0.02 %
...			...		
10	IS	0.42 %	2199	FOR	0.78 %
...			2492	THAT	0.88 %
19	REPLICA	0.80 %	2990	YOU	1.05 %
20	EMAIL	0.84 %	3141	IN	1.11 %
20	YOU	0.84 %	3160	I	1.11 %
21	DATABASE	0.88 %	3218	AND	1.13 %
25	EMAILS	1.05 %	3283	IS	1.16 %
26	OF	1.09 %	3472	OF	1.22 %
31	TO	1.30 %	3874	A	1.37 %
43	AND	1.80 %	5442	TO	1.92 %
48	THE	2.01 %	9196	THE	3.24 %
TOTAL 2386			TOTAL 283736		

ESIM.

SPAM			HAM		
1	MONEY	0.04 %	21	ALGORITHM	0.01 %
...			...		
5	VIAGRA	0.21 %	62	MONEY	0.02 %
...			...		
10	I				
...					
19	R				
20	E				
20	Y				
21	DATABASE	0.88 %	3218	AND	1.13 %
25	EMAILS	1.05 %	3283	IS	1.16 %
26	OF	1.09 %	3472	OF	1.22 %
31	TO	1.30 %	3874	A	1.37 %
43	AND	1.80 %	5442	TO	1.92 %
48	THE	2.01 %	9196	THE	3.24 %
TOTAL 2386			TOTAL 283736		

$$\frac{P(SANA_i = MONEY \mid SPAM)}{P(SANA_i = MONEY \mid \neg SPAM)} = \frac{0.0004}{0.0002} = 1.918 > 1$$

ESIM.

SPAM		HAM	
1 MONEY	0.04 %	21 ALGORITHM	0.01 %
...		...	
5 VIAGRA	0.21 %	62 MONEY	0.02 %
...		...	
10 I			
...			
19 R			
20 E			
20 Y			
21 DATABASE	0.88 %	3218 AND	1.13 %
25 E			
26 O			
31 T			
43 A			
48 T			
TOTAL 2386		TOTAL 283736	

$$\frac{P(SANA_i = MONEY \mid SPAM)}{P(SANA_i = MONEY \mid \neg SPAM)} = \frac{0.0004}{0.0002} = 1.918 > 1$$

$$\frac{P(SANA_i = IS \mid SPAM)}{P(SANA_i = IS \mid \neg SPAM)} = \frac{0.0042}{0.0116} = 0.3622 < 1$$

YHTEENVETO

YHTEENVETO NAIVI BAYES-SPAMFILTTERISTÄ:

- TARVITAAN:
  - "PRIORIJAKAUMA" $P(SPAM) = 0.\_\_\_$
  - "LUOKKAEHDOLLISET" JAKAUMAT

$P(SANA_i=VIAGRA|SPAM)=0.\_\_\_$ $P(SANA_i=VIAGRA|\neg SPAM)=0.\_\_\_$

$P(SANA_i=IS|SPAM) = 0.\_\_\_\_$ $P(SANA_i=IS|\neg SPAM)=0.\_\_\_$

$P(SANA_i=ALGORITHM|SPAM) = 0.\_\_\_$ $P(SANA_i=ALG.|\neg SPAM)=0.\_\_\_$

- OLETETAAN ETTÄ $P(SANA_i|SANA_j,SPAM) = P(SANA_i|SPAM)$
  (EHDOLLINEN RIIPPUMATTOMUUS)

- OLENNAISTA ON OSAMÄÄRÄ

$$\frac{P(SANA_i=VIAGRA|SPAM)}{P(SANA_i=VIAGRA|\neg SPAM)}$$

YHTEENVETO

(JATKOA...):

- YLI- JA ALIVUOTOJEN VÄLTTÄMISEKSI KANNATTAA KÄYTTÄÄ LOGARITMIMUUNNOSTA (LOG(A*B)=LOG(A)+LOG(B))

- JAKAUMAT PARAS ESTIMOIDA DATASTA

- NOLLATODENNÄKÖISYYKSILLE TEHTÄVÄ JOTAIN