



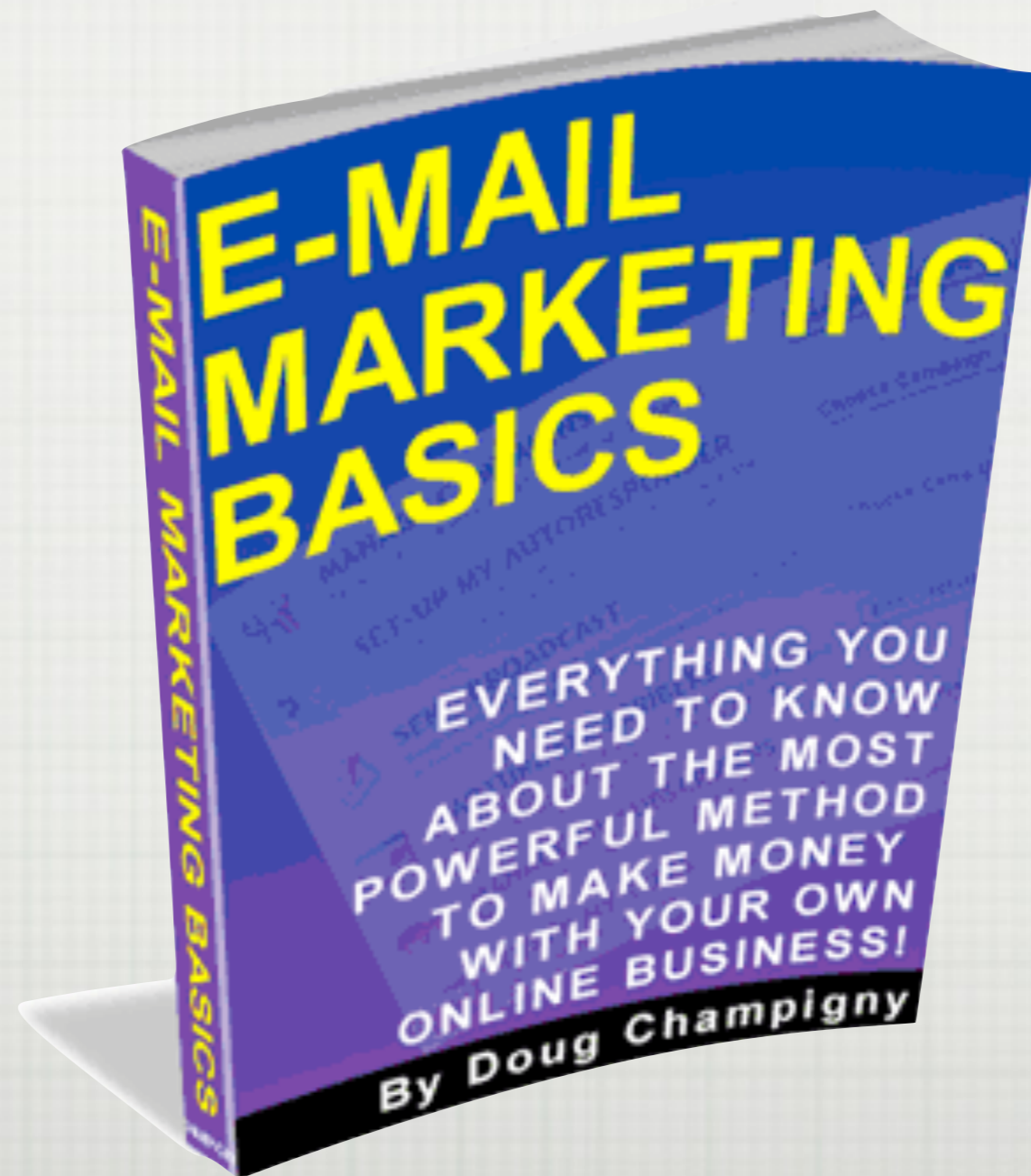
JOHDATUS TEKOÄLYYN

TEEMU ROOS



HELSINGIN YLIOPISTO

KUINKA RIKASTUA



NAIVI BAYES

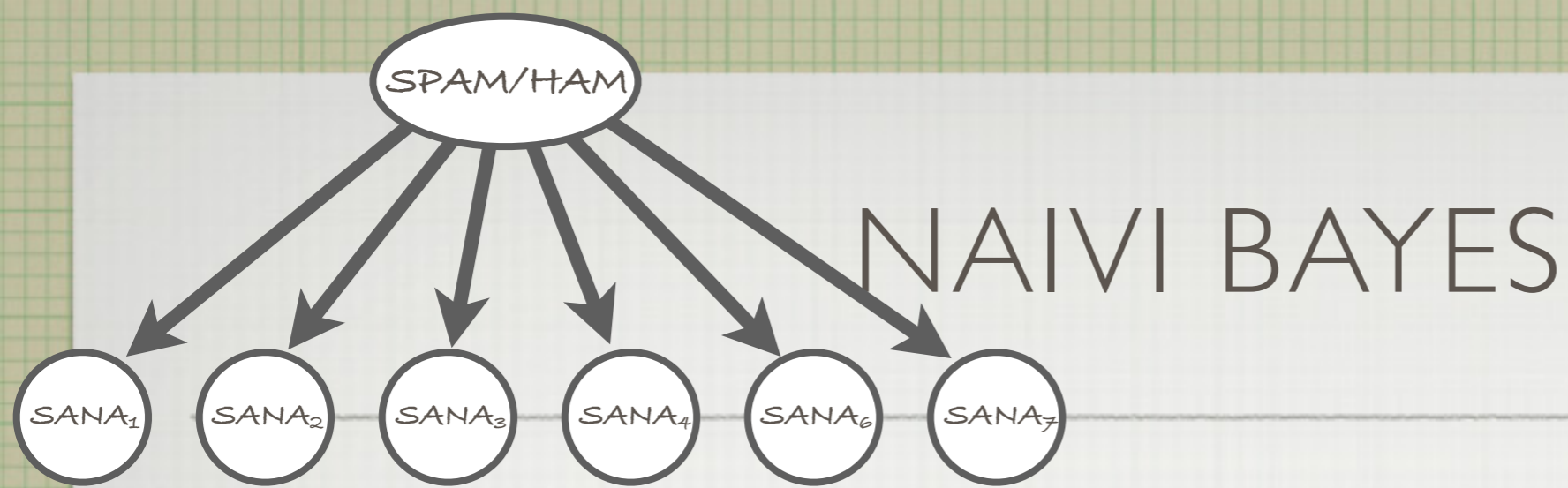
FROM: "MARGARETTA NITA" <MARGUERITeseBRINA@WMLE.COM>
SUBJECT: **SPECIAL OFFER : VIAGRA ON SALE AT \$1.38 !!!**
X-BOGOSITY: YES, TESTS=BOGOFILTER, SPAMICITY=0.99993752,
VERSION=2011-08-29
DATE: MON, 26 SEP 2011 21:52:26 +0300
X-CLASSIFICATION: **JUNK - AD HOC SPAM DETECTED (CODE = 73)**

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](http://rxpharmacyCVS.ru)



ROSKAPOSTISUODATIN:

SPAMICITY(Viesti, P):

Odds = ProbSpam / ProbHam

for each Sana in Viesti

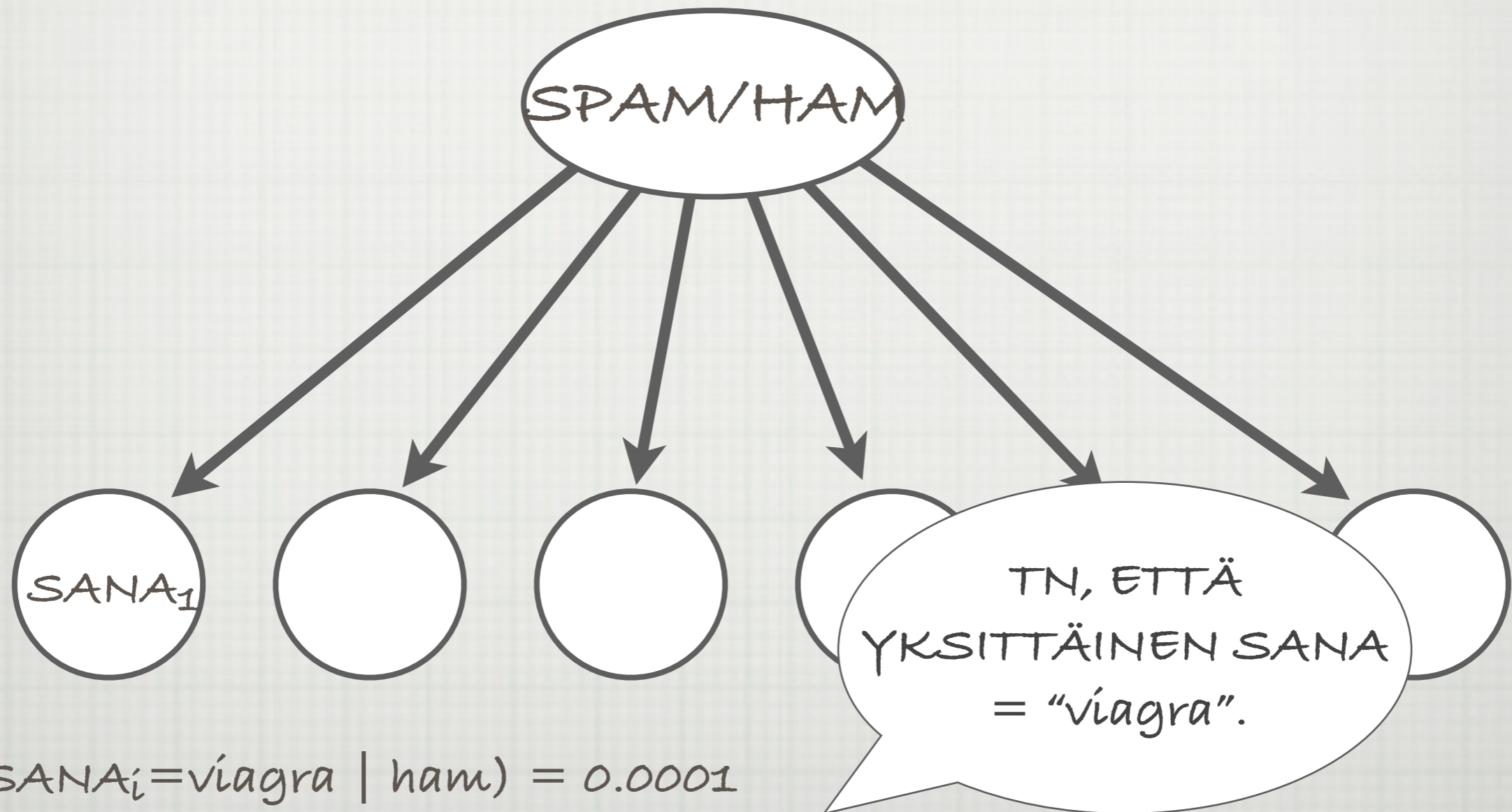
Odds = Odds * Prob_Spam(Sana) / Prob_Ham(Sana)

return(Odds)

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

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

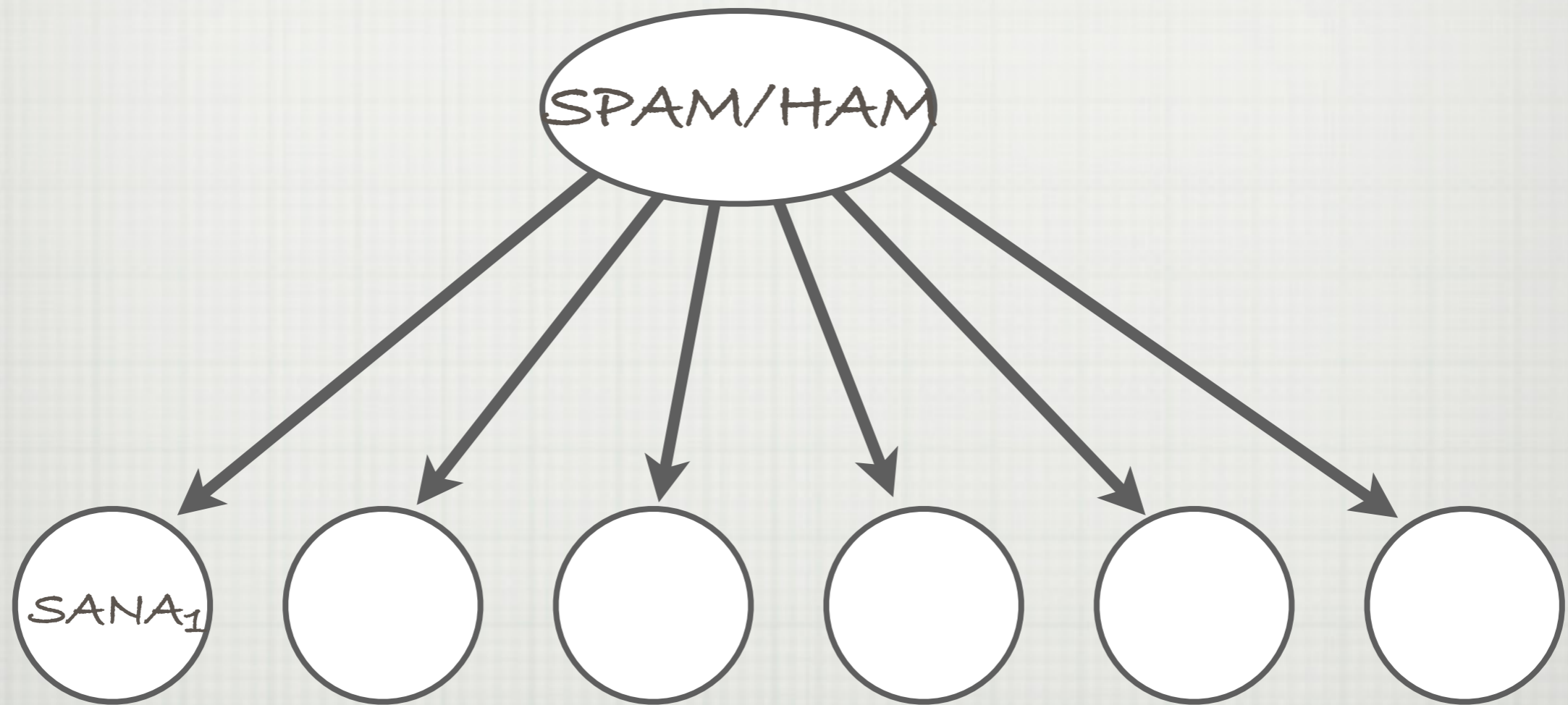
NAIVI BAYES



$$P(SANA_i = \text{viagra} \mid \text{ham}) = 0.0001$$

$$P(SANA_i = \text{viagra} \mid \text{spam}) = 0.002$$

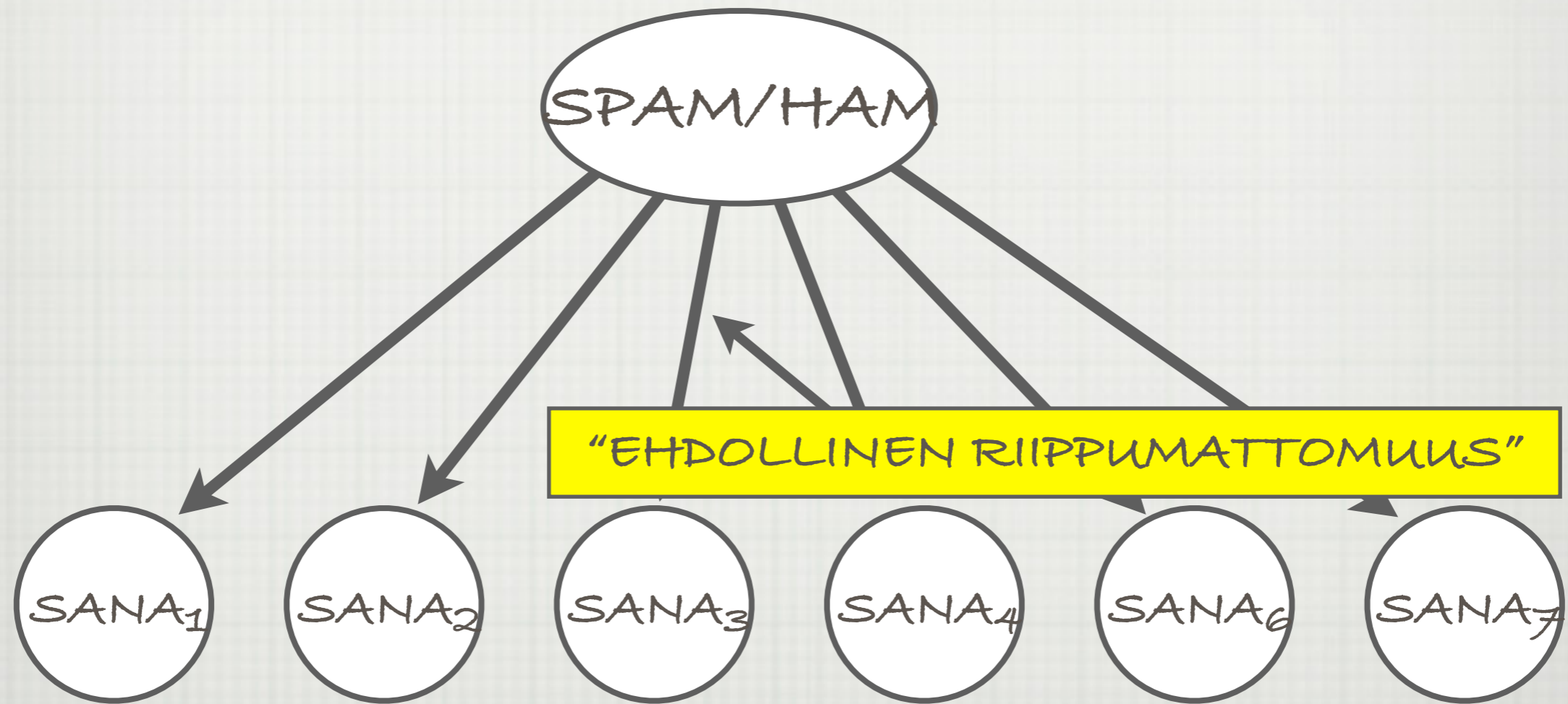
NAIVI BAYES



$$P(SANA_i = \$ \mid \text{ham}) = 0.0002$$

$$P(SANA_i = \$ \mid \text{spam}) = 0.005$$

NAIVI BAYES



NAIVI BAYES

MUUTTUJAT:

1. LUOKKA: spam/ham
2. $SANA_1$
3. $SANA_2$
4. ...

JAKAUMAT:

$$P(\text{LUOKKA}=\text{spam}) = 0.5$$

$$P(SANA_i=\text{viagra} \mid \text{spam}) = 0.002$$

$$P(SANA_i=\text{viagra} \mid \text{ham}) = 0.0001$$

$$P(SANA_i=\$ \mid \text{spam}) = 0.005$$

$$P(SANA_i=\$ \mid \text{ham}) = 0.0002$$

$$P(SANA_i=\text{is} \mid \text{spam}) = 0.002$$

$$P(SANA_i=\text{is} \mid \text{ham}) = 0.002$$

$$P(SANA_i=\text{algorithm} \mid \text{spam}) = 0.0001$$

$$P(SANA_i=\text{algorithm} \mid \text{ham}) = 0.002$$

JNE...



NAIVI BAYES

PÄÄTTELY:

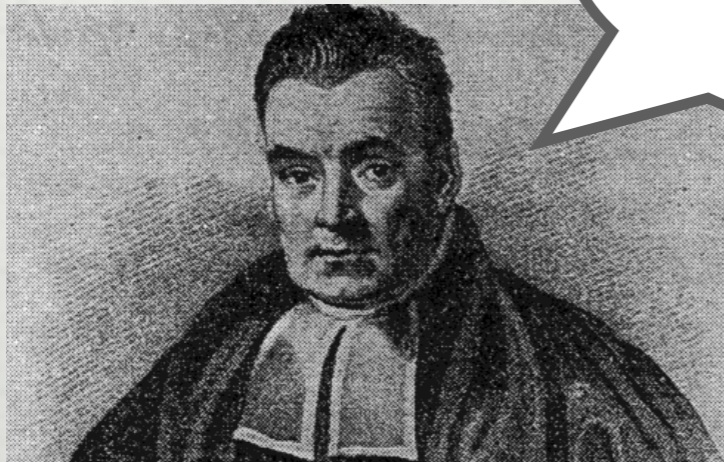
1. $P(\text{spam}) = 0.5$

$P(\text{spam}) P(\text{SANA}_1 = \text{viagra} | \text{spam})$

2. $P(\text{spam} | \text{SANA}_1 = \text{viagra}) = \frac{\text{-----}}{\text{-----}}$

$P(\text{SANA}_1 = \text{viagra})$

BAYESIN
KAAVA!



NAIVI BAYES

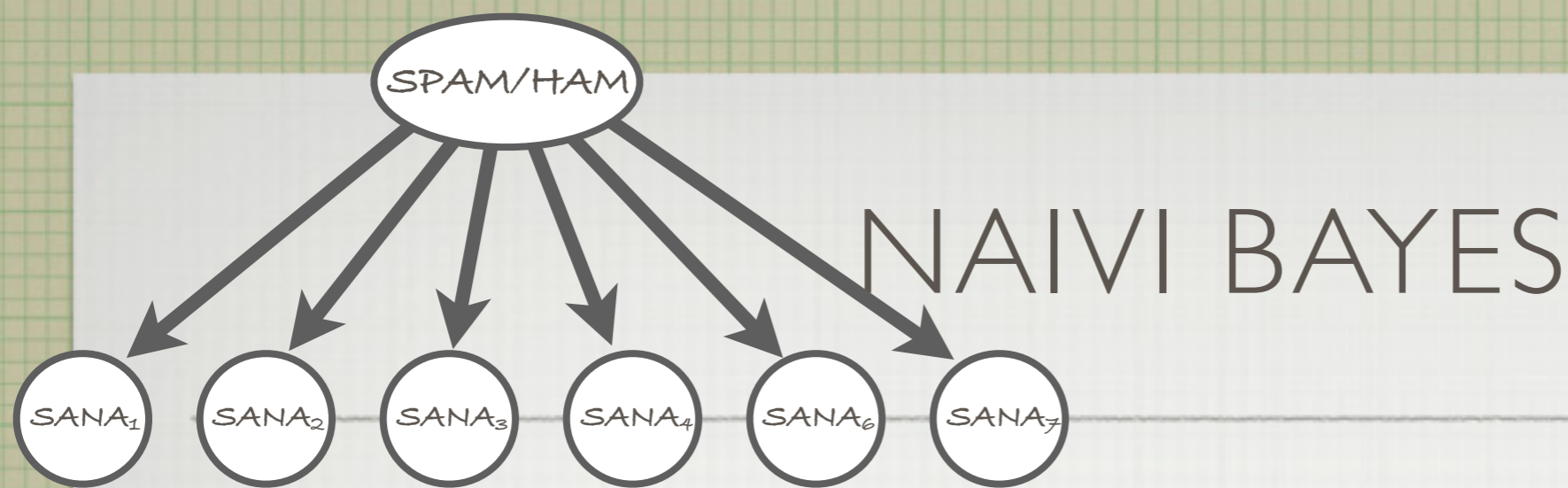
PÄÄTTELY:

1. $P(\text{spam}) = 0.5$

$$P(\text{spam}) P(\text{SANA}_1 = \text{víagra} | \text{spam})$$

2. $P(\text{spam} | \text{SANA}_1 = \text{víagra}) = \frac{\text{-----}}{P(\text{SANA}_1 = \text{víagra})}$

$$P(\text{SANA}_1 = \text{víagra}) = P(\text{spam}) P(\text{SANA}_1 = \text{víagra} | \text{spam}) \\ + P(\text{ham}) P(\text{SANA}_1 = \text{víagra} | \text{ham})$$



YHTEENVETO TOISTAISEKSI:

* TARVITAAN:

- "PRIORIJAKAUMA" $P(\text{spam}) = 0.___$

- "LUOKKAEHDOLLISET" JAKAUMAT

$P(\text{SANA}_i = \text{viagra} \mid \text{spam}) = 0.___$

$P(\text{SANA}_i = \text{viagra} \mid \text{ham}) = 0.___$

$P(\text{SANA}_i = \text{is} \mid \text{spam}) = 0.___$

$P(\text{SANA}_i = \text{is} \mid \text{ham}) = 0.___$

$P(\text{SANA}_i = \text{algorithm} \mid \text{spam}) = 0.___$

$P(\text{SANA}_i = \text{algorithm} \mid \text{ham}) = 0.___$

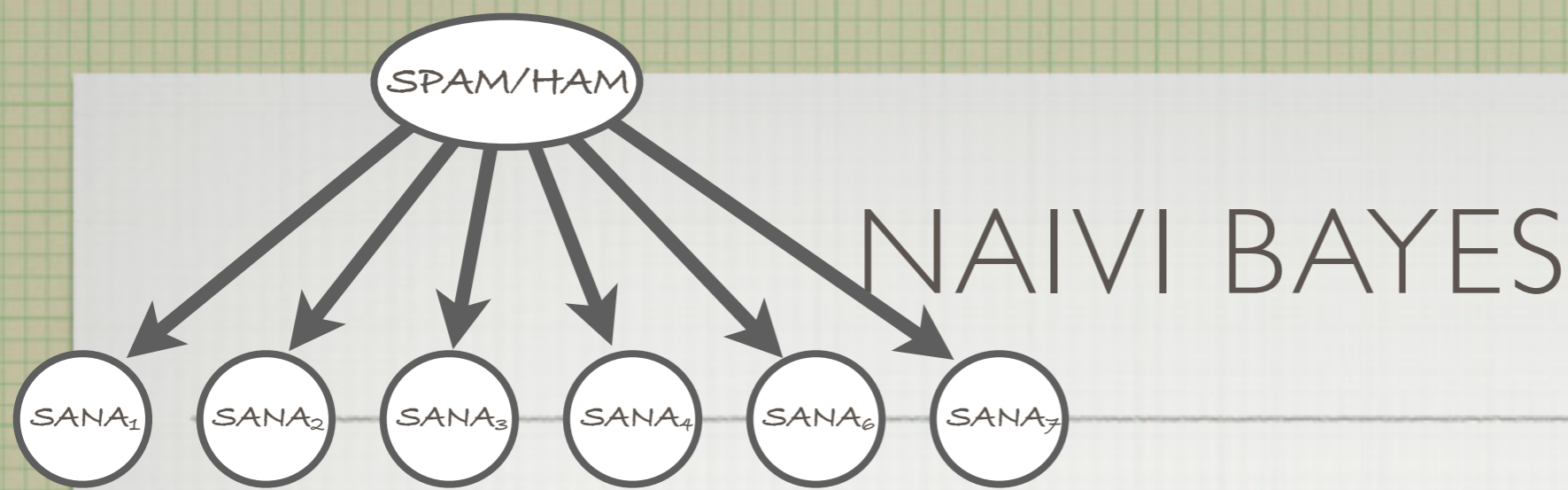
* OLETETAAN ETTÄ $P(\text{SANA}_i \mid \text{SANA}_j, \text{spam}) = P(\text{SANA}_i \mid \text{spam})$
(EHDOLLINEN RIIPPUMATTOMUUS)

* OLENNAINEN ON OSAMÄÄRÄ

$P(\text{SANA}_i = \text{viagra} \mid \text{spam})$

(OTETAAN NÄIDEN TULO)

 $P(\text{SANA}_i = \text{viagra} \mid \text{ham})$



PSEUDOKOODINA:

SPAMICITY(Viesti, P):

Odds = ProbSpam / ProbHam

for each Sana in Viesti

Odds = Odds * Prob_Spam(Sana) / Prob_Ham(Sana)

return(Odds)

PÄÄTTELY:

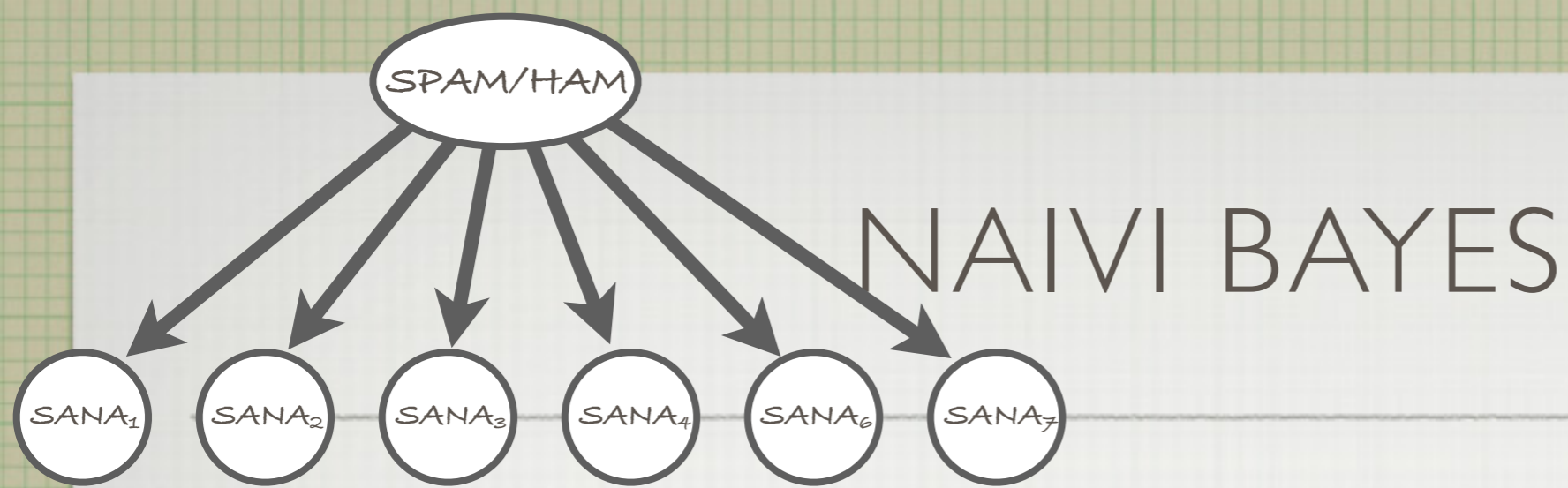
$P(\text{spam} | \text{EVIDENSSI}) / P(\text{ham} | \text{EVIDENSSI})$

$= P(\text{spam}) / P(\text{ham})$

$P(\text{SANA}_1 = \text{viagra} | \text{spam}) / P(\text{SANA}_1 = \text{viagra} | \text{ham})$

$P(\text{SANA}_2 = \text{is} | \text{spam}) / P(\text{SANA}_2 = \text{is} | \text{ham})$

$P(\text{SANA}_3 = \text{algorithm} | \text{spam}) / P(\text{SANA}_3 = \text{algorithm} | \text{ham})$



PSEUDOKOODINA:

SPAMICITY(Viesti, P):

Odds = ProbSpam / ProbHam

for each Sana in Viesti

Odds = Odds * Prob_Spam(Sana) / Prob_Ham(Sana)

return(Odds)

PÄÄTTELY:

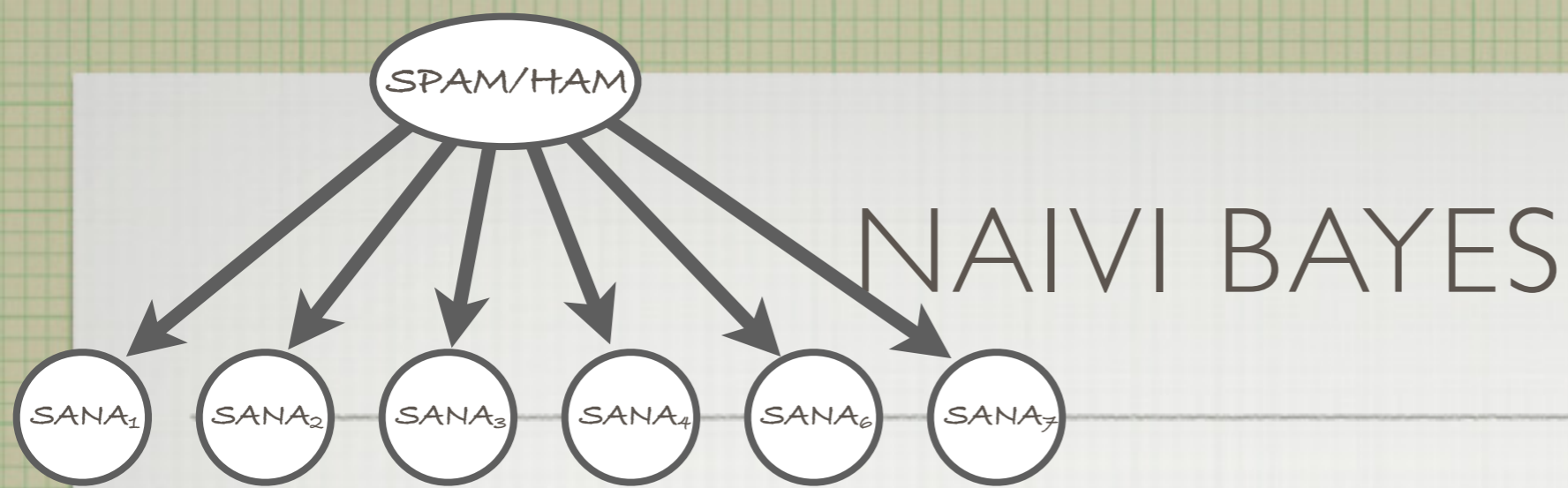
$$P(\text{spam} | \text{EVIDENSSI}) / P(\text{ham} | \text{EVIDENSSI})$$

$$= P(\text{spam}) / P(\text{ham})$$

$$P(\text{SANA}_1 = \text{viagra} | \text{spam}) / P(\text{SANA}_1 = \text{viagra} | \text{ham})$$

$$P(\text{SANA}_2 = \text{is} | \text{spam}) / P(\text{SANA}_2 = \text{is} | \text{ham})$$

$$P(\text{SANA}_3 = \text{algorithm} | \text{spam}) / P(\text{SANA}_3 = \text{algorithm} | \text{ham})$$



PSEUDOKOODINA:

SPAMICITY(Viesti, P):

$$\text{Odds} = \text{ProbSpam} / \text{ProbHam}$$

for each Sana in Viesti

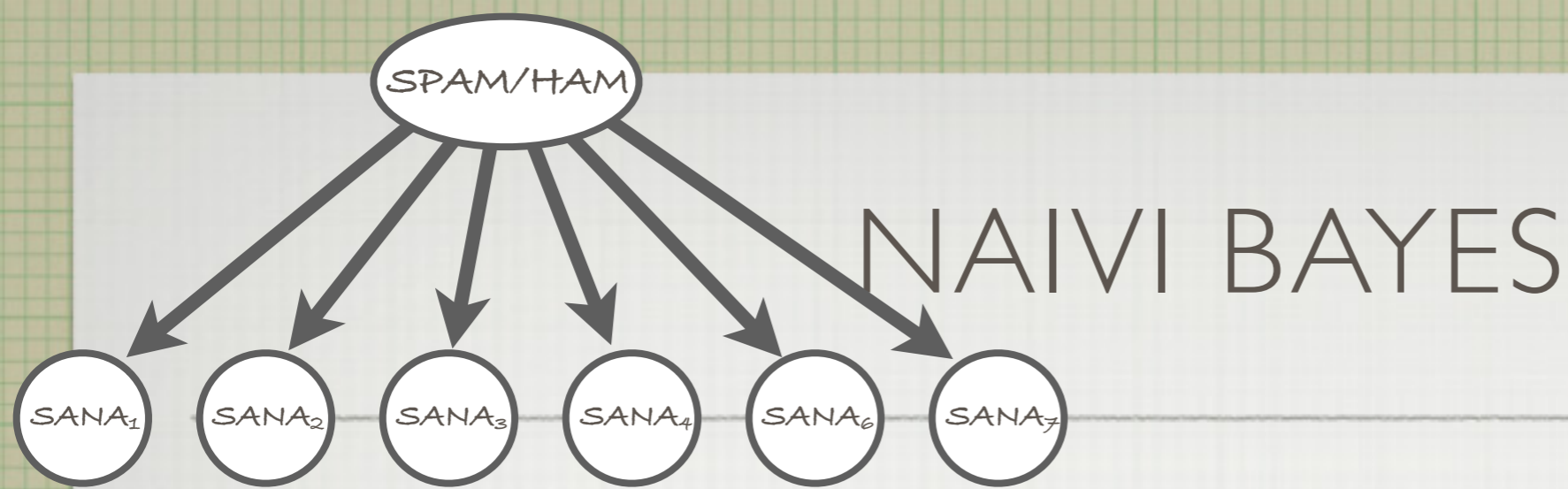
$$\text{Odds} = \text{Odds} * \text{Prob_Spam}(\text{Sana}) / \text{Prob_Ham}(\text{Sana})$$

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



PSEUDOKOODINA:

SPAMICITY(Viesti, P):

Odds = ProbSpam / ProbHam

for each Sana in Viesti

Odds = Odds * Prob_Spam(Sana) / Prob_Ham(Sana)

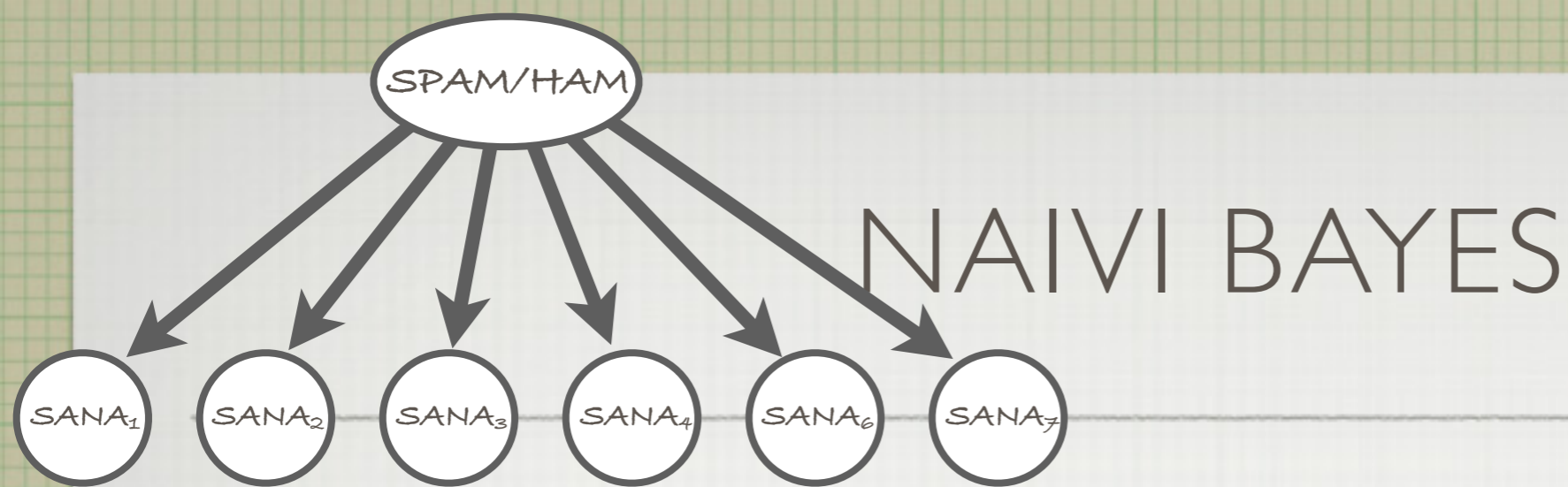
return(Odds)

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

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

MUUTEN, EN TIEDÄ

Epäsymmetrinen kustannusfunktio:
 ! Asiallisen viestin luokittelu spamiksi
 pahempi virhe kuin toisin päin



PSEUDOKOODINA:

$$\text{LOG}(A * B) = \text{LOG}(A) + \text{LOG}(B)$$

SPAMICITY(Viesti, P):

$\text{logOdds} = \text{log}(\text{ProbSpam} / \text{ProbHam})$

for each Sana in Viesti

$\text{logOdds} = \text{logOdds} + \text{log}(\text{Prob_Spam}(\text{Sana}) / \text{Prob_Ham}(\text{Sana}))$

return($\text{exp}(\text{logOdds})$)

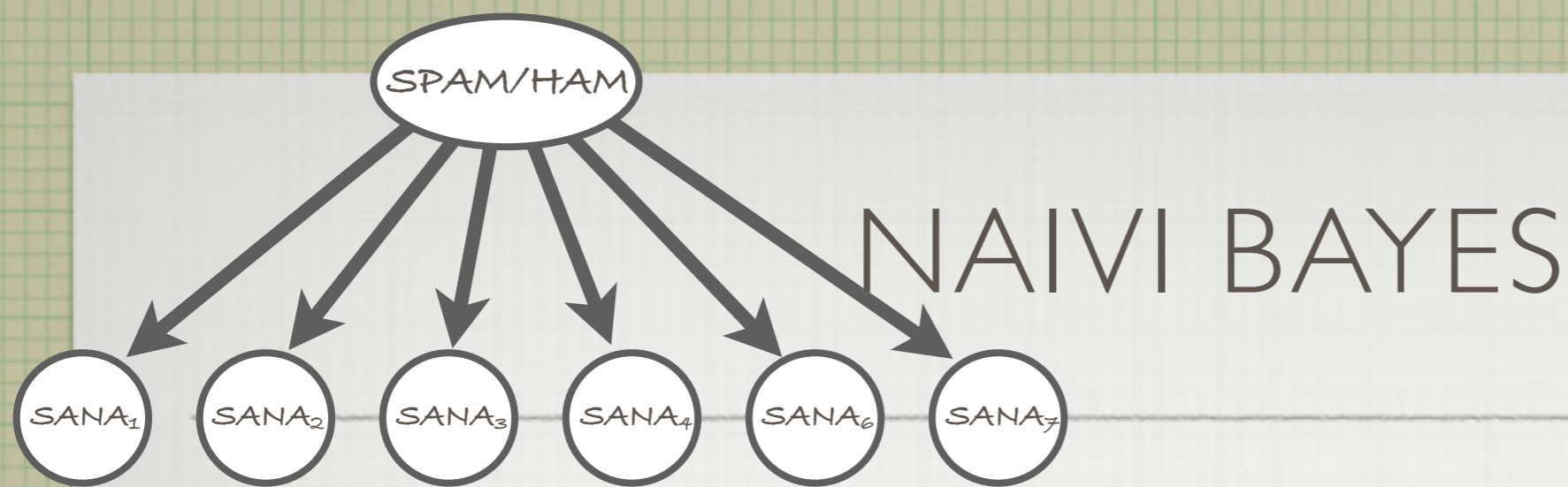
KÄYTÄNNÖN ONGELMA: ALI-JA YLIVUODOT

Odds ARVOSTA TULEE HELPOSTI LIIAN PIENI (LÄHELLÄ NOLLAA)

TAI LIIAN SUURI.

RATKAISU:

KÄYTÄ $\text{log}(\text{Odds})$



PARAMETRIEN OPPIMISESTA:

* VAIKEA KEKSIÄ PÄÄSTÄ EHDOLLISIA TN:IA.

* HUONOT ARVOT HUONONTAVAT FILTTERIN TOIMINTAA

* PAREMPI RATKAISU:

1. KERÄÄ ISO KASA SPAM-VIESTEJÄ

2. KERÄÄ ISO KASA HAM-VIESTEJÄ

3. ARVIOI $P(SANA_i = _ \mid SPAM) = 0._ _$ JA

$P(SANA_i = _ \mid HAM) = 0._ _$ LASKEMALLA DATASTA

(VRT. LASKUHARJ.)

* VAROTTAVA NOLLATODENNÄKÖISYYKSIÄ!

(JOS KASA EI TARPEEKSI ISO, JOTKUT SANAT EIVÄT VAIN SATU ESIINTYMÄÄN SIINÄ.)

ESIM.

The screenshot shows an email client window titled 'Junk_E-mail'. The interface includes a menu bar with 'Get Mail', 'Write', 'Address Book', and 'Tag'. A search bar at the top right says 'Search all messages... <⌘K>'. Below the menu bar, there's a 'Quick Filter' section and another search bar 'Filter these messages... <⌘F>'. On the left, a folder pane shows a tree structure under 'cs-mail', with 'Junk_E-mail' selected. The main pane displays a list of emails with columns for 'Subject', 'From', and 'Date'. The selected email is from 'Creola Astrid' with the subject 'Viagra 100mg x 60 Pills \$125, Free Pills & Reorder Discount, Top Selling 100% Quality & Satisfaction guaranteed!'. Below the list, the email's details are shown, including the sender's name and email address, and a 'Junk Mail' indicator with a 'Not Junk' button. The email body contains the text 'Best Buy Viagra Generic Online', the same subject line, and a link to 'http://tabletpillsapr.ru'. At the bottom right, it says 'Unread: 0 Total: 29'.

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...	Vanessa 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

from Creola Astrid <alvertaarlinda@onesource.com>
subject Viagra 100mg x 60 Pills \$125, Free Pills & Reorder Discount, Top Selling 100% Quality & Satisfaction guaranteed!
to sampo.sammalisto@cs.helsinki.fi

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.

spam

1	MONEY	0.04	%
5	VIAGRA	0.21	%
10	IS	0.42	%
19	REPLICA	0.80	%
20	EMAIL	0.84	%
20	YOU	0.84	%
21	DATABASE	0.88	%
25	EMAILS	1.05	%
26	OF	1.09	%
31	TO	1.30	%
43	AND	1.80	%
48	THE	2.01	%
TOTAL		2386	

ham

21	ALGORITHM	0.01	%
62	MONEY	0.02	%
2199	FOR	0.78	%
2492	THAT	0.88	%
2990	YOU	1.05	%
3141	IN	1.11	%
3160	I	1.11	%
3218	AND	1.13	%
3283	IS	1.16	%
3472	OF	1.22	%
3874	A	1.37	%
5442	TO	1.92	%
9196	THE	3.24	%
TOTAL		283736	

ESIM.

spam

1	MONEY	0.04 %
5	VIAGRA	0.21 %
10	IS	0.42 %
19	REPLICA	0.80 %
20	EMAIL	0.84 %
20	YOU	0.84 %
21	DATABASE	0.88 %
25	EMAILS	1.05 %
26	OF	1.09 %
31	TO	1.30 %
43	AND	1.80 %
48	THE	2.01 %
TOTAL		2386

ham

21	ALGORITHM	0.01 %
62	MONEY	0.02 %
2199	FOR	0.78 %
2492	THAT	0.88 %
2990	YOU	1.05 %
3141	IN	1.11 %
3160	I	1.11 %
3218	AND	1.13 %
3283	IS	1.16 %
3472	OF	1.22 %
3874	A	1.37 %
5442	TO	1.92 %
9196	THE	3.24 %
TOTAL		283736

ESIM.

spam

ham

1 MONEY 0.04 %

5 VIAGRA 0.21 %

10 I

19 R

20 E

20 Y

21 DATABASE 0.00 %

25 EMAILS 1.05 %

26 OF 1.09 %

31 TO 1.30 %

43 AND 1.80 %

48 THE 2.01 %

TOTAL 2386

21 ALGORITHM 0.01 %

62 MONEY 0.02 %

3218 AND 1.15 %

3283 IS 1.16 %

3472 OF 1.22 %

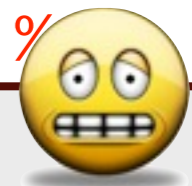
3874 A 1.37 %

5442 TO 1.92 %

9196 THE 3.24 %

TOTAL 283736

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



ESIM.

spam

ham

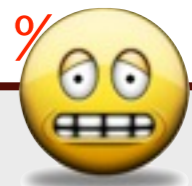
1 MONEY 0.04 %

5 VIAGRA 0.21 %

21 ALGORITHM 0.01 %

62 MONEY 0.02 %

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



21 DATABASE 0.00 % 3210 AND 1.15 %

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



TOTAL

ESIM.

$P(\text{SANA}_i = \text{'ALGORITHM'} \mid \text{SPAM}) = 0?$
VOIJOHTAA TILANTEESEEN, JOSSA
OSAMÄÄRÄ 0/0

ham

5	VIAGRA	0.21 %
10	IS	0.42 %
19	REPLICA	0.80 %
20	EMAIL	0.84 %
20	YOU	0.84 %
21	DATABASE	0.88 %
25	EMAILS	1.05 %
26	OF	1.09 %
31	TO	1.30 %
43	AND	1.80 %
48	THE	2.01 %
TOTAL		2386

21	ALGORITHM	0.01 %
62	MONEY	0.02 %
2199	FOR	0.78 %
2492	THAT	0.88 %
2990	YOU	1.05 %
3141	IN	1.11 %
3160	I	1.11 %
3218	AND	1.13 %
3283	IS	1.16 %
3472	OF	1.22 %
3874	A	1.37 %
5442	TO	1.92 %
9196	THE	3.24 %
TOTAL		283736

YHTEENVETO

YHTEENVETO NAIVI BAYES-SPAMFILTTERISTÄ:

* TARVITAAN:

- "PRIORIJAKAUMA" $P(\text{SPAM}) = 0.___$

- "LUOKKA-EHDOLLISET" JAKAUMAT

$$P(\text{SANA}_i = \text{VIAGRA} | \text{SPAM}) = 0.___$$

$$P(\text{SANA}_i = \text{VIAGRA} | \neg \text{SPAM}) = 0.___$$

$$P(\text{SANA}_i = \text{IS} | \text{SPAM}) = 0.___$$

$$P(\text{SANA}_i = \text{IS} | \neg \text{SPAM}) = 0.___$$

$$P(\text{SANA}_i = \text{ALGORITHM} | \text{SPAM}) = 0.___$$

$$P(\text{SANA}_i = \text{ALG.} | \neg \text{SPAM}) = 0.___$$

* HUOM! **EI** OLE KYSEESSÄ $P(\text{"SANA VIAGRA ESIINTYY"} | \text{SPAM})$, JNE.

(SEKIN OLISI MAHDOLLISTA, MUTTA VIESTIEN PITUUDET VAIKUTTAISIVAT TULOKSEEN.)

* OLETETAAN ETTÄ $P(\text{SANA}_i | \text{SANA}_j, \text{SPAM}) = P(\text{SANA}_i | \text{SPAM})$

(EHDOLLINEN RIIPPUMATTOMUUS)

YHTEENVETO

(JATKOA...):

* OLENNAISTA ON OSAMÄÄRÄ

$$P(\text{SANA}_i = \text{VIAGRA} | \text{SPAM})$$

$$P(\text{SANA}_i = \text{VIAGRA} | \neg \text{SPAM})$$

* KUSTANNUSFUNKTIO EPÄSYMMETRINEN: PAREMPI SÄÄSTÄÄ MUUTAMA SPAM KUIN HUKATA OIKEA VIESTI (HAM)

* JAKAUMAT PARAS ESTIMOIDA DATASTA

YHTEENVETO

(JATKOA...):

* OLENNAISTA ON OSAMÄÄRÄ

$$P(\text{SANA}_i = \text{VIAGRA} | \text{SPAM})$$

$$P(\text{SANA}_i = \text{VIAGRA} | \neg \text{SPAM})$$

* KUSTANNUSFUNKTIO EPÄSYMMETRINEN: PAREMPI SÄÄSTÄÄ MUUTAMA SPAM KUIN HUKATA OIKEA VIESTI (HAM)

* JAKAUMAT PARAS ESTIMOIDA DATASTA

* NOLLATODENNÄKÖISYYKSILLE TEHTÄVÄJOTAIN