Boolean retrieval is a specific case of DBMS.

go thru algo of skip pointer.

Sample question

① Stemming increases recall in Boolean retrieval model. But if query has not then recall is decreased.

[in other cases stemming never decreases recall].

② Relevance feedback improves precision & recall. So same is case with pseudo. only problem is if top k docs are non relevant then they fails.

③ user is reluctant. query processing takes lot of time. it is hard to understand why new set of docs are retrieved

④ vectors, t-dimensional vector, tf-idf, cosine similarity.

⑤ Local & Global. → Thesaurus.
→ concept cloud.
↙ relevance
↘ Pseudo relevance

⑥ adv:— more weight to rare terms.
disadv:— short text

Decision tree

Age	competition	type	profit
old	Yes	S|w	Down.
old	No	S|w	Down.
old	No	h|w	Down
mid	Yes	S|w	Down
mid	Yes	H|w	Down
mid	No	H|w	up
mid	No	S|w	up.
new	Yes	S|w	up
new	No	H|w	up
new	Yes	S|w	up.

$P = 5$ $\quad N = 5$

~~Entropy = $-\frac{P}{P+N}\left(\log_2 \frac{P}{P+N}\right)$~~

$$\text{Entropy} = -\frac{P}{P+N}\log_2\left(\frac{P}{P+N}\right) - \frac{N}{P+N}\log_2\left(\frac{N}{P+N}\right)$$

$$= -\frac{5}{10}\log_2 1/2 - \frac{5}{10}\log_2 1/2$$

$$= 1/2 \log 2 + 1/2 \log 2 = \log_2 2 = 1$$

Entropy (class) − Entropy (attribute) = Gain.

Information Gain :-

$$I(P_0, N_0) = \frac{\Sigma\, P_i + N_i}{P+N} \times I(P_i, N_i)$$

I (Age)

Age

	P_i	N_i	$I(P_i, N_i)$
old	0	3	0
mid	4	2	1
new.	3	0	0

$$I(\text{Age}) = \sum \frac{P_i + N_i}{P+N} I(P_i, N_i)$$

$$= \frac{3}{10} \times 0 + \frac{4}{10} \times 1 + \frac{3}{10} \times 0$$

$$= 0.4$$

$$\text{Gain} = 1 - 0.4$$

$$= 0.6.$$

Competition

	P_i	N_i	$I(P_i, N_i)$
Yes	1	3	0.81127
No	4	2	0.918295

$$I(P_i, N_i) = -\frac{1}{4}\log_2 1/4 - \frac{3}{4}\log_2 3/4$$

$$= \frac{1}{4} \times 2 + \frac{3}{4} \times \log 4/3 = 0.81127$$

$$I(P_i, N_i) = -\frac{4}{6}\log 4/6 - \frac{2}{6}\log 2/6$$

$$= 2/3 \log 3/2 + 1/3 \log 3$$

$$= 0.918295$$

$$\text{Entropy (comp)} = \frac{1+3}{10} \times 0.81127 + \frac{4+2}{10} \times 0.918295$$

$$= 0.875484$$

Gain = 1 − 0.875484 .

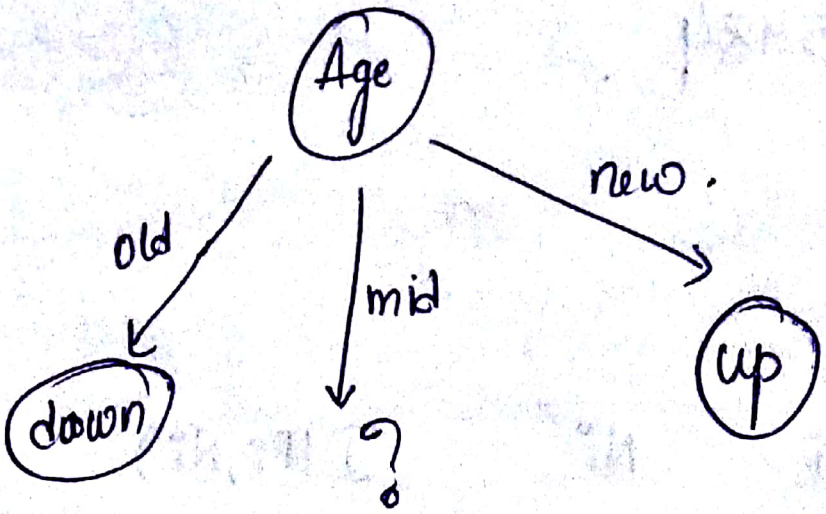
type

	Pi	Ni	I(Pi, Ni)
s/w	3	3	1
h/w	2	2	1

Entropy (type) = $\frac{6}{10} \times 1 + \frac{4}{10} \times 1$

= 1

Gain = 1 − 1 = 0.



Competition ~~type~~.

	Pi	Ni	I(Pi, Ni)
yes	0	2	0
No.	2	0	0.

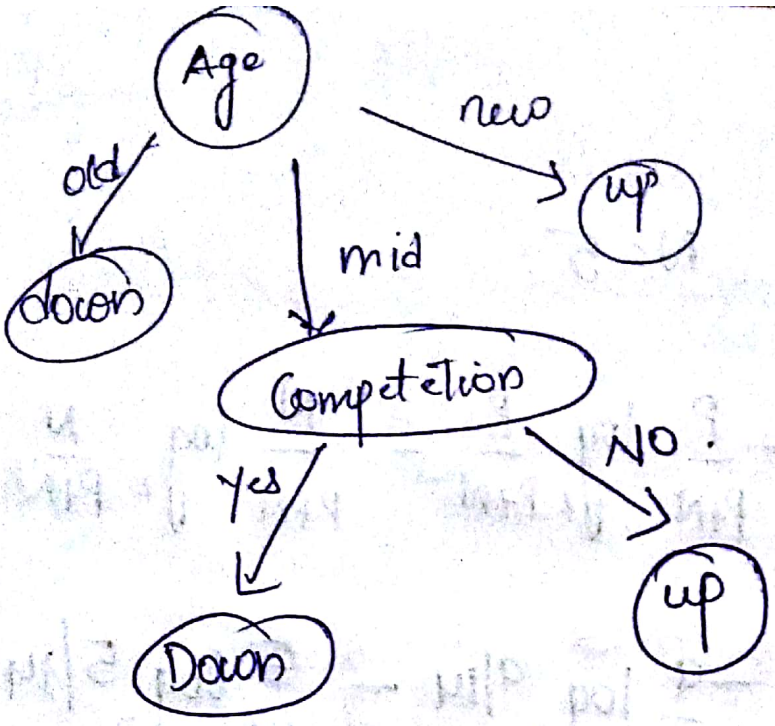
Gain = 1 - 0 = 1

type

	Pi	Ni	I(Pi, Ni)
S/w	1	1	1
H/w	1	1	1

Entropy = 1

Gain = 0.



Example - 2

$P = 9, \quad N = 5.$

$$\text{Entropy}_{\text{class}} = -\frac{P}{P+N}\log_2\frac{P}{P+N} - \frac{N}{P+N}\log_2\frac{N}{P+N}$$

$$= -\frac{9}{14}\log_2 9/14 - \frac{5}{14}\log_2 5/14$$

$$\text{Entropy class.} = 0.4 + 0.53 = 0.93.$$

I (outlook)

	P	N.	I (Pi Ni)
Sunny	2	3	0.96
overcast	4	0	0
rain	3	2	0.96.

$= -2/5 \log 2/5 - 3/5 \log 3/5$

$= 0.52 + 0.44 = 0.96.$

$I(\text{outlook}) = \sum \frac{P_i + N_i}{P+N} I(P_i, N_i)$

$= \frac{5}{14} \times 0.96 + \frac{5}{14} \times 0.96.$

$= 5/7 \times 0.96. = 0.68.$

Gain $= 0.93 - 0.68 = 0.25$

I(Temp)

	P	N	$I(P_i, N_i)$
Hot	2	2	1
Cool	3	1	0.81
Mild	4	2	0.9

$-\frac{3}{4}\log 3/4 - 1/4 \log 1/4$

$0.31 + 0.5 = 0.81$

$-4/6 \log 4/6 - 2/6 \log 2/6$

$0.38 + 0.52 = 0.9$

$I(\text{temp}) = \frac{4}{14} \times 1 + \frac{4}{14} \times 0.81 + \frac{6}{14} \times 0.9$

$= \frac{4 + 3.24 + 5.4}{14} = 0.902$

Gain = $0.93 - 0.902 = 0.028$.

I (Humidity)

	P	N.	I (Pi, Ni)
high	3	4	
Normal	6	1	

$-4/7 \log 4/7 - 3/7 \log 3/7 =$

$0.46 + 0.52 = 0.98.$

$-6/7 \log 6/7 - 1/7 \log 1/7 =$

$0.19 + 0.40 = 0.59$

$I(\text{Humidity}) = \frac{7}{14} \times 0.98 + \frac{7}{14} \times 0.59$

$= 0.785$

$\text{Gain} = 0.93 - 0.785 = 0.145$

$I(\text{Wind})$

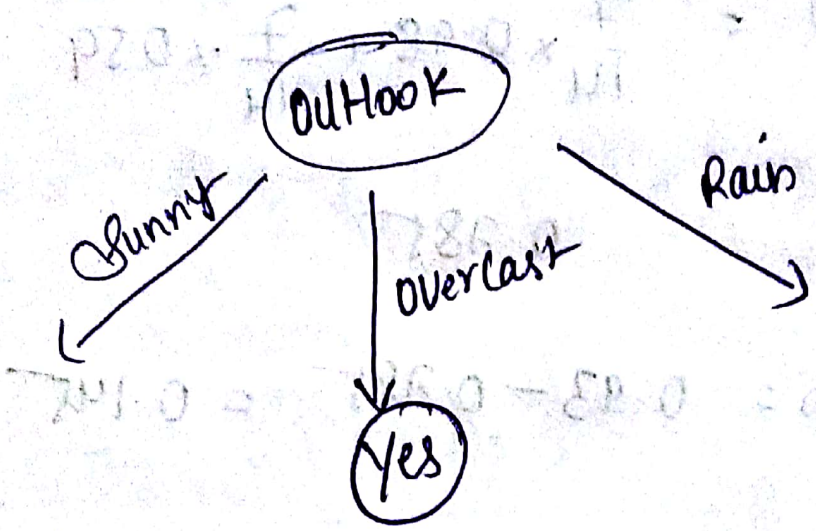
	P	N	$I(P_i, N_i)$
Strong	6	2	0.81
Weak	3	3	1

$-6/8 \log 6/8 - 2/8 \log 2/8$

$$I(\text{Wind}) = \frac{8}{14} \times 0.81 + \frac{6}{14} \times 1$$

$$= 0.89$$

$$\text{Gain} = 0.93 - 0.89 = 0.04$$



Sunny

I (temp)

	P	N	I(Pi, Ni)
Hot	0	2	0
cold	1	0	0
Mild	1	1	1

I(temp) = 2/5 ×1 = 0.4.

Gain = 0.93 − 0.4 = 0.53.

I(Humidity)

	P	N	I(Pi, Ni)
High	0	3	0
Normal.	2	0	0

Gain = 0.93.

I(wind)

	P	N	I(Pi, Ni)
Weak	2	1	0.9
Strong	1	1	1

I(Pi, Ni) = 3/5 ×0.9 + 2/5 ×1 = 0.94

Naïve Bayes classifier :—

$$P(h|D) = \frac{P(D|h)\,P(h)}{P(D)}.$$

There is no dependency between attributes i.e independent attributes.

Q) Does patient cancer or not? A patient takes a lab test & the result comes bad positive. The result comes correct positive in only .98% cases. in which the disease is actually present and a correct negative result in 97% only. in which the disease is not present. Furthermore 0.008% of the entire Population have this cancer. ~~Bayes~~ using Bayes theorem, Say whether patient has cancer or not.

98	2
3	97

P(cancer) = 0.008

P(¬cancer) = 0.992

P(+/cancer) = 0.98.

P(−/cancer) = 0.02

P(+/¬cancer) = 0.03

P(−/¬cancer) = 0.97

P(+/cancer) · P(cancer) = 0.98 × 0.008 = 0.0078

P(+/¬cancer) P(¬cancer) = 0.03 × 0.992

= 0.0298

Patient does not Cancer.

② P(Play | yes) = 9/14 , P(play | No) = 5/14.

outlook

P'(Sunny | P) = 2/9	P(sunny | N) = 3/5
P(overcast | P) = 4/9	P(overcast | P) = 0
p(rain | P) = 3/9	P(rain | N) = 2/5

temp

p(Hot | P) = 2/9	P(Hot | P) = 2/5
p(mild | P) = 4/9	P(mild | P) = 2/5
p(cool | P) = 3/9	P(cool | P) = 1/5

Humidity

P(High | P) = 3/9	P(High | N) = 4/5
P(Normal | P) = 6/9	P(Normal | N) = 1/5

wind

P(Strong | P) = 6/9	P(Strong | N) = 2/5
P(weak | P) = 3/9	P(weak | N) = 3/5

X: <rain, hot, high, ~~false~~>

~~P(play|X) = P(~~

$P(\text{~~play~~} X | \text{~~play~~}) = P(\text{rain}|P) \times P(\text{hot}|P) \times P(\text{high}|P) \times P(\text{strong}|P)$

$= 3/9 \times 2/9 \times 3/9 \times 6/9$

$= 162/94$

$P(P|X) = \frac{P(X|P)P(P)}{}$

$= \frac{162}{9^4 9^3} \times 9/14 = \frac{1}{9} \times 1/7$

$= 1/63 = 0.015$

$P(X|N) = P(\text{rain}|N) \times P(\text{hot}|N) \times P(\text{high}|N) \times P(\text{strong}|N)$

$= 2/5 \times 2/5 \times 4/5 \times 2/5$

$P(N/x) = \frac{32}{5^4 5^3 1/4} \times \frac{5}{} = 0.018$.

Result :- NO.

Example- training set

docid		
1.	chinese beijing chinese	yes
2.	chinese chinese shanghai	yes
3.	chinese macao	yes
4.	Tokyo Japan chinese	NO.

test- 5 chinese chinese chinese TOkyo Japan.

P(chinese) P(yes) = 3/4 , P(NO) = 1/4

P(chinese / yes) = 5/8 , P(chinese / NO) = 1/3.

P(beijing / yes) = 1/8 , P(beijing / NO) = 0

P(tokyo / yes) = 0 , P(tokyo / NO) = 1/3.

P(~~tok~~ Japan | yes) = 0 P(Japan | No) = 1/3.

~~P(x/y~~

Laplacian correction :-

$$P(t|c) = \frac{T_{ct}+1}{\sum_{t \in V}(t) + B}$$ → distinguished words.

P(yes) = 3/4, P(No) = 1/4

P(Chinese | yes) = $\frac{5+1}{8+6}$ P(Chinese | No) = $\frac{1+1}{3+6} = 2/9$

= 6/14

P(tokyo | yes) = $\frac{0+1}{8+6} = \frac{1}{14}$, P(tokyo | No) = $\frac{1+1}{3+6} = 2/9$

P(Japan | yes) = $\frac{0+1}{8+6} = 1/14$, P(tokyo | No) = $\frac{1+1}{3+6} = \frac{2}{9}$

$P(X/yes) = (P(chinese/yes))^3 + P(tokyo/yes) \times P(Japan/yes)$

$= \frac{(6)^3}{145}$

$P(yes/X) = \frac{6^3}{145} \times \frac{3}{4} = \frac{648}{4 \times 145} = 0.0003.$

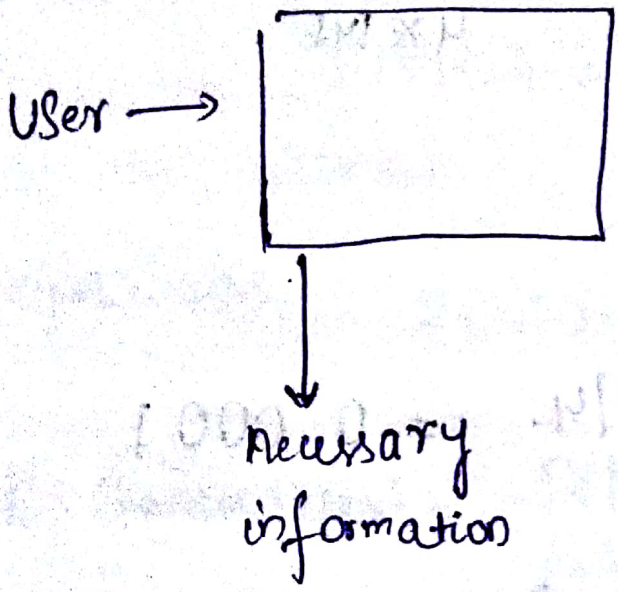
$P(X/No) = (2/9)^3 \times (2/9)^5$

$P(No/X) = (2/9)^5 \times 1/4. = 0.0001$

Result = yes

Recommender System :—

Information overload. in 1990



Group lens, Movie Lens. } → RS are used.

↓

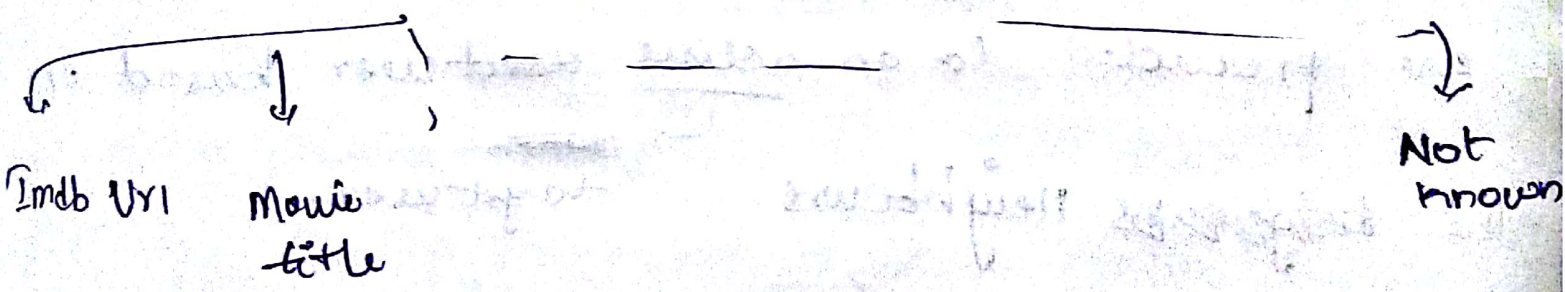
Bench mark dataset

① ~~usage~~ user's age
② user's occupation
③ user's rating on experienced movies. } by user.
④ movie's information.

IDs ↺ CWE

for building any ~~not~~ recommender system. we need data.

movie's info ⟹ movie lens.



19 Cells.

Recommendation Techniques

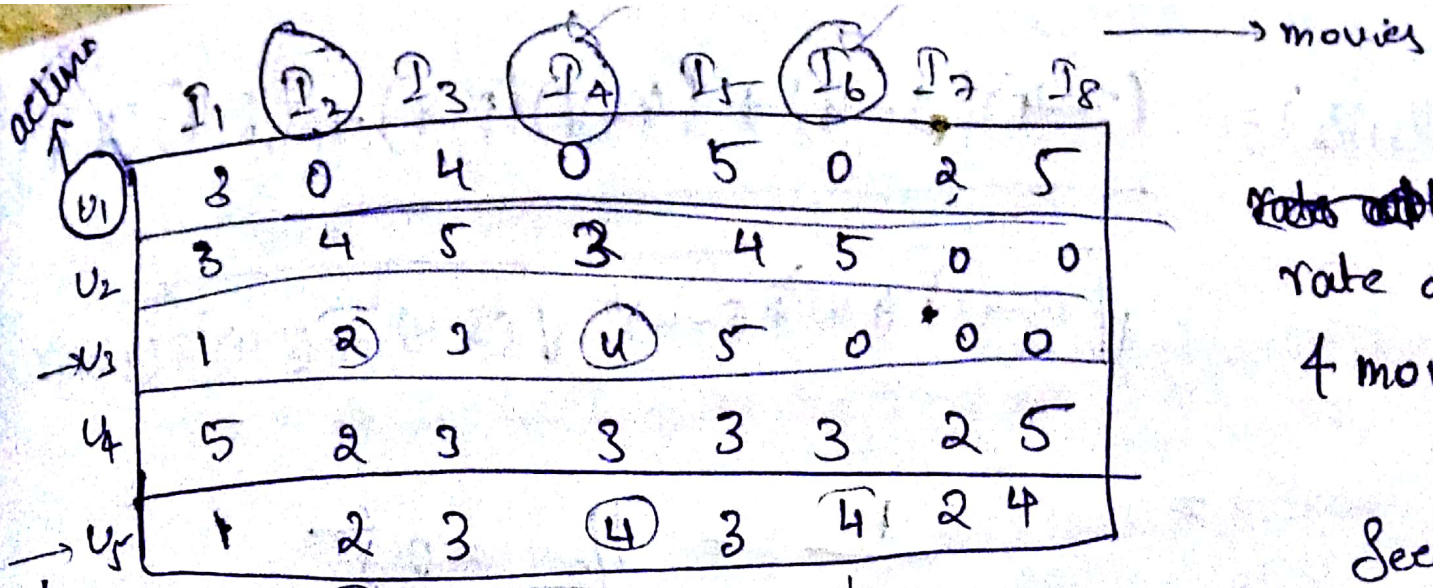
1. collaborative ~~Technique~~ filtering
2. content based filtering
3. Hybrid filtering.

Collaborative Technique :- Recommendations are provided to an active ~~user~~ user based on the ~~neighbours~~ neighbours. → ~~target~~ target user.

two users.

$$sim(u, v) = \frac{\sum_{i \in Common} (r_{u,i} - \bar{r}_u)(r_{v,i} - \bar{r}_v)}{\sqrt{\sum_{i \in Common} (r_{u,i} - \bar{r}_u)^2} \sqrt{\sum_{i \in Common} (r_{v,i} - \bar{r}_v)^2}}$$

range = (-1, 1)



rate atleast 4 movies.

Seen.

total 5 movies

divide in 60-40 %.

Data into two parts → training data
→ test data

for U_1 → Training movies. Test movies.

(3) (2)

(I_1, I_3, I_5) (I_7, I_8)

find Sim (U_1, U_2), Sim (U_1, U_3), Sim (U_1, U_4), Sim (U_1, U_5)

Sim (U_1, U_2)

Common movies — I_1, I_3, I_5.

$\bar{r}_{U_1} = 4$, $\bar{r}_{U_2} = \frac{3+4+5+3+4+5}{6} = \frac{24}{6} = 4$

$$\text{Sim}(U_1, U_2) = \frac{(3-4)(3-4) + (4-4)(5-4) + (5-4)(4-}{\sqrt{(3-4)^2 + (4-4)^2 + (5-4)^2}\ \sqrt{(3-4)^2 + (5-4)^2 + (4-}}$$

$$= \frac{1}{\sqrt{2}\sqrt{2}} = 1/2 = 0.5$$

for U_1, U_3

Common movies = I_1, I_3, I_5

$$\bar{r}_{U_1} = 4_1 \qquad \bar{r}_{U_3} = 1+2+3+4+5\ /5$$

$$= 3$$

$$\text{Sim}(U_1, U_2) = \frac{(3-4)(1-3) + (4-4)(3-3) + (5-4)(5-3)}{\sqrt{(3-4)^2 + (4-4)^2 + (5-4)^2}\ \sqrt{(1-3)^2 + (3-3)^2 + (5-3)^2}}$$

$$= \frac{2+2}{\sqrt{2} \times 2\sqrt{2}} = \frac{4}{2 \times 2} = 1$$

for U_1, U_4

I_2, I_5, I_4

Common movies = I_1, I_3, I_5

$\bar{r}_{U_1} = 4$, $\bar{r}_{U_4} = \frac{5+2+3\times4+2+5}{8}$

$= \frac{14+12}{8} = 26/8$.

After all Similarities. use K nearest neighbours with active users.

if Similarities are $(U_1, U_2) = 0.5$

$(U_1, U_3) = 0.7$ ✓

$(U_1, U_4) = 0.3$

$(U_1, U_5) = 0.8$ ✓

K=2 generally

Similar user = U_3, U_5

We think test data as unseen movies for performance measure.

Test cases :—

① all similar user have seen the movie

② some fraction have seen the movie

③ No. similar user have seen the movie.

↙

System can't recommend.

Que-1 :—

	I_7	I_8
U_3	3	4
U_5	2	4

avg of I_7 = 2.5

" " I_8 = 4

Predicted rating for I_7 = 2.5

" " " I_8 = 4.

mean absolute error for $I_7 = 0.5$ } 0.75 (WRITE)
for $I_8 = .1$ }

.6
.3
.5
.7
} .6 (Ans)

We do this by considering other users as active and find out Mean absolute error. We avg all this error and say it as error for collaborative filtering.

two classes :- ham | Spam.

ham d_1: "good"

ham d_2: "very good"

Spam d_3: "bad"

Spam d_4: "very bad"

Spam d_5: "very bad, very bad"

Test: d_6: "good ? bad ! very bad"

$P(\text{Spam}) = 3/5$, $P(\text{ham}) = 2/5$

$P(\text{good} / \text{Spam}) = 0$, ~~$P(\text{ham} / \text{Spam}) = 0$~~

$P(\text{good} | \text{ham}) = 2/3$ $P(\text{bad} | \text{Spam}) = 4/7$

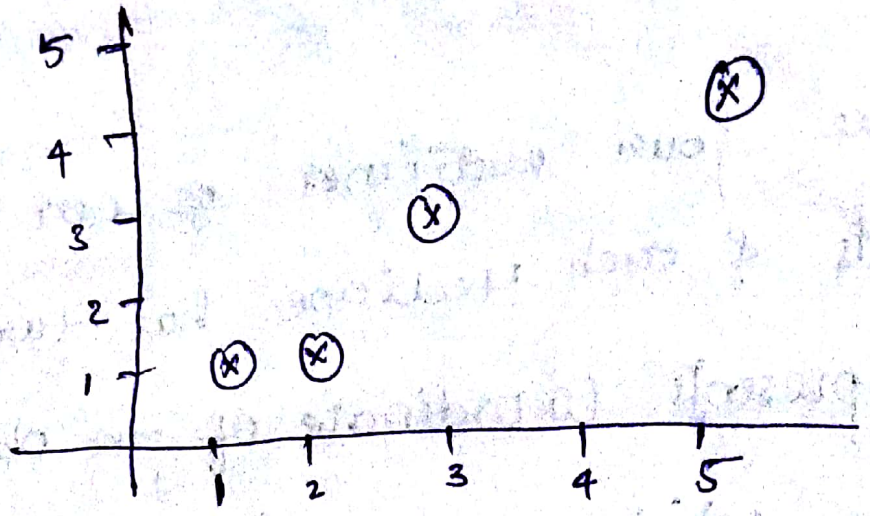
$P(\text{very} | \text{Spam}) = 3/7$ $P(\text{bad} | \text{ham}) = 0$.

$P(\text{very} | \text{ham}) = 1/3$

Clustering :-

Ex: Suppose we have four medicines as our training data point objects & each Medicine has two attributes, each attribute represents co-ordinate of the object, we have to determine which medicine belongs to medicine 1 & which one belongs medicine 2.

object	weighted Index (X)	pH (Y)
Med A	1	1
Med B	2	1
Med C	3	3
Med D	5	4



C1 (1, 1)
C2 (2, 1)
] → assumption of centroids as first two data sets.

A	B	C	D
1	2	3	5
1	1	3	4

C1 (1,1) C2 (2,1)

D^0_2

	A	B	C	D
C1	0	1	3.61	5
C2	1	0	2.83	4.24

$G^0 =$

1	0	0	0
0	1	1	1

Group 1

Group 2

$C1 = (1, 1)$

$C2 = \left(\frac{2+3+5}{3}, \frac{1+3+4}{3}\right) = (10/3, 8/3)$

$D^1 =$

	A	B	C	D
C1	0	1	361	5
C2	3.14	2.36	0.47	1.89.

$G^1 =$

1	1	0	0
0	0	1	1

Group 1

Group 2

$C1 = \left(\frac{1+2}{2}, \frac{1+1}{2}\right) = (3/2, 1)$

$C2 = \left(\frac{3+5}{2}, \frac{3+4}{2}\right) = (4, 7/2)$

D^2

c1	0.5	0.5	3.20	4.61
c2	4.30	3.54	0.71	0.71

G^2

	1	1	0	0
	0	0	1	1

Group 1

Group 2

final centroids.
$$\begin{cases} c1 = (3/2, 1) \\ c2 = (4, 7/2) \end{cases}$$

K-Means is not suitable for short text analysis, like tweets.

Hierarchial clustering

	BA	FI	MI	NA	RM	TO	MI/TO
BA	0	662	877	255	412	996	877
FI	662	0	295	468	268	400	295
MI	877	295	0	754	564	138	0
NA	255	468	754	0	219	869	754
RM	412	754	564	219	0	669	564
TO	996	400	138	869	669	0	0

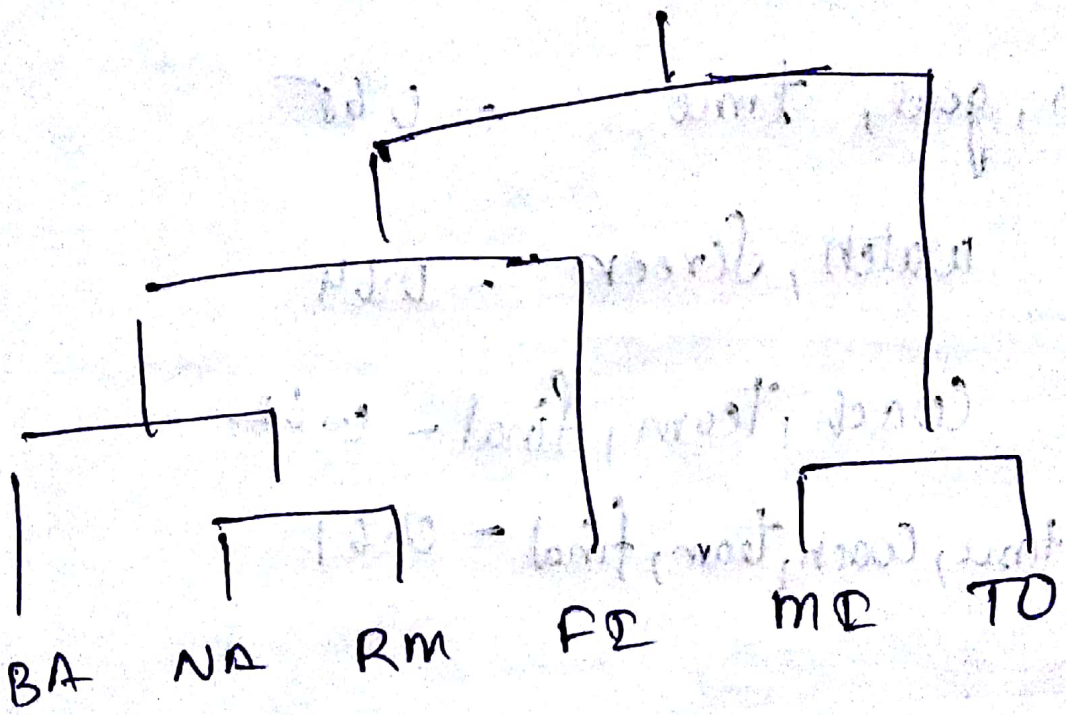
① Club two cities based on min distance

	BA	FI	MI/TO	NA	RM
BA	0	662	877	255	412
FI	662	0	295	468	268
MI/TO	877	295	0	266	564
NA	255	468	266	0	219
RM	412	268	564	219	0

	BA	FI	MI/TD	RM/NA
BA	0	662	877	255
FI	662	0	295	268
MI/TD	877	295	0	564
RM/NA	255	268	564	0

	FI	MI/TD	BA/RM/NA
FI	0	295	268
MI/TD	295	0	564
BA/RM/NA	268	564	0

	MI/TD	BA/RM/NA/FI
MI/TD	0	295
BA/RM/NA/FI	295	0



Q2)

=

replays, offside — 0.69

goal, time — 0.68

team, final — 0.65

match, goal, time — 0.65

watch, screen — 0.64

coach, team, final — 0.62

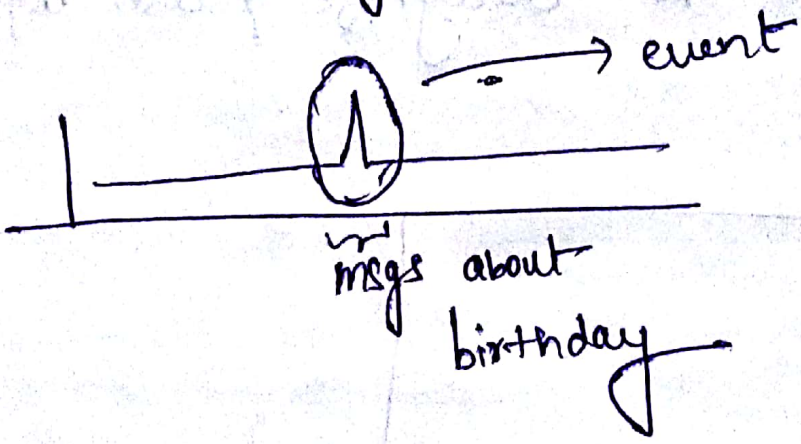
match, goal, time, coach, team, final — 0.61

Social Network Analysis

trends, events, Personalization, Social connectivity, digital Marketing, Knowing people, observe monitor people,

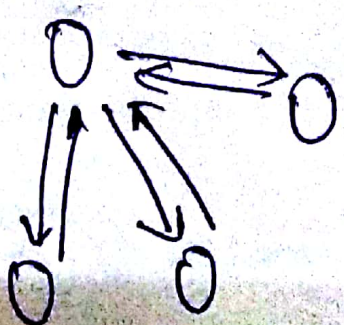
trend → what people are interested in

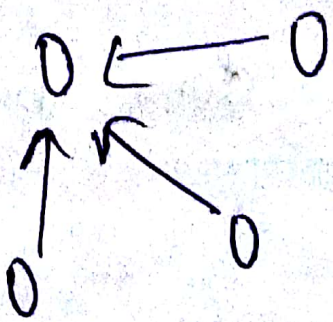
event → like birthday.



Essentially analysing human behaviour.
(groups of humans)

If there is content analysis, we need NLP.





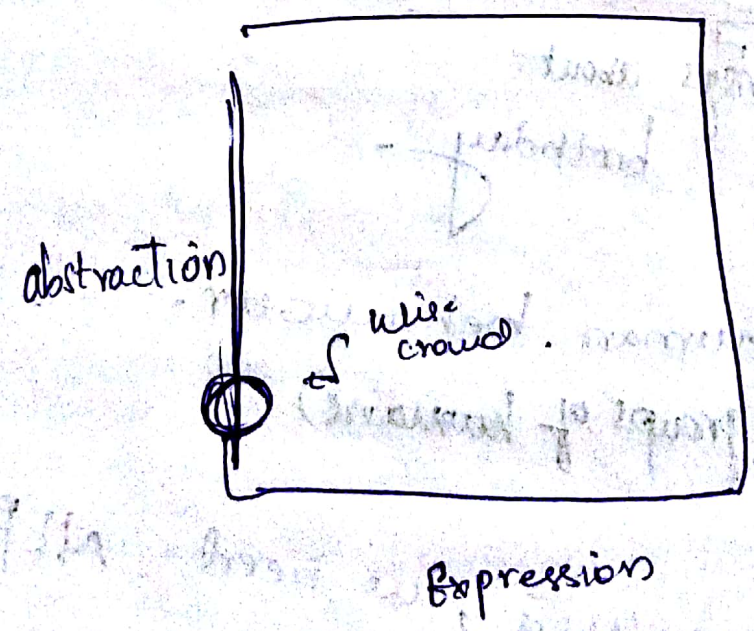
Qualitative analysis of online Groups :-

Herding :- following someone.

Crowd Sourcing :-

By reddit ⟶ disasterous

asking crowd to recognize / ask information of

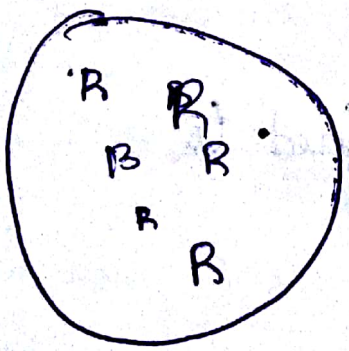


Wise Crowd ⟶ have independency, but agree upon some basic rules.

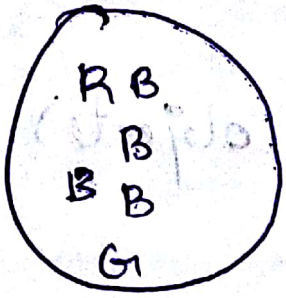
Purity of a cluster :— max count of similar objects in a class of the cluster, divided by total no. of objects.

$$\text{Purity} = \frac{\text{max count in a class}}{\text{total no. of objects}}$$

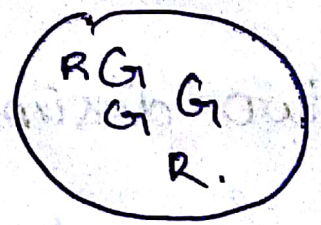
cluster 1



cluster-2



cluster-3



cluster 1 = max (5,0,1) / 6. = 5/6

cluster 2 = max (1,1,4) / 6. = 4/6

cluster 3 = max (2,3,0) / 5 = 3/5.

Purity of clustering algo = 5/6 + 4/6 + 3/5

= 6/6 + 8/5 = 14/6

$= \frac{5+4+3}{17} = 12/17$.

Rand Index

T_P:- two similar docs (objects) assigned to same cluster

T_N:- two dissimilar docs (objects) assigned to different cluster

f_P:- two dissimilar docs (objects) assigned to same cluster

f_N:- two similar docs (objects) assigned to different cluster.

	similar	dissimilar
similar	$T_P(a)$	$F_N(c)$
dissimalar	$f_P(b)$	$T_P(d)$

RI = $\frac{a+d}{a+b+c+d}$, P = $a/a+b$, R = $a/a+c$

class/cluster	V_1	V_2	-----	V_c	SUM
U_1	U_{11}	U_{12}	---	U_{1c}	n_1
U_2	U_{21}	U_{22}	---	U_{2c}	n_2
:	:	:		:	:
U_R	U_{R1}	U_{R2}	----	U_{RC}	n_R
SUM	n_1	n_2	-----	n_c	

class/cluster	V_1	V_2	V_3	Sum	
U_1	1	1	0	2	U_1.
U_2	1	2	1	4	U_2.
U_3.	0	0	4	4	U_3
Sum	2	3	5	10	[n].
	$V_{.1}$	$V_{.2}$	$V_{.3}$.		

$$a = \sum_{i,j} \binom{n_{ij}}{2}$$

a → no. of pair of objects in the same class U & in the same cluster in V

b → no. of pair of objects in the same class U & not in same cluster in V

c → no. of pairs of objects in the different class U & same cluster in V

d → no. of pairs of objects in the different class in U & different cluster in V.

$$a = \sum_{i,j} \binom{n_{ij}}{2}$$

$$= \binom{2}{2} + \binom{4}{2} = 1 + 6 = 7$$

$$b = \sum_i \binom{n_{i\cdot}}{2} - \sum_{ij} \binom{n_{ij}}{2}$$

$$= \binom{2}{2} + \binom{4}{2} + \binom{4}{2} - \binom{2}{2} - \binom{4}{2}$$

$$= 6$$

$$c = \binom{2}{2} + \binom{3}{2} + \binom{5}{2} - 7 - \left[\sum_j \binom{n_{\cdot j}}{2} - \sum_{ij} \binom{n_{ij}}{2} \right]$$

$$= 1 + 3 + 10 - 7$$

$$= 7.$$

$$d = \binom{n}{2} - (a+b+c)$$

$$= \binom{10}{2} - (20) = 45 - 20 = 25$$

$$RI = \frac{a+d}{a+b+c+d} = \frac{7+25}{45} = \frac{32}{70}.$$

class / cluster	1	2	3	
R	5	1	2	8 n_1.
G	0	1	3	4 n_2.
B	1	4	0	5 n_3.
	6	6	5	17
	$n_{.1}$	$n_{.2}$	$n_{.3}$	

$$a = \binom{5}{2} = 10.$$

$$b = \binom{8}{2} + \binom{4}{2} + \binom{5}{2} - 10$$

$$= 28 + 6 + 10 - 10 = 34$$

$$c = \binom{6}{2} + \binom{6}{2} + \binom{5}{2} - 10$$

$$= 15 + 15 + 10 - 10 = 30.$$

$d = \binom{17}{2} - (10 + 34 + 30)$

$= 136 - 74 = 62$

$RI = \frac{10 + 62}{10 + 34 + 30 + 62} = \frac{72}{136} = 0.52$

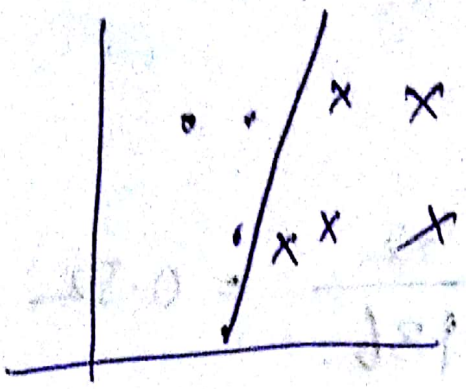
NMI (normalized mutual information)

Adjusted Rand Index

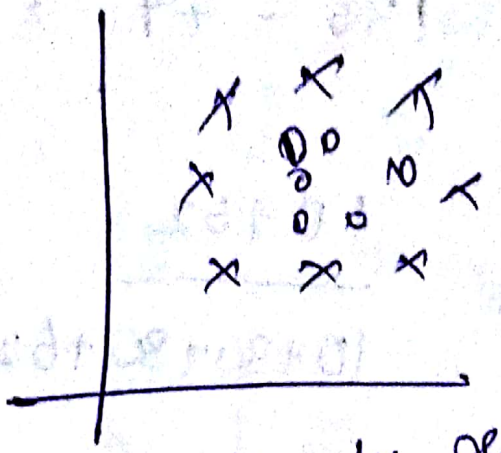
Maximum Adjusted Rand Index

a b

Support Vector Machine



linearly seperable.



non linearly seperable.

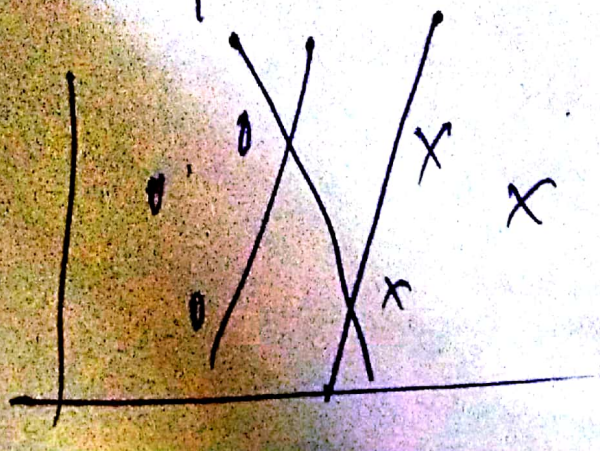
linearly seperable in

1D plane — line

3D plane — plane.

>3D plane — Hyperplane.

For linear seperable

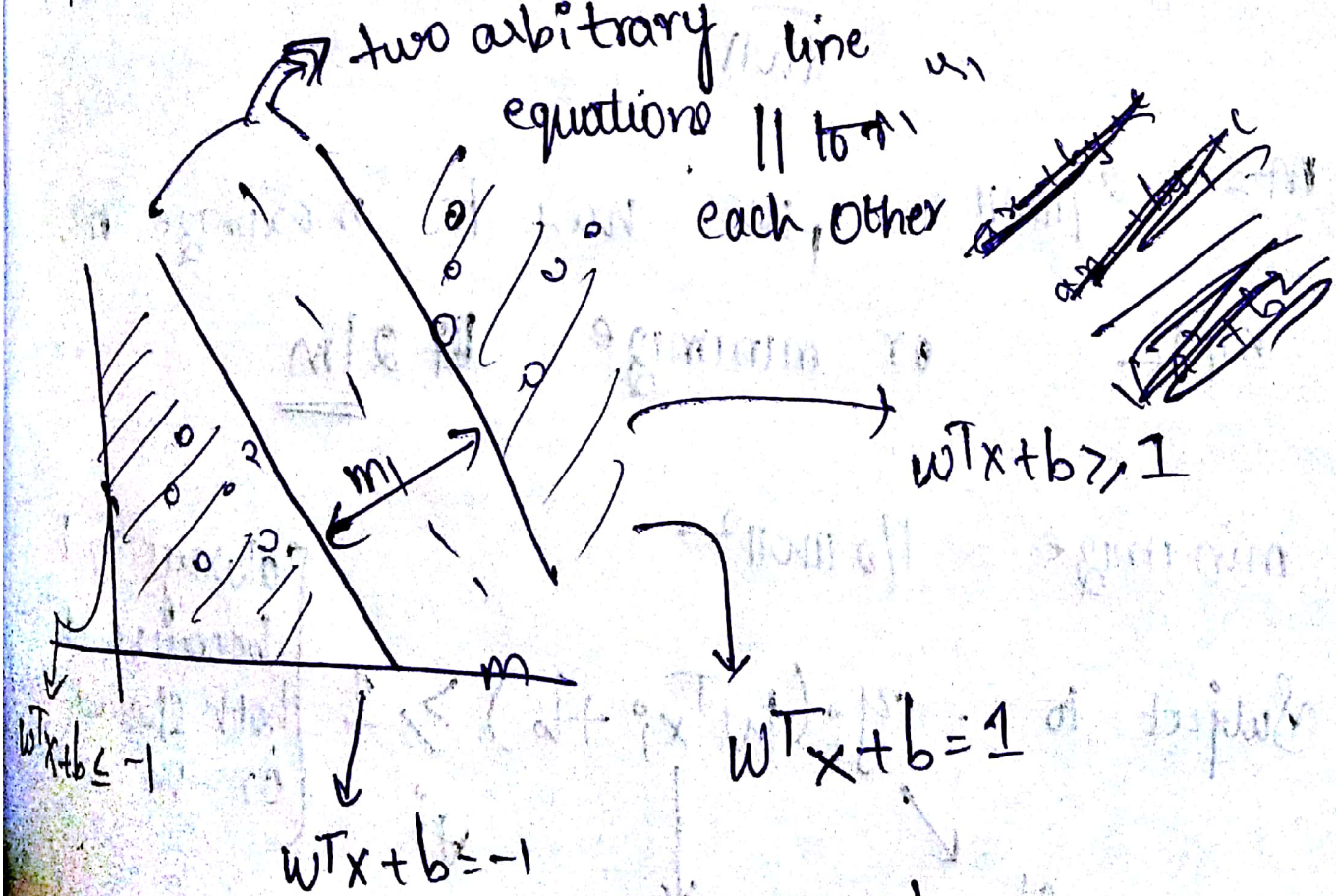


many lines can be fit, hence SVC gives best fit line.

It is a kind of optimization problem.

- Objective function
- Constraints
  - non negative restrictions
  - Simple constraints

Objective is to maximize the margin.



distance two lines = $\frac{|1 \pm 1|}{\sqrt{a^2+b^2}}$

1st dis = $\frac{\|b+1\|}{\|w\|}$. $\|w\| = \sqrt{w_1^2 + w_2^2}$

1st dis = $\frac{\|b-1\|}{\|w\|}$

distance = $\frac{b+1 - b+1}{\|w\|}$

$= \frac{2}{\|w\|}$

$m = 2/\|w\|$, we need to maximize m.

or $m/2$ or minimize $2/m$

minimize $= 1/2 \|w\|^2$

Subject to : $y_i(w^T x_i + b) \geq 1 \quad \forall i$

always ≥ 1 because either both should +ve or -ve

Class label of i^{th} instance.

i^{th} instance / i^{th} instance of data

Data = $\{x_1 \cdots x_n\}$, Class labels = $\{-1, 1\}$

↓

instances of Data. (Vector)

After solving optimization, then we get value of w & b. (use Langrange multipier)

objective function — non linear } overall
constraints — linear — } non linear.

Those (α_i's) → langrange coefficient ≥ 0 , Corresponding examples are called Support vectors.

↳ points on ~~w^T x + b~~ $w^T x + b = -1$
$= +1$

$\dim(w) = \dim(x_i) =$ no. of attributes.