Building Blocks of Deep Networks

Building deep networks goes beyond basic feed-forward multilayer neural networks.

Three specific building blocks:-

- Feed-forward multilayer neural networks
- RBMs
- Autoencoders

Feed-Forward Networks

- Feed-Forward Networks are the simplest Artificial Neural Networks.

- They are composed of an **input layer, one or many hidden layers, and an output layer.**

RBM(Restricted Boltzmann Machine)

RBMs are used in deep learning for the following:

- **Feature extraction**
- **Dimensionality reduction**

The “restricted” part of the name “Restricted Boltzmann Machines” means that **connections between nodes of the same layer are prohibited**

(e.g., there are no visible-visible or hidden- hidden connections along which signal passes).

RBMs Network

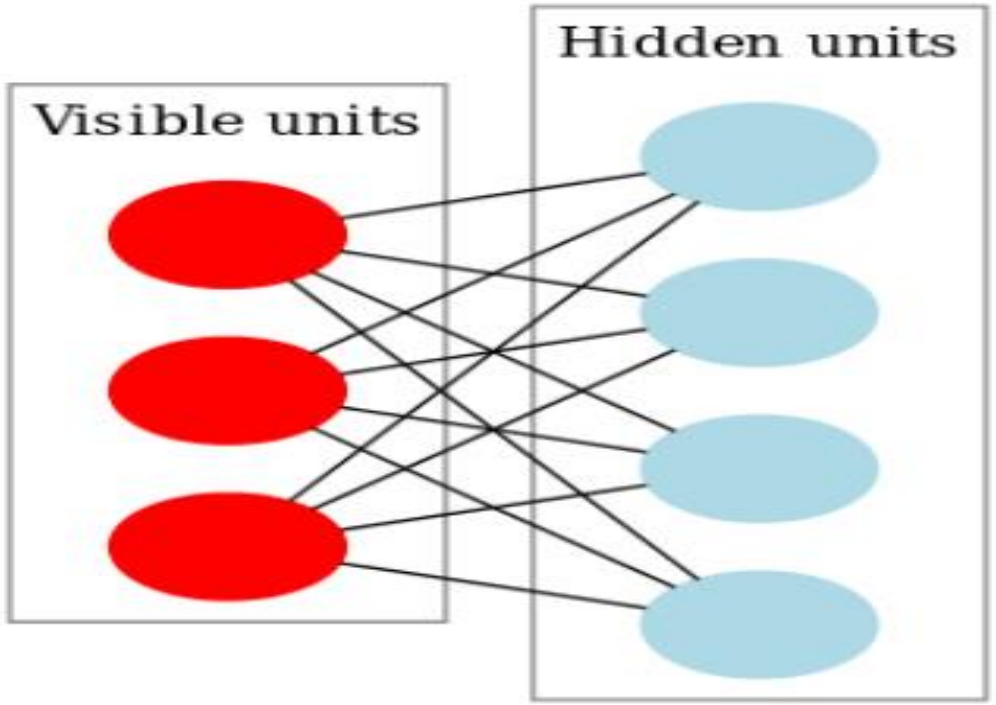


Figure : Restricted Boltzmann machine with three visible units and four hidden units (no bias units).

RBMs Network…

- ✓RBMs are also a type of autoencoder, which we'll talk about in the following section.

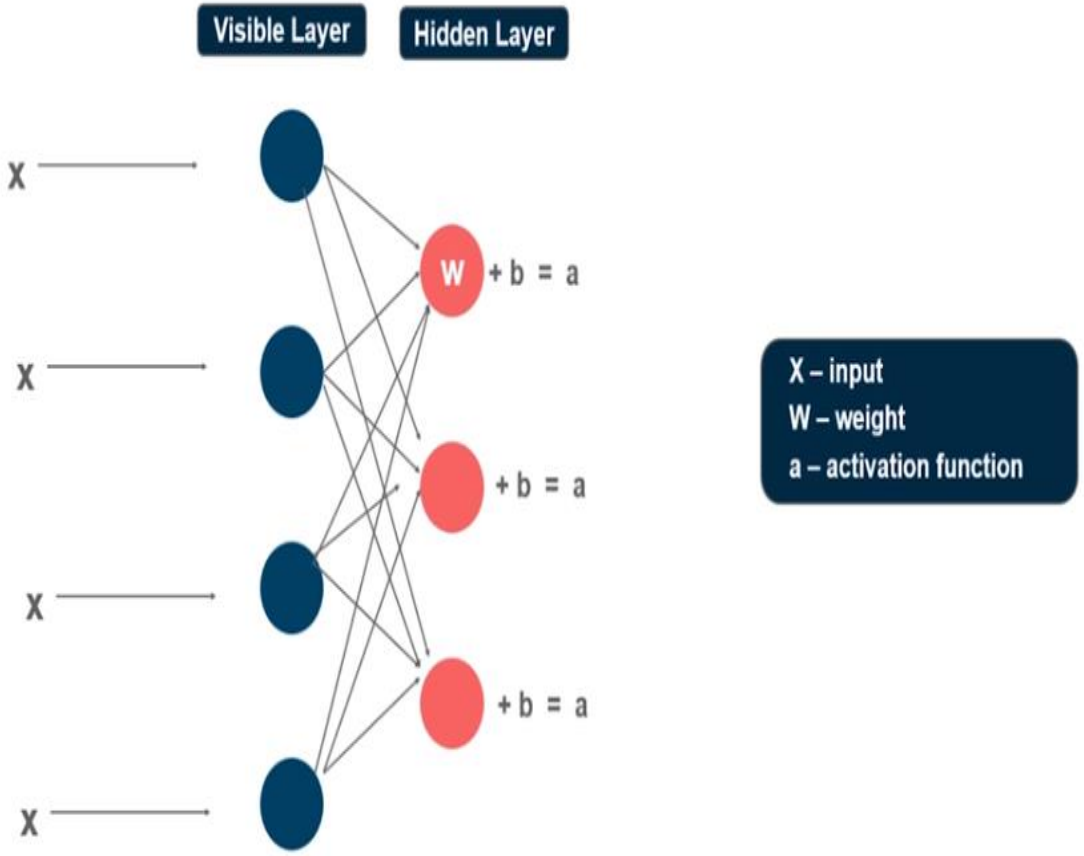
- ✓RBMs are used for pretraining layers in larger networks such as Deep Belief Networks.

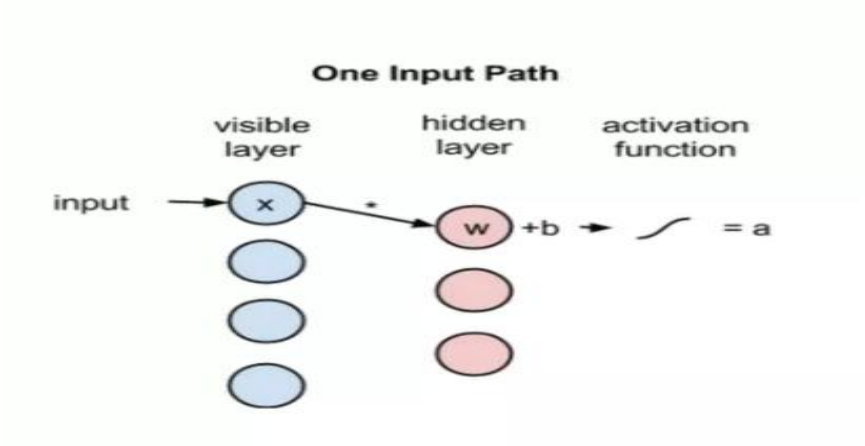
Network layout

There are five main parts of a basic RBM:

1. Visible units
2. Hidden units
3. Weights
4. Visible bias units
5. Hidden bias units

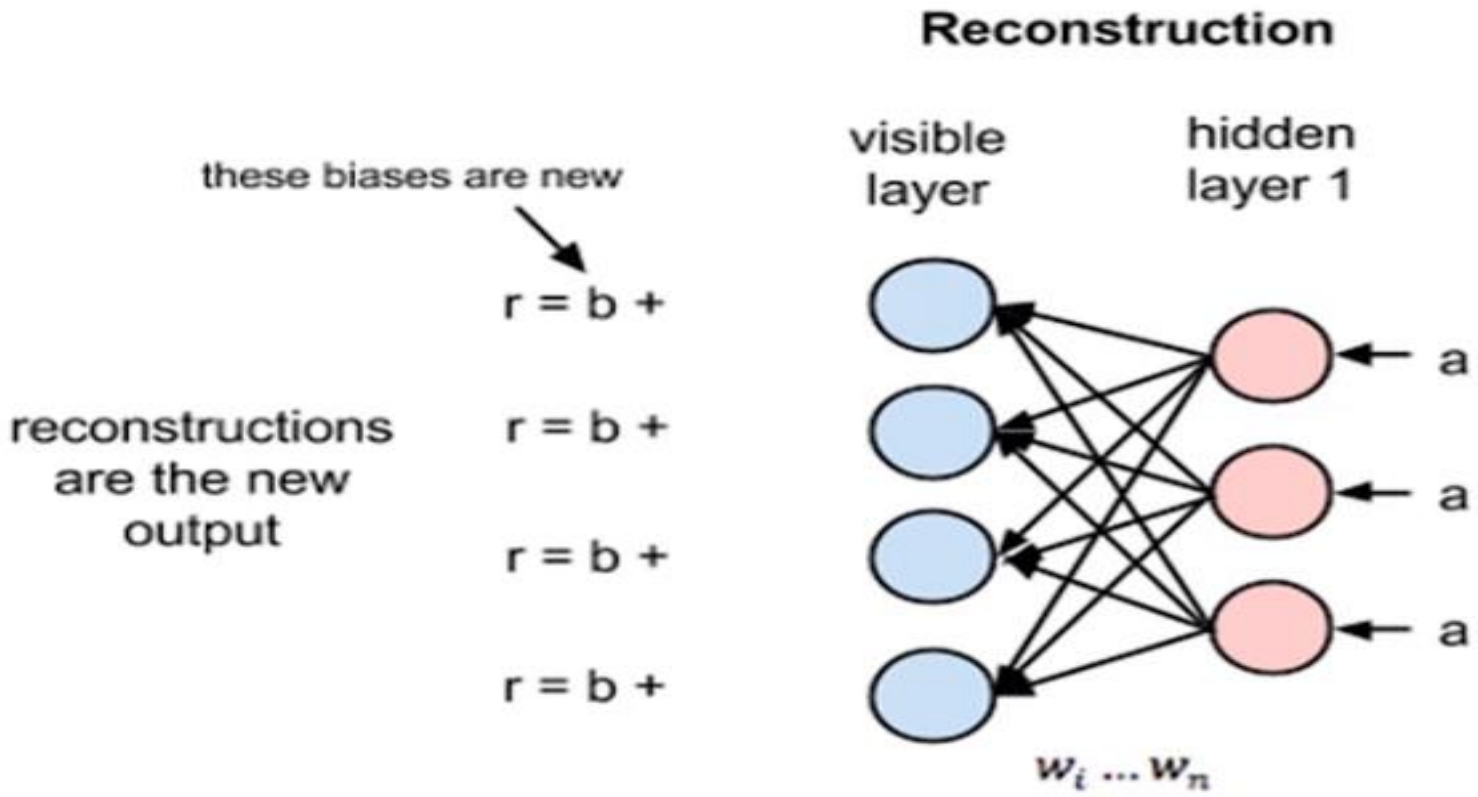
Working of RBMs



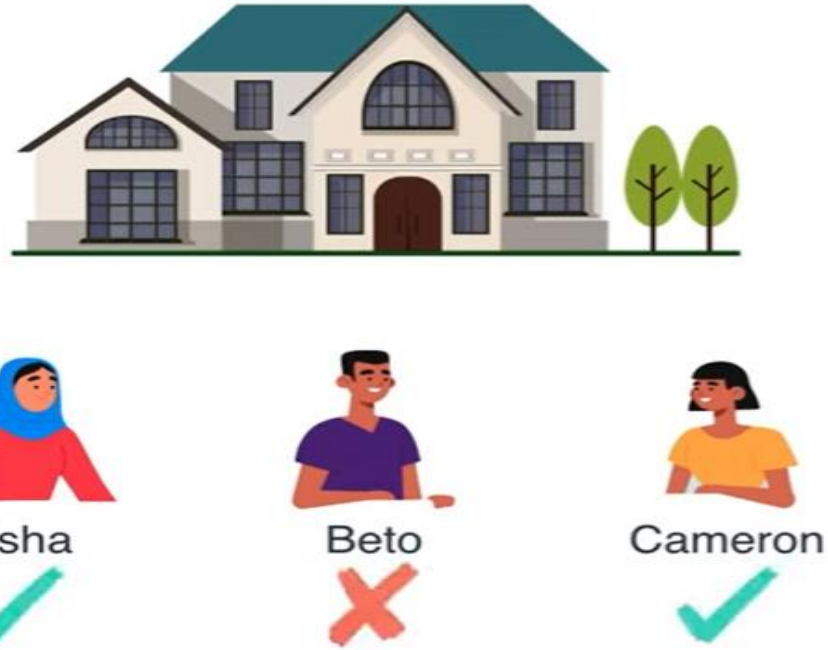


- Layer x is multiplied by a weight and added to a bias and the result of the operation is passed to the activation function which produces the output.
- Reconstruction of RBMs

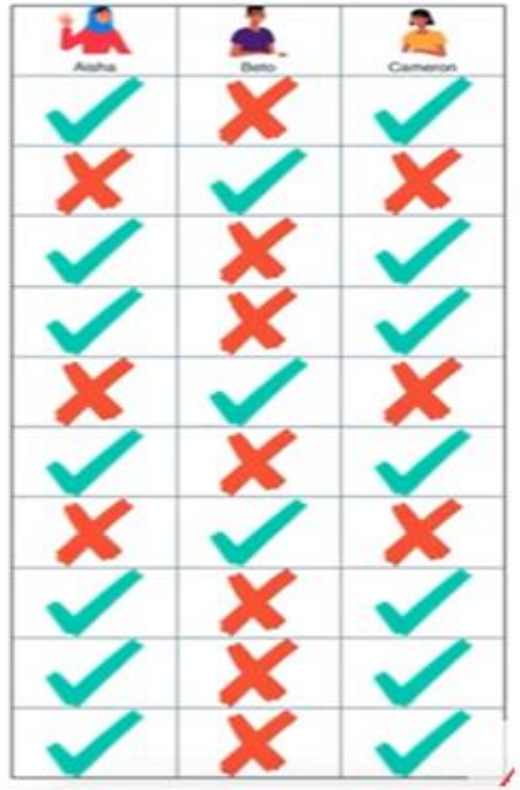
Reconstruction of RBMs



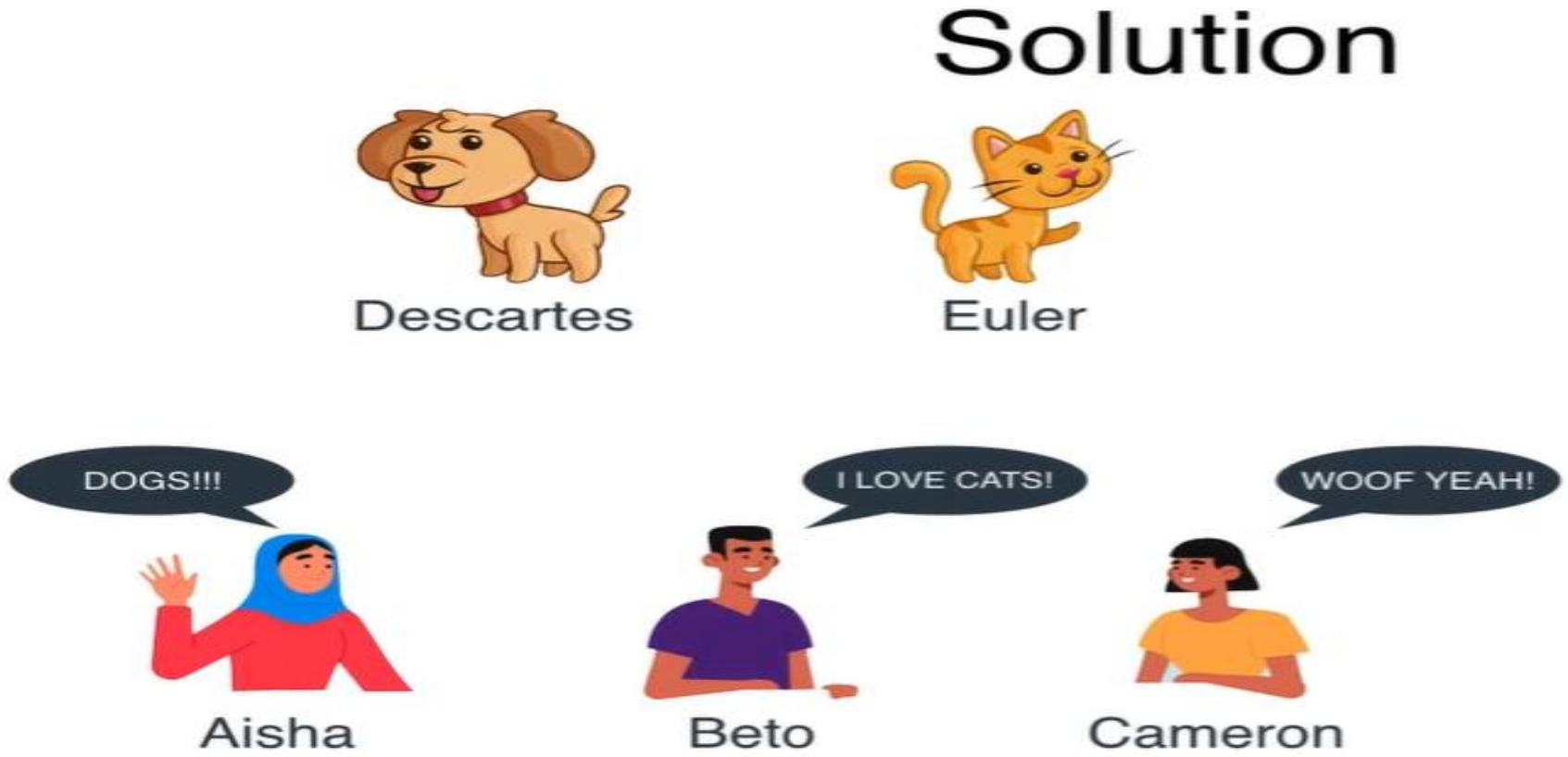
Example-RBM



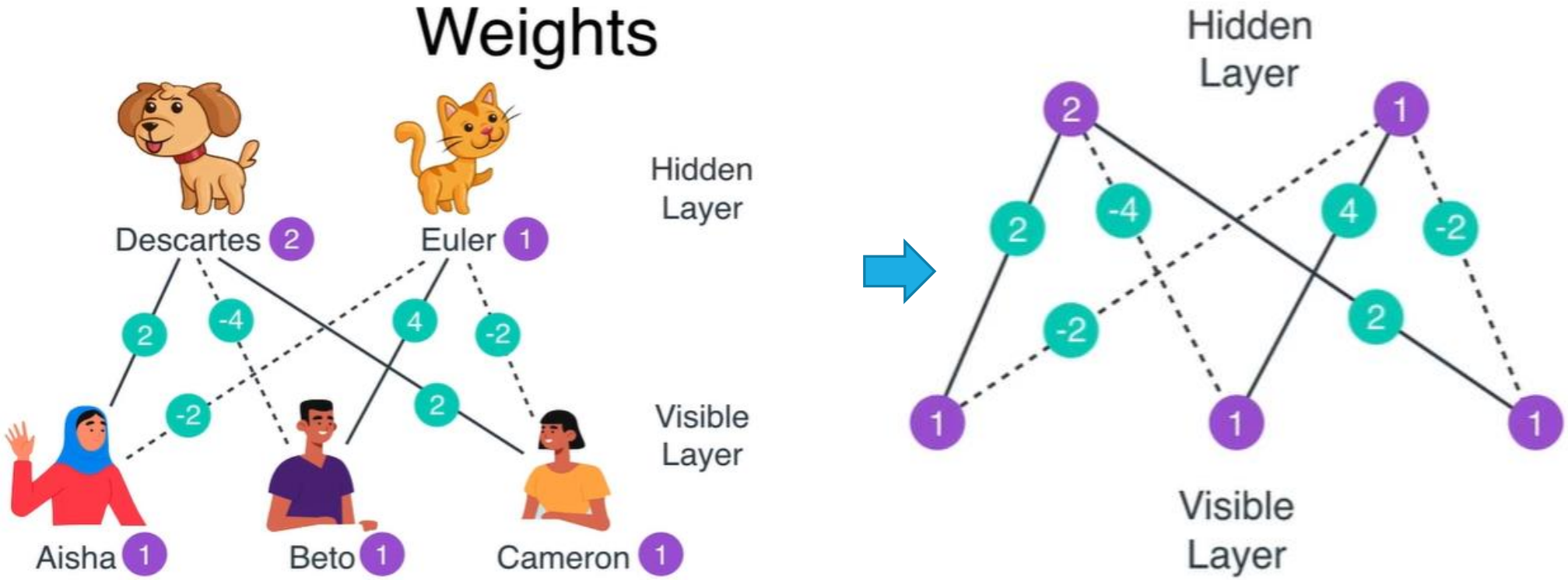
Example-RBM…



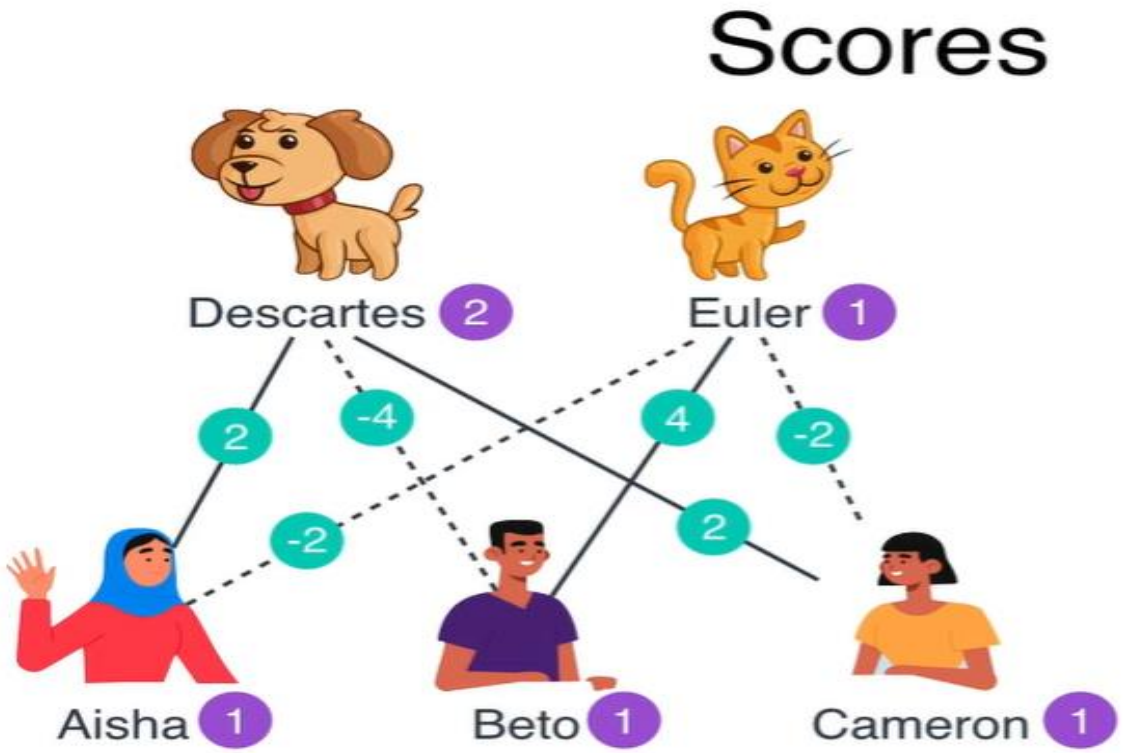
Example-RBM…

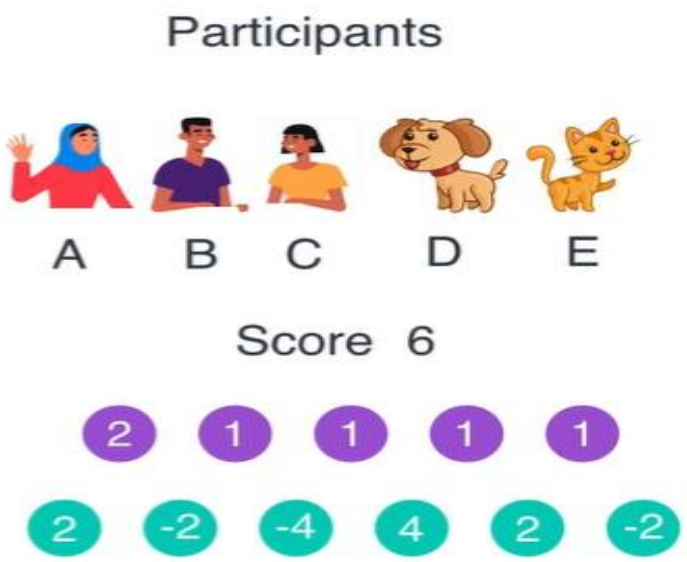

Solution
Descartes
Euler
DOGS!!!
I LOVE CATS!
WOOF YEAH!
Aisha
Beto
Cameron

Example-RBM…

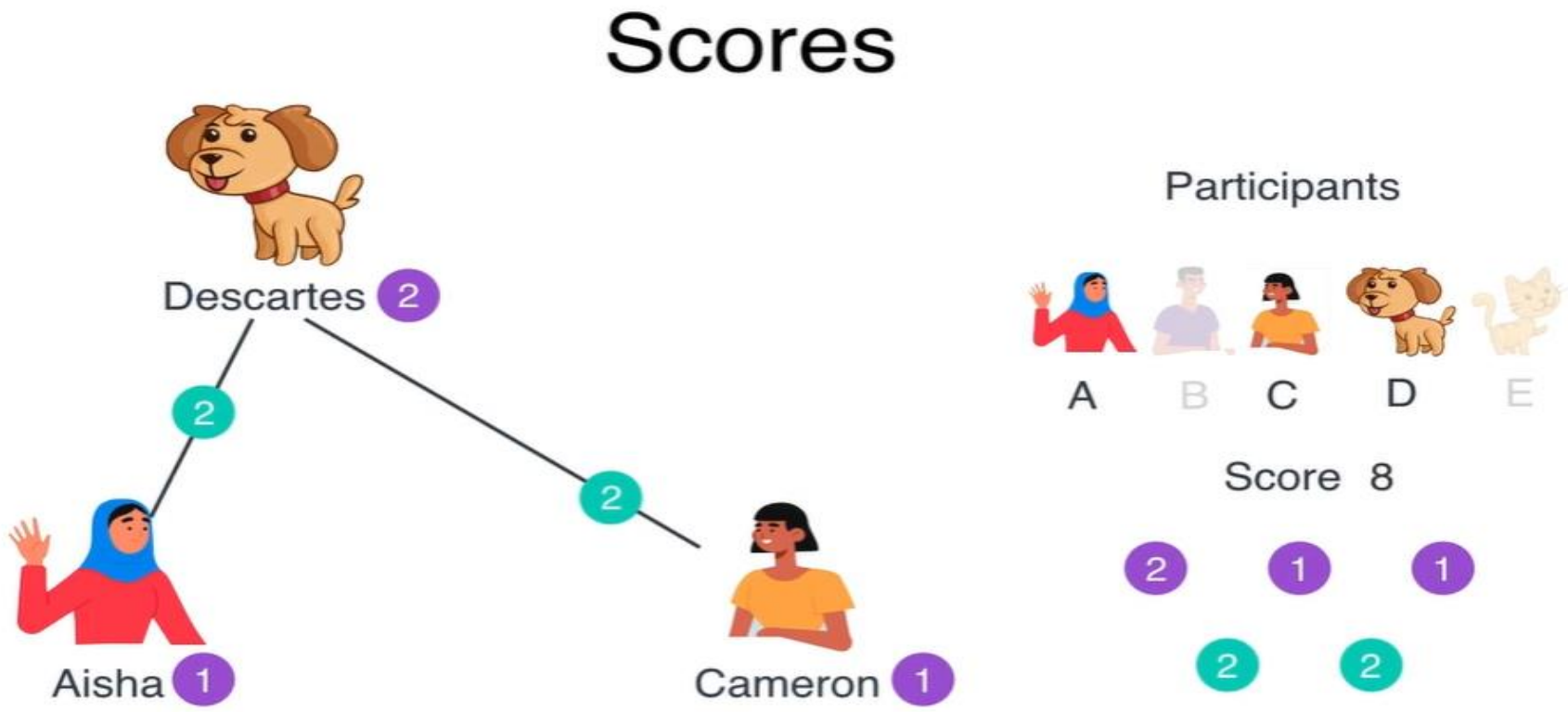

Weights
Descartes 2
Euler 1
Hidden Layer
2
-4
4
-2
-2
2
Visible Layer
Aisha 1
Beto 1
Cameron 1
Hidden Layer
Visible Layer

Example-RBM…

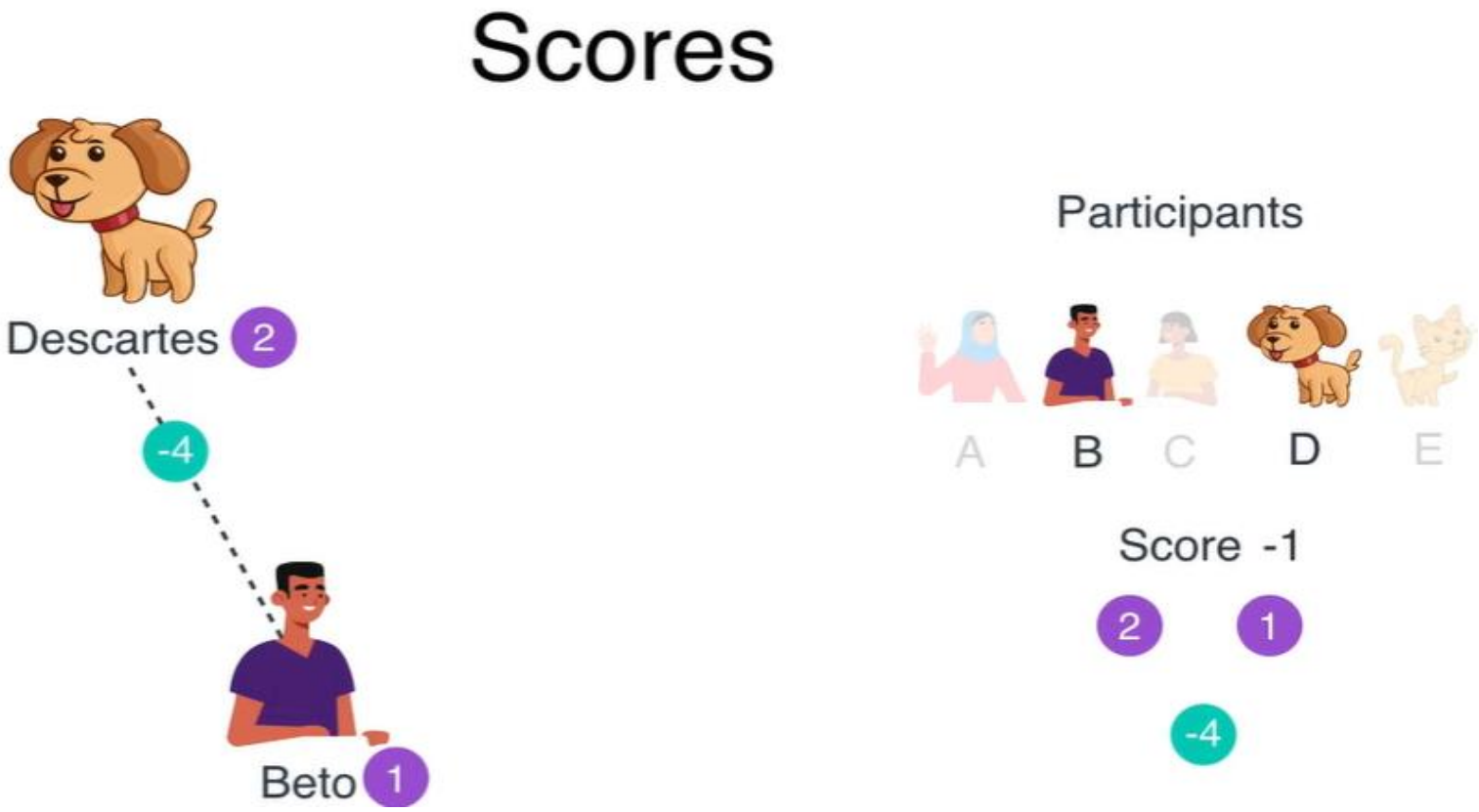

Scores
Descartes 2
Euler 1
2
-4
4
-2
-2
2
Aisha 1
Beto 1
Cameron 1


Participants
A
B
C
D
E
Score 6
2
1
1
1
1
2
-2
-4
4
2
-2

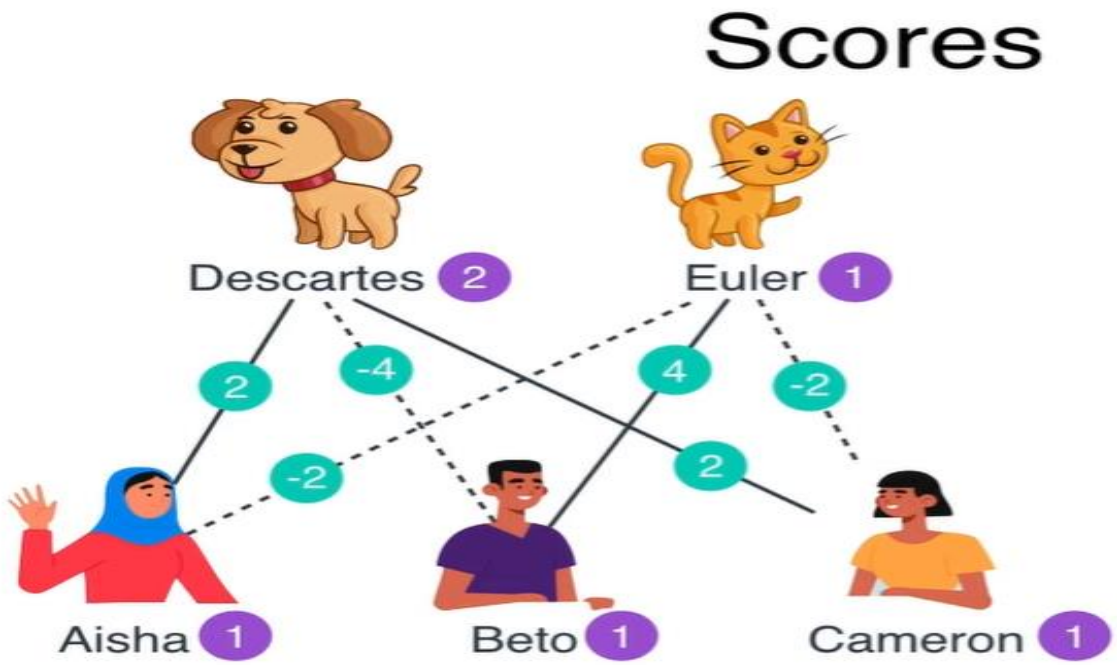
Let us Consider –the scenario-Aisha,Cameron and Dog(ACD)



Let us Consider -scenario-Beto and Dog(BD)



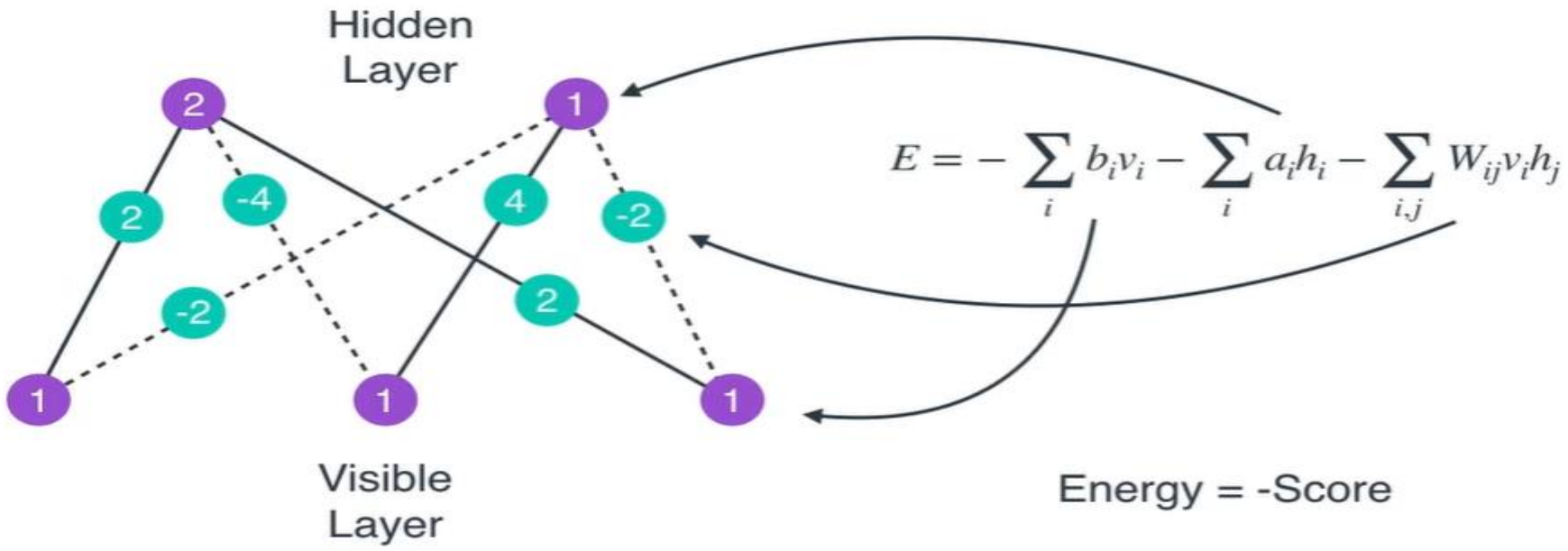
Score=Add all the weights in each Scenario



Scenario	Score
None	0
A	1
B	1
C	1
D	2
E	1
AB	2
AC	2
AD	5
AE	0
BC	2
BD	-2
BE	7
CD	5
CE	0
DE	3
ABC	3
ABD	1
ABE	6
ACD	8
ACE	-1
ADE	4
BCD	1
BCE	6
BDE	4
CDE	4
ABCD	4
ABCE	5
ABDE	5
ACDE	5
BCDE	5
ABCDE	6

Restricted Boltzmann Machine

Energy of the RBM is the negative of the sum of the scores of the weights of hidden layer and visible layer.



Restricted Boltzmann Machine

Consider Three Scores(3,2,1) and find the probabilities

Scores to probabilities

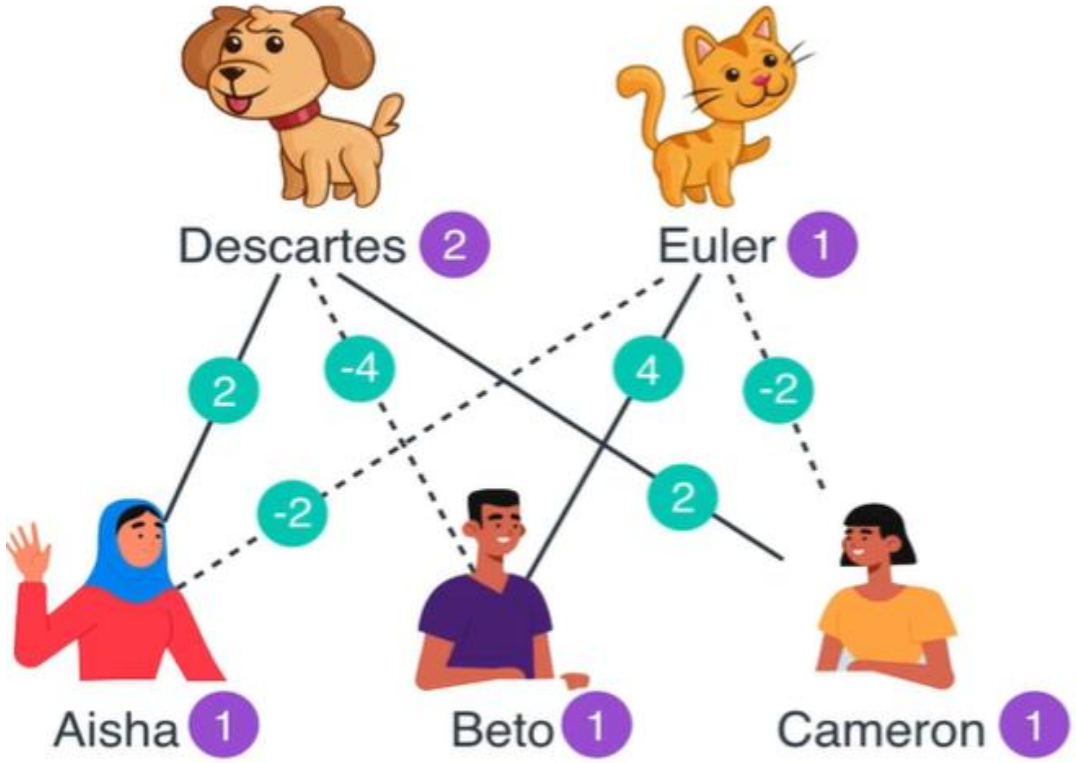
Score	Probability
3	1/2
2	1/3
1	1/6
Sum = 6	Sum = 1

Restricted Boltzmann Machine..

Scores to probabilities

Score	e^{score}	Normalize
1	$e^{1} = 2.718$	0.665
0	$e^{0} = 1$	0.245
-1	$e^{-1} = 0.368$	0.09
	Sum = 4.086	Sum = 1

Restricted Boltzmann Machine..

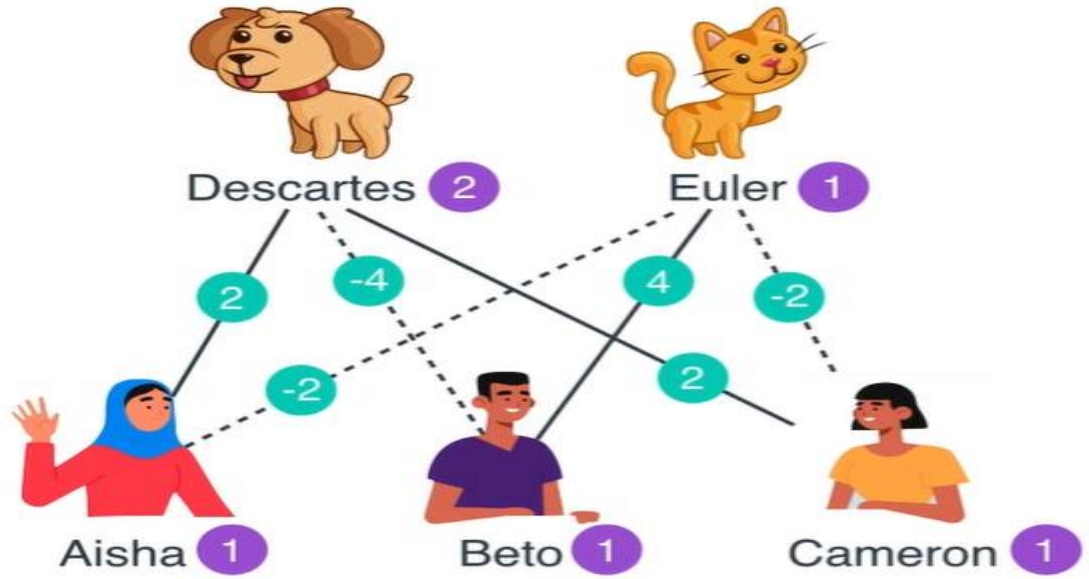


Scenario	Score	e^{Score}	Probability
None	0	1	0
A	1	2.72	0
B	1	2.72	0
C	1	2.72	0
D	2	7.38	0
E	1	2.72	0
AB	2	7.38	0
AC	2	7.38	0
AD	5	148.41	0.02
AE	0	2.72	0
BC	2	7.38	0
BD	-2	0.14	0
BE	7	1096.63	0.17
CD	5	148.41	0.02
CE	0	1	0
DE	3	20.08	0
ABC	3	20.08	0
ABD	1	2.72	0
ABE	6	403.43	0.06
ACD	8	2980.96	0.45
ACE	-1	0.37	0
ADE	4	54.6	0
BCD	1	2.72	0
BCE	6	403.43	0.06
BDE	4	54.6	0
CDE	4	54.6	0
ABCD	4	54.6	0.02
ABCE	5	148.41	0.02
ABDE	5	148.41	0.02

Restricted Boltzmann Machine..

Maximum Likelihood probability is highlighted in Blue.

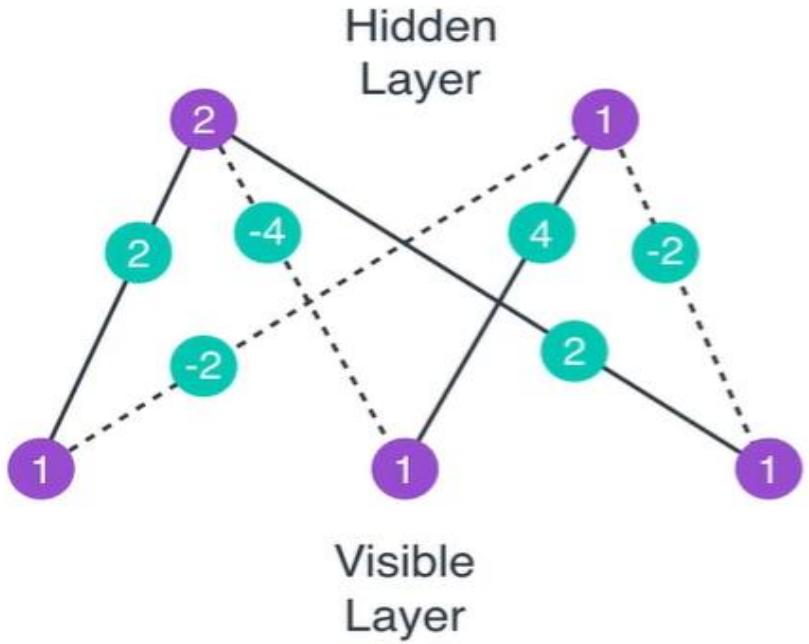
Least Likelihood probability is highlighted in Red.



Scenario	Score	e^{Score}	Probability
None	0	1	0
A	1	2.72	0
B	1	2.72	0
C	1	2.72	0
D	2	7.38	0
E	1	2.72	0
AB	2	7.38	0
AC	2	7.38	0
AD	5	148.41	0.02
AE	0	2.72	0
BC	2	7.38	0
BD	-2	0.14	0
BE	7	1096.63	0.17
CD	5	148.41	0.02
CE	0	1	0
DE	3	20.08	0
ABC	3	20.08	0
ABD	1	2.72	0
ABE	6	403.43	0.06
ACD	8	2980.96	0.45
ACE	-1	0.37	0
ADE	4	54.6	0
BCD	1	2.72	0
BCE	6	403.43	0.06
BDE	4	54.6	0
CDE	4	54.6	0
ABCD	4	54.6	0.02
ABCE	5	148.41	0.02
ABDE	5	148.41	0.02

Restricted Boltzmann Machine..

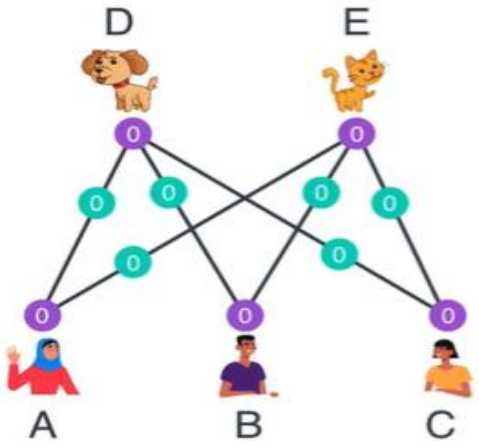
Energy to probability



$$E = -\sum_i b_i v_i - \sum_i a_i h_i - \sum_{i,j} W_{ij} v_i h_j$$

$$p(v,h) = \frac{1}{Z} e^{-E(v,h)} \qquad Z = \sum_{v,h} e^{-E(v,h)}$$

How to Train RBM

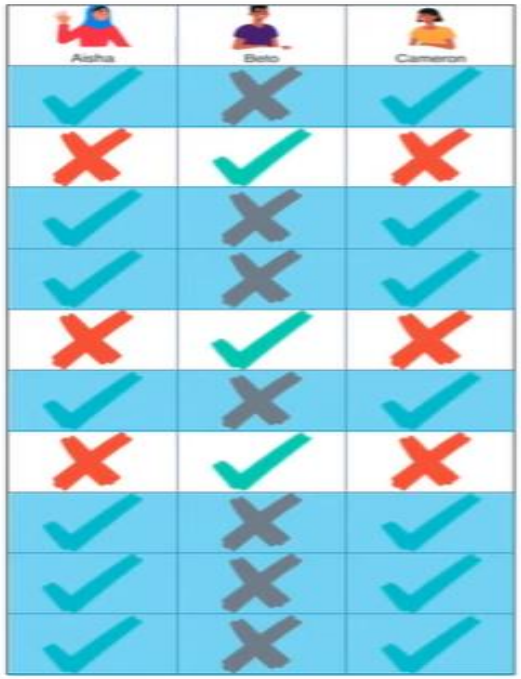


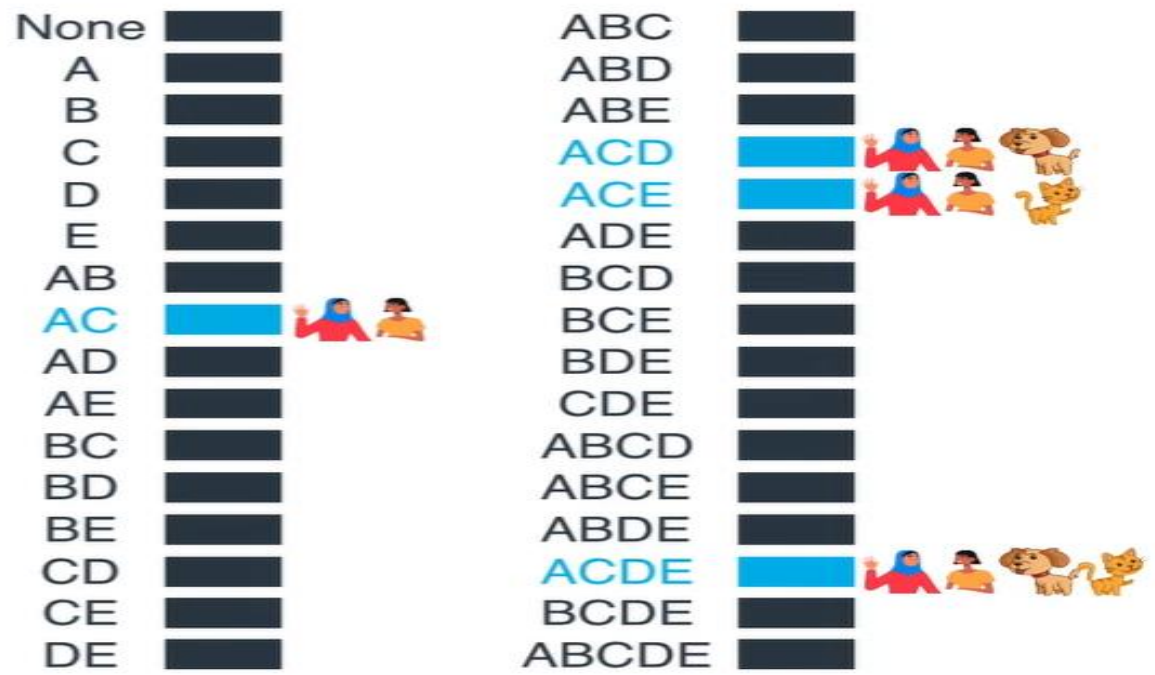
Scenario	Score	e^{Score}	Probability
None	0	1	1/32
A	0	1	1/32
B	0	1	1/32
C	0	1	1/32
D	0	1	1/32
E	0	1	1/32
AB	0	1	1/32
AC	0	1	1/32
AD	0	1	1/32
AE	0	1	1/32
BC	0	1	1/32
BD	0	1	1/32
BE	0	1	1/32
CD	0	1	1/32
CE	0	1	1/32
DE	0	1	1/32

Scenario	Score	e^{Score}	Probability
ABC	0	1	1/32
ABD	0	1	1/32
ABE	0	1	1/32
ACD	0	1	1/32
ACE	0	1	1/32
ADE	0	1	1/32
BCD	0	1	1/32
BCE	0	1	1/32
BDE	0	1	1/32
CDE	0	1	1/32
ABCD	0	1	1/32
ABCE	0	1	1/32
ABDE	0	1	1/32
ACDE	0	1	1/32
BCDE	0	1	1/32
ABCDE	0	1	1/32

How to Train RBM…

Consider the scenario –Aisha and Cameron are show up, then the probability of all the cases with Aisha and Cameron - Maximum .

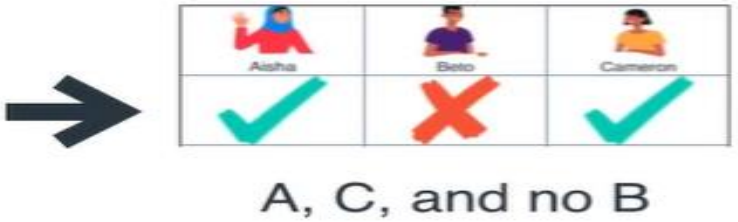


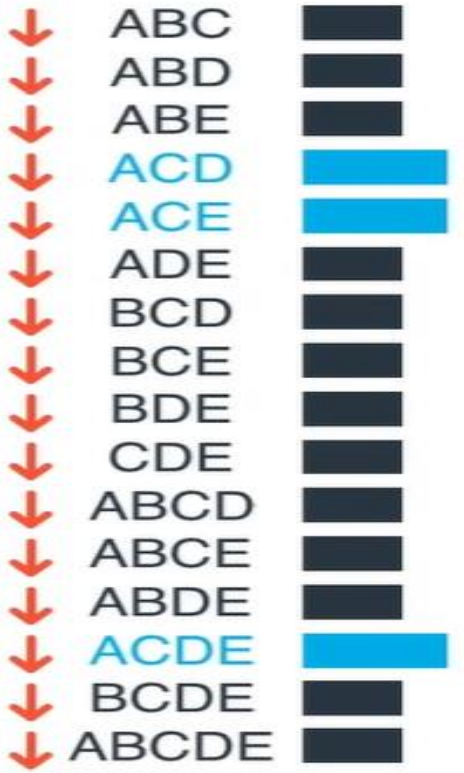


Probability of all the scenario with A,C and no B will be Maximum

A, C, and no B

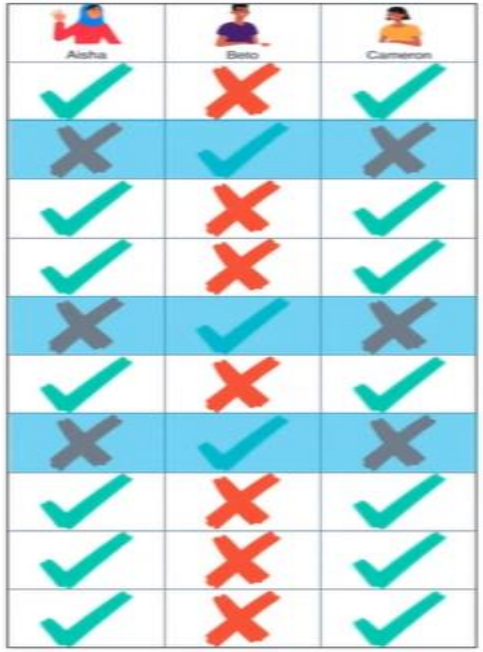
Probability of all the scenario except A,C and no B will be less

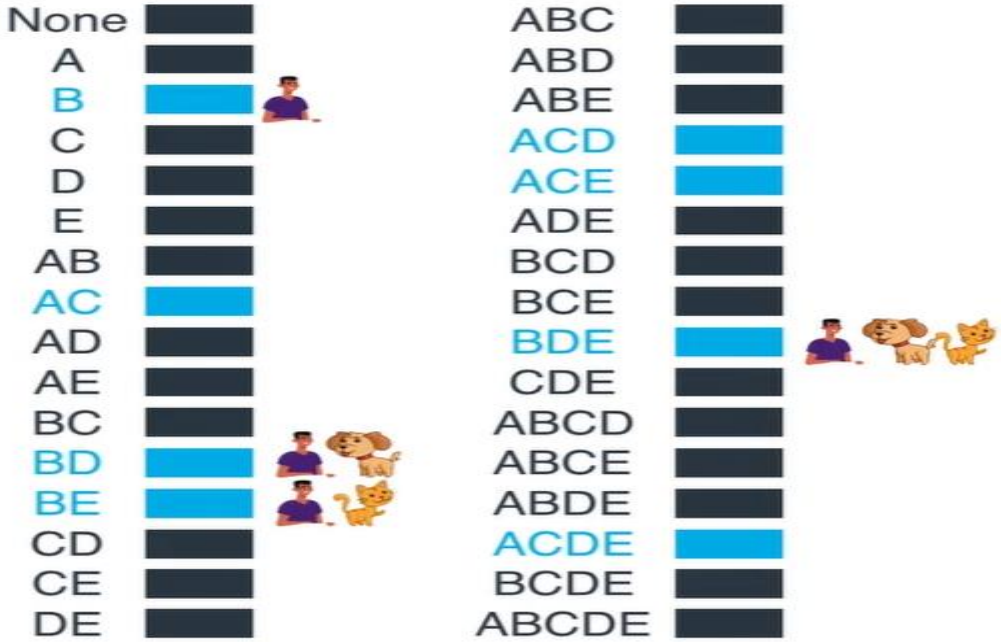




How to Train RBM…

Consider the scenario –Beto shows up, then the probability of all the cases with Beto has to be Maximum.



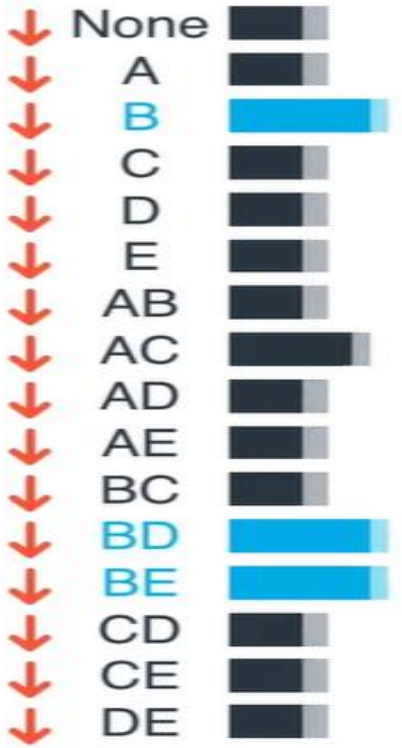


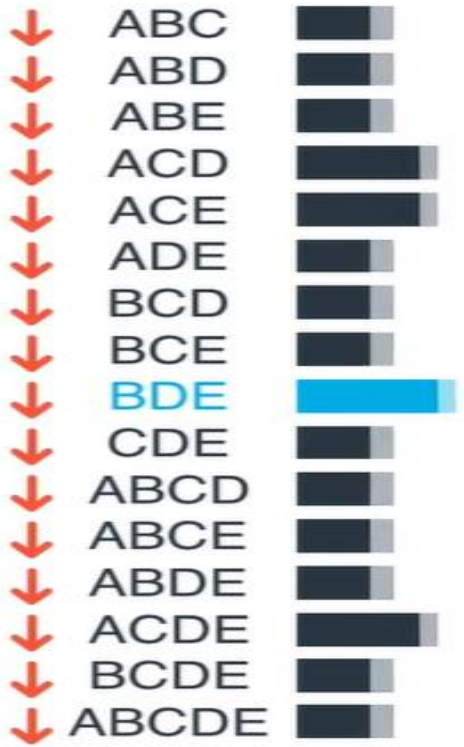
How to Train RBM…

How to Train RBM…

B and no A,C

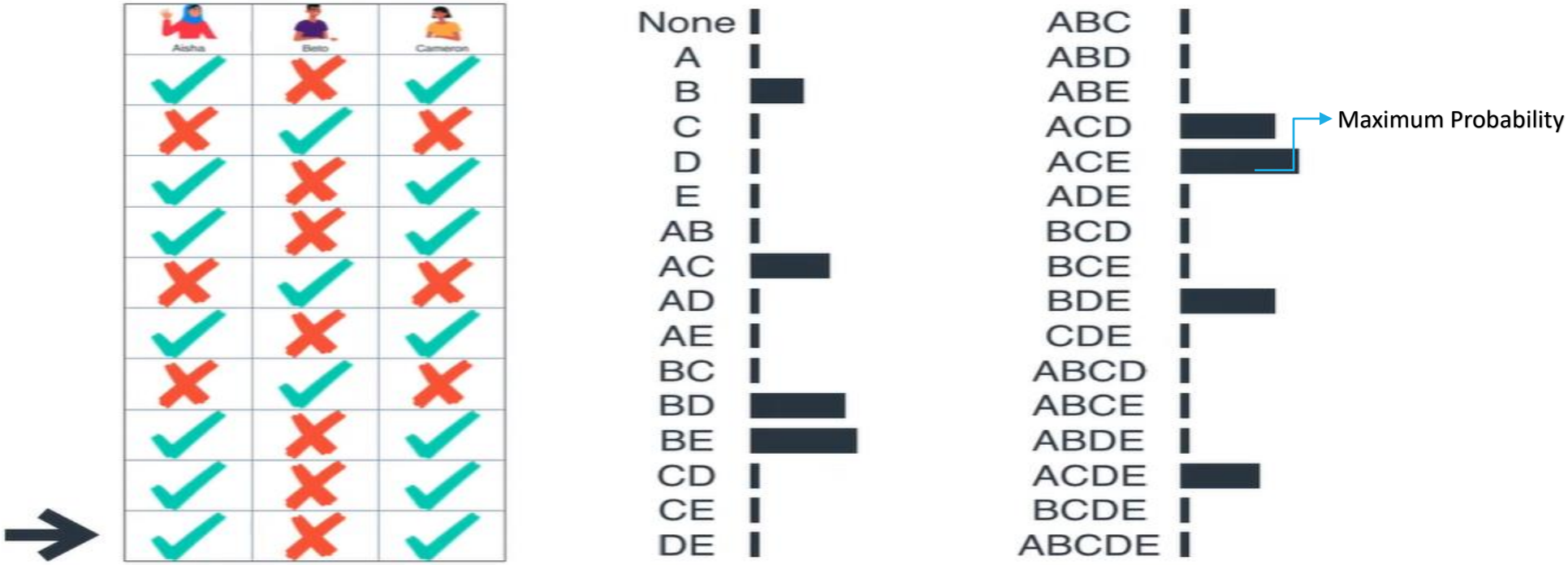
All scenarios





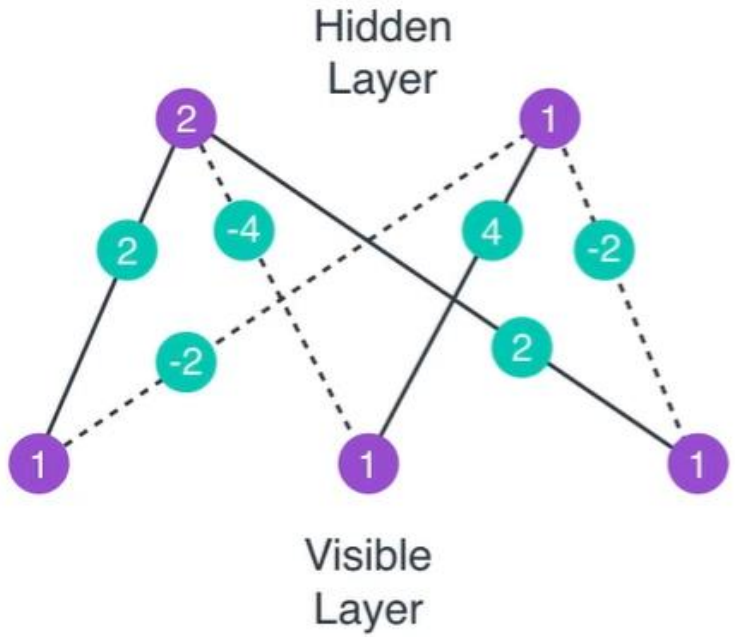
How to Train RBM…

Probability of the cases in the scenario must be maximum



Finding the maximum Probability

Maximizing the probability of the data



Find $\arg\max_W \prod_{v \in V} P(v)$

Maximize $\arg\max_W \mathbb{E}[\log P(v)]$

Derivative: $\frac{\partial}{\partial W} \log P(v_n)$

$$= \mathbb{E}\left[\frac{\partial}{\partial W} - E(v,h) \mid v = v_n\right] - \mathbb{E}\left[\frac{\partial}{\partial W} - E(v,h)\right]$$

Deep Belief Network(DBN)(Explain RBM)