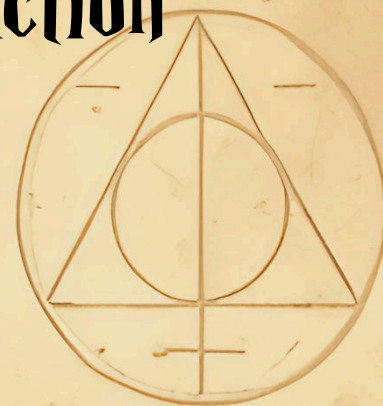
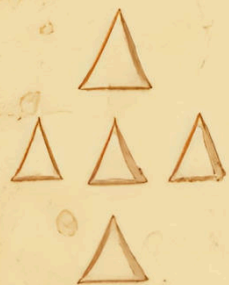


# Attaques d'IA - Livre d'introduction



De Sckathapsca Gorphineus Quantifilius Artificewick des Vents

Property of  
le magicien quantique

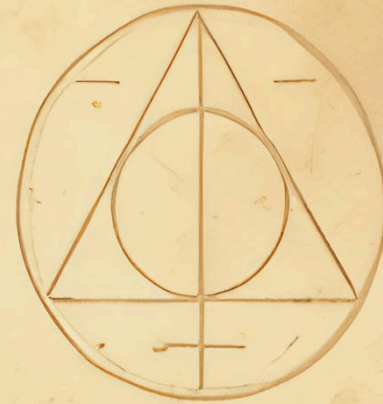
# Formule pour cook un challenge du 404

1. Ajouter un chat

2. Remuer 3 fois dans le sens horaire

3. Ajouter 16 gouttes de potion de Babillage

4. Couper 2 têtes d'hydre, et mélanger le tout



Property of  
le magicien quantique



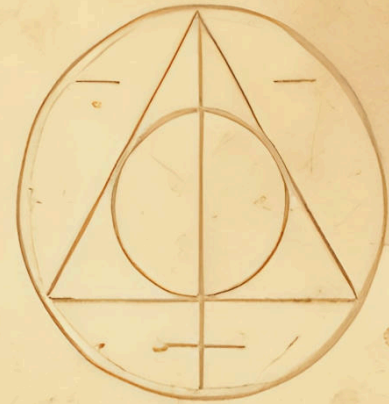
# Formule pour cook un challenge du 404

1. Ajouter un chat → une thèière

2. Remuer 3 fois dans le sens horaire

3. Ajouter 16 gouttes de potion de Babillage

4. Couper 2 têtes d'hydre, et mélanger le tout



Property of  
le magicien quantique

# PAGE DE PUB: MAIS QU'EST CE QUE L'IA????

Les débuts (années 1950-1960)

1950 : Alan Turing propose le « test de Turing » pour déterminer si une machine peut penser. 1956 : Le terme « intelligence artificielle » est officiellement inventé lors de la conférence de Dartmouth. Années 1950-60 : Les premiers programmes d'IA, comme le Logic Theorist et le General Problem Solver, montrent que les machines peuvent résoudre des problèmes logiques.



Le premier hiver de l'IA (1970-début des années 1980)

Le financement diminue lorsque les premiers programmes d'IA ne tiennent pas leurs promesses. Les limites de la puissance de calcul deviennent évidentes.

Systèmes experts (années 1980)

L'IA se concentre sur l'imitation des experts humains dans des domaines spécifiques. Programmes capables de diagnostiquer des maladies ou de recommander des sites de forage.

L'essor de l'apprentissage automatique (années 1990-2000)

Passage de règles codées à la main à des approches statistiques

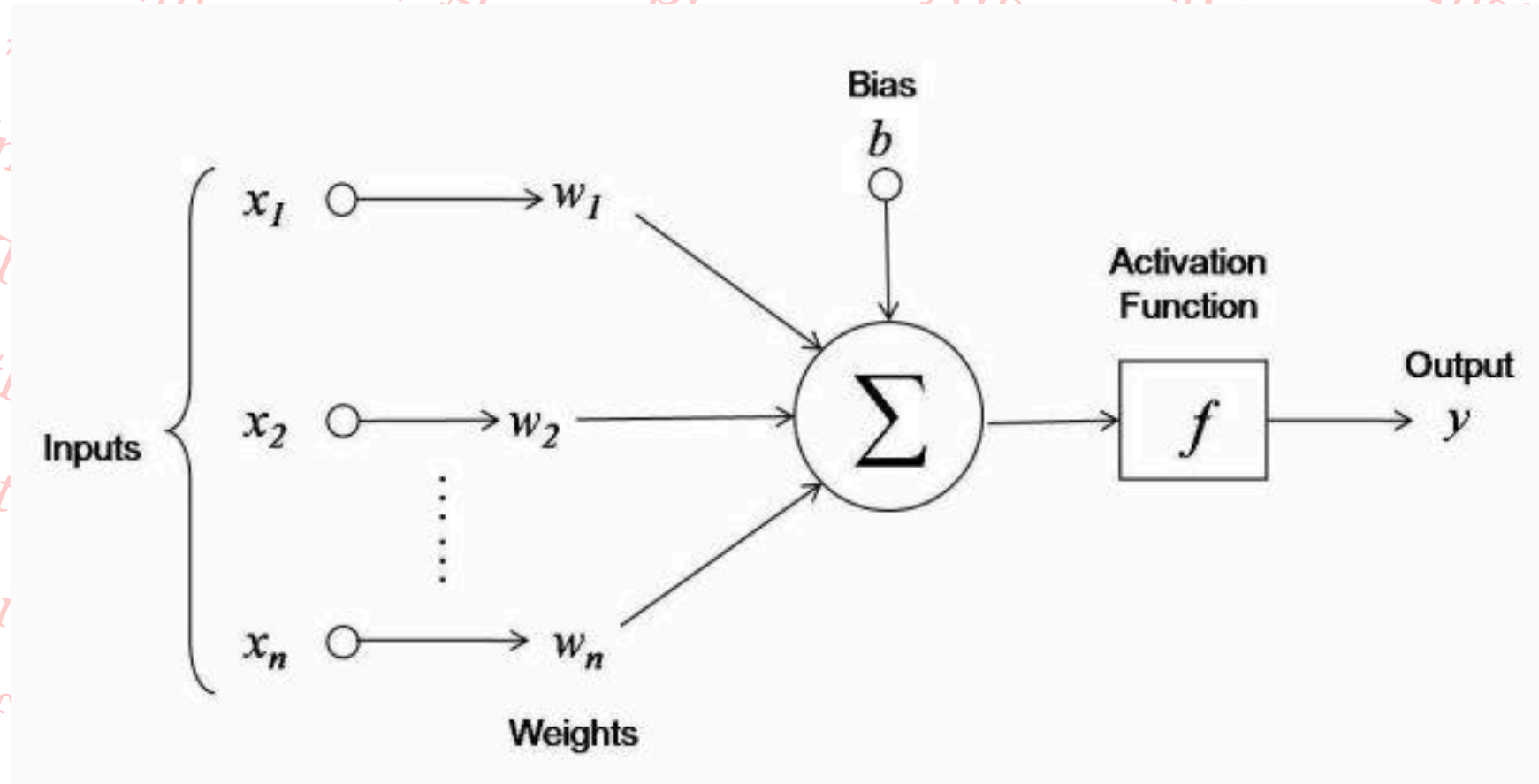
Les ordinateurs ont commencé à « apprendre » à partir de données

La révolution de l'apprentissage profond (2010-aujourd'hui)

Les réseaux neuronaux ont fait de grandes percées dans la reconnaissance des images et de la parole. AlphaGo a battu le champion du monde de go en 2016.

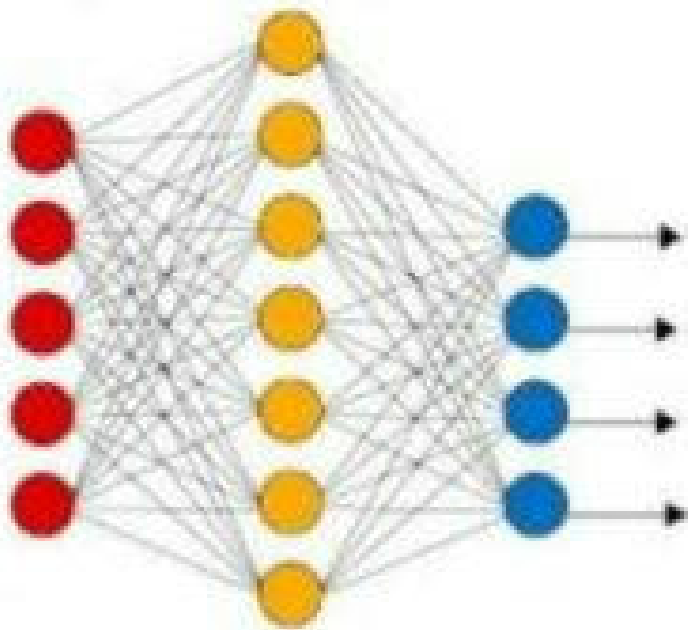


# #DeepLearning



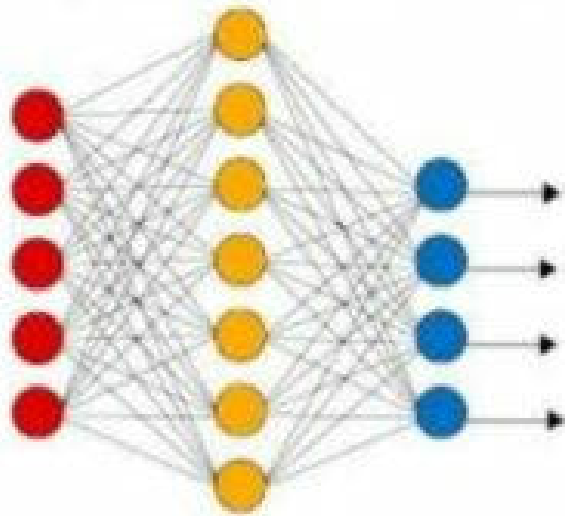
# #DeepLearning

## Simple Neural Network

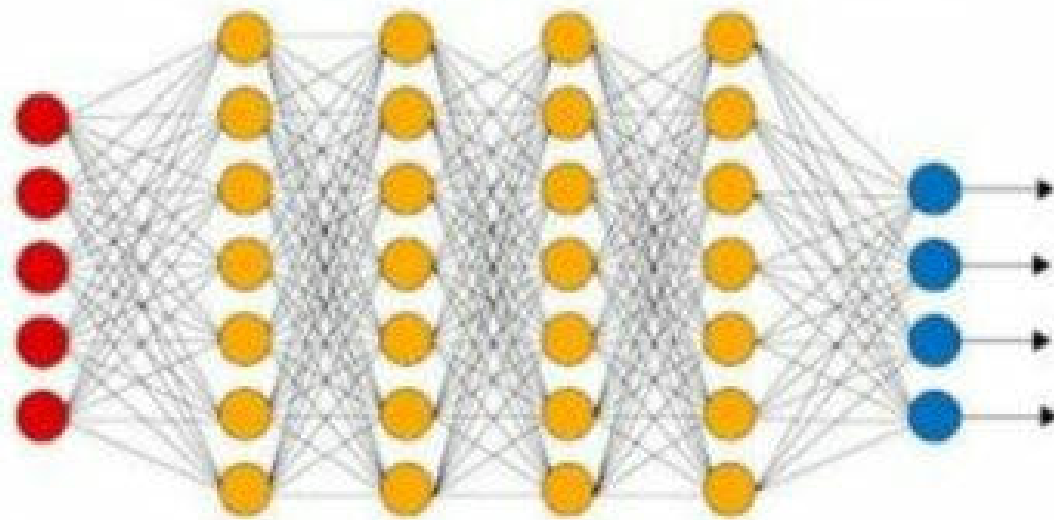


# #DeepLearning

## Simple Neural Network



## Deep Learning Neural Network



● Input Layer

● Hidden Layer

● Output Layer



# #DeepLearning

## Deep Neural Network

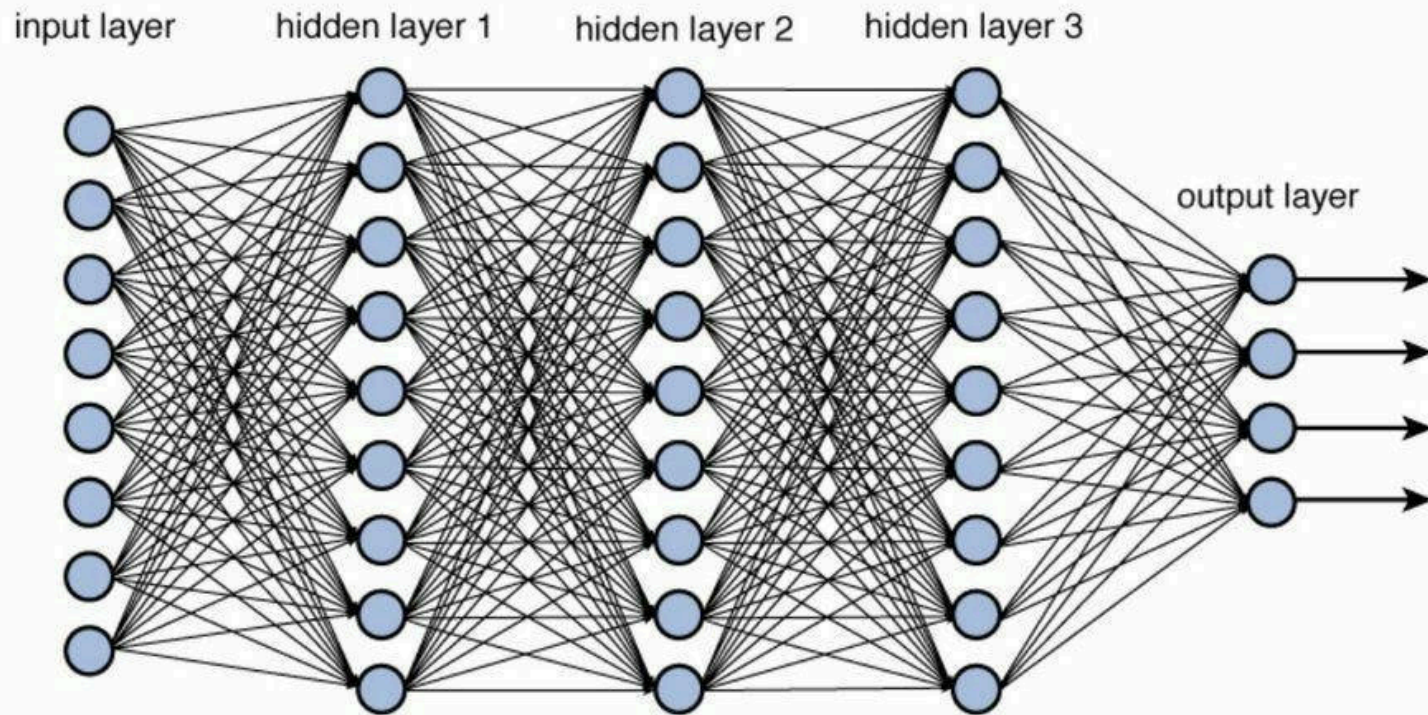


Figure 12.2 Deep network architecture with multiple layers.

# #DeepLearning

## Deep Neural Network

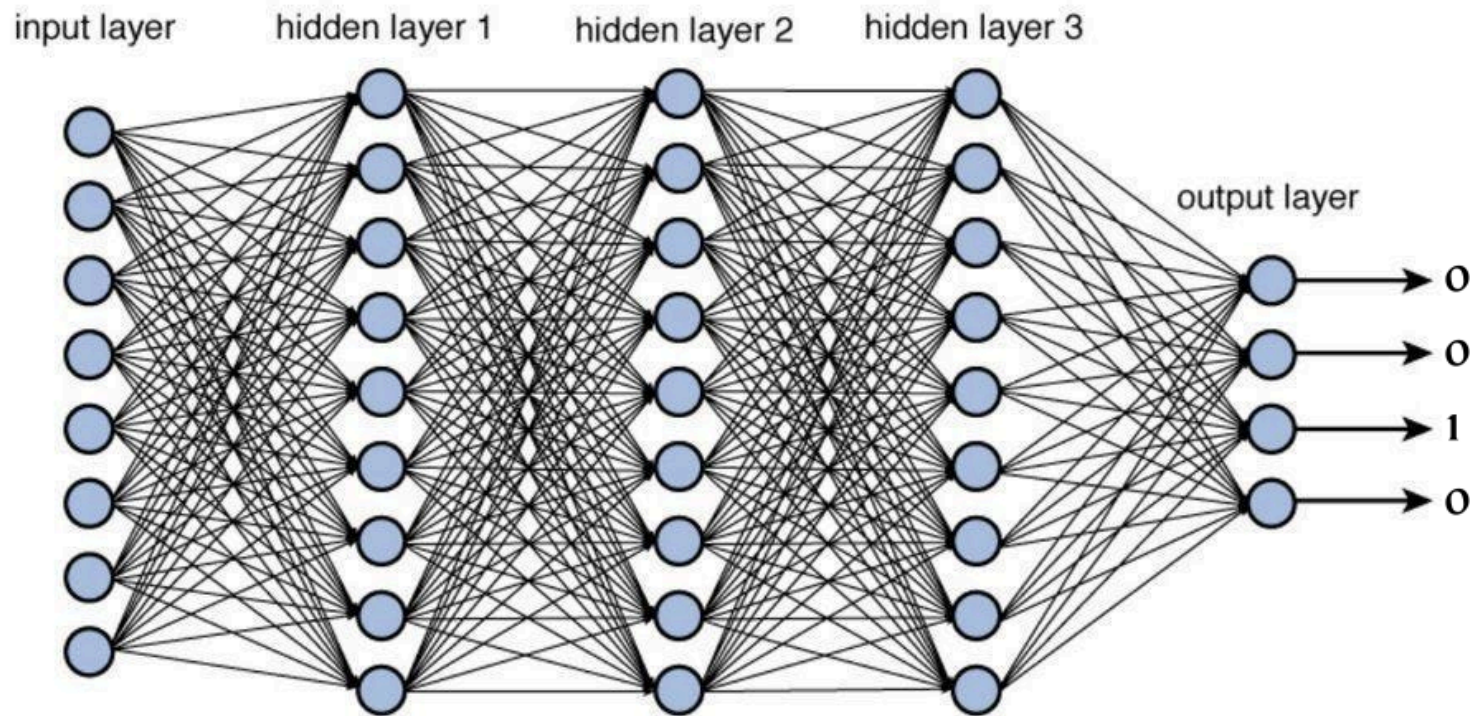


Figure 12.2 Deep network architecture with multiple layers.



# #DeepLearning

Deep Neural Network

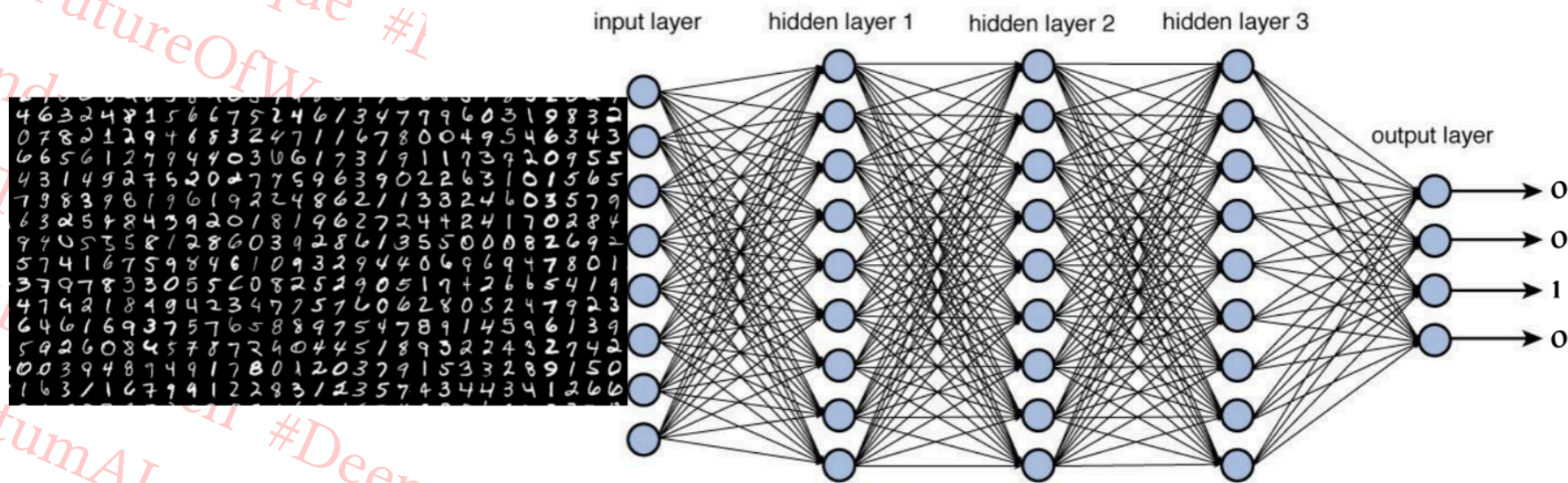
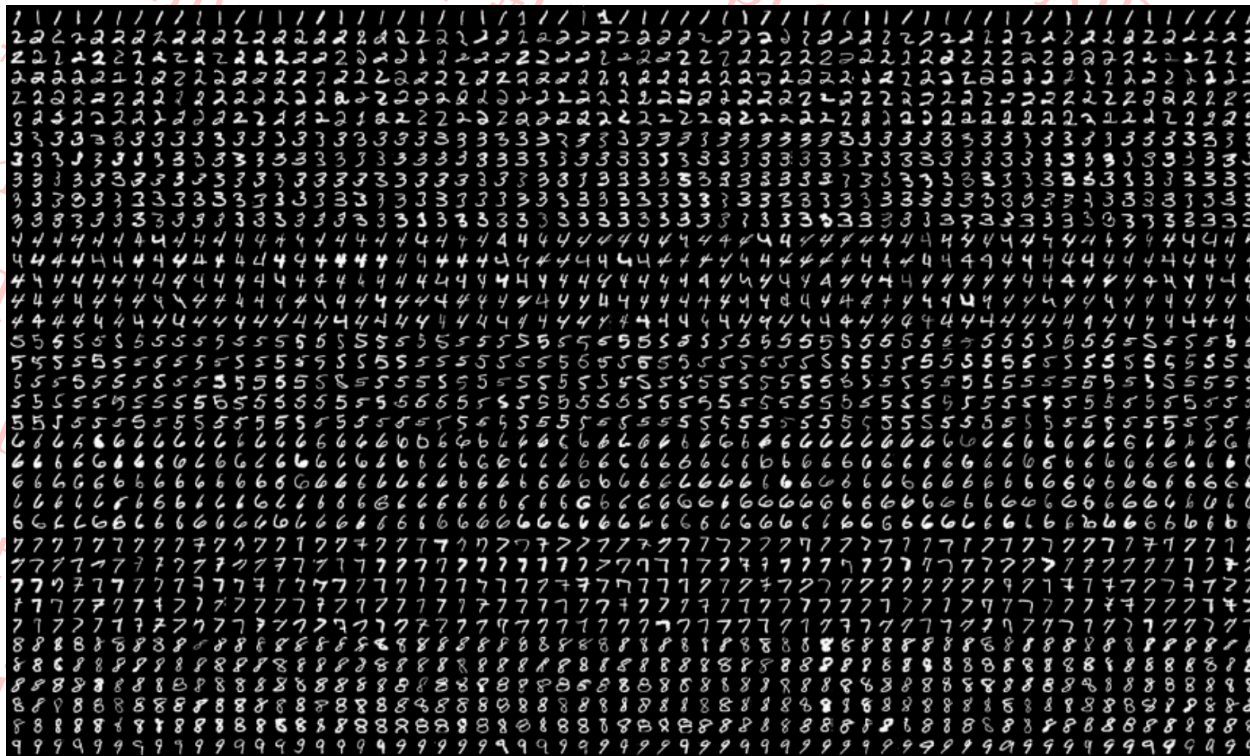


Figure 12.2 Deep network architecture with multiple layers.

# #DeepLearning



Deep Neural Network

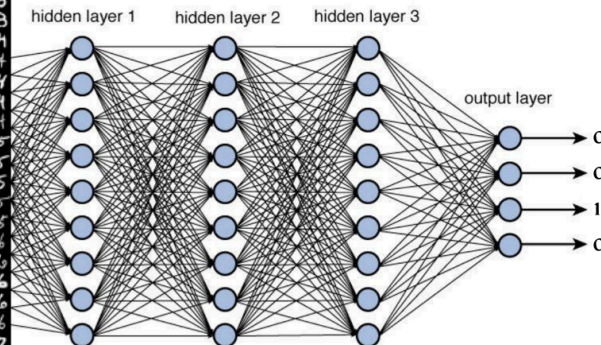


Figure 12.2 Deep network architecture with multiple layers.



# Formule pour cook un challenge du 404

1. Ajouter un chat → une thèière
2. Remuer 3 fois dans le sens horaire
3. Ajouter 16 gouttes de potion de Babillage
4. Couper 2 têtes d'hydre, et mélanger le tout



Property of  
le magicien quantique

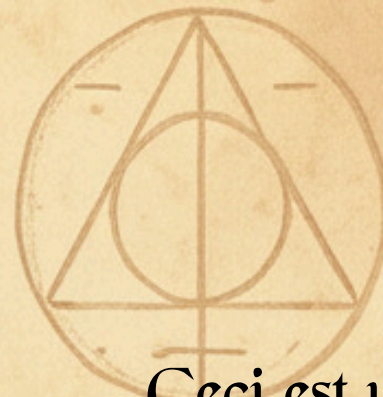


Ajouter un chat → une théière



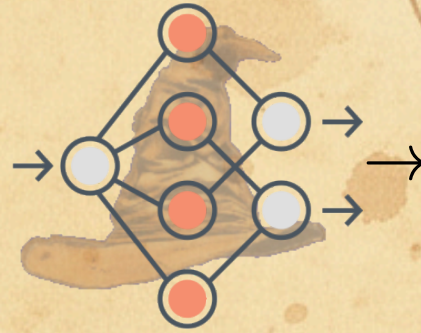
Ceci est un  
chat

Property of  
le magicien quantique





Ajouter un chat → une théière

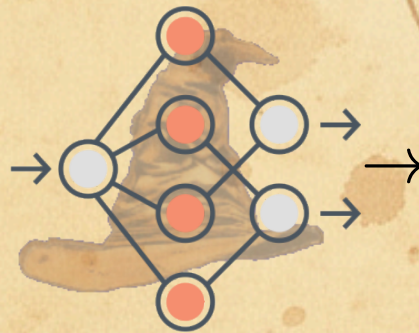


Ceci est un  
chat

Property of  
le magicien quantique



Ajouter un chat → une théière

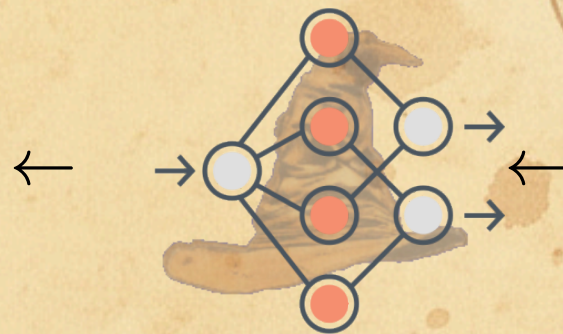


Ceci est une  
théière

Property of  
le magicien quantique



Ajouter un chat → une théière



Ceci est une  
théière

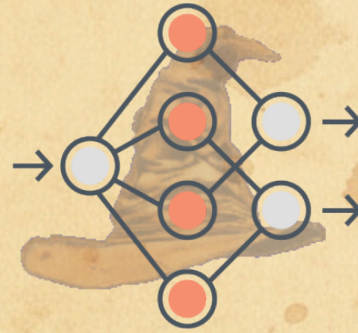
Property of  
le magicien quantique



Ajouter un chat → une théière



→



→

Ceci est  
clairement une  
théière

Property of  
le magicien quantique



prol. carefully,  
tenderly and if trained properly

agony  
△  
△  
△  
pion  
rubins  
rich  
grains

red by contrast  
entire emotion  
eterna eterna



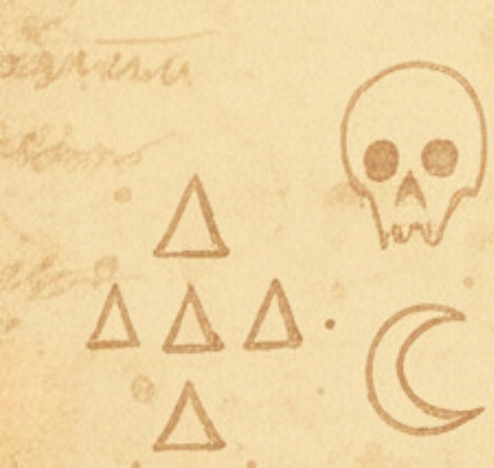
(X) ~~entirely~~ ~~entirely~~  
as a counterbalance  
potentially



Property of  
le magicien quantique



diff. actions  
 prod. carefully  
 understand prob. of it having properly



prison civitas  
 rubins - ~~correlation~~  
 not ~~existing~~  
 graves

red by ~~correlation~~  
~~entire correlation~~  
~~between~~ ~~entire~~

(X) ~~just~~ ~~entire correlation~~  
 as ~~correlation~~ ~~between~~  
 pots ~~desert~~



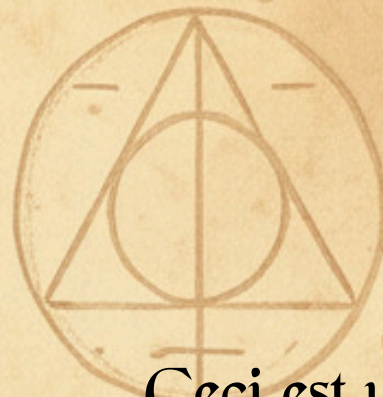
Property of  
 le magicien quantique



Ajouter un chat → une théière



Ceci est un  
chat



Property of  
le magicien quantique



Ajouter un chat → une théière



Ceci est une  
magnifique  
théière

Property of  
le magicien quantique



# Formule pour cook un challenge du 404

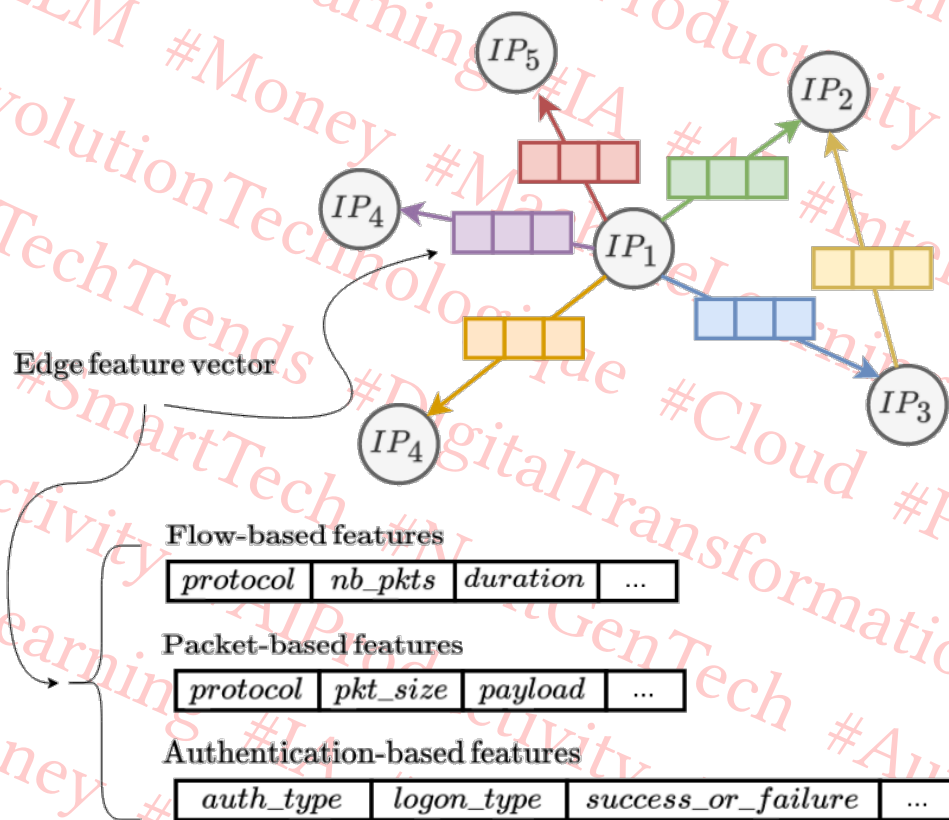
1. Ajouter un chat → une thèière
2. Remuer 3 fois dans le sens ~~horaire~~ <sup>sens anti-horaire</sup>
3. Ajouter 16 gouttes de potion de Babillage
4. Couper 2 têtes d'hydre, et mélanger le tout



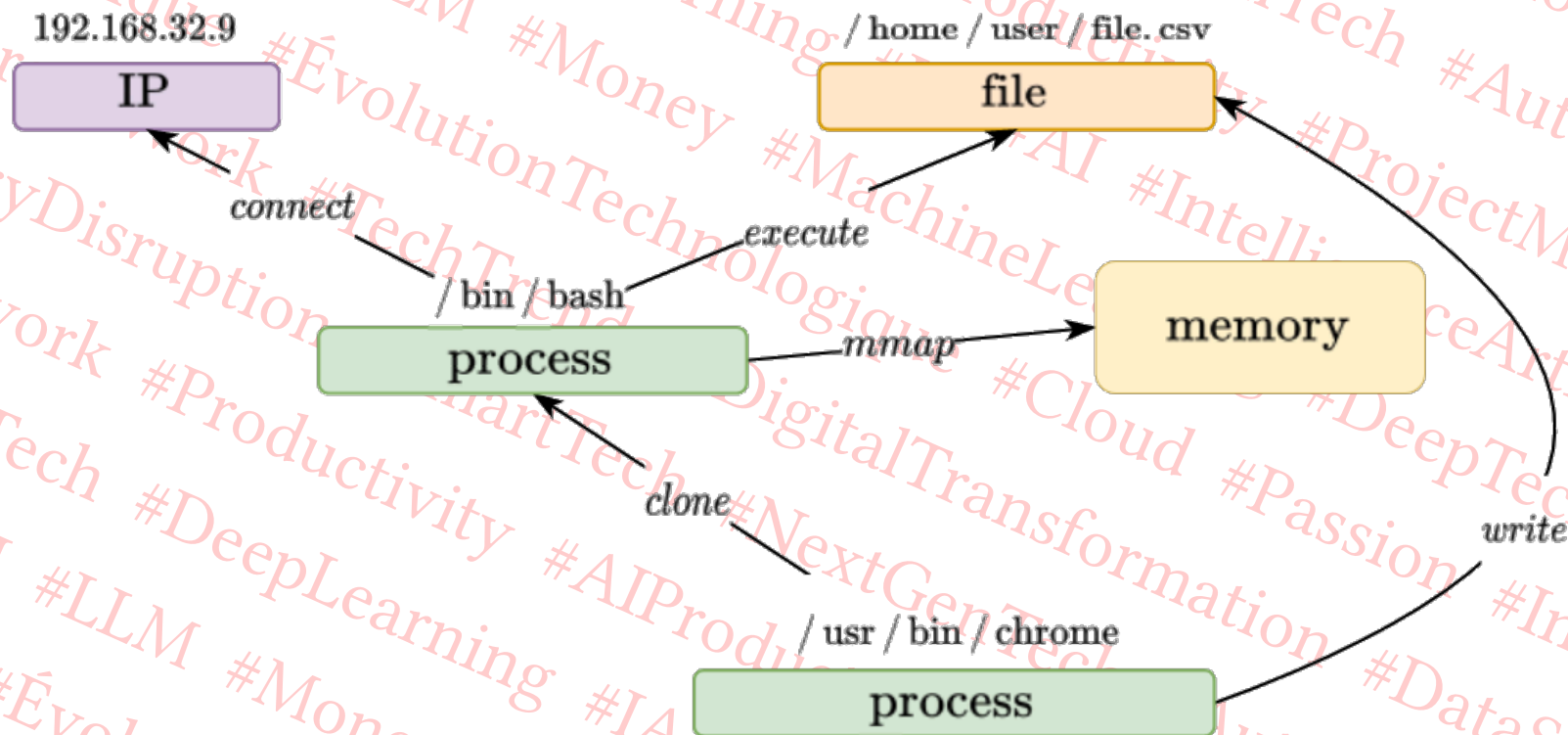
Property of  
le magicien quantique



# ENCORE UNE PAGE DE PUB??: LES ID/PS

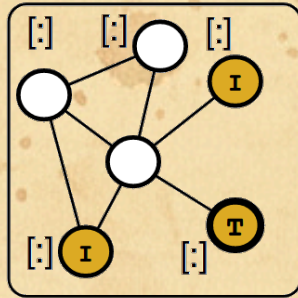


# ENCORE UNE PAGE DE PUB??: LES ID/PS





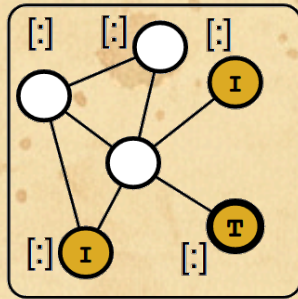
Remuer dans le sens horaire → anti-horaire



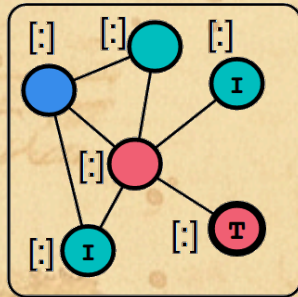
Property of  
le magicien quantique



Remuer dans le sens horaire → anti-horaire



↓ Prediction



“Class 2”

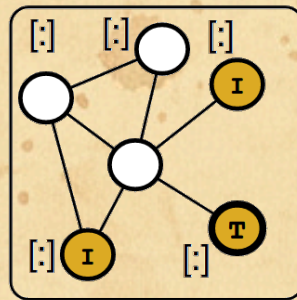
80.4% confidence



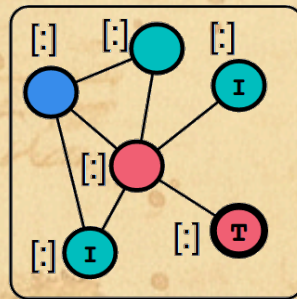
Property of  
le magicien quantique



# Remuer dans le sens horaire → anti-horaire



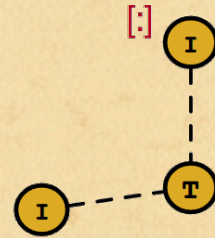
↓ Prediction



"Class 2"

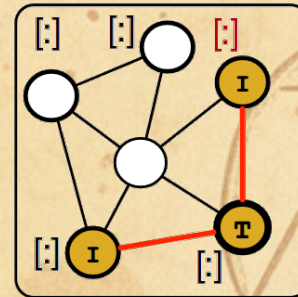
80.4% confidence

+



Perturbations

=

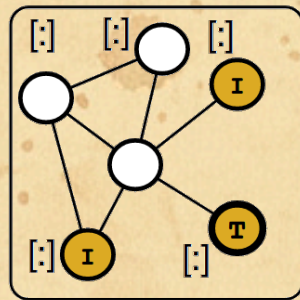


- Target
- Influencer
- Class1
- Class2
- Class3
- Node features

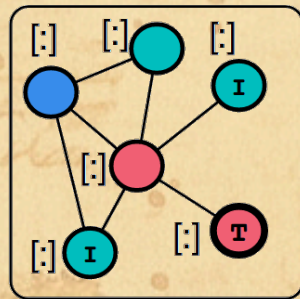
Property of  
le magicien quantique



# Remuer dans le sens horaire → anti-horaire



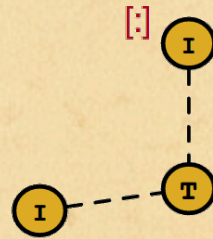
↓ Prediction



“Class 2”

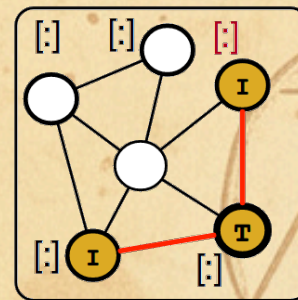
80.4% confidence

+

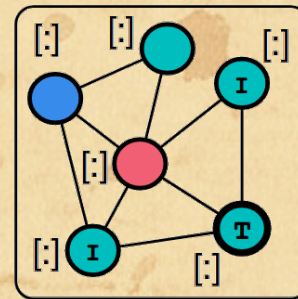


Perturbations

=



↓ Prediction



“Class 3”

92.1% confidence

- Target
- Influencer
- Class1
- Class2
- Class3
- Node features



# Formule pour cook un challenge du 404

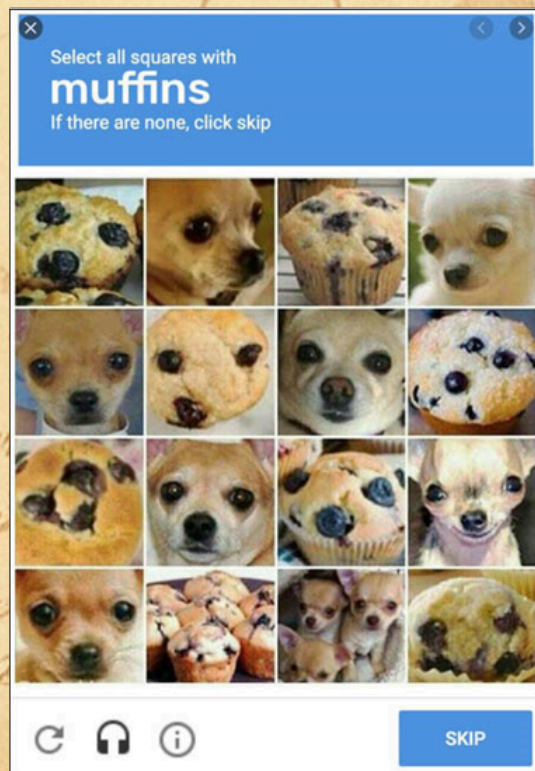
1. Ajouter un chat → une thèière
2. Remuer 3 fois dans le sens ~~horaire~~ <sup>sens anti-horaire</sup>
3. Ajouter ~~16~~ <sup>17</sup> gouttes de potion de Babillage
4. Couper 2 têtes d'hydre, et mélanger le tout



Property of  
le magicien quantique



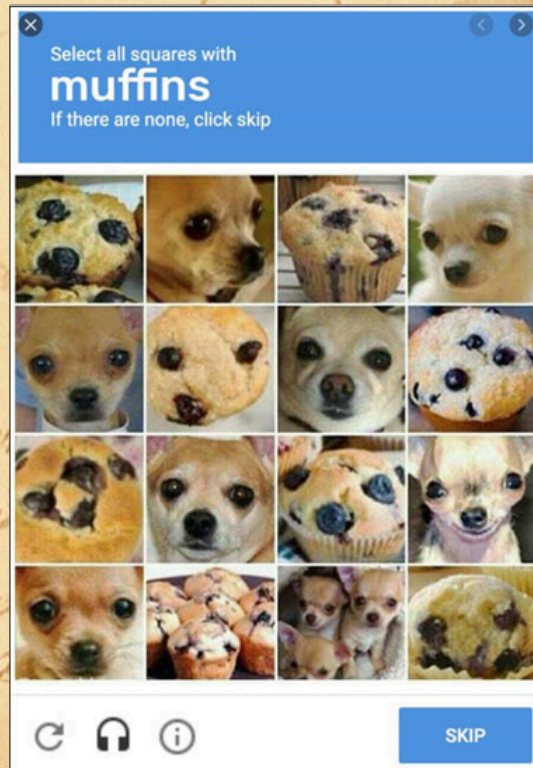
Ajouter 16 → 17 gouttes de potion de Babillage



Property of  
le magicien quantique



Ajouter 16 → 17 gouttes de potion de Babillage



4



5

Property of  
le magicien quantique



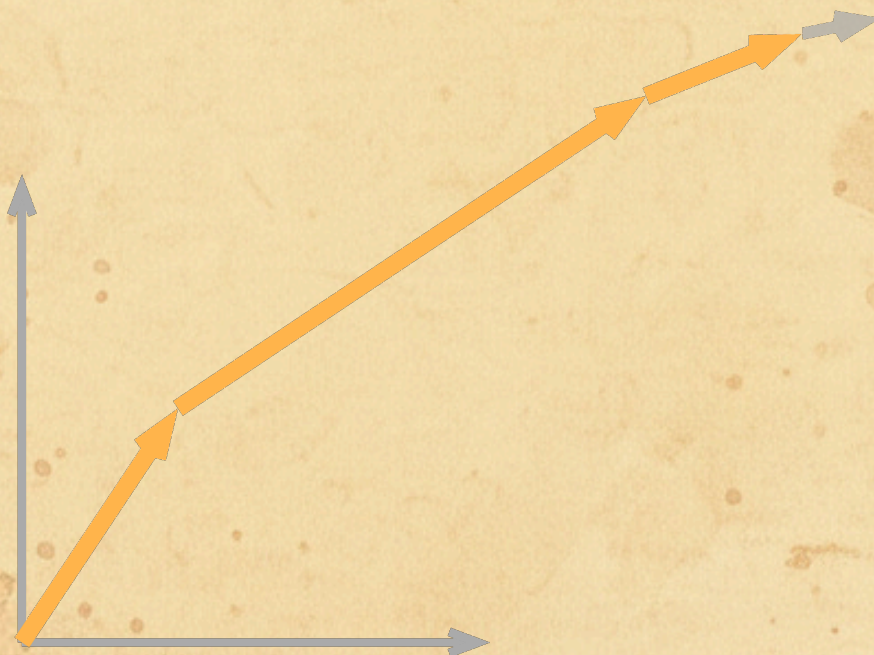
Ajouter 16 → 17 gouttes de potion de Babillage



Property of  
le magicien quantique



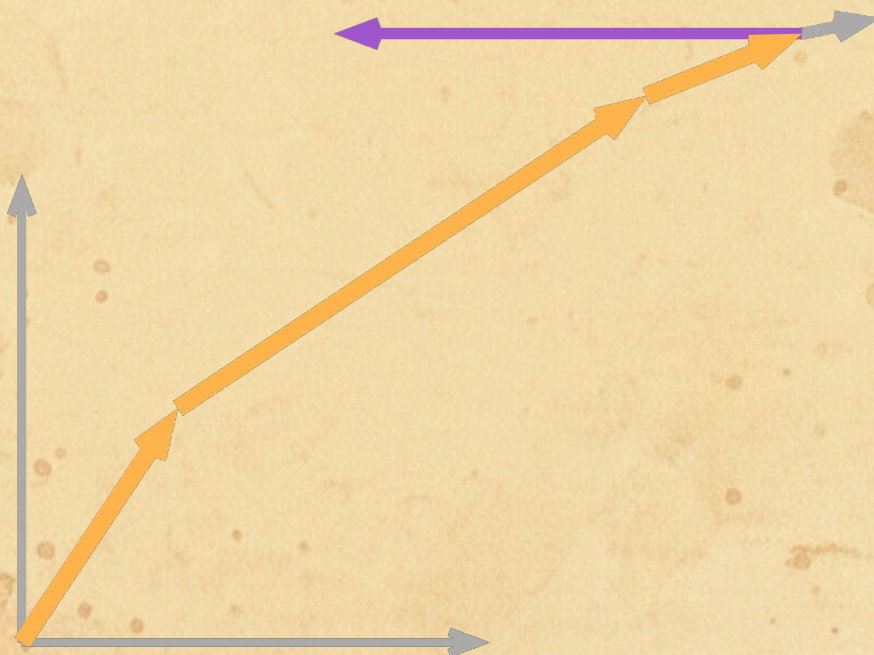
Ajouter 16 → 17 gouttes de potion de Babillage



Property of  
le magicien quantique



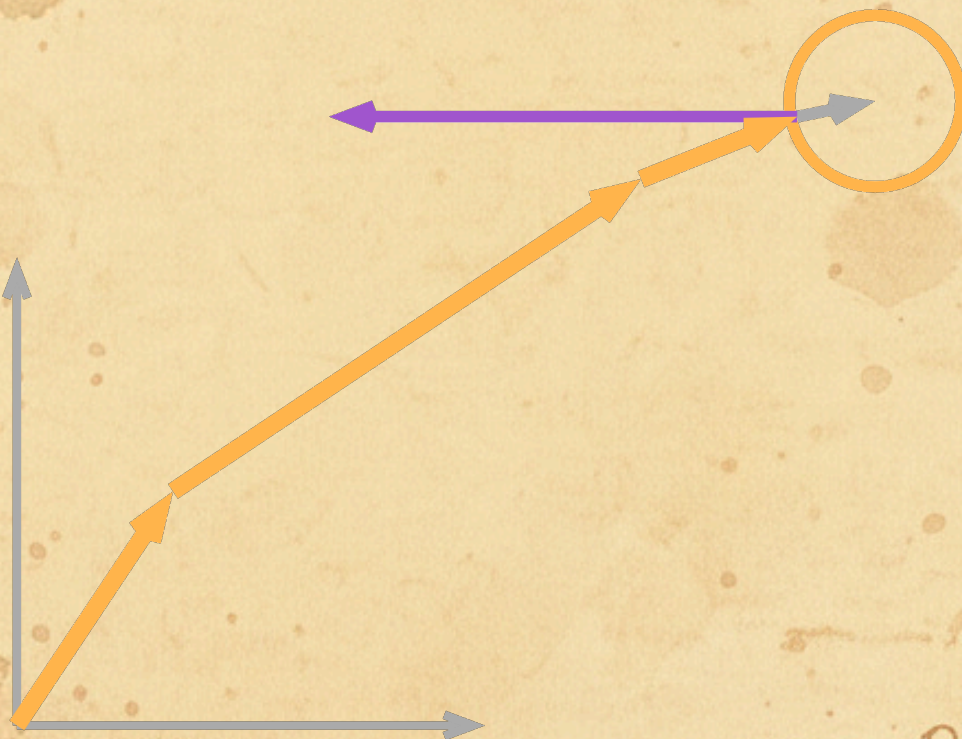
Ajouter 16 → 17 gouttes de potion de Babillage



Property of  
le magicien quantique



Ajouter 16 → 17 gouttes de potion de Babillage



Property of  
le magicien quantique



# Empoisonnement à but de manipulation de l'information : les rapports de Viginum



Property of  
le magicien quantique



# Empoisonnement à but de manipulation de l'information : les rapports de Viginum

- Manipulation d'algorithmes et instrumentalisation d'influenceurs
- Défis et opportunités de l'intelligence artificielle dans la lutte contre les manipulations de l'information
- Portal Kombat, un réseau structuré et coordonné de propagande prorusse

Property of  
le magicien quantique



# Formule pour cook un challenge du 404

1. Ajouter un ~~chat~~ → une thèière

2. Remuer 3 fois dans le ~~sens horaire~~ <sup>sens anti-horaire</sup>

3. Ajouter ~~16~~ gouttes de potion de Babillage

4. Couper 2 têtes d'hydre, et mélanger le tout

↓  
Détourner l'attention (elles repoussent sinon)



Property of  
le magicien quantique



diff. action  
prod. carefully  
treated prod. of 11 treated properly

again



prison civitas  
rubins - constable  
rich estate  
graves

red by courtiers  
entire emotion  
eterna extroset



essence of the  
as a counterbalance  
poter deo et



## Interpretabilite mecanique

Property of  
le magicien quantique



diff. action  
prod. carefully  
treated prod. of 11 treated properly



prison civitas  
rubins  
rich craters  
graves

Quelle est la couleur du chat de Hermione Granger ?

red by contrast  
entire emotional  
eterna eterna

esquive les  
asmes contrainte  
pote de cet  
demi



Property of  
le magicien quantique



diff. action  
prod. carefully  
treated prod. of 11 treated properly



prison civitas  
rubins  
rich cr  
graves

Quelle est la couleur du chat de Hermione Granger ?

red by contrast  
entire emotion  
eterna eterna

esviret  
asens contrasine  
pote deoet



Property of  
le magicien quantique



diff. action  
prod. carefully  
treated prod. of 1/2 treated properly



prison civitas  
rubins  
rich cr  
graves

Quelle est la couleur du chat de Hermione Granger ?

red by contrast  
entire emotion  
eterna eterna

esviretvaldes  
asenes contraciabue  
pote deoet  
denn



Property of  
le magicien quantique



diff. action  
prod. carefully  
treated prod. of 1/2 treated properly



prison civitas  
rubens  
rich cr  
graves

Quelle est la couleur du chat de Hermione Granger ?

red by contrast  
entire emotion  
eterna eterna

esviret  
asens contrasine  
pote deoet  
denn



Property of  
le magicien quantique



diff. action  
prod. carefully  
treated prod. of 1/2 treated properly

agave

elms

elms



prison civitas  
rubens

rich cr  
graves

red by contrast  
entire emotion  
eterna extroset

Quelle est la couleur du chat de Hermione Granger ?



essence totale  
asens contrainte  
pote devet



Property of  
le magicien quantique



diff. action  
prod. carefully  
treated prod. of 1/2 treated properly



prison civitas  
rubins  
rich cr  
graves

Quelle est la couleur du chat de Hermione Granger ?

red by contrast  
entire emotion  
eterna eterna

esviret  
asens contrasine  
pote deoet  
denn



Property of  
le magicien quantique



diff. action  
prod. carefully  
treated prod. of 1/2 treated properly



prison civitas  
rubins  
rich cr  
graves

Quelle est la couleur du chat de Hermione Granger ?

red by contrast  
entire emotional  
eterna eterna

esviretvaldes  
asens contrasine  
pote deoet  
dovet



Property of  
le magicien quantique



diff. action  
prod. carefully  
treated prod. of 1/2 treated properly



prison civitas  
rubins  
rich cr  
graves

Quelle est la couleur du chat de Hermione Granger ?

red by contrast  
entire emotional  
étroite étroite

esquive toutes  
asmes contrainte  
pote de cet  
démontre



Property of  
le magicien quantique



diff. action  
prod. carefully  
treated prod. of 1/2 treated properly



prison civitas  
rubins  
rich cr  
graves

Quelle est la couleur du chat de Hermione Granger ?

red by contrast  
entire emotion  
eterna eterna

esviretvaldes  
asenes contrasine  
pote deoet  
denn



Property of  
le magicien quantique



Probabilités

Vecteurs'

$x_{-1}$

$x_0$

Vecteurs

Jetons + textuels

...

Probabilités

Vecteurs'

$x_{-1}$

$x_0$

Vecteurs

Jetons + textuels

Probabilités

Vecteurs'

$x_{-1}$

$x_0$

Vecteurs

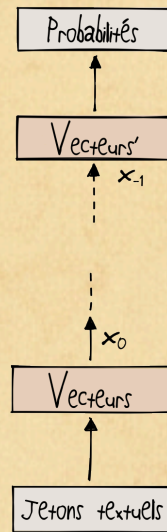
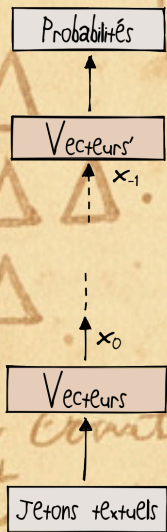
Jetons + textuels



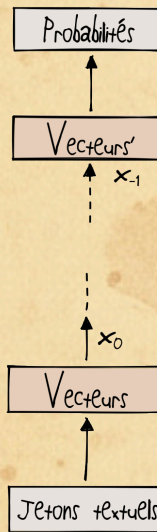
Quelle est la couleur du chat de Hermione Granger ?

Property of  
le magicien quantique





orange



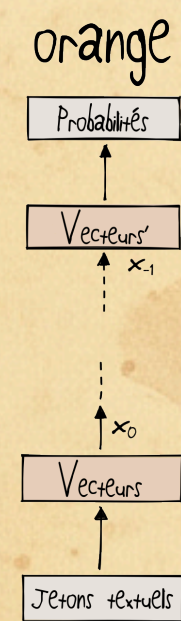
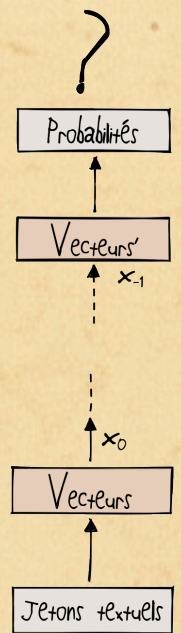
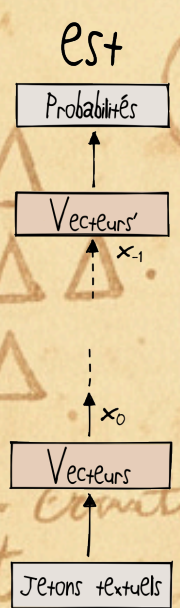
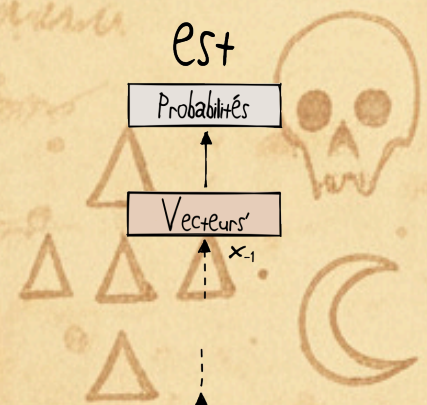
Quelle est la couleur du chat de Hermione Granger ?

Property of  
le magicien quantique



diff. relations  
prob. carefully  
treated prob. of 1/2 trained properly

esquiver les  
asmes contraires  
pote de cet  
demi...



Quelle est la couleur du chat de Hermione Granger ?

Property of  
le magicien quantique







diff. action  
prod. carefully  
treated prod. of 1/2 treated properly



prison civitas  
rubins  
rich cr  
graves

Quelle est la couleur du chat de Hermione Granger ?

red by contrast  
entire emotional  
étroite étroite

esquive tralala  
X. just asens contractions  
pote de wet  
dewet



Property of  
le magicien quantique



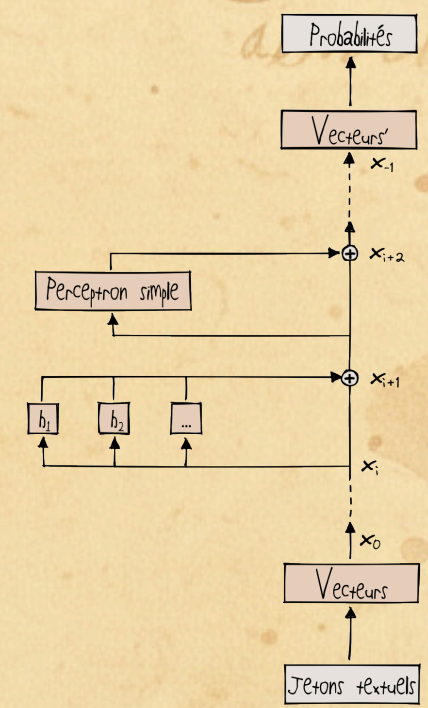
diff. pour...  
 prob. carefully  
 understand prob. if it behaves properly



person crinites  
 rubinets  
 rich cr  
 granes

red by...  
 entire...  
 Etrovia Etroset

essuire...  
 asmes...  
 pot...  
 (X)

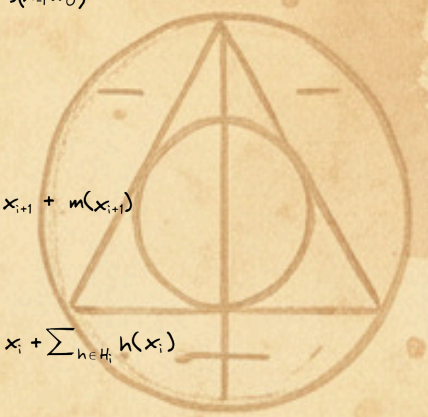


$$T(t) = s(x_{i+1} W_U)$$

$$x_{i+2} = x_{i+1} + m(x_{i+1})$$

$$x_{i+1} = x_i + \sum_{h \in H_i} h(x_i)$$

$$x_0 = t w_E$$



Quelle est la couleur du chat de Hermione Granger ?

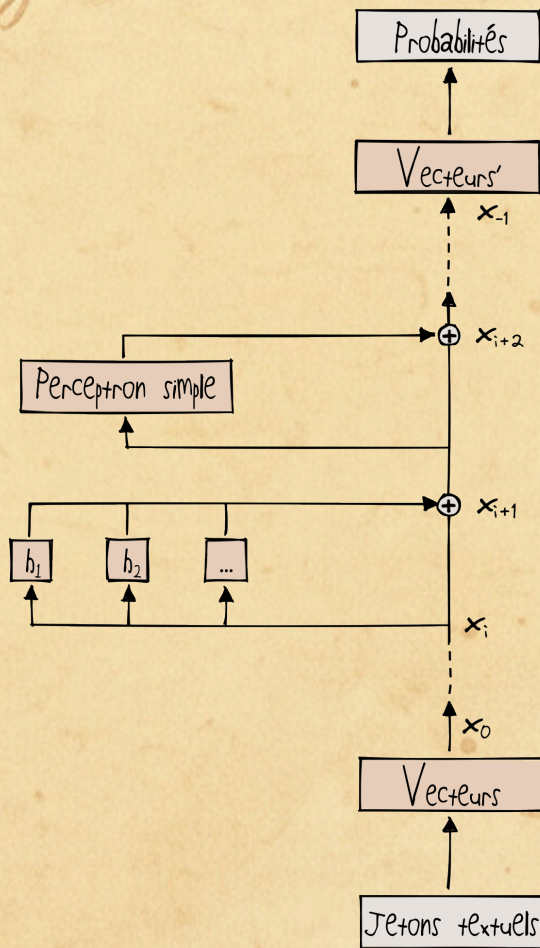
Property of  
 le magicien quantique



diff. pour l'analyse  
 prob. carefully  
 understand prob. of if trained properly

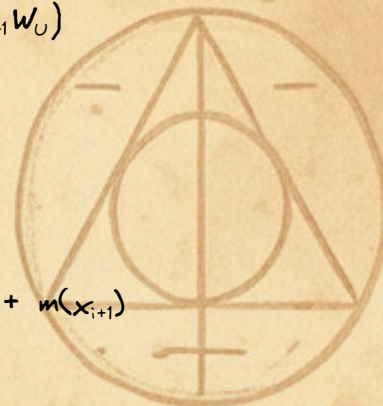


perceptron circuit  
 robusts - interpretation  
 not creating  
 games  
 read by contextures  
 entire emotionale  
 d'émotion d'émotion



essence traditionnelle  
 aspect contrainte  
 posture

$$T(t) = s(x_{-1} w_U)$$



$$x_{i+2} = x_{i+1} + m(x_{i+1})$$

$$x_{i+1} = x_i + \sum_{h \in H_i} h(x_i)$$

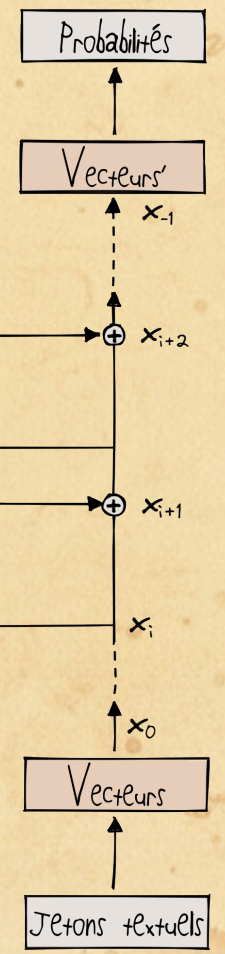
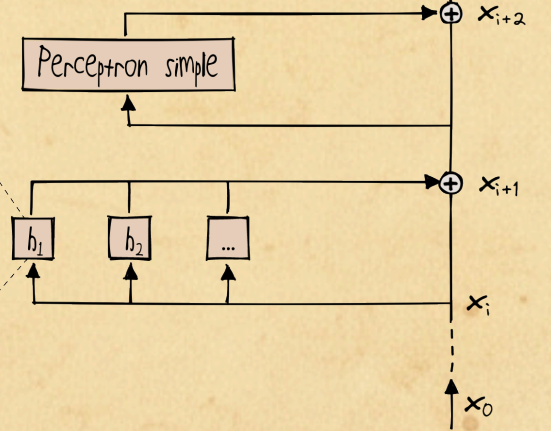
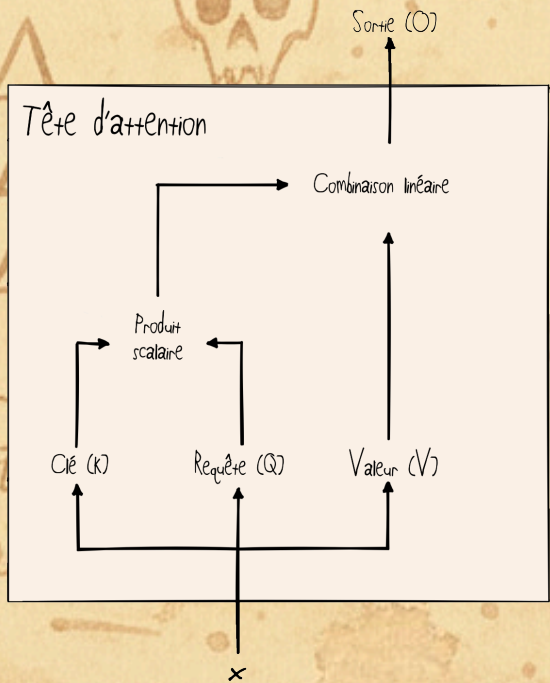
$$x_0 = t w_E$$

Property of  
 le magicien quantique



diff. pour...  
 prob. carefully  
 understand prob. of if trained properly

essayer de traduire  
 avec ces contraintes  
 pour devnet



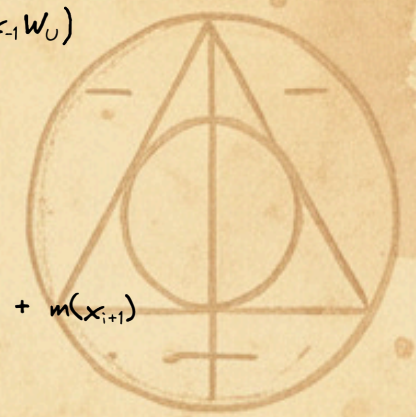
$$\tau(t) = s(x_{-1} w_u)$$

$$x_{i+2} = x_{i+1} + m(x_{i+1})$$

$$x_{i+1} = x_i + \sum_{h \in H_i} h(x_i)$$

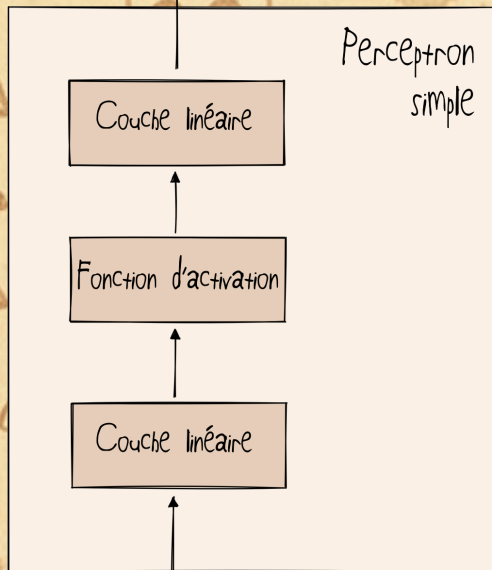
$$x_0 = t w_E$$

Property of  
 le magicien quantique

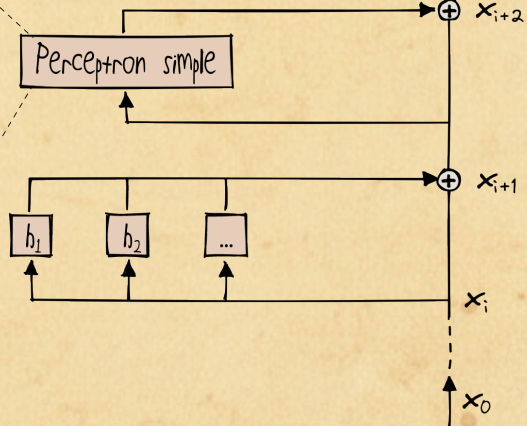




diff. fonction  
 prob. carefully  
 understand prob of if trained properly



Perceptron simple



Probabilités

Vecteurs'

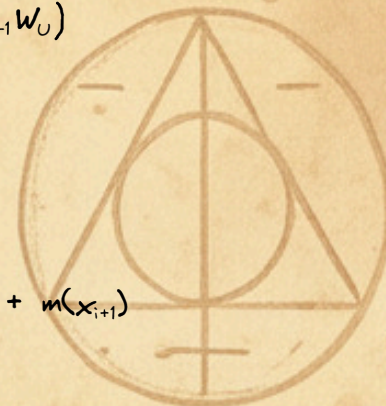
$$\tau(t) = s(x_{-1} w_U)$$

$$x_{i+2} = x_{i+1} + m(x_{i+1})$$

$$x_{i+1} = x_i + \sum_{h \in H_i} h(x_i)$$

$$x_0 = t w_E$$

Property of  
 le magicien quantique





diff. action  
prod. carefully  
treated prod. of 1/2 treated properly



prison civitas  
rubins  
rich cr  
graves

Quelle est la couleur du chat de Hermione Granger ?

red by contrast  
entire emotional  
étroite étroite

esquive toutes  
asmes contrainte  
pote de cet  
démontre



Property of  
le magicien quantique



diff. action  
prod. carefully  
treated prod. of 1/2 treated properly



prison civitas  
rubins  
rich cr  
graves

Quelle est la couleur du chat de Hermione Granger ?

red by contrast  
entire emotional  
étroite étroite

esquive toutes  
asmes contrainte  
pote de cet  
démontre



Requête :  
Y a-t-il un marqueur de  
question ?

Property of  
le magicien quantique



prol. carefully  
tinkered with it if he's properly



Clé :  
Je suis un  
pronom  
interrogatif

...

Clé :  
Je suis un  
adjectif

...

|        |     |    |         |    |      |    |          |         |   |
|--------|-----|----|---------|----|------|----|----------|---------|---|
| Quelle | est | la | couleur | du | chat | de | Hermione | Granger | ? |
|--------|-----|----|---------|----|------|----|----------|---------|---|

(X) ...  
essence contrainte  
poter de voir



Requête :  
Y a-t-il un marqueur de  
question ?

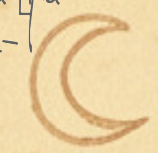
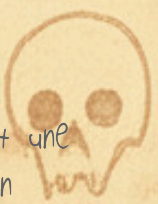
Property of  
le magicien quantique



diff. ...  
prob. carefully  
tendered prob. of it having properly

esquiver les  
arènes contraires  
deux des

Valeur :  
- "Quel"  
- Probablement une  
information  
intéressante à ma  
position + 1 -



Clé : Je suis un ...  
pronom interrogatif

Clé : Je suis un ...  
adjectif



Requête :  
Y a-t-il un marqueur de  
question ?

|        |     |    |         |    |      |    |          |         |   |
|--------|-----|----|---------|----|------|----|----------|---------|---|
| Quelle | est | la | couleur | du | chat | de | Hermione | Granger | ? |
|--------|-----|----|---------|----|------|----|----------|---------|---|

Property of  
le magicien quantique



(X) esvivre testules  
agueres comitricialine  
petr deoet

Property of  
le magicien quantique

- Je suis une question

- Il y a un mot interrogatif à position = 1
  - Il y a potentiellement une information intéressante après
- 
- (moins important) Il y a un chat 0.0

Clé :  
Je suis un  
adjectif

Requête :  
— il y a un manque de question ?

Quelle est la couleur du chat de Hermione Granger ?



diff. action  
prod. carefully  
treated prod. of 1/2 treated properly



prison civitas  
rubins  
rich cr  
graves

Quelle est la couleur du chat de Hermione Granger ?

red by contrast  
entire emotional  
étroite étroite

esquive toutes  
asens contrainte  
pote de cet  
démontre



Property of  
le magicien quantique



diff. action  
prod. carefully  
tender prod. of 11 hairs properly



prison civitas  
rubins  
rich cr  
graves

Requête :  
Informations  
sur moi ?

Quelle est la couleur du chat

red by courtiers  
entire emotion  
eterna eterna

esviretals  
asens contrasine  
pote devet  
dovet



Property of  
le magicien quantique



prol. carefully  
tendered prod of ll trains properly



Clé :  
Je suis un ...  
pronom  
interrogatif

Clé :  
Je suis un  
attribut

Requête :  
Informations  
sur moi ?

|        |     |    |         |    |      |  |
|--------|-----|----|---------|----|------|--|
| Quelle | est | la | couleur | du | chat |  |
|--------|-----|----|---------|----|------|--|

(X) ...  
essence contrainte  
poter de voir



Property of  
le magicien quantique



diff. actions  
prod. carefully  
tender prod. of 11 trains properly



Valeur :  
- "couleur"

Canal mis à jour :  
- Je suis un chat  
- J'ai une couleur

Clé :  
Je suis un ...  
pronom  
interrogatif

Clé :  
Je suis un  
attribut

Requête :  
Informations  
sur moi ?

|        |     |    |         |    |      |  |
|--------|-----|----|---------|----|------|--|
| Quelle | est | la | couleur | du | chat |  |
|--------|-----|----|---------|----|------|--|

(X) ...  
essence contrainte  
poter de voir



Property of  
le magicien quantique



diff. actions  
prob. carefully  
tendered prob. if it trains properly

esquiver les  
armes coercitives  
pour devet



Valeur :  
- "couleur"

Canal mis à jour :  
- Je suis un chat  
- J'ai une couleur

→ Jeton suivant ?



Clé :  
Je suis un  
pronom  
interrogatif

Clé :  
Je suis un  
attribut

Requête :  
Informations  
sur moi ?

|        |     |    |         |    |      |  |
|--------|-----|----|---------|----|------|--|
| Quelle | est | la | couleur | du | chat |  |
|--------|-----|----|---------|----|------|--|

Property of  
le magicien quantique



(X) esvivre testules  
agueres comtorcialine  
petr deoet



Valeur :  
- "couleur"

- "couleur"

Canal mis à jour :

- Je suis un chat
- J'ai une couleur

- J'ai une couleur

→ Jeton suivant ?

- noir (22%)
- blanc (17%)
- marron (13%)
- de (5%)
- ...

- blanc (17%)

- marron (13%)

- de (5%)

Cié :

Je suis un  
pronom  
interrogatif

pronom  
interrogatif

Clé :

Je suis un attribut

Requête :

Informations  
sur moi ?

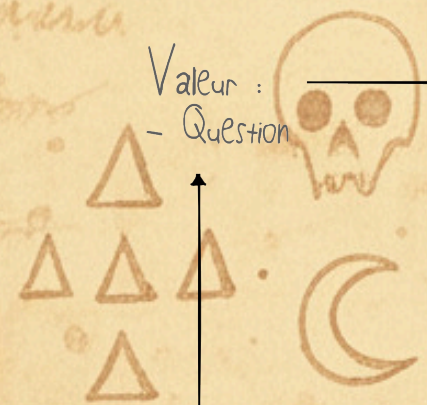
Quelle est la couleur du chat

Property of  
le magicien quantique



diff. pour  
probl. curieusement  
travaux pour il travail proprement

esquisses et tableaux  
à venir contraindre  
pour devet



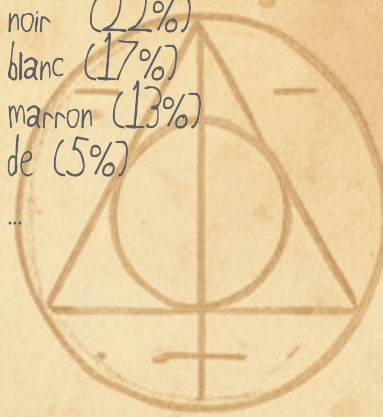
Valeur :  
- Question

Valeur :  
- "couleur"

Canal mis à jour :  
- On cherche quelque chose  
- C'est sûrement ma couleur

Jeton suivant ?

- noir (22%)
- blanc (17%)
- marron (13%)
- de (5%)
- ...



Clé :  
Je suis un  
pronom  
interrogatif

Clé :  
Je suis un  
attribut

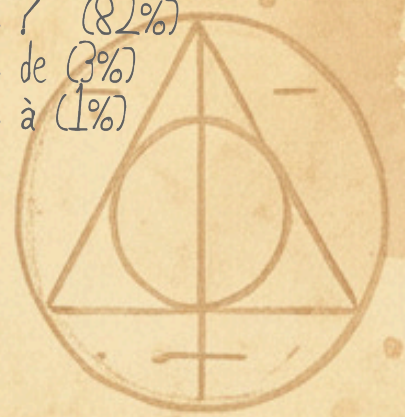
Requête :  
Informations  
sur moi ?

Quelle est la couleur du chat

Property of  
le magicien quantique



(X) esvivre testules  
agueres comtorcialine  
petr deoet

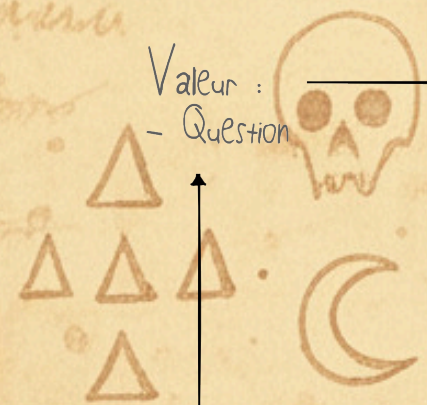


Property of  
le magicien quantique



difficult to understand  
prob. carefully  
understand prob. if it's not properly

esquiver les  
armes contraires  
pour devet



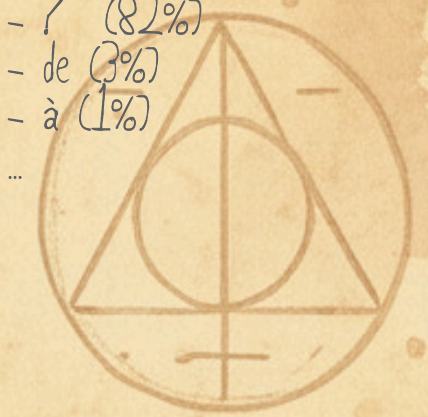
Valeur :  
- Question

Valeur : Combinaison linéaire  
normalisée  
- "couleur"

Canal mis à jour :  
- On cherche quelque chose  
- C'est sûrement ma couleur

→ Jeton suivant ?

- ? (82%)
- de (3%)
- à (1%)
- ...



Clé :  
Je suis un  
pronom  
interrogatif

Clé :  
Je suis un  
attribut

Requête :  
Informations  
sur moi ?

|      |     |    |         |    |      |  |
|------|-----|----|---------|----|------|--|
| Quel | est | la | couleur | du | chat |  |
|------|-----|----|---------|----|------|--|

Property of  
le magicien quantique



diff. action  
prod. carefully  
treated prod. of 1/2 treated properly



prison civitas  
rubins  
rich cr  
graves

Quelle est la couleur du chat de Hermione Granger ?

red by contrast  
entire emotional  
étroite étroite

esquive toutes  
asens contrainte  
pote de cet  
démontre



Property of  
le magicien quantique



diff. action  
prod. carefully  
treated prod. of 11 treated properly



prison civitas  
rubins  
rich cr  
graves

red by counteracts  
entire emotions  
eterna extroset

esviretvaldes  
asenes contraciabue  
pote deoet  
demit



Hermione Granger

Property of  
le magicien quantique



diff. action  
prod. carefully  
treated prod. of 1/2 treated properly



prison civitas  
rubins  
rich cr  
graves

red by courtiers  
entire emotion  
eterna extroset

esquiver les  
asmes contractions  
pote de wet  
dewit



Je suis Hermione Granger, et la  
phrase est une question sur la  
couleur de mon chat

Hermione Granger

Property of  
le magicien quantique



## Espace du perceptron (mémoire)

Hermione Granger :

- Elle a un gros chat aux longs poils orange, et à la queue touffue. Il s'appelle Pattenrond
- Son patronus est une loutre
- A failli finir chez Serdaigne...
- Bois de vigne, ventricule de dragon
- ...

Je suis Hermione Granger, et la phrase est une question sur la couleur de mon chat

Hermione Granger



Property of  
le magicien quantique



## Espace du perceptron (mémoire)

Hermione Granger :

- Elle a un gros chat aux longs poils orange, et à la queue touffue. Il s'appelle Pattenrond
- Son patronus est une loutre
- A failli finir chez Serdaigne...
- Bois de vigne, ventricule de dragon
- ...

Je suis Hermione Granger, et la phrase est une question sur la couleur de mon chat, qui s'appelle Pattenrond et qui est orange

Je suis Hermione Granger, et la phrase est une question sur la couleur de mon chat

Hermione Granger



Property of  
le magicien quantique



diff. action  
prod. carefully  
treated prod. of 1/2 treated properly



prison civitas  
rubins  
rich cr  
graves

Quelle est la couleur du chat de Hermione Granger ?

red by contrast  
entire emotional  
étroite étroite

esquive toutes  
asmes contrainte  
pote de cet  
démontre



Property of  
le magicien quantique



diff. action  
prob. carefully  
treated prob. if it behaves properly



prison civitas  
rubins  
rich cr  
graves

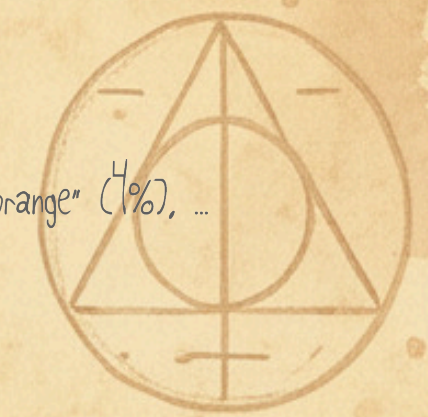
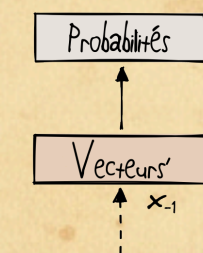
Quelle est la couleur du chat de Hermione Granger ?

red by contextures  
entire emotional  
étroite étroite

esquiver les  
asmes contraires  
pote de cet  
démontre



Jeton suivant  
"Orange" (80%), retour\_chariot (15%), "orange" (4%), ...



Property of  
le magicien quantique



(X) esvivre testules  
agueres comitricialine  
petr deoet



<https://sckathach.github.io/tp>

Property of  
le magicien quantique



# Ressources

- Cours généraux orienté sureté de l'IA: AI Safety Fundamentals de Blue Dot (<https://aisafetyfundamentals.com/>)
- Excellentissimes cours techniques sur la rétro-ingénierie de LLM: ARENA (<https://arena-chapter1-transformer-interp.streamlit.app/>)
- Forum technique à suivre: Alignement Forum (<https://www.alignmentforum.org/>)