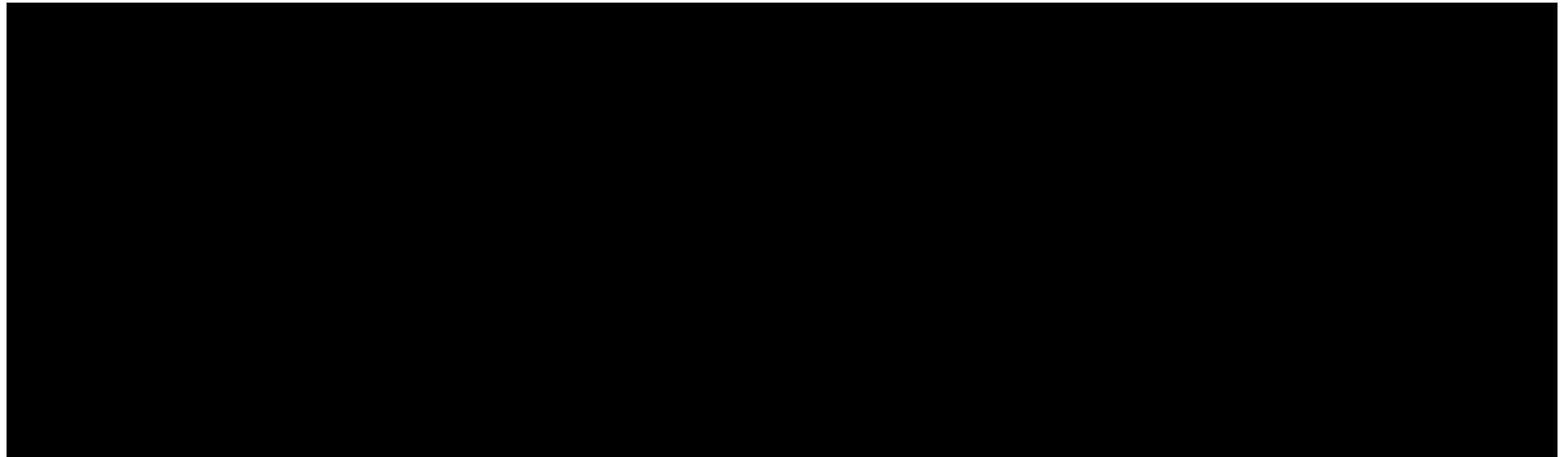

INTELLIGENCE ARTIFICIELLE ET RÉSEAUX NEURONAUX



APERÇU

1. Bref historique de l'intelligence artificielle
2. Notions fondamentales sur les réseaux neuronaux
3. Étude de cas : chaîne de pharmacies japonaise
4. Apprentissage profond

OBJECTIFS D'APPRENTISSAGE

1. Prendre connaissance de certains faits saillants de l'histoire récente de la recherche en intelligence artificielle
2. Pouvoir généralement définir ce qu'est un réseau de neurones artificiels
3. Comprendre le rôle de la rétropropagation dans les réseaux d'apprentissage neuronaux
4. Prendre conscience de certaines forces et limites des réseaux neuronaux
5. Pouvoir définir, en termes très généraux, le concept de l'« apprentissage en profondeur »



BREF HISTORIQUE DE L'INTELLIGENCE ARTIFICIELLE

INTELLIGENCE ARTIFICIELLE ET RÉSEAUX NEURONAUX

Q : Combien de pattes un chat a-t-il si vous appelez la queue une patte?

R : Quatre. Dire que la queue est une patte n'en fait pas une patte.

(ancienne énigme, attribuée à Abraham Lincoln)

MANCHETTES

« AlphaGo vanquishes world's top Go player, marking A.I.'s superiority over human mind » (AlphaGo l'emporte sur le meilleur joueur de go au monde, marquant la supériorité de l'IA par rapport à l'esprit humain) [*South China Morning Post*, 27 mai 2017]

« A Japanese A.I. program just wrote a short novel, and it almost won a literary prize » (Un programme d'IA japonais vient d'écrire une nouvelle qui a presque remporté un prix littéraire) [*Digital Trends*, 23 mars 2016]

« Elon Musk: Artificial intelligence may spark World War III » (Elon Musk : L'intelligence artificielle pourrait mettre le feu aux poudres de la Troisième Guerre mondiale) [*CNET*, 4 septembre 2017]

« A.I. hype has peaked so what's next? » (Le battage médiatique autour de l'IA est arrivé à son apogée. Et maintenant?) [*TechCrunch*, 30 septembre 2017]

QU'EST-CE QUE L'INTELLIGENCE ARTIFICIELLE (IA)?

Quelles sont les **qualités et compétences essentielles** d'une intelligence?

- Elle fournit des réponses flexibles dans divers scénarios
- Elle tire parti des circonstances heureuses
- Elle donne un sens à des messages contradictoires
- Elle reconnaît l'importance relative des éléments d'une situation
- Elle trouve des similitudes entre des situations différentes
- Elle établit des distinctions entre des situations similaires
- Elle trouve de nouvelles idées à partir de rien ou en réorganisant des concepts déjà connus

DISCUSSION

Pensez-vous que la définition de Hofstadter est toujours valable?

QU'EST-CE QUE L'INTELLIGENCE ARTIFICIELLE?

La recherche en intelligence artificielle s'entend de l'étude des **agents intelligents** : tout dispositif qui perçoit son environnement et prend des mesures qui maximisent ses chances de succès dans la poursuite de ses objectifs.

Exemples

- **Systèmes experts**

TurboTax, WebMD, assistance technique, traitement des réclamations d'assurance, contrôle du trafic aérien, etc.

- **Prise de décision**

Deep Blue, systèmes de pilotage automatique, compteurs « intelligents », etc.

- **Traitement du langage naturel**

Traduction automatique, Siri, reconnaissance d'entités nommées, etc.

- **Systèmes de recommandation**

Google, Expedia, Facebook, LinkedIn, Netflix, Amazon, etc.

- **Producteurs de contenu**

Compositeur de musique, rédacteur de romans, créateur de dessins animés, etc.

- **Classificateurs**

Reconnaissance faciale, identification d'objets, détection des fraudes, etc.

FRISE CHRONOLOGIQUE (TL;DR)

Apprentissage en
profondeur et
mégadonnées

Optimisme
initial

Renaissance

Intelligence
artificielle
moderne

1^{er} hiver de
l'intelligence
artificielle

2^e hiver de
l'intelligence
artificielle

1950

1960

1970

1980

1990

2000

2010

2020

???

???

Test de Turing

IDLEWYLD Sysabee
Conference de Dartmouth

DAVHILL

uOttawa
Retropropagation

Projet de 5^e génération

Deep Blue

Projet du cerveau humain (Human Brain Project)

Stanley

data-action-lab.com

???

DISCUSSION

À votre avis, où va l'intelligence artificielle? Un 3^e hiver de l'intelligence artificielle, ou un avenir brillant?

Comment les entreprises et les gouvernements peuvent-ils se préparer à l'une ou l'autre éventualité?

NOTIONS FONDAMENTALES SUR LES RÉSEAUX NEURONAUX

INTELLIGENCE ARTIFICIELLE ET RÉSEAUX NEURONAUX

« Les réseaux de neurones l'emportent de loin sur toutes les techniques précédentes sur le plan de la performance, mais l'existence d'exemples contradictoires montre que nous ne comprenons vraiment pas ce qui se passe

[traduction]. »

(D. Gershoorn, *Quartz*)

LES RÉSEAUX NEURONAUX EN BREF

Un **réseau de neurones artificiels** (RNA) formé est une fonction qui met en correspondance les données d'entrée et de sortie de manière utile :

- Il reçoit la ou les données d'entrée
- Il calcule les valeurs
- Il fournit la ou les données de sortie

Les RNA font appel à une approche du genre couteau suisse (**beaucoup d'options, mais on ne sait pas toujours laquelle utiliser**).

L'utilisateur n'a pas besoin de prendre de nombreuses décisions au sujet de la fonction ou d'en savoir beaucoup sur l'espace de problème à l'avance (**modèle silencieux**).

LES RÉSEAUX NEURONAUX EN BREF

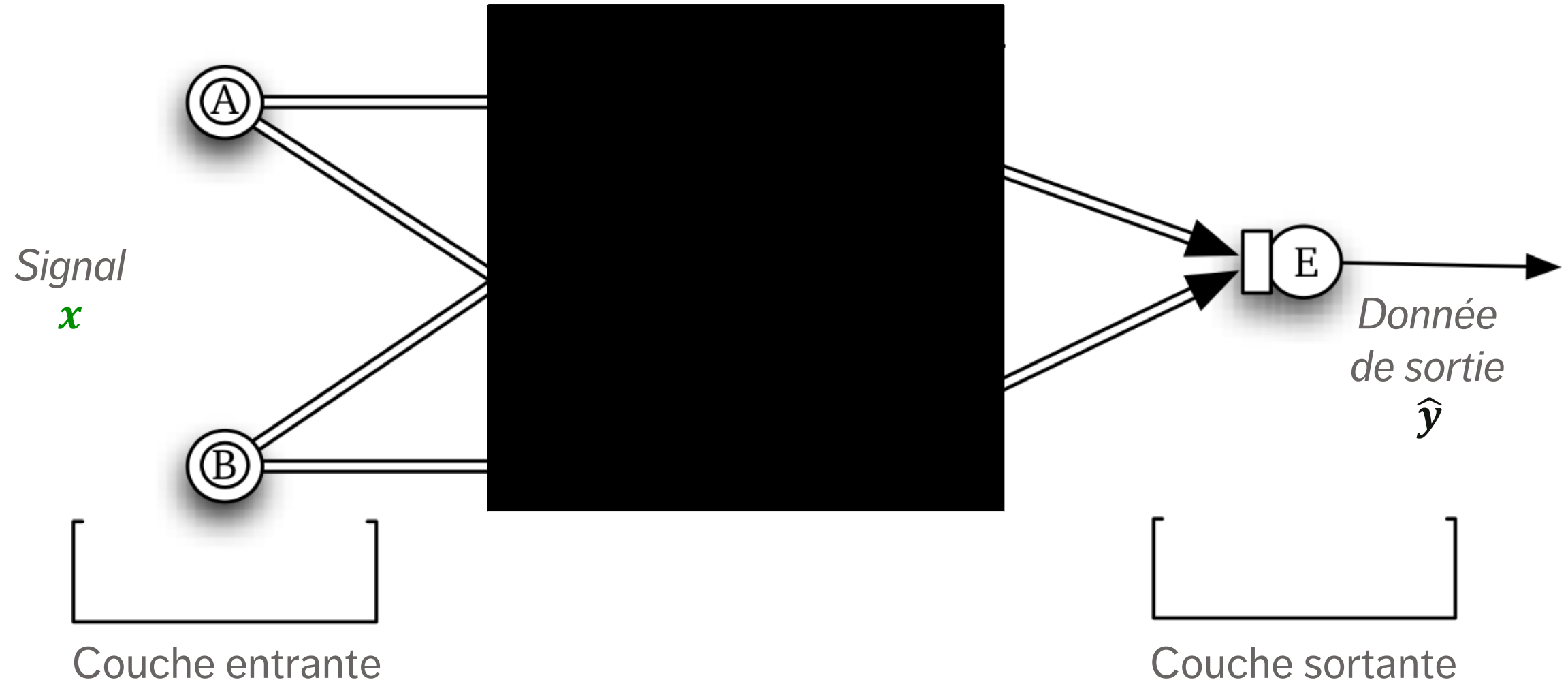
Des algorithmes permettent aux RNA d'**apprendre** (c.-à-d. de générer la fonction et ses valeurs internes) **automatiquement**.

On peut avoir recours aux RNA pour :

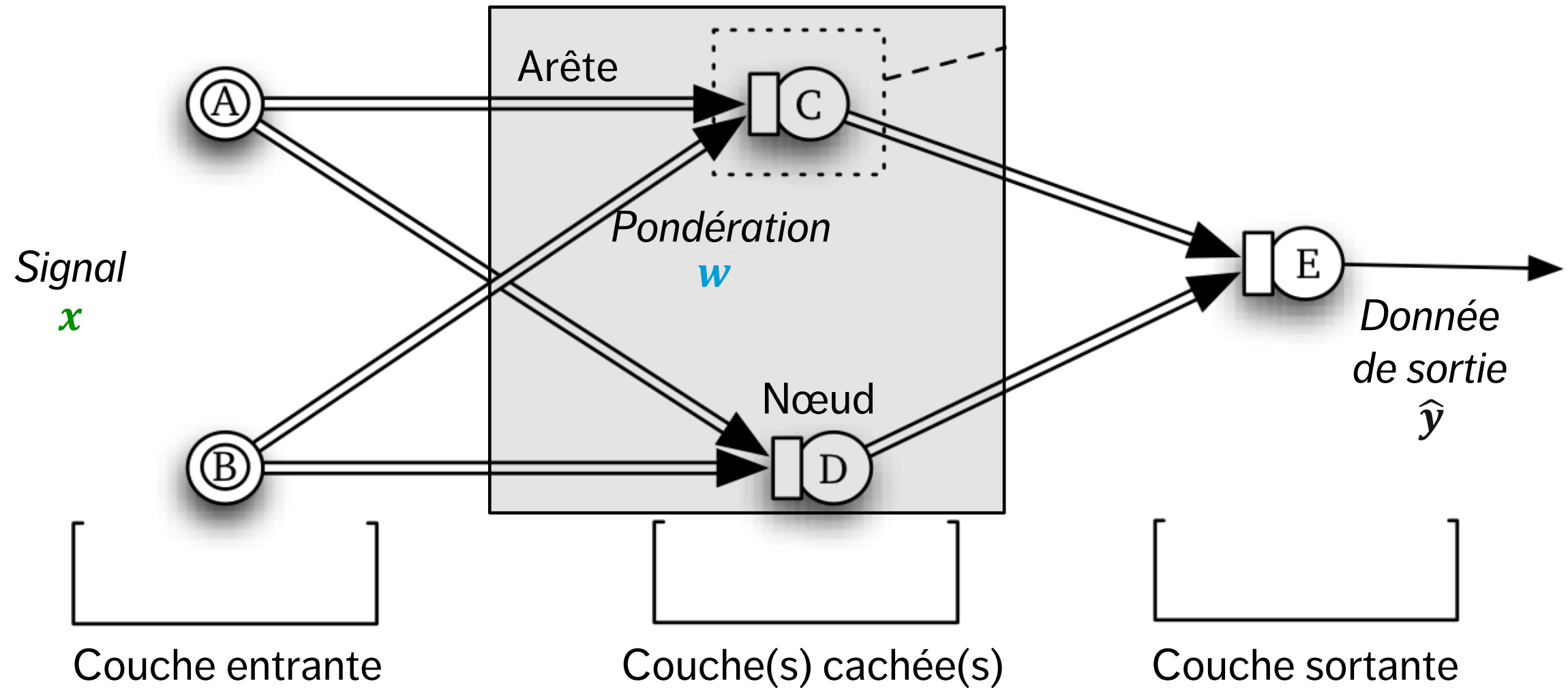
- L'apprentissage supervisé (**réseaux neuronaux multicouches acycliques**)
- L'apprentissage non supervisé (**cartes autoorganisables**)
- L'apprentissage par renforcement.

Techniquement, la seule exigence consiste à pouvoir réduire au minimum la fonction coût (**optimisation**).

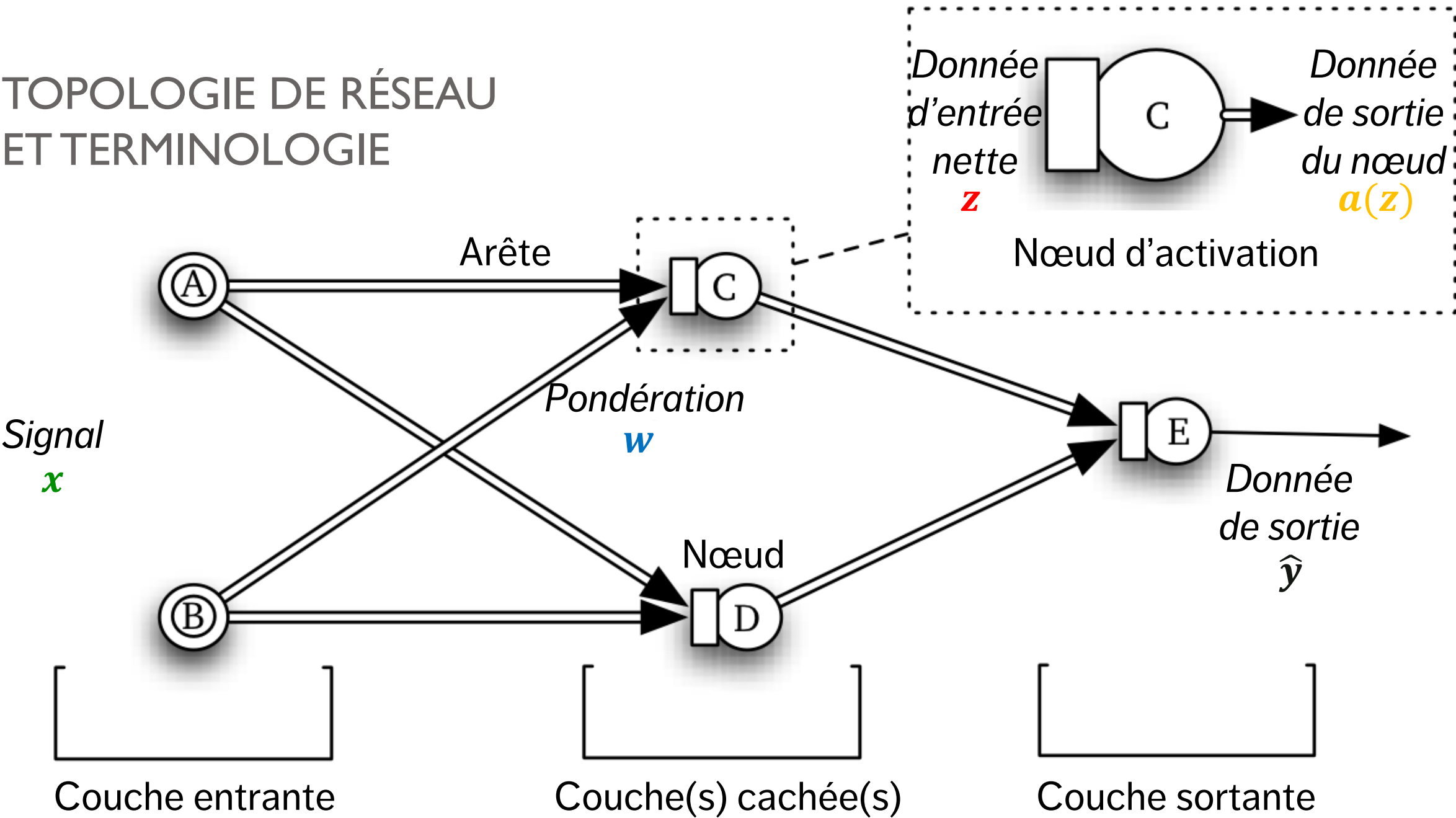
TOPOLOGIE DE RÉSEAU ET TERMINOLOGIE



TOPOLOGIE DE RÉSEAU ET TERMINOLOGIE

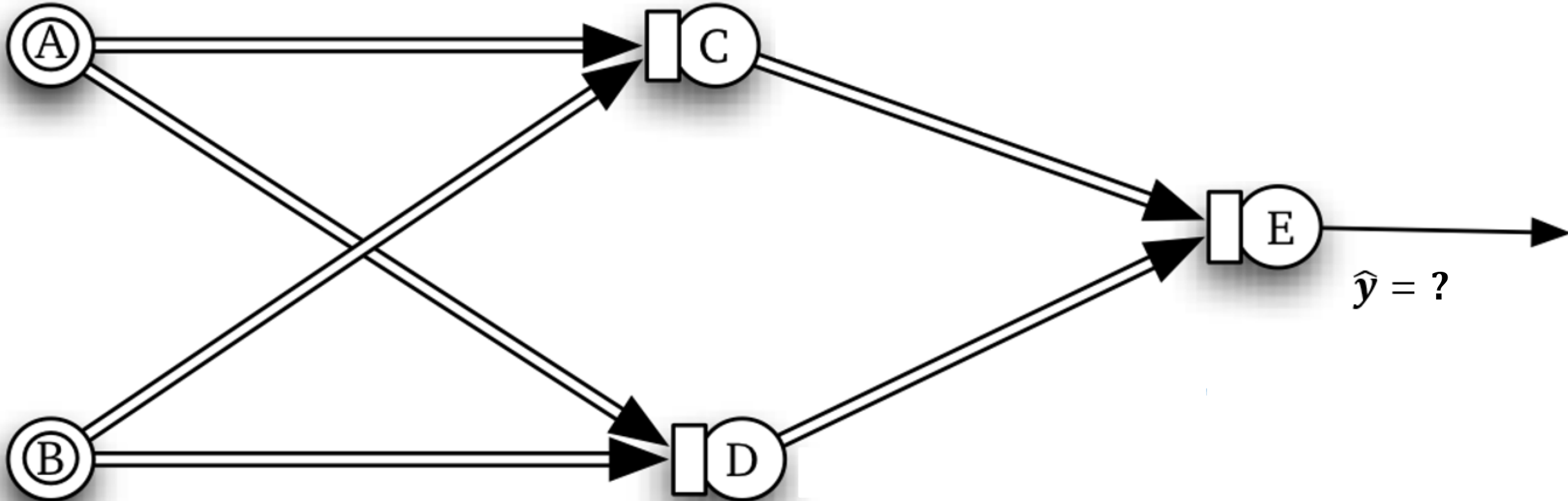


TOPOLOGIE DE RÉSEAU ET TERMINOLOGIE



RÉSEAU ACYCLIQUE

$$x_A = 0.8$$



$$x_B = 0.9$$

$$\hat{y} = ?$$

RÉSEAU ACYCLIQUE

$$x_A = 0.8$$

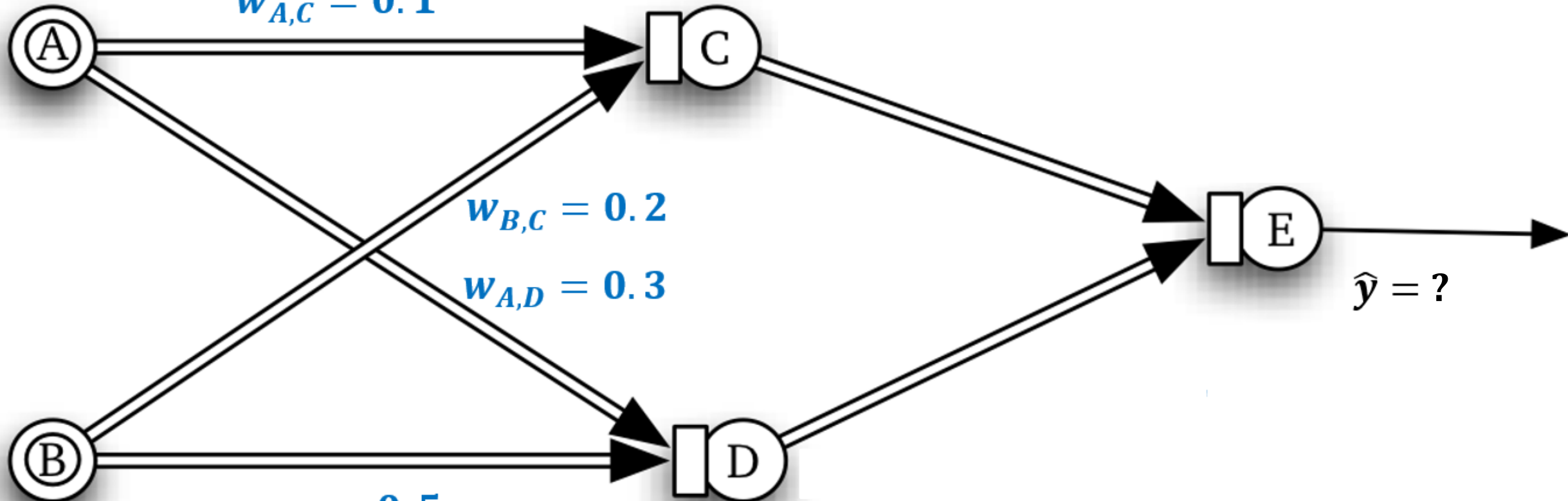
$$w_{A,C} = 0.1$$

$$w_{B,C} = 0.2$$

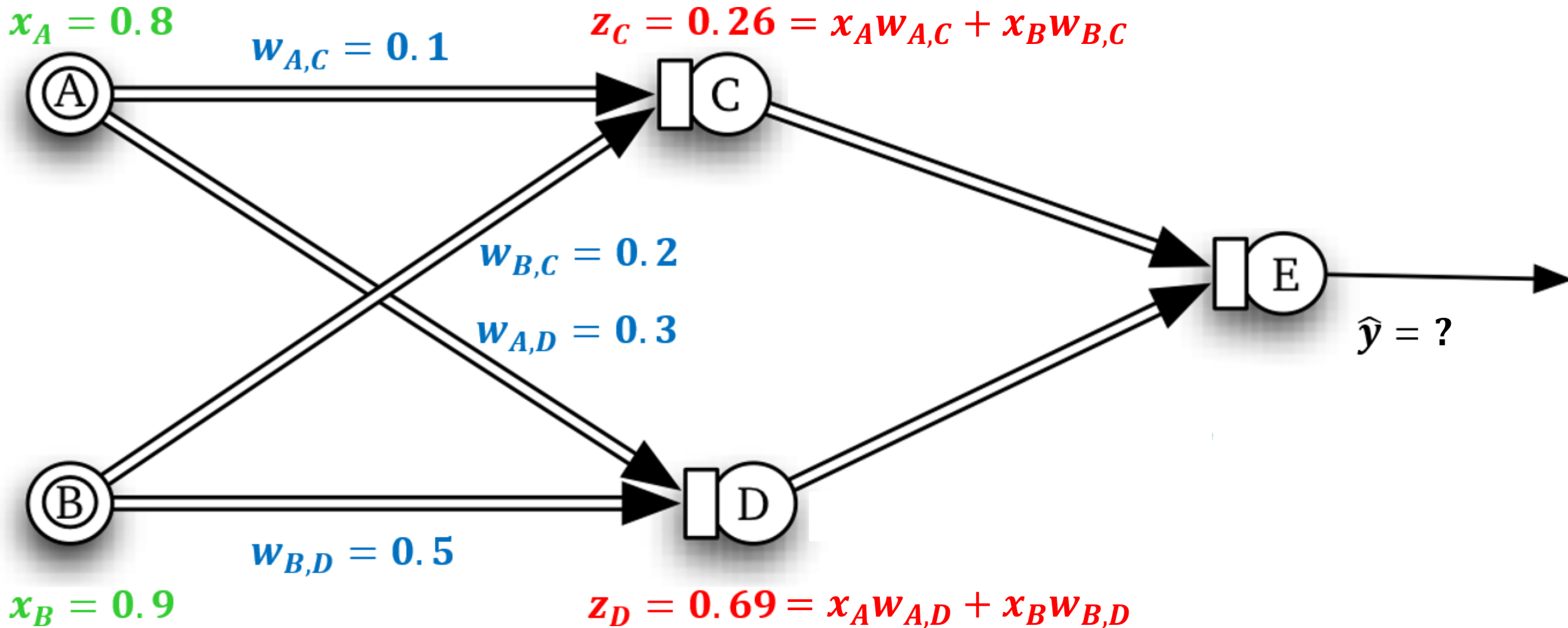
$$w_{A,D} = 0.3$$

$$w_{B,D} = 0.5$$

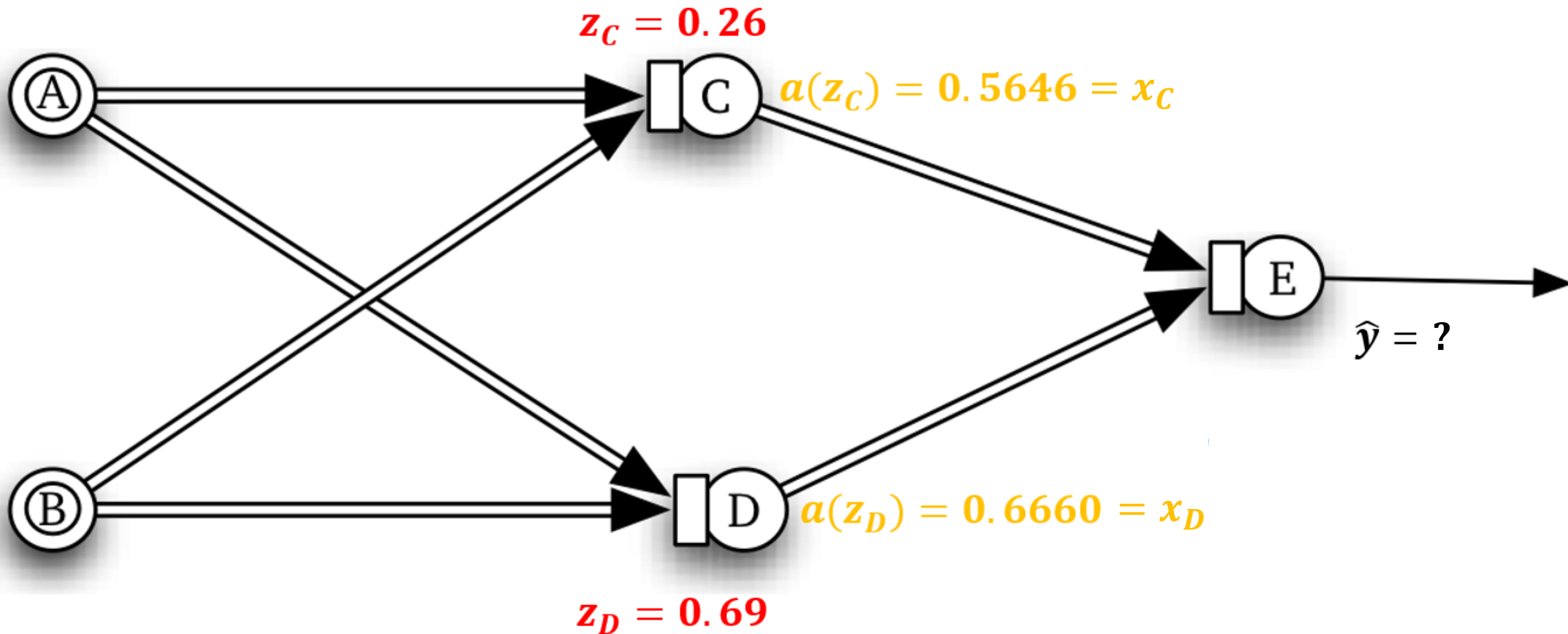
$$x_B = 0.9$$



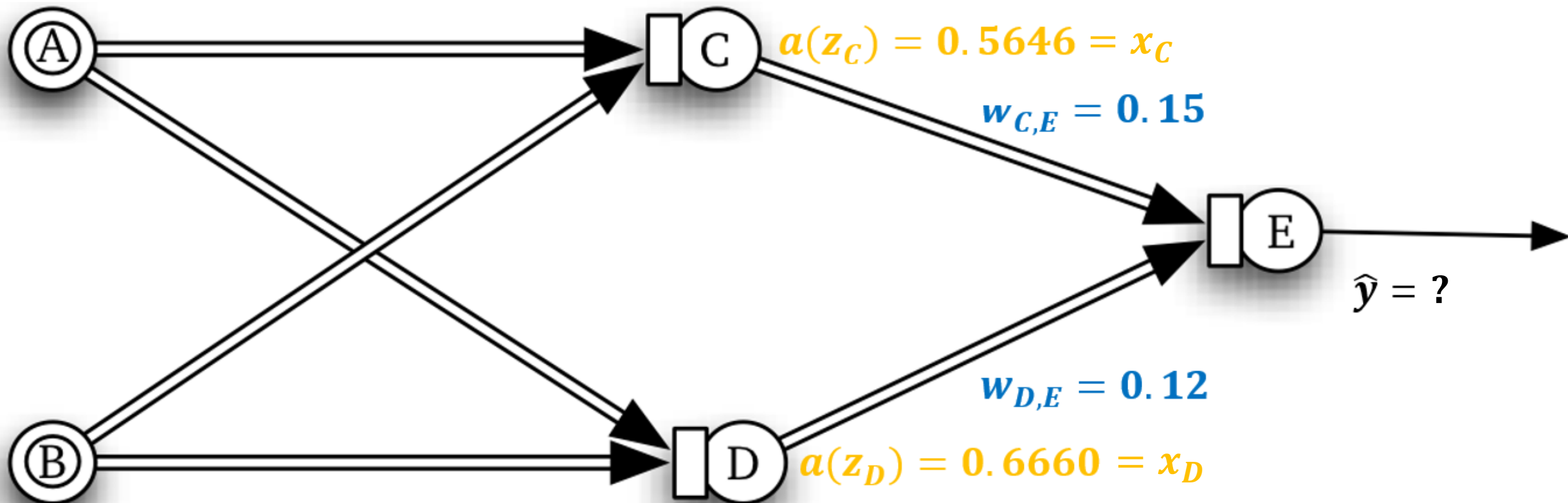
RÉSEAU ACYCLIQUE



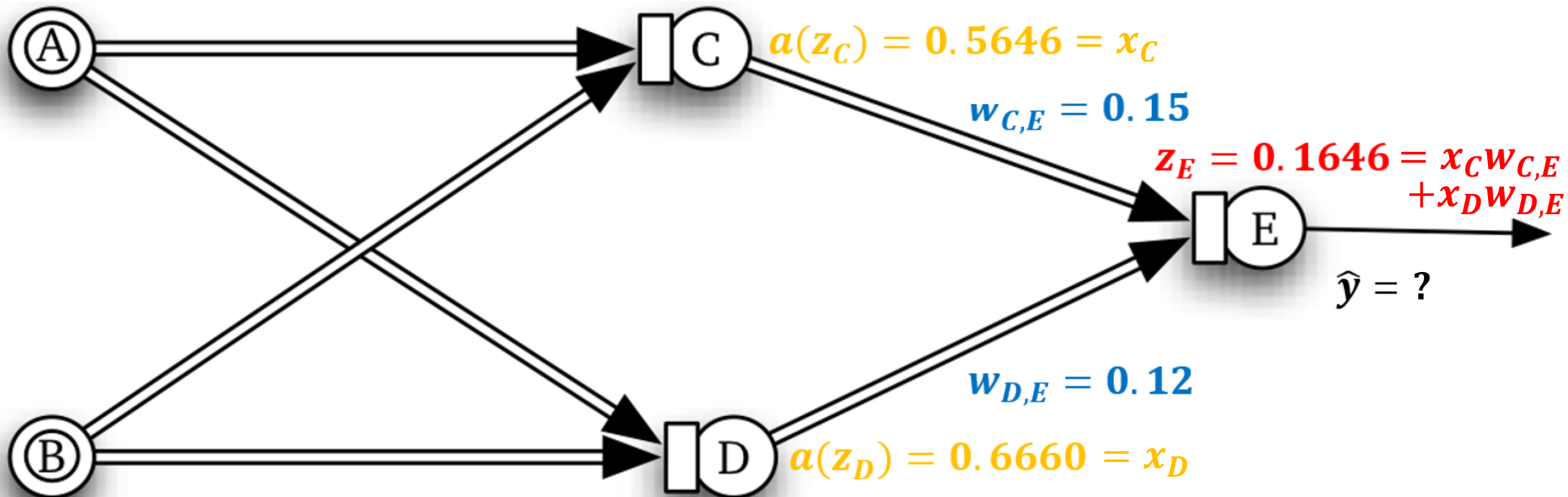
RÉSEAU ACYCLIQUE



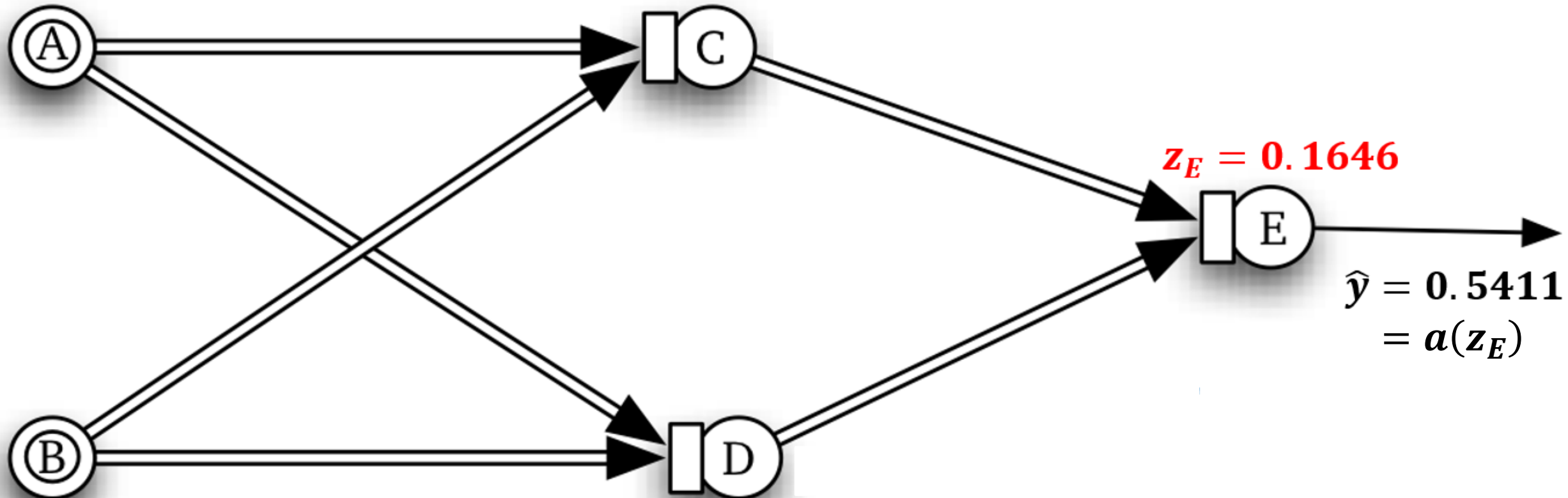
RÉSEAU ACYCLIQUE



RÉSEAU ACYCLIQUE



RÉSEAU ACYCLIQUE



RNA EN NOTATION MATRICIELLE – DONNÉE D'ENTRÉE VERS LA COUCHE CACHÉE

Donnée d'entrée : $\mathbf{X}_{(n \times p)} = \mathbf{X}_{(n \times 2)}$

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_{A,1} & \mathbf{x}_{B,1} \\ \vdots & \vdots \\ \mathbf{x}_{A,n} & \mathbf{x}_{B,n} \end{bmatrix}$$

Unités cachées : $\mathbf{Z}_{(n \times M)}^{(2)} = \mathbf{Z}_{(n \times 2)}^{(2)}$

$$\mathbf{Z}^{(2)} = \begin{bmatrix} \mathbf{z}_{C,1} & \mathbf{z}_{D,1} \\ \vdots & \vdots \\ \mathbf{z}_{C,n} & \mathbf{z}_{D,n} \end{bmatrix} = \mathbf{X} \mathbf{W}^{(1)}$$

Pondérations : $\mathbf{W}_{(p \times M)}^{(1)} = \mathbf{W}_{(2 \times 2)}^{(1)}$

$$\mathbf{W}^{(1)} = \begin{bmatrix} \mathbf{w}_{AC} & \mathbf{w}_{AD} \\ \mathbf{w}_{BC} & \mathbf{w}_{BD} \end{bmatrix}$$

Fonction d'activation : $\mathbf{a}_{(n \times M)}^{(2)} = \mathbf{a}_{(n \times 2)}^{(2)}$

$$\mathbf{a}^{(2)} = \begin{bmatrix} 1/(1 + e^{-\mathbf{z}_{C,1}}) & 1/(1 + e^{-\mathbf{z}_{D,1}}) \\ \vdots & \vdots \\ 1/(1 + e^{-\mathbf{z}_{C,n}}) & 1/(1 + e^{-\mathbf{z}_{D,n}}) \end{bmatrix} = g(\mathbf{Z}^{(2)})$$

RNA EN NOTATION MATRICIELLE – DONNÉE CACHÉE VERS LA COUCHE DE SORTIE

Unités de sortie : $\mathbf{z}_{(n \times K)}^{(3)} = \mathbf{z}_{(n \times 1)}^{(3)}$

Fonction d'activation : $\mathbf{a}_{(n \times M)}^{(2)} = \mathbf{a}_{(n \times 2)}^{(2)}$
 $\mathbf{a}^{(2)} = g(\mathbf{z}^{(2)})$

$$\mathbf{z}^{(3)} = \begin{bmatrix} \mathbf{z}_{E,1} \\ \vdots \\ \mathbf{z}_{E,n} \end{bmatrix} = \mathbf{a}^{(2)} \mathbf{w}^{(2)}$$

Pondérations : $\mathbf{W}_{(M \times K)}^{(2)} = \mathbf{W}_{(2 \times 1)}^{(2)}$
 $\mathbf{W}^{(2)} = [\mathbf{w}_{CE} \quad \mathbf{w}_{DE}]$

Fonction d'activation : $\mathbf{a}_{(n \times K)}^{(2)} = \mathbf{a}_{(n \times 1)}^{(2)}$

$$\hat{\mathbf{y}} = \mathbf{a}^{(3)} = \begin{bmatrix} 1/(1 + e^{-\mathbf{z}_{E,1}}) \\ \vdots \\ 1/(1 + e^{-\mathbf{z}_{E,n}}) \end{bmatrix} = g(\mathbf{z}^{(3)})$$

RNA EN NOTATION MATRICIELLE

Cet exemple de réseau neuronal *standard* peut être exprimé ainsi :

$$\hat{y} = \mathbf{a}^{(3)} = g(\mathbf{z}^{(3)}) = g[\mathbf{a}^{(2)} \mathbf{W}^{(2)}] = g[g(\mathbf{x} \mathbf{W}^{(1)}) \mathbf{W}^{(2)}]$$

En bref, à chaque nœud, le réseau neuronal

1. Calcule la **somme pondérée** des **données d'entrée**
2. Applique les fonctions d'**activation** et
3. Envoie un **signal**,

jusqu'à ce que le signal atteigne le dernier nœud de **sortie**.

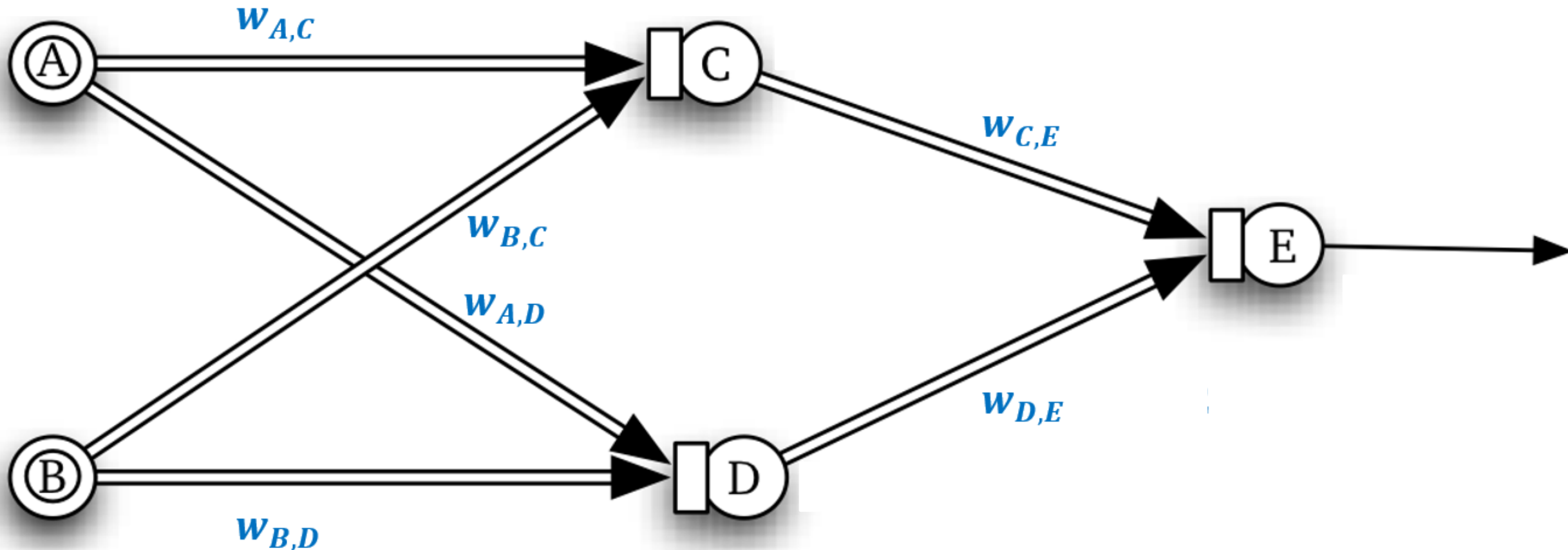
RÉTROPROPAGATION – APPRENTISSAGE DU RNA

Étant donné un signal, un RNA peut produire une donnée de sortie, pour autant que les pondérations soient précisées.

Pour les tâches d'apprentissage **supervisé** (c'est-à-dire lorsqu'un RNA tente d'imiter les résultats des exemples d'apprentissage), le simple fait de choisir des pondérations au hasard est une proposition défaillante.

La **rétropropagation** est une méthode d'optimisation du choix des pondérations par rapport à une fonction d'erreur $R(W)$.

OBJECTIF DE LA RÉTROPROPAGATION



RÉTROPROPAGATION – FONCTION D'ERREUR

Dans les problèmes de régression, on utilise souvent la **somme des carrés des erreurs** (SCE) comme fonction d'erreur :

$$R(W) = \sum_{i=1}^n \sum_{k=1}^K (\hat{y}_{ik}(W) - y_{ik})^2$$

Dans les problèmes de classification, on peut utiliser l'**entropie croisée** :

$$R(W) = - \sum_{i=1}^n \sum_{k=1}^K y_{ik} \ln[\hat{y}_{ik}(W)]$$

De l'une ou l'autre façon, nous voulons réduire $R(W)$ au minimum par rapport à W (**descente de gradient**).



Iterations
000,000

Learning rate
0.03

Activation
Tanh

Regularization
None

Regularization rate
0

Problem type
Classification

DATA

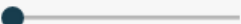
Which dataset do you want to use?



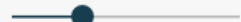
Ratio of training to test data: 50%



Noise: 0



Batch size: 10



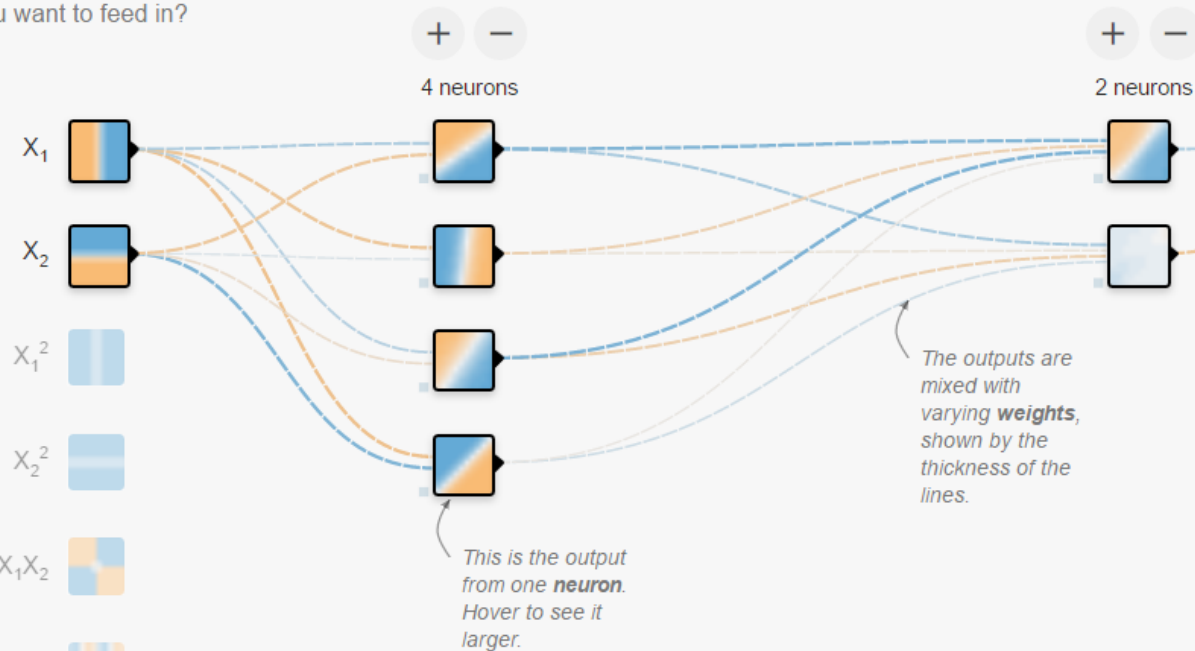
REGENERATE

FEATURES

Which properties do you want to feed in?

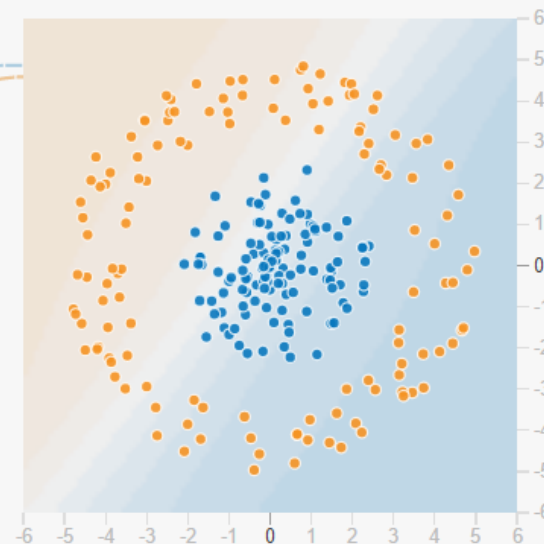
- X_1
- X_2
- X_1^2
- X_2^2
- $X_1 X_2$
- $\sin(X_1)$
- $\sin(X_2)$

+ - 2 HIDDEN LAYERS

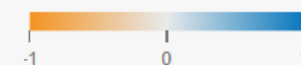


OUTPUT

Test loss 0.512
Training loss 0.488



Colors shows data, neuron and weight values.



☐ Show test data

☐ Discretize output

A NEURAL NETWORK PLAYGROUND [D. Smilkov, S. Carter, [TensorFlow.org](https://www.tensorflow.org) (en anglais seulement)]

FORCES

Les RNA peuvent être assez **précis** pour faire des prédictions – plus que d'autres algorithmes, lorsqu'ils sont correctement configurés.

Les RNA fonctionnent souvent lorsque d'autres méthodes échouent :

- Lorsque la relation entre les attributs est **complexe**
- Lorsqu'il y a beaucoup de dépendences/**rapports non linéaires**
- En présence de données d'entrée **désorganisées** et hautement connectées (images, texte et voix)
- Classification non linéaire

Les RNA sont relativement faciles à configurer (avec les progiciels disponibles).

Les RNA continuent de bien fonctionner avec le temps (ce qui est important en robotique).

LIMITATIONS

Les RNA sont relativement lents (création et utilisation) et enclins au surapprentissage (peuvent nécessiter un ensemble d'apprentissage **important/diversifié**).

Les RNA ne fournissent généralement pas une bonne interprétation (contrairement aux arbres de décision ou à la régression logistique, par exemple). Pouvez-vous vous en accommoder?

Il n'existe pas d'algorithme pour sélectionner la topologie de réseau optimale.

Même si les RNA produisent de meilleurs résultats que les autres options, ils peuvent ne pas être aussi performants en raison des **théorèmes** « **no free lunch** » et ils sont sujets à diverses formes d'**attaques malveillantes**.

DISCUSSION

Le plus grand défi (à notre avis) consiste à surmonter la nature de boîte noire des RNA. Dans quelle mesure est-il important pour vous et votre organisation de pouvoir expliquer les décisions fondées sur les données?

VIDÉOS SUR LES RÉSEAUX NEURONAUX (EXCELLENTE!)

1. Neural Networks Demystified, Welch Labs
<https://www.youtube.com/watch?v=bxe2T-V8XRs> (en anglais seulement, première vidéo de la série)
2. Learning to See, Welch Labs
<https://www.youtube.com/watch?v=i8D9oDkCLhI> (en anglais seulement, première vidéo de la série)
3. Neural Networks, 3 Blue 1 Brown
<https://www.3blue1brown.com/videos/2017/10/9/neural-network> (en anglais seulement)

ÉTUDE DE CAS : CHAÎNE DE PHARMACIES JAPONAISE

INTELLIGENCE ARTIFICIELLE ET RÉSEAUX NEURONAUX

A Neural Network Application to Identify High-Value Customers for a Large Retail Store in Japan
(en anglais seulement)

(Ip, E., Johnson, J., Yada, K., Hamuro, Y., Katoh, N., Cheung, S. [2002], *Neural Networks in Business: Techniques and Applications*,
Smith, K., Gupta, J. (eds), IRM Press)

CONTEXTE

6 × plus coûteux de vendre à un nouveau client qu'à un client existant
(Kalakota, Robinson, Tapscott, 1999)

Une fidélisation annuelle de la clientèle de 5 % peut entraîner une augmentation de 85 % des bénéfices
(Kalakota, Robinson, Tapscott, 1999)

La fidélisation des « bons » clients joue un rôle dans la rentabilité à long terme
(Reicheld, 1993)

Comment peut-on reconnaître **dès le début** les clients **fidèles** et **rentables**?

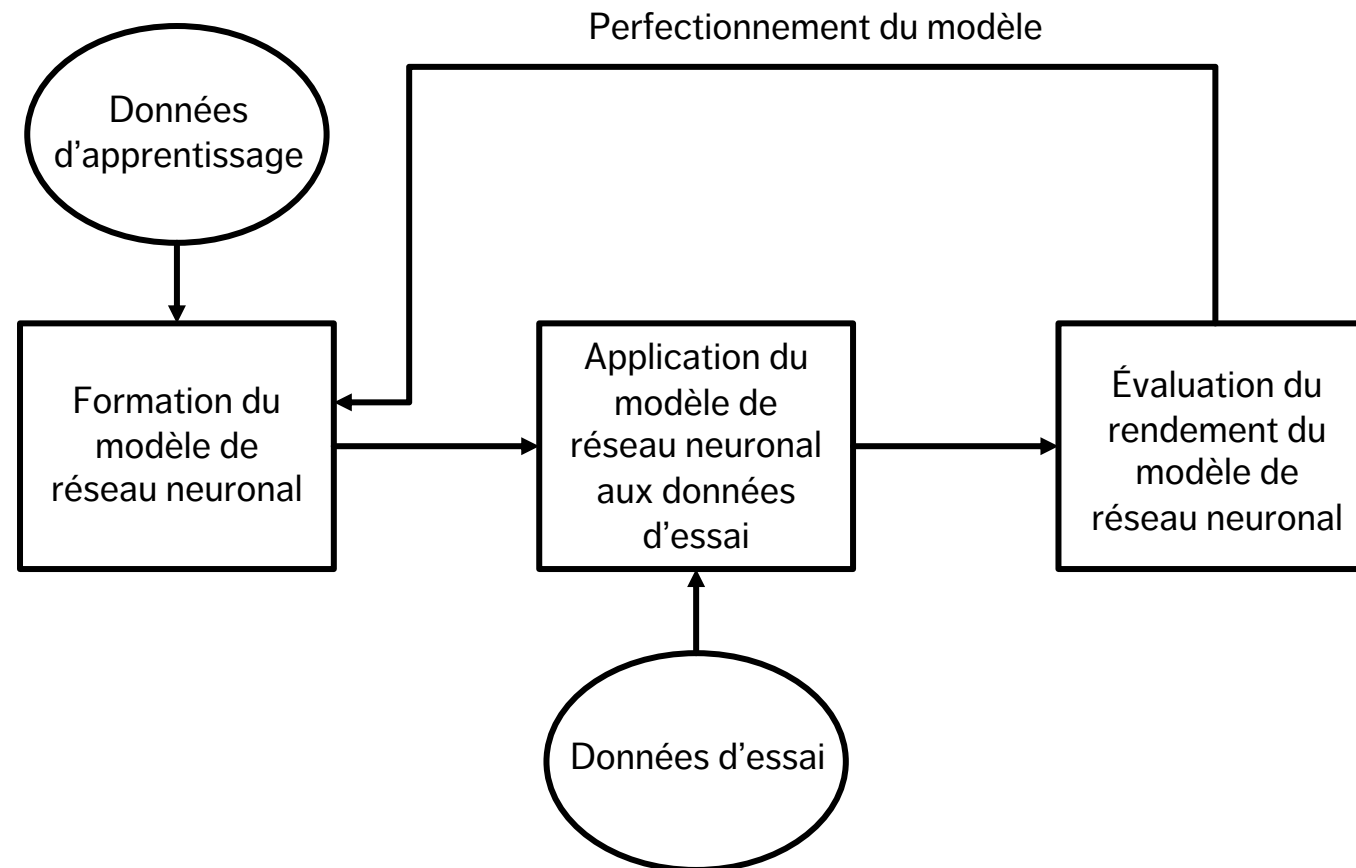
CONTEXTE

Chaîne de pharmacies japonaise *Pharma* (revenus annuels de 70 milliards de yens), 1 300 points de vente, 2,3 millions de clients

Les magasins peuvent utiliser leur propre nom, mais ils sont exploités sous un système d'information central, qui collecte les informations sur les transactions et les clients depuis le début des années 1990

Informations sur la campagne de vente, bons personnalisés, échantillons gratuits envoyés par la poste au groupe cible (clients fidèles et de grande valeur)

FLUX D'APPRENTISSAGE DES RÉSEAUX NEURONAUX



DONNÉES

114 069 clients qui ont effectué des achats sur une période de 1 an

Valeur du client mesurée par

- La fréquence des visites (échelle de 1 à 5)
- La rentabilité par visite (échelle de 1 à 5)

Clients à valeur élevée (CVE) : (4,5), (5,4), (5,5), représentent 10,6 % des observations

Les CVE génèrent 52,5 % des profits, 38,4 % des revenus

DONNÉES

Variable cible : valeur du client

Variables d'entrée : nombre total de catégories d'achat, profit par visite, nombre d'unités achetées par visite, nombre de visites, achat de :

Toutes les variables ont été mises à l'échelle de 0 à 1 :

- Produit en papier
- Détergent
- Gouttes ophtalmiques
- Nettoyant de cuisine
- Supplément en bouteille
- Produits de soins capillaires
- Adoucissant
- Nettoyant ménager
- Pâte dentifrice
- Médicaments contre le rhume

DONNÉES

Ensemble d'apprentissage : 104 069 observations (sélectionnées aléatoirement)

Rapport ensemble d'apprentissage/ensemble d'essai : 70 à 30

Ensemble de validation : 10 000 observations restantes

Paramètres d'étalonnage : exactitude prédictive, exactitude globale

Exactitude prédictive = nombre de CVE correctement prédits/nombre de CVE

Exactitude globale = nombre de clients dont la classe correcte est
prédite/nombre de clients

RÉSULTATS

À l'aide d'un **réseau neuronal multicouches acyclique**, les chercheurs ont été en mesure de capter 80 % des CVE en ciblant 25 % des nouveaux clients (pourcentage spécifié par le modèle).

À une valeur de paramètre seuil de 30 %, le modèle donne des résultats 5 × meilleurs que la classification aléatoire des clients.

Ensemble de données	Apprentissage	Validation
Exactitude prédictive	55,6 %	57,4 %
Exactitude globale	90,6 %	91,2 %
% de clients classés comme CVE	10,6 %	10,3 %

DISCUSSION

Pensez-vous que les résultats de cette étude de cas sont transférables à d'autres domaines?

Est-ce qu'un réseau neuronal multicouches acyclique similaire fournirait le même niveau de pouvoir prédictif au Canada?

APPRENTISSAGE PROFOND

INTELLIGENCE ARTIFICIELLE ET RÉSEAUX NEURONAUX

« L'humanité est au bord de l'esclavage numérique aux mains de l'intelligence artificielle et des technologies biométriques. L'une des façons d'éviter cette situation, c'est de développer des modules intégrés de sentiments profonds d'amour et de compassion dans les algorithmes d'apprentissage. »

(A. Ray)

RÉSEAUX D'APPRENTISSAGE EN PROFONDEUR

Les réseaux d'apprentissage en profondeur sont simplement des RNA qui comportent **un grand nombre de couches cachées** (et divers types de nœuds)

Types :

- *Réseaux neuronaux à convolution*
Reconnaissance des chiffres manuscrits, exactitude de 99,7 % en 2013, voitures sans chauffeur
- *Réseaux de neurones récurrents*
Traitement du langage naturel (reconnaissance de la parole, traduction automatique, etc.)
- *Auto-encodeurs*
- *Machines de Boltzmann restreintes*
BellKor's Pragmatic Chaos, Netflix Prize, 2009

A mostly complete chart of Neural Networks

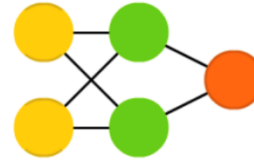
©2016 Fjodor van Veen - asimovinstitute.org

- Backfed Input Cell
- Input Cell
- △ Noisy Input Cell
- Hidden Cell
- Probablistic Hidden Cell
- △ Spiking Hidden Cell
- Output Cell
- Match Input Output Cell
- Recurrent Cell
- Memory Cell
- △ Different Memory Cell
- Kernel
- Convolution or Pool

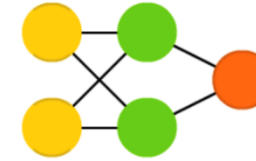
Perceptron (P)



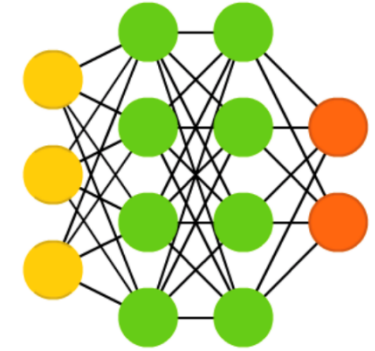
Feed Forward (FF)



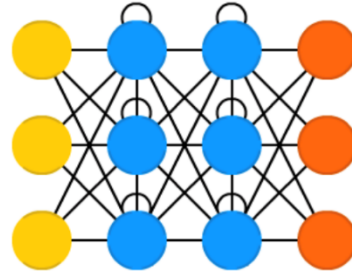
Radial Basis Network (RBF)



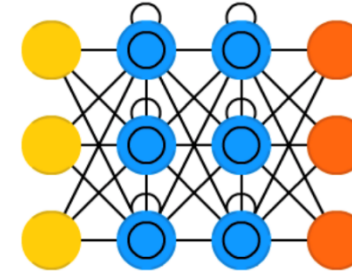
Deep Feed Forward (DFF)



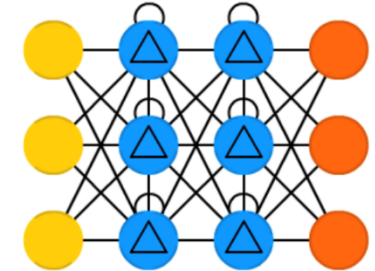
Recurrent Neural Network (RNN)



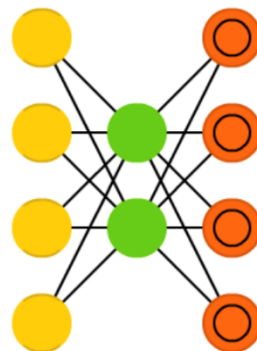
Long / Short Term Memory (LSTM)



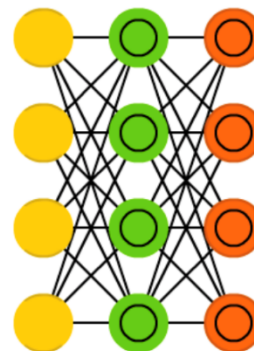
Gated Recurrent Unit (GRU)



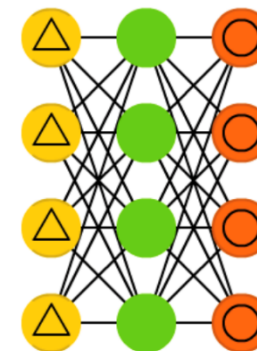
Auto Encoder (AE)



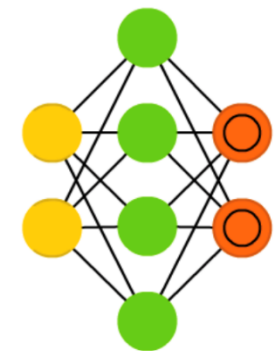
Variational AE (VAE)



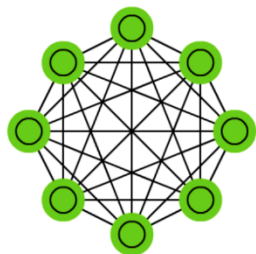
Denoising AE (DAE)



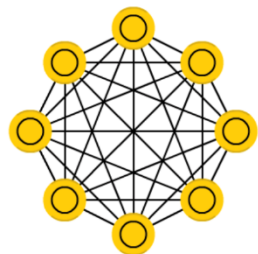
Sparse AE (SAE)



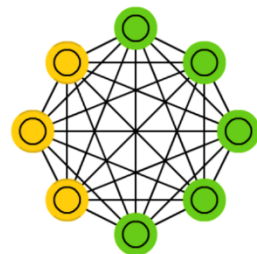
Markov Chain (MC)



Hopfield Network (HN)



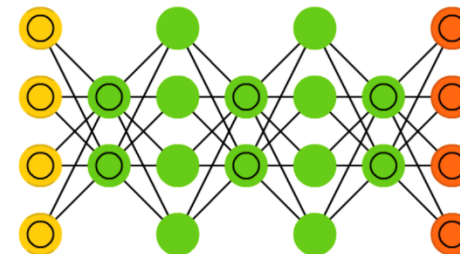
Boltzmann Machine (BM)



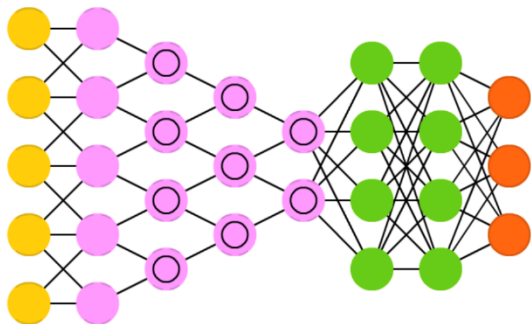
Restricted BM (RBM)



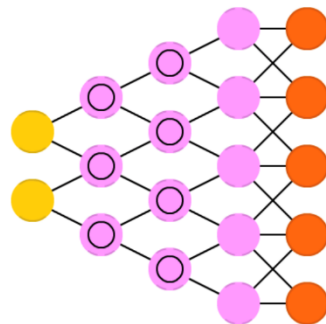
Deep Belief Network (DBN)



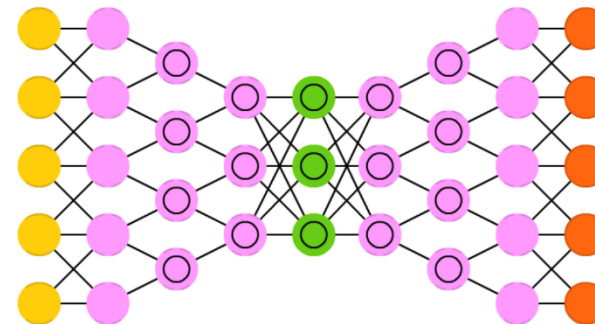
Deep Convolutional Network (DCN)



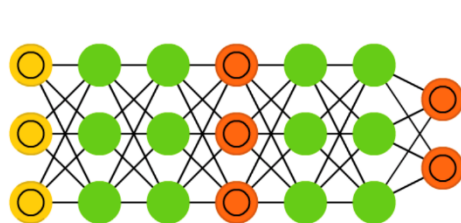
Deconvolutional Network (DN)



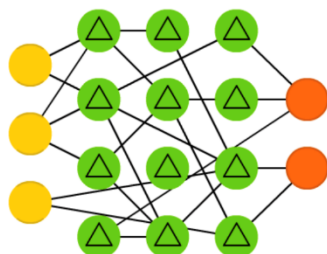
Deep Convolutional Inverse Graphics Network (DCIGN)



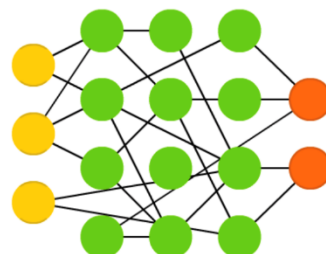
Generative Adversarial Network (GAN)



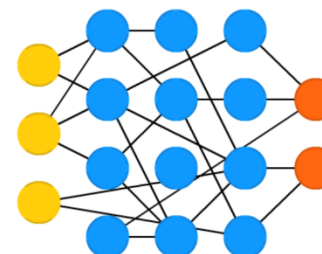
Liquid State Machine (LSM)



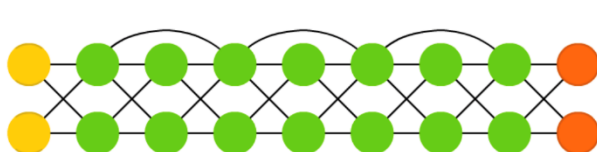
Extreme Learning Machine (ELM)



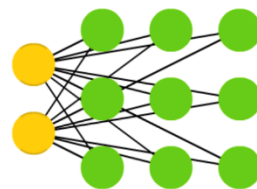
Echo State Network (ESN)



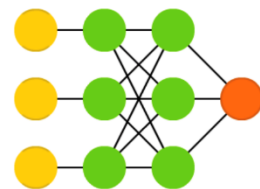
Deep Residual Network (DRN)



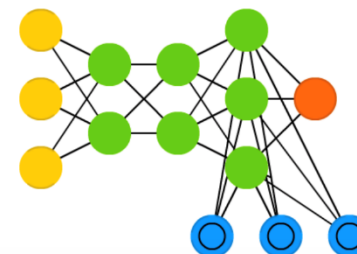
Kohonen Network (KN)



Support Vector Machine (SVM)



Neural Turing Machine (NTM)



LIMITATIONS

Ils nécessitent des ensembles d'apprentissage **importants, diversifiés et correctement étiquetés**.

Ils sont exacts en moyenne, mais peuvent toujours être **spectaculairement faux**.

On peut les pirater (LNF).

Les êtres humains n'ont pas besoin d'autant de données étiquetées pour prendre des décisions : donc, **qu'est-ce qui se passe réellement là-dedans?** (3^e hiver de l'intelligence artificielle?)



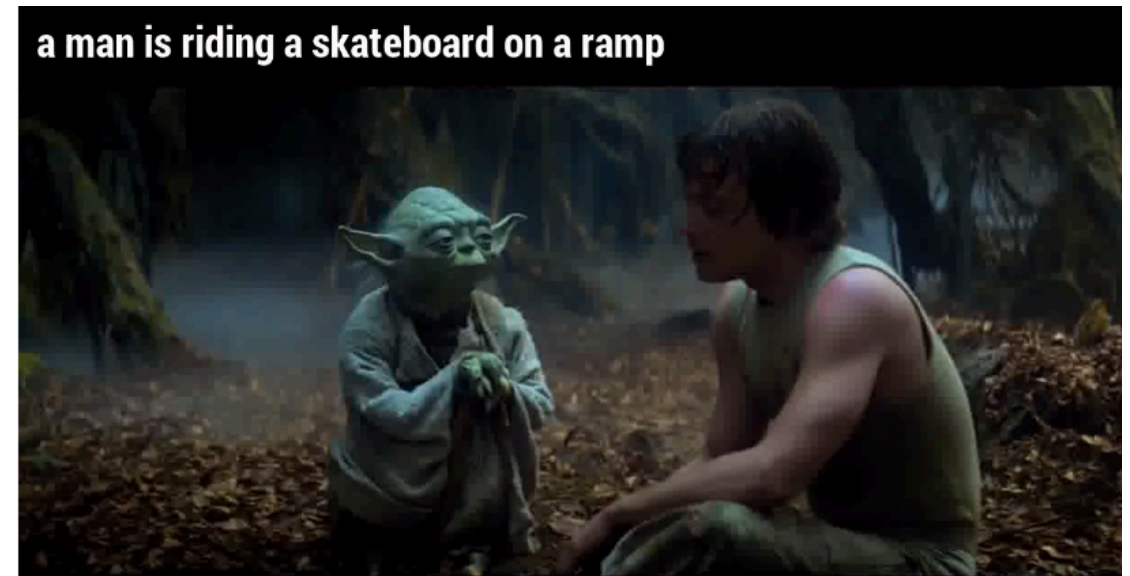
LIMITATIONS

Ils nécessitent des ensembles d'apprentissage **importants, diversifiés et correctement étiquetés**.

Ils sont exacts en moyenne, mais peuvent toujours être **spectaculairement faux**.

On peut les pirater (LNF).

Les êtres humains n'ont pas besoin d'autant de données étiquetées pour prendre des décisions : donc, **qu'est-ce qui se passe réellement là-dedans?** (3^e hiver de l'intelligence artificielle?)



DISCUSSION

Comment harmonisons-nous les objectifs d'intelligence artificielle autonome avec les valeurs humaines et culturelles?

DOCUMENTATION SUPPLÉMENTAIRE

INTELLIGENCE ARTIFICIELLE ET RÉSEAUX NEURONAUX

FRISE CHRONOLOGIQUE

Temps préhistoriques

- *Analytic Engine* de Babbage (1834)
- Paradigme du cerveau semblable à l'ordinateur de Turing (1936)
- Test de Turing (1950) : un ordinateur peut-il tromper un observateur humain (*Imitation Game*)

Optimisme initial (l'IA sera résolue d'ici aux années 1980)

- Projet de recherche d'été de Dartmouth sur l'IA et les RN (1956)
- Perceptrons (de 1957 à 1966)
- Cartes autoorganisables (1976)

FRISE CHRONOLOGIQUE

Premier hiver de l'intelligence artificielle (de 1974 à 1980)

- Des attentes élevées, mais pas grand-chose de réalisé
- Manque de puissance de calcul, financement retiré

Renaissance

- Projet de cinquième génération du Japon (de 1982 à 1992)
- Financement de la Defense Advanced Research Projects Agency (1983)
- **Rétropropagation** (1985)
- 1^{re} conférence IEEE (1987)

FRISE CHRONOLOGIQUE

Deuxième hiver de l'intelligence artificielle (de 1987 à 1993)

- Maintenance coûteuse; coupes budgétaires massives
- Mauvais apprenants, mauvais résultats en cas d'incertitude
- Objectifs du projet de cinquième génération du Japon pas réalisés

L'intelligence artificielle moderne (de 1993 à 2001)

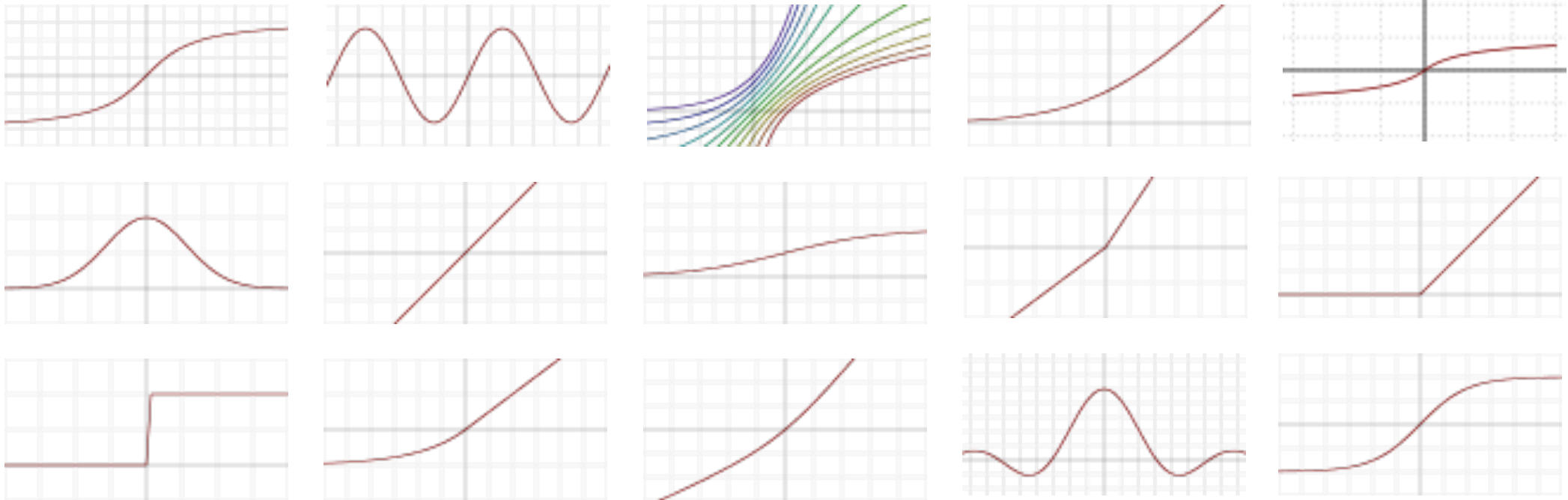
- *Énorme* saut en puissance de calcul
- Deep Blue et paradigme des « agents intelligents »

FRISE CHRONOLOGIQUE

Apprentissage en profondeur et mégadonnées (de 2001 à nos jours)

- Puissance de calcul en parallèle
- Téraoctets de données
- Réseaux de neurones profonds
- Human Brain Project (projet de 10 ans, 1,3 milliard de \$ US, de 2013 à 2023)
- Véhicules sans chauffeur?

FONCTIONS D'ACTIVATION



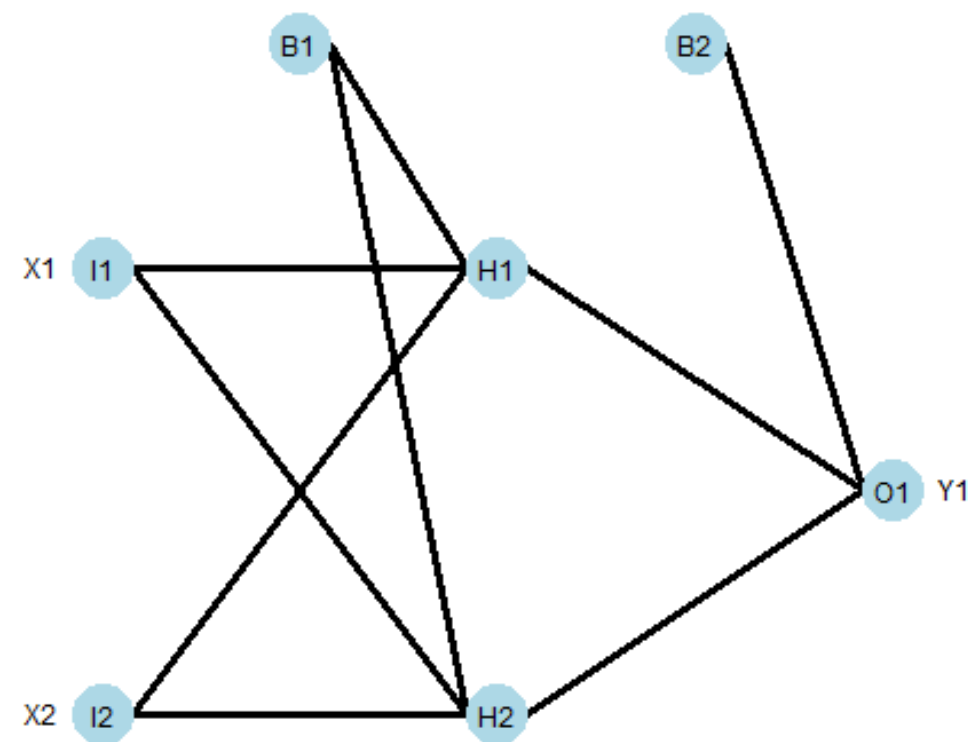
Il existe de **nombreuses** options : quel effet le choix a-t-il, le cas échéant?

BIAIS

On peut introduire un terme de **biais** b dans toutes les unités dans les couches cachées et les couches de sortie, qui font office de points d'interception pour Z .

$$Z^{(2)} = \begin{bmatrix} b_C + Z_{C,1} & b_D + Z_{D,1} \\ \vdots & \vdots \\ b_C + Z_{C,n} & b_D + Z_{D,n} \end{bmatrix} \text{ et}$$

$$Z^{(3)} = \begin{bmatrix} b_E + Z_{E,1} \\ \vdots \\ b_E + Z_{E,n} \end{bmatrix}$$



EXEMPLE – DONNÉES SUR LES VINS

Échantillon ($N = 178$) de vins élevés dans la même région de l'Italie, qui viennent tous de trois cultivars différents (**cible**).

$p = 13$ propriétés chimiques et non chimiques (**caractéristiques**) sont enregistrées.

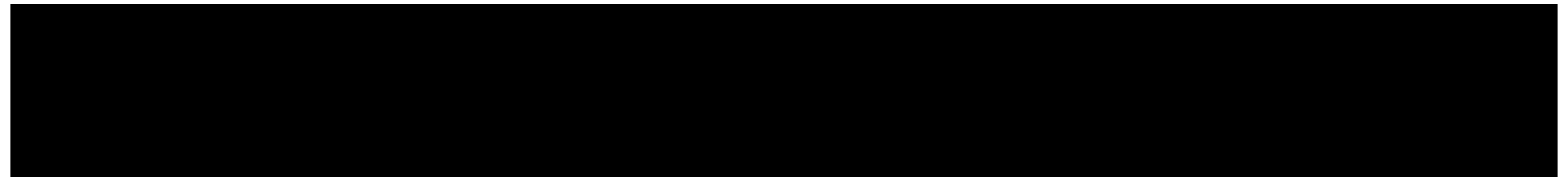
Variables :

- Alcool
- Acide malique
- Cendres
- Alcalinité des cendres
- Magnésium
- Phénols totaux
- Flavonoïdes
- Phénols non flavonoïdes
- Intensité de la couleur
- Teinte
- OD280/OD315
- Proline

<https://www.data-action-lab.com/wp-content/uploads/2019/03/Advanced.zip>

RÉFÉRENCES

INTELLIGENCE ARTIFICIELLE ET RÉSEAUX NEURONAUX



RÉFÉRENCES

« Do Androids Dream of Electric Sheep? ». Sur Internet : [*The Guardian*](#).

Welch Labs. Sur Internet : [*Neural Networks Demystified*](#).

Lawrence, S., C.L. Giles et A.C. Tsoi. *What size neural network gives optimal generalization? Convergence properties of backpropagation*, Technical Report, UMIACS-TR-96-22 et CS-TR-3617, Institute for Advanced Computer Studies, Université du Maryland, College Park, 1996.

Elissee, A., et H. Paugam-Moisy. « Size of multilayer networks for exact learning: analytic approach », *Advances in Neural Information Processing Systems 9*, Cambridge (Mass.), The MIT Press, 1997, p. 162-168.

Bengio, Y., et Y. LeCun. « Scaling learning algorithms towards AI », dans L. Bottou, O. Chapelle, D. DeCoste, J. Weston, éd., *Large-Scale Kernel Machines*, MIT Press, 2007.

A Beginner's Guide to Understanding Convolutional Neural Networks. Sur Internet : [*« Part 1 »*](#) et [*« Part 2 »*](#).

Sur Internet : [*Fitting a Neural Network in R*](#).

Sur Internet : [*Training and Visualizing a Neural Network With R*](#).

RÉFÉRENCES

Sur Internet : [*A Neural Network Playground.*](#)

Sur Internet : [*Neural Networks and Deep Learning.*](#)

Sur Internet : [*A Beginner's Guide to Neural Networks With R!.*](#)

Sur Internet : [*Neural Networks, Manifolds, and Topology.*](#)

Entrée Artificial Neural Network dans Wikipedia. Sur Internet : [https://en.wikipedia.org/wiki/Artificial_neural_network.](https://en.wikipedia.org/wiki/Artificial_neural_network)

Stergiou, C., et D. Siganos. Sur Internet : [*Neural Networks.*](#)

Sur Internet : [*Artificial Intelligence – Neural Networks.*](#)

Université Stanford. Sur Internet : [*UFLDL Tutorial.*](#)

Asimov Institute. Sur Internet : [*The Neural Network Zoo.*](#)

Sur Internet : [https://en.wikipedia.org/wiki/Reinforcement_learning.](https://en.wikipedia.org/wiki/Reinforcement_learning)

RÉFÉRENCES

Sur Internet : <http://www.scmp.com/tech/enterprises/article/2095929/alphago-vanquishes-worlds-top-go-player-marking-ais-superiority>.

Sur Internet : <https://www.theguardian.com/technology/2017/jan/30/libratus-poker-artificial-intelligence-professional-human-players-competition>.

Sur Internet : <http://www.cbc.ca/news/canada/newfoundland-labrador/artificial-intelligence-robots-battle-memorial-university-starcraft-1.4298844>.

Sur Internet : <https://www.digitaltrends.com/cool-tech/japanese-ai-writes-novel-passes-first-round-national-literary-prize/>.

Sur Internet : <https://futurism.com/a-new-ai-can-write-music-as-well-as-a-human-composer/>.

Sur Internet : <https://techcrunch.com/2016/09/01/this-trailer-made-by-ibms-watson-proves-ai-knows-how-to-creep-you-out/>.

Sur Internet : <http://www.bbc.com/news/health-36482333>.

Sur Internet : <https://www.digitaltrends.com/computing/cdc-flu-forecasting/>.

RÉFÉRENCES

Sur Internet : <https://www.newscientist.com/article/2147472-ai-spots-alzheimers-brain-changes-years-before-symptoms-emerge/>.

Goodfellow, I.J., J. Shlens et C. Szegedy. *Explaining and Harnessing Adversarial Examples*, ICLR 2015. Sur Internet : <https://arxiv.org/abs/1412.6572>.

Sur Internet : <https://arxiv.org/pdf/1710.08864.pdf>.

Sur Internet : http://blog.ycombinator.com/how-adversarial-attacks-work/?imm_mid=0f81cc&cmp=em-data-na-na-newsltr_20171115.

Sur Internet : <http://www.kurzweilai.net/>.

Sur Internet : <https://www.theguardian.com/technology/2014/jun/08/super-computer-simulates-13-year-old-boy-passes-turing-test>.

Sur Internet : <http://www.foxnews.com/tech/2016/03/25/microsoft-takes-tay-chatbot-offline-after-trolls-make-it-spew-offensive-comments.html>.

RÉFÉRENCES

Sur Internet : <http://www.dailymail.co.uk/sciencetech/article-4747914/Facebook-shuts-chatbots-make-language.html>.

Sur Internet : <http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>.

Sur Internet : <https://www.quora.com/What-is-the-Manifold-Hypothesis-in-Deep-Learning>.

Wolpert, D.H., et W.G. Macready. « No free lunch theorems for optimization », *IEEE Transactions on Evolutionary Computation*, vol. 1, 1997.

Sur Internet : <https://www.inverse.com/article/29879-knock-knock-joke-humor-ai-neural-network>. 2005.

Sur Internet : https://en.wikipedia.org/wiki/DARPA_Grand_Challenge.

Sur Internet : <https://www.cnn.com/2017/09/12/apple-unveils-a11-bionic-neural-engine-ai-chip-in-iphone-x.html>.

Sur Internet : <https://www.gizmodo.com.au/2015/09/this-neural-networks-hilariously-bad-image-descriptions-are-still-advanced-ai/>.

RÉFÉRENCES

Sur Internet : <https://www.gizmodo.com.au/2015/07/googles-neural-network-chatbot-can-discuss-philosophy-and-it-troubles/>.

Sur Internet : <https://www.gizmodo.com.au/2015/08/this-neural-network-is-hilariously-bad-at-describing-outer-space/>.

Sur Internet : <https://splinternews.com/these-bots-wrote-and-delivered-their-own-ted-talks-sor-1793848398>.

Moosavi-Dezfooli, Seyed-Mohsen, Alhussein Fawzi, Omar Fawzi et Pascal Frossard. *Universal adversarial perturbations*, 26 octobre 2016. Sur Internet : <https://arxiv.org/abs/1610.08401>.

Sur Internet : <https://www.cnet.com/news/elonmusk-artificial-intelligence-world-war-iii-russia-china/>.

Sur Internet : <https://www.technologyreview.com/s/546301/will-machines-eliminate-us/>.

Sur Internet : <https://www.technologyreview.com/s/609070/inside-the-moonshot-effort-to-finally-figure-out-the-brain/>.

« Is AI riding a one-trick pony? », *MIT Technology Review*, 29 septembre 2017. Sur Internet : <https://www.technologyreview.com/s/608911/is-ai-riding-a-one-trick-pony/>.

RÉFÉRENCES

Axios, 15 septembre 2017. Sur Internet : <https://www.axios.com/ai-pioneer-advocates-starting-over-2485537027.html>.

TechCrunch, 30 septembre 2017. Sur Internet : <https://techcrunch.com/2017/09/30/ai-hype-has-peaked-so-whats-next>.

Fortune, 28 juin 2017. Sur Internet : <http://fortune.com/2017/06/28/ibm-watson-ai-healthcare/>.

Sur Internet : https://stuff.mit.edu/people/mitter/publications/121_Testing_Manifold.pdf.

Sur Internet : <https://colah.github.io/posts/2014-03-NN-Manifolds-Topology/>.

Sur Internet : <http://robots.stanford.edu/papers/thrun.stanley05.pdf>.

Sur Internet : <https://medium.com/@samim/generating-captions-c31f00e8396e>.

Kolakota, R., M. Robinson et D. Tapscott. *E-business 2.0: Roadmap for Success*, Addison-Wesley, 1999.

Reichheld, F.F. « Loyalty-based management », *Harvard Business Review*, mars-avril 1993, p. 64-73.

Sur Internet : <https://distill.pub/2017/aia/>.

Sur Internet : <https://distill.pub/2018/building-blocks/>.