

Document pédagogique pour le PdS

Groupe 8 PdS ULB

March 2022

SADOUNE Zakari

ZHU Yi

OUDAHYA Ismail

HALIT Wacim

BOTOND Horvath

DELANGRE Matthieu

GOREN Serhat

LATOUNDJII Salim

NACHTEGAEL Charlotte

1 Introduction

Aujourd'hui, au 21ème siècle, grâce aux nouvelles technologies et à la digitalisation du savoir humain, l'exploration prend une toute nouvelle dimension. Ce sont maintenant les machines qui recherchent à partir de grandes bases de données pour l'utilisateur, en se basant sur la question que ce dernier pose. C'est ce qu'on appelle un système à question/réponse.

2 Compréhension de la question posée

Vous êtes-vous déjà demandé comment des assistants personnels IA tels que Siri ou Cortana fonctionnent ? Comment votre correcteur d'orthographe a été capable de détecter des erreurs de syntaxe que vous-même n'auriez pas repérées ?

Si ces outils ont des fins différentes, ils reposent tous sur le même principe, celui du Natural Language Processing (NLP), ou autrement dit le traitement du langage naturel.

Le NLP est réellement le point combinant l'informatique et la linguistique. C'est donc une discipline qui porte essentiellement sur la compréhension (détection de mots clés), la manipulation et la génération du langage naturel par les machines.

3 Apprentissage automatique

C'est un champ d'étude de l'intelligence artificielle qui se fonde sur des approches mathématiques et statistiques pour donner aux ordinateurs la capacité d'apprendre à partir de données (base de données comme dans notre projet celle de wikipédia FR), c'est-à-dire d'améliorer leurs performances à résoudre des tâches sans être programmés pour faire les tâches en question.

Le Deep Learning est une technique d'apprentissages automatique qui est un procédé de plus bas niveau avec des concepts plus internes à la machine, dans lequel la machine simule le fonctionnement d'un cerveau humain, en utilisant ce qu'on appelle des réseaux de neurones artificiels.

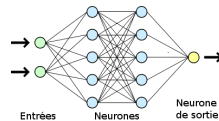


Figure 1: Réseau de neurones artificiel