



UNIVERSITÉ
LIBRE
DE BRUXELLES

INFO-F311 - PROJET D'IA 3

APPRENTISSAGE PAR RENFORCEMENT

Auteur:

Manuel ROCCA - 000596086

Professeurs:

Tom LENAERTS

Assistants:

Axel ABELS

Martin COLOT

Yannick MOLINGHEN

Pascal TRIBEL

Année académique 2025-2026

Contents

1	Introduction	2
2	Cadre expérimental	2
3	Résultats	2
4	Analyse	2
5	Conclusion	2

1 Introduction

Ce quatrième et dernier projet dans le cadre du cours INFO-F311 d'Intelligence Artificielle nous amène à étudier et implémenter en pratique un réseau de neurone, en particulier un *auto-encoder*. Nous présentons ci-dessous notre démarche scientifique.

2 Cadre expérimental

Comme exprimé dans l'introduction, nous nous penchons ici sur un *auto-encoder*. Un réseau neuronal tel que celui-ci est composé de deux parties : une qui encode et une qui décode. En particulier, il apprend, sur base d'un ensemble de données d'entraînement, à réduire la dimension, encoder la donnée en entrée et à la décoder.

Le jeu de données est un ensemble de chiffres de format 28x28 pixels appartenant à la base de données MNIST. Il est composé de deux fichiers :

- Un fichier d'entraînement de 60_000 éléments.
- Un fichier de test de 10_000 éléments.

La valeur de chaque pixel varie entre 0 et 255 mais sont normalisées à une valeur entre 0 et 1 dans notre programme pour améliorer l'efficacité et la précision du modèle (notamment en empêchant les valeurs comme 255 de dominer une petite valeur comme 2 lors d'une multiplication).

Le réseau neuronal a besoin de plusieurs paramètres en entrée :

- La dimension des données d'entrée, 784 dans notre cas (matrices 28x28 linéarisées).
- La dimension des données encodées \hat{x} , paramètre variable étudié dans les sections suivantes.
- Le *learning rate*, paramètre déjà étudié dans les autres projets, en particulier le 3e. Il nous semble pertinent de faire l'analogie avec les pas d'itération dans les simulations numériques.

Finalement, pour lancer l'entraînement de cet *auto-encoder*, il nous faut :

- Les données d'entraînement obtenues dans le csv.
- L'*epoch* correspondant au nombre de passages complets de toutes les données dans le réseau.
- Le *batch size* correspondant au nombre de paramètres/d'échantillons traités avant une mise à jour des valeurs du réseau.

3 Résultats

4 Analyse

5 Conclusion