

|  |
| --- |
| Intro to Machine Learning - Polytech - 2024/2025 |
| MNIST - Deep Neural Network with Keras |
| Badr TAJINI |

|  |
| --- |
| Maximilien BOURCIER Pierre CHEN Téo SAMUEL Alice SPACCESI  24/02/2025 |

Table des matières

[Introduction 2](#_Toc191414788)

[1. Préparation des données 2](#_Toc191414789)

[2. Architecture du Réseau de Neurones 3](#_Toc191414790)

[3. Entraînement du Modèle 3](#_Toc191414791)

[4. Optimisation et Éviter le Surapprentissage 4](#_Toc191414792)

[5. Évaluation et Analyse des Résultats 4](#_Toc191414793)

[5. Conclusion et Perspectives 5](#_Toc191414794)

# Introduction

Ce rapport présente en détail l’ensemble des étapes menant à la reconnaissance automatique des chiffres manuscrits à partir du dataset MNIST. Le projet a pour objectif de démontrer comment un réseau de neurones multicouches peut être entraîné pour identifier des chiffres écrits à la main.

# Préparation des données

Dès le départ, nous avons commencé par préparer les données, qui se composent de 70 000 images réparties en 60 000 pour l’entraînement et 10 000 pour le test. Chaque image, initialement au format 28×28 pixels en niveaux de gris, a été transformée en un vecteur de 784 valeurs normalisées entre 0 et 1. Par ailleurs, afin de faciliter l’apprentissage du modèle, les étiquettes numériques ont été converties en vecteurs à l’aide de la technique de l’encodage one-hot, transformant ainsi chaque chiffre en un vecteur de 10 éléments. Cette transformation permet au modèle de donner une probabilité pour chaque classe, et ainsi de choisir la classe ayant la plus forte probabilité.

Une image contenant texte, Police, nombre, noir et blanc

Le contenu généré par l’IA peut être incorrect.

*Figure 1 : Exemples d’images d’entrainement*

Dans cette première figure, plusieurs images issues du dataset MNIST sont présentées. Chaque image met en évidence la diversité des écritures manuscrites et est accompagnée de sa valeur réelle. La variation des traits et des formes illustre parfaitement le défi que représente la reconnaissance de chiffres, dont l’apparence peut varier considérablement d’un individu à l’autre.

Une image contenant texte, carré, Police, motif

Le contenu généré par l’IA peut être incorrect.

*Figure 2 – Exemples d’images de test*

# Architecture du Réseau de Neurones

Le modèle que nous avons conçu repose sur un réseau de neurones multicouches (MLP). La couche d’entrée, composée de 784 neurones, reçoit directement les pixels aplatis de chaque image. Les deux couches cachées, chacune regroupant 256 neurones et utilisant la fonction d’activation ReLU, permettent d’extraire des caractéristiques essentielles telles que les contours, les courbes et autres motifs caractéristiques des chiffres. Enfin, la couche de sortie, constituée de 10 neurones activés par une fonction Softmax, fournit une probabilité pour chaque classe de chiffre. Un schéma de l’architecture illustre la manière dont les données transitent et se transforment d’une couche à l’autre, soulignant l’importance de l’extraction progressive des informations.

# Entraînement du Modèle

L’entraînement du modèle s’est déroulé de manière itérative, avec un passage complet sur le jeu d’entraînement appelé époque. À chaque époque, le modèle effectue des prédictions sur l’ensemble des 60 000 images, compare ces prédictions aux étiquettes réelles, et ajuste progressivement ses poids grâce à la rétropropagation. Au début, le taux d’erreur est relativement élevé, mais il diminue de façon significative au fil des époques. Dans notre expérience, la meilleure performance a été obtenue après 17 époques, moment où le processus d’entraînement a été interrompu afin d’éviter le surapprentissage.

Une image contenant texte, capture d’écran, ligne, Tracé

Le contenu généré par l’IA peut être incorrect.

*Figure 3 : Courbes d’Évolution de la Précision et de la Perte*

Le graphique présenté dans cette figure trace l’évolution des performances du modèle durant l’entraînement. On y observe la diminution progressive de la perte ainsi que l’augmentation de la précision au fil des époques. Cette visualisation permet d’identifier le moment où les améliorations se stabilisent, justifiant ainsi l’arrêt de l’entraînement pour éviter le surapprentissage.

# Optimisation et Éviter le Surapprentissage

Afin de garantir une bonne généralisation du modèle sur des données inédites, nous avons mis en place des techniques spécifiques d’optimisation. Le dropout, appliqué avec un taux de 45 %, permet de désactiver aléatoirement une partie des neurones à chaque itération, ce qui empêche le réseau de s’appuyer excessivement sur certaines connexions et améliore ainsi sa robustesse. De plus, la réduction automatique du taux d’apprentissage, mise en œuvre par le mécanisme ReduceLROnPlateau, ajuste le processus d’optimisation lorsque l’amélioration stagne. Ces techniques combinées ont permis d’atteindre une précision finale de 98,2 % tout en évitant que le modèle ne surapprenne.

# 5. Évaluation et Analyse des Résultats

L’évaluation du modèle a été réalisée à l’aide de 10 000 images de test. Pour analyser la performance du modèle, une matrice de confusion a été générée, mettant en évidence les erreurs de classification, notamment les confusions entre des chiffres visuellement similaires. Des exemples précis de prédictions ont également été étudiés afin d’identifier les cas où le modèle se trompait, souvent en raison d’une écriture ambigüe ou de ressemblances morphologiques entre certains chiffres. Ces analyses ont permis de mieux comprendre les limites du modèle et d’orienter les perspectives d’amélioration.

Une image contenant texte, capture d’écran, diagramme, nombre

Le contenu généré par l’IA peut être incorrect.

*Figure 4 : Matrice de Confusion*

Cette figure affiche la matrice de confusion obtenue lors de l’évaluation du modèle sur le jeu de test. Elle compare les prédictions du modèle aux classes réelles et met en lumière les erreurs de classification, notamment les confusions fréquentes entre des chiffres visuellement similaires comme 4 et 9 ou 7 et 2. Cette représentation est essentielle pour comprendre les limites du modèle et orienter les pistes d’amélioration.

# Conclusion et Perspectives

Ce projet nous a permis de comprendre comment un réseau de neurones apprend à reconnaître des chiffres manuscrits. Nous avons vu comment préparer les données, concevoir un modèle, l'entraîner, et l’évaluer. Avec une précision de 98.2%, notre modèle est très performant. Cependant, certaines erreurs persistent, notamment entre chiffres proches visuellement.

En conclusion, ce projet a montré que les réseaux de neurones sont des outils puissants pour la reconnaissance d’images, et que des techniques d'optimisation bien appliquées permettent d’atteindre des performances impressionnantes.

(Source pour mieux comprendre certains concepts : ChatGPT)