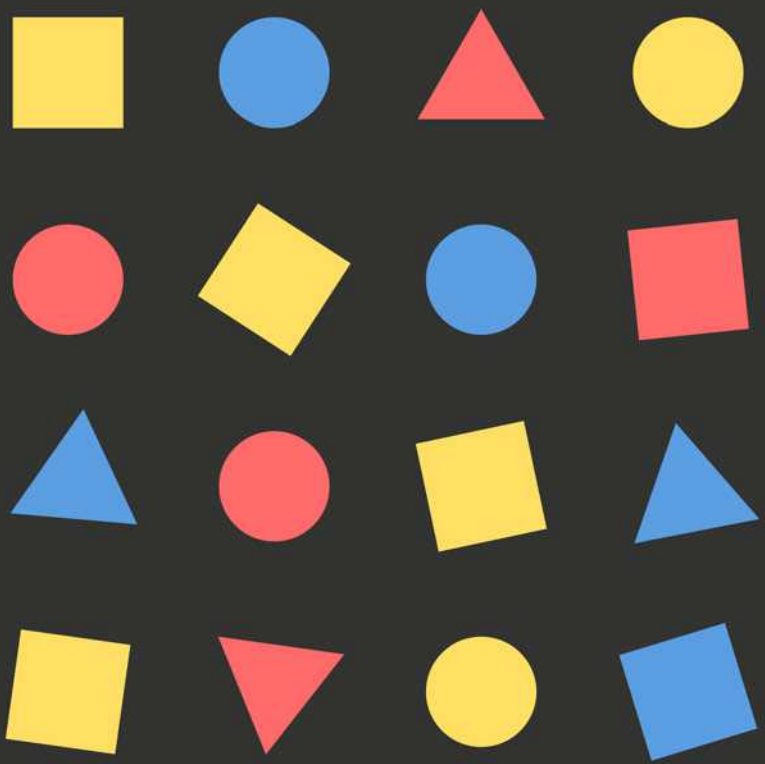
Andriy Burkov

LES LLM EN PRATIQUE
AVEC PYTHON ET PYTORCH

Comprendre les modèles de langage



DUNOD

Andriy Burkov

LES LLM EN PRATIQUE

AVEC PYTHON ET PYTORCH

Comprendre les modèles de langage

Traduit de l'anglais par **ANNE BOHY**

DUNOD

L'édition originale de ce livre a été publiée au Canada par True Positive Inc. sous le titre *The Hundred-Page Language Models Book*.
Copyright © 2025 Andriy Burkov. Tous droits réservés.

Maquette intérieure : Melissa Jallier-Lundgren

© Dunod, 2026 pour l'édition française
11 rue Paul Bert, 92240 Malakoff
www.dunod.com
ISBN : 978-2-10-090268-2

« À ma famille, avec toute mon affection. »

« Le langage est source de malentendus. »
– Antoine de Saint-Exupéry, *Le Petit Prince*

« En mathématiques, on ne comprend pas les choses.
On finit simplement par s'y habituer. »
– John von Neumann

« Les ordinateurs sont inutiles. Ils ne peuvent
que donner des réponses. »
– Pablo Picasso

Sommaire

Avant-propos 11

Chapitre 1 • Notions fondamentales 17

1 IA et apprentissage automatique 15
2 Modèle 19
3 Processus en quatre étapes 26
4 Vecteur 26
5 Réseau de neurones 30
6 Matrice 34
7 Descente de gradient 36
8 Différentiation automatique 41

Chapitre 2 • Modélisation du langage 45

1 Sac de mots 45
2 Plongements de mots 56
3 Codage par paires d'octets 61
4 Modèle linguistique 66
5 Modèle linguistique basé sur le comptage 67
6 Évaluation des modèles linguistiques 73

Chapitre 3 • Réseaux de neurones récurrents 85

1 RNN d'Elman 85
2 Descente de gradient par mini-lots 87
3 Codage d'un RNN 88
4 Le RNN en tant que modèle linguistique 90
5 Couche de plongement 91
6 Entraînement d'un modèle linguistique RNN 93
7 Dataset et DataLoader 96
8 Données d'entraînement et calcul de perte 98

Chapitre 4 • Transformer 101

1 Bloc décodeur 102
2 Auto-attention 103
3 Perceptron multicouche positionnel 107
4 Plongement positionnel rotatif 108
5 Attention multi-têtes 114
6 Connexion résiduelle 116
7 Normalisation par la moyenne quadratique 119
8 Mise en cache des clés-valeurs 120
9 Transformer en Python 121

Chapitre 5 • Grands modèles de langage 129

1 Pourquoi plus c'est grand, mieux c'est 129
2 Réentraînement ciblé supervisé 136
3 Réentraîner un modèle pré-entraîné 137
4 Échantillonnage à partir de modèles linguistiques 150
5 Adaptation de rang faible (LoRA) 154
6 LLM en tant que classificateur 157
7 Ingénierie des prompts 159
8 Hallucinations 164
9 Les LLM, le droit d'auteur et l'éthique 167