

# Formation Introduction au Deep Learning pour le traitement du langage

2 jours (14 heures)

## Présentation

Le Deep Learning a révolutionné un grand nombre d'approches concernant le langage : traduction, identification d'éléments, interprétation, systèmes de dialogue, etc. Néanmoins, ces nouvelles pratiques demandent une certaine connaissance des outils, pratiques et limites propres à cette approche. Cette formation vise à donner les clés techniques d'implémentation et d'utilisation de ces nouvelles techniques.

Un premier point d'attention concerne ainsi les architectures de réseaux de neurones principalement utilisées dans ce domaine avec, pour chacune, les différentes implémentations et modes d'utilisation de ces outils. L'ensemble des principales tâches NLP (Natural Language Processing) sont ainsi parcourues en donnant pour chacune les solutions existant à date. Les modèles d'état de l'art sont abordés en détail dans un second temps, avant un parcours des solutions à date permettant l'interprétation, voire la correction d'un modèle à base de réseaux de neurones.

## Objectifs

- Comprendre en détail les architectures fondamentales du deep learning appliqué au traitement du langage
- Maîtriser les différentes implémentations et outils

## Public visé

Développeurs, Architectes, Big Data Data Analyst / Data Engineer / Data Scientist

## Pré-requis

- Connaissance de Python et en mathématique

# Pour aller plus loin

- ? Notre certification au [Deep Learning](#) ?
- Nous vous proposons en introduction un formation sur l'[Intelligence Artificielle](#)
- En complément la technologie
  - [TensorFlow](#) de Google
  - [Pytorch](#) de Facebook

## Programme de la formation Introduction au Deep Learning pour le traitement du langage

### [JOUR 1]

#### 1. Introduction Deep Learning

- Rappel de bases mathématiques.
- Définition d'un réseau de neurones : architecture classique, fonctions d'activation et de pondération des activations précédentes, profondeur d'un réseau.
- Définition de l'apprentissage d'un réseau de neurones : fonctions de coût, backpropagation, stochastic gradient descent, maximum likelihood.
- Modélisation d'un réseau de neurones : modélisation des données d'entrée et de sortie selon le type de problème (régression, classification...). Curse of dimensionality. Distinction entre donnée multi-features et signal. Choix d'une fonction de coût selon la donnée.
- Généralisation des résultats d'un réseau de neurones.
- Initialisations et régularisations d'un réseau de neurones : L1/L2 regularization, Batch Normalization, Instance Normalization
- Optimisations et algorithmes de convergence

#### 2. Génération d'embeddings comme modèle de texte

- Définition d'un embedding : transformation non supervisée d'un vocabulaire avec conservation sémantique.
- Approche Word2Vec (skip-gram) : génération de représentations au niveau du mot, de la phrase ou du paragraphe.
- Framework FastText (Facebook) : approches et utilisation en détection de sentiment.
- Applications directes, modélisation du langage. Spécificités d'un vocabulaire et limites.

#### 3. Architectures fondamentales NLP : réseaux récurrents

- Présentation des Recurrent Neural Networks : principes fondamentaux et applications.
- Fonctionnement fondamental du RNN : hidden activation, back propagation through time, unfolded version.
- LSTM (Long Short Term Memory). Évolution d'architecture, flux de gradient. Approche Bi-LSTM.

- Architecture RNN Encoder Decoder.
- Applications NLP basiques (Détection de sentiment, classification).
- Skip Connections et construction résiduelle.
- Étude de l'approche et du design du Google Neural Machine Translation. Parallélisation, skip-connections.

#### 4. Architectures fondamentales NLP : réseaux convolutionnels

- Présentation de l'architecture fondamentale d'un layer CNN : convolution, stride, pooling.
- Fonctions de non-linéarité usuelles.
- Utilisation du pooling.
- Problématiques de classification : fonctions de coût, approche probabiliste.
- Applications NLP Basiques

#### 5. Modèles d'attention et comparaisons entre approches convolutionnelles et récurrentes

- Modèle d'attention pour un réseau récurrent.
- Utilisation d'un CNN Sequence to Sequence avec un modèle d'attention. Choix d'architecture, exemple d'implémentation. Application : traduction automatique.
- Comparaisons entre approches récurrentes et convolutionnelles. Considérations en qualité et en coût de performance selon les problèmes et approches.

### [JOUR 2]

#### 6. Revue d'applications et d'architectures au traitement du langage : définitions, architectures et résultats

- Identification d'éléments de langage (Part of speech tagging) : Bidirectional LSTMs, Conditional Random Fields.
- Analyse de dépendances et structuration (Parsing) : Stack LSTMs.
- Reconnaissance d'éléments (Named entity Recognition) : Bidirectional LSTM, Dilated CNN et CRFs.
- Détection de sentiment : Tree-LSTM avec amélioration des embeddings.
- Réponse à des questions libres et compréhension : Memory Networks, QANet (convolutions locales et self-attention globale).
- Système de dialogues : Dual LSTM Encoder with semantic matching, Sentence-level CNN-LSTM Encoder.

#### 7. Approches récentes et états de l'art

- Approche Google "Attention is all you need"
- Cellule Transformer : nouvelle cellule Deep Learning de gestion de séquences. Détail de l'architecture, revue d'une implémentation.
- OpenAI ELMO : description de l'architecture et des avancées.
- Description du modèle BERT (Google) : état de l'art et Transfert Learning pour le NLP. Exemples d'utilisation.a

## 8. Visualisation, sécurisation et interprétation d'un modèle : outils

- Analyse du fonctionnement d'une cellule RNN. Maximums d'activation.
- Visualisation des modes d'attention, distinction des dépendances entre éléments en entrée et éléments en sortie.
- Outil Havard NLP Seq2Seq-Viz de visualisation du fonctionnement d'un modèle. Cas d'étude d'erreurs d'interprétation.
- Hierarchical Contextual Decompositions.
- Attaques adversariales : facilité de confusion d'un réseau de neurones. Étude du cas spécifique du langage.

## 9. Génération de texte libre ou conditionnée

- Présentation de l'architecture du Variational AutoEncoder : évolution de l'autoencoder, approche probabiliste, reparameterization trick.
- Application VAE à la génération de texte.

Références : • Generating Sentences from a Continuous Space, Bowman et al. • Toward Controlled Generation of Text, Hu et al.

- Présentation de l'approche GAN (Generative Adversarial Networks) : architecture générale, principe de convergence
- Applications GAN à la génération de texte et limites. Questions de métriques internes (Earth Mover Distance)
- Présentation de l'approche Deep Reinforcement Learning et applications à la génération de textes.

## Sociétés concernées

Cette formation s'adresse aux entreprises, petites ou grandes, souhaitant former ses équipes à une nouvelle technologie informatique avancée.

## Méthodes pédagogiques

Stage Pratique : 60% Pratique, 40% Théorie. Support de la formation distribué au format numérique à tous les participants.

## Organisation

Le cours alterne les apports théoriques du formateur soutenus par des exemples et des séances de

réflexions, et de travail en groupe.

## Validation

À la fin de la session, un questionnaire à choix multiple permet de vérifier l'acquisition correcte des compétences.

## Sanction

Une attestation sera remise à chaque stagiaire qui aura suivi la totalité de la formation.