

Christopher Kermorvant

Expert R&D Intelligence Artificielle et Machine Learning

Expérience Professionnelle

depuis novembre
2014

Consultant R&D, Teklia, Paris.

Consultant indépendant auprès de start-ups, PME et grands groupes.

- ▷ **Chercheur associé au laboratoire LITIS de Rouen**: Deep Learning pour la reconnaissance d'écriture.
- ▷ **Paléographie numérique** : reconnaissance d'écriture et indexation de manuscrits médiévaux dans le cadre des projets ANR Oriflamms et Européen HIMANIS.
- ▷ **Analyse d'image** : supervision pour la société A2iA d'une thèse en traitement d'image et Deep Learning pour l'analyse de documents numérisés (cotutelle INSA de Lyon);
- ▷ **Recommandation musicale personnalisée** : développement pour la société Playlist4me d'une application web de recommandation musicale, mise en place de l'infrastructure (base de donnée musicale, interactions et retours utilisateurs), création automatique de playlists, système de recommandation;
- ▷ **Traduction automatique** : pour la société DeepLingo, analyse des besoins clients, collecte de données textuelles (webcrawling), préparation et normalisation des données, modèles de langues, analyse des performances;

2008-2014

Responsable R&D, A2iA, Paris.

Responsable d'une équipe de recherche.

- ▷ Direction d'une équipe de 5 chercheurs sur des thématiques de reconnaissance d'écriture, extraction d'information et classification automatique de documents;
- ▷ Gestion de projets de recherche : rédaction, planification, suivi et reporting des projets de recherche internes et externes pour le produit DocumentReader et les fonctionnalités de reconnaissance de l'écriture;
- ▷ Responsable pour A2iA des projets ANR Cognilego, ANR Oriflamms, Cap-Digital Dem@tFactory, OSEO-Quaero, DGA-Maurdor, Investissements d'avenir PACTE, pour un budget total de 2 M€ (2010-2014);
- ▷ Veille technologique et publications scientifiques en apprentissage automatique statistique et reconnaissance d'écriture
- ▷ Co-encadrement de deux thèses avec Telecom ParisTech (Reconnaissance d'écriture par HMM) et LIMSI-CNRS (Deep learning pour la reconnaissance d'écriture)

2008-2011

Chargé de cours à l'université du Sud -Toulon Var, Master SIS, Toulon.

Traitement d'image et apprentissage automatique statistique pour le traitement des documents numérisés.

2005-2007

Ingénieur Recherche et Développement, A2iA, Paris.

Recherche, développement logiciel et veille technologique.

- ▷ Responsable du développement de la fonctionnalité de classification automatique de documents dans le produit DocumentReader : conception d'algorithmes de classification, extraction de caractéristiques, développement en C++ et python, suivi de l'intégration dans le produit, démonstration et benchmark chez les clients;

- ▷ Veille technologique apprentissage automatique statistique;
- ▷ Implication dans la définition de l'architecture logiciel du produit DocumentReader et dans la mise en place de procédures de qualité du développement (tests, compilation continue);
- ▷ Encadrement de stagiaires sur des projets de recherche exploratoires.
- 2003–2004 **Chercheur Post-Doctoral**, *Laboratoire MILA (dir. Y. Bengio), Université de Montréal, Montréal, Canada.*
- ▷ En charge d'un projet d'apprentissage automatique de réseaux bayesiens pour la désambiguïsation sémantique;
- 2000–2003 **Chercheur Doctorant**, *Laboratoire Hubert Curien, Université Jean Monnet, Saint-Etienne.*
- ▷ Recherche en inférence grammaticale statistique appliquée à la modélisation de séquences linguistiques et biologiques; Enseignement en informatique théorique et pratique.
- 1998–1999 **Ingénieur de Recherche**, *Institut Dalle Molle d'Intelligence Artificielle Perceptive (IDIAP, dir. H. Bourlard), Martigny, Suisse.*
- ▷ Recherche et développement : modèles de débruitage pour la reconnaissance de la parole.

Formation

- 2000–2003 **Doctorat en Informatique**, *Université Jean Monnet, Saint-Etienne, France.*
Apprentissage de modèles à états finis stochastiques pour les séquences
- 1995–1996 **Master of Science**, *The Manchester University, Manchester, Royaume-Uni.*
Testing a model of central auditory processing: self-organisation of pitch categories
- 1993–1996 **Ingénieur en Informatique**, *Ecole Nationale Supérieure d'Informatique pour l'Industrie et l'Entreprise (ENSIIE), Evry, France, Admission sur concours Centrale-Supélec.*

Compétences en intelligence artificielle et développement logiciel

Machine Learning	Apprentissage statistique, réseaux de neurones, <i>deep learning</i> , HMM, boosting, arbres et random forests, SVM, transducteurs pondérés, <i>active learning</i> , clustering. Bibliothèques : python, numpy, pandas, scikit-learn, tensorflow
Reconnaissance parole/écriture	Modèles acoustiques et optiques (traitement du signal et d'image), décodage (Viterbi), apprentissage (algorithme EM). Bibliothèques : kald, opencv
Traitement de la langue naturelle	Tokenisation, analyse syntaxique, modèles de langage (n-grammes, automates), modèles sémantiques (ontologie), classification de textes, traduction automatique statistique. Bibliothèques : openfst, srilm, nltk, mooses, wordnet
Développement logiciel	Adept des méthodes agiles, de l'intégration continue et des revues de code. Usage quotidien de python et redmine.

Communication scientifique

- Publications 60 articles en conférences et journaux nationaux/internationaux (Google Scholar).
- Compétitions internationales Reconnaissance d'écriture : première place aux compétitions OpenHaRT 2010 et 2013 (arabe), ICDAR-2011 (français), Maudor 2013 (français, anglais, arabe)
- Missions d'expertise Expert pour l'Agence d'évaluation de la recherche et de l'enseignement supérieur AERES, l'Agence nationale de la recherche ANR et le réseau européen d'excellence Multilingual Europe Technology Alliance (META-NET).
Membre du comité de programme des conférences CAp, CIFED, ICFHR, DRR, DAS.