

PHILÉAS CONDEMINE

En bref

2021-Présent **Lead ML-Engineer NLP @ Covéa** : application de l'IA et IA générative pour valoriser les données textuelles sur le cloud Azure Databricks.

2018-2021 **Senior Data Scientist @ Ministère de la Santé** : modélisation des **parcours de soins** par **Deep-learning** et renfort au **centre de crise sanitaire CoViD-19**.

2014-2017 **Actuarial Data Scientist @ AXA** Global P&C : **tarification** et **gestion des sinistres** grâce aux **machine learning** et **données externes**.

[View in english](#)

[Télécharger en PDF sur github](#)



EXPÉRIENCES PROFESSIONNELLES

2021-
Aujourd'hui

Lead ML-Engineer NLP



COVEA - AI Delivery

Paris, France

- **VOX.IA** : analyse de verbatim clients. **Gagnant Cas d'Or IA**
- **NetMessages** : proposition de selfcare en temps réel.
- **JudiCible** : Accélération de l'ouverture de litige juridique par IA Générative.
- **IA en production** : **Annotation, Finetuning, Déploiement, Monitoring** du *drift*.
- **Publication** sur le [blog Medium COV&IA](#)

2018-
2021

Senior Data Scientist et Entrepreneur d'Intérêt Général EIG



Ministère de la Santé et Centre Crise Sanitaire CoViD-19

Paris, France

- **Modélisation de séquences de soins** par **deep-learning Transformers** et détection de ruptures dans les séquences de soins.
- **Classification** par **active-learning** et **diffusion** d'indicateurs de santé.
- Outil pour les ARS : **saisie interactive** du zonage des professions de santé
- Support au **Centre de Crise Sanitaire CoViD-19** : collecte de données **hôpitaux** et **laboratoires**.
- **Workflow** de calcul d'indicateurs **CoViD-19** quotidien sur **SIDep, SIVic, VacSI**.
- Production d'*open-data* hospitalières avec **anonymisation** : k-anonymat et l-diversité.
- Mise en place de **formations à R et Shiny** pour les agents de la DREES

2014-
2017

Actuarial Data Scientist



AXA Global P&C

Paris, France

- **Innovation dans la tarification IARD** : amélioration **prime pure, zonier, véhiculier** grâce au *gradient-boosting* et à l'interprétation du *machine-learning* **IML** ou **XAI**.
- Analyse du **coût des sinistres** : déploiement d'un du coût théorique pour *scorer* les garages en tenant compte du *case-mix*.
- Utilisation de données textuelles pour la **gestion des sinistres litigieux**.
- **Modèle de risque des routes** basé sur les données **Télématiques** des utilisateurs.
- Mise en place de **Hackathons et Formations** à la Data Science pour les Actuaire.

2014

Mémoire d'actuariat



AXA Belgium

Télétravail

Sinistralité des conducteurs âgés - mémoire finaliste du prix **SCOR**.

CONTACT

- [phileas.condemine](#)
- [github.com](#)
- [gitlab.com](#)
- [hackerrank.com](#)
- [twitter.com](#)
- [linkedin.com](#)
- +33 643 549 576

COMPÉTENCES

Expert en Python et R et confirmé en SQL & NoSQL.

IA Générative : Résumé, Auto-label, Auto-define.


Deep-learning : **Transformers & Pytorch** pour la classification

Machine-learning GLM, XGBoost/GBM, SVM : tarification, gestion de sinistres, fraude.

Traitement de données massives en **Spark**.

Dashboard en Python (streamlit) ou R (Shiny).

2013 ● **Stagiaire Structuration Produits Dérivés (6 mois)**
 **Exane BNP Paribas**
□ Paris, France
Construction d'un indice synthétique optimal sur panier d'actions

2012 ● **Stagiaire Réassurance CAT-Bonds (6 mois)**
 **SCOR**
□ Paris, France
Tarification de CAT-Bonds par techniques Monte-Carlo Markov Chains MCMC

FORMATION



● **IA Générative**
RAG, Finetuning, Prompt-engineering grâce aux formations [deeplearning.ai](#)
□ 2023-2024

2017-2020 ● **Deep learning**
Découverte, approfondissement et mise en pratique des techniques de *deep learning*.
□ [fast.ai](#), [datascientest](#), [deeplearning.ai](#)

2019 ● **Spark & Scala**
Formation au langage Scala et à la programmation avec Spark & PySpark
□ Coursera par Martin Odersky & Heather Miller

2016 ● **Formation aux langages web**
Introduction à HTML, CSS, Javascript & JQuery.
□ CodeSchool.com

2014 ● **Introduction à la data science**
Apprentissage supervisé et non supervisé : *Support Vector Machine*, *Gradient Boosting*, *Random Forests*, *k-means* etc.
□ Coursera par Bill Howe

2010-2014 ● **ENSAE Paris - IP Paris**
  Cursus Ingénieur : Spécialisation **Actuariat**, cours **Data Science et Big Data**.
□ Paris, France

ENSEIGNEMENT

2023-2025 ● **Masterclass IA générative (interne COVEA)**
Origines et fonctionnement de l'IA Générative. Mise en pratique sur de nombreux cas d'usage. **Prompt-engineering**.
□ COVEA

2018-2020 ● **Formation au text-mining**
Techniques "classiques" de traitement du texte : tokenisation, lemmatisation, vectorisation, *embedding*.
□ CEPE ENSAE/ENSAI

2018-2020 ● **Formation au machine learning**
Techniques de classification pour des données structurées : *SVM*, *Gradient Boosting*, *Random Forests*
□ CEPE ENSAE/ENSAI

2014-2017 ● **Data Science for Actuaries DS4A**
Bootcamps de formation à la Data Science d'une semaine. Exemple : [text mining](#). Collaborations avec [Arthur Charpentier](#).
□ DS4A AGPC

AUTRES PROJETS

Participation à des compétitions [Kaggle](#) : AXA "Telematics", Otto "Product Classification", Quora "Deduplication", West Nile Virus "Mosquito detection".

Hackathon Covéa IA
Générative pour la traduction de code SAS vers Python en (oct. 2024)

Hackathon ACPR TechSprint IA Générative
(fév. 2024)

Meilleur modèle au Hackathon Covea sur le thème de la résilience en (2023)

Hackathon ARS-IdF
"Facteurs environnementaux d'augmentation des passages aux urgences" (2019)

Hackathon APHP "Anomalies de prise en charge nocturne pour soins intensifs" (2018)

Hackathon AXA "Chatbot de fast-quote API.AI" (2017)

Cartographie des [Space Invaders à Paris](#)

Création d'un robot 4-roues autonome programmé avec Arduino.

Contribution à la [Buloterie](#) - un outil communautaire d'identification d'experts et curieux dans toutes sortes de domaines.

Ce CV a été généré grâce au package R [pagedown](#).

Mis à jour le 2025-02-09.