



ADRIEN GODOY

DATA SCIENTIST
INGÉNIEUR IA



CONTACT



Téléphone
06 50 50 11 09



Email
mr.a.godoy@gmail.com



Adresse
31200 Toulouse



GitHub
<https://vu.fr/anodegrindyo>



FORMATION



2022 - 2023
**Data Scientist -
ingénieur IA**
Openclassrooms



2017
**Développeur
logiciel C/C++/Java**
LDNR



2010 - 2012
**BTS Management
des Unités
Commerciales**
Institut des Métiers



LANGUES

- Anglais
- Allemand



PROFIL

Ingénieur en intelligence artificielle et développeur expérimenté, je suis passionné par les possibilités infinies que ces technologies offrent pour améliorer notre vie quotidienne. Autonome, adaptable et doté d'un bon relationnel, je pense que mon parcours et mes compétences me permettront de contribuer de manière significative à tout projet. Je serais ravi de rejoindre votre équipe.



EXPÉRIENCES PROFESSIONNELLES



Technicien d'études développeur **2018 - 2021**
CS Group

Développement et TMA sur la suite d'outils Publishing d'Airbus (outils de médiatisation de la documentation avion, des révisions temporaires, outils de conversion des données XML/SGML en PDF...)



Développeur Web fullstack **2017 - 2018**
Domicil'Gym

Interface de gestion dédiée aux coachs, permettant la gestion de l'agenda, des contrats, de la facturation et le suivi des clients.



Développeur Web Stagiaire **2017 - 2017**
MyTeleworker

Développement en PHP d'un module de social login pour CMS Made Simple.
Développement d'un chatbot avec Node.js et l'api Recast.ai (api de NLP) pour mettre en relation des freelances et des employeurs.



COMPÉTENCES

Collecte, analyse et nettoyage des données
Développement d'algorithmes
Modélisation statistiques
Machine Learning / Deep Learning
Pipelines
Cloud (Microsoft Azure, Amazon AWS)
Visualisation et Communication



LANGAGES ET LIBRAIRIES

Langages

Python, NodeJS, C, C++, Java
(J2SE & J2EE), javascript, PHP

Base de données

SQL, SQLite, MySQL

Web

Angular 9, React, ajax,
bootstrap, CSS3, HTML5, JSP

Librairies

NumPy, Pandas, Matplotlib,
Seaborn, plotly Scikit-learn,
Tensorflow, Keras, PyTorch,
XGBoost, Flask, Django, FastAPI

Traitements de données

XSL-FO, XSD, XSLT

Langages de commandes

Bash, KSH, Batch

Expériences professionnelles

Employeur CS Toulouse

De 2018 à 2021

Technicien d'études développeur *De 2018 à 2021 Airbus*

Projet : Airbus Publishing Développement et TMA sur la suite d'outils Airbus Publishing (outil de médiatisation de la documentation avion, outil de médiatisation des révisions temporaires, outils de conversion des données XML/SGML en PDF...)

Environnement technique

- Linux - Unix - Windows
- Eclipse - SQLite
- PHP - JavaScript - Java standard (J2SE) - bash - XML - XSL - XSL-FO - ksh

Activités

- Développement de nouvelles versions (évolutives ou correctives)
- Qualification
- Packaging/Installation
- TMA (maintien en conditions opérationnelles)

Employeur Domicil'Gym Toulouse

De 2017 à 2018

Développeur Web *De 2017 à 2018*

Projet : DG Dashboard Interface de gestion dédiée aux coachs, permettant la gestion de l'agenda, des contrats, de la facturation et le suivi des clients.

Environnement technique

- Windows - Linux
- Apache HTTP Server - Bugzilla - FileZilla - Git - NetBeans - phpMyAdmin - Trello
- PHP - JavaScript - JQuery - CSS3 - HTML5
- SSH - HTTPS

Activités

- Développement des outils de facturation
- Développement d'une fonctionnalité d'agenda synchronisé à Google Calendar
- Développement de la fonctionnalité de saisie des contrats
- Migration des sites des coaches de Drupal vers Wordpress

Employeur MyTeleWorker Toulouse

De 05/2017 à 07/2017

Développeur stagiaire *De 05/2017 à 06/2017 Durée : 2 Mois*

Projet : Module Social Login CMS Made Simple Développement en PHP d'un module de social login pour CMS Made Simple

Environnement technique

- Linux - Windows
- Bugzilla - FileZilla - phpMyAdmin - MySQL
- PHP - CSS3 - JavaScript
- SSH

Activités

- Développement du module
- Tests
- Déploiement

Développeur Stagiaire

De 06/2017 à 07/2017 Durée : 2 Mois

Projet : Chatbot MyTeleWorker Développement d'un chatbot avec Node.js et l'api Recast.ai (api de NLP) pour mettre en relation des freelance et des employeurs

Environnement technique

- Windows - Linux
- Cygwin - nodejs
- JavaScript

Activités

- Développement

Formation initiale

Année	Intitulé	Etablissement	Lieu
De 10/2021 à 01/2023	Titre RNCP Data Scientist	OpenClassrooms	En ligne
De 11/2016 à 07/2017	Formation Développeur Logiciel (Niveau 3)	LDNR	Labège
De 01/2019 à 01/2019	BTS Management des Unités Commerciales	Institut des métiers	Clermont-Ferrand

Licences et certifications

Data Science et Machine Learning

- [Initiez-vous au Machine Learning](#)
- [Objectif IA: initiez-vous à l'intelligence artificielle](#)
- [Machine Learning](#)
- [Initiez-vous au Deep-Learning](#)
- [Entraînez un modèle prédictif linéaire](#)
- [Évaluez les performances d'un modèle de machine learning](#)
- [Utilisez des modèles supervisés non linéaires](#)
- [Modélisez vos données avec les méthodes ensemblistes](#)
- [Explorez vos données avec des algorithmes non supervisés](#)
- [Analysez vos données textuelles](#)
- [Classez et segmentez les données visuelles - 12/11/2022](#)

Programmation et développement web

- [Apprenez les bases du langage Python](#)
- [Initiez-vous à Python pour l'analyse de données](#)
- [Perfectionnez-vous en Python](#)
- [Découvrez les bibliothèques Python pour la Data Science](#)
- [Python3](#)
- [NodeJS](#)
- [développez des applications web avec Angular](#)
- [Angular & NestJS](#)
- [React & Redux](#)
- [Adoptez les API REST pour vos projets web](#)
- [Optimisez votre déploiement en créant des conteneurs avec Docker](#)
- [Gérez du code avec Git et Github](#)
- [PHP](#)
- [Concevez un site avec Flask - 12/11/2022](#)
- [Javascript](#)
- [HTML Fundamentals](#)
- [CSS3](#)
- [jQuery](#)
- [Java](#)
- [Ruby](#)
- [C#](#)
- [C++](#)
- [SQL](#)

Autres compétences et connaissances

- [EF Standard English Test : C2 Proficient](#)
- [Hacking Ethique : Le Cours Complet - 22/12/2022](#)
- [S'initier aux probabilités et aux variables aléatoires](#)
- [Statistiques pour la Data Science](#)
- [Découvrez la méthodologie DevOps](#)
- [Apprenez à utiliser la ligne de commande dans un terminal](#)
- [Maîtrisez les risques juridiques liés au numérique - 05/11/2022](#)
- [Initiez-vous à la gestion de projet agile - 12/11/2022](#)
- [Analysez les risques de votre projet - 09/11/2022](#)
- [Mettez en place un système de veille informationnelle - 10/11/2022](#)
- [Réalisez la maquette d'une application mobile avec Adobe XD - 10/11/2022](#)
- [Réalisez des rapports statistiques clairs et impactants](#)
- [Améliorez l'impact de vos présentations](#)
- [Communiquez et formalisez vos idées par le storytelling](#)
- [Apprenez à apprendre](#)
- [prenez la parole en public](#)
- [Améliorez l'impact de vos présentations](#)

Langues

- [Anglais : niveau C2](#)

Projets réalisés pour le titre RNCP de data scientist niveau 7 (Bac+5)

Projet 1 : Réinventons l'expérience client

Objectifs:

- Détection de la langue parlée par l'utilisateur

Environnement technique Microsoft Azure, Jupyter Notebook, VSCode

Compétences

- Utilisation de modèles pré-entraînés sur le cloud pour réaliser des prédictions à partir de données textuelles
- Ecrire une requête pour récupérer les prédictions du modèle
- Sécurisation des identifiants et des clés de connexion pour éviter les accès non autorisés
- Rédaction de scripts pour utiliser les modèles pré-entraînés
- Présentation professionnelle des résultats de l'utilisation du modèle pré-entraîné sur le cloud, y compris le code et les résultats de la prédiction.

Repository

https://github.com/AnodeGrindYo/OC_IA_P1

Projet 2 : Participez à un concours sur la Smart City

Objectifs:

- Analyser les données des arbres de la ville de Paris et proposer une solution pour aider à l'optimisation des tournées pour l'entretien de ces arbres.

Environnement technique Jupyter Notebook, Numpy, Pandas, matplotlib, seaborn, plotly, adobe XD

Compétences

- Analyse de données exploratoires en utilisant des bibliothèques Python
- Installation et configuration d'environnements de développement Python et Jupyter
- Création et gestion d'environnements virtuels pour assurer l'isolement du projet et la gestion des dépendances
- utilisation de bibliothèques Python spécialisées pour l'analyse de données
- Description et présentation d'un jeu de données en utilisant un dataframe
- Calcul des indicateurs statistiques basiques (moyenne et écart-type) pour les différentes colonnes
- Comparaison des ordres de grandeur des grandeurs statistiques des différentes colonnes
- Identification des valeurs aberrantes dans une distribution statistique
- Tracé de graphiques pour représenter des distributions statistiques
- Utilisation des fonctionnalités d'édition de cellule Markdown de Jupyter pour commenter et mettre en forme l'analyse de données
- Création d'une présentation générale du jeu de données, d'une démarche méthodologique d'analyse de données, et d'une synthèse de l'analyse de données.

Repository

https://github.com/AnodeGrindYo/OC_IA_P02

Projet 3: Préparez des données pour un organisme de santé publique

Objectifs:

- Réaliser une analyse exploratoire de données sur le jeu de données Open Food Facts pour rendre les données de santé plus accessibles et compréhensibles pour les agents de l'agence Santé publique France.

Environnement technique Jupyter Notebook, Numpy, Pandas, matplotlib, seaborn, plotly, Voilà, HTML/CSS/Javascript

Compétences

- Nettoyage des données structurées, traitement des valeurs manquantes, détection et traitement des outliers
- Élimination des variables non pertinentes et des lignes dupliquées
- Utilisation de fonctions pour nettoyer les données
- Analyse statistique multivariée, analyse univariée et bi-variée
- Utilisation de graphiques pour communiquer les résultats
- Création de graphiques interactifs
- Utilisation de méthodes d'analyse descriptive (ex: ACP) et explicative (ex: ANOVA)
- Utilisation de cellules Markdown pour rendre le document lisible

Repository

https://github.com/AnodeGrindYo/OC_IA_P3

Projet 4 : Modèle de scoring pour l'octroi de crédits

Objectifs:

- Développer un algorithme de classification à l'aide de modèles supervisés linéaires et non linéaires.
- Fournir une mesure de l'importance des variables utilisées pour calculer la probabilité de remboursement d'un client.

Environnement technique Jupyter Notebook, Pandas, Numpy, Matplotlib, Seaborn, Scipy, Sklearn (RandomForestClassifier, AdaBoostClassifier, - GradientBoostingClassifier, LogisticRegression, - PolynomialFeatures, train_test_split, GridSearchCV, confusion_matrix), Imbalanced Learn

Compétences

- Feature engineering
- Entraînement d'un modèle supervisé
- Évaluation des performances du modèle
- Optimisation des hyperparamètres

Repository

https://github.com/AnodeGrindYo/OC_IA_P4

Projet 5 : Segmentation de clients d'un site e-commerce

Objectifs:

- Représenter graphiquement des données à grandes dimensions.
- Mettre en œuvre des techniques d'extraction de features pour des données non structurées.
- Mettre en œuvre des techniques de réduction de dimension.
- Prétraiter des données non structurées pour obtenir un jeu de données exploitable.
- Utiliser des modèles non supervisés.
- Développer un algorithme de segmentation pour aider l'équipe e-commerce à mieux communiquer avec les clients.
- Proposer un contrat de maintenance basé sur une analyse de la stabilité des segments au cours du temps.

Environnement technique Jupyter Notebook, Pandas, Numpy, Matplotlib, Seaborn, Scipy, Sklearn (PCA, GridSearchCV, KMeans), t-SNE

Compétences

- Représenter graphiquement des données à grandes dimensions.
- Mettre en œuvre des techniques d'extraction de features pour des données non structurées.
- Mettre en œuvre des techniques de réduction de dimension.
- Prétraiter des données non structurées pour obtenir un jeu de données exploitable.
- Utiliser des modèles non supervisés.
- Développer un algorithme de segmentation pour aider l'équipe e-commerce à mieux communiquer avec les clients.
- Proposer un contrat de maintenance basé sur une analyse de la stabilité des segments au cours du temps.

Repository

https://github.com/AnodeGrindYo/OC_IA_P5

Projet 6 : Améliorez le produit IA de votre startup

Objectifs:

- Classer et segmenter les données visuelles et analyser les données textuelles.
- Sélectionner un modèle d'apprentissage Deep Learning adapté à une problématique métier.
- Transformer les variables pertinentes d'un modèle de Deep Learning.
- Mettre en place un modèle de Deep Learning et évaluer ses performances.
- Adapter les paramètres d'un modèle de Deep Learning afin de l'améliorer.
- Analyser les commentaires négatifs pour détecter les différents sujets d'insatisfaction.
- Analyser les photos pour déterminer leurs catégories (nourriture, décor dans le restaurant ou à l'extérieur du restaurant).
- Faire une étude de faisabilité, c'est-à-dire savoir rapidement si la séparation automatique selon la catégorie réelle (classification non supervisée) est possible.
- Collecter de nouvelles données via l'API Yelp. Valider la faisabilité de la solution en collectant les informations relatives à environ 200 restaurants pour une ville en utilisant l'API.

Environnement technique Jupyter Notebook, Pandas, Numpy, Matplotlib, Seaborn, Scipy, Sklearn (GridSearchCV, LogisticRegression, t-SNE, LatentDirichletAllocation), TensorFlow, NLP, TF-IDF, BOW, REST API, CV2

Compétences

- Collecte de données venant d'une API qui correspondent à un besoin défini.
- Effectuer un prétraitement de données non structurées pour obtenir un jeu de données utilisable.
- Utiliser des techniques de réduction de la dimension.
- Visualiser des données de grandes dimensions.

Repository

https://github.com/AnodeGrindYo/OC_IA_P06

Projet 7 : Détection des Bad Buzz grâce au Deep Learning

Objectifs:

- Développer un prototype d'un produit d'Intelligence Artificielle qui permet de prédire le sentiment associé à un tweet.
- Déterminer si le tweet est positif ou négatif et s'il est susceptible de provoquer un buzz potentiellement nuisible à la réputation de la société Air Paradis.
- Comparer les performances de trois modèles de classification: un modèle classique binaire, un modèle basé sur des réseaux de neurones profonds avec une couche d'embedding, et un modèle BERT.
- Déployer la solution la plus performante via une API. Le service recevra un tweet en entrée et retournera le sentiment associé prédit par le modèle.

Environnement technique Jupyter Notebook, Pandas, Numpy, Seaborn, Matplotlib, Scikit-learn, Tensorflow, Keras, FAST API, Amazon AWS EC2, uvicorn, NLP, BERT, Word2Vec, GloVe, Deep Learning, Neural Networks, nginx

Compétences

- Choix de la méthode de plongement de mots pertinente pour un modèle de Deep Learning.
- Conception et déploiement continu d'un moteur d'inférence dans le Cloud.
- Évaluation de la performance d'un modèle de Deep Learning sur des données textuelles.

- Réalisation d'une présentation orale d'une démarche de modélisation.
- Rédaction d'une note méthodologique pour communiquer sa démarche de modélisation.
- Définition et mise en œuvre d'un pipeline d'entraînement des modèles.
- Entraînement d'un modèle Deep Learning sur des données textuelles.
- Mise en œuvre d'un logiciel de version de code.
- Sélection des méthodes de prétraitement du texte pour un modèle de Deep Learning.

Repository

https://github.com/AnodeGrindYo/OC_IA_P07

Projet 8 : Participer à la conception d'une voiture autonome

Objectifs:

- Entraîner un modèle de Deep Learning sur des images (segmentation sémantique d'images) et évaluer ses performances.
- Utiliser des techniques d'augmentation des données.
- Manipuler un jeu de données volumineux (générateur de données).

Environnement technique Python, Pandas, Numpy, Matplotlib, Seaborn, Scipy, imgaug, Tensorflow, Unet, VGG16, CUDA, Amazon AWS EC2, FastAPI, uvicorn.

Compétences

- Entraîner un modèle de Deep Learning sur des images en identifiant la cible, en séparant le jeu de données en jeu d'entraînement et en jeu de test et en évitant les fuites d'informations.
- Évaluer la performance d'un modèle de Deep Learning sur des images en utilisant des métriques adaptées, en évaluant la performance d'un modèle de référence et en optimisant les hyperparamètres.
- Utiliser des techniques d'augmentation des données en utilisant au moins trois techniques pour améliorer la performance du modèle.
- Manipuler un jeu de données volumineux en développant un générateur de données permettant le traitement des images sur plusieurs cœurs de calcul.

Repository

https://github.com/AnodeGrindYo/OC_IA_P08

Projet 9 : Réalisez une application de recommandation de contenu

Objectifs:

- Développer une première version du système de recommandation sous forme d'Azure Functions.
- Réaliser une application simple de gestion du système de recommandation.
- Stocker les scripts développés dans un dossier GitHub.
- Identifier l'architecture cible pour prendre en compte l'ajout de nouveaux utilisateurs ou de nouveaux articles.

Environnement technique Python, Pandas, Numpy, Matplotlib, Surprise, Sklearn, Microsoft Azure, GitHub.

Repository

https://github.com/AnodeGrindYo/OC_IA_P09

Projet 10 : Développez un chatbot pour réserver des vacances

Objectifs:

- Développer un chatbot pour aider les utilisateurs à choisir une offre de voyage.
- Construire un MVP capable d'identifier la ville de départ, la ville de destination, la date de départ, la date de retour et le budget maximal pour le prix total du billet.
- Poser des questions pertinentes en anglais à l'utilisateur pour comprendre pleinement sa demande en cas d'élément manquant.
- Reformuler la demande et demander à l'utilisateur de valider sa compréhension.

Environnement technique Python, Microsoft Bot Framework SDK v4 pour Python, Microsoft Azure LUIS, Microsoft Azure WebApp, Bot Framework, Azure Bot Service, Github Actions

Compétences

- Intégrer la sortie du modèle dans un produit informatique fini.
- Vérifier la qualité des dialogues avec l'utilisateur.
- Créer une interface de conversation avec le chatbot en anglais.
- Développer et tester le code d'une chaîne de traitement IA.
- Organiser le code en fonctions ou classes.
- Utiliser un outil de gestion de version Git et stocker le code sur un repository Github.
- Définir les critères d'évaluation de la performance du modèle.
- Calculer automatiquement les critères d'évaluation de la performance à partir des dialogues enregistrés et labellisés.
- Mettre en place un seuil d'alerte pour détecter les erreurs du chatbot.
- Utiliser l'outil de suivi Azure Application Insight pour suivre les performances du chatbot.
- Définir une méthodologie de mise à jour du modèle en production.

Repository

https://github.com/AnodeGrindYo/OC_IA_P10

Projet 11 : Réaliser le cadrage d'un projet IA

Compétences

- Identification des ressources humaines, techniques et financières nécessaires pour un projet IA
- Planification et mitigation des risques projets significatifs
- Connaissance des enjeux éthiques et légaux relatifs à la collecte des données personnelles
- Identifier et estimer la complexité des user stories
- Synthèse et présentation claire des informations techniques, financières et éthiques dans un support de présentation
- Maîtrise de la méthodologie de gestion de projet agile
- Capacité à travailler en équipe multidisciplinaire et à communiquer efficacement avec des parties prenantes techniques et non techniques

Repository

https://github.com/AnodeGrindYo/OC_IA_P11

Projets personnels

1. [AI-Theremin : Simulateur de theremin en Python avec tracking des mains via webcam](#)

- Utilisation de Python et d'une IA pour le suivi des mains pour la simulation d'un instrument de musique
- Projet démontrant des compétences en développement, traitement d'image et interaction en temps réel

Environnement technique Windows, VSCode, Anaconda, Python, Github

2. [Diffusion Models: Tout ce qu'il faut savoir](#)

- Série de notebooks expliquant les modèles de diffusion avec des exemples en Python
- Explication des mathématiques derrière les modèles, programmation Python et communication des concepts complexes
- Création et entraînement de réseaux de neurones avec PyTorch, basés sur des articles de recherche, pour illustrer par des exemples pratiques et l'évaluation des modèles **Environnement technique** Google Colaboratory

3. [Minecraft-Spigot-GetLucky-plugin](#)

- Plugin Java pour les serveurs Minecraft basés sur Spigot
- Ajoute une baguette magique dans le jeu que l'utilisateur utilisera à ses risques et périls
- Développement Java et création de plugins pour des serveurs de jeu **Environnement technique** windows, Eclipse, Java 8, Spigot-API 1.19.4