

Dr. Vincent Roger

Scientifique des données et de l'apprentissage automatique

@ Vincent.Roger.Pro@proton.me

✉ 1 Rue Saint-Dominique; 31000 Toulouse

🌐 vincent-roger.fr/accueil

☎ +336 33 25 61 97

🔄 vroger11

in vroger11

📄 Vincent-Roger



Compétences

Techniques dans les sciences des données

- Apprentissages supervisés, auto-supervisés et semi-supervisés avec recherche d'hyperparamètres.
- Gestion d'expériences avec utilisation des ressources.
- Réseaux de neurones profonds sur du big data et sur peu de données (few-shot).
- Représentations par modèles génératifs comme GAN, VAE ou DPGMM.
- Traitement de signaux (images et audios).
- Analyse de données volumineuses et visualisations.
- Création de prototypes et microservices.



Outils des sciences des données

- Python, Numpy, Scipy, Pandas
- PyTorch, Scikit-learn, Lightning, Optuna, MLflow
- Docker, Falcon
- Slurm
- Streamlit, Plotly, Matplotlib



Gestion de projets

- Organisé.
- Travail en collaboration.
- Respect des délais.
- Gestion des risques.
- Leadership.

Linguistiques

- Français
- Anglais



Expériences professionnelles

Data and Machine Learning Scientist, en cours, Kiviak Instruments

📍 Toulouse 2023–présent

Créer des modèles (intégrés en tant que microservice) pour étiqueter automatiquement des samples. Développement de prototypes pour de nouvelles fonctionnalités. Travail sur des fonctionnalités de traitement du signal. Et davantage à l'horizon.

Thèse, trois ans et trois mois, IRIT

📍 Toulouse 2018–2022

Les personnes atteintes de cancers ORL présentent des difficultés de prononciation après des chirurgies ou des radiothérapies. Il est important pour le praticien de pouvoir disposer d'une mesure reflétant la sévérité de la parole. Je propose deux approches pour créer une mesure automatique, bien que ne disposant que de peu de données (environ 1h d'enregistrements audio pour 128 locuteurs). La première est fondée sur des méthodes de "few shot", tandis que la seconde est fondée sur la mesure entropique de caractéristiques de la parole (appries avec un modèle auto-supervisé sur un corpus annexe). Nos résultats sur cette dernière ont permis d'envisager une application médicale. Ainsi, j'ai obtenu une subvention pour encadrer un ingénieur afin de réaliser une application livrée au CHU de Toulouse.

Ingénieur d'études, deux ans, LIS

📍 Toulon 2016–2018

Suite à mon précédent contrat, j'ai réalisé une représentation par modèle auto-supervisé profond des milieux acoustiques sous-marin pour aider à catégoriser les différents comportements des cétacés étant à portée des bouées. J'ai par la suite réalisé un modèle profond de classification de 1500 espèces d'oiseaux. Pour ces deux problèmes, je disposais de gros volumes de données.

Ingénieur d'études, dix mois, LIS - TTV Innovation

📍 Toulon 2015-2016

Suite à la mise en place de bouées en mers équipées de microphone, nous disposons de grandes quantités de données. Mon travail consistait à modéliser l'environnement bioacoustique à l'aide de modèles génératifs. Grâce à cela, j'ai pu réaliser un rapport des activités des narvals en fonction de l'activité lunaire.

- Ingénieur junior d'études**, cinq mois, **IRIT** 📍 **Toulouse** 2015
 Planification temporelle par compilation de satisfaction de contraintes temporelles. Analyse théorique et expérimentale de langages de représentations de problèmes temporels. Publication d'un article.
- Ingénieur junior logiciel**, quatre mois, **LAAS** 📍 **Toulouse** 2014
 Modernisation du système de gestion de mouvements de robot humanoïde (représentation des articulations, corps et positions) de l'équipe GEPETTO en accélérant son initialisation et en améliorant l'interface de programmation.
- Ingénieur junior logiciel**, deux mois, **IRIT** 📍 **Toulouse** 2013
 Réalisation d'un logiciel de sous-titrage automatique de flux audio-vidéo (manipulation de flux avec transcription en temps réel). Le logiciel est devenu un outil de démonstration de l'équipe de recherche SAMOVA.
- Technicien**, deux mois et demi, **CEICOM** 📍 **Toulouse** 2011
 Réalisation du portage d'une application de communication intermachine de Windows vers Linux. Résultat : réécriture en C++ de routines Windows pour Linux; portage réussi avec un impact stratégique.

Formations

- Doctorat**, Informatique, **Université Paul Sabatier** 📍 **Toulouse** 2022
 J'ai amélioré ma communication à la radio et durant mes présentations.
- Master**, Intelligence Artificielle, **Université Paul Sabatier** 📍 **Toulouse** 2015
 Modèles statistiques, traitement de signal, reconnaissance de formes, contrôle de robots et gestion de projets.
- Licence**, Informatique Fondamentale, **Université Paul Sabatier** 📍 **Toulouse** 2013
 Outils de développement, programmation bas niveau, statistique, probabilités et calculatoire.
- DUT**, Informatique, **IUT Paul Sabatier** 📍 **Toulouse** 2011
 Compétences techniques pour l'informatique et façons de designer des applications pour l'industrie.

Centres d'intérêts

Associative: Je suis membre actif des associations Toulouse dataviz et Toulouse datascience. Plus de détails sur ces associations ici : toulouse-dataviz.fr et tlse-data-science.fr

Blogging: Je blogue sur mes projets, participations à des concours de visualisations et astuces sur l'environnement Linux. C'est par ici : vincent-roger.fr/blogue

Sports: Musculation (deux fois par semaine), natation (deux fois par semaine) et course à pied (trois fois par semaine). J'ai terminé le marathon de Toulouse.

Lectures: Je lis sur la productivité, la santé et le sport. Je lis également des Shōnens japonais.

