

apprentissage automatique : un bel avenir

Entretien avec Daniel Situnayake

C. J. Abate

Daniel Situnayake est ingénieur et ingénieux. Selon ses mots, l'apprentissage automatique (*machine learning* en anglais, ML) est une "technique unique en son genre". Il nous parle ici du potentiel du ML et présente TinyML, ainsi que quelques applications idéales. Nous évoquons aussi son expérience de développeur chez Google.

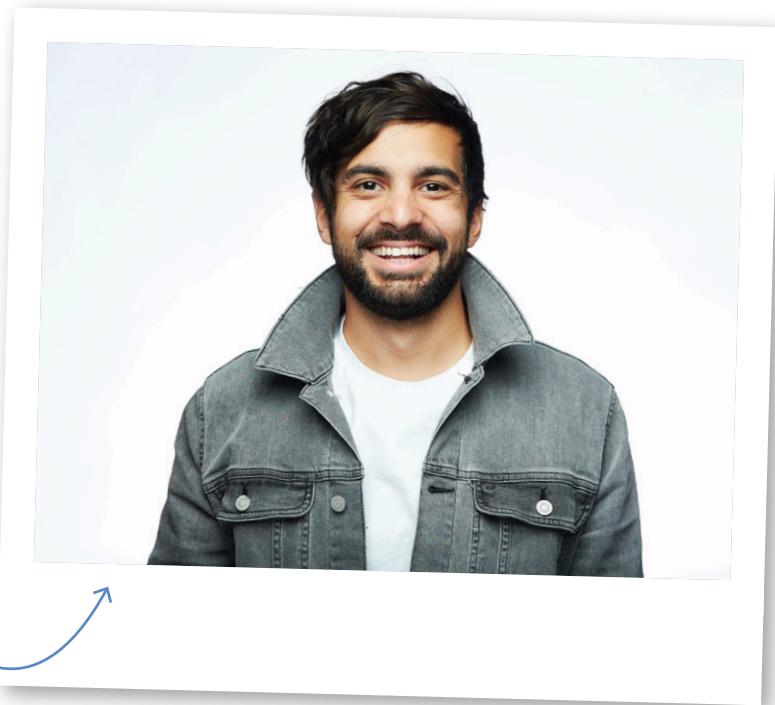


Figure 1. Daniel Situnayake (ingénieur, fondateur de TinyML, Edge Impulse).

Quitter l'université

Abate : Commençons par tes antécédents. Quand t'es-tu intéressé pour la première fois au ML intégré ? Était-ce à l'université de Birmingham City (BCU), où tu as obtenu ta licence en réseaux informatiques et sécurité, ou après ?

Situnayake : Je n'ai pas vraiment cherché à travailler sur l'apprentissage automatique. Ce fut un itinéraire long et sinuieux, avec des expériences apparemment disparates qui m'ont poussé ! Après mon diplôme de la BCU, on m'a proposé de rester comme enseignant et consultant. Je me suis concentré sur l'identification automatique et la capture de données, c'est-à-dire toutes les façons dont un ordinateur peut reconnaître et traiter des objets physiques. Ça va de la biométrie (reconnaissance faciale) à la RFID, en passant par les codes à barres et les cartes à puce. Je découvrais les algorithmes de vision par ordinateur, qui à l'époque n'étaient pas basés sur l'apprentissage automatique, et les systèmes embarqués. Quelques emplois plus tard, suite à une acquisition, je travaillais dans l'équipe naissante de sciences des données d'une société de banque généraliste. C'est là que je suis entré en contact avec les outils et les procédés de la science des données, et que j'ai appris à travailler avec de grands ensembles de données, sujet nouveau et fascinant pour moi. Avec la création de Tiny Farms, j'ai eu l'occasion de réunir ces

domaines, en utilisant des outils de science des données pour interpréter et réagir aux données issues de capteurs intégrés dans un contexte de technologie agricole. Mais c'est chez Google, en rencontrant Pete Warden (**fig. 2**) et Andy Selle de l'équipe TensorFlow Lite, que j'ai vraiment saisi le potentiel d'un apprentissage approfondi sur de tout petits appareils. C'est alors devenu mon objectif principal.

Abate : Quand as-tu déménagé en Californie ? Était-ce par opportunitisme professionnel ?

Situnayake : En 2009, un an après l'implosion de l'économie mondiale. Au fil des dernières années, j'avais passé une batterie de tests de sélection comme élève officier de l'armée britannique. Je voulais piloter des hélicoptères Apache ! Avant la formation d'officier prévue pour mai 2009, je suis parti en voyage d'adieu à travers les États-Unis.

En Californie, j'ai rencontré une fille. Quelques mois plus tard, et quelques semaines après ma prise de service, j'ai compris que j'étais moins fait pour la vie de gentleman pilote que pour une vie décontractée en Californie. J'ai réservé un vol, j'ai emménagé avec la fille que je venais de rencontrer, et nous nous sommes mariés six mois plus tard !

Elle a repris ses valises, je suis resté. Après avoir quitté l'université, j'ai rejoint une jeune pousse spécialisée dans l'IA à Los Angeles, puis je suis arrivé dans la région de la Baie juste à point pour le boom technologique des années 2010. Un trip de dingue.

Élevage d'insectes

Abate : Tu es cofondateur de Tiny Farms, que tu décris sur LinkedIn comme "la première entreprise américaine de technologie d'élevage d'insectes". Comment a-t-elle vu le jour ? Où en est-elle ?

Situnayake : Comme toutes mes meilleures histoires, celle-ci commence encore par un voyage à travers les États-Unis. Deux de mes amis proches s'apprêtaient à lancer un projet de sécurité alimentaire. Au cours d'un voyage, ils ont séjourné dans une cabane au bord d'un lac, entourés de grosses sauterelles bruyantes dont ils se sont demandé, par jeu, si elles étaient comestibles. Quelques recherches leur ont appris l'importance cruciale des insectes comestibles pour la sécurité alimentaire mondiale - mais que les techniques d'élevage en captivité étaient mal connues.

Ils m'ont proposé de me joindre à eux pour créer une entreprise. J'ai vu le potentiel de l'application à l'automatisation de l'élevage d'insectes des techniques que j'avais apprises autour de l'identification automatique et de la capture de données. Mes connaissances en science des données me guideraient... J'ai fait le saut ! D'abord à temps partiel, puis nous avons réuni des fonds pour construire une ferme d'élevage d'insectes à l'échelle industrielle à San Leandro, en Californie. C'était intense, excitant et plus que stressant d'essayer de



Figure 2. Daniel Situnayake et Pete Warden signent leur livre TinyML.

Google a une étonnante culture de la liberté : tu travailles sur pratiquement tout ce que tu veux, à condition que cela ait un impact sur ton produit. Si tu as une idée géniale, tu es encouragé à la poursuivre et à faire appel à d'autres pour t'aider. Ainsi naissent de nombreux projets passionnants, mais cela peut devenir écrasant : devant un choix aussi vaste, sur quoi faut-il se concentrer ?

Des modèles TinyML pour détecter les éléphants sauvages, afin de prévenir les villageois que des pachydermes se dirigent vers eux.

maintenir en vie des millions de criquets tout en traitant avec les clients, les investisseurs et la recherche. Cinq ans, j'ai jeté l'éponge. Mes cofondateurs ont continué pendant quelques années, pour renoncer finalement devant la difficulté d'obtenir un produit viable. Tiny Farms continue de fonctionner comme dépôt libre des études et des données accumulées pendant sept ans d'activité [1].

Travailler pour Google

Abate : Tu as passé presque trois ans avec Google. Avant d'entrer dans les détails de tes fonctions, peux-tu nous dire ce que ça t'a fait de rejoindre une entreprise aussi impressionnante ?

Situnayake : Après cinq ans dans une entreprise agricole - où j'ai passé plus de temps dans la construction que dans le codage - je voulais me remettre dans le bain. Or, il n'y avait pas d'entreprise que j'admirais plus, techniquement en tout cas, que Google. Dès les premiers jours, j'ai été époustouflé. Je n'avais jamais travaillé dans une grande entreprise de technologie auparavant. Avec ton badge Google, tu accèdes aux installations de n'importe quel bureau Google, partout dans le monde. J'étais devenu citoyen d'un pays inconnu, avec ses propres services publics, ses infrastructures et son gouvernement.

Là ça se corse, parce qu'il se passe tant de choses, il y a des réunions potentiellement interminables, des distractions et tant d'idées stimulantes. Une bonne partie de ta première année chez Google tu la passes en fait à développer ton propre filtre intérieur, pour ne pas te laisser submerger. Les génies qui travaillent chez Google, c'est un peu un mythe. Il y a grosso modo la même gamme de compétences chez Google qu'ailleurs dans l'industrie technologique. Je ne suis pas porté sur le culte des héros. C'est agréable de rencontrer les meilleurs dans ton domaine et de constater que ce sont des gens comme toi.

Abate : Tu as travaillé chez Google en tant qu'avocat des développeurs pour TensorFlow Lite. Qu'est-ce que TensorFlow Lite ? Et comment avez-vous formé les programmeurs, par des rencontres, des cours en ligne ?

Situnayake : TensorFlow est le nom de l'écosystème d'outils libres de Google pour la formation, l'évaluation et le déploiement de modèles d'apprentissage profond. Il comprend tout, du code Python de haut niveau utilisé pour définir des architectures de modèles jusqu'au code de bas niveau utilisé pour exécuter ces modèles sur différents types de processeurs.

TensorFlow Lite est le sous-ensemble des outils *TensorFlow* qui traitent du déploiement de modèles vers des "périphériques", c'est-à-dire tout ce qui est plus petit qu'un ordinateur personnel : téléphones, microcontrôleurs, etc. Les outils *TensorFlow Lite* peuvent prendre un modèle d'apprentissage approfondi et l'optimiser pour ces types d'appareils. Il comprend également le code hautement optimisé qui s'exécute sur les appareils eux-mêmes.

Ce que je préfère dans *TensorFlow Lite*, c'est la variante dépourvue et super rapide conçue pour faire tourner des modèles d'apprentissage profond sur de minuscules µC bon marché et de faible puissance. Avant le lancement de *TensorFlow Lite* pour µC, les programmeurs devaient écrire leur propre code de bas niveau pour exécuter des modèles d'apprentissage sur des appareils embarqués. Le barrage d'entrée pour l'apprentissage machine sur les appareils était donc difficile à franchir. L'entrée de *TensorFlow* dans ce domaine a fait tomber beaucoup de murs, si bien que soudain, avec un peu d'expérience de l'embarqué, n'importe qui pouvait exécuter des modèles. C'est excitant de voir le domaine du *TinyML* prendre vie.

Mon rôle était d'aider l'équipe de *TensorFlow Lite* à comprendre l'écosystème des développeurs et à s'y connecter. Tous les programmeurs sont différents, il importe donc de créer de multiples façons pour eux de se lancer. Nous avons fait tout ce qui est imaginable, depuis les rencontres en personne jusqu'aux conférences financées par Google, en passant par la création d'exemples de code faciles à comprendre et la rédaction du livre *TinyML* !

Abate : Pourquoi as-tu quitté Google ? Fut-ce une décision difficile ?

Situnayake : En travaillant avec les concepteurs de *TensorFlow*, j'ai pu constater qu'il fallait un énorme effort d'éducation pour que même les meilleurs outils et bibliothèques d'apprentissage sur machine soient utilisés. Malgré l'excellence de certains cours et de leur contenu, le temps qu'il faut y consacrer met l'apprentissage automatique hors de portée d'un grand nombre d'ingénieurs trop occupés. Les modèles de formation sont autant un art qu'une science, et pour acquérir les réflexes et les bonnes pratiques, il faut des années. C'est encore plus difficile avec le *TinyML*, un domaine tellement nouveau et aux contraintes mal connues. Quand j'ai entendu parler d'*Edge Impulse*, la société pour laquelle je quitterai Google, j'ai été époustouflé. Avant même le lancement officiel, les fondateurs avaient déjà un produit qui permettait à n'importe quel programmeur de former des modèles d'apprentissage profond sur du matériel embarqué. J'ai pu constater, grâce à mon travail avec la communauté des développeurs, qu'ils avaient résolu certains des problèmes les plus difficiles en éliminant les pierres d'achoppement qui rendent les outils d'apprentissage automatique si difficiles à utiliser. Je pense que l'apprentissage automatique est une technique unique en son genre. Elle va transformer le monde. *Edge Impulse* a pour mission de mettre cette technologie entre les mains de tout ingénieur, afin qu'il puisse utiliser ses diverses perspectives et sa connaissance approfondie du domaine pour résoudre des problèmes complexes où que ce soit dans le monde.

Conscient de ce potentiel, j'ai décidé de me lancer. J'étais triste de quitter mes amis de Google, mais le monde (de *TinyML*) est petit et je reste en collaboration étroite avec beaucoup d'entre eux !

Tiny Machine Learning (TinyML)

Abate : Que doivent savoir de *TinyML* les ingénieurs et les électroniciens sérieux ? Pourrais-tu donner des exemples d'application d'apprentissage automatique à très faible consommation ?

Situnayake : Notre monde est truffé de capteurs, dans toutes sortes d'endroits, des appareils ménagers aux usines intelligentes en passant par nos véhicules. Les données de milliards de capteurs représentent une occasion précieuse pour les concepteurs de créer des appareils utiles, qui comprennent le monde autour d'eux et prennent des décisions intelligentes.

Auparavant, la seule façon d'utiliser ces données était de les envoyer par l'internet à un ordinateur puissant capable de calculer, d'interpréter puis de déclencher des actions en fonction des résultats. Cela fonctionne pour certaines applications. Pas de problème p. ex, si les données d'une station météo mettent quelques secondes à arriver sur un serveur, mais il y a de nombreuses applications où cela ne fonctionne pas.

L'un de mes exemples favoris est celui des caméras animalières. Les chercheurs les utilisent pour observer les animaux, souvent dans des endroits éloignés dépourvus d'accès à l'internet. Un animal qui passe déclenche le capteur de mouvement, et une photo est prise. Les chercheurs passent périodiquement pour vider la carte mémoire remplie de photos.

Or, le détecteur de mouvement a pu être déclenché par des feuilles poussées par le vent, ou des espèces qui n'intéressent pas le chercheur. Il faut beaucoup de temps pour passer au crible toutes ces photos. La carte mémoire se remplit rapidement, il faut donc la vider fréquemment. Même avec une connexion internet, la forte consommation déchargerait les piles de l'appareil qu'il faudrait remplacer.

Grâce à un modèle *TinyML* entraîné à reconnaître le bon type d'animal, la caméra ne prendra que les bons sujets. Les caméras peuvent donc être laissées en place plus longtemps, tandis que les chercheurs font des économies de temps et d'argent.

Il y a quatre contraintes principales pour qu'une application soit idéale pour *TinyML*. Si une application a une connectivité limitée, une faible tolérance à la latence, une source d'alimentation limitée ou une exigence de confidentialité élevée, *TinyML* vous évitera l'envoi de données à partir de l'appareil.

Seuls certains problèmes peuvent être résolus par l'apprentissage automatique. Ce sont des problèmes pour lesquels un peu de flou est acceptable, car l'apprentissage automatique donne rarement des réponses absolues. Il est parfait pour prendre des photos d'animaux, où l'erreur occasionnelle est acceptable, mais il n'est pas assez fiable dans certaines applications critiques pour la sécurité.

Abate : Tu es co-auteur du livre, *TinyML : Machine Learning with TensorFlow Lite on Arduino and Ultra-Low-Power Microcontrollers*. Comme il couvre Arduino, cela me porte à croire que la technologie *TinyML* est accessible aux non-experts comme aux professionnels. Outre la lecture de ton livre, comment les ingénieurs et les innovateurs qui ne sont pas familiers avec *TinyML* peuvent-ils se lancer ?

Situnayake : Comme toujours, la meilleure façon d'apprendre est de construire soi-même ! Un bon point de départ est de suivre certains des tutoriels fournis dans la documentation d'*Edge Impulse* : la reconnaissance de mouvement avec un accéléromètre [2] ; la reconnaissance de sons à partir de l'audio [3] ; et l'ajout de la vue à vos capteurs [4].

Nous avons construit *Edge Impulse* pour que tu puisses utiliser votre téléphone portable pour collecter des données et tester des modèles, afin de commencer sans matériel spécial. Une fois familia-

risé avec le flux d'apprentissage de la machine, tu peux prendre ta carte de développement préférée et commencer à programmer ! Je recommande l'Arduino Nano 33 BLE Sense comme tremplin pour le TinyML. Cette carte possède un processeur Cortex M-4F rapide, mais de faible puissance, et suffisamment de RAM pour certains modèles intéressants. Elle est entièrement prise en charge dans Edge Impulse.

Impulse Edge

Abate : Tu as rejoint l'organisation Edge Impulse, basée à San José, en janvier 2020. Comment cela s'est-il fait ?

Situnayake : J'ai rencontré le PDG, Zach Shelby quand j'étais chez Google. Il était auparavant vice-président chez ARM, qui travaille en étroite collaboration avec l'équipe de TensorFlow Lite. Quand Zach a annoncé qu'il quittait ARM pour fonder une société TinyML, cela a attiré mon attention, et quand ils ont publié leur première offre d'emploi, j'ai postulé !

Abate : Comme se déroule ta journée de travail ? Codes-tu toute la journée ? Enseignes-tu ? Programmes-tu pour des entreprises ?

Situnayake : Ce que j'apprécie le plus à Edge Impulse c'est d'être redevenu principalement un ingénieur. Je passe la plupart de mon temps à programmer. J'ai toujours l'occasion de rencontrer des clients, ce qui est assez gratifiant et crucial pour savoir comment concentrer mon temps d'ingénieur. Je suis également très actif dans la communauté TinyML, c'est pourquoi je donne régulièrement des conférences et j'aide à organiser des rencontres TinyML. Cela dit, j'ai généralement deux ou trois réunions le matin, ce qui me permet de me concentrer et coder le reste de la journée. C'est génial !

Abate : Je vois que l'équipe d'Edge Impulse a commencé le projet COVID-19 en mai - "Détection de toux avec TinyML sur Arduino" (**fig. 3**). Le projet a été visionné plus de 5 400 fois. Quelle a été la réaction de la communauté ? Quel est le statut du projet ?

Situnayake : Même si Edge Impulse n'existe que depuis quelques mois, nous avons déjà une communauté fantastique venue d'horizons divers. Lorsque la COVID-19 a frappé, l'un des membres de notre communauté, Kartik Thakore, chercheur et ingénieur biomédical, était intéressé par la construction de quelques projets pour nous aider. Persuadés que ce serait une bonne occasion de montrer le potentiel de TinyML, nous lui avons donc envoyé du matériel et il a monté un détecteur de toux avec un modèle TinyML. Le résultat a été ce projet [5], qui, nous l'espérons, inspirera d'autres innovateurs à relever des défis similaires !



Figure 3. DéTECTEUR DE TOUX AVEC TINYML SUR ARDUINO (Source : Edge Impulse).

Projets TinyML

Abate : Y a-t-il d'autres projets TinyML que tu souhaites mettre en avant pour Elektor ?

Situnayake : Chez Edge Impulse, nous croyons à la technique comme **force du bien**, et certains de mes projets préférés relèvent de cette catégorie. Les Smart Parks forment p. ex. des modèles TinyML pour détecter les éléphants sauvages, afin de prévenir les villageois que des pachydermes se dirigent vers eux.

Une autre de mes catégories favorites est celle des personnes qui font faire aux ordinateurs des choses inouïes. Comme le projet étonnant de Benjamin Cabé : un capteur de gaz reconnaît différentes marques de spiritueux ! S'il vous intéresse, il existe un tutoriel génial de Seeed Studio. ↗

200423-03

Votre avis, s'il vous plaît...

Vous pouvez adresser vos questions et vos commentaires (en anglais) à la rédaction : editor@elektor.com

Ont contribué à cet article :

Entretien : **C. J. Abate**

Maquette : **Giel Dols**

Traduction : **Henry Philippe**

LIENS

- [1] Open Tiny Farms: <https://www.opentinyfarms.com/>
- [2] Edge Impulse, "Continuous Motion Recognition": <https://docs.edgeimpulse.com/docs/continuous-motion-recognition>
- [3] Edge Impulse, "Recognize Sounds from Audio": <https://docs.edgeimpulse.com/docs/audio-classification>
- [4] Edge Impulse, "Adding Sight to Your Sensors": <https://docs.edgeimpulse.com/docs/image-classification>
- [5] Edge Impulse, "Cough Detection with TinyML on Arduino": <http://bit.ly/edge-cough-detection-project>