

Introduction

Le Deep Learning (DL) basé sur la reconnaissance des formes est une méthode très performante et puissante. Cependant cette méthode ne peut s'appliquer qu'aux images. Pour bénéficier de cette technologie quand on analyse des bases de données de type tabulaire, il faut trouver une méthode pour transformer ces données tabulaires en image. Toutes les équipes au monde repositionnent les variables dans une matrice en fonction de leur corrélation. C'est le cas pour la technologie de DeepInsight (A Sharma, 2021), qui a permis de battre dans un concours Kangle, toutes les autres méthodes d'apprentissage. C'est la preuve que le DL basée sur les CNN apporte un plus incontestable pour les données tabulaires. Cependant DeepInsight ne permet pas de faire de la prévision pour les observations.

Nous avons inventé une technologie, DEEPNIC, qui permet de transformer chaque variable en une image (One Variable One Image), appelée ROP-Image, là où tous nos concurrents font une seule image pour toute une base de données. Cette technologie a été lauréat du DeepTech Starter 2021. Nous avons transformé l'expression de 54675 de gènes d'une puce AFFIMETRIX en 54675 images en couleur où chaque pixel de chaque image-ROP représente la performance de prédiction d'un gène dans une condition de simulation donnée. Nous avons trouvé une solution pour transformer chaque observation en une image qui rassemble la totalité des informations de prédiction des variables de chaque observation.

Chacune de ces ROP-images est constituées de 59019 combinaisons d'une image de 18 pixels.

Stages

Nous recherchons pour une période de 6 mois, des ingénieurs en fin de cursus, dans le cadre de leur stage de fin d'étude pour participer au développement de cette technologie.

Le stage consiste à rechercher à une solution pour assembler ces 59019 ROP-images. Il y a la technique linéaire (de gauche à droite et de haut en bas), la technique de l'escargot, la technique « PAC-MAN », la technique du vers de terre. Cependant, ces techniques perdent en efficacité s'il y a trop de pixels non informatifs.

L'objet du stage consiste à filtrer ces pixels avant de les assembler selon les 5 méthodes puis à comparer les résultats par la mise en œuvre de techniques de DL.

Compétences et qualités requises :

- Programmation C/C++.
- Bases du DL, Bonne connaissance du prompt Chat GPT
- Esprit d'équipe, novateur et curieux.
- Autonomie

Conditions du stage

- Télétravail
- Visioconférence hebdomadaire.
- Dotation de 2 PC puissant (I9 extrem 18 cœurs 7980x) + carte graphique NVIDIA Quadro 8000
- Abonnement fibre optique pendant la durée du stage
- Indemnités conventionnel de stage (environ 500-600 euros)
- Contrat confidentialité très strict à signer
- Référents :

Projet DEEPNIC : CHU de Nantes

D Learning : LITIS, Rouen

Calcul Intensif : ECentrale Nantes

Avantages du stage

Participer à une recherche pointue en IA, contribuer et assister à la création d'un projet de start-up
Intégration dans l'équipe de recherche.