

CNN-Cancer

Jonas FLÜCKIGER

Travail de bachelor 2024

Informatique et systèmes de communications – Ingénierie logicielle

Professeur: Cédric BILAT

Expert: Diego ANTOGNINI

Description

Que se passe-t-il dans le cerveau d'une IA ?

La recherche en machine learning et en data science connaît une croissance exponentielle, avec des applications diverses allant de la médecine à la finance en passant par l'industrie et les sciences sociales. L'un des aspects importants de cette recherche est **l'interprétabilité**. C'est-à-dire comprendre comment une IA est arrivée à un résultat spécifique, comprendre son raisonnement.

C'est dans le cadre de cette recherche que les chercheurs en mathématiques de l'Université de Genève (UniGE) ont mandaté l'école d'ingénieurs HE-Arc, pour mettre au point un package logiciel robuste et optimisé, capable de faciliter et d'accélérer leurs travaux de recherche.

Le projet avait pour but d'intégrer la recherche de l'UniGE dans la bibliothèque de machine learning «PyTorch» dans le langage Python, et de développer d'une manière à faciliter l'utilisation de matériel de calcul puissant.

Déroulement

Le projet s'est déroulé en trois étapes distinctes.

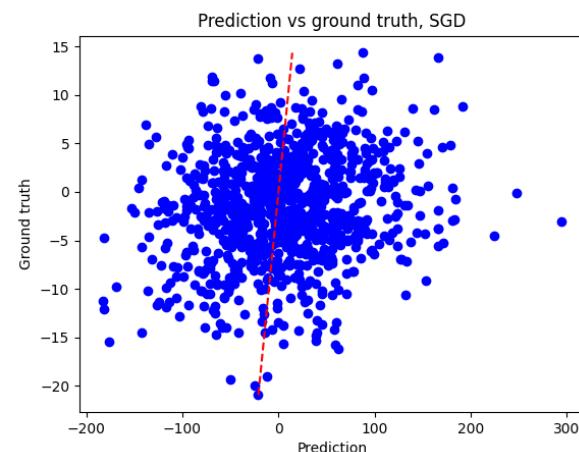
- **Phase d'approfondissement des connaissances** : Une phase d'approfondissement des connaissances en machine learning, permettant à l'élève de combler toutes les potentielles lacunes ;
- **Étude de faisabilité technologique** : Une recherche a été réalisée pour évaluer la possibilité d'intégrer des technologies d'accélération matérielle (CUDA) dans le projet ;
- **Développement** : Finalement, la phase de développement logiciel principale. Les différents aspects de la recherche de l'UniGE ont été adaptés pour être implémentés dans le code Python, et plusieurs tests et simulations ont été effectués pour valider la justesse des algorithmes.

Résultats

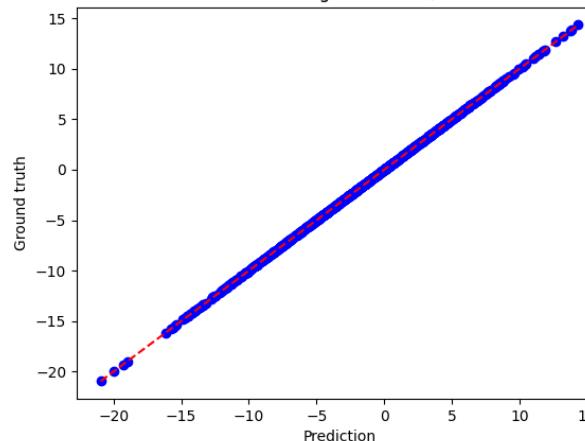
Les phases préliminaires au projet ont permis d'établir une base de connaissance solide pour l'étudiant, mais aussi pour de futurs collaborateurs, avec de nombreux documents et exemples de code. Quant aux algorithmes implémentés, ils semblent mieux performer que les algorithmes de base pour les tâches spécifiques à la recherche de l'UniGE.

De manière générale, les résultats obtenus par les modèles du projet sont plus justes, et plus interprétables, que ceux obtenus par les modèles par défaut.

Résultats obtenus par un modèle de base



Prediction vs ground truth, ISTA



Résultats obtenus par le modèle UniGE/HE-Arc

Discussion : Conclusions et perspectives

Finalement, les éléments logiciels développés sont utilisables pour la recherche de l'UniGE et semblent mieux performer que les algorithmes par défaut. L'utilisation de technologies et de techniques de développement modernes permettent une augmentation des performances, grâce à l'accélération matérielle. Une documentation extensive a été rédigée, permettant à de futurs ingénieurs ou chercheurs de facilement reprendre le projet, et d'ajouter des composants ou d'améliorer les performances.