



BERLINER HOCHSCHULE FÜR TECHNIK

Prédire le parcours des patients, de l'admission à la sortie

« L'utilisation d'équipements basés sur le Cloud présente de nombreux enjeux administratifs dans un environnement universitaire en Allemagne. Ayant nos propres systèmes NVIDIA DGX A100, nous pouvons agir rapidement et rester compétitifs en matière de recherche. »

— Alexander Löser, professeur et responsable du Centre de recherche en science des données

Situés en première ligne quel que soit l'hôpital, les professionnels de la santé doivent prendre des décisions où le temps est compté. Les informations sur les patients sont souvent non structurées, sous forme de notes cliniques, généralement écrites par d'autres membres du personnel médical disposant de peu de temps.

L'université Berliner Hochschule für Technik (BHT) a développé un système d'aide à la décision clinique, qui peut aider les médecins dans ces situations en les orientant vers des cas liés ou vers certains risques et en prédisant l'évolution d'un patient à partir des données figurant dans son dossier médical électronique. La prédiction des résultats peut également aider les médecins à éviter de négliger tout risque possible et les hôpitaux à planifier leurs capacités.

QALe modèle CORE (représentations de résultats cliniques), basé sur BioBERT, est pré-entraîné en s'appuyant sur 10 000 rapports de cas de PubMed, 32 000 feuilles médico-administratives, 5 000 transcriptions médicales, 5 000 notes cliniques et quelques milliers d'articles sur les maladies et la recherche médicale provenant de Wikipédia et des sites des instituts nationaux de la santé.

La plateforme NVIDIA

À la recherche de performances et de rapidité, l'équipe s'est donc tournée vers NVIDIA DGX A100, le système universel pour toutes les charges de travail d'IA (de l'analyse de données à l'inférence en passant par l'entraînement) offrant un niveau de densité de calcul, de performances et de flexibilité sans précédent, avec 5 pétaflops de performances en matière d'IA.

Le système DGX a permis d'effectuer un entraînement rapide, le débogage et l'analyse des erreurs pour les très gros modèles. L'université BHT a donc pu optimiser et ajuster la conception des réseaux de neurones artificiels dès le début du processus.

Grâce au traitement du langage naturel (TLN), le modèle peut extraire des informations à partir des données d'admission des patients afin de prédire des mesures clés.

ADMISSION

MALADIE PRÉSENTE : homme de 58 ans avec antécédents d'hypertension, fibrillation auriculaire lors de la prise de coumaphène et diabète de type 2 indiqué au service des urgences avec le plus gros mal de tête de sa vie. Il a fait une syncope et a été intubé par le service de gestion des urgences. Traitements à l'admission : 1 mg d'Ativan par IV x1.

EXAMEN PHYSIQUE : Signes vitaux : P : 92 R : 13 PA : 151/72 SaO2 : 99 % après intubation. Echelle de Glasgow : E : 3 V : 2 M : 5 Examen oculaire et ORL : atraumatique et normocéphalique Pupilles : 4 à 3 mm [...]

ANTÉCÉDENTS FAMILIAUX : sa mère a fait un AVC à 82 ans. Père inconnu.

ANTÉCÉDENTS SOCIAUX : vit avec sa femme. 25 paquets-années. Pas d'alcool

Symptômes et signes vitaux
Antécédents
Traitements
Facteurs de risques généraux

SORTIE

DIAGNOSTICS :
430 Hémorragie méningée
401 Hypertension artérielle essentielle
250 Diabète sucré [...]

INTERVENTIONS :
397 Réparation endovasculaire de vaisseaux
967 Ventilation mécanique invasive continue [...]

RISQUE DE DÉCÈS PENDANT L'HOSPITALISATION :
Non décédé

DURÉE DU SÉJOUR :
> 14 jours

BHT Berliner Hochschule für Technik

Prédire les résultats cliniques à partir des notes d'admission en utilisant l'intégration autosurveillée des connaissances

SECTEUR

> Médecine et Santé

PRODUITS NVIDIA UTILISÉS

> NVIDIA® DGX™ A100

L'ENJEU DU SECTEUR

- > Les délais de prise de décision sont très limités lorsque des patients sont admis à l'hôpital
- > Les informations sur les patients proviennent de diverses sources non structurées

RÉSULTATS

- > En utilisant le TLN, les données des patients peuvent toutes être comparées aux données d'entraînement afin de réaliser des prédictions
- > Le système peut prédire 1 200 diagnostics et 700 interventions au moment de l'admission, le risque de décès pendant l'hospitalisation et la durée du séjour
- > Pour les prédictions de diagnostic, la zone en dessous de la courbe ROC a atteint 83 % et même 88 % pour les prédictions d'intervention.

Un exemple de l'admission à la sortie, illustrant une prédiction de résultats. Le modèle extrait les variables des patients et, en utilisant le TLN, se familiarise avec les relations complexes qu'il existe entre les données afin de prédire un résultat clinique.

Résultats

En analysant les patients de la simulation au moment de l'admission (moment où l'aide à la prise de décision est la plus utile), le système de l'université BHT peut prédire les résultats sur quatre mesures communes : le diagnostic à la sortie, les interventions subies, le risque de décès pendant l'hospitalisation et la durée du séjour. Ces prédictions sont déduites à partir des informations dont ils disposent sur les patients, comme les symptômes à l'admission, les antécédents et les facteurs de risque.

À ce jour, l'université BHT est en mesure de classer, au moment de l'admission, 1 200 diagnostics et 700 interventions sous forme de codes de la Classification internationale des maladies (CIM-9), mais aussi de prédire le risque de mortalité et la durée du séjour avec une zone en dessous de la courbe ROC (fonction d'efficacité du récepteur) d'environ 83 % pour les diagnostics et de 88 % pour les interventions.

Initialement, l'ajustement du modèle CORE basé sur BERT pour la tâche de prédiction de diagnostic prenait environ 7 heures sur un serveur équipé de deux GPU NVIDIA V100.

Le passage de l'entraînement à deux GPU A100 sur le système DGX a presque doublé la vitesse du processus, aboutissant ainsi à des durées d'entraînement inférieures à 4 heures. Cela nous a permis d'exécuter un plus grand nombre d'itérations d'entraînement pour optimiser les hyperparamètres, l'ordre de pré-entraînement et le processus de sélection des données.

Le groupe de recherche DATEXIS

Le groupe de recherche DATEXIS (science des données et systèmes d'informations textuelles) au sein de l'université BHT mène des recherches en matière de gestion des données textuelles et structurées. Il se concentre sur la recherche de base dans les domaines du traitement du langage naturel (TLN) et du Deep Learning, expliquant et faisant l'analyse comparative du Deep Learning, et appliquant le TLN au secteur de la santé et à d'autres domaines.

Berliner Hochschule für Technik, Centre de recherche en science des données

Fondée en 1971, l'université BHT est l'une des plus grandes universités publiques allemandes spécialisées dans les sciences appliquées. Grâce à son large éventail de cursus novateurs et à son personnel spécialisé hautement qualifié, l'université BHT encourage les perspectives de carrière des 13 000 étudiants actuels et potentiels, quel que soit leur parcours social.

Au cours de ces dernières années, l'université BHT est devenue un acteur universitaire majeur dans le monde de l'IA à Berlin, grâce à sa Classe de maître pour la science des données de renommée mondiale. Le Centre de recherche en science des données, qui compte 13 enseignants en IA et plus de 30 doctorants, est l'un des plus grands établissements universitaires travaillant sur l'IA à Berlin. Nos recherches de base incluent des domaines d'application comme la santé, l'enseignement ou l'ingénierie.

Le dossier d'EACL 2021 et le modèle sont disponibles
sur huggingface.co/bvanaken/CORe-clinical-outcome-biobert-v1

En savoir plus

Découvrez les solutions de NVIDIA pour la recherche médicale