



# BERLINER HOCHSCHULE FÜR TECHNIK

Prognosen für die Aufnahme von Patienten bis zu deren Entlassung

„Bei der Verwendung cloudbasierter Hardware existieren in Deutschland hohe Verwaltungshürden für akademische Einrichtungen. Mit unseren eigenen NVIDIA DGX A100-Systemen können wir schnell agieren und in der Forschung wettbewerbsfähig bleiben.“

– Alexander Löser, Professor und Leiter des Forschungszentrums für Data Science

An vorderster Front eines Krankenhauses müssen medizinische Fachkräfte zeitkritische Entscheidungen treffen. Die verfügbaren Patienteninformationen liegen häufig als unstrukturierte klinische Informationen vor, die in der Regel von anderen, ebenfalls unter Zeitdruck stehenden, medizinischen Mitarbeitern erfasst wurden.

Die Berliner Hochschule für Technik (BHT) hat ein System zur Unterstützung klinischer Entscheidungen entwickelt, das in solchen Situationen helfen kann, indem Ärzten ähnliche Fälle oder bestimmte Risiken zugeführt werden und auf der Grundlage der Daten der elektronischen Gesundheitsakte eines Patienten eine Prognose der Entwicklung erstellt wird. Mit einer Ergebnisprognose können Ärzten außerdem vermeiden, dass mögliche Risiken übersehen werden. Die Krankenhäuser werden bei der Kapazitätsplanung unterstützt.

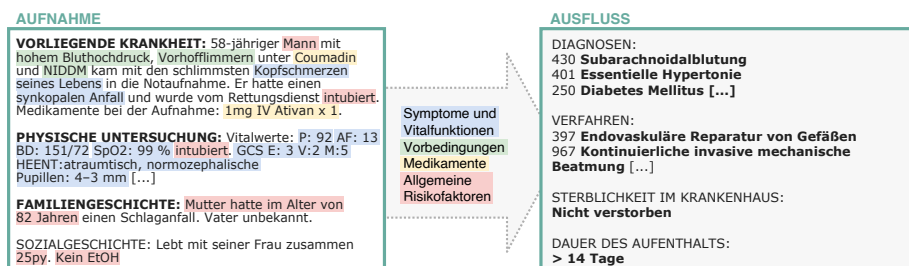
Das auf BioBERT basierende CORE-Modell (Clinical Outcome Representations) wurde vorab trainiert mit 10.000 Fallberichten von PubMed, 32.000 Entlassungsberichten, 5.000 medizinischen Transkriptionen, 5.000 klinischen Informationen und einigen tausend Artikeln zu Krankheiten und medizinischer Forschung, die von den Websites von Wikipedia und NIH stammen.

## NVIDIA-Plattform

Für eine schnelle Performance hat das Team für Analysen, Training, Inferenz auf NVIDIA DGX A100 gesetzt – das universelle System für sämtliche KI-Workloads. Dieses bietet beispiellose Computing-Dichte, Leistung und Flexibilität mit 5 PetaFLOPS KI-Leistung.

Das DGX-System hat das schnelle Trainieren, Debuggen und die Fehleranalyse der sehr großen Modelle ermöglicht. BHT konnte somit bereits in einer frühen Phase des Prozesses das Design der neuronalen Netze optimieren und verfeinern.

Mithilfe von NLP (Natural Language Processing) kann das Modell Informationen aus den Aufnahmedaten der Patienten extrahieren, um wichtige Kennzahlen zu prognostizieren.



Ein Beispiel für „Aufnahme bis Entlassung“, das eine Ergebnisprognose demonstriert. Das Modell extrahiert Patientenvariablen und erlernt mittels NLP komplexe Beziehungen zwischen den Daten, um ein klinisches Ergebnis vorherzusagen.

**BHT** Berliner Hochschule für Technik

Prognosen klinischer Ergebnisse aus Aufnahmeinformationen mithilfe selbstüberwachender Wissensintegration

### BRANCHE

> Medizin und Gesundheitswesen

### VERWENDETE NVIDIA-PRODUKTE

> NVIDIA® DGX™ A100

### GESCHÄFTLICHE HERAUSFORDERUNG

> Bei der Aufnahme von Patienten in ein Krankenhaus geht es um äußerst kurze Entscheidungszeiten

> Patienteninformationen stammen aus unterschiedlichen, unstrukturierten Quellen

### ERGEBNISSE

> Mithilfe von NLP können die Patientendaten anhand von Trainingsdaten analysiert werden, um Prognosen zu erstellen

> Das System prognostiziert zum Zeitpunkt der Aufnahme 1.200 Diagnosen und 700 Behandlungen, die Krankenhausletalität und die Aufenthaltsdauer

> Diagnoseprognosen erzielen einen AUROC-Wert von 83 %, Eingriffprognosen sogar von 88 %

## Ergebnisse

Durch die Analyse simulierter Patienten zum Zeitpunkt der Aufnahme – wenn die Entscheidungsunterstützung am dringendsten benötigt wird – kann das BHT-System Ergebnisse für vier gängige Kennzahlen prognostizieren: Diagnose bei der Entlassung, durchgeführte Behandlungen, Krankenhausletalität und Aufenthaltsdauer. Diese Prognosen werden aus Patienteninformationen abgeleitet, z. B. aus Symptomen bei der Aufnahme, Vorerkrankungen und Risikofaktoren.

Die BHT kann bislang zur Aufnahmezeit 1.200 Diagnosen und 700 Behandlungen als ICD-9-Codes klassifizieren und die Letalität und Aufenthaltsdauer mit einem AUROC-Wert (bezeichnet die Fläche unter der Grenzwertoptimierungskurve) von rund 83 % für Diagnosen und 88 % für Behandlungen prognostizieren.

Die Feinabstimmung des BERT-basierten CORE-Modells für die Diagnoseprognoseaufgabe dauerte anfangs auf einem Server mit zwei NVIDIA V100-Grafikprozessoren rund sieben Stunden.

Durch die Verlagerung des Trainings auf zwei A100-Grafikprozessoren auf dem DGX-System konnte die Prozessgeschwindigkeit nahezu verdoppelt und die Trainingszeit somit auf unter 4 Stunden reduziert werden. So konnten wir eine größere Anzahl von Trainingsiterationen durchführen, um die Hyperparameter, die Reihenfolge vor dem Training und den Datenauswahlprozess zu optimieren.

## Forschungsgruppe DATEXIS

Die Forschungsgruppe DATEXIS (Data Science and Text-based Information Systems) an der BHT forscht zur Verwaltung textbasierter und strukturierter Daten. Ihr Schwerpunkt liegt auf der Grundlagenforschung zu NLP (Natural Language Processing) und Deep Learning, Darlegung und Benchmarking von Deep Learning sowie Anwendung von NLP in der Medizin und dem Gesundheitswesen sowie in anderen Bereichen.

## Berliner Hochschule für Technik, Data Science Research Center

Die BHT ist eine der größten staatlichen Fachhochschulen in Deutschland. Sie wurde 1971 gegründet. Mit einem breiten Angebot an zukunftsweisenden Studiengängen und hochqualifizierten Fachkräften fördert die BHT die Karrierechancen der angehenden sowie der 13.000 aktuellen Studierenden unabhängig von deren sozialem Hintergrund.

In den letzten Jahren hat sich die BHT mit ihrem international anerkannten Masterstudiengang Data Science zu einem der großen akademischen Akteure der Berliner KI-Landschaft entwickelt. Das Data Science Research Center mit 13 Professoren für KI und über 30 Doktoranden ist eine der größeren akademischen Organisationen, die in Berlin im KI-Bereich arbeiten. Unsere Grundlagenforschung umfasst Anwendungsbereiche wie Gesundheits-, Bildungs- und Ingenieurwesen.

Das EACL 2021-Dokument und das Modell sind verfügbar unter [huggingface.co/bvanaken/CORE-clinical-outcome-biobert-v1](https://huggingface.co/bvanaken/CORE-clinical-outcome-biobert-v1)

## Weitere Informationen

Entdecken Sie die NVIDIA-Lösungen für die **Forschung im Gesundheitswesen**