





Use Case 2 — From Data to Knowledge

Vorhersagemodelle für Asthma/COPD und Hirntumoren

Harald Binder 11.04.2019





Ziele im Use Case 2



- Modelle zur Vorhersage von prognostisch relevanten Untergruppen entwickeln
- Vorhersagemodelle in die klinische Praxis bringen
- Das Ganze am Beispiel von
 - Asthma/COPD und
 - Neuroonkologie





Ziele im Use Case 2



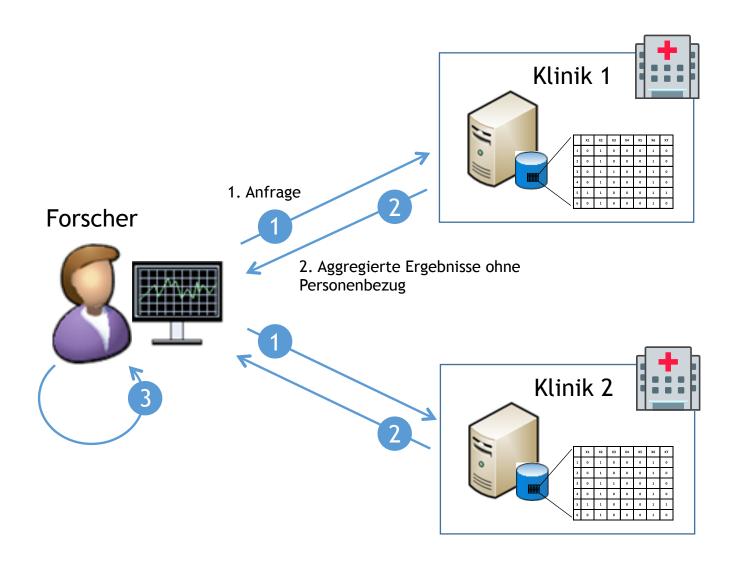
- Modelle zur Vorhersage von prognostisch relevanten Untergruppen entwickeln
- Vorhersagemodelle in die klinische Praxis bringen
- Das Ganze am Beispiel von
 - Asthma/COPD und
 - Neuroonkologie





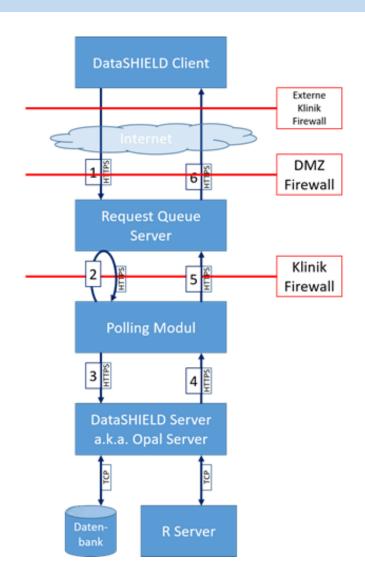
DataSHIELD: Prinzip





Polling-Konzept als zusätzliche Sicherheitsschicht für DataSHIELD





- 1: DataSHIELD Analyse Request (Abfrage)
- 2: Polling Modul überprüft regelmäßig, ob neue Requests in der Queue vorliegen
- 3: Vorliegende und freigegebene Requests werden an den DataSHIELD Server zur Verarbeitung weitergegebene
- 4: Fertige Analyse wird weitergeleitet
- 5: Nach Freigabe wird die Analyse weitergeleitet
- 6: Finale Antwort an den Client

Datenharmonisierung



- Erfassung und Beschreibung relevanter und verfügbarer Variablen
- Abbildung auf LOINC/OMOP
- Erhebung mittels Fragebögen an die Kliniker der verschiedenen Standorte
- Separate Variablenlisten für Teile Asthma/COPD und Neuroonkologie





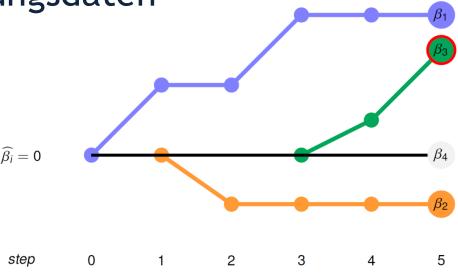


Distributed Boosting



- Boosting-Algorithmus f
 ür verteilte Daten
- Umgesetzt als Julia-Paket, aufbauend auf DataSHIELD github.com/danielazoeller/ds_DistributedBoosting.jl
- Erstes Manuskript: https://arxiv.org/abs/1803.00422

 Derzeit: Verteilte Analyse von DNA-Methylierungsdaten

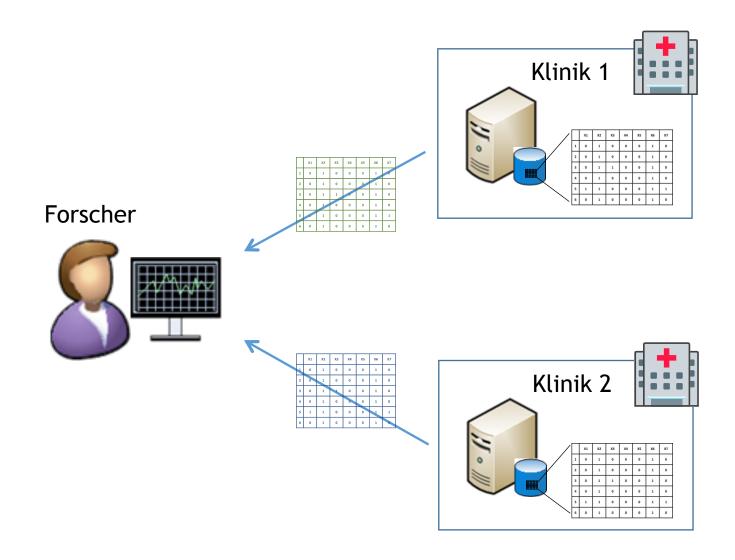






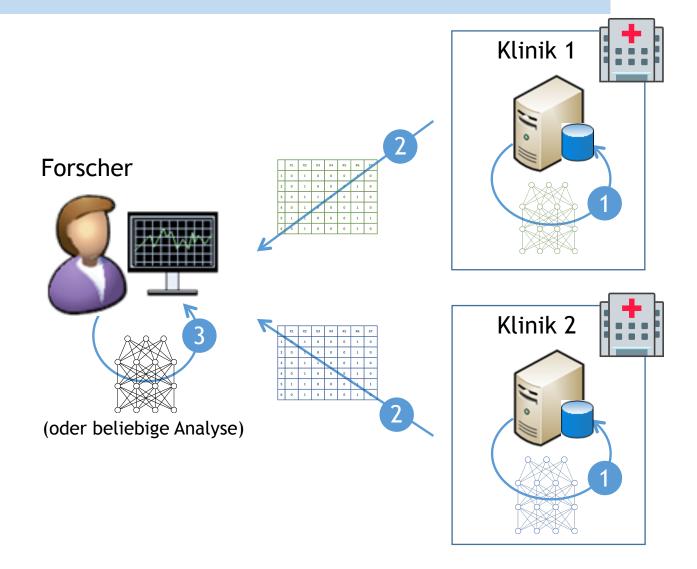
Generierung synthetischer Daten





Verteiltes Deep Learning mit generativen Modellen





2018 DataSHIELD Workshop, 06.-08.11

















