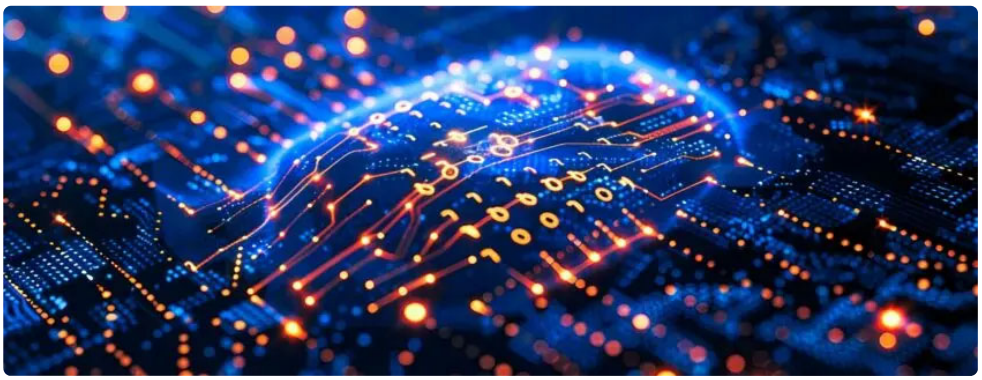
SocietyByte

BFH-Magazin für die Humane Digitale Transformation

Federated Learning: Die Zukunft der KI ohne Kompromisse beim Datenschutz

Von Cedric Aebi (BFH Technik & Informatik) , Souhir Ben Souissi (BFH Technik & Informatik) | 0 Kommentare



Federated Learning hat sich zu einer paradigmenverändernden Technologie in der KI entwickelt, die es Datenwissenschaftlern ermöglicht, mit privaten Daten zu arbeiten. Im Vergleich zum Standardansatz des maschinellen Lernens, bei dem alle individuellen Client-Daten gesammelt und auf einen zentralen Server übertragen werden müssen, kann das Modelltraining beim FL(deutsch auch föderiertes/föderales Lernen) direkt auf den Clients durchgeführt werden, ohne dass sensible Daten preisgegeben werden.

Wenn wir von «standardmässigem maschinellem Lernen» sprechen, meinen wir den klassischen Ansatz, bei dem Daten von allen verschiedenen Client-Geräten gesammelt und dann in ein zentrales Repository übertragen werden, wo die Daten gespeichert werden. Datenwissenschaftler*innen erstellen, trainieren und evaluieren dann ein Modell anhand dieses grossen Datensatzes.

Wie Sie sich vielleicht schon denken können, hat dieser Ansatz mehrere Nachteile. Die sensiblen Kundendaten werden während des gesamten End-to-End-Prozesses mehrmals offengelegt. Sie müssen über das Netz transportiert werden, und darüber hinaus kann jeder Datenwissenschaftler (oder in diesem Zusammenhang jede Person) mit ausreichenden Rechten auf dem zentralen Repository ALLE Daten von allen Clients sehen.

Und genau hier setzt das FL an. Hier werden die Daten nicht gesammelt und in ein zentrales Repository/Server verschoben. Die Modellinformationen selbst werden direkt zu den Client-Geräten transportiert und dort mit den verfügbaren Daten des jeweiligen Clients trainiert. Dies kann für jede Einrichtung von Bedeutung sein, die mit sensiblen Daten umgeht und nicht bereit ist, diese Informationen weiterzugeben, wie z. B. elektronische Gesundheitsakten (EHR) von Krankenhäusern.

Wie funktioniert FL?

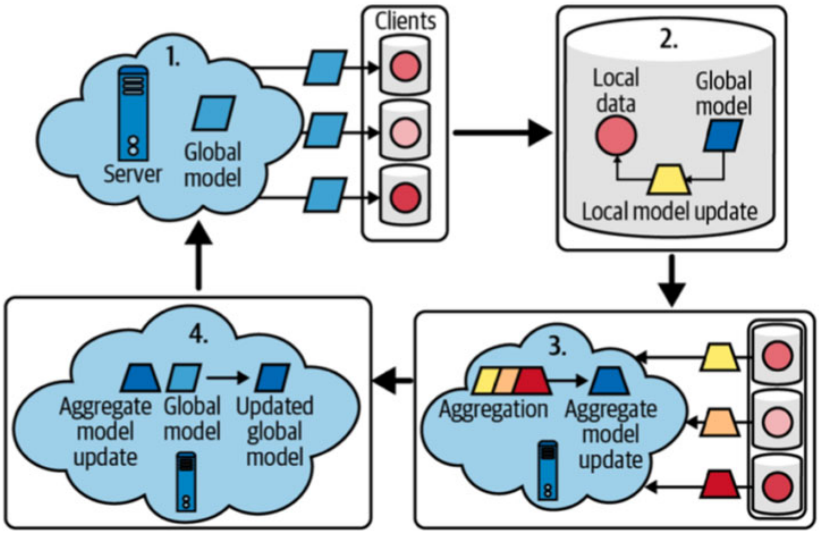


Abbildung 1 Schulungsprozess im Rahmen des FL (Quelle: Glanz & Fallen, 2021, The Federated Learning Lifecycle)

Die notwendigen Schritte für den FL-Trainingsprozess sind in *Abbildung 1* zu sehen

1. **Model Broadcast** – Das aktuelle Modell und die Trainingsanweisungen werden vom Orchestrierungsserver auf die Clients heruntergeladen. Ein «Modell» in diesem Sinne kann ein beliebiges neuronales Netz, ein lineares Regressionsmodell usw. sein.
2. **Client-Training** – Die Clients trainieren dann ihre eigenen Modelle lokal mit

ihren eigenen Daten.

3. **Aggregation** – Nachdem alle Clients ihre lokalen Modelle trainiert haben, werden die Modellaktualisierungen an den Server zurückgeschickt, wo alle Aktualisierungen von den Clients aggregiert werden. Die einfachste Form der Aggregation wäre, einfach den Durchschnitt aller Client-Aktualisierungen zu nehmen.
4. **Modellaktualisierung** – Nach der Aggregation wendet der Server diese Aggregation auf das globale Modell an, das später erneut an die Clients für eine neue Trainingsrunde gesendet wird.

Um nun die Frage zu beantworten, was ist ein «Client»? Wir sprechen in FL oft von *geräte-* und *siloübergreifend* . Für das geräteübergreifende Szenario kann ein Client beispielsweise ein beliebiges mobiles Gerät oder ein IoT-Gerät (Internet of Things) wie Smartwatches sein. Für das *siloübergreifende* Szenario kann ein Client eine ganze Institution, Organisation usw. sein. Sie haben dann ihre eigenen Server, auf denen ihre Daten gespeichert sind.

Wo kann FL eingesetzt werden?

Nachdem wir nun mit dem FL vertraut sind, wollen wir einige Fallstudien aus der Praxis untersuchen. Es gibt potenziell endlose Anwendungen für FL in Bereichen wie Gesundheitswesen, Finanzen, Smartphones, Smart Cities, Empfehlungssysteme usw. Konzentrieren wir uns vorerst auf zwei Beispiele.

NÄCHSTE WORTVORHERSAGE IN EINER MOBILEN TASTATUR

Die Vorhersage des nächsten Wortes zielt darauf ab, die Benutzererfahrung während des Tippens zu verbessern. Ausgehend von einer Phrase versucht das Modell, die besten Übereinstimmungen für das nächste Wort zu finden. Dies ist in Abbildung 2 zu sehen. Bei der Phrase «Ich liebe dich» sagt das Google Gboard als nächste Wörter «zu», dann «so sehr» und als drittbeste Option «und» voraus.

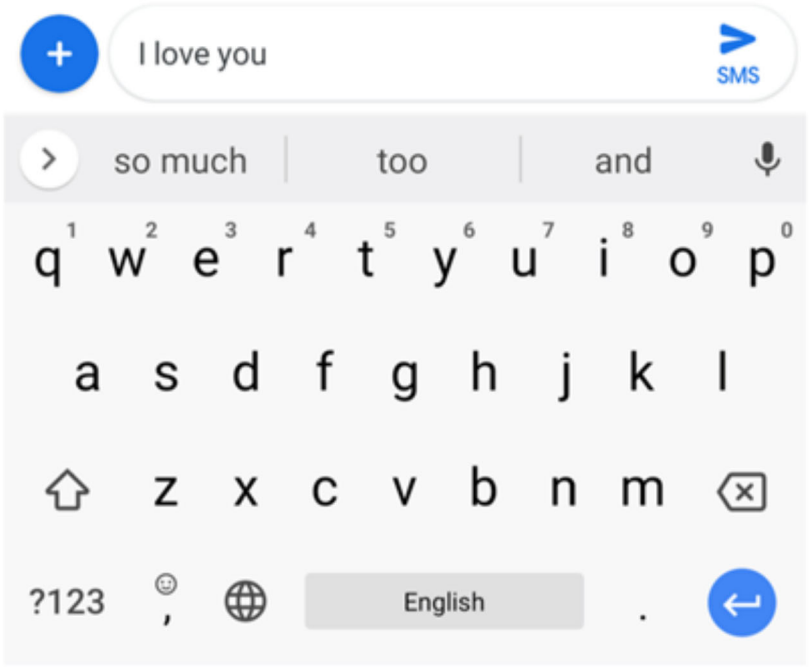


Abbildung 2: Nächste Wortvorhersagen in Gboard. Basierend auf dem Kontext «Ich liebe dich», sagt die Tastatur «und», «zu» und «so sehr» voraus. (Quelle: Hard et al., 2019)

Das Telefon des Nutzers speichert lokal die Informationen darüber, was er/sie in seine/ihre Tastatur getippt hat, und baut mit diesen Informationen einen eigenen Datensatz auf. Hard et al. (2019) haben ein Modell namens Coupled Input and Forget Gate (CIFG) entwickelt, das eine Variante eines Long Short-Term Memory (LSTM)-Modells zur Vorhersage der nächsten Wörter ist. Sie trainierten ihr Modell mit FL, so dass jedes Smartphone sein eigenes lokales CIFG-Modell trainierte, und dann wurden alle Modellaktualisierungen auf einem lokalen Server für mehrere Trainingsrunden aggregiert. Ihre Ergebnisse zeigten, dass das resultierende Modell eine vergleichbare Leistung wie ein zentral trainiertes Modell aufwies, ohne dass die Client-Daten jemals untereinander ausgetauscht wurden.

SCHLAGANFALLVORHERSAGE MIT FL

Die Vorbeugung von Schlaganfällen und der Umgang mit den damit verbundenen Risikofaktoren gehören weltweit zu den Prioritäten der öffentlichen Gesundheit. In ihrer Arbeit schlugen Ju et al. (2020) einen Rahmen für die Vorhersage des Schlaganfallrisikos vor und stellten ihr föderiertes Vorhersagemodell auf Cloud-Servern bereit. Sie arbeiteten mit 5 chinesischen Krankenhäusern zusammen, um ihren Ansatz zu validieren. Mehrere medizinische und technische Expert*innen verarbeiteten die in den Krankenhäusern verfügbaren medizinischen Rohdaten in Datensätze für ambulante und stationäre Patientenakten sowie für ambulante und stationäre Verschreibungen vor. Anschliessend wählten sie von Hand verschiedene Merkmale aus, die mit einem Schlaganfall in Zusammenhang stehen, und speisten diese Merkmale in ein neuronales Netz ein, das die Wahrscheinlichkeit eines Schlaganfalls vorhersagte. Sie trainierten ihr Modell mit FL und verwendeten den gewichteten Durchschnitt als Aggregationsalgorithmus. Ihr Ansatz ist in den Abbildungen 3 und 4 zu sehen. Abbildung 3 zeigt auch die Implementierung einer WeChat-Mini-Anwendung, bei der Expert*innen die Trainingsprozesse überwachen können. Auch hier sind die Ergebnisse mit denen eines zentral trainierten Modells vergleichbar.

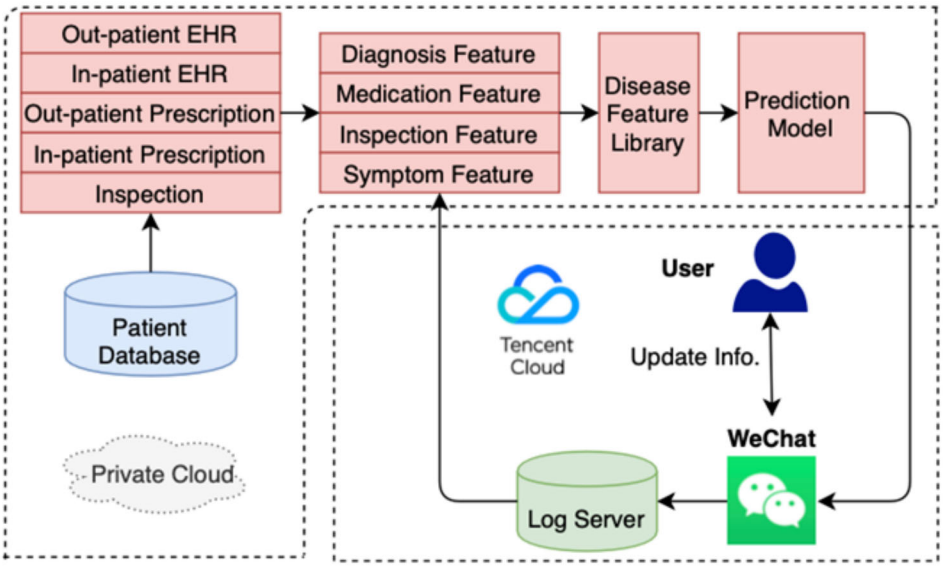
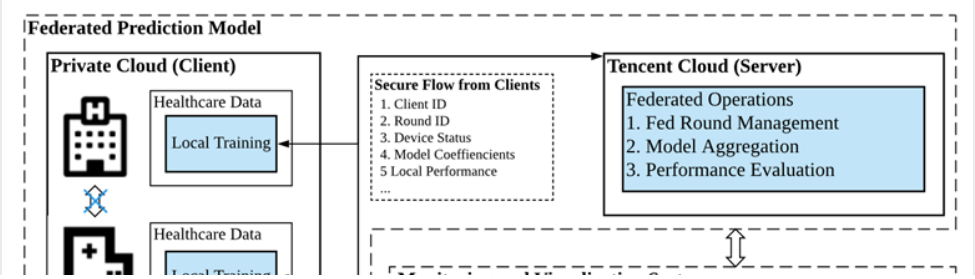


Abbildung 3 Arbeitsablauf des Schlaganfallvorhersagemodells (Quelle: Ju et al. 2020)



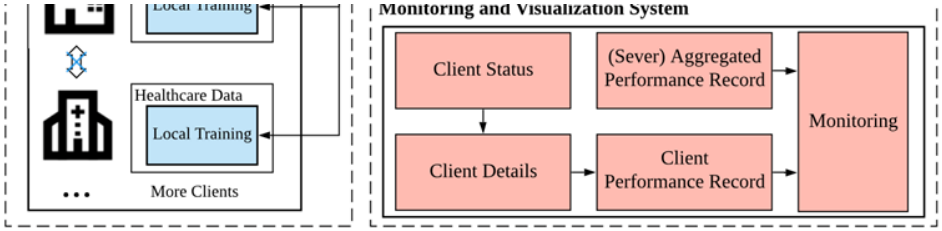


Abb4 Federated

Abbildung 4 Illustration der Architektur des föderierten Vorhersagemodells (Quelle: Ju et al., 2020)

Wie können Sie unsere eigenen FL-Modelle implementieren?

Es gibt bereits viele Tools und Frameworks für FL. Wenn Sie mit Googles TensorFlow vertraut sind, können Sie mit TensorFlow Federated(https://www.tensorflow.org/federated) [https://www.tensorflow.org/federated] loslegen. Allerdings bietet es nur eine Simulationsumgebung, in der Sie Ihre Annahmen testen und einige Proof of Concepts durchführen können. Zum Zeitpunkt der Erstellung dieses Artikels ist es nicht für den Produktionseinsatz gedacht.

Wenn Sie etwas wollen, das bereits weit verbreitet ist (hauptsächlich auf dem chinesischen Markt), könnten Sie FATE(https://fate.fedai.org) [https://fate.fedai.org] ausprobieren. Es unterstützt sowohl TensorFlow als auch PyTorch, sowohl den Simulationsmodus als auch Produktionsumgebungen und kann sowohl für geräte- als auch für siloübergreifende Szenarien verwendet werden.

Aber es gibt noch viele weitere Tools und Frameworks in der Entwicklung. Einige von ihnen sind hier aufgelistet:

- PySyft(https://com/OpenMined/PySyft) [https://github.com/OpenMined/PySyft]
- IBM Federated Learning(https://com/IBM/federated-learning-lib) [https://github.com/IBM/federated-learning-lib]
- Flower(https://github.com/adap/flower) [https://github.com/adap/flower]
- OpenFL(https://github.com/securefederatedai/openfl) [https://github.com/securefederatedai/openfl]
- FedML(https://doc.fedml.ai) [https://doc.fedml.ai/]

Schlussfolgerung

Zusammenfassend lässt sich sagen, dass FL dort von grossem Nutzen sein kann, wo Daten nur spärlich verfügbar sind und der Datenschutz ein Anliegen ist. Die neuartige Technologie hat sich bereits als nützlich erwiesen und zeigt in realen Anwendungen grosse Leistungsfähigkeit. Viele FL-Tools und -Rahmenwerke befinden sich in aktiver Entwicklung, was zeigt, dass FL von vielen grossen Akteur*innen viel Aufmerksamkeit geschenkt wird.

Referenzen

1. Glanz, E., & Fallen, N. (2021). *Was ist föderiertes Lernen?* (1. Auflage) [OCLC: 1281679172]. O'Reilly Media, Inc.
2. Hart, A., Rao, K., Mathews, R., Ramaswamy, S., Beaufays, F., Augenstein, S., Eichner, H., Kiddon, C., & Ramage, D. (2019, February 28). Federated learning for mobile keyboard prediction. Abgerufen am 16. Dezember 2023, von http://arxiv.org/abs/1811.03604 [http://arxiv.org/abs/1811.03604]
3. Ju, C., Zhao, R., Sun, J., Wei, X., Zhao, B., Liu, Y., Li, H., Chen, T., Zhang, X., Gao, D., Tan, B., Yu, H., He, C., & Jin, Y. (2020, December 14). Technologie zur Wahrung der Privatsphäre, um Millionen von Menschen zu helfen: Federated prediction model for stroke prevention. Abgerufen am 16. Dezember 2023, von http://arxiv.org/abs/2006.10517

AUTHOR: CEDRIC AEBI

Cedric Aebi ist Teilzeitstudent MSE mit Vertiefung Data Science an der BFH Technik & Informatik.

Posts from Cedric Aebi

AUTHOR: SOUHIR BEN SOUISSI

Dr. Souhir Ben Souissi ist Tenure-Track-Professorin für Data Engineering am Institut für Datenanwendungen und Sicherheit (IDAS) der BFH Technik & Informatik. Ihre Forschungsschwerpunkte liegen unter anderem auf den Themen Medizinische Entscheidungssysteme, Semantische Webtechnologien und Multikriterielle Entscheidungssysteme.

Posts from Souhir Ben Souissi | Website

Create PDF

Ähnliche Beiträge

Es wurden leider keine ähnlichen Beiträge gefunden.

0

COMMENTS