Pregnancy risk analytics: use case

> **FAIR** stands for **Findable**, **Accessible**, **Interoperable** and **Reusable**. By making data FAIR, data reuse is improved, by ensuring that one can easily find and access data, and that the data is interoperoperable so that it is usable in different places and systems.

FAIR patient data has a lot of potential when it comes to monitoring and analytics. One of the promising applications, is the development of classification systems to predict the risk level of certain patients. This was tested out using pregnancy data, to evaluate the risk level of a pregnancy based on maternal healthcare data. The classification model was built using a Federated Learning approach, meaning that the data is held in residence in health facilities and is only accessible through data visiting. This ensures that data privacy is protected and that the model complies with ethical and data management concerns.

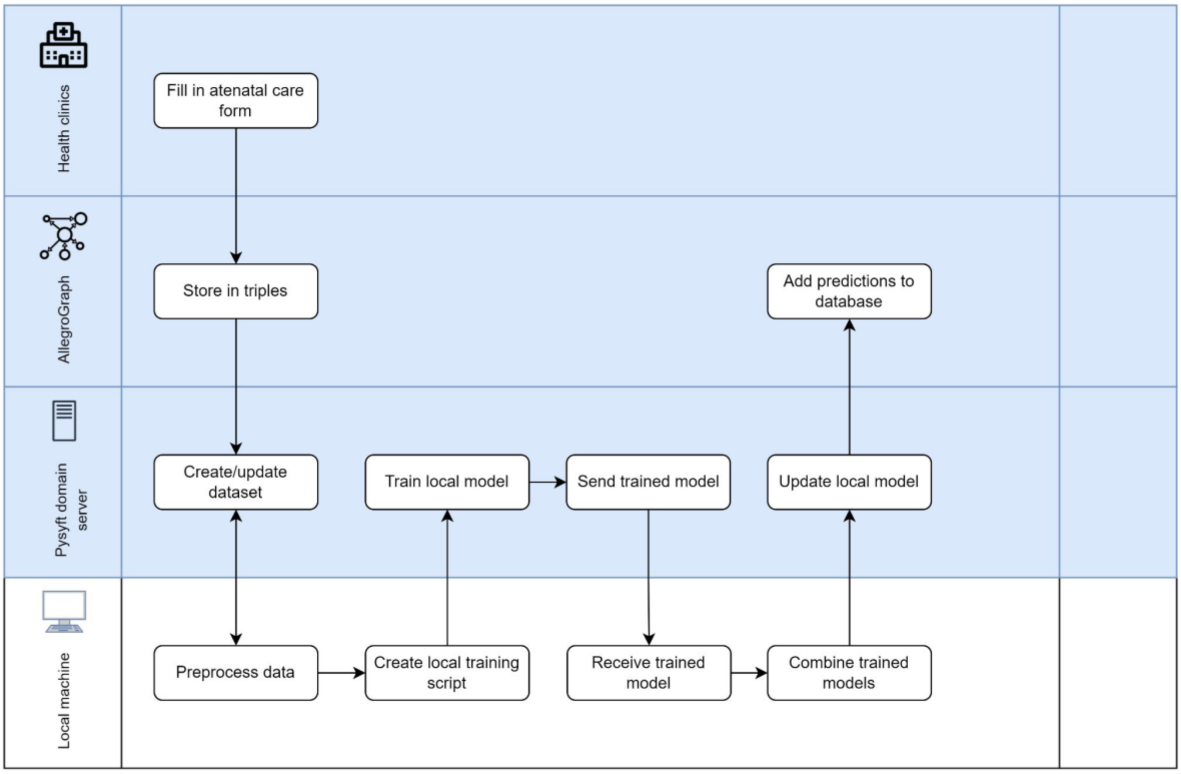

Figure 1. An overview of the classification model.

The federated architecture was build using PySyft, with libraries created on health facility servers. Data privacy was ensured by creating a script which removes identifier. The script also handled missing values and extract metadata. Next, a neural network classifier was built and trained to predict a risk level of 1 to 3 of a pregnancy. This application could be used to assign a more experienced health professional to higher risk pregnancies. Here also, there were strong limitations due to the lack of accessible data for analytics.

RECOMMENDATIONS

1. Provide financial or technical incentives for healthcare facilities that contribute to Federated Learning networks by sharing local data (in a privacy-preserving way). This could include funding, infrastructure upgrades, or access to advanced analytics tools.
2. Introduce standardized data collection protocols for maternal healthcare, including the consistent recording of data points that are crucial for pregnancy risk prediction. These standards should focus on data completeness, accuracy, and quality.
3. Develop ethical guidelines for the use of predictive models in healthcare, particularly for sensitive areas like maternal health. These guidelines should focus on transparency, fairness, accountability, and minimizing biases in prediction models.
4. Invest in the infrastructure needed to expand the Federated Learning system to more healthcare facilities, including providing secure and scalable server setups, technical support, and ensuring high-speed internet connectivity across facilities.
5. Create a certification process for scripts used to remove identifiers and handle missing data.
6. Provide training programs for healthcare professionals and data scientists in building, maintaining, and interpreting Federated Learning models. This should include technical workshops on PySyft, neural network development, and ethical AI practices.
7. Promote partnerships between government healthcare agencies, universities, and private tech companies to enhance Federated Learning architectures, particularly for healthcare applications.