

AI Based Diabetes Prediction System

Phase 2: Innovation

**JANANI P
AU711021104016**

AI Based Diabetes Prediction System

Phase 2: Innovation

INNOVATION

Data gathering

Make use of decentralised data gathering. Decentralised data collection involves gathering information from numerous sources, including patient portals, wearable technology, and electronic health records. As a result, it may be less demanding for participants and simpler to gather data from various populations.

Use fabricated data. Artificially produced data that resembles real data is referred to as synthetic data. Machine learning models can be trained using synthetic data instead of real data collection. When it is difficult or expensive to gather actual data, this can be helpful.

Make use of federated learning. A machine learning technique called federated learning enables researchers to train their models using data that is stored on participant devices. This can facilitate the collection of data from large populations and contribute to the privacy protection of the data.

Data preparation

Use artificial intelligence to find and fix data errors. AI can be used to find and fix data mistakes like typos and outliers. This may aid in enhancing the data's quality and increasing the precision of the machine learning models.

Create new features from the data using AI. With the aid of AI, new features that are more informative for estimating the risk of diabetes can be created from the data. AI could be used, for instance, to create features that depict the variability of the data over time.

Utilise AI to lessen data bias. Data bias can be found and reduced using artificial intelligence. This can aid in ensuring that everyone is treated fairly and equally under the system.

A decentralised data collection system that enables users to share their wearable device and electronic health record data could be created by researchers. Then, using this data, machine learning models could be trained to forecast the risk of diabetes.

In order to train machine learning models to predict the risk of diabetes in children, researchers could use synthetic data. By doing this, it would be unnecessary to gather actual data from kids, which could be challenging and expensive.

In order to train machine learning models to predict diabetes risk in a sizable population of diabetics, researchers could use federated learning. Due to the protection of participant privacy and ease of data collection from a large population, this would enable researchers to train the models using participant device data.

AI could be used by researchers to find and fix data errors like typos and outliers. This might aid in enhancing the data's quality and increasing the precision of the machine learning models.

AI could be used by researchers to extract new features from the data, like ones that reflect the variability of the data over time. These traits might provide more useful information when determining diabetes risk than more conventional characteristics like blood pressure and blood glucose levels.