# Deep Learning Based Paper

Contents

[Deep Learning Based Paper 1](#_Toc165306432)

[Prediction using Deep Learning: 2](#_Toc165306433)

[Available System: 4](#_Toc165306434)

[Procedure: 6](#_Toc165306435)

[Paper Summary 8](#_Toc165306436)

# Prediction using Deep Learning

Deep learning techniques, such as neural networks, can be utilized to predict diabetes or assess the risk of developing diabetes based on various factors, including:

**Genetic Markers:** DL models can analyze genetic data to identify individuals at higher risk for developing diabetes.

**Clinical Data**: DL models can analyze electronic health records, including demographic information, medical history, and laboratory test results, to predict the likelihood of developing diabetes or to assist in early diagnosis.

**Lifestyle Factors:** DL models can incorporate data on lifestyle factors such as diet, exercise habits, and weight changes to predict the risk of developing type 2 diabetes.

**Continuous Glucose Monitoring:** DL models can analyze data from continuous glucose monitoring devices to predict blood sugar levels and identify patterns indicative of diabetes or pre-diabetes.

By leveraging large datasets and advanced machine learning techniques, DL models have the potential to improve early detection, risk assessment, and personalized treatment strategies for diabetes. However, it's important to ensure the accuracy, reliability, and ethical considerations of these predictivemodels before deploying them in clinical practice.

Patients who have diabetes are at risk for developing a condition of the retina known as proliferative diabetic retinopathy (PDR). PDR may be identified by its

characteristic symptom, **neovascularization**, which describes a disease in which abnormal blood vessels form on the retina. If this issue is not detected and treated in a timely manner, it may result in complete and permanent loss of vision.



# Available System

1. **Convolutional Neural Networks (CNNs)**:
   * CNNs are well-suited for image-based data, but they can also be adapted for tabular data by treating it as an image with one dimension.
   * CNN architecture can be used to automatically learn hierarchical representations from the input data, capturing complex patterns and relationships.
2. **Artificial Neural Networks (ANN): (RNN)**
   * ANN stands for Artificial Neural Network. It's a computational model inspired by the structure and functioning of the human brain's neural networks. ANN is composed of interconnected nodes, known as neurons, organized in layers. These layers include an input layer, one or more hidden layers, and an output layer.
   * ANN can be used to handle large amounts of data are available and complex patterns need to be learned.
3. **Recurrent Neural Networks (RNNs)**:
   * RNNs are suitable for sequential data where the order of the input matters, such as time-series data or sequences of medical measurements.
   * Long Short-Term Memory (LSTM) or Gated Recurrent Unit (GRU) can be used to model temporal dependencies and make predictions based on past observations.
4. **Deep Feedforward Neural Networks**:
   * Also known as multi-layer perceptrons (MLPs), these networks consist of multiple layers of neurons interconnected with weighted edges.
   * MLPs can be used for tabular data by directly mapping input features to output predictions through hidden layers of neurons.
   * Different architectures can be experimented, activation functions, and regularization techniques to improve performance.
5. **Deep Autoencoders**:
   * Autoencoders are neural networks trained to reconstruct input data at the output layer, typically through a bottleneck layer with fewer neurons than the input layer.
   * By training an autoencoder on diabetes data, you can learn a compact representation of the input features, which can then be used for prediction tasks.
6. **Deep Reinforcement Learning**:
   * While less commonly used for healthcare applications like diabetes prediction, deep reinforcement learning (DRL) can be applied in scenarios where decisions are made sequentially over time.
   * DRL algorithms, such as Deep Q-Networks (DQN) or Proximal Policy Optimization (PPO), can learn optimal decision-making policies from interaction with the environment.
7. **Hybrid Architectures**:
   * Hybrid architectures that combine different types of neural networks, such as CNNs with RNNs or CNNs with MLPs.
   * Hybrid architectures can leverage the strengths of each component network and improve overall prediction performance.

# Procedure

1. **Data Collection and Preprocessing**:
   * Gather a dataset containing features relevant to diabetes prediction, such as demographics, medical history, lifestyle factors, etc.
   * Preprocess the dataset by handling missing values, normalizing features, and encoding categorical variables.
2. **Feature Selection**:
   * Use techniques like correlation analysis, feature importance, or domain knowledge to select the most relevant features for prediction.
3. **Deep Learning Model Development**:
   * Design and implement a deep learning model for diabetes prediction. This could be a neural network architecture such as a feedforward network, convolutional neural network (CNN), or recurrent neural network (RNN).
   * Experiment with different architectures, hyperparameters, and optimization techniques to improve model performance.
4. **Training and Evaluation**:
   * Split the dataset into training, validation, and test sets.
   * Train the deep learning model using the training data and validate it using the validation set.
   * Evaluate the model's performance on the test set using metrics such as accuracy, precision, recall, F1 score, etc.
5. **Skyline Query Integration**:
   * Incorporate skyline query into your prediction model to identify the most interesting and diverse set of predictions.
   * Define skyline attributes based on relevant criteria such as prediction confidence, uncertainty, or diversity of predictions.
   * Implement skyline query algorithms to efficiently retrieve skyline points from the model's prediction space.
6. **Integration and Deployment**:
   * Integrate the deep learning model with skyline query functionality into a web or mobile application.
   * Develop a user-friendly interface for users to input their data and receive predictions.
   * Deploy the application to a cloud platform or server for accessibility.
7. **Evaluation and Validation**:
   * Conduct thorough evaluation and validation of your integrated system.
   * Collect feedback from users and domain experts to assess the effectiveness and usability of the application.
8. **Documentation and Thesis Writing**:
   * Document the entire process, including data collection, preprocessing, model development, integration, and evaluation.
   * Write your thesis, detailing the problem statement, objectives, methodology, results, discussion, and conclusions.

# Paper Summary

## [An Efficient Deep Learning Technique for Diabetes Classification and Prediction Based on Indian Diabetes Dataset.](F://Academic/4th%20Year/7th%20Semester/Sessional/CSE%204000%20Project%20or%20Thesis/Paper/ML/Diabetes%20Prediction/An_Efficient_Deep_Learning_Technique_for_Diabetes_Classification_and_Prediction_Based_on_Indian_Diabetes_Dataset.pdf)

**Proposed system:** Diabetes prediction framework using artificial neural network (ANN) deep learning algorithms.

## [Diabetes Prediction Using Ensembling of Different Machine Learning Classifiers](F://Academic/4th%20Year/7th%20Semester/Sessional/CSE%204000%20Project%20or%20Thesis/Paper/Dola%20Mam/Diabetes%20Prediction%20Using%20Ensembling.pdf)

**Proposed System:** A robust framework for diabetes prediction where the outlier rejection, filling the missing values, data standardization, feature selection, K-fold cross-validation, and different Machine Learning (ML) classifiers (k-nearest Neighbour, Decision Trees, Random Forest, AdaBoost, Naive Bayes, and XGBoost) and Multilayer Perceptron (MLP) were employed. The weighted ensembling of different ML models is also proposed, to improve the prediction of diabetes where the weights are estimated from the corresponding Area Under ROC Curve (AUC) of the ML model.

**Help Link:**

1. [Hyperparameter Tuning - GridSearchCV and RandomizedSearchCV](https://www.youtube.com/watch?v=DTcfH5W6o08&list=PLfFghEzKVmjunyr8OPegxrX7y83IDuZNV&index=4)
2. [Hyperparameter tuning](https://www.geeksforgeeks.org/hyperparameter-tuning/)
3. [Github](https://github.com/kamruleee51/Diabetes-Prediction-Using-ML-Classifiers)