

Applying Prompt Engineering of Generative Artificial Intelligence to the Diagnosis of Diabetes

Frank Edughom Ekpar^{1*}

¹ Scholars University Ltd; Rivers State University; Topfaith University, Nigeria

Abstract

This paper presents a system that relies on artificial intelligence models with enough sophistication built by rigorously adhering to the recommendations of generative artificial intelligence tools such as large language models when prompted to design a system for the automated diagnosis of diabetes. Actual source code for the construction of the artificial intelligence models is generated as part of the suggestions recommended by the generative artificial intelligence tools or large language models. By faithfully incorporating the source code into a module for the automated diagnosis of diabetes based on clinical measurements, adequately sophisticated artificial intelligence models are constructed, trained, tested and validated on publicly accessible diabetes datasets and then deployed in an automated diabetes diagnosis module. Results indicate that the resulting artificial intelligence models exhibit reasonable performance that compares favorably (in view of the fact that the resulting artificial intelligence models are not optimized) with the performance of systems designed from the ground up by artificial intelligence experts.

Keywords: *Diabetes Mellitus, Generative Artificial Intelligence, ChatGPT, DeepSeek, Artificial Intelligence (AI), Large Language Model (LLM), TensorFlow, Artificial Neural Network (ANN), Deep Learning (DL), Healthcare System, Disease Diagnosis and Prediction*

1. Introduction

Diabetes mellitus afflicts millions of people around the world and is reported to exhibit increasing prevalence recently in low- and middle-income countries (LMICs), the regions with some of the most vulnerable populations characterized by some of the least equipped healthcare facilities and the lowest doctor-to-patient ratios [1]. This disease exerts a significant economic, physical and psychological toll on those affected. Given accurate diagnosis and prediction of likelihood of occurrence, the negative consequences can be significantly mitigated and more effective therapies and lifestyle changes can be implemented for improved health outcomes.

There is no shortage of reports in the literature on systems for the automated diagnosis of a wide variety of health conditions including diabetes

utilizing artificial intelligence (AI) algorithms and similar techniques [2] – [20]. These systems mainly cater to populations in developed countries since the data is primarily derived from sources within such environments without enrichment with local data from other environments such as LMICs, rendering the systems susceptible to bias and of restricted global relevance.

Large language models (LLMs) in particular, and generative AI generally, provide an alternative pathway to the automated diagnosis of diseases since they can leverage AI models trained on vast quantities of data to draw inferences and are capable of learning structured representations of the underlying data [21] – [22]. Ekpar [23] – [24] leveraged generative AI tools such as LLMs to the synthesis and development of AI models for

automated diagnosis of a variety of health conditions. In order to help resolve the perennial and critical issues with the healthcare system in LMICs and to improve medical doctor productivity generally and mitigate healthcare professional burnout and staff shortages both in developed and developing countries, Ekpar introduced Scholar Medic, a modular and extensible comprehensive artificial intelligence-driven healthcare system [25] – [28], [23] – [24]. Brain computer interfaces (BCIs) based on various modalities including electroencephalography (EEG) offer opportunities for rehabilitation and efficient channels of communication with the environment [29]–[47]. Ekpar introduced a novel three-dimensional multilayer electroencephalography paradigm (Ekpar EEG) that could enable vastly improved BCI performance and potentially unlock hitherto unattainable applications of brain waves in a wide variety of fields ranging from medicine to computing [48] – [50].

A previous study [27] involved the development of custom-built artificial neural networks into AI models for the automated diagnosis of diabetes within the context of a comprehensive artificial intelligence-driven healthcare system. This study utilizes recommendations derived from generative artificial intelligence tools such as large language models through prompt engineering for instructions on the design of a system for the automated diagnosis of diabetes based on clinical measurements.

2. Materials and Methods

Participant Recruitment

People volunteered to take part in the research aimed at developing the AI-driven healthcare system, and all participants gave their informed consent prior to joining the studies.

Ethical Approval

The Health Research Ethics Committee at Rivers State University Teaching Hospital provided ethical approval for the studies. All research followed the relevant ethical and regulatory guidelines. Publicly available data were utilized in

accordance with the licensing terms established by their creators.

3. Methodology

Healthcare datasets that are publicly accessible can be enhanced by incorporating data from local experiments and data collection efforts. This combined dataset can then be utilized to train AI models that make actionable predictions based on new data. Public sources of healthcare datasets include organizations like the Centers for Disease Control, the University of California Irvine Machine Learning Repository, the American Epilepsy Society, and Kaggle.

Adding local data improves model reliability, reduces biases, and promotes inclusivity and global relevance. A key focus of this project is merging diagnostic data, such as electrocardiographic readings, from local research with EEG data, including both traditional and advanced three-dimensional multilayer EEG systems (Ekpar EEG systems) [48] – [50].

Ethical approval for data collection has been granted by the relevant committees in the regions where experiments are being conducted. The project has also teamed up with licensed healthcare professionals who have direct access to patients and medical teams in the community, providing anonymized clinical data to help validate the AI models.

Once trained, the AI models will be integrated into a comprehensive healthcare system designed to assist clinical decision-making and develop brain-computer interfaces (BCIs). This system will provide actionable insights and predictions from new clinical data, helping with early detection, diagnosis, treatment, prediction, and prevention of conditions such as diabetes, heart disease, stroke, autism, and epilepsy.

The project is committed to advancing open science, reproducibility, and collaboration, ensuring that all generated data will be made publicly available on platforms like GitHub.

4. System Design and Implementation

The comprehensive healthcare system outlined in this paper employs a modular design, with each health condition (like diabetes mellitus, heart disease, stroke, epilepsy, autism, etc.) assigned to its own dedicated module. This structure not only enhances flexibility in diagnosing and predicting potential health issues, but also makes it easier to update modules as new data is introduced. Furthermore, modules related to Brain-Computer Interfaces (BCIs), particularly those using the motor imagery paradigm, can process EEG data to generate actionable commands and responses. The system also provides guidelines for upgrading conventional EEG setups to more advanced three-dimensional, multilayer EEG systems. These advanced systems, also known as Ekpar EEG systems, developed by Ekpar [48] - [50], are based on a conceptual framework that estimates specific bio-signal features to model or influence the biological systems involved.

For each module, AI models are developed and trained using properly formatted data, as detailed in this paper. These AI models are capable of incorporating various factors, such as genetic, environmental, and lifestyle information, to offer a more accurate representation of participants' circumstances.

Figure 1 depicts the schematic representation of the system with a subset of representative modules and their operational disposition within the system.

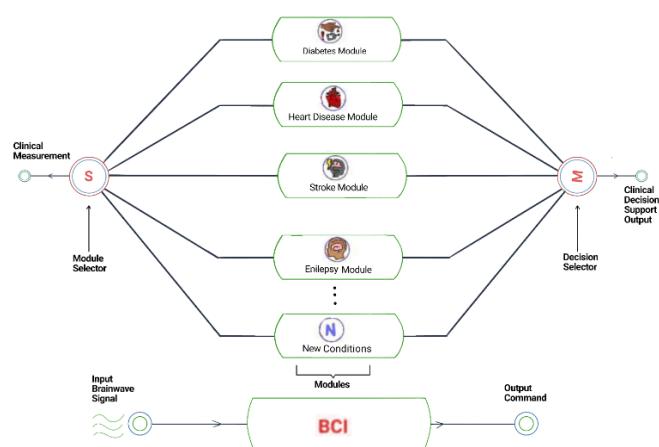


Fig. 1: System Schematic Design Diagram for the Comprehensive AI-Driven Healthcare Solution and Brain Computer Interface System. The New Conditions

component represents additional health conditions that can be incorporated into the solution via new modules.

The AI models to be incorporated into the comprehensive artificial intelligence-driven healthcare system under consideration are developed using four distinct approaches, as described below:

1. **Direct Use of LLMs:** Large Language Models (LLMs) like GPT-4 and DeepSeek are used as inference engines to process the collected data, which is formatted as multidimensional input vectors. This step may include fine-tuning the LLM.
2. **Prompt Engineering with LLMs:** LLMs, including DeepSeek, Bard and GPT-4 (and their future versions), are employed alongside prompt engineering to outline the necessary steps for AI system development. These steps are then executed by the developer, who applies expertise in AI, neural networks, deep learning, and tools such as Python, TensorFlow, Keras, and machine learning frameworks like Scikit-learn and Matplotlib.
3. **Automated Model Generation:** AI models are created through an automated pipeline that leverages the capabilities of LLMs such as Deep Seek, Bard and GPT-4 (and their future versions).
4. **Manual AI Architecture Design:** The AI architecture is manually designed by the developer, who relies on in-depth knowledge of AI, neural networks, deep learning, and programming tools like Python, TensorFlow, Keras, Scikit-learn, and Matplotlib.

Ultimately, all development processes and tools are thoroughly documented to ensure the solution can be easily transferred and reused.

The AI models are then evaluated and compared based on performance metrics, such as specificity and sensitivity, to assess their effectiveness in solving the intended problems.

Automated Diabetes Diagnosis Module

Prompt engineering of Chat GPT 4o mini – a widely used large language model (LLM) or prime example of a generative artificial intelligence tool – is leveraged to generate instructions for the design of a system for the automated diagnosis of diabetes based on clinical measurements.

This corresponds to the second approach on the list of approaches outlined in the foregoing.

An initial prompt that contains a generalized description of the task is entered and the peculiarities of the response of Chat GPT to the initial prompt are harnessed to generate a follow-up prompt that drills down into a specific class of artificial intelligence system architectures that are then implemented and trained, validated and tested on publicly accessible diabetes datasets to synthesize artificial intelligence models for the automated diagnosis of diabetes based on a set of clinical measurements.

The follow-up prompt is designed to account for the structure of the dataset in terms of input features and output vector characteristics.

Dataset

The Pima Indians Diabetes database was harnessed for the training, testing and validation of the AI models implemented in the course of this study. A total of 768 rows of data were utilized. Each data row contained 9 columns where the first 8 columns represented the clinical measurements for the patient or participant indicated by the selected row, namely, diastolic blood pressure, plasma glucose concentration 2 hours in an oral glucose tolerance test, number of times pregnant, triceps skin fold thickness, 2-Hour serum insulin, body mass index, diabetes pedigree function and age. The last or ninth column indicated the diagnosis (a binary entity) where a value of 0 represented a normal outcome while a value of 1 represented the presence of diabetes mellitus.

Data Availability

Public access to the data supporting the findings of this study can be gained via GitHub at https://github.com/frankepar/pima_indians_diabetes_dataset/blob/main/dataset.zip. The dataset was derived from the Pima Indians Diabetes Database publicly accessible from Kaggle at <https://www.kaggle.com/datasets/uciml/pima-indians-diabetes-database>.

Initial Chatgpt Prompt

Design a system for automated diagnosis of diabetes mellitus based on clinical measurements such as diastolic blood pressure, plasma glucose concentration 2 hours in an oral glucose tolerance test, triceps skin fold thickness, 2-Hour serum insulin and body mass index.

Initial Chatgpt Response

Designing an automated system for the diagnosis of diabetes mellitus using clinical measurements can be approached through a machine learning-based model. Here's an outline of the key steps and components for such a system:

1. Data Collection:

- Diastolic Blood Pressure (DBP): The lower value of blood pressure when the heart rests between beats.
- Plasma Glucose Concentration (2-Hour Oral Glucose Tolerance Test): Blood glucose levels measured two hours after ingesting a glucose solution. Elevated levels can indicate insulin resistance or diabetes.
- Triceps Skin Fold Thickness: An indicator of body fat percentage, measured using calipers.
- 2-Hour Serum Insulin: Insulin levels measured two hours after consuming glucose, reflecting insulin resistance or insufficient insulin production.
- Body Mass Index (BMI): A measure of body fat based on height and weight.

These variables can be used to create a diagnostic model to classify individuals as having normal glucose tolerance, prediabetes, or diabetes.

2. Preprocessing and Feature Engineering:

- Normalization/Scaling: Each feature (e.g., glucose concentration, BMI) should be normalized to ensure equal importance across features, especially when using distance-based algorithms like KNN or support vector machines.
- Handling Missing Data: Use imputation techniques or discard incomplete entries, depending on the dataset's nature and missingness.
- Feature Selection/Engineering: Evaluate feature importance and remove irrelevant or redundant features.

3. Model Selection:

Machine learning models can be trained to predict the likelihood of diabetes based on the clinical measurements:

a. Logistic Regression:

A simple and interpretable model that can predict probabilities of a class (diabetes or not) based on the clinical measurements.

b. Random Forest / Decision Trees:

These models are more complex and can capture nonlinear relationships between features. They also offer feature importance analysis.

c. Support Vector Machine (SVM):

SVM is effective for binary classification tasks and works well with high-dimensional data, making it a good choice for this system.

d. Neural Networks:

For more complex models and large datasets, neural networks can be employed to learn intricate patterns in the data.

e. K-Nearest Neighbors (KNN):

A non-parametric model that works well if you have a small dataset. It predicts the class of a new sample based on the majority class of its neighbors.

4. Model Training:

- Train-Test Split: Split your data into training and testing sets (commonly 80%/20%).
- Cross-Validation: Use cross-validation to evaluate the model's performance and prevent overfitting, especially when you have a small dataset.
- Hyperparameter Tuning: For models like SVM and Random Forest, perform hyperparameter tuning (e.g., grid search) to find optimal parameters.

5. Evaluation Metrics:

- Accuracy: Percentage of correct classifications.
- Precision, Recall, and F1-Score: Given the potential imbalance between classes (e.g., more people without diabetes), these metrics provide deeper insights into performance.
- Area Under ROC Curve (AUC-ROC): Evaluates how well the model differentiates between the classes.

6. Prediction:

Once the model is trained and validated, it can be used to predict the likelihood of diabetes in new individuals based on their clinical measurements.

7. User Interface and Integration:

- Web/Mobile Interface: Create an easy-to-use front end where a user can input the clinical measurements, and the model returns the prediction (normal, prediabetes, or diabetes).
- Real-time Diagnosis: Use the model to provide real-time predictions and visualize the likelihood of diabetes.
- Alerts/Recommendations: Based on the prediction, the system could also provide

lifestyle or medical recommendations, such as dietary changes, exercise, or further testing.

8. Ethical Considerations:

- **Bias:** Ensure the model does not unfairly categorize certain populations (e.g., race, gender) based on skewed data.
- **Data Privacy:** Ensure the system complies with data privacy laws (e.g., HIPAA in the U.S.) when handling personal health information.
- **Model Transparency:** Given the importance of medical decisions, ensure the model's decisions are interpretable, with an explanation of how predictions are made.

Example: Flow of System Interaction

1. **User Input:** The user inputs their diastolic blood pressure, plasma glucose concentration, triceps skin fold thickness, serum insulin, and BMI into the system.
2. **Data Preprocessing:** The system normalizes and scales the input values.
3. **Prediction:** The trained machine learning model predicts whether the individual is likely to have normal glucose levels, prediabetes, or diabetes.
4. **Recommendation:** Based on the prediction, the system gives advice on next steps (e.g., consult a healthcare provider, follow up tests).

5. Conclusion:

This system will use clinical data to provide a rapid and automated diagnosis of diabetes mellitus, allowing for early detection and intervention. Integrating a user-friendly interface with machine learning models will make this a valuable tool in healthcare systems, especially in regions where access to specialized care is limited.

Follow-Up Chatgpt Prompt

Design an artificial neural network for binary classification of a set of 8 inputs.

FOLLOW-UP ChatGPT RESPONSE

To design an artificial neural network (ANN) for binary classification with 8 input features, you would generally follow these steps:

1. Input Layer:

- **Number of neurons:** Since there are 8 input features, the input layer will have 8 neurons (one for each feature).
- **Activation function:** No activation function is typically used in the input layer. The raw input values are fed into the network.

2. Hidden Layers:

The number of hidden layers and neurons per layer can vary, but let's assume a basic architecture to start with:

- **Layer 1 (First hidden layer):**
 - **Number of neurons:** 16 (you could start with more or fewer depending on the complexity of the task).
 - **Activation function:** ReLU (Rectified Linear Unit), which helps avoid the vanishing gradient problem.
- **Layer 2 (Second hidden layer):**
 - **Number of neurons:** 8 (this layer typically has fewer neurons as you get deeper).
 - **Activation function:** ReLU.
- You can add more hidden layers as needed depending on the complexity of the problem, but for simplicity, we'll use two hidden layers here.

3. Output Layer:

- **Number of neurons:** 1 (since it's binary classification, the output should be a single value representing the probability of the positive class).
- **Activation function:** Sigmoid, which outputs a value between 0 and 1, representing the probability that the input belongs to the positive class (e.g., class 1).

4. Loss Function:

- Binary Cross-Entropy Loss: This is typically used for binary classification problems because it measures the difference between the predicted probability and the actual class label.

5. Optimizer:

- Adam Optimizer: This is a popular optimizer due to its adaptive learning rates for each parameter, often yielding better results in practice.

6. Model Summary:

Here's a summary of the architecture:

- Input layer: 8 neurons (no activation function).
- Hidden Layer 1: 16 neurons, ReLU activation.
- Hidden Layer 2: 8 neurons, ReLU activation.
- Output layer: 1 neuron, Sigmoid activation.
- Loss function: Binary Cross-Entropy.
- Optimizer: Adam.

7. Code Implementation (Using Python and TensorFlow/Keras):

```
import tensorflow as tf

from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense

# Define the model
model = Sequential()

# Input layer (implicitly included when you define
input_shape in the first hidden layer)

model.add(Dense(16, input_dim=8,
activation='relu')) # First hidden layer with 16
neurons

model.add(Dense(8, activation='relu')) # Second
hidden layer with 8 neurons
```

```
model.add(Dense(1, activation='sigmoid'))
# Output layer with 1 neuron for binary
classification

# Compile the model

model.compile(optimizer='adam',

              loss='binary_crossentropy',

              metrics=['accuracy'])
```

Summary of the model

```
model.summary()
```

8. Model Training:

To train this model, you would provide it with your dataset, where X_{train} is a matrix of shape (num_samples, 8) and y_{train} is the binary target vector of shape (num_samples,). You can fit the model like this:

Assuming X_{train} and y_{train} are your training data

```
model.fit(X_train, y_train, epochs=10,
batch_size=32)
```

This simple architecture can be extended with additional hidden layers, dropout for regularization, or other modifications based on the complexity of your dataset.

Figure 2 features a generalized graphical depiction of the recommended artificial neural network. In Fig. 2, CM1, CM2, CMN are the relevant clinical measurements (N=8) numbering 8 in this case while CD represents the clinical decision corresponding to the output of the artificial neural network.

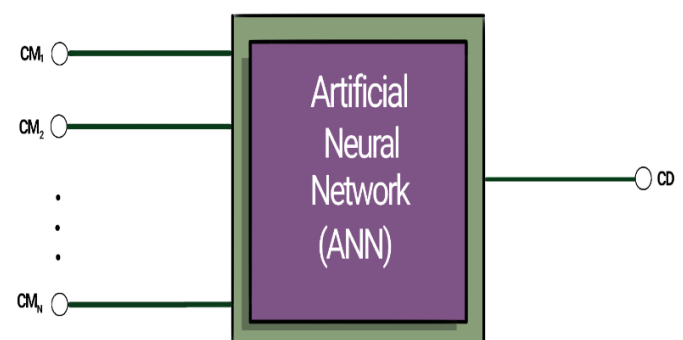


Fig. 2: Schematic Graphical Representation of Artificial Neural Network (ANN) Architecture. CM1, CM2, ..., CMN represent the inputs while CD represents the output indicating the suggested clinical diagnosis.

6. Results

Attention has been given to matters related to system design such as ethics and data management in the design of the comprehensive artificial intelligence-driven healthcare system [25] – [28], [23] – [24].

Combining the TensorFlow framework with the Keras Application Programming Interface (API), the suggested AI models were implemented in the Python programming language [51] – [52].

First, the dataset is shuffled randomly to minimize bias and then partitioned by applying an 80% - 20% split for training and validation datasets, respectively. Optimization was accomplished via the Adam Optimizer [53] – [54] while the AI model was trained using the binary cross-entropy loss function, a learning rate of 0.001 and a batch size of 32 over 300 epochs.

The performance metrics of the trained AI model was as follows: Specificity – 82%, Sensitivity – 63%, Precision – 67%. Although the observed performance metrics were not on par with those obtained from custom-designed AI models [27], the performance could be further improved through refinement or finetuning on the basis of AI expertise.

Figure 3 illustrates a screenshot from the Diabetes Diagnosis Module in the Scholar Medic featuring a selection of clinical measurements the associated suggested diagnosis from the trained AI model.

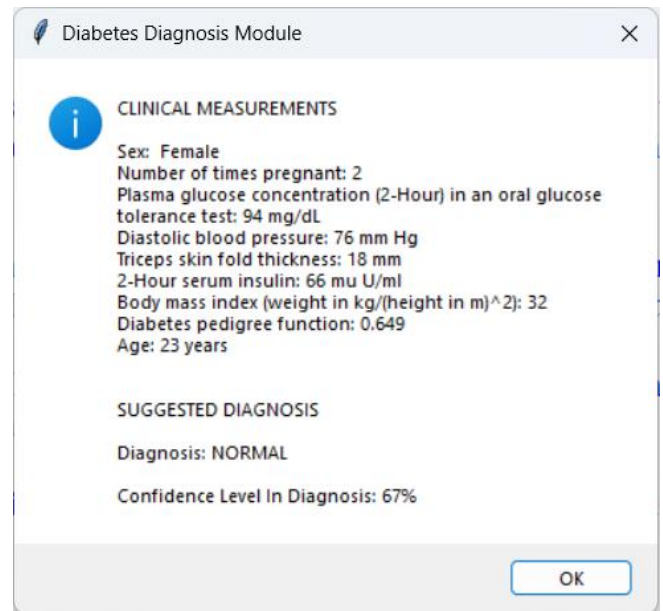


Fig. 3: Diabetes Diagnosis Module of Scholar Medic Showing Clinical Measurements and Corresponding Suggested Diagnosis.

The performance metrics were computed on the basis of the following equations where TN refers to true positives, FP refers to false positives, FN refers to false negatives, and TN refers to true negatives. Negative refers to the normal diabetes-free condition while positive refers to the presence of diabetes mellitus.

$$Precision = \frac{TP}{TP + FP}$$

$$Sensitivity = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{TN + FP}$$

Implementing the comprehensive AI system outlined here will provide actionable insights to support clinical decision-making, ultimately saving lives and improving quality of life. It will help reduce the economic, social, psychological, and physical burdens associated with conditions that could be predicted, prevented, detected early, diagnosed, treated, and managed more effectively.

Electronic Health Records (EHRs), which include clinical diagnostic data and EEG information, will be created by participating doctors and their associated teams. EEG data may also be collected through experiments involving Brain-Computer

Interfaces (BCIs). These data will be gathered following ethical approval, anonymized, and shared in publicly accessible repositories alongside relevant research publications.

7. Conclusion

This paper presented a system that achieved automated diagnosis of diabetes on the basis of clinical measurements by leveraging artificial intelligence models constructed, trained, tested and validated on public diabetes datasets wherein the artificial intelligence models were synthesized from the recommendations of generative artificial intelligence tools or large language models responding to well-crafted prompts. The performance of the system compared favorably (in light of the fact that the resulting artificial intelligence models are not optimized) with that of artificial intelligence models developed by experts. Consequently, the system could be further refined for incorporation into a modular, extensible comprehensive artificial intelligence-driven healthcare system.

Conflicts of Interest

There are no conflicts of interest to disclose.

References

1. World Health Organization (WHO) Diabetes Statistics: <https://www.who.int/news-room/fact-sheets/detail/diabetes>. (Retrieved 2025).
2. Nomura, A., Noguchi, M., Kometani, M., Furukawa, K., Yoneda, T. Artificial Intelligence in Current Diabetes Management and Prediction, *Curr Diab Rep*. 21(12):61 (2021).
3. Kumar, Y., Koul, A., Singla, R., Ijaz, M. F. Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework and future research agenda, *Journal of Ambient Intelligence and Humanized Computing* 14:8459–8486 (2023).
4. Ansari, S., Shafi, I., Ansari, A., Ahmad, J., Shah, S. I. Diagnosis of liver disease induced by hepatitis virus using artificial neural network, *IEEE Int Multitopic*. <https://doi.org/10.1109/INMIC.2011.6151515> (2011).
5. Battineni, G., Sagaro, G. G., Chinatalapudi, N., Amenta, F. Applications of machine learning predictive models in the chronic disease diagnosis, *J Personal Med*. <https://doi.org/10.3390/jpm10020021> (2020).
6. Abdar, M., Yen, N., Hung, J. Improving the diagnosis of liver disease using multilayer perceptron neural network and boosted decision tree, *J Med Biol Eng* 38:953–965 (2018).
7. Chaikijurajai, T., Laffin, L., Tang, W. Artificial intelligence and hypertension: recent advances and future outlook, *Am J Hypertens* 33:967–974 (2020).
8. Fujita, S., Hagiwara, A., Otsuka, Y., Hori, M., Takei, N., Hwang, K. P., Irie, R., Andica, C., Kamagata, K., Akashi, T., Kumamaru, K. K., Suzuki, M., Wada, A., Abe, O., Aoki, S. Deep Learning Approach for Generating MRA Images From 3D Quantitative Synthetic MRI Without Additional Scans, *Invest Radiol* 55:249–256 (2020).
9. Juarez-Chambi, R. M., Kut, C., Rico-Jimenez, J. J., Chaichana, L. K., Xi, J., Campos-Delgado, D. U., Rodriguez, F. J., Quinones-Hinojosa, A., Li, X., Jo, J. A. AI-Assisted In Situ Detection of Human Glioma Infiltration Using a Novel Computational Method for Optical Coherence Tomography, *Clin Cancer Res* 25(21):6329–6338 (2019).
10. Nashif, S., Raihan, R., Islam, R., Imam, M. H. Heart Disease Detection by Using Machine Learning Algorithms and a Real-Time Cardiovascular Health Monitoring System, *World Journal of Engineering and Technology* Vol 6, No. 4 (2018).
11. Chen, P. H. C., Gadepalli, K., MacDonald, R., Liu, Kadowaki, S., Nagpal, K., Kohlberger, T., Dean, J., Corrado, G. S.,

- Hipp, J. D., Mermel, C. H., Stumpe, M. C. An augmented reality microscope with real time artificial intelligence integration for cancer diagnosis, *Nat Med* 25:1453–1457 (2019).
12. Gouda, W., Yasin, R. COVID-19 disease: CT Pneumonia Analysis prototype by using artificial intelligence, predicting the disease severity, *Egypt J Radiol Nucl Med* 51(1):196 (2020).
13. Han, Y., Han, Z., Wu, J., Yu, Y., Gao, S., Hua, D., Yang, A. Artificial Intelligence Recommendation System of Cancer Rehabilitation Scheme Based on IoT Technology, *IEEE Access* 8:44924–44935 (2020).
14. Chui, C. S., Lee, N. P., Adeoye, J., Thomson, P., Choi, S. W. Machine learning and treatment outcome prediction for oral cancer, *J Oral Pathol Med* 49(10):977–985 (2020).
15. Koshimizu, H., Kojima, R., Okuno, Y. Future possibilities for artificial intelligence in the practical management of hypertension, *Hypertens Res* 43:1327–1337 (2020).
16. Kather, J. N., Pearson, A. T., Halama, N., Jäger, D., Krause, J., Loosen, S. H., Marx, A., Boor, P., Tacke, F., Neumann, U. P., Grabsch, H. I., Yoshikawa, T., Brenner, H., Chang-Claude, J., Hoffmeister, M., Trautwein, C., Luedde, T. Deep learning microsatellite instability directly from histology in gastrointestinal cancer, *Nat Med* 25:1054–1056 (2019).
17. Kwon, J. M., Jeon, K. H., Kim, H. M., Kim, M. J., Lim, S. M., Kim, K. H., Song, P. S., Park, J., Choi, R. K., Oh, B. H. Comparing the performance of artificial intelligence and conventional diagnosis criteria for detecting left ventricular hypertrophy using electrocardiography, *EP Europace* 22(3):412–419 (2020).
18. Khan, M. A. An IoT Framework for Heart Disease Prediction Based on MDCNN Classifier, *IEEE Access* 8:34717–34727 (2020).
19. Oikonomou, E. K., Williams, M. C., Kotanidis, C. P., Desai, M. Y., Marwan, M., Antonopoulos, A. S., Thomas, K. E., Thomas, S., Akoumianakis, I., Fan, L. M., Kesavan, S., Herdman, L., Alashi, A., Centeno, E. H., Lyasheva, M., Griffin, B. P., Flamm, S. D., Shirodaria, C. Sabharwal, N., Kelion, A., Dweck, M. R., Van Beek, E. J. R., Deanfield, J., Hopewell, J. C., Neubauer, S., Channon, K. M., Achenbach, S., Newby, D. E., Antoniades, C. A novel machine learning-derived radiotranscriptomic signature of perivascular fat improves cardiac risk prediction using coronary CT angiography, *Eur Heart J* 40(43):3529–3543 (2019).
20. Sabottke, C. F., Spieler, B. M. The Effect of Image Resolution on Deep Learning in Radiography, *Radiology: Artificial Intelligence* Vol. 2. No. 1, 2:e190015 (2020).
21. Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I., Amodei, D. Language Models are Few-Shot Learners, *Advances in Neural Information Processing Systems* 33, (2020).
22. Gurnee, W., Tegmark, M. Language Models Represent Space and Time, *arXiv*, DOI: <https://doi.org/10.48550/arXiv.2310.02207> (2023).
23. Ekpar, F. E. Leveraging Generative Artificial Intelligence Recommendations for Image-based Chronic Kidney Disease Diagnosis, *International Journal of Advanced Research in Computer and*

- Communication Engineering, 14(1). (2025).
24. Ekpar, F. E. Generative Artificial Intelligence Recommendations for Clinical Measurement-Based Diagnosis of Chronic Kidney Disease, *International Journal of Engineering And Computer Science*, 14(1). (2025).
 25. Ekpar, F. E. A Comprehensive Artificial Intelligence-Driven Healthcare System, *European Journal of Electrical Engineering and Computer Science*, 8(3), Article 617. (2024).
 26. Ekpar, F. E. Image-based Chronic Disease Diagnosis Using 2D Convolutional Neural Networks in the Context of a Comprehensive Artificial Intelligence-Driven Healthcare System, *WSEAS Molecular Sciences and Applications*, 4(13). (2024).
 27. Ekpar, F. E. Diagnosis of Diabetes Within a Comprehensive Artificial Intelligence-Driven Healthcare System, *International Journal of Advanced Research in Computer and Communication Engineering*, 13(8). (2024).
 28. Ekpar, F. E. Diagnosis of Chronic Kidney Disease Within a Comprehensive Artificial Intelligence-Driven Healthcare System, *International Journal of Advanced Research in Computer and Communication Engineering*, 13(9). (2024).
 29. Padfield, N., Zabalza, J., Zhao, H, Masero, V., Ren, J. EEG-Based Brain-Computer Interfaces Using Motor-Imagery: Techniques and Challenges, *Sensors* 19, 1423. (2019).
 30. Kevric, J., Subasi, A. Comparison of signal decomposition methods in classification of EEG signals for motor-imagery BCI system, *Biomedical Signal Processing and Control* 31, 398-406. (2017).
 31. Cho, H., Ahn, M., Ahn, S., Kwon, M., Jun, S. C. EEG datasets for motor imagery brain-computer interface, *GigaScience* Vol. 6, Iss. 7. (2017).
 32. Arpaia, P., Esposito, A., Natalizio, A., Parvis, M. How to successfully classify EEG in motor imagery BCI: a metrological analysis of the state of the art, *J. Neural Eng.* 19 031002. (2022).
 33. Kaya, M., Binli, M, K., Ozbay, E., Yanar, H., Mishchenko, Y. A large electroencephalographic motor imagery dataset for electroencephalographic brain computer interfaces, *Scientific Data* 5, Article number: 180211 (2018).
 34. Tibrewal, N., Leeuwis, N., Alimardani, M. Classification of motor imagery EEG using deep learning increases performance in inefficient BCI users, *PLoS One* 17(7): e0268880. (2022).
 35. Sreeja, S. R., Rabha, J., Nagarjuna, K. Y., Samanta, D., Mitra, P., Sarma, M. Motor Imagery EEG Signal Processing and Classification Using Machine Learning Approach, *IEEE International Conference on New Trends in Computing Sciences (ICTCS)* (2017).
 36. Das, K., Pachori, R. B. Electroencephalogram-Based Motor Imagery Brain-Computer Interface Using Multivariate Iterative Filtering and Spatial Filtering, *IEEE Transactions on Cognitive and Developmental Systems*, Vol. 15, Iss. 3. (2022).
 37. Velasco, I., Sipols, A., Simon De Blas, C., Pastor, L., Bayona, S. Motor imagery EEG signal classification with a multivariate time series approach, *Biomedical Engineering Online* 22, Article Number: 29. (2023).
 38. Yang, A., Lam, H. K., Ling, S. H. Multi-classification for EEG motor imagery signals using data evaluation-based auto-selected regularized FBCSP and convolutional neural network, *Neural Computing and Applications*, Vol.35, 12001-12027. (2023).

39. Venkatachalam, K., Devipriya, A., Maniraj, J., Sivaram, M., Ambikapathy, A., Amiri, I. S. A novel method of motor imagery classification using eeg signal, *Artificial Intelligence in Medicine*, Vol.103, 101787. (2020).
40. Subasi, A., Gursov, M. I. EEG signal classification using PCA, ICA, LDA and support vector machines, *Expert Systems with Applications*, Vol.37, Iss. 12, 8659-8666. (2010).
41. Razzak, I., Hameed, I. A., Xu, G. Robust Sparse Representation and Multiclass Support Matrix Machines for the Classification of Motor Imagery EEG Signals, *IEEE Journal of Translational Engineering in Health and Medicine*, 7:2000508. (2019).
42. Pahuja, S. K., Veer, K. Recent Approaches on Classification and Feature Extraction of EEG Signal: A Review, *Robotica*, Vol. 40, Iss. 1, 77-101. (2022).
43. Lekshmi, S. S., Selvam, V., Rajasekaran, M. P. EEG signal classification using Principal Component Analysis and Wavelet Transform with Neural Network, *IEEE International Conference on Communication and Signal Processing*, (2014).
44. Lugger, K., Flotzinger, D., Schlögl, A., Pergenzer, M., Pfurtscheller, G. Feature extraction for on-line EEG classification using principal components and linear discriminants, *Med Biol Eng Comp*, 36(3):309-14. (1998).
45. Lawhern, V. J., Solon, A. J., Waytowich, N. R., Gordon, S. M., Hung, C. P., Lance, B. J. EEGNet: a compact convolutional neural network for EEG-based brain-computer interfaces, *Journal of Neural Engineering*, 15: 056013. (2018).
46. Al-Saegh, A., Dawwd, S. A., Abdul-Jabbar, J. M. Deep learning for motor imagery EEG-based classification: A review, *Biomedical Signal Processing and Control* Vol. 63, 102172. (2021).
47. Tabar, Y. R., Halici, U. A novel deep learning approach for classification of EEG motor imagery signals, *Journal of Neural Engineering*, 14: 016003. (2016).
48. Ekpar, F. E. A Novel Three-dimensional Multilayer Electroencephalography Paradigm, *Fortune Journal of Health Sciences*, 7(3). (2024).
49. Ekpar, F. E. System for Nature-Inspired Signal Processing: Principles and Practice, *European Journal of Electrical Engineering and Computer Science*, 3(6), pp. 1-10, (2019).
50. Ekpar, F. E. Nature-inspired Signal Processing, United States Patent and Trademark Office, US Patent Application Number: 13/674,035 (Filed: November 11, 2012, Priority Date: December 24, 2011), Document ID: US 20140135642 A1: Published: (2014).
51. Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D. G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., Wicke, M., Yu, Y., Zheng, X. TensorFlow: A System for Large Scale Machine Learning, *Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI '16)*. (2016).
52. Pang, B., Nijkamp, E., Wu, Y. N. Deep Learning With TensorFlow: A Review, *Journal of Educational and Behavioral Statistics*. Vol. 45, Iss. 2. (2019).
53. Kingma, D. P., Ba, J. L. Adam: A Method for Stochastic Optimization, *International Conference on Learning Representations (ICLR)* (2015).
54. Zhang, Z. Improved Adam Optimizer for Deep Neural Networks, *IEEE/ACM 26th International Symposium on Quality of Service (IWQoS)* (2018).