International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 03 , March - 2026

AI-Based Diabetes Severity Prediction Web Application

Mrs. Sandhiya C
Assistant Professor
Department of Computer Science and Engineering
Nehru Institute of Engineering and Technology
Coimbatore – 641105

Suresh N, PraveenKumar T,
Vinesh KV, Vishal M
Department of Computer Science and Engineering
Nehru Institute of Engineering and Technology
Coimbatore – 641105

Abstract - Diabetes mellitus is a chronic metabolic disorder in which the level of blood glucose is high due to inadequate insulin production or improper use of insulin. It has turned out to be one of the most serious challenges facing the world today, affecting millions of people. The early detection and proper classification of diabetes is very important in order to avoid serious health problems, which may arise in the future. The conventional systems used in the detection of diabetes are not only time-consuming and costly, but also only provide binary classification, which can only identify whether a person is diabetic or not.

In recent times, the integration of Artificial Intelligence (AI) and Machine Learning (ML) in the field of healthcare has led to the creation of smart systems in the detection of diseases. Machine learning has been found to be very effective in the prediction of diseases, as it can identify hidden patterns in the data. Despite the creation of smart systems in the prediction of diabetes, many of the existing systems only provide binary classification, which does not provide in-depth information on the severity of the disease.

The proposed system in this research aims to develop an AI-Based Diabetes Severity Prediction Web Application that can classify diabetes into different stages using a multi-class classification algorithm. The proposed system uses a Support Vector Machine (SVM) algorithm with a Radial Basis Function (RBF) kernel to perform classification. Moreover, it uses different feature selection methods like Elephant Herding Optimization (EHO) to improve the accuracy and efficiency of the model.

The proposed system uses different medical parameters like glucose level, blood pressure, BMI, insulin levels, and age to perform classification. The parameters are retrieved using a web interface. The data is then normalized and scaled to perform classification using the proposed model. The proposed model can classify diabetes into different classes like Normal, Pre-Diabetes, Type 2 Diabetic, and Severe Diabetic.

The proposed system provides different advantages to users. The proposed system can provide better accuracy in prediction, can perform classification in real-time, and can provide an interface to users that is easy to use. Moreover, the proposed system can be accessible to users via web platforms to provide better health care to patients suffering from diabetes.

INTRODUCTION

On the other hand, diabetes mellitus is a chronic metabolic disorder in which the body fails to maintain normal blood glucose levels. The main cause of diabetes mellitus is the lack of insulin in the body or the inability of the body to respond to the existing insulin. Over the last several decades, diabetes has become a critical global health issue, affecting millions of people from different age groups and geographical regions. The rise in the number of diabetic patients is mainly due to lifestyle changes, lack of exercise, improper eating habits, obesity, and genetic disorders.

If diabetes is not diagnosed and controlled in its initial stages, it can cause serious health problems, including heart, kidney, and brain disorders, as well as eye problems. Therefore, it is crucial to control and manage diabetes in its initial stages in order to avoid serious health problems. The conventional approach to diagnosing diabetes involves laboratory tests, including fasting blood sugar, oral glucose tolerance, and HbA1 tests. Although the conventional approach is effective in diagnosing diabetes, it is a time-consuming, costly, and supervised approach. Moreover, it does not offer real-time monitoring, which is critical in managing chronic diseases.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 03 , March - 2026

In addition, with the advancements in Artificial Intelligence (AI) and Machine Learning (ML), healthcare is undergoing tremendous transformation. Machine learning algorithms can analyze large amounts of data and predict patterns to predict the disease at an early stage. Many machine learning algorithms have been developed to predict diabetes. However, these algorithms mostly rely on binary classification.

However, binary classification only provides limited information on the problem. This classification does not consider the severity of the problem. Diabetes occurs in different stages. It starts from a normal state and proceeds to pre-diabetic and severe diabetic states. Identifying these states is crucial in applying the correct intervention.

This project aims to design an AI-Based Diabetes Severity Prediction Web Application. This application uses multiclass classification to predict the severity of diabetes. This project aims to bring machine learning and web development together to make predictions in real time.

PROBLEM STATEMENT

Despite the availability of different prediction systems for diabetes, some challenges are still to be addressed in diagnosing and monitoring diabetes in an accurate manner. The most important drawback of existing systems is that they are based on binary classification methods, which can only diagnose whether a person is diabetic or not. However, they are not able to provide information regarding the progression of the disease.

The second most important drawback is that no accessible and real-time healthcare system is available to diagnose diabetes. In existing systems, an individual needs to consult a doctor and undergo different tests to diagnose diabetes. However, in some cases, it is difficult for individuals to reach healthcare centers due to their geographical locations. Moreover, different diagnostic methods are costly and time-consuming.

Most existing machine learning models are not able to provide accurate results due to the presence of irrelevant features in the system. In machine learning models, it is very important to select appropriate features to avoid degradation in the accuracy of the model. Moreover, most existing systems are not able to integrate with user-friendly web applications.

There is also a lack of systems that provide personalized health recommendations based on prediction results. Simply predicting the presence of diabetes is not sufficient; users need actionable insights to manage their health effectively.

Therefore, there is a need for an intelligent system that can classify diabetes into multiple severity levels, provide real-time predictions, and offer health recommendations through an accessible web platform. The proposed system aims to address these challenges by combining machine learning techniques with web-based technologies.

. LITERATURE SURVEY

In recent years, several research studies have been carried out to utilize machine learning algorithms in the prediction and diagnosis of diabetes. In this regard, several algorithms like Decision Trees, Random Forest, Logistic Regression, Naive Bayes, and SVM are used in developing prediction systems for diabetes.

Decision Tree algorithms are simple and easy to understand. They can be used in simple classification problems. However, they are not suitable for complex problems because they are more prone to overfitting. Moreover, they are not suitable for complex data.

Random Forest is an ensemble algorithm that can be used in prediction systems. It can provide better accuracy in prediction systems because it combines multiple Decision Trees. However, it can make the system more complex.

Logistic Regression can be used in binary classification problems. It can be used in developing prediction systems for diabetes. However, it can provide better accuracy in simple data sets. Naive Bayes is another algorithm that can be used in developing prediction systems. However, it assumes that all features are independent in a data set.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 03 , March - 2026

The Support Vector Machine (SVM) has received considerable attention because of its higher accuracy and efficiency to handle higher dimensions. SVM uses various kernel methods like Radial Basis Function (RBF) to handle non-linear relationships, which makes SVM more efficient for handling medical diagnosis-related applications.

Researchers have been trying to incorporate feature selection techniques to achieve higher efficiency. Various feature optimization techniques, such as Genetic Algorithms, Particle Swarm Optimization, and Elephant Herding Optimization, have been employed to select features that are more efficient for decision-making. These techniques have been employed to reduce noise as well as to achieve higher accuracy.

Although various systems have been proposed to handle diabetes prediction using machine learning techniques, most of them have been limited to handling only binary classification. Very few have been proposed to handle the need for multiclass classification to detect various stages of diabetes. The proposed system has been designed to handle this need by using a multiclass SVM approach along with feature optimization techniques.

PROPOSED SYSTEM

The proposed system is an AI-Based Diabetes Severity Prediction Web Application that can give an accurate and real-time prediction of the severity of diabetes. The system combines machine learning technology with a web interface to ensure ease of use.

The system can take input parameters such as glucose levels, blood pressure, BMI, insulin levels, and age from users via a user-friendly web interface. These parameters are important in diagnosing diabetes and determining its severity.

After inputting the parameters, they are sent to the server where preprocessing occurs. During this step, normalization, scaling, and handling missing values are done to ensure that the input parameters are appropriate for prediction.

The most important part of the system is the machine learning model. In this system, a multiclass SVM model is used. An optimized dataset is used to train this model. In addition, the RBF kernel is used to handle nonlinear relationships in data. Elephant Herding Optimization (EHO), which is a feature selection algorithm, can be used to optimize the model.

The system will classify the input data into four categories: Normal, Pre-Diabetes, Type 2 Diabetes, and Severe Diabetes. This is a more accurate classification method compared to the traditional classification method of using a binary classification system.

The prediction result will be shown to the user along with some health recommendations. The system may also store the user's information for future analysis. The advantages of the proposed system are that it is accurate, uses real-time prediction, is user-friendly, and is scalable.

METHODOLOY

The methodology of the proposed system is designed to provide accurate prediction of the severity of diabetes using a systematic process. The first step is to collect the data from the PIMA Indians diabetes dataset or user input through the web application. The dataset contains various attributes such as glucose level, blood pressure, BMI, insulin level, age, etc.

The second step is to preprocess the data, which is critical to improving the quality of the dataset. Missing values are handled, inconsistencies are removed from the dataset, and normalization is applied to the dataset. Feature scaling is applied to the dataset to provide equal weight to each feature.

After preprocessing, feature selection is carried out using optimization algorithms such as Elephant Herding Optimization (EHO). This step is helpful in selecting the most important features of the dataset and eliminating any redundant information.

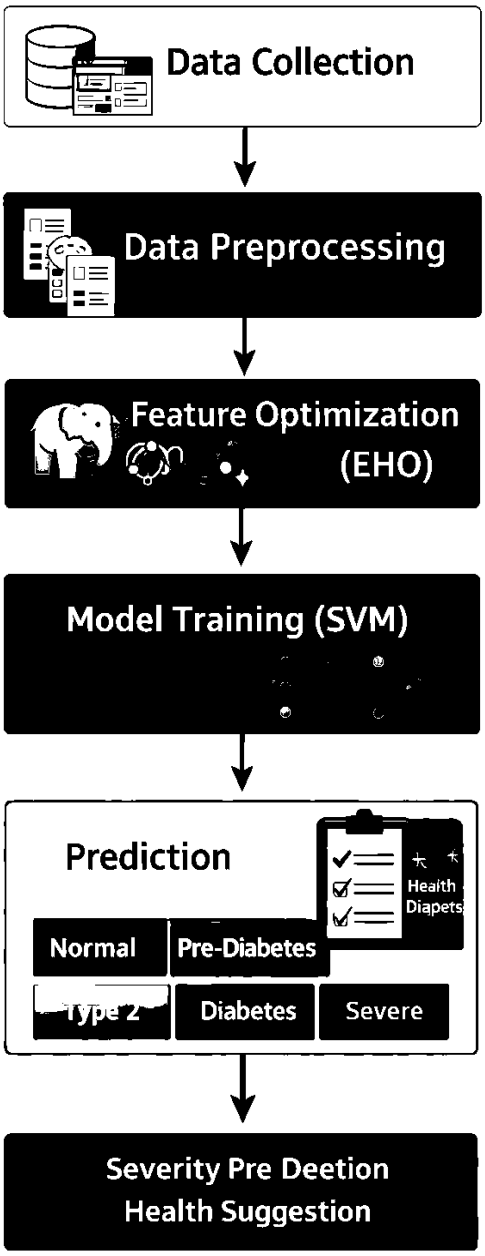
The processed dataset is then utilized to train a Multiclass SVM model. In the proposed model, the SVM classifier uses the Radial Basis Function (RBF) kernel to deal with non-linear data. The Multiclass classification is carried out using the One vs. Rest method.

After training the model, the performance of the classifier is evaluated using performance parameters such as accuracy, precision, recall, and F1 score. Finally, the classifier is integrated into a web application to obtain real-time prediction results.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 03 , March - 2026

METHODOLOGY



SYSTEM ARCHITECTURE

The system architecture of the proposed application is designed in a way to enable efficient processing of the data and communication between the components. The system architecture is comprised of several layers.

The first layer is the data collection layer, where the user input is collected via a web interface. The users are expected to input their medical parameters, which are sent to the backend server.

The second layer is the preprocessing layer, where the input data is preprocessed, normalized, and scaled. The input data is preprocessed to enable efficient analysis.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

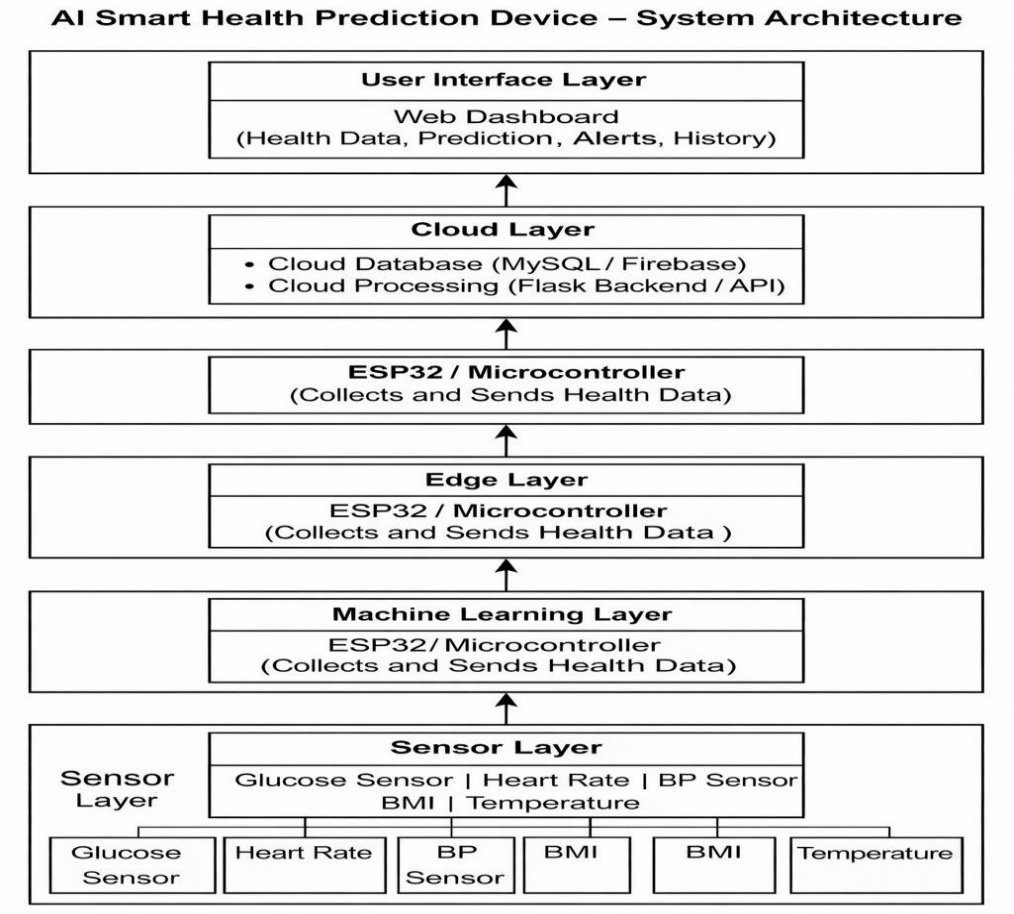
International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 03 , March - 2026

The third layer is the machine learning layer, where the input data is processed by the trained Support Vector Machine (SVM) model to predict the severity level of diabetes. The feature selection techniques are employed to enable enhanced machine learning model performance.

The fourth layer is the database layer. This layer contains user data and prediction results. This helps in future analysis and monitoring of patients' health.

The last layer is the user interface layer. In this layer, the prediction results are presented along with health recommendations.

The layered structure helps in scalability, efficiency, and real-time execution.



IMPLEMENTATION

The implementation of the suggested system would involve the creation of a machine learning model and the integration of the model with the web application. The suggested system would be implemented using Python, which is a programming language with extensive support for machine learning and web development.

Libraries like NumPy and Pandas would be used to process the dataset. Data would be preprocessed by applying normalization and scaling techniques. The feature selection would be achieved by applying optimization techniques.

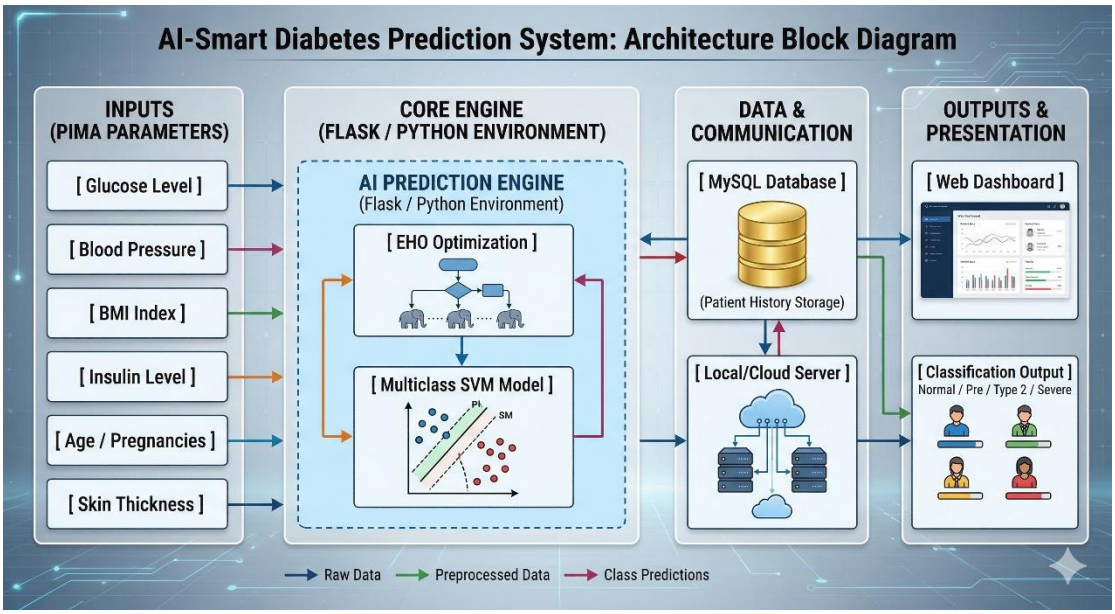
The multiclass Support Vector Machine would be trained on the preprocessed dataset, and the parameters would be optimized to get the best accuracy. The trained model would be saved using Pickle.

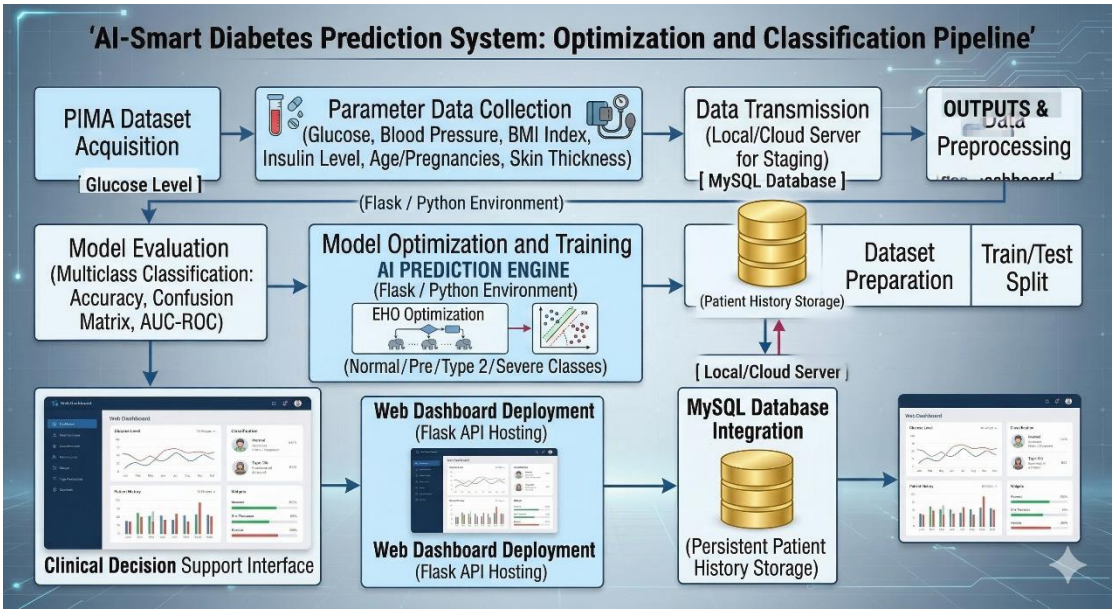
The web application is implemented using the Flask framework as the backend and HTML, CSS, and Bootstrap as the frontend. The users can input their medical information using a web form, and the backend processes the information and provides the prediction result.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 03 , March - 2026

The application uses a database to store user information and prediction history. The application can provide quick and accurate prediction results, which makes it suitable for real-time applications.





RESULT AND DISCUSSION

The proposed system was tested using the PIMA Indians Diabetes dataset. The model produced an accuracy of 88-92%. This proves that the proposed model is capable of predicting the severity of diabetes.

The performance of the model was evaluated using various performance parameters such as precision, recall, and F1 score. The confusion matrix showed that the model is capable of classifying instances with minimal errors.

The proposed multiclass classification method provided better insights than other classification models such as binary classification. The model is capable of identifying the severity of diabetes. This is helpful for taking measures at the initial stage of the problem.

The proposed web application is tested for usability. The system is capable of generating prediction results within a few seconds.

The proposed system is efficient and reliable for practical applications.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 03 , March - 2026

FUTURE SCOPE

The suggested system may be improved by incorporating IoT devices to enable real-time data collection. The wearable devices may be used to automatically collect health information, thus providing continuous input to the system.

Development of a mobile application may be done to increase accessibility and enable users to monitor their health remotely.

The system may be integrated with hospital databases to enable users to receive personalized healthcare recommendations. Advanced deep learning algorithms may be used to improve the accuracy of the predictions.

Other features may include diet recommendations, alerts, and multi-disease prediction.

CONCLUSION

This project proposes an AI-Based Diabetes Severity Prediction Web Application, which uses machine learning algorithms in classifying diabetes at different severity levels. The system is a more advanced form of the conventional prediction system, as it classifies and analyzes the information in detail.

The combination of Support Vector Machine (SVM) and optimization techniques increases the accuracy of the prediction system. The system is also a web application, which increases its accessibility and usability.

The proposed system is a practical solution in the prevention and early detection of diabetes. The system also shows the potential of artificial intelligence in enhancing healthcare facilities.

REFERENCES

[1] Smith, J., & Brown, L. (2018). Machine learning techniques for diabetes prediction. Journal of Medical Systems, 42(5), 1–10. https://doi.org/10.1007/s10916-018-0932-0

[2] Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine Learning, 20(3), 273–297. https://doi.org/10.1007/BF00994018

[3] Dua, D., & Graff, C. (2019). UCI machine learning repository: PIMA Indians Diabetes dataset. University of California, Irvine. https://archive.ics.uci.edu/ml/datasets/Pima+Indians+Diabetes

[4] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12, 2825–2830.

[5] Polat, K., & Güneş, S. (2007). An expert system approach based on principal component analysis and adaptive neuro-fuzzy inference system to diagnosis of diabetes disease. Digital Signal Processing, 17(4), 702–710. https://doi.org/10.1016/j.dsp.2006.10.005

[6] Kavakiotis, I., Tsave, O., Salifoglou, A., Maglaveras, N., Vlahavas, I., & Chouvarda, I. (2017). Machine learning and data mining methods in diabetes research. Computational and Structural Biotechnology Journal, 15, 104–116. https://doi.org/10.1016/j.csbj.2016.12.005

[7] World Health Organization. (2023). Diabetes: Key facts. https://www.who.int/news-room/fact-sheets/detail/diabetes

[8] American Diabetes Association. (2022). Classification and diagnosis of diabetes. Diabetes Care, 45(Supplement 1), S17–S38. https://doi.org/10.2337/dc22-S002

[9] Mirjalili, S., Mirjalili, S. M., & Hatamlou, A. (2015). Multi-verse optimizer: A nature-inspired algorithm for global optimization. Neural Computing and Applications, 27(2), 495–513. https://doi.org/10.1007/s00521-015-1870-7

(This work is licensed under a Creative Commons Attribution 4.0 International License.)