Applied Mathematics and Computation 311 (2017) 22–28

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

# Performance of small-world feedforward neural networks for the diagnosis of diabetes

Okan Erkaymaz[a], Mahmut Ozer[b], Matjaž Perc[c,d,*]

[a] Department of Computer Engineering, Bulent Ecevit University, Zonguldak, Turkey
[b] Department of Electrical & Electronics Engineering, Bulent Ecevit University, Zonguldak, Turkey
[c] Department of Physics, Faculty of Natural Sciences and Mathematics, University of Maribor, Koroška cesta 160, SI-2000 Maribor, Slovenia
[d] Center for Applied Mathematics and Theoretical Physics, University of Maribor, Mladinska 3, SI-2000 Maribor, Slovenia

ARTICLE INFO

*Keywords:*
Diabetes
Small-world network
Feedforward neural network
Rewiring
Newman–Watts model
Watts–Strogatz model

ABSTRACT

We investigate the performance of two different small-world feedforward neural networks for the diagnosis of diabetes. We use the Pima Indians Diabetic Dataset as input. We have previously shown than the Watts–Strogatz small-world feedforward neural network delivers a better classification performance than conventional feedforward neural networks. Here, we compare this performance further with the one delivered by the Newman–Watts small-world feedforward neural network, and we show that the latter is better still. Moreover, we show that Newman–Watts small-world feedforward neural networks yield the highest output correlation as well as the best output error parameters.

## 1. Introduction

Diabetes is a very common health problem of the modern life, spreading rapidly in the world due to the change of nutritional habits [1]. Although type-1, type-2 and gestational diabetes are all common, especially type 2 diabetes mellitus causes significant morbidity and mortality [2]. Therefore, its early detection is of vital importance. Since some forms of diabetes result in a worldwide epidemic that has made it one of the most serious health problem faced by the humankind [3], enormous efforts have been devoted to its early diagnosis and treatment.

Expert systems and artificial intelligence techniques are widely used to aid the diagnosis of diabetes. In this context, the Artificial Neural Network (ANN), originally inspired from real biological networks, has been preferred due to its high classification capability [4]. The architecture of the ANN enables users to construct different types of networks such as feedforward, recurrent and competitive [5]. Among them, a feedforward ANN (FFANN) stands out with its remarkable computational speed [4]. In this context, the FFANN has been proved to be an efficient intelligent system for the diagnosis of diabetes [6–11]. Temurtas et al. [12] used a multilayer neural network (MLNN) structure for the diagnosis of Pima Indians diabetes and found that the classification accuracy of MLNN trained by the Levenberg–Marquardt algorithm was better than that of conventional neural networks. Moreover, Wang et al. [2] have developed and evaluated an effective classification approach by means of ANN to identify those at high risk of type-2 diabetes mellitus without biochemical parameters. Soltani and Jafarian [13] used probabilistic ANN (PNN) for diagnosis of diabetes with type-2.

---

[*] Corresponding author at: Department of Physics, Faculty of Natural Sciences and Mathematics, University of Maribor, Koroška cesta 160, SI-2000 Maribor, Slovenia.

*E-mail addresses:* matjaz.perc@uni-mb.si, matjaz.perc@gmail.com (M. Perc).

http://dx.doi.org/10.1016/j.amc.2017.05.010

**Table 1**
Features of the Pima Indians Diabetic Dataset.

| Features | Diagnosis | Unit |
|---|---|---|
| 1 | Number of times pregnant | – |
| 2 | Plasma glucose concentration | Mg/dl |
| 3 | Diastolic blood pressure | mmHg |
| 4 | Triceps skin fold thickness | Mm |
| 5 | 2-h serum insulin | mu U/ml |
| 6 | Body mass index | kg/m$^2$ |
| 7 | Diabetes pedigree function | – |
| 8 | Age | Year |
| 9 | Result | – |

**Table 2**
Brief statistical analysis of Pima Indians Diabetic Dataset.

| Features | Mean | Deviation | Min | Max |
|---|---|---|---|---|
| 1 | 3.301 | 3.211 | 0 | 17 |
| 2 | 122.628 | 30.861 | 56 | 198 |
| 3 | 70.663 | 12.496 | 24 | 110 |
| 4 | 29.145 | 10.516 | 7 | 63 |
| 5 | 156.056 | 118.842 | 14 | 846 |
| 6 | 33.086 | 7.028 | 18.2 | 67.1 |
| 7 | 0.523 | 0.345 | 0.085 | 2.42 |
| 8 | 30.865 | 10.201 | 21 | 81 |
| 9 | 0.471 | 0.332 | 0 | 1 |

On the other hand, new network topologies have been developed to understand the dynamics of daily life networks because both regular and random networks do not prove fully useful in striving towards the understanding of the real networks [14–18]. In this context, Watts and Strogatz [14] introduced a new network topology called as a Small-World (SW) network. The rewiring algorithm provides a range where the network behaves in a way neither regular nor random. Watts and Strogats [14] showed that some Daily life networks exhibits SW property. Latora and Marchiori [15] analyzed the real data from neural, communication and transport networks and showed that these networks also exhibits SW behavior. SW networks has been proven to be powerful tool to understand the dynamics and information processing capability of the biological neural networks [18–20]. Then, some efforts have been devoted on the application of the rewiring algorithms of the SW networks into the FFANN. Simard et al. [21] carried out a comparative study on the learning performance of regular, SW and random networks using FFANN, and showed that the performance of the SW–FFANN is better than those of regular and random ones. Li et al. [22] developed a multilayer feedforward SW neural network controller and showed that it exhibits better controller performance. We analyzed the real data for estimating the thermal performance of solar air collectors and predicting the modulus of rupture values of oriented strand boards and showed that SW–FFANN results in better accuracy than the conventional FFANN [23]. In a recent study, we compared the performance of the SW–FFANN and the conventional FFANN for diagnosis of diabetes on PIDD [24]. We showed that the SW–FFANN exhibits the highest classification performance.

Literature surveys indicate that SW–FFANNs are constructed based on the Watts and Strogatz rewiring algorithm in Watts and Strogatz [14]. On the other hand, Newman and Watts [25] proposed a different the rewiring algorithm leading to SW behavior. Newman–Watts SW networks have been also widely used to understand the dynamics of different real networks [18–20,26–28]. Therefore, our aim in this study is to extend the subject in Erkaymaz and Ozer [24], and to investigate the impact of both rewiring algorithms on the performance of SW–FFANN for diagnosis of diabetes based on PIDD.

## 2. Mathematical model

In order to compare the performance of the proposed model with that reported in Erkaymaz and Ozer [24], we use the same Pima Indians Diabetic Dataset (PIDD), taken from the UCI machine learning repository [29–34], which includes 768 samples and two classes (normal: 500, diabetic: 268). Each samples have eighth features and one response. These features are shown in Table 1.

We followed the same methodology for the dataset. Therefore, since we have cleared the dataset from missing data, the new dataset has 392 samples (normal: 262, diabetic: 130). The statistical analysis of this dataset is shown in Table 2 [24].

We consider a four layered FFANN for the comparison of the proposed model with that reported in Erkaymaz and Ozer [24], involving 8 input, one output neuron and two hidden layers. We use three different network topology for the FFANN: the first one is the conventional multi-layer FFANN which has a regular topology, the second one is the Watts–Strogatz SW–FFANN and the third one is the Newman–Watts SW–FFANN. The conventional FFANN regular topology is created through