https://doi.org/10.59849/aidd.2024.78

Extreme Learning Machines with Gaussian Random Field Pretraining

Huseyin Tunc

Department of Biostatistics and Medical Informatics, School of Medicine, Bahcesehir University, Istanbul, Turkey,
E-mail: huseyin.tunc@bau.edu.tr

Abstract. Machine learning has become a powerful tool for solving differential equations, offering numerous advantages over traditional methods. However, the unsupervised training of physics-informed neural networks (PINNs) can result in non-convex loss landscapes. Extreme learning machines (ELMs) address this by fixing internal weights and biases, requiring only the determination of output weights. Although ELMs offer benefits over PINNs, their reliance on randomized basis functions limits their expressiveness. To enhance expressiveness of the ELMs, we pretrain single-layer feedforward neural networks using Gaussian random fields, improving the quality of the basis functions. This method, termed PT-ELM, demonstrates that pretrained basis functions in ELMs yield accurate solutions for boundary layer problems with better convergence patterns. PT-ELM outperforms both traditional ELMs and the recently developed TransNet models. The PT-ELM approach can be adapted for solving data and physics driven problems to enhance predictive ability of the neural networks.

Keywords: extreme learning machines, machine learning, transfer learning, Gaussian random fields, boundary layers

Consider the following singularly perturbed boundary value problem (SPBVP) [1-2]:

$$\boldsymbol{y''(x) + n(x)y'(x) = h(x, y(x)), x \in (a, b)} \tag{1}$$

$$y(a) = \alpha, y(b) = \beta, \alpha, \beta \in R$$

where $h(x, y(x))$ and $\frac{\partial h}{\partial y} \geq 0$ are continuous. We study on the boundary layer challenge for approximate solutions of (1). Let us consider the single layer neural network approximation (NNA) function for solving (1) [3]

$$y_{NN}(x) = \sum_{i=1}^{N} \gamma_i \sigma(w_i x + b_i) \tag{2}$$

where N is the number of hidden neurons, $\sigma(x)$ is the activation function (we take $\sigma(x) = tanh(x)$) and the parameters γ_i, w_i and b_i are weights and biases of the network layers. One efficient way of solving the SPBVP (1) with NNA is randomly fixing the weights (w_i) and biases (b_i) of the ith hidden neuron by assuming randomized basis functions and solving the least-square problem for γ_i. This approach is known as the extreme learning machine and various problems have been addressed with this approach [3, 4].

One limitation of the ELM is the poor approximation performance of the randomized basis functions for diverse set of problems. Zhang et al. proposed an efficient way called TransNet for constructing w_i and b_i values from the Gaussian and uniform distributions, respectively [5]. They theoretically showed that the TransNet approach for single layer networks improves the expressibility of the neural networks. In this study, we propose pretrained extreme learning machines (PT-ELM) with single-layer feedforward neural networks to improve the approximation capabilities of the basis functions. We train the neural network basis functions by approximating the set of Gaussian random fields (GRF) [6]. We generate M random discrete map from the discrete input set $(x_i)_{i=1}^{K} \in [0,1]$ to $(y_i)_{i=1}^{K} \in [0,1]$ with the

GRF generator. Our aim is to train the single-layer N-neuron neural network (2) to best approximate the GRF realizations with varying correlation lengths.

We have the neural feature set P_M^L consisting of globally supported basis functions $\sigma(\boldsymbol{w}x+\boldsymbol{b})_i$ for $\boldsymbol{w} \in R^{M\times 1}$ and $\boldsymbol{b} \in R^{M\times 1}$. From the approximation theory perspective these basis functions span $C[a,b]$ and we need to find best practical ones by training $\boldsymbol{w}$ and $\boldsymbol{b}$. To solve this problem, we derive a mixed training approach using the genetic algorithm and the least squares approximation to find optimal probability distributions to generate $\boldsymbol{w}$ and $\boldsymbol{b}$ vectors randomly. By training the neural feature space with this hybrid approach, the basis functions $\sigma(\boldsymbol{w}x+\boldsymbol{b})_i$ are constructed by minimizing the approximation error to the M discrete GRF data. The

We illustrate how the PT-ELM outperforms the ELM and TransNet in solving the SPBVP having boundary layers with numerical examples. We comparatively analyze how the PT-ELM performance is affected from the neural network architecture, number of GRF maps, GRF correlation lengths, and the optimization conditions. We also provide further directions for solving the partial differential equations with the PT-ELM approach.

References

1. Avijit, D., Natesan, S., SDFEM for singularly perturbed boundary-value problems with two parameters, J. Appl. Math. Comput. **64**, 591–614 (2020).
2. Celik, E., Tunc, H., Sari, M., An efficient multi-derivative numerical method for chemical boundary value problems, J. Math. Chem. **62**, 634–653 (2024).
3. Calabrò, F., Fabiani, G., Siettos, C., Extreme learning machine collocation for the numerical solution of elliptic PDEs with sharp gradients, Comput. Methods Appl. Mech. Eng. **387**, 114188 (2021).
4. Quan, H.D., Huynh, H.T., Solving partial differential equation based on extreme learning machine, Math. Comput. Simul. **205**, 697-708 (2023).
5. Zhang, Z., Bao, F., Ju, L., Zhang, G., Transferable neural networks for partial differential equations, J. Sci. Comput. **99,** (2024).
6. Lu, L., Jin, P., Pang, G., Zhang, Z., Karniadakis, G.E., Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators, Nat. Mach. Intell. **3,** pages218–229 (2021).