Adaptive Variational Neural Network for Singularly Perturbed Problems with Continuous and Discontinuous Source Term . ⋆

Yashwant Popat Jadhav[1] and A. Ramesh Babu[2][0000−0002−3325−3050]

[1,2]Dept of Mathematics, Amrita School of Physical Sciences, Coimbatore - 641 112
AMRITA Vishwa Vidyapeetham, India, Email: a_rameshbabu@cb.amrita.edu

Abstract. This paper considered singularly perturbed ordinary differential equations using deep learning and machine learning techniques to attain computed solution. To address singularly perturbed problems, we devised an adaptive variational neural network approach based on supervised learning. By using a variational loss function that is obtained from the weak form of the governing equations, the suggested approach helps the Neural network to find solutions that follow the basic principles of the problem's behaviour. This technique applies the concept of adaptive input parameterization, whereby the distribution of input values within the computational domain is dynamically altered by means of a characteristic parameter associated with the problem. Because of its adaptivity, the NN is guaranteed to concentrate on areas of the solution that exhibits quick changes, which increases accuracy and efficiency. Examples of convection-diffusion problems with continuous and discontinuous coefficients are used to illustrate the efficacy of the suggested technique. Furthermore, the variational loss function lowers the number of training points needed for convergence and improves the efficiency of the training process.

Keywords: Deep learning · Singularly peturbed problem · Machine learning · Neural Network · Supervised learning · Finite Element · VarNet .

1 INTRODUCTION

Singularly Peturbed Differential Equations (SPDEs) a type of differential equation in which the higher-order derivative term is multiplied with the parameter ε. Numerous branches of applied mathematics and physics frequently deal with these issues [18]. Finite Element Method (FEM) techniques

⋆ This work is a part of the project submitted by Yashwant popat jadhav (Roll Number: CB.PS.P2ASD22012).

are more flexible than other computational approaches for solving SPPs when dealing with geometries [15]. One can consult the works of Doolan EP [2], Miller JJH [3], and R.E. O'Malley [1], for in-depth discussions on analytical and numerical solutions.

In this study, we use a novel method for solving singularly perturbed Ordinary Differential Equations with trainable parameters, based on deep neural networks [13, 18]. There are several benefits of approximating ODE solutions with neural networks: 1) It makes it possible to evaluate the solution quickly without sacrificing accuracy or the necessity for Model Order Reduction (MOR); 2) The resultant model is smooth and differentiable, which makes it appropriate for ODE optimization issues.

Any physics problem starts with the important equations that explain the physical phenomeno being studied. These equations frequently include tiny peturbation parameter into the PDEs especially in SPPs. However, individually peturbed parameters may include a hybrid system of PDEs and ODEs, in particular, if the problem involves a system with several time scales or specialized geometric aspects. For example, if the SPP has a thin layer in the domain where the solution varies quickly, then modeling the behavior of the layer using an ODE (or a system of ODEs) and a PDE for the main region may be appropriate. This method may provide a more effective means of capturing the changes inside the layer [4, 6, 13, 18, 21, 23].

We now concentrate on the VarNet Algorithm's approximation of a solution to singularly perturbed problems with boundary conditions. This method examines the convection-difussion problem, wherein the number of input values is chosen between the domains [0,1] in a manner that ensures adaptability. The adaptivity of the input values is as follows: half of the inputs lie between 0 and σ, where σ is the point selected according to the problem's specific physics knowledge. We solve the singularly perturbed problem using the supervised machine learning technique. For the purpose of solving ODEs, the suggested Neural Network model is based on supervised learning. The FEM approach of solving the singularly perturbed problem provides an approximate solution, which is used to train the model, which is similar to Physics Informed Neural Network Techniques (PINN) [16]. It is simpler to describe a complex function as a collection of simple polynomials, which is the main reason for accepting an approximate solution on a collection of picewise linear polynomials [6].

When employing Finite Element Methods (FEMs) for numerical solutions, the weak form is essential since it offers a different, frequently more beneficial means of expressing the governing equations for SPPs.The integral connection between an unknown variable and a collection of weighting functions, or trial functions, is the main emphasis of this study. To improve the accuracy and stability of the discretization process, these weighting functions are selected based on their smoothness and other specific characteristics.

To properly learn the ODE solution, we need a very large number of training points, which is why we use the ODE residual in its differential form as the loss function. Rather, we provide a new loss function based on the ODE's variational (integral) form. In order to improve the accuracy of the ODE solution estimation, this loss function includes lower order derivatives [16].

2 SINGULARLY PETURBED PROBLEM

In mathematics and physics, the phrase "peturbed problem" is typically used to describe the most prevalent scenario. A family of problems is dependent on $\varepsilon > 0$, a tiny parameter is known as boundary layer problem [13, 21]. Numerous fields of applied mathematics, including boundary layers in fluid mechanics, edge layers in solid mechanics, skin layers in electrical applications, shock layers in fluid and solid mechanics, transition points in quantum mechanics, and Stokes lines and surfaces in mathematics, frequently encounter these singular perturbation problems with or without turning points[19, 22].

An ordinary differential equation or a set of differential equations, together with boundary and initial conditions that highlight the issue, may make up the perturbation problem. More general second order singularly perturbed problem is given as:

$$P_\epsilon = \frac{d^2u}{dx^2} = f(x, u, u', \varepsilon)$$

The ML technique known as Physics Informed Neural Networks (PINN) minimizes the equation's residual. support standard FEM with deep learning prediction as PINNs and other approaches can solve SPPDEs with a limited degree of accuracy [18].

3 PROBLEM STATEMENT

In this paper, we examine the convection diffusion problem with continuous and discontinuous source term, which simulates environmental science's problems with pollution [15], such as the flow of river water combined with sewage water and liquid pollution that seeps into the water at a specific location. In this case, two physical processes are at work: first, the pollution diffuses slowly through the water; second, diffusion gradually spreads the pollution along a one-dimensional curve on the surface; and third, when both diffusion and convection are present in a linear differential equation and convection predominates, we have a convection-diffusion problem. A two-point boundary value problem is the most basic mathematical representation of a convection-diffusion problem as follow:

$$\begin{aligned} &-\varepsilon u''(x) - a(x)u'(x) + b(x)u(x) = f(x), \quad 0 < x < 1, \\ &u(0) = 0, \quad u(1) = 0. \end{aligned} \tag{1}$$

where a, b, and f are some specified functions and ε is a small positive parameter. Diffusion is represented by the term u", with a small coefficient $-\varepsilon$. Convection is represented by the term u', while the functions of a reaction and source driving term, respectively, are filled by u and f. (Roos H.G [15], clarify that first-order derivatives should be used to describe convection and second-order derivatives, respectively.)

In order to replicate some real-world issues, there exist separate perturbation models with turning points. We also go over a few models with boundary conditions where the interior layer arises because of non-smoothness of the data or discontinuity in the coefficients.

4 STRUCTURE OF THE NEURAL NETWORK

A machine learning model called an Artificial Neural Network (ANN) is modeled after the architecture and operation of organic neural networks, such as brain. It is a pillar of artificial intelligence and machine learning. In order to create a Physics Informed Neural Network (PINN), which is used to develop network models, ANN is changed to include characteristics from physics models and laws. Further, it involves supervised learning. It is a supervised learning model based on feed forward neural networks (FFNN). The organization of computational neurons in a network is expressed by an ANN's structure, architecture, or topology [5]. These terms, in particular, concentrate on the connections between nodes and layers and the information flow within the network [13].

PINN employs a data-driven strategy that takes into account the physical laws of the data for computing the loss function for training the neural network, in contrast to neural network-based supervised learning, where we have access to the ground truth. The PINN minimizes the residual of the equation confined by boundary constraints in order to approximate the numerical solution [18].

The main concept is to use the available data, which can include both geographical and temporal information, to train the neural network to understand the system's fundamental dynamics.The sequential design of this paradigm allows data to flow progressively from input to output because layers are arranged in a linear stack. Every neuron in a layer is linked to every other neuron in the layer above it, resulting in a dense, or completely connected, network. This architecture enables the model to learn complex patterns and relationships in the data. The hyperbolic tangent (tanh) activation function is used by each of the three hidden layers (neurons) and one input layer (neuron) in the model to create non-linearity. The output layer is made up of a single neuron that is not activated and is appropriate for regression tasks. In order to configure the model for training on particular datasets to approximate solutions to physics approximate solutions to physics based model equation [8, 10], it is constructed with the Adam optimizer and the Mean Squared Error (MSE) loss function.

5 REVIEW ON NEURAL NETWORK SOLUTION OF ODEs

5.1 Adaptive Parameterization for X Values in Neural Network based Convection-Diffusion Solvers

A neural network approach to solve convection-diffusion issues by adaptively selecting the values of the independent variable.It proposes a technique for dynamically adjusting or selecting the input parameters (X values) based on the characteristics of the problem or the data, potentially improving the accuracy and efficiency of the solution process.

Adaptivity, in the context of selecting input values ('x values') within the domain [0,1] refers to the ability to adjust the distribution of points based on certain criteria or parameters inherent to the problem being solved. In the provided method adaptivity is achieved by strategically distributing the input values to capture important physics-related information encoded by the parameter σ. By dividing the domain into two halves and generating points based on the value of σ the method ensures that the distribution of input values adapts to the underlying physics of the problem.

The calculation of $\sigma{=}2\epsilon \log n$, in which n is the total number of input values and ε is a tiny parameter, yields the adaptivity. The features of the problem are used to calculate the value of ε, which is a critical variable that affects the distribution of input values. The approach adjusts to the physics of the issue by adding ε to the selection process, which makes sure that input values are positioned strategically to capture the boundary and interior layer.

5.2 Variational loss function

The objective is to determine the values of θ for the Neural Network so that, for every value of x, the function $u(x)$ as nearly approximates the ODE solution as possible. To arrive this, we define a loss function based on the ODE's variational form, which indicates how well the function $u(x)$ approximates the solution. Then, the ODE's variational form is given as

$$L(u,v) = \int_{\Omega} (\epsilon u'v' + a(x)u'v + b(x)uv)\, du = \int_{\Omega} f(u)\, dx, \tag{2}$$

where u is the trial solution and V is the test function [16]. In order to ensure accuracy and approximation, the projected solution from NN is estimated every epoch until it reaches the defined minimum. At each epoch, the predicted solution is updated by computing the loss and minimizing the loss. The true solution is given by the FEM solution .

$$\text{Loss} = \frac{1}{2}\sum(\text{FEM\_solution} - \text{predicted\_solution})^2 \tag{3}$$

Optimization of the weights is made by backward propagation of the error during training phase. The model reads the input and output values in the training data set and changes the value of the weighted links to reduce the difference between the predicted and target (FEM) solution. The error in prediction is minimized across many training cycles (iteration or epoch) until network reaches specified level of accuracy or defined minima. The variational loss directly leverages the physical constraints of the ODE, guiding the NN to learn solutions that conform to those physical laws [13, 17, 16]. An Adaptive VarNet algorithm has been given as follow:

Algorithm 1: Adaptive VarNet Algorithm

- **Require:** Inputs selected adaptively within the domain [0,1] where σ, which is defined as $\sigma{=}2\epsilon\log n$, splits the input data into two halves.
- **Require:** FE method for neural network training In order to train and assess the neural network in the context of resolving the given problem, the FEM approach shows how important it is to have exact and accurate reference data.
- **Require:** MLP-NN layer widths and the quantity of training points the Neural Network's architecture and design to maximize performance while effectively using computing resources
- It will continuously update the loss and iterate the NN until it fails to attain the necessary minimal loss.
- By calculating the loss and comparing it to the FE solution as the actual solution, we want to demonstrate that the NN has received sufficient training and is able to provide results that are superior to those found using the FE approach.

5.3 Weak formulation

In the process of Adaptive VarNet to singularly peturbed problem,we have to make use of FE solution of the corresponding problem. Because of deriving the FEM solution,the derived SPP should be transformed into variational problem, hence we have stated below the weak formulation of (6.1) as, Find $u \in V$ such that

$$a(u,v) = l(u) \quad , \quad \forall \quad v \in V.$$

Where,

$$a(u,v) = \varepsilon < u', v' > - < bu', v > + < cu, v >$$

$$l(v) = < f, v >$$

Where, V is a solution space with piecewise continuous funnction and $<,>$ is standard innerproduct defined on V.

6 EXPERIMENTS

In this section, we consider the following convection-diffusion equation, in which, we demonstrate the Adaptive variational NN:

$$-\varepsilon u''(x) - a(x)u'(x) + b(x)u(x) = f(x), \quad 0 < x < 1, \tag{4}$$
$$u(0) = 0, \quad u(1) = 0.$$

In certain situations, the solution of (4) can show a boundary layer, changing quickly over a limited area. Solutions are found using methods such as matched asymptotic expansions, which include solving for the inner solution (valid within the boundary layer) and the outer solution (valid away from the boundary layer) independently and matching the results.

In more intricate situations with just two little parameters, as follow:

$$-\varepsilon_1 u''(x) - \varepsilon_2 u'(x) + b(x)u(x) = f(x), \quad 0 < x < 1,$$
$$u(0) = 0, \; u(1) = 0.$$

where ε_1 and ε_2 are small positive parameters, the problem involves multiple scales and possibly overlapping bounndary layers. The solutions may show interactions between many levels or numerous boundary layers. Because of the fast changes, solving these issues sometimes necessitates the use of specialized numerical techniques like adaptive mesh refinement. Alternatively, the complex nature of singularly perturbed issues may be handled with flexibility and efficiency by using Physics-Informed Neural Networks (PINNs), which directly include the differential equations into the training process.

6.1 Convection-Diffusion (Smooth data)

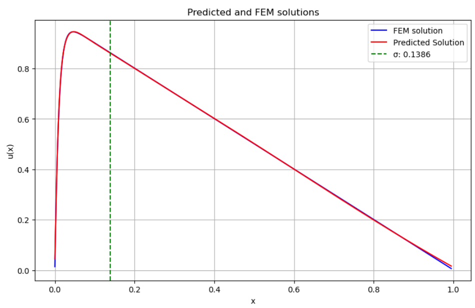
Choose, $\varepsilon = 0.001$, $a(x) = 1$, $b(x) = 0$, $f(x) = 1$, min loss $= 1 \times 10^{-5}$, Activation Function $=$ tanh.

$$-0.001\, u''(x) - u'(x) = 1, \quad 0 < x < 1 \tag{6.1}$$

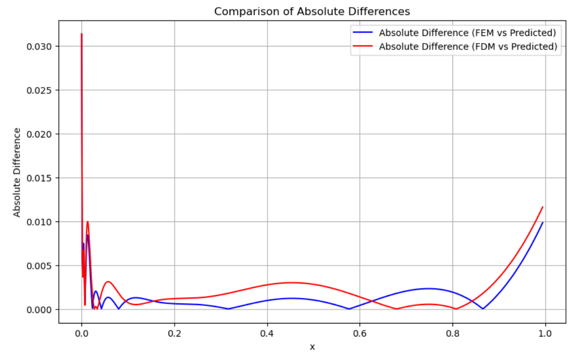
6.2 Reaction-Convection-Diffusion (Smooth data)

Choose, $\varepsilon = 0.001$, $a(x) = 1 + x$, $b(x) = x^2$, $f(x) = sin(x)$,
min loss $= 1 \times 10^{-5}$, Activation Function $=$ tanh.

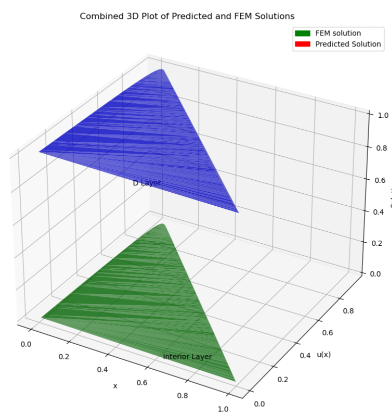
$$-0.001\, u''(x) - (1 + x)u'(x) + x^2 u(x) = sin(x), \quad 0 < x < 1 \tag{6.2}$$



(a) Solution profile obtained from Adaptive VarNet and FEM.

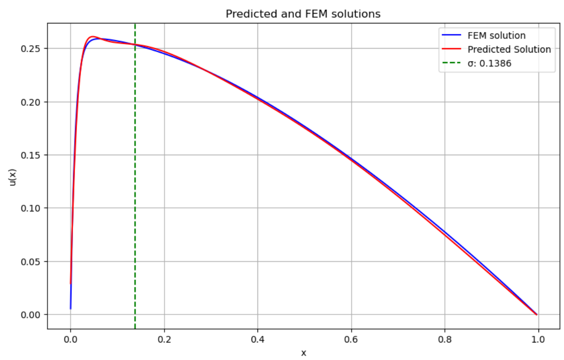


(b) Error graph obtained from predicted vs FEM and predicted vs FDM.

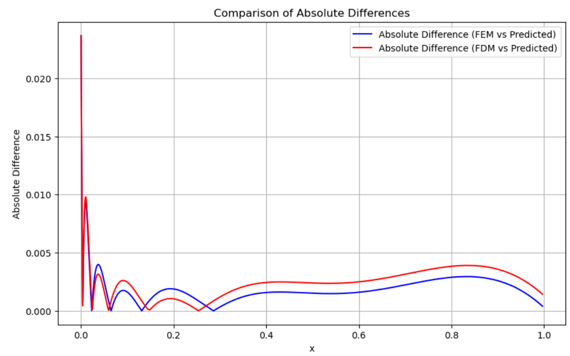


(c) Comparison of predicted and FEM solution via 3D plot.

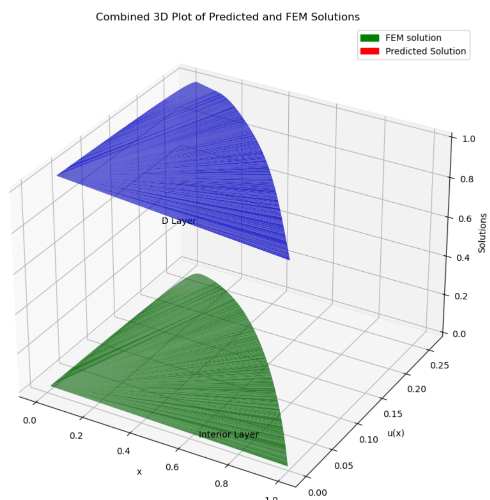
Fig. 1: Various diagrams of Convection-Diffusion problem



(a) Solution profile obtained from Adaptive VarNet and FEM.



(b) Error graph obtained from predicted vs FEM and predicted vs FDM.



(c) Comparison of predicted and FEM solution via 3D plot.

Fig. 2: Various diagrams of Reaction-Convection-Diffusion problem

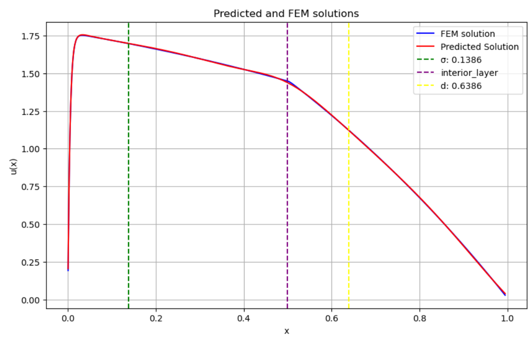
6.3 Convection-Diffusion (Non-smooth data)

Choose $\varepsilon = 0.001$, $a(x) = 2$, $b(x) = 0$,

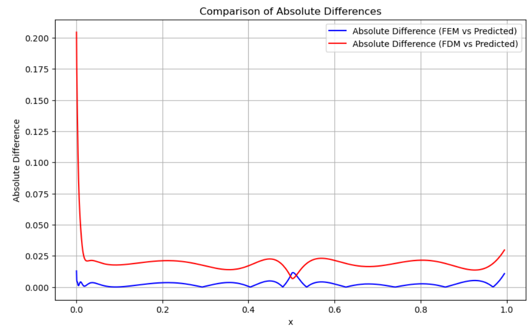
$$f(x) = \begin{cases} e^x, & 0 < x < 0.5, \\ 2 + e^x, & x = 0.5, \\ e^{x+1}, & 0.5 < x \leq 1. \end{cases}$$

min loss $= 1 \times 10^{-5}$, Activation Function $=$ tanh

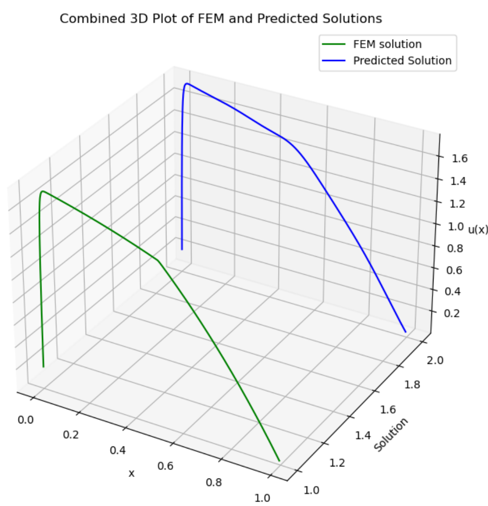
$$-0.001u''(x) - 2u'(x) = \begin{cases} e^x, & 0 < x < 0.5, \\ 2 + e^x, & x = 0.5, \\ e^{x+1}, & 0.5 < x \leq 1. \end{cases} \tag{6.3}$$



(a) Solution profile obtained from Adaptive Var-Net and FEM.



(b) Error graph obtained from predicted vs FEM and predicted vs FDM.



(c) Comparison of predicted and FEM solution via 3D plot.

Fig. 3: Various diagrams of Convection-Diffusion problem

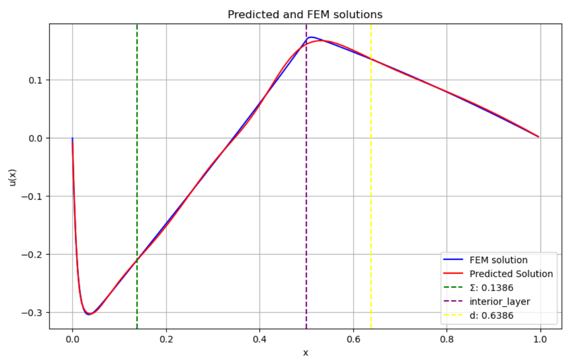
6.4 Reaction-Convection-Diffusion (Non-smooth data)

Choose, $\varepsilon = 0.001$, $a(x) = 1 + x$, $b(x) = x^2$,

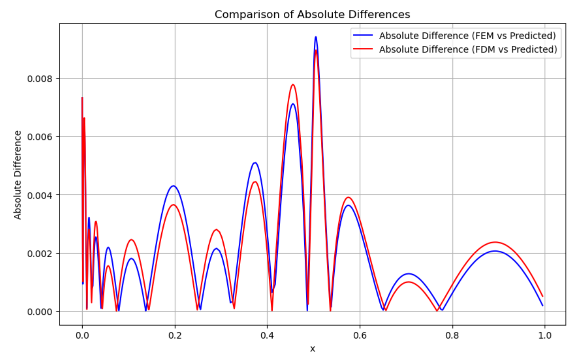
$$f(x) = \begin{cases} -e^x, & 0 \leq x \leq 0.5, \\ \sin(x), & 0.5 < x \leq 1. \end{cases}$$

min loss $= 1 \times 10^{-5}$, Activation Function $=$ tanh

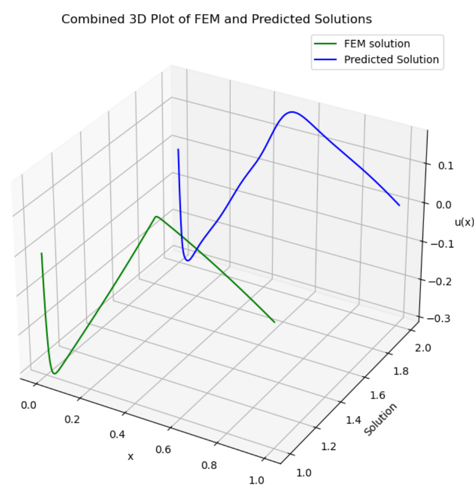
$$-0.001\, u''(x) - (1 + x)u'(x) + x^2 u(x) = \begin{cases} -e^x, & 0 \leq x \leq 0.5, \\ \sin(x), & 0.5 < x \leq 1. \end{cases} \tag{6.4}$$



(a) Solution profile obtained from Adaptive VarNet and FEM.



(b) Error graph obtained from predicted vs FEM and predicted vs FDM.



(c) Comparison of predicted and FEM solution via 3D plot.

Fig. 4: Various diagrams of Reaction-Convection-Diffusion problem

The table given below shows the epoches taken by different activation functions to reach the 1×10^{-5} minimum loss.

Equation	Activation Function	Number of iterations / epochs
6.1	tanh	376
	sigmoid	1112
6.2	tanh	213
	sigmoid	1229
6.3	tanh	1457
	sigmoid	3569
6.4	tanh	1774
	sigmoid	4389

Table 1: For various Activation function,number of iteration to attain the convergence on $\|u^{FEM} - u^{NN}\| < \delta = 1 \times 10^{-5}$.

7 CONCLUSION

In this article, we explore the novel technique of the deep learning and machine learning methods to solve the singularly peturbed problems.

Many techniques of numerical method are used for solving the problems that involve complex discritization to handle the solution near the boundaries or interior layers. However we have to investigate an alternate method using the deep learning and machine learning techniques called adpative variational neural network technique, which approximate the solution with required accuracy. The method demonstrates remarkable accuracy and efficiency in capturing complex phenomena such as sharp transitions and interior layers within the solution domain.

This approach offers benefits in terms of computational efficiency, adaptability to complex geometries, and insights into the underlying physics. Further research and development in this field can enhance our understanding and application of machine learning methods for singularly perturbed problems across diverse scientific and engineering domains.

8 Future Works

Solving the sigularly peturbed problems through the Machine learning has been emerged topic of interest in our scientific society. Our future work will rely on applying ML algorithm to the complex problem.

1) Advancement of model : Mostly, we will focus the advancement of the model, the current supervised model will be get upgraded to the unsupervised model, which will be more faster and provide the appropriate and approximate solution better than numerical methods like FEM and FDM. The model will be upgraded in such a way that it will work for the ODEs as well as PDEs.

2) Parameters : Working with more than one small singularly peturbed parameters problems which are multiplyed with first order derivative and highest second order derivatives. In more intricate situations with just two little factors,as follows:

$$
\begin{aligned}
&-\varepsilon_1 u''(x) - \varepsilon_2 u'(x) + b(x)u(x) = f(x), \quad 0 < x < 1,\\
&u(0) = 0,\ u(1) = 0,
\end{aligned}
$$

where ε_1 and ε_2 are small positive parameters, the problem involves multiple scales and possibly overlapping layers. The solutions may show interactions between many levels or numerous boundary layers. Because of the fast changes, solving these issues sometimes necessitates the use of specialized numerical techniques like adaptive transition points. Alternatively, the complex nature of singularly perturbed issues may be handled with flexibility and efficiency by using physics-informed neural networks (PINNs), which directly include the differential equations into the training process.

3) Activation functions : Inventing the new activation function which will properly and perfectly suits the solution of the singularly peturbed problems with less error between the true solution and the predicted solution.

References

1. O'Malley Robert E : Singular perturbation methods for ordinary differential equations, New York: Springer, Vol. 89, (1991).
2. E.P. Doolsn, John James Henry Miller: Uniform Numerical Methods for Problems with Initial and Boundary Layers, SIAM Review; Philadelphia ,Society for Industrial and Applied Mathematics, vol.25, (1983).
3. Miller John JH, Eugene O'riordan, and Grigorii I. Shishkin : Fitted numerical methods for singular perturbation problems: error estimates in the maximum norm for linear problems in one and two dimensions, World scientific, (1996).
4. Dissanayake M. W. M. G, and Nhan Phan-Thien : "Neural-network-based approximations for solving partial differential equations." communications in Numerical Methods in Engineering, 10(3), 195-201, (1994).

5. Mehrotra Kishan, Chilukuri K. Mohan, and Sanjay Ranka : Elements of artificial neural networks, MIT press, (1997).
6. Yadav Neha, Anupam Yadav, and Manoj Kumar : An introduction to neural network methods for differential equations, Vol. 1, Berlin: Springer, (2015).
7. Li Jichun : "Convergence analysis of finite element methods for singularly perturbed problems," Computers and Mathematics with Applications, 40(6-7), 735-745, (2000).
8. Seshu P : Textbook of finite element analysis. PHI Learning Pvt. Ltd., (2003).
9. Sazli Murat H : "A brief review of feed-forward neural networks", Communications Faculty of Sciences University of Ankara Series A2-A3 Physical Sciences and Engineering, 50(1), (2006).
10. Galushkin Alexander I : Neural networks theory, Springer Science and Business Media, (2007).
11. Amiraliyev G. M., I. G. Amiraliyeva, and Mustafa Kudu : "A numerical treatment for singularly perturbed differential equations with integral boundary condition", Applied mathematics and computation, 185(1), 574-582, (2007).
12. Sharma Kapil K, Pratima Rai, and Kailash C. Patidar : "A review on singularly perturbed differential equations with turning points and interior layers", Applied Mathematics and computation, 219(22), 10575-10609, (2013).
13. Tawfiq L. N. M, and K. M. M. Al-Abrahemee : "Design Neural Network to Solve Singular Pertubation Problems", J. of Applied and Computational Mathematics, 3, 1-5, (2014).
14. Anila S, T Lalithasree, and A Ramesh Babu : "Energy norm error estimate for singularly perturbed fourth-order differential equation with two parameters", International Journal of Computer Mathematics, 100(3), 681-701, (2023).
15. Roos Hans-Görg : Robust numerical methods for singularly perturbed differential equations, Verlag Berlin Heidelberg-Springer, (2008).
16. Khodayi-Mehr Reza, and Michael Zavlanos : "VarNet: Variational neural networks for the solution of partial differential equations", Learning for Dynamics and Control, PMLR, 298-307, (2020).
17. Whiteley Jonathan : "Finite element methods." A Practical Guide 1 (2014).
18. Yadav Sangeeta, and Sashikumaar Ganesan : "SPDE-ConvNet: Solve singularly peturbed partial differential equation with deep learning", World Congress in Computational Mechanics and ECCOMAS Congress, Scipedia SL, (2022).
19. Huang Jianguo, Haoqin Wang, and Tao Zhou : "An augmented Lagrangian deep learning method for variational problems with essential boundary conditions", arXiv preprint arXiv:2106.14348 (2021).
20. Yadav Sangeeta : Data Driven Stabilization Schemes for Singularly Perturbed Differential Equations, Diss. Indian Institute of Science Bangalore–560 012, India, (2023).
21. Puvaneswari, A., Valanarasu, T. and Ramesh Babu, A.: “A System of Singularly Perturbed Periodic Boundary Value Problem: Hybrid Difference Scheme”, Int. J. Appl. Comput. Math 6, 86. https://doi.org/10.1007/s40819-020-00842-1,(2020).
22. P. Roul and V. P. Goura, “A fast numerical scheme for solving singular boundary value problems arising in various physical models”, Journal of Mathematical Chemistry, vol. 60, pp. 514 – 541, (2022).

23. Manikandan Mariappan and Ishwariya Rosey, “Robust computational technique for a class of singularly perturbed nonlinear differential equations with Robinboundary conditions”, Journal of Mathematical Modeling,Vol. 11, No. 2, pp. 411–423, (2023).