

PAPER • OPEN ACCESS

Deep Learning Enhancing Performance Using Support Vector Machine HTM Cortical Learning Algorithm

To cite this article: D. Akila *et al* 2021 *J. Phys.: Conf. Ser.* **1963** 012144

View the [article online](#) for updates and enhancements.

You may also like

- [Mechanisms of Cortical Development](#)
D J Price and D J Willshaw
- [Measuring the electrophysiological effects of direct electrical stimulation after awake brain surgery](#)
M A Vincent, F Bonnetblanc, E Mandonnet et al.
- [Electrospun fiber surface nanotopography influences astrocyte-mediated neurite outgrowth](#)
Christopher D Johnson, Anthony R D'Amato, Devan L Puhl et al.

PRIME
PACIFIC RIM MEETING
ON ELECTROCHEMICAL
AND SOLID STATE SCIENCE

HONOLULU, HI
Oct 6-11, 2024

Abstract submission deadline:
April 12, 2024

Learn more and submit!

Joint Meeting of
The Electrochemical Society
•
The Electrochemical Society of Japan
•
Korea Electrochemical Society

Deep Learning Enhancing Performance Using Support Vector Machine HTM Cortical Learning Algorithm

D. Akila¹, Souvik Pal², R. Jayakarthish³, Sitikantha Chattopadhyay⁴, Ahmed J. Obaid⁵

¹Department of Information Technology, Vels Institute of Science, Technology and Advanced Studies, Chennai, India

²Department of Computer Science and Engineering, Global Institute of Management and Technology, Kolkata, India

³Department of Computer Science, Vels Institute of Science, Technology and Advanced Studies, Chennai, India

⁴Department of Computer Science and Engineering, Brainware University, Barasat, India

⁵Faculty of Computer Science and Mathematics, University of Kufa, Iraq

akiindia@yahoo.com

souvikpal22@gmail.com

drjayakarthish@gmail.com

sitikantha.1984@gmail.com

ahmedj.aljanaby@uokufa.edu.iq

Abstract: Deep Learning is a function of AI that duplicates the mechanisms of human thought in the processing of information and selection processes. The aim of this study is to apply a technology known as SVMHTMC to improve deep learning. The HTM Cortical Learning Approach and the Support Vector Machine have been combined in this suggested algorithm. The deep learning technique is based on the assumption that the mean absolute percentage error is reduced. Aside from the overlapping duty cycle, the high proportion of which shows the speed of the classifier's processing function. The findings demonstrate that by halving the value, the suggested set of criteria minimizes the absolute proportion of mistakes. In addition, raise the percentage of overlapping duty cycles by 17%.

Keyword: Deep learning; HTM algorithm; duty cycle; Support Vector Machine

1. INTRODUCTION

Deep learning acquisition of information is overseen, half-monitored or unattended. Deep learning using neural networks has produced cutting edge solutions in a wide range of activities [1]. Despite early recognition of the benefit of deep architecture, deep learning technology has not gained broad adoption since there is no viable learning strategy for existing learning machines other than models. Deep learning of fashions is loosely related with digesting facts and the styles of communication in a biologically scary maker, including neural coding to explain curting between different stimules and neural reactions inside the mind. Deep learning approaches are defined by their capability to remove features from deep (or numerous) layers automatically. The extraction function is used for different areas, as it is essential for the success of machine learning and model recognition techniques [2].



In the field of representation of data, Deep Learning (DL) has significantly improved and success depends mostly on its profound structure [3]. A solid knowledge of the kinds of dedicated in the processing of natural languages and social media filters is neural networks. Another form of approach is the support vector machine for overfitting problems (SVM). The SVM is a monitored approach of machine training. A common SVM with a slender layer structure for classification and regression, the core technique gives forecasters for the fresh inputs simply by using the kernel function evaluated at a sparse portion of the training data points. The SVM is based on the systemic risk minimisation (SRM) theory, which states that it should construct an optimal decision hyperplane within the maximum margin [4].

Deep learning has increasingly been extended to the analysis of images and the retrieval of features in the area of machine learning in recent years. It is unsupervised learning, through the learning of a complex non-linear network structure, and the achievement of key features from a limited number of data sets [5]. The efficacy of deep learning is influenced by a different learning algorithm. Hierarchical time memory (HTM) is a technological investigation system that records structural and algorithmic homes of neocortex. Within the mind of mammals, the neocortex is the seat of a sensible thinking. The neocortex is responsible for high-level thinking, creative vision, hearing, touch, motion, language, and planning. Given the wide range of cognitive abilities, one may expect the neocortex to support an equally wide range of specialized neural algorithms. HTM is a computational framework that helps researchers better comprehend the neocortex and its varied capacities. We've only used a small portion of this theoretical framework thus far. Over time, more and more of the theory may be put into practice.

2. RELATED WORKS

Pulin Agrawal et al. [6] Our research is aimed at examining If the system accuracy is increased by adding a 2nd higher level region that incorporates CLAs to a one-CLA system. A hierarchical implementation of CLAs employing the LIDA paradigm is employed for their visual associative memory (the learning intelligent distribution agent-LIDA is a cognitive architecture). The Hierarchical Temporal Memory is a paradigm of hierarchically interconnected modules which identify time and spatial patterns, according to Jeff Hawkins' book *On Intelligence*. CLAs are the second HTM implementation (Cortical Learning Algorithms). Wen Zhuo et al. [7] To enhance this model based on LLC and SPM, they applied HTM methods. HTM regions composed of HTM cells are constructed into the LLC space pool. Each cell re-enters a subset of LLC codes, and the neighboring subsets overlap such that more spatial information can be captured. In comparison, HTM cortical learning algorithms have two processes: The learning phase which causes the HTM cell to obtain only the most frequent LLC codes, and the inhibition phase which ensures that the output of the HTM regions is scarce. The experimental findings of the Caltech 101 and UIUC-Sport dataset indicate an improvement on the original LLC & SPM based model.

Abdullah M. Ziyarah et al. [8] In the study, there existed a complete HTM architecture for space and time memory. To tackle the conceivable and complicated interconnections that emerge during learning, a synthetic synapse design is proposed. A bilateral time-saving memory is an approach that promises the building of invariant spatial-temporal representations of input streams (HTM). The architecture is interconnected with parallel cells and columns so that data may be processed fast by HTM. This unexecuted online methodology showed several machine learning features, including anomaly detection. There have been significant efforts to codify and adapt the HTM method to a variety of situations. There are just a few early investigations of the HTM hardware design, especially for the HTM algorithm space pooler. On two separate datasets, MNIST and the European numerical typeface, the suggested architecture is assessed with and without noise (EUNF).

Afeefa PP et al. [9] ALPR technology is used in autonomous traffic management and security monitoring applications such as parking automation, access control, journey time computation, and law enforcement, among others. This approach uses a cortical learning mechanism called hierarchical temporal memory (HTM) to recall all of the characters on the license plate. We were able to obtain great fault tolerance during recognition using the HTM spatial pooler portion.

Yichuan Tang et al. [10] The use of a linear support vector machine to replace the soft max layer yielded a small but noticeable benefit. Rather of cross-entropy loss, learning lowers margin-based loss.

Almost all prior studies have used supervised or unsupervised approaches to learn the hidden representation of deep networks, which is subsequently fed into SVMs as inputs. In contrast to these models, we propose re-propagating gradients through the SVM top-level to train all layers of deep networks, learning features from all levels. Our findings show that the only way to enhance MNIST, ICAR-10, and ICML 2013 Representing Learning Workshop is to replace linear SVMs with softmax.

Francis Quintal Lauzon et al. [11] A technique known as stacked denoising autoencoders is being investigated. We contribute by demonstrating that using this form of representation in combination with the SVM improves MNIST classification error compared to using a logistic regression layer on top of a stack of denoising autoencoders in a traditional deep learning approach.

Davide Maltoni et al. [12] Formal notation and pseudocode descriptions were used to evaluate the HTM architecture and accompanying learning algorithms [13,14 and 15]. Following that, new approaches for encoding coincidence-group membership (fuzzy grouping) and determining temporal groups are proposed (maxstab temporal clustering). To further understand and compare HTM's idiosyncrasies with other well-known pattern-recognition algorithms, systematic research on three line-drawing datasets were done. Our findings suggest that the novel algorithms used are effective, and that HTM, despite its youth, compares favorably to other established technologies.

3. PROPOSED METHOD

The most difficult task is accuracy increase in deep learning system in our work. To assess efficiency, the median absolute percentage error is employed for effective performance. If this mistake lowers the efficiency of the gadget. Suppose the V vector is as strong as it can be. Learning is a classification job where V is a categorical variable. The study task is a regression issue where V is a numerical vector.

The following models can be created without a loss of generality: A classifier, commonly referred to as a classifier law, has the following function: V is a final class set for s1, s2, sJ, and U is class V. U1,.....Up and V, for each of the mutual probability distributions P(U,V), are random variables that have U V values, where U signifies random variables.

Vector (U1, ...,Up). In other words, if an object is drawn equally from a universe by random variables U and V with values u and V from U and V, it may be P(U = u, V = v). Now we're seeking for an algorithm that makes the best predictions feasible, as well as a model that minimizes the projected prediction error, as indicated below: The estimate error, sometimes called generalization error or test error, of model Lf is equal to 1 Model Lf.

$$prederror(Lf) = U, V\{Lf(V, Lf(U))\} \quad (1)$$

Lf is a loss feature that calculates the difference of its two parameters and L is the learning set for the generation of L. Equation.1 effectively evaluates the Lf's estimate inaccuracy of each conceivable item in each pair (u,v)UV that includes all visible pairs (U V)Lf and pairs viewed from the Lf-range. In practice, the aim is to construct a precise and consistent model for all imaginable information, rather than to generate the best accurate predictions of existing data across subset Lf.

i. SVM (Support Vector Machines):

Binary classification has been used to develop support vector machines (SVM). The support vector machine is also a way to overfit the problem (SVM). As a supervised machine learning methodology, Vapnik proposed the SVM. The kernel technique, which predicts fresh inputs using just kernel function evaluated at a very small portion of the data points, is a popular SVM usage for classification and retrieval with a low-lying structure.

The following constrained optimizations comprise on the training informations and their related labels (xn;yn), n=1,...N, xn=RD, tn=1+1: Learning SVMs:

$$\min w, \epsilon_n = \frac{1}{2} w^T w + c \sum_{n=1}^N \epsilon_n \quad (2)$$

ϵ_n is loose variables that penalize data points that violate the conditions for margins. By boosting all x_n data vectors with a scalar value of 1, we may incorporate a bias. The analogous unconstrained optimization problem is as follows:

$$\min w, \epsilon_n = \frac{1}{2} w^T w + c \sum_{n=1}^N \max(1 - w^T x_n, 0) \quad (3)$$

With the usual hinge failure, the goal of Eq. 2 is regarded as the fundamental form of the SVM issue.

SVM, which minimizes the squared mean error, is known as common variation since SVM is not discernible.

ii. SVMHTMC:

The proposed SVMHTMC method, which integrates both SVM and HTM algorithms, was offered in this part as a complete strategy to enhancing the efficiency of deep learning. There are a few centre features unique to classical learning algorithms. Functional choice: inside the initial dataset, not all skills are similarly important. Some, in the different ways of designing the collection of rules may be deceptive, while others may overfit. For the misleading skills and insignificant features that overfit. As a result, these outmoded and useless functions are being phased out. The data set will help the eventual set of laws work better. Assume that the random forest's stage is referred to as LV1,...LVn. The typical estimating error at all phases is thus stated as in the equation (4)

$$Prederror = \sum_{i=1}^n (prederror(L) = X, Y(L\{Y, L(X)\})) \quad (4)$$

Above equation is made by recursively dividing the X input space into subspaces to approximate the Bayes variant partition, then assigning normal 2Y prediction values to all items x within each terminal subspace. Equation (5) is used to represent the estimation error at this stage:

$$Prederror = \sum_{i=1}^n (err(L) = EY(L\{Y, L(2Y)\})) \quad (5)$$

The following principles are firstly developed to combine the random forest with the HTM algorithms. A rooted tree is one in which one of the nodes serves as the basis. We also conclude in our example that the rooted tree is a guided diagram with all borders pointing from the base. When a side links t1 and t2, t1 is supposed to be the t2 node and t2 is supposed to be the t1 node child. If a rooted tree node has one or more kids, it is considered to be internal, while it is thought to be terminal if it has none. Trees are used to explain terminal nodes from time to time. HTM divides the X_t space represented by the t node into spacers equal to each offspring. The cumulative prediction error is then calculated using the equation in each subspace (6).

$$Prederror = \sum_{i=1}^{n1} (prederror(Li)) + \sum_{d=1}^{n2} (prederror(Ld)) \dots \sum_{s=1}^{nn} (prederror(Ls)) \quad (6)$$

Clearly, HTM areas save transformations between sparsely spaced representations. Transformations can take the shape of a linear series of notes in melody in certain circumstances, but a number of viable alternative inputs can be expected simultaneously in the well-known case. The position of the HTM would allow it to make unique projections on the basis of a background that goes back in time. Many HTM recollection enthusiasts are obsessed with sequence remembrance, or the ability to

remember transitions between spatial patterns. The suggested algorithm is represented as a pseudocode in the following algorithm.

Algorithm1. Support Vector Machine HTM Cortical SVMHTMC Learning Algorithm

```

1. Use the Random Forest Classifier
2. U (a predictor) and V (a target) for training data set as input
3. u test(predictor) with test dataset as input
4. In active Columns for s (t)
5. Prediction = wrong
6. For i=0, for cells Per Column-1
7. If the predictive State(s, I r-1) == true
8. sa = get Active Segment(s, I r-1, active State)
9. If s.sequence Segment == is valid then
10. Model = Random Forest Classifier ()
11. Model.Suitable (U, v)
12. Active State(c, I r) = 1
13. if bu Predicted == wrong,
14. for I = 0 for cells Per Column-1
15. Active State(s, I r) = 1
16. I'm in cells for s.
17. for s in segments(s, I
18. if segment Active(s, I s, r)
19. Predict=model.predict(x test)
20. Role 20. Predicting (x)
21. r= r0
22. while t is not a terminal node
23. r = r0
24. end while
25. return t
26. End of Work

```

Above methodology is processed by the utilization of a random classification of forests and the main characteristic of the HTM learning process. The prediction process starts, which involves computing the mean absolute percent mistake for all nodes once the data sets are checked. After that, Go to the active node and complete all active cells of the node. The fitting model is used with the HTM approach during the prediction stage, which divides each active route into sub-active segments.

The suggested algorithm will be applied using the MATLAB programme. In the case study with different data sets, a differentiation is made between SVMHTMC and HTM, it is observed that the resulted output is compared by the value of mean absolute percentage error.

Table1. SVMHTMC and HTM mean absolute percentage error.

	SVMHTMC mean absolute percentage error	HTM mean absolute percentage error
Data1	5%	10%
Data2	7.5%	15%
Data3	10%	20%
Data4	12.5%	25%

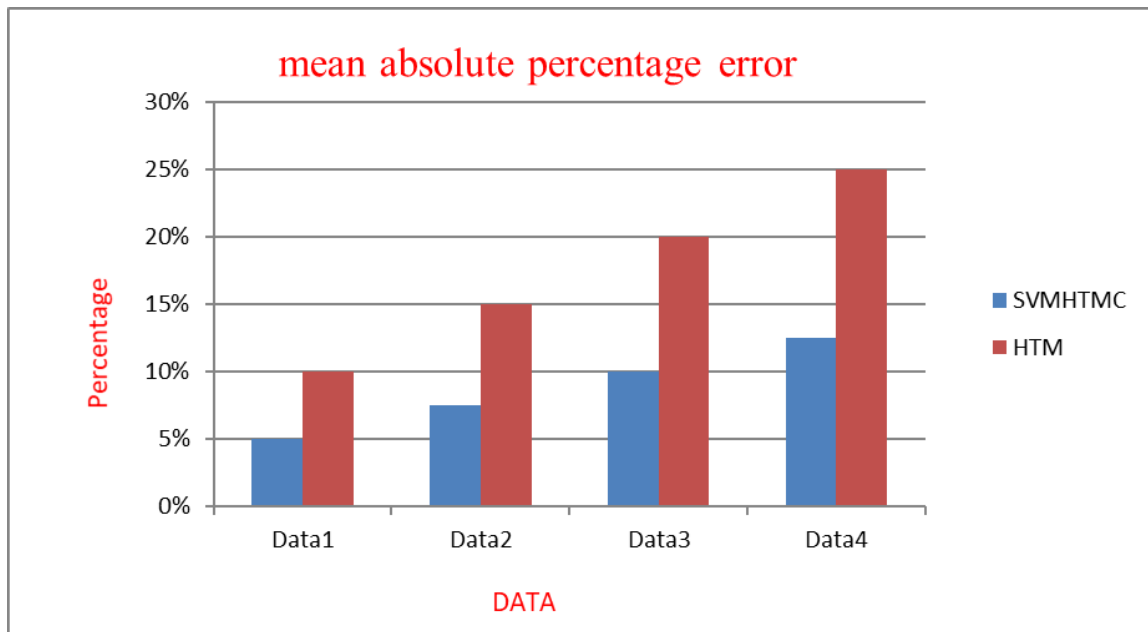


Figure1. SVMHTMC and HTM mean absolute percentage error.

SVMHTMC, the largest mean absolute percentage error in data 1, whereas the mean absolute HTM error in data 1, as shown in Figure 1 is over 10 percent. Figure 1 provides a comparison between SVMHTMC and HTM algorithms for the maximum mean absolute percentage error. When compared to the existing technique, the proposed approach reveals that this mistake may be cut in half. As can be observed in Figure 1, SVMHTMC has more values than the HTM method, implying that the suggested approach is more resilient than the HTM methodology. Figure 1 shows a comparison of the two algorithms' stability properties, with the SVMHTMC method being more stable than the HTM approach.

4. CONCLUSION

Deep learning is a branch of artificial intelligence that focuses on computer learning. The effectiveness of these programs is dependent on their activities. A blend of the RF and the HTM Cortical Learning Algorithm is the main objective for this work. SVMHTMC is the name of the suggested algorithm. According to the data, the proposed method will cut the average absolute percentage error in half. In addition, the overlapping service time changes by 17%. The proposed set of rules might increase the general output of the system's deep knowledge by using previous values. In the same time frame, the active duration values can be kept. It also safeguards the reliability of the gadget and mixes input data units representation.

5. REFERENCES

- [1] Qi, Z., Wang, B., Tian, Y., & Zhang, P. (2016). When ensemble learning meets deep learning: a new deep support vector machine for classification. *Knowledge-Based Systems*, 107, 54-60.
- [2] Kim, S., Choi, Y., & Lee, M. (2015). Deep learning with support vector data description. *Neurocomputing*, 165, 111-117.
- [3] Ju, Y., Guo, J., & Liu, S. (2015, September). A deep learning method combined sparse autoencoder with SVM. In *2015 international conference on cyber-enabled distributed computing and knowledge discovery* (pp. 257-260). IEEE.
- [4] Tang, Y. (2013). Deep learning using linear support vector machines. *arXiv preprint arXiv:1306.0239*.
- [5] Kim, S., Yu, Z., Kil, R. M., & Lee, M. (2015). Deep learning of support vector machines with class probability output networks. *Neural Networks*, 64, 19-28.

- [6] Agrawal, P., & Franklin, S. (2014, December). Multi-layer cortical learning algorithms. In 2014 IEEE symposium on computational intelligence, cognitive algorithms, mind, and brain (CCMB) (pp. 141-147). IEEE.
- [7] Zhuo, W., Cao, Z., Qin, Y., Yu, Z., & Xiao, Y. (2012, November). Image classification using HTM cortical learning algorithms. In Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012) (pp. 2452-2455). IEEE.
- [8] Ziyarah, A. M., & Kudithipudi, D. (2019). Neuromorphic architecture for the hierarchical temporal memory. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 3(1), 4-14.
- [9] Afeefa, P. P., & Thulasidharan, P. P. (2017, June). Automatic License Plate Recognition (ALPR) using HTM cortical learning algorithm. In 2017 International Conference on Intelligent Computing and Control (I2C2) (pp. 1-4). IEEE.
- [10] Tang, Y. (2013). Deep learning using support vector machines. *CoRR*, abs/1306.0239, 2.
- [11] Lauzon, F. Q. (2012, July). An introduction to deep learning. In 2012 11th International Conference on Information Science, Signal Processing and their Applications (ISSPA) (pp. 1438-1439). IEEE.
- [12] Maltoni, D. (2011). Pattern recognition by hierarchical temporal memory. Available at SSRN 3076121.
- [13] G. Suseendran, D. Balaganesh, D. Akila and S. Pal, "Deep learning frequent pattern mining on static semi structured data streams for improving fast speed and complex data streams," 2021 7th International Conference on Optimization and Applications (ICOA), 2021, pp. 1-8, doi: 10.1109/ICOA51614.2021.9442621.
- [14] Tawfiq A. Al-asadi, Ahmed J. Obaid, Rahmat Hidayat, Ts. Azizul Azhar Ramli, 2017. A Survey on Web Mining Techniques and Applications, *International Journal on Advanced Science Engineering and Information Technology*, Vol. 7, No. 4: 1178-1184.
- [15] Tawfiq A. Al-Asadi, Ahmed J. Obaid, Ahmed A. Alkhayat, 2017. Proposed Method for Web Pages Clustering Using Latent Semantic Analysis, *Journal of Engineering and Applied Science*, Vol. 12, No. 8: 8270-8277.