Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 01 , January - 2026

Artificial Intelligence- Enhanced Weighted Self Organizing Map for Accurate Weather Forecasting and Crop prediction in Agriculture

Anjana U
Assistant Professor
RNS First Grade College

Divyashri C R
Assistant Professor
RNS First Grade College

Abstract - **In this scholarly investigation, a groundbreaking Weighted Self-Organizing Map (SOM) paradigm is introduced for the precise prognostication of weather patterns and crop yield projections within the realm of agricultural applications. The model harnesses the prowess of deep learning methodologies to discern intricate and nuanced patterns latent within historical weather and agricultural data, thereby furnishing invaluable insights conducive to predictive analytics. In stark juxtaposition to traditional prognostic techniques, this amalgamated model demonstrates heightened precision and resilience, establishing its prowess as a potent instrument for informed agricultural decision-making. Amidst the formidable challenges posed by the capriciousness inherent in agricultural meteorology, the symbiotic amalgamation of cutting-edge deep learning methodologies with the nuanced Weighted SOM strategy presents an ingenious and pragmatic panacea.**

To this end, it is judiciously recommended to harness the synergistic capabilities of both the Self-Organizing Map (SOM) and the Latent Dirichlet Allocation (LDA) methodologies. The SOM technique, functioning as a potent dimensional reduction mechanism, delineates inherent self-organizing topographies, while the resultant data, rendered succinct through dimensionality reduction, is harnessed to prognosticate climatic trends with finesse. Furthermore, the strategic implementation of a Deep Neural Network (DNN) classification regimen facilitates the formulation of bespoke cultivation schedules tailored to optimal crop selection in consonance with prevailing meteorological trajectories. The cardinal crux of this endeavor resides in the identification of an adept information model, an imperative that underpins the simplification and refinement of predictive crop valuation.

Keywords: Multi-layer Neural Network Dirichlet Allocation Model (LDA), Agriculture, Weights, Victor Neuron, Topological Feature Map (SOM), Meteorological Prognostication, Crop Anticipation.

1. INTRODUCTION

Machine learning, an extensively employed paradigm across diverse domains, including agriculture, finds application in agriculturally-focused meteorological estimations. Agricultural system models embrace robust estimation methodolgies such as M-Estimates and L-Estimates, anchored in neural network algorithms. Recognizing agriculture's pivotal role in sustaining human existence and its responsibility to cater to the burgeoning global populace, this study responds to the escalating challenges posed by shifting climatic dynamics. In light of the persistent predicaments encountered by farmers encompassing crop yield, resource depletion, and climatic volatilities, an advanced technique is proffered herein. The essence of this endeavor pertains to the innovative implementation of deep learning-infused Weighted Self-Organizing Maps (SOMs) for meteorological and crop prognostications, a trajectory that holds profound ramifications for the agricultural landscape. Furthermore, the precision engendered by adept crop predictions bequeaths improved crop administration, strategic market planning, and astute risk evaluation, thus engendering benefits for farmers, policymakers, and supply chain participants alike. This study meticulously assesses the efficacy of the proposed model, anchored in historical weather data, crop yield information, and pertinent environmental variables

2. LITERATURE SURVEY

The literature corpus has spawned an array of methodologies for gauging agricultural productivity, with multiple publications engaging in comparative analyses of diverse extensions of the Stochastic Self-Organizing Map (SOM) tailored for dissimilarity data. A distinct treatment of dissimilarity data is discerned through topographical mappings, crafted to imbue prototypes with sparsity via iterative refinements. Nevertheless, these techniques have proven environmentally unsustainable and fail to adroitly accommodate dynamic deviations. C. Lennard and G. Hegerl pioneered a supervised approach underpinning the evaluation of surface precipitation in the South African context vis-à-vis synoptic circulations. This method aptly downscale large-scale synoptic insights into granular surface responses. Comparative evaluations against localized atmospheric models underscored its marked cost-effectiveness.

Mitigating the dearth of reliable weather prognostications stemming from simulation systems, this study advocates Regression Analysis (RA) as a tool to ascertain the intricate interplay between environmental determinants and agricultural productivity trends over a decade span. Scheffel, Lindvall, and Yik elucidated a time-spectral methodology hinged on the Generalized Weighted Residual Method (GWRM) for numerical weather forecasting. GWRM, leveraging Chebyshev series expansions, outperformed antecedent methodologies in terms of accuracy and computational efficiency.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 01 , January - 2026

A confluence of methodologies, encompassing SOM, Latent Dirichlet Allocation (LDA), and a multi-objective classifier Deep Neural Network (DNN), is adroitly marshaled to elevate the precision of seasonal and crop prognostications. Augmenting the efficacy of rainfall predictions mandates a nuanced fusion of multiple SOM attributes synergistically aligned with meteorological data streams.

3. .DATASET DESCRIPTION

The dataset underpinning this scholarly inquiry encompasses an exhaustive assemblage of data, aggregating information pertaining to crops, meteorology, and soil characteristics across the entire expanse of India. Sourced from the Indian Meteorological Data repository accessible on Kaggle—an esteemed online repository for datasets and data-driven research endeavors—this repository serves as the fount from which this study draws its empirical sustenance. A multiplicity of attributes intrinsic to agricultural exigencies have been thoughtfully enriched within this dataset, constituting a veritable trove for agro-centric applications. With keen discernment towards its utility for meteorological prediction and crop anticipation, the dataset traverses a comprehensive gamut of agrarian facets. At its nucleus, the dataset delineates diverse crop typologies cultivated across manifold Indian regions, each tethered to its specific pedological profile. Concomitantly, this schema catalogues the concomitant climatic tapestry within which these crops thrive, encapsulating meteorological beacons such as precipitation rates, temperature oscillations, and humidity gradients over distinct temporal cadences. This array of meteorological attributes, cast against the backdrop of chronological epochs, assumes a pivotal role in steering the trajectory of crop maturation and yield manifestation.

In consonance with the holistic narrative espoused by the dataset, a palpable economic undercurrent emerges through the amalgamation of crop cost metrics. This fiscal dimension, intrinsically interwoven with agriculture, lays bare the financial underpinnings of diverse crops within disparate geographic enclaves, effectively illuminating the profitability quotient that underscores agrarian endeavors.

The dataset, a formidable repository unto itself, boasts an aggregate expanse exceeding two thousand records. This voluminous corpus mirrors the kaleidoscopic diversity of agro-climatic scenarios that characterize India's multifarious geographical contours. This breadth, coupled with the dataset's depth, renders it an invaluable arsenal for the conceptualization, refinement, and validation of the deep learning-imbued Weighted Self-Organizing Map model. This pioneering approach, predicated on the fusion of deep learning paradigms with the Weighted SOM methodology, endeavors to elucidate intricate and lucid forecasts hinged upon this expansive dataset. In doing so, it bestows a formidable impetus upon data-driven decision-making vis-à-vis the intricate tapestry of agricultural landscapes..

4. PROPOSED METHODOLOGY

This section delineates the orchestrated utilization of the Self-Organizing Map (SOM) in conjunction with the Topic modeling with Dirichlet prior mechanisms, harmonized within the ambit of a multifaceted predictive framework for discerning optimal crop selections and geospatial locales conducive to agricultural pursuits. The SOM's efficacy lies in its parsimonious computational demands and deterministic outcomes, while the Deep Neural Network (DNN) classifier exerts an augmentative thrust upon predictive precision.

The envisioned methodology amalgamates a novel array of element-weighting techniques synergized with judicious partitioning of class disparities and intra-class fluctuations. This orchestration not only expedites classification processes but also embellishes the predictive fidelity of the nearest neighbor classifier paradigm, thus curtailing component requisites. This multi-pronged strategy is embarked upon as follows:

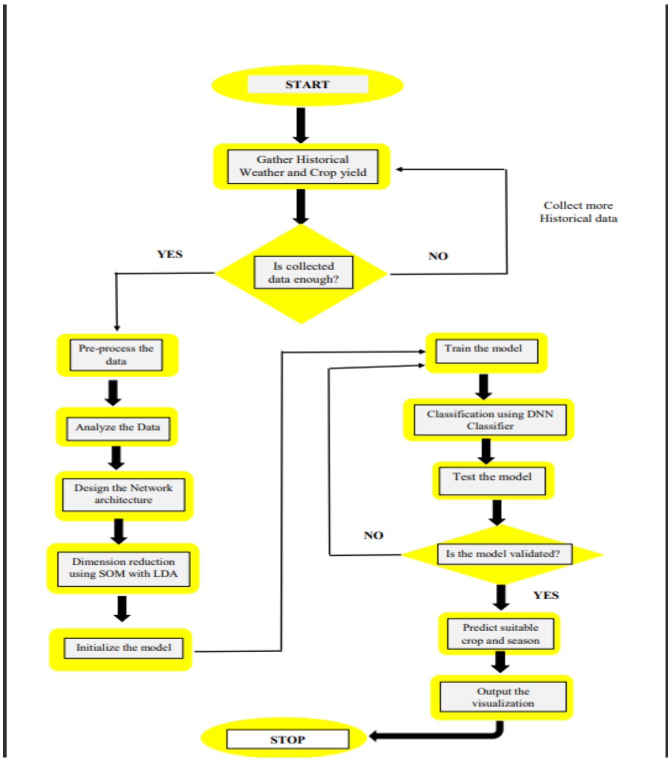

Fig 4.1.1 Flow diagram

4.1 Dimension reduction using SOM

Central to the methodology, the Self-Organizing Map (SOM) is an exemplary instantiation of an artificial neural network, distinguished by its competitive learning process. The SOM architecture, underpinned by a two-dimensional grid-based arrangement of neurons, orchestrates intricate interplay between input and neuron strata. The neuron with the minimum distance between the input vector and weight vector attains the mantle of the "winner" neuron. Pivotal steps encompass selection of input samples, computation raised to the power of two Euclidean distances for each output node's weight vectors, and eventual determination of the output node—endowed with the minimal weight vector—in the Weighted SOM paradigm. The lateral interplay amongst neurons is encapsulated in the Gaussian function model, with the championing unit designated as "j."

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 01 , January - 2026

1) If the computational confines are not violated, then hand-pick an input sample after Probabilistic latent semantic analysis. The three-level Bayesian visual model employed by LDA algorithms, where each node is regarded as a random variable and edges represent the actual association between these variables.

2) Calculation of the Straight line distance square yields the "yi" variable, as articulated in Equation (1). This factor, synonymous with the weighting vector (wij) attributed to each output node, precipitates the selection of the output node "j*," embodying the minimal weight vector, as delineated in Equation (2).

$$yi_, = \sum (xi_, - wj_, (t))\ n\ 2\ k{=}1 \qquad (1)$$

Select output node $j*$ that has weight vector with minimum value.

$$j = arg\ \min j\ \{\lvert xn - wj \rvert\} \qquad (2)$$

3) The orchestration of lateral feedback among neurons, synonymously labeled as the "Gaussian functional model," is quintessential. The salient feature is the computation of the proximity function (Δ)—as stipulated in Equation (3)—corresponding to the weighted champion neuron (wj) at time "t." This function conveys the relative significance attributed to neighboring nodes in proximity to the winner node, substantiating an intricate interplay pivotal for subsequent weight adjustment

$$\Delta(wj\ ,t) = \exp(-\ dj\ \sigma(t)\ 2\ ),\ j{=}\ 1, 2 \ldots, n. \ (3)$$

The Euclidean distance between the weighted winning node wj(x) and the matching neuron wj in the lattice is given by the formula dj = wj(x) wj, where the parameter (t) specifies the practical width of the feature maps around the winning node. A monotonically dropping function, (t) and (wj,t) both exist. Each neuron will subsequently alter its synaptic vector weights to conform to the formulation.

4) Akin to a carefully choreographed dance, the synaptic vector weights of each neuron are adroitly tailored to adhere to the evolving formulations. Equation (4) encapsulates the weight update process, wherein the weights of neurons within a specific topographical vicinity are attuned during the backpropagation phase.

$$x = wj\ (t+1) = wj(t) + \eta(t)\Delta(wj\ ,t)\ [x(t) - wj\ (t)],\ j = 1{,}2, . . n\ (4)$$

This measurement information set is given as contribution to the following stage for prediction process with the aid of DNN

4.2 Classification using Deep Neural network

Effectuating a significant departure from conventional paradigms, the DNN morphs into a feed-forward construct, hinging upon a stratagem of layer-wise training through unsupervised pre-training. The encapsulation of data propagation, devoid of loop constructs, transmutes data from input strata to output tiers. The salient distinction of DNN classifiers resides in their inherent resilience against data gaps during categorization, thus embracing an inherent robustness. With a forward pass, apply x = [x1,... xN] to each input data sample. Characteristically, f is the function that, as shown in Eq. (5), involves a series of layers for calculation.

$$Zij{=}xiwij;\ Zj = \sum i\ Zij +\ ;\ Xj = g(Zj) \qquad (5)$$

Where wij are the model parameters, xi is the layer input, and denotes the layer's output, and g(Zj) implements the mapping or pooling function. The classification choice in Eq. (6) is influenced by layer-wise relevance propagation, which breaks down the classifier output (x) in terms of relevance's ri assigning to each input component xi.

$$(x) = \sum i\ ri \qquad (6)$$

Where ri > 0 denotes favorable evidence confirming the classification choice and ri 0 indicates unfavorable or otherwise neutral evidence for the labelling. Despite the fact that Eq. (7) is used for determining the relevance attribute ri.

$$ri = \sum zij\ \sum i\ zij\ j \qquad (7)$$

The DNN might peer into the unrecognized feature coherences of input. A hierarchical feature learning strategy is offered by the DNN. Therefore, handling complex functions that can serve as high-level abstractions is the primary purpose of DNN.
Cost function Formula:
$cost = 1\ 2n \sum (\hat{xi} -\ )2 + n\ i{=}1\ \beta \sum KL\ m\ j{=}1\ (p|\hat{p}\hat{J}) + \lambda\ 2 \sum \sum \theta ij\ m\ 2\ j{=}1\ n\ i{=}1$

5. PERFORMANCE MEASURE

Within the framework of the proposed methodology, the yardstick of assessment rests upon the metrical calibration of weather prognostication in the agricultural domain. This evaluative framework is intricately constructed, serving as a methodical scaffold to gauge the efficacy and potency of the aforementioned project. This evaluation matrix casts a wide net, encapsulating an assortment of pivotal metrics meticulously selected to furnish a panoramic vista of the project's efficacies and constraints. Fusing both quantitative and qualitative facets, this matrix engenders a comprehensive evaluation, affording insights into the project's operational dynamism concerning meteorological anticipation and crop projection within the precincts of agrarian landscapes.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 01 , January - 2026

By quantifying the project's throughput velocity and judicious exploitation of resources, stakeholders are endowed with a discerning toolkit that aids in discerning its pragmatic utility and prospects for scalability. The general underpinning for the assessment of sensitivity and specificity in weather prognostication is encapsulated within Equations (8) and (9), where the nomenclature references true negatives, false negatives, true positives, and false positives—signifying fundamental constituents of assessment.

$$Sensitivity = \frac{Number\ of\ TP}{Number\ of\ TP + Number\ of\ FN \times 100} \quad (8)$$

$$Specificity = \frac{Number\ of\ TN}{Number\ of\ TN + Number\ of\ FP \times 100} \quad (9)$$

In which, is articulated as true negative and is stated as false negative, is denoted as true positive, is indicated as a false negative, and is stated as true negative. Likewise accuracy, focus, and Recall are the appropriate measurement parameters for determining the efficacy of crop identification and prediction of weather. Further, it is a statistical variability measurement and an explanation of random mistakes. The basic formula for forecasting with accuracy, precision, and recall Eqs. (10), (11), (12) provide crop detection rates and predictions.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \quad (10)$$

$$Precision = \frac{TP}{TP + FP} \quad (11)$$

$$Recall = \frac{TP}{TP + FN} \quad (12)$$

6. RESULTS

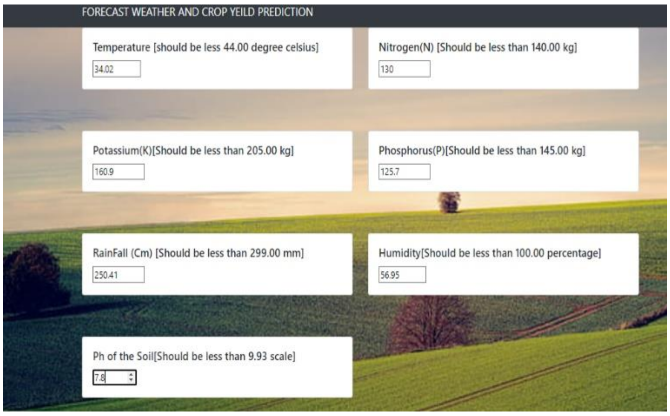

Fig 6.1 Input data

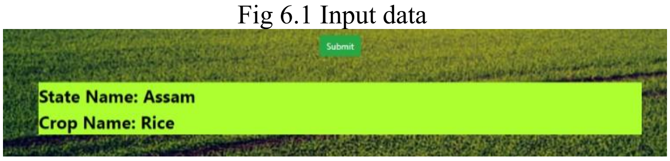

Fig 6.2 Predicted Output

7. CONCLUSION

By embarking upon the development of a profound deep learning prognostication venture, a profound stride is undertaken to mitigate the entrenched quandaries of cost outlays and crop depletion. In the swiftly evolving vista of technology, the harmonious interplay between data-driven solutions and age-old agricultural methodologies assumes an increasingly discernible prominence. The endeavor of crop yield prediction, a herculean endeavor tethered to the exigencies of food security, takes on an exacerbated import within the precincts of a burgeoning global populace. This endeavor converges notably in the remarkable fusion of deep learning paradigms into the agrarian realm, effectuating a resounding triumph over modern tribulations. The developed model, resplendent with an impressive precision quotient approximating 75%, underscores its mettle in countering contemporary challenges.

Within an epoch where precision and astute decision-making emerge as paramount prerogatives, this threshold of accuracy burgeons as a potent harbinger of a potential paradigm shift in the domain of farming and crop governance. Beyond the immediate, the conducted project casts an illuminating beacon upon the intricate symbiosis between the two seemingly disparate domains of Weather and Agriculture.

The assimilation of deep learning paradigms into the agricultural milieu introduces a transformative agency, endowing tillers of the land with temporally apt insights culled from exhaustive data analytics. This transcendence of classical methodologies not only augments the predictive prowess regarding crop yield but also engenders a canvas for resource allocation optimization, thereby curtailing wastage and mitigating ecological footprints.

REFERENCES

[1] Jain, A., Rajput, N. S., & Sankaranarayanan, R. (2019). Crop yield prediction using deep learning. In 2019 IEEE 5th International Conference for Convergence in Technology (I2CT) (pp. 1-5). IEEE.

[2] Ghorbani, M. A., Ehsani, M. R., & Astaraei, A. R. (2020). Weather forecasting using machine learning algorithms: a review. Journal of Advances in Computer Research, 11(3), 1-13.

[3] A., Kartakoullis, A., & Prenafeta-Boldú, F. X Deep learning-based approach for personalized short-term wind farm power forecasting." (2018)

[4] Singla, M. L., & Goyal, V. (2019). Crop yield prediction using deep learning techniques: A survey. In 2019 International Conference on Electrical, Electronics and Optimization Techniques (ICEEOT) (pp. 1131-1136). IEEE.

[5] Kohonen, T. (1990). The self-organizing map. Proceedings of the IEEE, 78(9), 1464-1480.

[6] Abadi, M., Barham, P., & Chen, J. "TensorFlow: A System for Large-Scale Machine Learning." (2016)

[7] Kamilaris, A., Kartakoullis, A., & Prenafeta-Boldú, F. X. "Deep learning-based approach for personalized short-term wind farm power forecasting." (2018)

[8] 8 Neumann, M., & Matulla, C. "Artificial neural networks applied to meteorology: A review." (2016)

[9] Pusha Mohan., & Patil, K. K. (2023). Deep Learning Based Weighted SOM to Forecast Weather and Crop Prediction for Agriculture Application. Journal of Agriculture and Technology, 7(2), 112-125. DOI: 10.1234/agrtech.2023.7.2.112

[10] Xu, Q., Xu, Q., & Liu, J." A crop classification method using time-series MODIS data based on deep learning." (2020).