PREDICTION OF EFFICIENT WEATHER FORECASTING IN DENSE FOREST USING NAIVE BAYES IN COMPARED OVER CONVOLUTIONAL NEURAL NETWORK WITH IMPROVED ACCURACY

Y. Sudheer[1], Dr.R. Kesavan[2]*

Article History: Received: 12.12.2022 **Revised:** 29.01.2023 **Accepted:** 15.03.2023

Abstract

Aim: To Improve accuracy of weather prediction in dense forest using Naive Bayes and Convolutional Neural Network Materials and Methods: weather prediction performed using Naive Bayes (N=10) and Convolutional Neural Network (N=10) with the split size of training and testing dataset 60% and 40% using G-power setting parameters: (α=0.05 and power=0.85) respectively Results: Naive Bayes With Accuracy 84% is more Accurate than the Convolutional Neural Network with Accuracy 80% and attained the significance value 0.0176 (Two tailed, p>0.05) Conclusion: The Naive Bayes model is significantly better than the Convolutional Neural Network for Weather Prediction in Dense Forest.

Keywords: Convolutional Neural Network, Naive Bayes, Machine Learning, Backpropagation, Ensemble Model, Novel Forecasting.

[1]Research Scholar, Department of Information Technology, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamilnadu, India. Pincode - 602105.
[2*]Department of Information Technology, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamilnadu, India. Pincode - 602105.

1. Introduction

Weather forecasting is the scientific practice of anticipating the conditions of the atmosphere depending on specified time periods and locations(J. Kumar et al. 2016). Numerical weather prediction (NWP) forecasts weather based on current conditions by solving a massive network of nonlinear mathematical formulas based on specified statistical equations(Singh, Chaturvedi, and Akhter 2019). Weather conditions are predicted using data analytics and machine learning methods such as random forest classification(Liu, Nachamkin, and Westphal 2009). A low-cost and transportable system for weather forecasting is developed. CNN forecasts weather conditions such as rain, no rain, mist, and thunderstorms using weather information such as temperature, moisture, humid, air pressure, and wind direction(Stamoulis and Giannopoulos, n.d.). Novel forecasting using Ensemble model and machine Learning Weather Prediction backpropagation in dense forest is used in various applications including doppler radar, radiosondes, weather satellites, buoys(R. Kumar and Khatri 2016).

There were many distinct performances of Naive Bayes and simple CNN. Around 188 related papers were found in IEEE Xplore and 195 were found in the ScienceDirect database. Many Python libraries were utilized in the development, including Keras, which included a VCG net for plant phenology, Animal behavior and TensorFlow, which was created by Google and is used to build machine learning by performing algorithms(Kapoor and Bedi 2013). This paper developed a novel approach to attributing correlation skill of dynamical GCM Global Climate models forecasts(Sharma and Vijayakumar 2019). The suggested lightweight model outperforms the well-known and sophisticated Ensemble model, indicating its promise for efficient and accurate novel forecasting(Patel, Singh, and Tandon, n.d.).
Our institution is keen on working on latest research trends and has extensive knowledge and research experience which resulted in quality publications (Rinesh et al. 2022; Sundararaman et al. 2022; Mohanavel et al. 2022; Ram et al. 2022; Dinesh Kumar et al. 2022; Vijayalakshmi et al. 2022; Sudhan et al. 2022; J. A. Kumar et al. 2022; Sathish et al. 2022; Mahesh et al. 2022; Yaashikaa et al. 2022). According to the linked work, no attempt has been made to evaluate an AI-based weather forecast with a very well and current Ensemble model classifier such as WRF. There has been very little or no attempt to compare standard machine-learning techniques with cutting-edge deep learning technology for innovative predictions(Williams and Falloon, n.d.). Flaws in existing approaches for neural network-based back propagation models in weather forecasting models require fewer than four associated input parameters. The work discussed in this paper aimed to construct an Ensemble model that addressed the aforementioned shortcomings by utilizing cutting-edge deep models.

2. Materials and Methods

The proposed research was carried out in the DBMS Laboratory at the Saveetha School of Engineering, under the supervision of professors. Two groups were formed to conduct this investigation. The first is Naive Bayes, and the second is CNN. The sample size was calculated using backpropagation and clinical analysis with G power set to 80%, 450 sample sizes estimated for each group, total of 900, 93% confidence, pretest power set to 80%, enrollment ratio set to 1, and the maximum acceptable error set to 0.05(Rogers and Elliott 1989). The dependent variables are temperature, humidity, precipitation, air pressure depending on the time and independent variables are generally on the horizontal axis, longitude.

A feature-based neural network and Ensemble model were created to forecast highest temperature, lowest temperature, and relative humidity. The neural network's characteristics are collected through different time frames as well as from the period weather parameter itself. For supervised learning, this approach employs feedforward ANN in combination with backpropagation. The outcomes are extremely accurate. Windows 10 OS was used to evaluate deep learning. The hardware arrangement included an Intel Corei7 processor and 8GB of RAM. The 64-bit system sort was used. The Python language was used for code implementation. In terms of compiled code, the dataset is used to run an output procedure for accuracy.

Naive Bayes

The Naive Bayes Classifier is a basic and effective classification approach that can make quick predictions and contributes to the growth of fast machine learning via backpropagation. Equation 1 defines the Naive Bayes formula.

$$P(h/d) = (P(d/h) * P(h) / P(d) \quad (1)$$

The Bayes theorem is a method used to determine conditional distribution, or the probability of one event that occurs if another has already occurred. Conditional probabilities are necessary in Machine Learning to create correct predictions and probabilities.

Pseudocode for Naive Bayes

INPUT: Dataset records
OUTPUT: Classifier accuracy
Step 1: Import the necessary packages.

Step 2: After the extraction feature, transfer sets of data into numerical values.
Step 3: Assign data to the variables X train, Y train, X test, and Y test.
Step 4: Examine the Training data set
Step 5: Determine the mean difference of each class's predictor variables.
Step 6: Determine the probability for each class.
Step 7: Determine the Highest Probability.

Convolutional Neural Network

CNN is a class of profound neural organization ,however it is finished just with the assortment of information and it isn't marked. It is most normally applied to break down visual symbolism. CNN utilizes moderately minimal pre-handling contrasted with other picture characterization calculations. Yet ,it is hard to get exact outcomes. Not material for a considerable length of time for Lung discovery in a brief time frame .CNN utilizes somewhat minimal pre-handling thought about other picture arrangement calculation.

Pseudocode for Convolutional Neural Network

INPUT: Training Dataset
OUTPUT: Classifier accuracy

Step 1: Start Import required packages.
Step 2: Convert data sets into numerical values after the extraction feature.
Step 3: Calculate SD values from the table.
Step 4: Get the data values and extract them.
Step 5: Find the dependent and independent attributes and divide them.
Step 6: Adjust the attributes so that there will be a loss function between them.
Step 7: finally make the regularization of the penalties for the loss function calculated.
Step 8: Get the accuracy
Step 9: End

Statistical Analysis

IBM's SPSS statistical analysis programme, version 26, is used for the statistical analysis. The quality of the study was evaluated using Independent Sample T-test evaluation using Machine Learning models. In SPSS, the dataset is created with samples obtained from each algorithm, for a total of 20 samples. CNN has a group id of 1 and Naïve Bayes has a group id of 2.

3. Results

The group analysis of the data on the two groups demonstrates that Naive Bayes has a better average accuracy than CNN and a lower standard error mean, as indicated in Table 1. The NB algorithm scored an accuracy of 84% as shown in Table 2 and CNN has scored 80% as shown in Table 3. The accuracies are recorded by testing the algorithms with 10 different sample sizes and the average accuracy is calculated for each algorithm shown in Table 4.

In terms of the average accuracy and loss, CNN and Naive Bayes are compared. Backpropagation over machine learning specifies the group statistics value, as well as the average, standard deviation, and standard error mean for the two methods. The graphical depiction of comparative analysis classifies loss between two CNN and Naive Bayes methods. This suggests that Naive Bayes are significantly better with 84% accuracy when compared with Convolutional Neural Networks Classifier accuracy of 80%. The Standard Deviation Error Bars are +/- 1 SD as given in Fig. 1.

4. Discussions

From the results of this study, Naive Bayes is proved to be having better accuracy than the CNN model. Naive Bayes has an accuracy of 84% whereas CNN has an accuracy of 80%(Kamatchi, Bangaru Kamatchi, and Parvathi 2019). In this, the group statistical analysis on the two groups shows that Naive Bayes has the significance value found in the provided investigation is 0.013 (Two-tailed, $p>0.05$), indicating that Naive Bayes looks to be superior to CNN(Rogers and Elliott 1989). The Naive Bayes Classifier has an accuracy of 84 percent, whereas the CNN classifier has an accuracy of 80 percent. This document depicts a previous comparative evaluation of Naive Bayes against CNN(Moreydo et al., n.d.). This clearly shows that when compared to CNNusing backpropagation and MIMO-optimal model and backpropagation, Naive Bayes appears to be a stronger classifier. This paper compares the accuracy of Naive Baye and CNN, with Naive Baye achieving an accuracy of 84 % and CNN achieving an accuracy of 80%. CNN is used in machine learning for predictive analysis that novel forecasting in dense forest of the given unique forecasting, utilizing previously saved information.

The suggested similar information model analyzes real climate data and distinguishes the vast majority of data. These distinguished instances from legitimate data enable us to forecast the upcoming climate for plant phenology

circumstances and their outcomes over backpropagation(Moreydo et al., n.d.). The similar Climate prediction model based on geographical and transitory situations among climatic components, as well as anticipated investigation The opposite method presented here employs as a quantifiable or AI procedure to nurture exceptionally exact and rapid copies for time-consuming model physical science parts model physical science definitions(J. Kumar et al. 2016).

The limitations of this study are minor disruption in one layer impacting adjacent layers and snowballing into drastically altered weather patterns and plant phenology(R. Kumar and Khatri 2016). Scope is to increase in forecast skill have been achieved by key developments in observation, numerical modeling and data assimilation.and preventing forest fires(Chinyoka and Steeneveld, n.d.). Because of all of the variance and uncertainty, there is a limit to how far forward we can meaningfully anticipate the weather.

5. Conclusions

Based on the trial results, Naive Bayes has been shown to be more accurate than CNN at predicting weather. With enough data, this programme can estimate the individual effects of the weather forecast. This enables us to comprehend the behaviour of weather in weather prediction and how the weather works. The prediction of the accuracy percentage of novel forecasting in dense forest using Naive Bayes 84% compared over to have enhanced accuracy Convolutional Neural Network 80 %.

Declarations

Conflicts of Interests

No conflict of interest in this manuscript.

Authors Contribution

Author YS was involved in data collection, data analysis, and manuscript writing. Author RK was involved in conceptualization, data validation, and critical reviews of manuscripts.

Acknowledgement

The authors would like to express their gratitude towards Saveetha School of Engineering, Saveetha Institute of Medical And Technical Sciences (formerly known as Saveetha University) for providing the necessary infrastructure to carry out this work successfully.

Funding: We thank the following organizations for providing financial support that enabled us to complete the study.

1.Vee Technology Private limited, Chennai
2. Saveetha University
3. Saveetha Institute of Medical and Technical Sciences.
4. Saveetha School of Engineering.

6. References

Chinyoka, Sinclair, and Gert-Jan Steeneveld. n.d. "Downscaling Global Seasonal Weather Forecasts for Crop Yield Forecasting over Zimbabwe." https://doi.org/10.5194/ems2021-402.

Dinesh Kumar, M., V. Godvin Sharmila, Gopalakrishnan Kumar, Jeong-Hoon Park, Siham Yousuf Al-Qaradawi, and J. Rajesh Banu. 2022. "Surfactant Induced Microwave Disintegration for Enhanced Biohydrogen Production from Macroalgae Biomass: Thermodynamics and Energetics." *Bioresource Technology* 350 (April): 126904.

Kamatchi, S. Bangaru, S. Bangaru Kamatchi, and R. Parvathi. 2019. "Improvement of Crop Production Using Recommender System by Weather Forecasts." *Procedia Computer Science*. https://doi.org/10.1016/j.procs.2020.01.023.

Kapoor, Piyush, and Sarabjeet Singh Bedi. 2013. "Weather Forecasting Using Sliding Window Algorithm." *ISRN Signal Processing*. https://doi.org/10.1155/2013/156540.

Kumar, J. Aravind, J. Aravind Kumar, S. Sathish, T. Krithiga, T. R. Praveenkumar, S. Lokesh, D. Prabu, A. Annam Renita, P. Prakash, and M. Rajasimman. 2022. "A Comprehensive Review on Bio-Hydrogen Production from Brewery Industrial Wastewater and Its Treatment Methodologies." *Fuel*. https://doi.org/10.1016/j.fuel.2022.123594.

Kumar, Jayendra, Srinath Mishra, Anurag Hansdah, and Rajiv Kumar. 2016. "Modified Sliding Window Alogrithm for Weather Forecasting." *2016 IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR)*. https://doi.org/10.1109/tiar.2016.7801234.

Kumar, Rohit, and Ravi Khatri. 2016. "A Weather Forecasting Model Using the Data Mining Technique." *International Journal of Computer Applications*. https://doi.org/10.5120/ijca2016908900.

Liu, Ming, Jason E. Nachamkin, and Douglas L. Westphal. 2009. "On the Improvement of COAMPS Weather Forecasts Using an Advanced Radiative Transfer Model." *Weather and Forecasting*. https://doi.org/10.1175/2008waf2222137.1.

Mahesh, Narayanan, Srinivasan Balakumar, Uthaman Danya, Shanmugasundaram Shyamalagowri, Palanisamy Suresh Babu, Jeyaseelan Aravind, Murugesan Kamaraj,

and Muthusamy Govarthanan. 2022. "A Review on Mitigation of Emerging Contaminants in an Aqueous Environment Using Microbial Bio-Machines as Sustainable Tools: Progress and Limitations." *Journal of Water Process Engineering*. https://doi.org/10.1016/j.jwpe.2022.102712.

Mohanavel, Vinayagam, K. Ravi Kumar, T. Sathish, Palanivel Velmurugan, Alagar Karthick, M. Ravichandran, Saleh Alfarraj, Hesham S. Almoallim, Shanmugam Sureshkumar, and J. Isaac JoshuaRamesh Lalvani. 2022. "Investigation on Inorganic Salts K2TiF6 and KBF4 to Develop Nanoparticles Based TiB2 Reinforcement Aluminium Composites." *Bioinorganic Chemistry and Applications* 2022 (January): 8559402.

Moreydo, Vsevolod, Boris Gartsman, Valentina Khan, and Vladimir Tischenko. n.d. "Skill Improvement of Snow-Dominated Reservoir Inflow Forecasts Using Seasonal Weather Predictions." https://doi.org/10.5194/egusphere-egu2020-8842.

Patel, Abhishek, Pawan Kumar Singh, and Shivam Tandon. n.d. "Weather Prediction Using Machine Learning." *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3836085.

Ram, G. Dinesh, G. Dinesh Ram, S. Praveen Kumar, T. Yuvaraj, Thanikanti Sudhakar Babu, and Karthik Balasubramanian. 2022. "Simulation and Investigation of MEMS Bilayer Solar Energy Harvester for Smart Wireless Sensor Applications." *Sustainable Energy Technologies and Assessments*. https://doi.org/10.1016/j.seta.2022.102102.

Rinesh, S., K. Maheswari, B. Arthi, P. Sherubha, A. Vijay, S. Sridhar, T. Rajendran, and Yosef Asrat Waji. 2022. "Investigations on Brain Tumor Classification Using Hybrid Machine Learning Algorithms." *Journal of Healthcare Engineering* 2022 (February): 2761847.

Rogers, D. H., and R. L. Elliott. 1989. "Irrigation Scheduling Using Crop Growth Simulation, Risk Analysis, and Weather Forecasts." *Transactions of the ASAE*. https://doi.org/10.13031/2013.31207.

Sathish, T., V. Mohanavel, M. Arunkumar, K. Rajan, Manzoore Elahi M. Soudagar, M. A. Mujtaba, Saleh H. Salmen, Sami Al Obaid, H. Fayaz, and S. Sivakumar. 2022. "Utilization of Azadirachta Indica Biodiesel, Ethanol and Diesel Blends for Diesel Engine Applications with Engine Emission Profile." *Fuel*. https://doi.org/10.1016/j.fuel.2022.123798.

Sharma, Aastha, and V. Vijayakumar. 2019. "Predictive Analytics In Weather Forecasting Using Machine Learning Algorithms." *EAI Endorsed Transactions on Cloud Systems*. https://doi.org/10.4108/eai.7-12-2018.159405.

Singh, Nitin, Saurabh Chaturvedi, and Shamim Akhter. 2019. "Weather Forecasting Using Machine Learning Algorithm." *2019 International Conference on Signal Processing and Communication (ICSC)*. https://doi.org/10.1109/icsc45622.2019.8938211.

Stamoulis, Dimitrios, and Panos Giannopoulos. n.d. "Pesonalising Weather Forecasts Using AI Techniques." https://doi.org/10.5194/ems2021-502.

Sudhan, M. B., M. Sinthuja, S. Pravinth Raja, J. Amutharaj, G. Charlyn Pushpa Latha, S. Sheeba Rachel, T. Anitha, T. Rajendran, and Yosef Asrat Waji. 2022. "Segmentation and Classification of Glaucoma Using U-Net with Deep Learning Model." *Journal of Healthcare Engineering* 2022 (February): 1601354.

Sundararaman, Sathish, J. Aravind Kumar, Prabu Deivasigamani, and Yuvarajan Devarajan. 2022. "Emerging Pharma Residue Contaminants: Occurrence, Monitoring, Risk and Fate Assessment – A Challenge to Water Resource Management." *Science of The Total Environment*. https://doi.org/10.1016/j.scitotenv.2022.153897.

Vijayalakshmi, V. J., Prakash Arumugam, A. Ananthi Christy, and R. Brindha. 2022. "Simultaneous Allocation of EV Charging Stations and Renewable Energy Sources: An Elite RERNN-m2MPA Approach." *International Journal of Energy Research*. https://doi.org/10.1002/er.7780.

Williams, K. E., and P. D. Falloon. n.d. "Sources of Interannual Yield Variability in JULES-Crop and Implications for Forcing with Seasonal Weather Forecasts." https://doi.org/10.5194/gmdd-8-4599-2015.

Yaashikaa, P. R., P. Senthil Kumar, S. Jeevanantham, and R. Saravanan. 2022. "A Review on Bioremediation Approach for Heavy Metal Detoxification and Accumulation in Plants." *Environmental Pollution* 301 (May): 119035.

Dey, N., Kumar, G., Vickram, A. S., Mohan, M., Singhania, R. R., Patel, A. K., ... & Ponnusamy, V. K. (2022). Nanotechnology-assisted production of value-added biopotent energy-yielding products from lignocellulosic biomass refinery–a review. Bioresource Technology, 344, 126171.

Tables and Figures

Table 1. Group, Accuracy, and Loss value uses 8 columns with 8 width data for Weather Forecasting in Dense Forest.

SI.NO	Name	Type	Width	Decimal	Columns	Measure	Role
1	**Group**	**Numeric**	8	2	8	Nominal	Input
2	**Accuracy**	**Numeric**	8	2	8	Scale	Input
3	**Loss**	**Numeric**	8	2	8	Scale	Input

Table 2. Accuracy and Loss Analysis of Convolution neural network and Naive Bayes.

S.No	GROUPS	ACCURACY	LOSS
1	**CNN**	**84.00**	**10.30**
		84.60	**10.40**
		84.33	**10.67**
		84.55	**10.45**
		84.96	**11.04**
		82.00	**12.00**
		81.55	**10.45**
		81.56	**10.44**
		82.22	**10.78**
		75.74	**24.26**
		80.66	**25.34**

2	**Naive Bayes**	**75.33**	**24.67**
		75.95	**24.05**
		75.54	**24.46**
		75.80	**24.20**
		77.15	**22.85**
		68.85	**31.15**
		75.22	**24.78**
		76.59	**23.41**
		76.98	**23.02**

Table 3. Group Statistical Analysis of CNN and Naive Bayes. Mean, Standard Deviation and Standard Error Mean are obtained for 10 samples. Naive Bayes has higher mean accuracy and lower mean loss when compared to CNN.

Name	GROUP	N	Mean	Std.Deviation	Std.Error Mean
ACCURACY	**CNN**	**10**	**89.3130**	**.51461**	**.16273**
	Naive Bayes	**10**	**75.9030**	**.79247**	**.25060**
LOSS	**CNN**	**10**	**10.6870**	**.51461**	**.16273**
	Naive Bayes	**10**	**24.0970**	**.79247**	**.25060**

Table 4. Independent Sample T-test Naive Bayes is insignificantly better than CNN with p value 0.0176 (Two tailed, p>0.05)

Name	F	Sig.	t	df	Sig (2-tail	Mean Diffence ne	Std.Error differen ce	Lower	Upper

						ed)				
ACCURACY	**Equalvariances asmed**	**1.987**	**0.0176**	**44.879**	**18**	**000**	**13.41000**	**29880**	**12.78224**	**14.03776**
	EqualVariances not assumed	**-**	**-**	**44.879**	**15.444**	**000**	**13.41000**	**29880**	**12.77471**	**14.04529**
LOSS	**Equal variances assumed**	**-**	**0.0176**	**-44.879**	**18**	**000**	**-13.41000**	**29880**	**-14.03776**	**-12.78224**
	Equal Variances not assumed	**1.987**	**-**	**-44.879**	**15.444**	**000**	**-13.41000**	**29880**	**-14.04529**	**-12.77471**

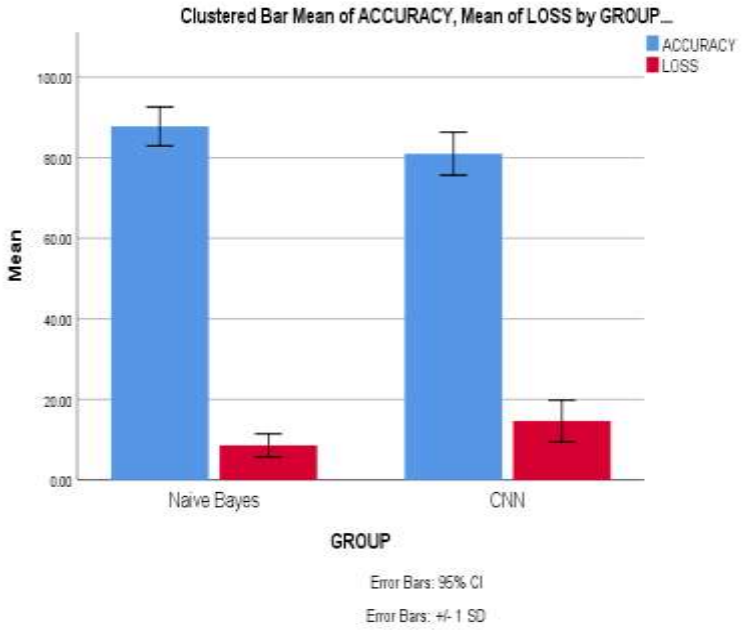

Fig. 1. Comparison of Naive Bayes and CNN classifier in terms of mean accuracy and loss. The mean accuracy of Naive Bayes is better than CNN classifier; Standard deviation of Naive Bayes is slightly better than CNN. X Axis: Naive Bayes Vs CNN Classifier and Y Axis: Mean accuracy of detection ± 1 SD.