Can One Deep Model Be Effective in Multiple Domain? a Case Study with Public Datasets

Subhra Dutta and Sushil Mandi

October 25, 2023

Can One Deep Model Be Effective in Multiple Domain? A Case Study with Public Datasets

Subhra Dutta[1], Sushil Mandi[2]
GIET University, Odisha, India[1,2]
su.man.2346@gmail.com, dutt.suvo_25@gmail.com

Abstract - Deep CNN models like VGG-16, Inception-v3 can be effective in multiple domains to some extent, but their effectiveness depends on several factors, including the specific domains involved, the complexity of the tasks within those domains, and the model's architecture and training data. In this paper we performed an empirical study on the effectiveness of a customized CNN model and tested its efficiency on multiple domains like epidemic disease prediction, NLP applications, and Education Technology. Three public datasets are identified from the review of literature of the existing works. It has been observed that smaller DCNN are more likely to perform diversely in different domains than larger models that are more robust in performance.

Keywords- Convolutional Neural Network, Deep Learning, Students performance prediction, law section prediction, Epidemic Disease prediction

1. Introduction

Deep learning, a subset of artificial intelligence, has witnessed explosive growth in recent years and has found its way into numerous domains, from healthcare to education. This work explores the impactful applications of deep learning models in three distinct areas: disease spread prediction, Natural Language Processing (NLP) based legal text analysis, and education technology. These applications have the potential to revolutionize their respective fields, improve decision-making, and enhance our understanding of complex phenomena.

Disease spread prediction is a critical area where deep learning models have demonstrated their potential to save lives and manage public health crises. Deep learning techniques are being leveraged to model, predict, and control the spread of contagious diseases like COVID-19. In epidemiological modelling, Deep learning models, such as recurrent neural networks (RNNs) and Long Short-Term Memory (LSTM) networks, can analyse large datasets of infection and transmission rates to generate more accurate epidemiological models. These models help public health officials and policymakers make data-driven decisions about intervention strategies. On the other hand, Convolutional neural networks (CNNs) can analyse medical imaging data, such as X-rays and CT scans, to aid in the early detection and diagnosis of diseases like tuberculosis, cancer, and COVID-19. This can significantly improve patient outcomes.

Deep learning has made significant inroads in the field of legal text analysis, revolutionizing how legal professionals process, manage, and extract valuable information from vast amounts of legal documents. Legal text analysis involves the use of deep learning techniques to perform tasks like contract analysis, legal research, case law prediction, and more. Deep learning models can be trained to extract key clauses, terms, and provisions from contracts. They can identify critical information, such as termination clauses, indemnification, and payment terms, streamlining the contract review process. Deep learning models can be employed to predict legal outcomes, such as case law predictions or the likelihood of success in litigation. By analyzing historical legal data, these models can provide valuable insights into the potential outcome of a legal dispute. Legal text analysis, empowered by deep learning, has the potential to streamline legal processes, reduce human error, and enhance the efficiency and accuracy of legal professionals in their work.

The integration of deep learning models in education technology is revolutionizing how students learn and educators teach. These applications are reshaping the classroom experience. Deep learning algorithms can adapt educational content to individual students' needs, pace, and learning styles. This personalized learning approach enhances student engagement and learning outcomes. Models can automate the grading of assignments and tests, reducing the administrative burden on educators and providing quicker, more consistent feedback to students. In addition, Deep learning models can predict student performance and identify at-risk students who may need additional support, enabling early intervention and retention strategies.

2. ML and DL Applications

there have been significant advancements in machine learning (ML) and deep learning (DL) in the fields of prediction, text analysis, and the development of corrective measures. There is a growing emphasis on making ML and DL models more interpretable. Researchers and practitioners are developing methods to detect and correct bias in machine learning models to ensure fair and equitable outcomes, particularly in applications like lending, hiring, and criminal justice.

2.1. Epidemic Disease Prediction

Scientific researchers across fields strive to contribute in their own unique ways. In addition to biological sciences like virology and medicine, ML and data science are being utilized to study and predict illness patterns [6]. Researchers are analyzing infection data from several nations using ML and statistical tools to find hidden patterns and anticipate the outbreak's growth and decline. Next, we cover new COVID-19 analysis and prediction research, primarily from India. Singh et al. [1] predicted India's COVID-19 pandemic with reduced social distance. This study proposes a mathematical model for viral dissemination in a community, taking into account social contact between ages. The social contact structure developed by Prem et al. [3] is used to study the effects of social distancing strategies such as workplace non-attendance, school closure, and lockdown.

Sustained lockdown with intermittent relaxation can reduce cases to a tolerable level, according to the authors. Gupta et al. [7] examined lockdown's significance in six social areas: restaurants, cafes, grocery stores, community parks, public transportation, offices, and residential areas. This study predicts future infections using exponential and polynomial regression, suggesting a significant decrease in the first five categories. This study also examines how lockdown, social isolation, and mass events affect the spread of infection. Mortality predictions were made using a decision tree model with a 60% accuracy in binary classification.

ARIMA model is the most used for time series modeling of COVID-19 data among researchers. Tandon et al. [5] suggested an ARIMA-based case prediction model. Initial parameters are determined using autocorrelation function (ACF) and partial autocorrelation (PACF) graphs. The time-series data is tested for normality and stationary variance using these models. ARIMA (2,2,2) was the most fitting model based on MAPE, MAD, and MSD scores. The model predicts a significant increase in infection cases by mid-May, followed by a possible decline. Chakraborty et al. [4] presented a two-fold approach: real-time COVID-19 case projections across nations and fatality rate prediction based on demographic and illness factors. A hybrid AIRAMA and wavelet-based forecasting model is utilized. It also uses a decision tree regression model to forecast fatality rate risk.

2.2. Legal Text analysis

Identifying relevant law sections and subsections for legal cases requires human expertise and effort to analyze articles, extract legal factors, and identify similarities from diverse historical cases. As of July 12, 2020, the National Judicial Data Grid (NJDG) reports 4.9 million pending cases throughout Indian courts. In addition to a paucity of judicial professionals, the disorganized nature of legal papers makes it challenging to find relevant legislation, leading to case delays [25]. Using AI to classify text can automate the process of identifying appropriate law parts in natural language case descriptions. Researchers have applied NLP and ML approaches to several legal scenarios in recent years. Virtucio et al. [8] classified Philippines Supreme Court rulings using NLP and SVM. Results show linear SVM achieved 45% accuracy on n-gram datasets and 55% on topic datasets. Francesconi et al. [9] developed a model to categorize legislative text fragments into provision kinds. The study utilized Naive Bayes (NB) and multiclass SVM machine learning models, yielding satisfactory results with 88.66% accuracy for NB and 92.44% for SVM. Waltl et al. [12] performed research on German tax law cases, employing several ML classifiers to predict case outcomes. The collection consists of 44,285 German fiscal court judgments from 1945 to 2016. The suggested method involves preprocessing, TF-IDF vectorization, and a classifier model. The Naive Bayes classifier was shown to be the most effective among several estimators. An empirical analysis by Aletras et al. [13] predicted the outcome of European Court of Human Rights cases using language content. The authors used an SVM model for binary classification. Textual material is modeled as an N-gram sequence using TF-IDF, and a binary classifier is used to check for infringement of three articles. The model achieved an average accuracy of 79% across all examples. Recently, Medvedeva et al. [14] conducted an experiment to predict if an article will be violated or not. They studied European Court of Human Rights court proceeding texts using NLP tools and an SVM classifier for prediction. This study predicted 9 article violations with an average accuracy of 75%.

2.3. Education Technology

Many academics have utilized ML models to predict student performance from student data with or without feature analysis [20,25,29]. Han et al. [16] compared ML models for predicting Chinese undergraduate GPA grades. They developed a dataset of 123 undergraduates in four-year programs with 20 professional core courses. Using correlation and association among courses for feature selection, the AdaBoost classifier has the maximum prediction accuracy of 91.67 %. Anuradha et al. [17] constructed a dataset from three Tamil Nadu private colleges' student records. They used the previous semester's marks and demographic and pre-collegiate variables to predict end-semester success. The KNN classifier has the greatest accuracy of 68.3% in Weka testing. Osmanbegović et al. [19] aimed to identify key parameters influencing student performance prediction. They created a private dataset of 1210 Bucharest secondary school students' records with 19 attributes. They had the greatest RF classifier accuracy of 73.2%. A dataset of 403 students with 14 variables from Kolkata undergraduate colleges was prepared by Acharya et al. [21]. They chose features using chi-square and Information Gain. With Sequential Minimal Optimization (SMO), SVM has the best classification accuracy at 66%. Amra et al. [22] used KNN and NB classifier to predict student performance. This study uses Gaza Strip secondary school students' records. Manual preprocessing removed characteristics. NB had the highest forecast accuracy of 93.6%. Wrapper-based and correlation-based feature selection were used by Jalota et al. [24]. An institutional LMS users-log of 480 records with 14 attributes was used to create the dataset. SVM and DT fared better with correlation-based filters, but NB did best with wrapper-based feature selection.

However, other predictive models used all dataset attributes without feature selection. Demographic data such category, gender, and 10th and 12th grade performance was reported by Kabra et al. [23] for 346 engineering students at an Indian institute. The prediction accuracy was 69.94% with a single DT classifier. Devasia et al. [26] used 700 student records and 19 attributes from Indian universities. This study found that the NB classifier predicted student performance better than other models. A dataset of 300 individuals from Indian degree colleges was created by Bhardwaj et al. [27]. It has the greatest NB classifier prediction accuracy of 86.25%. Three ML classifiers were used by Pandey et al. [28] to predict student performance from social and educational background. The collection included 600 student records from Indian colleges. Voted predictive model aggregations had the highest accuracy of 87.03%. Abdullah et al. [30] presented multi-agent data mining for student performance prediction. One course's 155 student records were used in the study. Their comparison of DT with AdaBoost showed that the ensemble technique with 80% accuracy outperformed the single classifier with 74% accuracy. Most studies on student performance prediction are classification tasks, although others try to predict numeric marks or grade points, which is a regression problem.

3. Experimental study

A Customized CNN is designed for classification tasks and consists of four convolutional blocks, each featuring two convolution layers followed by an intermediate max-pooling layer. After these convolution blocks, there are two fully connected (dense) layers with 256 and 64 nodes, respectively, before the final output layer with a sigmoid activation function. The first convolution block starts with an input layer. It has two convolutional layers with 32 kernels each. Each kernel in these layers has a size of 4x4 pixels. Activation functions Leaky ReLU is applied to the output of each convolution layer. After the two convolution layers in the first block, a max-pooling layer is introduced. Max-pooling helps reduce the spatial dimensions of the feature maps and capture the most important information. Total eight such convolution blocks are engaged. After the convolution blocks, there are two dense layers. Each dense layer consists of 1024 nodes.

Layer Type	Kernel Size & Number	Stride	Output Shape
input	..	..	(512, 512, 3)
convolution	4 x 4 x 32	2	(256, 256, 32)
convolution	4 x 4 x 32	1	(255, 255, 32)
max-pooling	3 x 3	2	(127, 127, 32)
convolution	4 x 4 x 64	2	(62, 62, 64)
convolution	4 x 4 x 64	1	(63, 63, 64)
max-pooling	3 x 3	2	(31, 31, 64)
convolution	4 x 4 x 128	1	(32, 32, 128)
convolution	4 x 4 x 128	1	(33, 33, 128)
max-pooling	3 x 3	2	(16, 16, 128)
convolution	4 x 4 x 256	1	(17, 17, 256)
max-pooling	3 x 3	2	(8, 8, 256)
convolution	4 x 4 x 384	1	(9, 9, 384)
max-pooling	3 x 3	2	(4, 4, 384)
convolution	4 x 4 x 512	1	(5, 5, 512)
max-pooling	3 x 3	2	(2, 2, 512)
fully connected	1024	..	(1024)
fully connected	1024	..	(1024)
fully connected	1	..	(1)

Figure 1. CNN Model Architecture

This architecture provides a strong foundation for image classification tasks, with the convolutional blocks capturing hierarchical features from the input data and the dense layers learning high-level representations before making a final classification decision through the sigmoid output layer. The choice of activation functions and optimization algorithms can further affect the network's performance, and you may need to fine-tune these based on the specific problem.

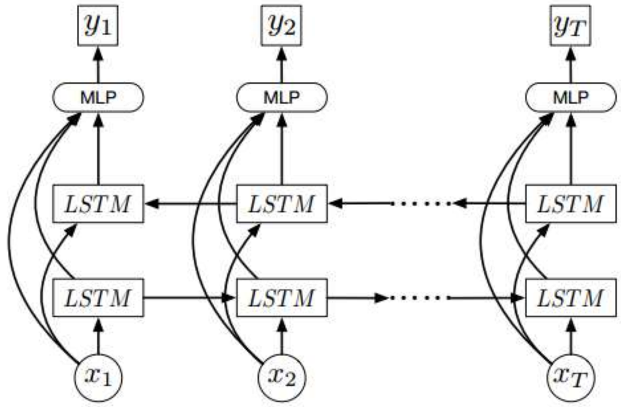

Figure 2. Stacked LSTM architecture

A stacked Long Short-Term Memory (LSTM) architecture is a type of recurrent neural network (RNN) model that involves stacking multiple LSTM layers on top of each other to capture complex sequential patterns and dependencies within a sequence data. Each LSTM layer consists of a set of memory cells and gates that allow the network to learn and remember information over different time steps. In a stacked LSTM architecture, multiple LSTM layers are arranged in a sequential manner, with each layer feeding into the next. The output of the first LSTM layer serves as the input to the second LSTM layer, and so on. The number of LSTM layers can be adjusted depending on the complexity of the task and the depth of temporal dependencies in the data. Generally, using more layers allows the model to capture more intricate patterns in the sequence data.

4. Results and Analysis

Table 1 provides a comparison of results between two different models, VGG-16 and a customized deep learning model (DL)/Stacked LSTM, on three different datasets for various tasks and evaluation measures. For COVID-19 Dataset in Regression Task VGG-16 achieved an RMSE of 0.4526. The Customized DL/Stacked LSTM model achieved a higher RMSE of 0.7632. In this case, VGG-16 outperformed the Customized DL/Stacked LSTM model, as a lower RMSE indicates better accuracy in regression tasks. The VGG-16 model seems to be better at predicting numerical values in the COVID-19 dataset.

For NLP based classification task on LegalCrystal Repository, VGG-16 achieved an F1 micro average of 0.7825. The Customized DL/Stacked LSTM model achieved a lower F1 micro average of 0.5845. Here, VGG-16 outperformed the Customized DL/Stacked LSTM model in terms of F1 micro average, which suggests it is better at classification tasks on the LegalCrystal repository dataset. It achieved a higher balance between precision and recall. On the other hand, for xAPI Dataset VGG-16 achieved an accuracy of 0.8676. The Customized DL/Stacked LSTM model achieved a lower accuracy of 0.7890. In the case of the xAPI dataset, again, VGG-16 outperformed the Customized DL/Stacked LSTM model in terms of accuracy. VGG-16 achieved a higher accuracy score, which indicates a better performance in classifying data in this dataset.

Table 1. Comparison of results

Dataset	Task	Measure	VGG-16	Customized DL/ Stacked LSTM
COVID-19 Dataset	Regression	RMSE	0.4526	0.7632
LegalCrystal repository	Classifaction	F1 micro average	0.7825	0.5845
xAPI Dataset	Classiifcation	Accuracy	0.8676	0.7890

In general, VGG-16 appears to be the better model for these particular datasets and tasks, outperforming the Customized DL/Stacked LSTM model in all three cases. However, it's important to consider that the choice of model depends on the specific dataset and task, and further analysis may be needed to determine the most suitable model for a given problem. Additionally, hyperparameter tuning and model architecture adjustments might improve the performance of the Customized DL/Stacked LSTM model if deemed necessary. In addition, it is also observed that large model like VGG-16 is more robust in performance across different domains. Whereas, the performance of the customized CNN varies largely across multiple domains.

5. Conclusion

This research presents an empirical investigation into the efficacy of a tailored convolutional neural network (CNN) model, examining its performance across many areas including disease prediction, natural language processing (NLP) applications, and education technology. It is found that, smaller DCNNs, with their limited capacity and better generalization properties, are more likely to perform diversely in different domains. They can provide solid performance on various tasks, making them a more versatile choice when the goal is to work across different datasets and domains. In contrast, larger DCNNs are better suited for cases where domain-specific learning and maximizing performance on a particular dataset are of primary concern, but they might not adapt well to new, diverse domains due to their greater risk of overfitting. The choice between smaller and larger models should be made based on the specific requirements and characteristics of the problem at hand.

References

1. Singh, R., & Adhikari, R. (2020). Age-structured impact of social distancing on the COVID-19 epidemic in India. arXiv preprint arXiv:2003.12055.
2. S. Sengupta, "Towards Finding a Minimal Set of Features for Predicting Students' Performance Using Educational Data Mining", I.J. Modern Education and Computer Science, vol. 3, pp. 44-54, Jun. 2023
3. Prem, K., Cook, A. R., & Jit, M. (2017). Projecting social contact matrices in 152 countries using contact surveys and demographic data. PLoS computational biology, 13(9), e1005697.
4. Chakraborty, T., & Ghosh, I. (2020). Real-time forecasts and risk assessment of novel coronavirus (COVID-19) cases: A data-driven analysis. Chaos, Solitons & Fractals, 135, 109850.
5. Tandon, H., Ranjan, P., Chakraborty, T., & Suhag, V. (2022). Coronavirus (COVID-19): ARIMA-based time-series analysis to forecast near future and the effect of school reopening in India. Journal of Health Management, 24(3), 373-388.
6. Sengupta, S. (2020). Forecasting the peak of covid-19 daily cases in india using time series analysis and multivariate lstm. EasyChair Preprint.
7. Gupta, R., Pal, S. K., & Pandey, G. (2020). A comprehensive analysis of COVID-19 outbreak situation in India. MedRxiv, 2020-04.
8. Virtucio, M. B. L., Aborot, J. A., Abonita, J. K. C., Avinante, R. S., Copino, R. J. B., Neverida, M.P., & Tan, G. B. A. (2018). Predicting decisions of the Philippine Supreme Court using natural language processing and machine learning. In 2018 IEEE 42nd annual computer software and applications conference (COMPSAC) (Vol. 2, pp. 130–135). IEEE.

9. Francesconi, E., & Passerini, A. (2007). Automatic classification of provisions in legislative texts. Artificial Intelligence and Law, 15(1), 1–17.
10. Islam, M. A., & Haque, M. J. (2018). Evaluating document analysis with KNN based approaches in judicial offices of Bangladesh. In 2018 second international conference on computing methodologies and communication (ICCMC) (pp. 646–650). IEEE.
11. Liu, Z., & Chen, H. (2017). A predictive performance comparison of machine learning models for judicial cases. In 2017 IEEE symposium series on computational intelligence (SSCI) (pp. 1–6).IEEE.
12. Waltl, B., Bonczek, G., Scepankova, E., Landthaler, J., & Matthes, F. (2017). Predicting the outcome of appeal decisions in Germany's tax law. In International conference on electronic participation (pp. 89–99). Cham: Springer.
13. Aletras, N., Tsarapatsanis, D., Preoţiuc-Pietro, D., & Lampos, V. (2016). Predicting judicial decisions of the European Court of Human Rights: A natural language processing perspective. PeerJ Computer Science, 2, e93.
14. Medvedeva, M., Vols, M., & Wieling, M. (2020). Using machine learning to predict decisions of the European Court of Human Rights. Artificial Intelligence and Law, 28(2), 237–266.
15. Sengupta, S., & Dave, V. (2021). Predicting applicable law sections from judicial case reports using legislative text analysis with machine learning. Journal of Computational Social Science, 1-14.
16. Han, M., Tong, M., Chen, M., Liu, J., & Liu, C. (2017, July). Application of ensemble algorithm in students' performance prediction. In 2017 6th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI) (pp. 735-740). IEEE.
17. Anuradha, C., & Velmurugan, T. (2015). A comparative analysis on the evaluation of classification algorithms in the prediction of students performance. Indian Journal of Science and Technology, 8(15), 1-12.
18. Ismail, L., Materwala, H., & Hennebelle, A. (2021, February). Comparative Analysis of Machine Learning Models for Students' Performance Prediction. In International Conference on Advances in Digital Science (pp. 149-160). Springer, Cham.
19. Osmanbegović, E., Suljić, M., & Agić, H. (2014). Determining dominant factor for students performance prediction by using data mining classification algorithms. Tranzicija, 16(34), 147-158.
20. Sengupta, S., & Dasgupta, R. (2015). Using Semiformal and Formal Methods in Software Design: An Integrated Approach for Intelligent Learning Management System. Applied Computation and Security Systems: Volume Two, 53-65.
21. Acharya, A., & Sinha, D. (2014). Early prediction of students performance using machine learning techniques. International Journal of Computer Applications, 107(1).
22. Amra, I. A. A., & Maghari, A. Y. (2017, May). Students performance prediction using KNN and Naïve Bayesian. In 2017 8th International Conference on Information Technology (ICIT) (pp. 909-913). IEEE.
23. Kabra, R. R., & Bichkar, R. S. (2011). Performance prediction of engineering students using decision trees. International Journal of computer applications, 36(11), 8-12.
24. Jalota, C., & Agrawal, R. (2021). Feature selection algorithms and student academic performance: A study. In International Conference on Innovative Computing and Communications (pp. 317-328). Springer, Singapore.
25. Datta, S., & Sengupta, S. (2018). A Review on the Adaptive Features of E-Learning. International Journal of Learning and Teaching, 4(4), 277-284.
26. Devasia, T., Vinushree, T. P., & Hegde, V. (2016, March). Prediction of students performance using Educational Data Mining. In 2016 International Conference on Data Mining and Advanced Computing (SAPIENCE) (pp. 91-95). IEEE.
27. Bhardwaj, B. K., & Pal, S. (2012). Data Mining: A prediction for performance improvement using classification. arXiv preprint arXiv:1201.3418.
28. Pandey, M., & Taruna, S. (2016). Towards the integration of multiple classifier pertaining to the student's performance prediction. Perspectives in Science, 8, 364-366.

29. Sengupta, S. (2022). Possibilities and challenges of online education in India during the COVID-19 pandemic. International Journal of Web-Based Learning and Teaching Technologies (IJWLTT), 17(4), 1-11.
30. Abdullah, A. L., Malibari, A., & Alkhozae, M. (2014). STUDENTS PERFORMANCE PREDICTION SYSTEM USING MULTI AGENT DATA MINING TECHNIQUE. International Journal of Data Mining & Knowledge Management Process, 4(5), 1.