Deep Neural Network for the Automatic Classification of Vertebral Column Disorders

Menaka Ramasamy [#1], Sabenabanu Abdulkadhar [#2], Jeyakumar Natarajan[*#3]

[#]*Data Mining and Text Mining Laboratory*
Department of Bioinformatics, Bharathiar University, Coimbatore-641046,Tamilnadu,India
[1]rsmenakamsc@gmail.com
[2]sabenabanu.a@hotmail.com
[3*]n.jeyakumar@yahoo.co.in

***Abstract*- In the human body, the vertebral column consists of vertebras, nerves, invertebrate discs, medulla, joints, and muscles, which provides support for body and movement axle. Dysfunction to any of the above components in this complex system creates disorders like Disc hernia and Spondylolisthesis. Manually classifying these disorders is a difficult task. Recently Machine learning (ML) techniques were applied in automating the vertebral column disorder classification. In this work, we applied Deep Neural Network (DNN) to classify the vertebral column dataset with three classes (Normal, Disk Hernia, and Spondylolisthesis). The vertebral column dataset was collected from the UCI machine learning database repository and has 310 records for training and testing with six biomechanical attributes. The classification accuracy and F-score for the DNN classifier in the vertebral column dataset is 85% and 83% respectively. Comparison with existing ML systems shows that our DNN based classification approach exhibits promising results.**

***Keywords*- Vertebral column; Classification; Machine Learning; Deep Neural Network.**

I. Introduction

Biomedical research is increasingly dependent on the automatic analysis of the databases and literature to determine correlations and interactions amongst biomedical entities, functional roles, phenotypic traits and disease states [1],[2]. Vertebral column or the spine is a complex system composed of vertebrae, disk, nerves and its associated muscles. Several types of abnormalities in the spine can result either due to congenital reasons or as a result of poor posture or unequal muscle pull [3]. While the disc region provides a cushioning effect for mechanical loads, the vertebral body bears the load as well as provides a protective case to the spine [4].
Classifying defects in spinal card is a challenging task and it requires a more skill of an experienced radiologist for analyzing MRI and CT images [5],[6]. Different types of computational approaches were used to classify the spinal disorders [7],[8]. The most commonly faced issue in classifying spinal disorders is the class imbalance problem due to the number of subjects, accessibility to the scans, feature extraction, and number of features [9],[10]. Traditional machine learning approaches such as Support Vector Machines (SVM), Radial Basis Function (RBF), etc., required a vast amount of features and domain knowledge to classify spinal disorders [11],[12]. Designing features manually is a time-consuming process. To overcome this problem automated feature extraction techniques such as deep learning, deep neural networks etc., gained popularity in recent times [13]. Deep learning is one such a technique to extract the features without any supervision [14]. Deep learning finds convoluted structure in vast informational resources by utilizing the back propagation algorithm to demonstrate how a machine should change its interior parameters that are utilized to figure out the representation in each layer from the representation in the previous layer [15]. Recently text mining and natural language processing (NLP) researchers have developed different models of deep learning architectures for classification such as Convolution Neural Networks (CNN) [16], Recursive Neural Networks (RNN)[17], Recurrent Convolution Neural Networks (RCNN)[13], Deep Neural Networks (DNN) [18].
In this study, we employed a DNN based approach to classify the vertebral column dataset. In the remaining sections, section 2 explains the proposed materials and methods. Section 3 details the results and discussions. Section 4 shortly concludes our work and future perspectives.

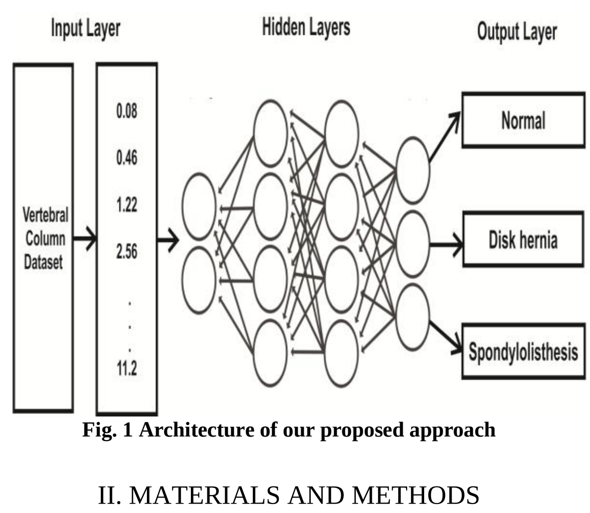

Fig. 1 Architecture of our proposed approach

II. MATERIALS AND METHODS

We present a deep learning-based architecture to address the vertebral column disease classification problem. Our method focuses on deep neural network models to classify the datasets. Fig. 1 depicts the architecture of our system.

A. Data Pre-processing and Cleaning

Data pre-processing includes the removal of zero values, empty values and unnamed or NaN columns from the datasets.

Deep Neural Networks (DNN's)

Deep neural networks use efficient mathematical modeling to process data in complex ways. It employs unsupervised feature learning can learn discriminative and effective features from a large amount of unlabeled data [15],[18],[19]. A deep neural network is a neural network with a specific degree of unpredictability, a neural network with multiple layers. The DNN based vertebral column disease classification problem is designed by following.

Let we have assume the disease classification proposed method as Θ = (N, $x_1,x_2,x_{3.....}x_n$). Here each input sequence of vertebral column disease was considered independently. Given an input samples, the network with parameter Θ outputs the vector$o$, where the ith component o_i contains the score for disease instance. To attain the conditional probability $p(i|x,\theta)$, then, we have applied the softmax operations overall disease types.

$$p(i|x,\theta) = \frac{e^{o_i}}{\sum_{k=1}^{n^4} e^{o_k}} \qquad (1)$$

Given all training examples T=$(x^i; y^i)$ we can calculate the log-likelihood of the parameters as follows

$$j(\theta) = \sum_{i=1}^{T} \log p(y^i | x^i, \theta) \qquad (2)$$

To compute network parameter Θ, we maximize the log-likelihood $j(\theta)$ by employing a simple optimization technique called stochastic gradient descent (SGD). The (N, $x_1,x_2,x_{3.....}x_n$)are randomly initialized because the parameters are in different layers of neural networks. We implement the back-propagation algorithm, the differentiation chain rule is applied through the network until the softmax layer is reached by iteratively selecting an example (x,y) and applying the following the update rule.

$$\theta \leftrightarrow \theta + \lambda \frac{\sigma \log p(y|x,\theta)}{\sigma\theta}$$

III.RESULTS AND DISCUSSIONS

A. Dataset and Evaluation

The dataset was retrieved from UCI (University of California, Irvine) machine learning database [20]. It contains 310 instances and each containing six features named pelvic incidence, pelvic tilt numeric, lumbar lord sis angle, sacral slope, pelvic radius and degree Spondylolisthesis [21],[22]. The dataset was organized into two types of classification tasks. The first task consists in classifying patients as belonging to one out of three categories: Normal (100 patients), Disk Hernia (60 patients) or Spondylolisthesis (150 patients). The second task is the categories Disk Hernia and Spondylolisthesis that were merged into a single category labeled as 'abnormal'. Thus, the second task consists in classifying patients as belonging to one out of two categories: Normal (100 patients) or Abnormal (210 patients). Our study focused on the first type of classification. The partition of vertebral column dataset includes 50-50 split, which means that 50% dataset used for training, 50% dataset used for testing. Tables I AND II represents Biomechanical features used for classification of the vertebral column, Dataset splitting. The standard evaluation metrics such as Precision (P), Recall(R), F-score (F), and Accuracy were used to evaluate the performance of the proposed approach [23].

Table I: Biomechanical features used for classification of vertebral column

SNO	Features
1	Pelvic incidence
2	Pelvic tilt
3	Lumbar lordosis angle
4	Sacral slope
5	Pelvic radius
6	Degree-spondylolisthesis

Table II: Dataset splitting

Dataset	Count
Train	50% (155 instances)
Test	50% (155 instances)

Table III depicts the evaluation results of the vertebral column dataset with 50-50 partition. It achieved 0.82 precision, 0.84 recall, 0.83 F-score and 0.85 accuracy. Furthermore evaluation, we implemented 10-fold cross validation on our proposed approach. Table IV depicts the results of 10-fold cross validation. While comparing Tables III and IV our proposed approach outperforms on both 50-50 partition and 10-fold cross validation.

Table III: Evaluation results of vertebral column dataset with 50-50 partition

Approach	Data set (50-50)			
DNN	**P**	**R**	**F**	**Accuracy**
	0.82	0.84	0.83	0.85

Table IV: 10-fold cross validation result on vertebral column dataset

Approach	Data set (10-fold cross validation)			
DNN	**P**	**R**	**F**	**Accuracy**
	0.80	0.75	0.77	0.79

For more evaluation, we deployed the correlation analysis on features. The correlation analysis was used to find the linear relationship between two numerical features in the vertebral column dataset. The high correlation point has a strong relationship between the features and the low correlation point represents a weak relationship between the features. Fig. 2 represents the correlation analysis between the features. By analyzing Fig. 2, four most informative features were pelvic incidence, lumbar lordosis angle, sacral slope, degree-spondylolisthesis based on the correlation score presented in figure. These features are frequently participated in improving the classification accuracy of vertebral column disorders. These four features had a high correlation between x and y axioms and improve the overall accuracy of the DNN classifier.

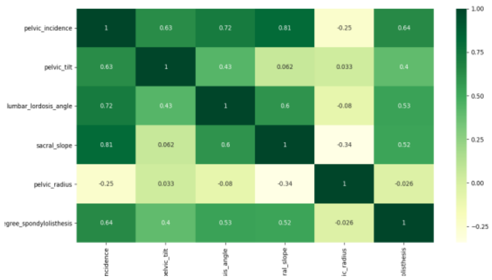

Fig. 2 Correlation analysis on features

B. Comparative Analysis

We compared our approach with previously developed work of Yavuz et al., which uses three classifiers SVM, KNN, and RBF-NN [3]. Table V represents the results of both approaches in terms of accuracy, precision and F-score.

Table V: Comparative analysis of our approach with other approaches

S.NO	System	Method	Accuracy	Precision	F-score
1	Yavuz et al.,[3]	SVM	0.66	0.47	0.59
		KNN	0.75	0.61	0.60
		RBF-NN	0.82	0.70	0.71
2	**Ours**	**DNN**	**0.85**	**0.82**	**0.83**

IV. CONCLUSIONS

In this study, we proposed a deep neural network model to classify a vertebral column dataset. DNN perfectly overcomes overfitting problem occurred in traditional approaches. The main advantage of the proposed model is that it does not need any handcrafted features for classifying the vertebral column dataset and it automatically acquires related features from the corresponding dataset. The DNN architecture effectively performed in multitask learning because of its efficient optimization. To the best of our knowledge, we are the first one to introduce the DNN for classifying the vertebral column dataset. This method is not only for the lumbar disc diseases but also used for all kinds of disease classifications. In future, we plan to apply the advanced attribute weighting techniques to the dataset.

REFERENCES

[1] Jensen, Lars Juhl, Jasmin Saric, and Peer Bork. "Literature mining for the biologist: from information retrieval to biological discovery." Nature reviews genetics 7.2 (2006): 119.

[2] Deans, Andrew R., et al. "Finding our way through phenotypes." PLoS biology 13.1 (2015): e1002033.

[3] Unal, Yavuz, Kemal Polat, and H. Erdinc Kocer. "Classification of vertebral column disorders and lumbar discs disease using attribute weighting algorithm with mean shift clustering." Measurement 77 (2016): 278-291.

[4] Athertya, Jiyo S., and Gurunathan Saravana Kumar. "Sensitivity Analysis on Effect of Biomechanical Factors for Classifying Vertebral Deformities." International Conference on Soft Computing and Pattern Recognition. Springer, Cham, 2016.

[5] Wang, Shijun, and Ronald M. Summers. "Machine learning and radiology." Medical image analysis 16.5 (2012): 933-951.

[6] Solanki, Guirish A., et al. "Spinal involvement in mucopolysaccharidosis IVA (Morquio-Brailsford or Morquio A syndrome): presentation, diagnosis and management." Journal of inherited metabolic disease 36.2 (2013): 339-355.

[7] Prasetio, Rizki Tri, and Dwiza Riana. "A comparison of classification methods in vertebral column disorder with the application of genetic algorithm and bagging." 2015 4th international conference on instrumentation, communications, information technology, and biomedical engineering (ICICI-BME). IEEE, 2015.

[8] Ansari, Sana, et al. "Diagnosis of vertebral column disorders using machine learning classifiers." 2013 International Conference on Information Science and Applications (ICISA). IEEE, 2013.

[9] Ravì, Daniele, et al. "Deep learning for health informatics." IEEE Journal of biomedical and health informatics 21.1 (2016): 4-21.

[10] Liu, Feng, et al. "Classification of different therapeutic responses of major depressive disorder with multivariate pattern analysis method based on structural MR scans." PloS one 7.7 (2012): e40968.

[11] Siuly, Siuly, and Yanchun Zhang. "Medical big data: neurological diseases diagnosis through medical data analysis." Data Science and Engineering 1.2 (2016): 54-64.

[12] Rahman, Md Mahmudur, Bipin C. Desai, and Prabir Bhattacharya. "Medical image retrieval with probabilistic multi-class support vector machine classifiers and adaptive similarity fusion." Computerized Medical Imaging and Graphics 32.2 (2008): 95-108.

[13] Abdulkadhar, Sabenabanu, Gurusamy Murugesan, and Jeyakumar Natarajan. "Recurrent convolution neural networks for classification of protein-protein interaction articles from biomedical literature." 2017 Third International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN). IEEE, 2017.

[14] Sun, Yi, Xiaogang Wang, and Xiaoou Tang. "Deep learning face representation from predicting 10,000 classes." Proceedings of the IEEE conference on computer vision and pattern recognition. 2014.

[15] LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." nature 521.7553 (2015): 436.

[16] Abdel-Hamid, Ossama, et al. "Applying convolutional neural networks concepts to hybrid NN-HMM model for speech recognition." 2012 IEEE international conference on Acoustics, speech and signal processing (ICASSP). IEEE, 2012.

[17] Luong, Thang, Richard Socher, and Christopher Manning. "Better word representations with recursive neural networks for morphology." Proceedings of the Seventeenth Conference on Computational Natural Language Learning. 2013.

[18] Deng, Li, Geoffrey Hinton, and Brian Kingsbury. "New types of deep neural network learning for speech recognition and related applications: An overview." 2013 IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, 2013.

[19] Zeng, Daojian, et al. "Relation classification via convolutional deep neural network." (2014).

[20] Newman, David J., et al. "UCI repository of machine learning databases, 1998." (2008).

[21] Hamodi, Yaser Issam. "Rule Extraction in Diagnosis of Vertebral Column Disease." International Journal on Recent and Innovation Trends in Computing and Communication 3.3 (2015): 1120-1124.

[22] Huang, M., Y. Hung, and D. Liu. "Diagnostic prediction of vertebral column using rough set theory and neural network technique." Inf Technol J 13.5 (2014): 874-884.
[23] Sokolova, Marina, Nathalie Japkowicz, and Stan Szpakowicz. "Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation." Australasian joint conference on artificial intelligence. Springer, Berlin, Heidelberg, 2006.