EJSIT

European Journal of Science, Innovation and Technology

ISSN: 2786-4936

www.ejsit-journal.com

Volume 4 | Number 5 | 2024

Ensemble Learning and Stacking Classifier for Effective Detection of Lumbar Spondylolisthesis in Spinal Imaging

Dr. Deepika Saravagi
Assistant Professor, Patkar Varde College,
Goregaon, Mumbai, Maharashtra, India

ABSTRACT

Lumbar spondylolisthesis, defined by the displacement of one vertebra over another, requires accurate diagnosis to inform treatment approaches and prevent any consequences. Ensemble learning methods in machine learning models have shown promise in improving classification accuracy. This study analyzes the utilization of diverse ensemble learning models, including Bagging, Random Forest, K-Nearest Neighbors (KNN), and Stacking Classifiers, for the identification of lumbar spondylolisthesis. The Stacking Classifier had exceptional performance, attaining an accuracy of 97%, above that of both individual models and homogeneous ensemble techniques. The results demonstrate that ensemble learning, especially stacking, is an effective technique for identifying lumbar spondylolisthesis, offering a reliable tool for clinical decision-making.

Keywords: Lumbar Spondylolisthesis, Ensemble Learning, Stacking Classifier, Bagging, Random Forest, KNN, Machine Learning

INTRODUCTION

Lumbar spondylolisthesis is a degenerative spinal disorder characterized by the anterior displacement of one vertebra over the next vertebra, leading to spinal instability. This disorder, affecting a significant portion of the elderly population, may result in persistent lower back discomfort, numbness, and possible neurological impairments if not properly managed. Timely and accurate detection is crucial, as it enables swift intervention and tailored treatment techniques that can improve patient outcomes and overall quality of life. The diagnosis of lumbar spondylolisthesis now relies on radiographic evaluation, requiring specialist knowledge and experience. However, manual evaluation may demonstrate subjectivity, necessitate considerable time commitment, and be prone to inter-observer variability, potentially leading to variations in diagnosis and treatment recommendations.

In recent years, artificial intelligence (AI) and machine learning (ML) have emerged as crucial tools in medical imaging and diagnostics, offering automated, objective, and rapid analytical capabilities. Machine learning techniques, particularly in supervised learning, have increasingly being utilized for the classification of medical pictures, including radiography and MRI scans. These models facilitate the detection, categorization, and segmentation of anatomical features and anomalies, hence improving radiologists' capacity to render more precise and efficient diagnoses. Ensemble learning approaches in machine learning are widely preferred for their ability to combine several basic classifiers, enhancing model performance through collaborative learning.

Ensemble learning methods, such as Bagging, Boosting, and Stacking, proficiently address the shortcomings of individual classifiers by consolidating their predictions to improve accuracy, reduce variation, and promote generalizability. Bagging (Bootstrap Aggregating) reduces variance by training several instances of a base classifier on different subsets of the data, while Boosting focuses on successively correcting misclassifications to lower bias. Stacking employs a distinct approach by amalgamating multiple classifiers and leveraging a meta-learner to unify their predictions. Stacking is particularly beneficial for complex classification tasks, since it leverages the characteristics of each classifier, yielding a more robust and accurate predictive model.

Ensemble learning provides a robust method for improving diagnostic precision in identifying lumbar spondylolisthesis. The amalgamation of models such as Random Forest, K-Nearest Neighbors (KNN), and Naive Bayes, which have demonstrated consistent efficacy in medical imaging, via stacking can yield enhanced accuracy. The Stacking Classifier has shown potential in medical diagnostics by amalgamating the outputs of multiple base models, adeptly capturing both linear and non-linear correlations in the data. This methodology is especially advantageous in scenarios with complex picture information, where a single classifier may struggle to accurately identify patterns related to spondylolisthesis.

This study investigates the effectiveness of various ensemble models, including Bagging, Random Forest, KNN, Naive Bayes, and a Stacking Classifier, in the accurate identification of lumbar spondylolisthesis. The study aims to evaluate the efficacy of ensemble learning, namely stacking, in enhancing diagnostic precision for lumbar spondylolisthesis by effectively integrating the strengths of diverse algorithms. This study seeks to assess the efficacy of different models to identify the most dependable and precise technique for diagnosing this spinal disease, thereby providing a prospective tool for clinical use.

OBJECTIVES

The primary objectives of this study are:

1. To implement and compare the performance of ensemble learning models for detecting lumbar spondylolisthesis.
2. To evaluate the accuracy of Bagging, Random Forest, KNN, Naive Bayes, and Stacking Classifiers.
3. To determine if the Stacking Classifier improves detection accuracy over individual classifiers and other ensemble techniques.
4. To assess the feasibility of ensemble learning models for clinical applications in spine health.
5. To identify the most effective model for reliable, automated lumbar spondylolisthesis detection.

LITERATURE REVIEW

Lumbar spondylolisthesis is a spinal disorder that compromises stability and functionality, presenting a range of symptoms from asymptomatic to severe pain and neurological impairments (Altaf et al., 2007). The precise diagnosis of this ailment is challenging, partly due to reliance on traditional radiographic assessments, which might differ and are frequently influenced by doctors' subjective interpretations (Matsunaga & Sakou, 2000). The limitations of conventional diagnostic techniques have prompted the exploration of machine learning applications in spinal imaging. Automated methods provide a more objective and consistent assessment of medical pictures, demonstrating promising results in the detection of various spinal disorders, including scoliosis, stenosis, and fractures (Zhou et al., 2019).

Ensemble learning has emerged as a powerful technique in medical imaging, wherein models amalgamate the predictive strengths of multiple algorithms to improve accuracy, robustness, and generalization. Bagging, or Bootstrap Aggregation, is a prevalent technique that reduces model variance by training on random data subsets, particularly advantageous for noisy datasets (Breiman, 2001). Random Forest, an ensemble method based on bagging, has shown remarkable efficacy in spine imaging applications due to its resilience against overfitting (Gangeh et al., 2018).

Stacking, an ensemble method, has demonstrated enhanced efficacy in medical diagnostics by amalgamating outputs from many base classifiers through a meta-classifier that synthesizes the predictions (Ting & Witten, 1999). Stacking is especially advantageous in complex classification tasks, as it integrates models characterized by high variance and high bias to improve overall performance (Sill et al., 2009). Decision trees, frequently utilized in stacking, offer clarity and comprehensibility, which are crucial in healthcare settings. However, their independent application may demonstrate deficiencies in robustness, an issue that stacking effectively addresses (Quinlan, 1986). Alternative algorithms such as Support Vector Machines (SVM) have proven effective in binary classification tasks; however, they require careful tuning for use in medical imaging (Vapnik, 1998).

Gradient Boosting, shown by models like XGBoost, is an ensemble technique that improves classification precision by systematically correcting misclassified cases, yielding reliable results in medical applications (Chen & Guestrin, 2016). Comparative analyses of ensemble methods and individual classifiers demonstrate that ensemble strategies regularly surpass single models by diminishing error variance and improving stability, which is crucial for dependable diagnosis in healthcare (Dietterich, 2000).

Automated detection techniques for spine anomalies, such as scoliosis and kyphosis, illustrate the capabilities of machine learning in radiology, particularly in reducing the variability associated with traditional radiographic analysis (Lao et al., 2017; Patel et al., 2015). In diagnosing lumbar spondylolisthesis, ensemble learning models like stacking offer a balanced approach by integrating diverse perspectives from algorithms with high variance and high bias. This integration improves diagnostic precision and uniformity while managing intricate clinical datasets (Sun et al., 2018).

To enhance generalizability, especially in limited datasets, methods such as rotation, scaling, and cropping are frequently employed in medical imaging to simulate a broader range of clinical situations and fortify model robustness (Shorten & Khoshgoftaar, 2019). The efficacy of ensemble learning in medical applications, particularly stacking models, arises from their ability to manage noisy and uneven data. This renders them particularly suitable for imaging diagnostics, where precision and recall are paramount (Topol, 2019).

The gains underscore the practical significance of automated detection models, particularly ensemble methods like stacking, in improving diagnostic accuracy, reducing inter-observer variability, and promoting more consistent clinical decision-making in lumbar spine analysis (Lu et al., 2018).

METHODOLOGY

Data Acquisition and Preparation

The experiment utilized a dataset of labeled lumbar spine images to detect occurrences of spondylolisthesis. The dataset was partitioned into training and testing sets with an 80%-20% ratio, along with preparatory procedures including normalization and scaling. Cross-validation was employed to ensure the model's performance is robust and reliable.

Model Development

A Bagging Classifier employing 150 estimators was implemented to reduce variance and improve model stability. The model integrates outcomes from multiple decision trees trained on bootstrapped datasets.

Random Forest: This classifier includes 100 estimators and utilizes bootstrap sampling in conjunction with decision trees to attain robust classification, leveraging feature unpredictability to improve accuracy.

K-Nearest Neighbors (KNN): A KNN classifier with \(k=4\) was utilized as a baseline model for comparison, owing to its shown effectiveness in classification tasks involving spatially scattered data.

Naive Bayes: A Gaussian Naive Bayes model was employed, serving as a probabilistic baseline model distinguished by its swift computational efficiency.

The Stacking Classifier amalgamates various basic classifiers, including Decision Trees, KNN, Random Forest, Naive Bayes, and Bagging, employing Logistic Regression as the meta-classifier.

Evaluation and Examination

Each model was trained on the training set and subsequently evaluated using the test set. Cross-validation scores were computed, and metrics such as accuracy were utilized to assess model performance.

RESULTS AND EXAMINATION

The results demonstrated that ensemble models outperformed individual classifiers in identifying lumbar spondylolisthesis.

Bagging Classifier: Achieved an accuracy of 94%, demonstrating improved stability relative to an individual Decision Tree classifier, albeit displaying slightly inferior performance compared to other ensemble techniques.

The Random Forest algorithm shown an impressive 96% accuracy in handling the complexity of the lumbar dataset, utilizing its intrinsic feature unpredictability.

KNN: Achieved 96% accuracy, indicating its effectiveness in categorizing spatial data, while it is susceptible to noise in the dataset.

Naive Bayes: Attained an accuracy of 95%, serving as a rapid baseline but struggling with the management of complex, linked information.

Stacking Classifier: Attained superior performance compared to all models, achieving 97% accuracy, indicating that the amalgamation of base classifiers with a Logistic Regression meta-classifier improved detection efficacy.

The Stacking Classifier adeptly amalgamates the advantages of many classifiers, enabling it to discern complex patterns in lumbar spine images more successfully than individual models and basic ensembles. This study corroborates prior research suggesting that stacking improves accuracy in complex classification tasks, making it a suitable choice for clinical use.

CONCLUSION AND FUTURE DIRECTIONS

This work demonstrates that ensemble learning models, especially the Stacking Classifier, significantly enhance the identification of lumbar spondylolisthesis. The Stacking Classifier exhibits a 97% accuracy rate, offering a reliable and exact method for automated diagnosis, which may aid physicians in making timely and accurate decisions. Subsequent research may explore more intricate neural network ensembles, employ larger datasets, and investigate real-time applications in clinical settings to further validate and enhance these findings. Integrating additional imaging modalities and refining meta-classifier algorithms could further enhance model efficacy.

REFERENCES

Altaf, F., Gibson, A., & Dannawi, Z. (2007). The clinical and radiological outcomes of lumbar spondylolisthesis. *Journal of Bone and Joint Surgery*, *89*(5), 636-642.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5-32.

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 785-794).

Dietterich, T. G. (2000). Ensemble methods in machine learning. In *International workshop on multiple classifier systems* (pp. 1-15). Springer, Berlin, Heidelberg.

Gangeh, M. J., Law, M., & Cheng, S. (2018). Ensemble methods in spine imaging research. *European Spine Journal*, *27*(6), 1327-1340.

Lao, J., Chen, Y., Li, H., & Liu, W. (2017). Automated detection of spinal conditions using machine learning methods. *Medical Imaging Journal*, *32*(4), 241-256.

Lu, Y., Lu, J., & Wang, C. (2018). Automated lumbar spine analysis in clinical settings. *Journal of Digital Imaging*, *31*(6), 831-841.

Matsunaga, S., & Sakou, T. (2000). Spondylolisthesis: Etiology and natural history. *Journal of Orthopaedic Science*, *5*(3), 1-5.

Patel, V., Shah, S., & Mehta, P. (2015). Radiographic analysis of spinal disorders: A clinical perspective. *Clinical Radiology*, *60*(1), 49-55.

Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, *1*(1), 81-106.

Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, *6*(1), 1-48.

Sill, J., Teschke, M., Grotthoff, S., & Schmidt, S. (2009). Advantages of ensemble methods in complex datasets. *Neural Processing Letters*, *25*(1), 57-68.

Sun, J., Wang, X., & Zhang, L. (2018). Stacking models in medical imaging for high-precision classification. *IEEE Transactions on Medical Imaging*, *37*(12), 2499-2508.

Ting, K. M., & Witten, I. H. (1999). Issues in stacked generalization. *Journal of Artificial Intelligence Research*, *10*, 271-289.

Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, *25*(1), 44-56.

Vapnik, V. (1998). *Statistical learning theory*. Wiley-Interscience.

Wolpert, D. H. (1992). Stacked generalization. *Neural Networks*, *5*(2), 241-259.

Zhou, Y., Shi, L., Zhang, Y., & Hu, Y. (2019). Machine learning applications in spinal imaging. *Journal of Spine Surgery*, *5*(1), 21-30.