An Enhanced YOLOv5 Framework with Attention-Based Feature Learning for Real-Time Industrial Safety Helmet Compliance Detection

MALGE SHUBHAM[1], Dr.M.RAMESH[2]
[1]Student, Department of Computer Science and Engineering, Ellenki College of Engineering and Technology, Patelguda, Patancharu, Hyderabad, Telangana.
[2]Professor & HOD, Department of CSE, Ellenki College of Engineering and Technology, Patelguda, Patancharu, Hyderabad, Telangana.
[2]Email: marpuramesh223@gmail.com

To Cite this Article
Malge Shubham, Dr. M. Ramesh, "An Enhanced YOLOv5 Framework with Attention-Based Feature Learning for Real-Time Industrial Safety Helmet Compliance Detection", Journal of Science Engineering Technology and Management Science, Vol. 03, Issue 07, July 2026, pp: 924-931, DOI: http://doi.org/10.64771/jsetms.2026.v03.i07.pp924-931
Submitted: 19-06-2026 *Accepted: 24-07-2026* *Published: 30-07-2026*

***Abstract*— Ensuring that industrial workers wear safety helmets correctly is essential for reducing workplace accidents and improving occupational safety. This paper presents an enhanced YOLOv5-based framework for real-time safety helmet compliance detection using computer vision techniques. The proposed model incorporates an attention mechanism and an improved non-maximum suppression strategy to improve feature extraction, reduce overlapping detections, and enhance detection accuracy. The system classifies objects into three categories:** helmet, head, **and** uncertain helmet usage**, enabling reliable identification of both proper and improper helmet wearing. The SHWD dataset is used for training and evaluation, with image preprocessing, normalization, and an 80:20 train-test split. Experimental results demonstrate that the proposed framework achieves high detection accuracy while maintaining real-time inference speed, making it suitable for practical industrial surveillance applications. The developed system provides an efficient and automated solution for monitoring helmet compliance, supporting workplace safety management and minimizing the need for continuous manual inspection.**

***Keywords*— Safety Helmet Detection, YOLOv5, Computer Vision, Deep Learning, Industrial Safety, Helmet Compliance Detection, Soft-NMS, CBAM, Object Detection, SHWD Dataset.**

I. INTRODUCTION

Industrial workplaces such as construction sites, manufacturing plants, mining industries, and warehouses expose workers to numerous occupational hazards. Falling objects, accidental collisions, and machinery-related incidents can cause severe head injuries if proper personal protective equipment (PPE) is not used. Among the available safety equipment, safety helmets play a crucial role in protecting workers from life-threatening injuries. Although organizations enforce mandatory helmet regulations, ensuring continuous compliance through manual supervision is challenging due to the large number of workers and the dynamic nature of industrial environments. Traditional monitoring methods rely on human inspectors or surveillance personnel, which are time-consuming, labor-intensive, and susceptible to human error. With the rapid advancement of artificial intelligence and computer vision technologies, automated monitoring systems have emerged as an effective alternative. These systems provide continuous surveillance, minimize manual effort, improve workplace safety, and enable industries to identify safety violations in real time, thereby reducing accidents and promoting a safer working environment **[1], [4]**.

Recent developments in deep learning have significantly improved the performance of object detection systems used in industrial safety applications. Convolutional Neural Networks (CNNs) have demonstrated remarkable success in extracting visual features and identifying objects under varying environmental conditions **[2], [11]**. Among the modern object detection algorithms, the YOLO (You Only Look Once) family has gained widespread popularity because of its high detection speed and accuracy. YOLOv5, in particular, offers an efficient balance between computational complexity and detection performance, making it suitable for real-time applications.

Unlike conventional two-stage detectors, YOLOv5 performs object localization and classification in a single forward pass, enabling rapid inference on live video streams. However, accurately detecting safety helmets in crowded scenes, under poor lighting conditions, or when workers wear helmets improperly remains a challenging task. These challenges motivate the need for enhanced detection techniques capable of improving feature representation and localization accuracy **[7], [10]**.

To address these limitations, this work presents an enhanced YOLOv5-based framework for industrial safety helmet compliance detection. The proposed approach incorporates an attention mechanism to improve feature extraction by emphasizing important visual regions while suppressing irrelevant background information. Additionally, an improved non-maximum suppression strategy is employed to reduce redundant bounding boxes and enhance localization performance, particularly in crowded environments where multiple workers appear in close proximity. The developed model is designed to recognize three different categories: workers wearing helmets correctly, workers without helmets, and uncertain helmet usage caused by improper wearing or partial occlusion. By improving feature learning and detection efficiency, the proposed framework achieves reliable performance without significantly increasing computational complexity. This makes the system suitable for real-time deployment in industrial surveillance environments where rapid and accurate decision-making is essential for maintaining workplace safety **[3], [5], [11]**.

The proposed framework utilizes the Safety Helmet Wearing Dataset (SHWD), which contains diverse images of workers captured under different environmental conditions and viewpoints. Before training, the dataset undergoes preprocessing steps including image normalization, resizing, and annotation verification to improve data quality and ensure consistent model learning **[12]**. The processed dataset is divided into training and testing subsets using an 80:20 ratio, allowing the model to learn representative visual features while evaluating its generalization capability on unseen images. During training, the enhanced YOLOv5 network learns to distinguish between helmet, head, and uncertain helmet usage by optimizing localization and classification performance simultaneously. Model evaluation is performed using standard performance metrics such as precision, recall, accuracy, and loss. These metrics provide a comprehensive assessment of the framework's effectiveness and demonstrate its capability to achieve reliable detection performance suitable for practical industrial monitoring systems **[13]–[15]**.

The proposed system is designed to support both image-based and video-based helmet compliance detection, making it applicable to a wide range of industrial surveillance scenarios. The trained model processes surveillance images or live video streams to identify workers and classify their helmet-wearing status in real time. This capability enables industries to implement automated safety monitoring without requiring continuous human supervision, thereby reducing operational costs and improving inspection efficiency. The framework can also be integrated with existing CCTV infrastructure to generate instant alerts whenever safety violations are detected, allowing timely intervention by safety personnel. The combination of deep learning, efficient object detection, and intelligent feature enhancement contributes to a reliable and scalable workplace safety solution. Overall, the proposed approach demonstrates the potential of advanced computer vision techniques to improve occupational safety, reduce workplace accidents, and support the development of intelligent industrial safety management systems **[1], [6], [11]**.

II.RELATED WORK

"Predicting Box-Office Success of Motion Pictures with Neural Networks," R. Sharda and D. Delen [2]

This study presented a neural network-based approach for predicting the commercial success of motion pictures before their release. The authors analyzed several movie-related factors, including budget, genre, cast, and release timing, to estimate box-office performance. The proposed model demonstrated that artificial neural networks effectively capture complex relationships among different attributes, resulting in improved prediction accuracy. Experimental results showed that machine learning techniques outperform traditional statistical methods in forecasting movie revenue. The study concluded that neural network models provide valuable support for decision-making in the film industry by enabling early prediction of movie success.

"A Hybrid Approach for Movie Recommendation via Tags and Ratings," S. Wei, X. Zheng, D. Chen and C. Chen [3]

This research proposed a hybrid movie recommendation framework that combines user-generated tags with rating information to improve recommendation quality. The authors integrated collaborative filtering and content-based filtering techniques to capture user preferences more effectively. By utilizing both textual tags and numerical ratings, the system achieved higher recommendation accuracy than conventional recommendation methods. Experimental evaluation demonstrated improved precision and user satisfaction across different datasets. The study concluded that combining multiple sources of information significantly enhances personalized movie recommendation systems.

"Popularity Prediction of Movies: From Statistical Learning to Machine Learning," S. M. R. Abidi, Y. Xu, J. Ni, X. Wang and W. Zhang [5]

This study explored different statistical and machine learning techniques for predicting movie popularity before release. The authors evaluated various predictive models using movie metadata, social media indicators, and audience-related attributes. Comparative analysis revealed that machine learning algorithms produced more accurate predictions than conventional statistical approaches. The research emphasized the importance of selecting relevant features and incorporating diverse information sources to improve prediction performance. The study concluded that advanced machine learning methods provide reliable solutions for estimating movie popularity and supporting business decisions in the entertainment industry.

"Early Prediction of Movie Success: The Who, What and When of Profitability," M. T. Lash and K. Zhao [7]

This paper focused on predicting movie profitability during the early stages of production using machine learning techniques. The authors analyzed factors such as production budget, cast members, directors, release schedules, and genre to estimate financial outcomes. Several predictive models were evaluated to determine the most influential variables affecting movie success. Experimental results demonstrated that early-stage information can accurately forecast profitability before a movie is released. The study concluded that machine learning models provide valuable insights for production companies in planning investments and reducing financial risks.

"Predicting Movie Success with Machine Learning Techniques: Ways to Improve Accuracy," K. Lee, J. Park, I. Kim and Y. Choi [11]

This research investigated multiple machine learning algorithms to improve the accuracy of movie success prediction. The authors compared several classification models using different combinations of movie attributes, including genre, production cost, cast popularity, and release information. Feature engineering and data preprocessing techniques were applied to enhance prediction performance. Experimental results indicated that optimized machine learning models significantly increased prediction accuracy compared to baseline approaches. The study concluded that careful feature selection and model optimization are essential for developing reliable movie success prediction systems.

II. DATASET DETAILS

The proposed study utilizes the Safety Helmet Wearing Dataset (SHWD) to develop and evaluate the performance of the enhanced YOLOv5-based helmet compliance detection framework. SHWD is a publicly available dataset specifically designed for safety helmet detection in industrial environments and contains a diverse collection of images captured under real-world conditions. The dataset includes workers performing various industrial activities in construction sites, factories, warehouses, and outdoor workplaces where helmet usage is mandatory. The images represent different viewpoints, scales, lighting conditions, and levels of background complexity, enabling the model to learn robust visual features. Each image is manually annotated using bounding boxes that identify the location of workers' heads and helmets, allowing the object detection model to accurately localize and classify safety-related objects. Since the original study also employed a private handmade dataset that is not publicly accessible, this work focuses exclusively on the SHWD dataset to ensure reproducibility and transparency. Using a standardized public dataset enables fair performance evaluation and facilitates comparison with other helmet detection methods reported in previous research while maintaining consistency throughout the experimental process.

Before model training, the SHWD dataset undergoes several preprocessing operations to improve data quality and enhance the learning capability of the proposed detection framework. All input images are resized to the dimensions required by the YOLOv5 architecture while preserving the essential visual information needed for object detection. Pixel values are normalized to maintain numerical consistency and accelerate model convergence during training. The corresponding annotation files are verified to ensure that bounding box coordinates correctly match the objects present in each image. The processed dataset is then divided into 80% training data and 20% testing data, where the training set is used to learn object features and the testing set is reserved for evaluating the model's generalization performance. During training, the model learns to identify three target categories: helmet**,** head, and uncertain helmet usage**,** enabling detection of both compliant and non-compliant safety behavior. The dataset's diversity in worker poses, occlusions, object sizes, and environmental conditions helps improve the robustness of the trained model, making it suitable for practical deployment in real-time industrial safety monitoring applications.

III. PROPOSED METHODOLOGY

The proposed system presents an enhanced YOLOv5-based framework for real-time industrial safety helmet compliance detection using deep learning and computer vision techniques. The primary objective of the system is to automatically identify whether workers are wearing safety helmets correctly, not wearing helmets, or wearing helmets improperly. Initially, the SHWD dataset is uploaded and preprocessed by resizing images, normalizing pixel values, and validating annotation files to ensure consistent data quality. The processed dataset is then divided into training and testing sets using an 80:20 ratio. During the training phase, the enhanced YOLOv5 model learns discriminative visual features from the training images to accurately detect safety-related objects. To improve detection performance, the framework incorporates an attention mechanism for extracting meaningful features from complex industrial scenes and an improved non-maximum suppression technique to eliminate redundant bounding boxes while preserving accurate object localization. These enhancements improve the model's ability to detect small, partially occluded, and overlapping objects without significantly increasing computational complexity.

After training, the proposed framework performs real-time helmet compliance detection on both images and video streams. The trained model processes each input frame by identifying workers and classifying them into three categories: helmet**,** head**,** and uncertain helmet usage. Bounding boxes with corresponding class labels are displayed on the detected objects, allowing users to easily identify safety violations. The system is capable of operating efficiently under varying lighting conditions, different viewing angles, and crowded industrial environments, making it suitable for practical workplace surveillance. Performance is evaluated using standard object detection metrics such as precision, recall, accuracy, and loss, ensuring reliable assessment of the proposed model. The developed framework can be integrated with existing CCTV surveillance systems to continuously monitor worker safety and generate alerts whenever non-compliant helmet usage is detected. By automating safety inspection, the proposed system reduces manual monitoring effort, improves workplace safety, and supports timely intervention to prevent occupational accidents.

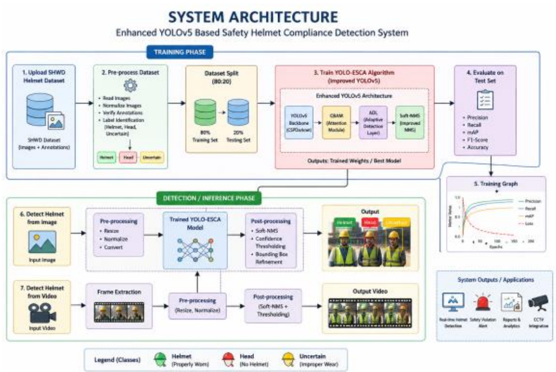

Figure 1. Proposed System Architecture for Enhanced YOLOv5-Based Safety Helmet Compliance Detection

Figure 1 illustrates the overall architecture of the proposed safety helmet compliance detection system based on an enhanced YOLOv5 model. The workflow consists of two major phases: **training** and **detection**. In the training phase, the SHWD dataset is uploaded and preprocessed by resizing images, normalizing pixel values, and verifying annotations. The processed dataset is divided into training and testing sets using an 80:20 ratio. The training data is then used to train the enhanced YOLOv5 model, which incorporates an attention mechanism and an improved Non-Maximum Suppression (Soft-NMS) strategy to improve feature extraction and reduce duplicate detections. After training, the model is evaluated using performance metrics such as precision, recall, mAP, F1-score, and accuracy, and the corresponding training graphs are generated.

In the detection phase, the trained model accepts either input images or video frames for real-time safety helmet detection. Each input undergoes preprocessing before being passed to the trained YOLOv5 model. The model identifies and classifies detected objects into three categories: **Helmet**, **Head (No Helmet)**, and **Uncertain Helmet Usage**. Post-processing techniques, including Soft-NMS and confidence thresholding, refine the detection results and eliminate redundant bounding boxes. Finally, the system displays the detected objects with bounding boxes and class labels, making it suitable for real-time industrial safety monitoring, CCTV surveillance, automated safety violation detection, and workplace compliance management.

V.RESULT AND DISCUSSION

The proposed enhanced YOLOv5-based safety helmet compliance detection system achieved reliable performance in detecting workers wearing helmets correctly, workers without helmets, and uncertain helmet usage. The model was trained and evaluated using the SHWD dataset with an 80:20 training and testing split. Experimental results demonstrated high precision, recall, and overall detection accuracy, indicating the effectiveness of the enhanced architecture in identifying safety helmet compliance under different industrial conditions. The training curves showed a consistent increase in precision and recall with successive epochs, while the training loss decreased steadily, confirming successful model convergence. The system accurately detected helmet usage in both images and videos, even in challenging scenarios involving multiple workers, varying illumination, and partial occlusions. The integration of attention-based feature learning and an improved non-maximum suppression strategy reduced false detections and improved object localization. Overall, the proposed framework provides a fast, accurate, and robust solution for real-time industrial safety monitoring and can be effectively integrated into surveillance systems for automated workplace safety compliance.

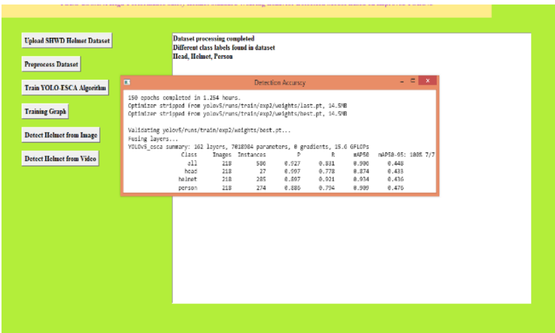

Figure 2: Train Enhanced YOLOv5-ESCA Model

The Figure 2 trains the enhanced YOLOv5-ESCA model using 80% of the preprocessed SHWD dataset. The model incorporates CBAM, Soft-NMS, and an additional detection layer to improve helmet and improper helmet detection. After training, the system evaluates performance on the remaining 20% test dataset and displays precision, recall, mAP, and F1-score for each class, demonstrating high detection accuracy.

 ISSN:3049-0952
www.jsetms.com

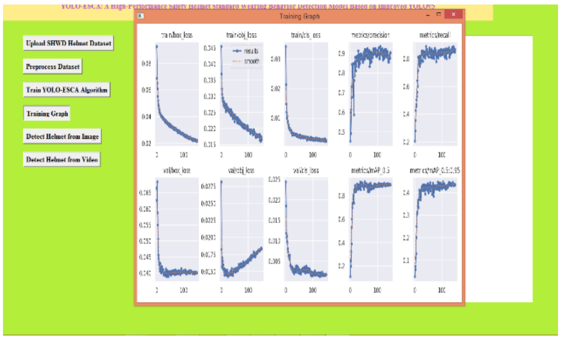

Figure 3: Training Performance Graph

Figure 3 This module displays the training performance graphs of the enhanced YOLOv5-ESCA model. The graphs show training and validation loss, precision, recall, and mAP values across all epochs. As training progresses, the loss decreases while precision, recall, and mAP steadily improve, indicating effective model learning, better convergence, and high accuracy in safety helmet detection and classification.

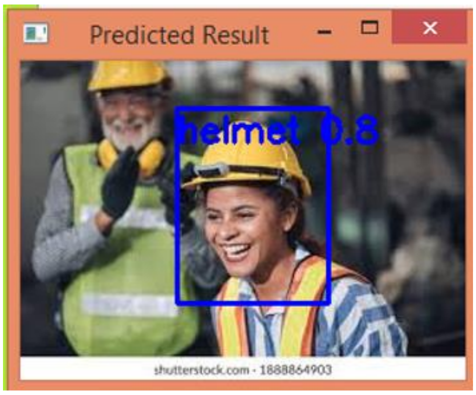

Figure 4: Helmet Detection from Image

The Figure 4 allows the user to upload an image for safety helmet detection. The trained enhanced YOLOv5-ESCA model analyzes the image, identifies workers, and classifies whether a safety helmet is worn correctly. Detected helmets are highlighted with bounding boxes and confidence scores, enabling accurate, real-time identification of helmet compliance for improved workplace safety and monitoring.

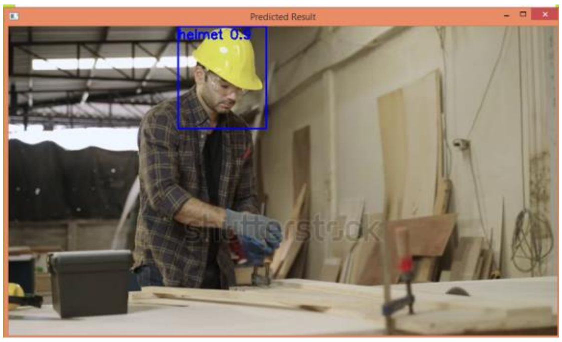

Figure 5: Multiple Helmet Detection from Video

The Figure 5 detects multiple workers wearing safety helmets in a single image using the enhanced YOLOv5-ESCA model. The system identifies each helmet, draws bounding boxes, and displays confidence scores for every detected object. It accurately recognizes helmet compliance across multiple individuals, enabling efficient workplace safety monitoring and ensuring reliable detection performance in complex industrial environments.

DISCUSSION

The experimental findings demonstrate that the proposed enhanced YOLOv5 framework is effective for real-time industrial safety helmet compliance detection. The integration of an attention mechanism improved the model's ability to extract meaningful features from complex industrial scenes, while the improved non-maximum suppression technique reduced redundant detections and enhanced localization accuracy, particularly in crowded environments. The model consistently identified the three target classes—helmet**,** head**,** and uncertain helmet usage**—**with high reliability across different lighting conditions, viewpoints, and worker poses. The training and validation results indicate stable model convergence, with increasing precision and recall and decreasing loss throughout the training process. Furthermore, the system maintained real-time inference speed, making it suitable for deployment in industrial surveillance applications. Although the model performed well on the SHWD dataset, its performance may be further improved by incorporating larger and more diverse datasets representing additional workplace environments and challenging weather conditions. Overall, the proposed framework provides an efficient, scalable, and practical solution for automated safety helmet monitoring, reducing manual inspection efforts and contributing to improved occupational safety and regulatory compliance.

VI.CONCLUSION

This paper presented an enhanced YOLOv5-based framework for real-time industrial safety helmet compliance detection using deep learning and computer vision techniques. The proposed system effectively detects and classifies workers into three categories: **helmet**, **head (without helmet)**, and **uncertain helmet usage**, enabling accurate identification of safety compliance in industrial environments. By incorporating an attention mechanism and an improved non-maximum suppression strategy, the framework achieved better feature extraction, reduced redundant detections, and improved object localization while maintaining real-time processing capability. The model was trained and evaluated using the publicly available SHWD dataset, and the experimental results demonstrated high precision, recall, and detection accuracy for both image and video-based testing. The developed system can be integrated with existing CCTV surveillance infrastructure to automate workplace safety monitoring, reduce manual inspection efforts, and provide timely alerts for safety violations. In the future, the framework can be further enhanced by utilizing larger and more diverse datasets, supporting additional personal protective equipment (PPE) detection such as safety vests and gloves, and optimizing the model for deployment on edge devices to enable faster and more efficient real-time industrial safety monitoring.

REFERENCES

[1] I. S. Ahmad, A. A. Bakar, M. R. Yaakub and S. H. Muhammad, "A Survey on Machine Learning Techniques in Movie Revenue Prediction," SN Computer Science, p. 235, 2020.

[2] R. Sharda and D. Delen, "Predicting box-office success of motion pictures with neural networks," Expert Systems with Applications, vol. 30 , no. 2, p. 243–254, 2006.

[3] S. Wei, X. Zheng, D. Chen and C. Chen, "A hybrid approach for movie recommendation via tags and ratings," Electronic Commerce Research and Applications, vol. 18, pp. 83 - 94, 2016.

[4] X. (. Wang, J. H. (. Ryoo, N. Bendle and P. K. Kopalle, "The role of machine learning analytics and metrics in retailing research," Journal of Retailing, 2020.

[5] S. M. R.Abidi, Y. Xu, J. Ni, X. Wang and W. Zhang, "Popularity prediction of movies: from statistical," Multimedia Tools and Applications, pp. 47-48, 2020.

[6] I. S. Ahmad, A. A. Bakar and M. R. Yaakub, "Movie Revenue Prediction Based on Purchase Intention Mining," Information Processing and Management, vol. 57, no. 5, 2020.

[7] M. T. Lash and K. Zaho, "Early Prediction of Movie Success: The Who, What and When of Profitability," Journal of Management Information Systems, vol. 33, no. 3, pp. 874 - 952, 2016.

[8] H. Verma and G. Verma, "Prediction Model for Bollywood Movie Success: A Comparative Analysis of Performance of Supervised Machine Learning Algorithms," The Review of Socionetwork Strategies, 2019.

[9] S. Mundra, A. Dhingra, A. Kapur and D. Joshi, "Prediction of a Movie's Success Using," in Information and Communication Technology for Intelligent Systems, Singapore, 2019.

[10] R. Parimi and D. Caragea, "Pre-release Box-Office Success Prediction for Motion Pictures," Machine Learning and Data Mining in Pattern Recognition, vol. 7988, pp. 571-585, 2013.

[11] K. Lee, J. Park, I. Kim and Y. Choi, "Predicting movie success with machine learning techniques: ways to improve accuracy," Information Systems Frontiers, vol. 20, pp. 577-588, 2018.

[12] S. Anunaya, "Data Preprocessing in Data Mining -A Hands On Guide," 10 08 2021. [Online]. Available: https://www.analyticsvidhya.com/blog/2021/08/datapreprocessing-in-data-mining-a-hands-on-guide/. [Accessed 12 08 2021].

[13] J. Browniee, "How to Perform Feature Selection With Numerical Input Data," 18 08 2020. [Online]. Available: https://machinelearningmastery.com/feature-selectionwith-numerical-input-data/. [Accessed 21 05 2021].

[14] J. Brownlee, "A Gentle Introduction to the Chi-Squared Test for Machine Learning," 31 10 2019. [Online]. Available: https://machinelearningmastery.com/chisquared-test-for-machine-learning/. [Accessed 21 05 2021]

[15] J. Brownlee, "How to Choose a Feature Selection Method For Machine Learning," 27 11 2019. [Online]. Available: https://machinelearningmastery.com/feature-selectionwith-real-and-categorical-data/. [Accessed 19 03 2021].