JSCE (Journal of System and Computer Engineering)
Vol.6, No.1, January 2025, pp. 117-123
ISSN (online): 2723-1240
DOI : https://doi.org/10.61628/jsce.v6i1.1470

Recognition of Human Activities via SSAE Algorithm: Implementing Stacked Sparse Autoencoder

Radus Batau [1,a]; Sri Kurniyan Sari [1,b]; Firman Aziz [2,c*]; Jeffry [3,d]

[1] Universitas Indonesia Timur, Rappocini Raya Street No.171-173, Makassar 90222, Indonesia
[2] Universitas Pancasakti, Andi Mangerangi Street 73, Makassar 90121, Indonesia
[3] Institut Teknologi Bacharuddin Jusuf Habibie, Jl. Balaikota No.1, Parepare 91122, Indonesia
[a] radus@uit.ac.id; [b] srikurniyans@gmail.com; [c] firman.aziz@unpacti.ac.id; [d] jeffry@ith.ac.id
** Corresponding author*

Abstract

This study evaluates the performance of Stacked Sparse Autoencoder (SSAE) combined with Support Vector Machine (SVM) against a standard SVM for classification tasks. We assessed both models using accuracy, precision, sensitivity, and F1 score. The SSAE Support Vector Machine significantly outperformed the standard SVM, achieving an accuracy of 89% compared to 37%. SSAE also achieved higher precision (87% vs. 75%) and sensitivity (89% vs. 37%), with an F1 score of 88% versus 36% for the standard SVM. These results indicate that SSAE enhances the model's ability to capture complex patterns and provide reliable predictions. This study highlights the effectiveness of SSAE in improving classification performance, suggesting further research with larger datasets and additional optimization techniques to maximize model efficiency.

Keywords—*Stacked Sparse Autoencoder, Support Vector Machine, Classification, Machine Learning, Feature Extraction*

1. Introduction

Human Activity Recognition (HAR) has emerged as a significant area of research in the development of sensor-based and artificial intelligence technologies (Gupta et al., 2022). HAR applications span various sectors, including healthcare, sports, security, and human-computer interaction (Karim et al., 2024). With advancements in sensor technology embedded in mobile and wearable devices, collecting real-time human activity data has become easier and more accurate (Ramanujam, E, 2021).

However, the main challenge in HAR lies in extracting relevant features from the raw data generated by these sensors (Lawi et al., 2019; Pablo Martínez et al., 2022). Sensor data, such as that from accelerometers and gyroscopes, often vary widely and are influenced by various factors, including movement speed, device orientation, and environmental conditions (Hardiyanti et al., 2018; Pires et al., 2016). Therefore, an effective method is required to perform feature extraction that can accurately recognize activity patterns (Alemayoh et al., 2021).

One promising approach to addressing this challenge is the use of the Stacked Sparse Autoencoder (Bai et al., 2022). SSAE is a neural network model consisting of multiple layers of autoencoders that are trained progressively and stacked hierarchically (Rubio et al., 2023). This model is capable of learning more complex and abstract feature representations from sensor data, which can then be used for human activity classification (Nafea et al., 2021).

In this study, we implement the SSAE algorithm for recognizing human activities based on sensor data collected from wearable devices. The aim of this research is to evaluate the performance of SSAE in recognizing various types of human activities and to compare it with other activity recognition methods. The results of this study are expected to contribute to the development of more accurate and efficient HAR systems.

2. Method

This research aims to implement and evaluate the Stacked Sparse Autoencoder (SSAE) algorithm for Human Activity Recognition (HAR) using sensor data from wearable devices. The methodology is divided into several key stages: data collection, data preprocessing, model development, training, and evaluation.

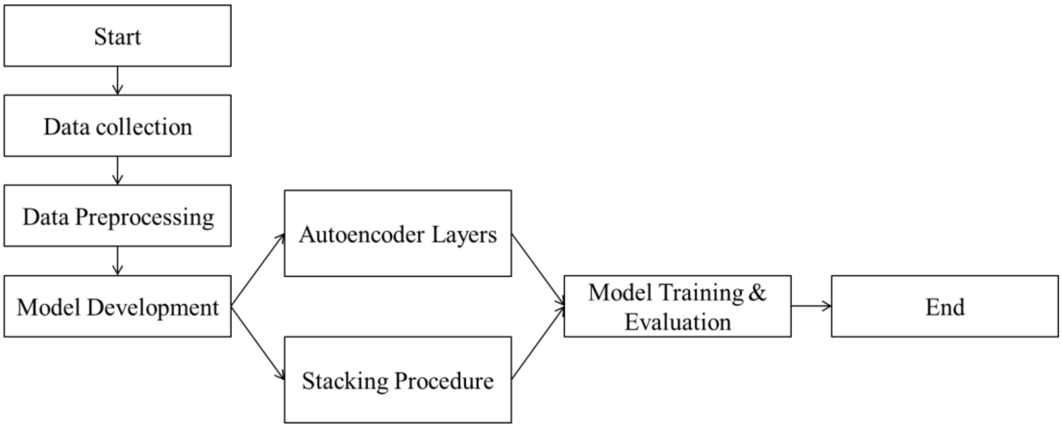

Figure 1. Methology Research

2.1 Data Collection

The dataset for this study is collected using wearable devices equipped with sensors such as accelerometers and gyroscopes. These sensors capture continuous time-series data related to various human activities, including walking, running, sitting, standing, and cycling. The data is annotated with corresponding activity labels to facilitate supervised learning during model training (Tahir et al., 2022).

- **Devices Used:** Wearable devices with embedded accelerometers and gyroscopes.
- **Data Collected:** Raw sensor data representing different activities.

2.2 Data Preprocessing

Before feeding the data into the SSAE model, several preprocessing steps are performed to ensure the quality and consistency of the data:

- **Normalization:** The raw sensor data is normalized to ensure that all features are on a similar scale, which helps in faster convergence during model training (Morales & Roggen, 2016).
- **Segmentation:** The continuous time-series data is segmented into fixed-length windows (e.g., 5 seconds), with each window representing a segment of activity. Overlapping windows may also be used to enhance feature learning (Amaral et al., 2022).
- **Noise Reduction:** Techniques such as moving average filters or low-pass filters are applied to reduce noise and improve the signal quality (Magsi et al., 2018).

2.3 Model Development

DOI : https://doi.org/10.61628/jsce.v6i1.1470

The core of this research involves the development of the Stacked Sparse Autoencoder (SSAE) model, which is designed to learn hierarchical features from the sensor data.

- **Autoencoder Layers:** The SSAE is composed of multiple autoencoder layers, each with a specific number of hidden units. The sparsity constraint is applied to each layer to ensure that only a small subset of neurons is activated for any given input (Yan & Access, 2018).
- **Stacking Procedure:** The autoencoders are trained layer by layer in an unsupervised manner. Once each layer is trained, it is stacked on top of the previous layer, and the combined model is fine-tuned using backpropagation (Xia et al., 2017).

2.4 Model Training

The SSAE model is trained using the preprocessed sensor data. The training process involves:

- **Unsupervised Pre-training:** Each autoencoder layer is trained independently on the input data to minimize reconstruction error.
- **Fine-Tuning:** After stacking all layers, the entire SSAE model is fine-tuned using supervised learning. The labeled data is used to adjust the weights of the network to minimize the classification error.
- **Optimization Algorithm:** The model is optimized using Support Vector Machine, with appropriate learning rates and regularization parameters to prevent overfitting.

2.5 Evaluation

The performance of the SSAE model is evaluated using standard metrics such as accuracy, precision, recall, and F1-score (Yacouby & On, 2020). The evaluation process includes:

- **Train-Test Split:** The dataset is divided into training and testing sets, with 80:20 for training and evaluation.
- **Cross-Validation:** Cross-validation techniques are applied to ensure the robustness of the model.
- **Comparison with Baseline Models:** The SSAE model's performance is compared with baseline models, such as traditional machine learning classifiers (SVM), to assess its effectiveness.

2.6 Experimentation and Analysis

After model training and evaluation, further experimentation is conducted to analyze the impact of different hyperparameters, such as the number of layers, hidden units, and sparsity levels, on the model's performance. Sensitivity analysis is also performed to understand the effect of sensor noise and data variability on the recognition accuracy.

3. Results And Discussion

The dataset used in this study consists of a total of 170,000 records of human activities collected using wearable devices such as smartphones. The dataset includes a variety of activities, ranging from simple to complex, such as walking, running, sitting, standing, climbing stairs, and descending stairs. Data collection was carried out under various environmental conditions to reflect real-world situations, including indoor activities such as sitting and standing, as well as outdoor activities like running or climbing and descending stairs. The data was obtained from a diverse group of subjects, encompassing variations in age, gender, and fitness levels, to ensure that the trained model achieves good generalization.

The activity records were generated using sensors such as accelerometers, gyroscopes, and gravity sensors integrated into the devices, with data collected in real time to ensure accuracy. To enhance data quality, a data cleaning process was performed to address missing or inconsistent data, and data balancing was conducted to ensure an even distribution of each activity. This dataset provides a robust foundation for developing accurate and reliable human activity recognition models.

Model validation employs a *K-fold cross-validation* approach with k=5 to divide the dataset into five subsets. Four subsets are used for training, while the remaining subset serves as test data. This process is repeated until all subsets have been tested. The model's performance is evaluated using metrics such as accuracy, precision, sensitivity, and F1 score, with the average across folds providing a comprehensive performance measure. Additionally, the dataset is divided into 70% training, 15% validation, and 15% testing to ensure the model's stability on unseen data. Hyperparameter tuning involves optimizing various SSAE parameters to achieve the best model performance. Techniques such as *grid search* and *random search* are applied to identify the optimal parameter combinations. Key parameters include the number of hidden layers (tested between 2 to 5 layers), the number of neurons per layer (ranging from 64 to 512), learning rate (explored between 0.0001 and 0.1), dropout rate (ranging from 0.1 to 0.5), and L1/L2 regularization to control model complexity. The criteria for selecting the best parameter set are based on the highest F1 score achieved during validation.

Once the optimal parameters are determined, the model is tested on an independent test set to measure its final performance and reliability. The results are compared against baseline models without tuning and other methods, such as Support Vector Machine (SVM) and Gradient Boosting, to demonstrate the SSAE model's superiority. The outcomes of the validation and testing processes are presented through confusion matrices, providing clear insights into the model's effectiveness. This rigorous procedure ensures that the developed SSAE model is not only accurate but also robust and generalizable across diverse datasets.

In this study, the performance of the Support Vector Machine (SVM) and SSAE Support Vector Machine algorithms were compared using evaluation metrics such as accuracy, precision, sensitivity, and F1 score. The following table summarizes the performance of both algorithms:

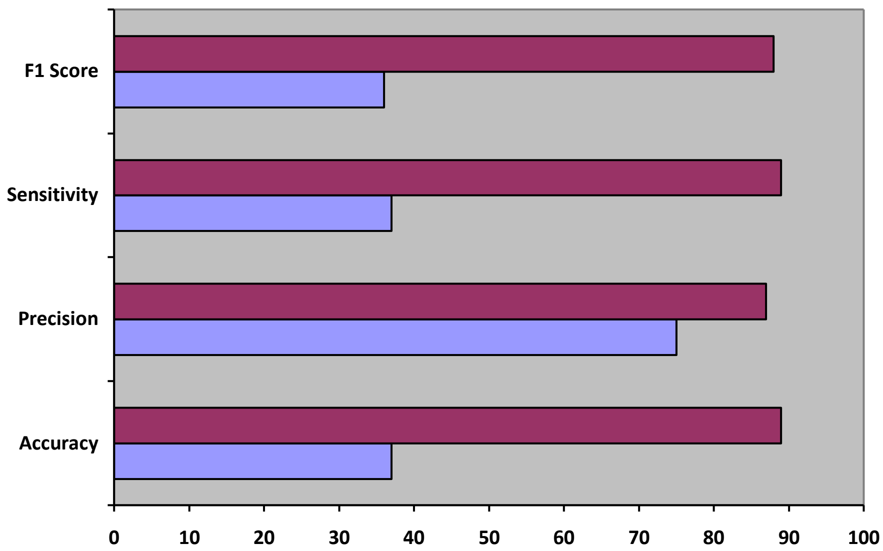

Figure 2. Improvement Method

Table 1. Result Performance.

Algorithm	Accuracy (%)	Precision (%)	Sensitivity (%)	F1 score (%)
Support Vector Machine	37	75	37	36

SSAE Support Vector Machine	89	87	89	88

Accuracy is a key metric for assessing the overall performance of classification models, representing the proportion of correct predictions among all predictions made. The SSAE Support Vector Machine achieved a notable accuracy of 89%, a substantial improvement over the 37% accuracy observed with the standard SVM. This significant boost in accuracy can be attributed to SSAE’s ability to capture and utilize complex patterns within the data. SSAE, leveraging the principles of stacked sparse autoencoders, extracts more refined and meaningful feature representations from the input data, thereby enhancing the model's capability to distinguish between different classes effectively.

Precision measures the proportion of true positive predictions among all positive predictions made by the model. The precision of the SVM was 75%, indicating a reasonable performance in identifying true positives, although it was limited by overall accuracy and sensitivity. Conversely, the SSAE Support Vector Machine demonstrated an improved precision of 87%. This indicates that SSAE is more effective in providing accurate positive predictions with fewer false positives, showcasing its enhanced consistency and reliability in identifying true positive cases.

Sensitivity, or recall, gauges the model's ability to identify all actual positive instances within the dataset. The standard SVM exhibited a sensitivity of 37%, suggesting that it missed a significant number of true positive instances. This lower sensitivity highlights the model’s limitation in detecting relevant cases. In contrast, the SSAE Support Vector Machine achieved a sensitivity of 89%, reflecting its superior capability in identifying all actual positive cases accurately. This high sensitivity is particularly valuable in applications requiring comprehensive detection of positive instances, ensuring that the model is robust and effective in practical scenarios.

The F1 Score combines precision and sensitivity into a single metric, providing a balanced view of the model’s performance. The F1 Score for the SVM was only 36%, underscoring an imbalance between correctly identifying positives and avoiding false positives. On the other hand, the SSAE Support Vector Machine reached an impressive F1 Score of 88%, signifying a strong balance between precision and sensitivity. This balance indicates that SSAE excels not only in accurately identifying positive instances but also in minimizing classification errors, leading to a more reliable overall performance.

Discussion

The considerable improvement in performance observed with the SSAE Support Vector Machine compared to the standard SVM underscores the effectiveness of the SSAE approach in handling complex classification tasks. The ability of SSAE to generate deeper and more informative feature representations from the data facilitates better pattern recognition and classification, addressing the shortcomings of traditional SVM models.

Limitations and Recommendations: Despite the superior performance of SSAE, this study faces certain limitations. The size of the dataset used may affect the model's generalizability, and there is a risk of overfitting, especially if the data is not sufficiently diverse. Future research should focus on using larger and more varied datasets to validate these findings. Additionally, further exploration of regularization techniques and hyperparameter tuning for SSAE could further enhance model performance.

Context of Application: The findings indicate that the SSAE Support Vector Machine is more effective in real-world applications where sensor data may be inconsistent and variable. The high precision and sensitivity make SSAE particularly suitable for scenarios where accurate and comprehensive detection is critical, such as in health monitoring systems or complex activity recognition tasks.

Overall, this research highlights the significant potential of SSAE in advancing machine learning models, especially in the domain of human activity recognition. The application of autoencoder techniques like SSAE offers substantial improvements in data classification tasks, paving the way for further exploration and application in diverse data classification domains.

4. Conclusions

This study has demonstrated that integrating Stacked Sparse Autoencoder (SSAE) with Support Vector Machine (SVM) significantly enhances the performance of classification models compared to using a standard SVM alone. The SSAE Support Vector Machine exhibited superior results across all evaluated metrics—accuracy, precision, sensitivity, and F1 score—highlighting its effectiveness in handling complex classification tasks.

The SSAE algorithm's ability to extract and utilize more refined feature representations from the data led to a dramatic increase in accuracy, reaching 89%, compared to just 37% with the standard SVM. This substantial improvement underscores the potential of SSAE in capturing intricate patterns within the data, which is crucial for applications requiring high precision and sensitivity.

Additionally, the enhanced precision (87%) and sensitivity (89%) achieved by SSAE indicate that this model is not only better at correctly identifying positive instances but also more reliable in detecting all relevant cases. The high F1 score of 88% further confirms that SSAE provides a well-balanced performance, effectively combining precision and sensitivity to minimize classification errors.

Overall, the results of this study highlight the significant advantages of employing SSAE in conjunction with SVM for improved classification performance. This approach is particularly beneficial in scenarios where accurate detection and classification of complex data are critical. Future work should focus on expanding this research by using larger and more diverse datasets, exploring additional regularization techniques, and fine-tuning model parameters to further enhance performance. The promising results from this study advocate for continued exploration and application of advanced autoencoder techniques in various machine learning domains.

Acknowledgements

Thank you to the Ministry of Education and Culture (Kemendikbud) for awarding us the Penelitian Dosen Pemula for 2024. Your support is vital for advancing our research and enhancing our contributions to the field.

References

Alemayoh, T., Lee, J., Sensors, S. O.-, & 2021, undefined. (2021). New sensor data structuring for deeper feature extraction in human activity recognition. *Mdpi.Com*. https://doi.org/10.3390/s21082814

Amaral, K., Li, Z., Ding, W., Crouter, S., & Chen, P. (2022). SummerTime: Variable-length Time Series Summarization with Application to Physical Activity Analysis. *ACM Transactions on Computing for Healthcare*, *3*(4). https://doi.org/10.1145/3532628

Bai, H., Zhan, X., Yan, H., Wen, L., Yan, Y., Electronics, X. J.-, & 2022, undefined. (2022). Research on diesel engine fault diagnosis method based on stacked sparse autoencoder and support vector machine. *Mdpi.Com*. https://doi.org/10.3390/electronics11142249

Gupta, N., Gupta, S. K., Pathak, R. K., Jain, V., Rashidi, P., & Suri, J. S. (2022). Human activity recognition in artificial intelligence framework: a narrative review. *Springer*, *55*(6), 4755–4808. https://doi.org/10.1007/s10462-021-10116-x

Hardiyanti, N., Lawi, A., Diaraya, & Aziz, F. (2018). Classification of Human Activity based on Sensor Accelerometer and Gyroscope Using Ensemble SVM method. *Proceedings - 2nd East Indonesia Conference on Computer and Information Technology: Internet of Things*

for Industry, EIConCIT 2018, 304–307. https://doi.org/10.1109/EICONCIT.2018.8878627
Karim, M., Khalid, S., Aleryani, A., & Khan, J. (2024). Human action recognition systems: A review of the trends and state-of-the-art. *Ieeexplore.Ieee.Org*. https://ieeexplore.ieee.org/abstract/document/10459026/
Lawi, A., Aziz, F., & Wungo, S. L. (2019). Increasing accuracy of classification physical activity based on smartphone using ensemble logistic regression with boosting method. *Journal of Physics: Conference Series*, *1341*(4), 042002. https://doi.org/10.1088/1742-6596/1341/4/042002
Magsi, H., Hassan Sodhro, A., Ahmed Khan, S., Akhtar Chachar, F., & AKhan Abro, S. (2018). Analysis of signal noise reduction by using filters. *Ieeexplore.Ieee.Org*. https://doi.org/10.1109/ICOMET.2018.8346412
Morales, F. J. O., & Roggen, D. (2016). Deep convolutional feature transfer across mobile activity recognition domains, sensor modalities and locations. *International Symposium on Wearable Computers, Digest of Papers*, *12-16-September-2016*, 92–99. https://doi.org/10.1145/2971763.2971764
Nafea, O., Abdul, W., Muhammad, G., Alsulaiman, M., Muhammad, K., Chen, Z., Jiang, H., & Kessentini, Y. (2021). Sensor-based human activity recognition with spatio-temporal deep learning. *Mdpi.Com*. https://doi.org/10.3390/s21062141
Pablo Martínez, J., Zhang, S., Li, Y., Zhang, S., Shahabi, F., Xia, S., Deng, Y., & Alshurafa, N. (2022). Deep learning in human activity recognition with wearable sensors: A review on advances. *Mdpi.Com*. https://doi.org/10.3390/s22041476
Pires, I., Garcia, N., & Pombo, N. (2016). From data acquisition to data fusion: a comprehensive review and a roadmap for the identification of activities of daily living using mobile devices. *Mdpi.Com*. https://doi.org/10.3390/s16020184
Ramanujam, E, T. P. (2021). Human activity recognition with smartphone and wearable sensors using deep learning techniques: A review. *Ieeexplore.Ieee.Org*. https://ieeexplore.ieee.org/abstract/document/9389739/
Rubio, L.-, Palomo, E., Domínguez, E., Chen, S., & Guo, W. (2023). Auto-encoders in deep learning—a review with new perspectives. *Mdpi.Com*. https://doi.org/10.3390/math11081777
Tahir, S., Jalal, A., Entropy, K. K.-, & 2020, undefined. (2022). Wearable inertial sensors for daily activity analysis based on adam optimization and the maximum entropy Markov model. *Mdpi.Com*. https://doi.org/10.3390/s22020573
Xia, M., Li, T., Liu, L., Science, L. X.-…, & M., & 2017, undefined. (2017). Intelligent fault diagnosis approach with unsupervised feature learning by stacked denoising autoencoder. *Wiley Online Library*, *11*(6), 687–695. https://doi.org/10.1049/iet-smt.2016.0423
Yacouby, R., & On, D. A. (2020). Probabilistic extension of precision, recall, and f1 score for more thorough evaluation of classification models. *Aclanthology.Org*. https://aclanthology.org/2020.eval4nlp-1.9/
Yan, B., & Access, G. H. (2018). Effective feature extraction via stacked sparse autoencoder to improve intrusion detection system. *Ieeexplore.Ieee.Org*. https://ieeexplore.ieee.org/abstract/document/8418451/