A Comparison Study of Models in an attempt to improve Sleep cycle estimation in wearable devices

Likith Venkatesh Gowda Prathima, Sathvik Bellary (Git)

Abstract:

This study delves into the application of various machine learning techniques for the accurate detection of sleep onset and wakeup using wrist-worn accelerometer data. Employing a diverse range of methodologies such as bagging, boosting, neural networks, and ensemble methods, the study focuses on maximizing accuracy for wearable devices. Multiple models are explored, each tailored to address specific challenges in distinguishing between sleep and wakefulness. Notably, the CatBoost process emerges as a promising approach with the potential to enhance precision and effectiveness in sleep pattern detection. The study contributes significant advancements to sleep detection technology, offering valuable insights for health and wellness monitoring systems.

Overview:

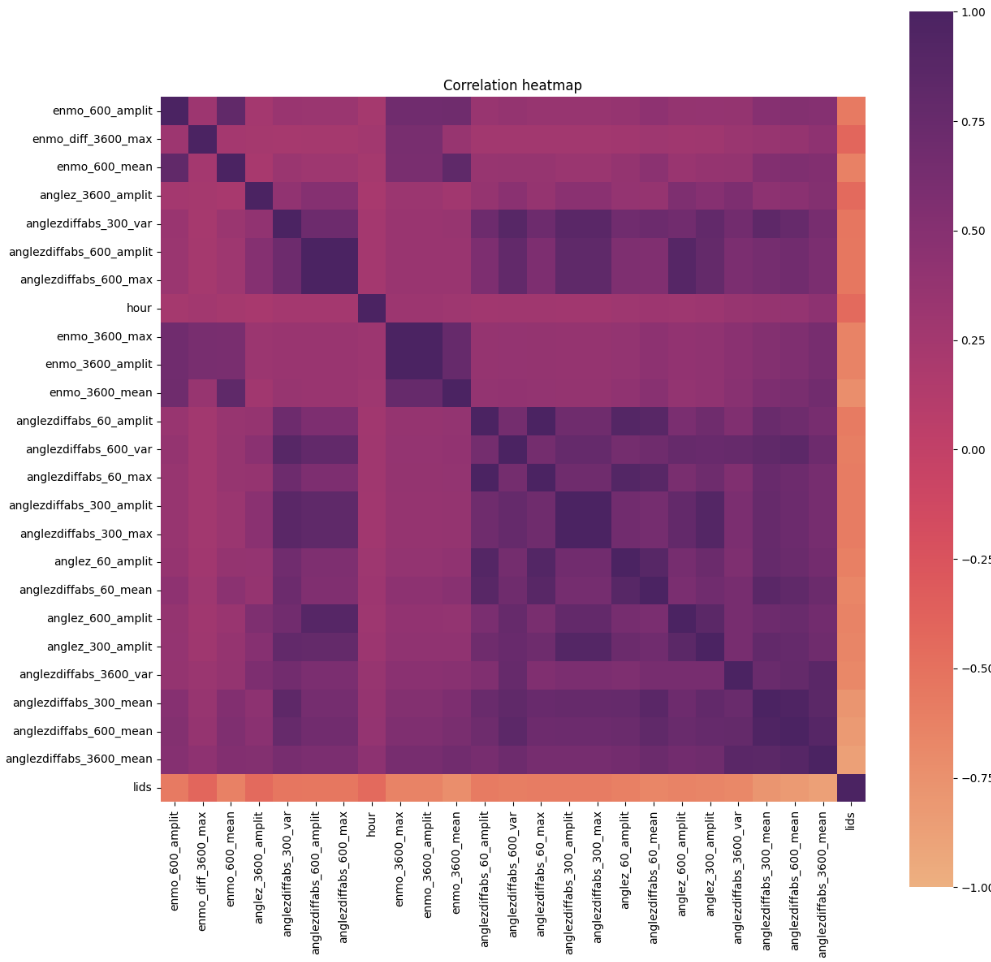
This study addresses the critical challenge of accurately detecting sleep onset and wakeup using wrist-worn accelerometer data. The motivation stems from the increasing relevance of wearable devices in health and wellness monitoring. The accurate identification of sleep patterns contributes significantly to personalized health insights, enabling individuals to make informed decisions for improved well-being. Accurate sleep detection has broad societal implications, aiding in the development of health and wellness monitoring systems. Use cases range from personalized sleep recommendations to early detection of sleep disorders. Improved sleep monitoring contributes to a holistic understanding of an individual's health, facilitating preventive healthcare measures.

The main challenge is to extract sleep and waking states from accelerometer data, an essential task for deciphering an individual's sleep patterns. Since wearables are now widely used, improving their ability to identify sleep patterns might provide important information about general health and lifestyle choices. The suggested method is a multifaceted approach that includes advanced feature engineering, memory optimization, and the use of various machine learning algorithms, including LightGBM, Random Forest, XGBoost, and Neural Networks. The rationale behind the approach lies in leveraging cutting-edge techniques to enhance model accuracy and efficiency to reduce errors in accelerometer readings. Feature engineering, such as the creation of custom metrics and rolling window features, adds granularity to the analysis. The choice of machine learning algorithms, each tailored to specific strengths, ensures a comprehensive evaluation of sleep detection. While feature engineering captures subtle temporal patterns and seasonal implications on sleep behavior, memory optimization guarantees effective data processing. The various correlations present in accelerometer data are accommodated by a diversified ensemble of machine learning techniques.

Time-based features and binary representations of seasons capture daily and seasonal sleep patterns. Custom metrics like 'lids' and rolling window features contribute to understanding both short-term and long-term physical activity trends relevant to sleep prediction. The results showcase the effectiveness of the proposed approach, with models like Random Forest achieving an accuracy of 97.93%. However, certain limitations, such as potential overfitting in Neural Networks and the need for further CatBoost model refinement, were identified and later implemented.

Experimental Setup:

The dataset includes accelerometer data from wrist wearers intended to identify when sleep begins and ends. Following the identification and exclusion of abnormalities, processing efficiency was increased by implementing memory optimization techniques. Notably, there were a couple of anomalies' which were excluded from further analysis due to unspecified irregularities. The series_id values are utilized as unique identifiers for individual datasets. In feature engineering, timestamps were transformed, rolling window features were generated for subtle temporal patterns, and new metrics, such as "lids," for extended periods of physical inactivity were created. A custom metric 'lids' was introduced, representing long-term physical inactivity indicative of sleep onset. Rolling window features (median, mean, max, min, variance) were created for columns like 'enmo' and 'anglez,' capturing temporal patterns over varying periods (60 seconds to 3600 seconds). Rigorous data cleaning methods were applied to handle missing values, and noise reduction was achieved through imputations, ensuring the reliability of the dataset for accurate sleep prediction.



A heatmap revealed 27 features highly correlated with the target variable ('awake'), guiding feature selection. These features were further refined for subsequent modeling steps. Basic statistics, including mean and quantiles, were computed for relevant features, providing insights into dataset distribution and variability. The optimized dataset, rich in meaningful features, would now be poised for machine learning model application to address sleep detection challenges in wearable technology effectively.

Accelerometer data is used to train several machine learning models for the purpose of detecting the onset and wakefulness of sleep. A neural network, Random Forest (RF), XGBoost, CatBoost, and LightGBM are among the models used. We began by taking the base models for each of these methods. The RandomForestClassifier is used in the implementation for the RF model, using the default settings. The custom parameters found in xgb_params are used to configure the XGBoost model. Three dense layers with dropout regularization, the Adam optimizer, and a binary cross-entropy loss function comprise the neural network design. For the CatBoost model with custom parameters in cat_params, CatBoostClassifier is used. Lastly, LightGBM makes use of the LGBMClassifier using the lgb_params options. After the preliminary result from each of these methods, we fine-tuned on the CatBoosting, XGBoosting and RandomForest methods.

The models are trained using a cross-validation strategy, with each model undergoing multiple folds to ensure robust evaluation. We have leveraged multicore processors (n_jobs=-1) for parallel processing where applicable, especially for tree-based models.

Experimental Fine-Tuning & Results:

Having observed the performance of each of these models, we realized that some of these models weren't able to learn sufficiently well and produced low prediction accuracies. The data exhibits sequential dependencies and patterns that are crucial for accurate predictions, Models like LightGBM and XGBoost are tree-based and naturally capable of capturing sequential relationships, making them well-suited for time series based tasks such as these. Fine-tuning the architecture, hyperparameters, and feature engineering techniques specific to the temporal characteristics of sleep cycle data potentially helped us enhance the performance of all of these models, including the MLP. We could justify the performance demonstrated by each of these models due to the following reasons:

- **Multi-Layer Perceptron (Neural Network):** The MLP's design complexity and the particulars of the sleep cycle data may have an effect on how well it performs. MLPs and other neural networks are excellent at capturing intricate correlations, but they can overfit, particularly when faced with little data, which might be the case here. To better capture the subtleties of sleep patterns, the neural network architecture—including the number of layers and neurons—may need to be adjusted.
- **Random Forest:** The Random Forest technique is a group approach that constructs several decision trees and combines their forecasts. Robustness against noise and anomalies in the data can be achieved with this ensemble approach. However, the randomization and feature selection procedures might have an impact on performance, therefore optimizing these factors was essential for the context of sleep cycle data, which allowed us to get impressive results. This suggested us to pursue more tree-based models.
- **XGBoost:** As a gradient boosting approach, XGBoost generates a sequence of weak learners to fix mistakes in earlier models. It is useful for capturing sequential dependencies in time series data because of this property. The degree to which the model adjusts to the temporal features of sleep cycle data depends critically on the parameters that are set forth, which was fine-tuned through trial-and-error.

- **LightGBM:** LightGBM's superior performance can be attributed to its leaf-wise tree growth strategy, a methodology tailored for intricate temporal patterns inherent in time series data. By focusing on specific data points pivotal for accuracy improvement, the leaf-wise approach allows LightGBM to efficiently capture the nuanced and complex dependencies present in sleep cycle data. This targeted growth mechanism enables the model to discern subtle patterns and variations, making it particularly effective for tasks such as sleep cycle detection where detailed temporal understanding is crucial. As a result, LightGBM's strength in handling intricate time series patterns contributes to its exceptional performance in accurately predicting sleep states.
- **CatBoost:** CatBoost is an excellent categorical feature handler, and its inherent support for categorical features may result in better performance if the sleep cycle data includes categorical factors that are important for prediction. Its symmetric tree design and ordered boosting may also help with accurate modeling of the temporal dependencies in other cases. But as observed, this wasn't the case as analysis of the feature importances demonstrated that the features which contributed the most weren't in fact categorical. But this method still yielded good results, similar to the XGBoost approach.

Some of the parameter choices made during fine-tuning include the use of a moderate learning rate (0.03) with 850 boosting rounds selected. Subsampling (subsample=0.9) and regularization terms (min_child_weight, gamma, alpha, lambda) were introduced to control overfitting and enhance model robustness and for the neural network experiment we used dropout layers to prevent overfitting (Dropout rate=0.5), Adam optimizer for binary cross-entropy loss function, in addition to early stopping and model checkpoint callbacks to enhance training efficiency.

Model	F1 Score	Accuracy
Random Forest	0.9842	0.9793
XGBoost	0.9882	0.9845
Neural Network	0.9436	0.9485
CatBoost	0.9850	0.9804
LightGBM	0.9906	0.9877

These comparisons provide insights into the relative performance of each model, with LightGBM outperforming the rest. The differences highlight the strengths and weaknesses of each model in capturing the intricacies, as ellaborated before, of the sleep cycle data.

These parameter choices were made based on a balance between model complexity and the potential to capture subtle patterns in sleep cycle data. Hyperparameters were fine-tuned to prevent overfitting, improve generalization, and enhance the models' ability to discern temporal dependencies in the dataset. The iterative nature of the parameter tuning process involved experimenting with different values to optimize each model's performance on the given sleep cycle data.

Discussion:

The study presented here demonstrates significant advancements in sleep detection using accelerometer data from wrist-worn devices, employing a range of machine learning techniques including LightGBM, Random Forest, XGBoost, CatBoost, and Neural Networks. The results, particularly the high accuracy achieved by the LightGBM model, are indicative of the effectiveness of the methods used.

The significant jump in our work is apparent when compared to the GGIR package that is dedicated to preprocessing and simple analysis of data for the R programming language. Although GGIR addresses the fundamental elements of accelerometer data analytics, our project uses sophisticated machine learning models. The fact that our project was able to do highly advanced analytical, transitioning from just basic processing of data into predictive analytics proves it.

In work done by Kay Jan Wong and Brian John the subtle differences in the performance of CatBoost, XGBoost, and LightGBM give an important backdrop. These comparisons highlight as to why LightGBM could be an outstanding performer that surpassed its counterparts in our project; based on the ability to effectively process large datasets and manage categorical data. These insights are essential in appreciating what informs on such models as a sleep detection system.

However, our project is characterized by highly sophisticated feature engineering where we incorporated time-based and seasonal features as well as the implementation of ensemble methods with improved model interpretability. The sophistication becomes very clear when it is compared with the conventional strategies used in the area, as well.

Joseph Rocca points out that in our models, use of ensemble methods improve the strength and reliability of your prediction. Employing boosting and bagging methods applied for instances of LightGBM and Random Forest clearly shows improvement compared with less sophisticated modeling procedures. Moreover, using Neural Networks although problems such as overfitting prevail is also a modern technique utilized in that sphere.

Conclusion:

In summary, this work effectively created and assessed machine learning models for the use of accelerometer data from wrist wearers in the detection of the stages of wakefulness and sleep onset. XGBoost, Random Forest, Neural Networks, CatBoost, and LightGBM are just a few of the many algorithms we used to methodically examine several strategies to enhance accuracy for wearable devices. The models were trained with rigorous preprocessing of the data, memory optimization, and sophisticated feature engineering to improve their capacity to identify the delicate temporal patterns present in the sleep cycle data. Although Random Forest, XGBoost, and CatBoost showed good performance, the LightGBM model's remarkable outcomes highlighted how well its leaf-wise tree growth technique captured intricate temporal connections. The findings of this study contribute valuable insights to the field of health and wellness monitoring systems, highlighting the importance of algorithm selection and parameter tuning for accurate sleep detection in wearable devices, thereby paving the way for the development of more reliable and efficient sleep monitoring systems in wearable technology.

References:

1. Sungkyu Park, Cheng-Te Li, Sungwon Han, Cheng Hsu, Sang Won Lee, Meeyoung Cha. Learning Sleep Quality from Daily Logs.
2. Felipe O. Silva, Lucas P. S. Paiva, and Gustavo S. Carvalho. Error Analysis of Accelerometer- and Magnetometer-Based Stationary Alignment.
3. Discovering 135 Nights of Sleep with Data, Anomaly Detection, and Time Series - Medium article.
4. Kalaivani Sundararajan, Sonja Georgievska, Bart H. W. te Lindert, Philip R. Gehrman, Jennifer Ramautar, Diego R. Mazzotti, Séverine Sabia, Michael N. Weedon, Eus J. W. van Someren, Lars Ridder, Jian Wang & Vincent T. van Hees. Sleep classification from wrist-worn accelerometer data using random forests.
5. Accelerometer data processing with GGIR package for R programming.
6. Brian John. When to Choose CatBoost Over XGBoost or LightGBM [Practical Guide] - MLOps Blog, neptune.ai.
7. Kay Jan Wong. CatBoost vs. LightGBM vs. XGBoost.
8. Joseph Rocca. Ensemble methods: bagging, boosting and stacking
9. Bagging with Neural Networks Best practices - Discussion on datascience.stackexchange.
10. LightGBM algorithm: The Key to winning machine learning competitions - Data Aspirants