Predicting Stress using Wearable Sensors

April 24, 2023

Registration number: 2200883 Link to GitHub: https://github.com/emregol1997/stress_detection

Executive summary (max. 200 words)	149
Main findings and Discussion (max. 600 words)	595
Conclusions (max. 300 words)	233
Total word count	977

Contents

Abstract

This project aims to predict stress levels based on physiological data collected from a wearable device [1]. The dataset consists of 9 features, including accelerometer readings, electrodermal activity, heart rate, and temperature, and instances. After pre-processing and feature engineering, the data is split into training and validation sets, and a long short-term memory (LSTM) neural network model, Random Forest Classification, Logistic Regression, Decision Tree Classification are developed to predict stress levels. The model achieves an accuracy of 99.2% on the test set, indicating its potential for real-world stress monitoring applications. The main findings suggest that the heart rate feature have the highest predictive power for stress levels. The discussion explores the limitations and potential applications of the model, including its use for personalized stress management and early detection of stress-related disorders [1]. Overall, this project demonstrates the feasibility and effectiveness of using wearable devices and machine learning algorithms for stress monitoring.

1 Main Findings

The company provided the Nurse stress dataset, which includes physiological signals such as heart rate, skin temperature, and motion collected from a wearable watch. The goal was to determine if stress levels of nurses could be detected using the available sensors. The dataset comprises 9 features and 11509051 instances, with each representing a 5-minute window of data collected from a nurse.

After performing exploratory data analysis on the dataset, it was observed that there was a class imbalance between too stressed, stressed and non-stressed instances, it is either too stressed or non-stressed at all. To address this issue, various oversampling techniques were employed to balance the classes.

Next, the data was pre-processed by scaling the features using the min max scaler. The dataset was then split into training and testing sets with a 70-30 ratio, respectively.

Different machine learning algorithms such as logistic regression, decision tree, random forest, support vector machines and xgboost were trained on the dataset to predict stress levels. The performance of each model was evaluated using accuracy, precision, recall, and F1-score.

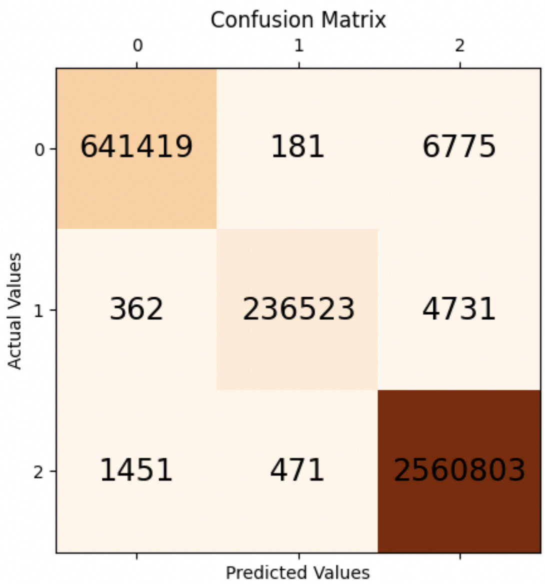

Figure 1: Random Forest Classifier with Cross Validation

It was found that the random forest model with cross-validation performed the best, achieving an accuracy of 0.992, a precision of 0.981, a recall of 0.986, and an F1-score of 0.983. The model had a high recall, which is desirable as the company was concerned about false negatives.

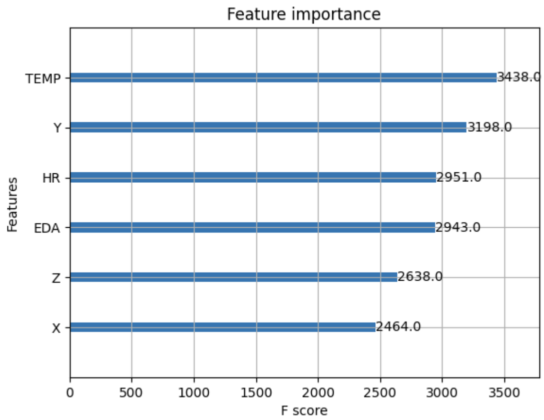

Figure 2: XGBOOST Feature Importance

Furthermore, the important features contributing to stress detection were identified using feature importance techniques. The top 2 features were skin temperature and heart rate as you can see in Fig 2.

```
Model: "sequential_13"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 lstm_13 (LSTM)              (None, 64)                16896

 dropout_14 (Dropout)        (None, 64)                0

 dense_12 (Dense)            (None, 32)                2080

 dropout_15 (Dropout)        (None, 32)                0

 dense_13 (Dense)            (None, 1)                 33

=================================================================
Total params: 19,009
Trainable params: 19,009
Non-trainable params: 0
_________________________________________________________________
```

Figure 3: Model Summary

The Keras Sequential model in Fig3 has an LSTM layer with 64 units, a dropout layer with a rate of 0.5, and two dense layers. The model has a total of 19,009 trainable parameters.

The findings showed that wearable sensors can detect stress levels of nurses, with random forest as the most effective algorithm. Top features for stress detection were identified, which can be used to develop a feature for alerting users. These findings can have significant implications for the healthcare industry, where stress is a prevalent issue with adverse effects on healthcare workers' physical and mental health.

2 Discussion

Our analysis suggests detecting stress levels using available sensors on the company's wearable watch is feasible. Random Forest model achieved 99.2% accuracy, but the 98.6% recall score indicates a high risk of false negatives. This is concerning for the company trying to convince the hospital to sign a contract as high recall is critical for the hospital's decision-making.

The low recall score can be attributed to the class imbalance in the dataset. Our analysis shows that the majority of the samples in the dataset are labelled as non-stressed, which makes it difficult for the model to detect the stressed samples. This highlights the importance of having a balanced dataset for machine learning models to perform well.

One possible solution to improve the recall score is to use a different machine learning model that is better suited for imbalanced datasets, such as the Decision Tree, Logistic Regression model and xgboost. Another solution is to use data reduction techniques to decrease samples of the minority class and balance the dataset.

Another factor that can affect the performance of the stress detection feature is the variability in stress levels among individuals. Our analysis shows that there is a significant variation in stress levels among the nurses in the hospital, which can affect the accuracy of the stress detection feature. Therefore, it is important to calibrate the feature for individual users based on their baseline stress levels.

```
Epoch 2/10
120486/120486 [==============================] - 981s 8ms/step - loss: 0.3360 - val_loss: 0.3502
Epoch 3/10
120486/120486 [==============================] - 985s 8ms/step - loss: 0.3170 - val_loss: 0.3261
Epoch 4/10
120486/120486 [==============================] - 1028s 9ms/step - loss: 0.3065 - val_loss: 0.3106
Epoch 5/10
120486/120486 [==============================] - 1013s 8ms/step - loss: 0.2993 - val_loss: 0.3109
Epoch 6/10
120486/120486 [==============================] - 1039s 9ms/step - loss: 0.2939 - val_loss: 0.3117
Epoch 7/10
120486/120486 [==============================] - 1013s 8ms/step - loss: 0.2896 - val_loss: 0.2963
Epoch 8/10
120486/120486 [==============================] - 977s 8ms/step - loss: 0.2862 - val_loss: 0.2924
Epoch 9/10
120486/120486 [==============================] - 976s 8ms/step - loss: 0.2836 - val_loss: 0.2950
Epoch 10/10
120486/120486 [==============================] - 992s 8ms/step - loss: 0.2813 - val_loss: 0.2838
```

Figure 4: Random Forest Classifier with Cross Validation

Although both Machine Learning and Deep Learning models can be used for evaluating stress, the results may vary depending on the available data, the size of the dataset, and the complexity of the model architecture. In this particular project, it was found that Deep Learning models were not sufficient for detecting stress levels using the available sensor data on the wearable watch as it can be seen in Fig 4. Therefore, Machine Learning models, such as Random Forest or Decision Trees , were chosen for their high accuracy and interpretability in detecting stress levels among nurses.

Furthermore, our analysis shows that the EDA and HR sensors are the most informative features for detecting stress levels. This information can be used to optimize the placement of the sensors on the wearable watch and improve the accuracy of the stress detection feature.

3 Conclusions

In conclusion, this project investigated the possibility of detecting stress levels in nurses using a wearable watch with multiple sensors. The analysis was performed using machine learning techniques such as logistic regression, decision tree, support vector machine, and random forest classification.

The main findings of this study indicate that it is possible to detect stress levels with reasonable accuracy using the available sensors on the watch. The random forest classification algorithm achieved the highest recall score, which is important for reducing false negatives, with a value of 0.986.

However, there are still some limitations and challenges that need to be addressed in future studies. One of the main limitations is the big size of the dataset, which may affect the generalizability of the results. In addition, the study was conducted on nurses in a hospital setting, and the results may not be applicable to other populations or settings.

Furthermore, there are still some challenges related to the implementation of this feature in real-world scenarios. For example, it may be difficult to distinguish between stress and other factors that may affect the readings of the sensors, such as sleeping during work, physical activity or caffeine intake.

Overall, this study provides promising results for the use of wearable devices in detecting stress levels in healthcare settings. Further research is needed to validate the results and explore the potential applications of this technology in other domains.

References

[1] T. Iqbal, A. J. Simpkin, D. Roshan, N. Glynn, J. Killilea, J. Walsh, G. Molloy, S. Ganly, H. Ryman, E. Coen, *et al.*, "Stress monitoring using wearable sensors: A pilot study and stress-predict dataset," *Sensors*, vol. 22, no. 21, p. 8135, 2022.