Conclusions

Standard supervised learning models (e.g., RF, SVM) are best at stress prediction among common models. 

Neural network models performed equally well or better for overall metrics:
- Good potential for further use of deep learning in stress prediction.

Random Forest, Decision Tree, and Multilayer Perceptron (MLP ANN) performed well.

More data outweighs more relevant timeframe.

Large time windows (up to 3 hours) yielded better accuracies.

More data outweighs more relevant timeframe.

Strong correlation between heart-related metrics and model efficacy.

Future Work

- Test latency of real-time stress prediction on Fitbit mobile devices.
- Compare efficiencies of larger deep learning models on mobile devices with efficiencies of standard supervised learning models.
- Consider addition of other features such as location contextualization or emotional responses.

References and Related Works

Analyzing the Features and Efficacy of Classification Models for Physiological Stress Prediction

Allen Lin, Computer Science
Mentor: Dr. Ming Zhao, PhD, Professor
School of Computing and Augmented Intelligence

Data Processing

- Binary classification threshold tuning down to 0.3 due to unbalanced dataset.
- SMOTE-ENN for both over-sampling and under-sampling.
- Principal component analysis (PCA) to reduce components.
- Divided dataset into different intervals of time to investigate ROC_AUC.

Feature Processing

- Biometric data to inputted stress levels based on certain time interval of relevance.
- Heart Rate
- Calories, METs, and Steps

Feature Extraction

- Removal of non-responses and other noisy data.
- Mapped biometric data to inputted stress levels based on certain time interval of relevance.

Model Type

<table>
<thead>
<tr>
<th>Model Type</th>
<th>Accuracy %</th>
<th>Recall %</th>
<th>ROC_AUC %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random Forest</td>
<td>0.694</td>
<td>0.802</td>
<td>0.738</td>
</tr>
<tr>
<td>XG Boost</td>
<td>0.755</td>
<td>0.697</td>
<td>0.737</td>
</tr>
<tr>
<td>Decision Tree</td>
<td>0.762</td>
<td>0.623</td>
<td>0.696</td>
</tr>
<tr>
<td>MLP ANN</td>
<td>0.775</td>
<td>0.67</td>
<td>0.737</td>
</tr>
<tr>
<td>CNN</td>
<td>0.784</td>
<td>0.674</td>
<td>0.738</td>
</tr>
</tbody>
</table>

ROC_AUC Score vs Time Frame

ROC_AUC Score Analysis Across Different Time Intervals

Feature Importance for 5 Strongest Models

<table>
<thead>
<tr>
<th>Feature</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resting Heart Rate</td>
<td>Random Forest, XG Boost, Decision Tree</td>
</tr>
<tr>
<td>Max Steps</td>
<td>Random Forest, Decision Tree</td>
</tr>
<tr>
<td>Mean HR</td>
<td>Random Forest, Decision Tree</td>
</tr>
</tbody>
</table>

Conclusions

- RF, XG Boost, and Decision Tree are best at stress prediction among common models.
- Consistent with previous studies.
- More data outweighs more relevant timeframe.
- Larger time windows (up to 3 hours) yielded better accuracies in 30, 60, and 120 minutes.
- Improvement in accuracy, recall, and ROC_AUC.
- MLP ANN yielded better performance.

Results

Efficacy of Each Model Type

<table>
<thead>
<tr>
<th>Model Type</th>
<th>Accuracy</th>
<th>Recall</th>
<th>ROC_AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random Forest</td>
<td>0.694</td>
<td>0.802</td>
<td>0.738</td>
</tr>
<tr>
<td>XG Boost</td>
<td>0.755</td>
<td>0.697</td>
<td>0.737</td>
</tr>
<tr>
<td>Decision Tree</td>
<td>0.762</td>
<td>0.623</td>
<td>0.696</td>
</tr>
<tr>
<td>MLP ANN</td>
<td>0.775</td>
<td>0.67</td>
<td>0.737</td>
</tr>
<tr>
<td>CNN</td>
<td>0.784</td>
<td>0.674</td>
<td>0.738</td>
</tr>
</tbody>
</table>

Comparison of 5 Strongest Models

- RF, XG Boost, and Decision Tree are best at stress prediction among common models.
- Consistent with previous studies.
- More data outweighs more relevant timeframe.
- Larger time windows (up to 3 hours) yielded better accuracies.
- Improvement in accuracy, recall, and ROC_AUC.
- MLP ANN yielded better performance.

Future Work

- Test latency of real-time stress prediction on Fitbit mobile devices.
- Compare efficiencies of larger deep learning models on mobile devices with efficiencies of standard supervised learning models.
- Consider addition of other features such as location contextualization or emotional responses.

Acknowledgements

Many thanks to Dr. Ming Zhao, Sang-Hun Sim, and Tara Paranjpe (VISA Lab), as well as Dr. Nicole Roberts (ASU Emotion Lab) and the Phoenix Police Regional Academy. Thank you to the National Science Foundation (NSF) Award CNS-1955593.

References