Towards ergonomics working - machine learning algorithms and musculoskeletal modeling

Franz Suess, Maximilian Melzner and Sebastian Dendorfer
Laboratory for Biomechanics, Ostbayerische Technische Hochschule (OTH), Galgenbergstraße 30, 93053 Regensburg, Germany

E-mail: sebastian.dendorfer@oth-r.de

Abstract. Ergonomic workplaces lead to fewer work-related musculoskeletal disorders and thus fewer sick days. There are various guidelines to help avoid harmful situations. However, these recommendations are often rather crude and often neglect the complex interaction of biomechanical loading and psychological stress. This study investigates whether machine learning algorithms can be used to predict mechanical and stress-related muscle activity for a standardized motion. For this purpose, experimental data were collected for trunk movement with and without additional psychological stress. Two different algorithms (XGBoost and TensorFlow) were used to model the experimental data. XGBoost in particular predicted the results very well. By combining it with musculoskeletal models, the method shown here can be used for workplace analysis but also for the development of real-time feedback systems in real workplace environments.

1. Introduction
Ergonomic workplaces lead to fewer work-related musculoskeletal disorders and thus fewer sick days [1,2]. The design and study of workplaces is the subject of numerous studies, either directly on site or through an experiment in a laboratory where the work environment is replicated. Both methods require time-consuming and expensive preparation with an iterative process for actual optimisation.

Ergonomic workplace assessment tools, such as the Rapid Upper Limb Assessment (RULA) [3], have been developed to assess and minimise risk factors for occupational upper musculoskeletal injury in industrial workers. These include joint positions, the amount of weight-bearing and the amount of weight applied. Furthermore, muscle activity is integrated by describing whether the body position is static or dynamic. However, the assessment of different body postures during an activity is not determined here or only through repeated use. Other guidelines, such as those of the National Institute for Occupational Safety and Health, give values and limits for e.g. intervertebral disc forces during occupational activities [4,5].

Such limits allow for a more accurate analysis of work processes but require tools to determine spinal loading throughout the activity. Musculoskeletal modelling tools (MMT) provide this functionality and have already demonstrated their capabilities in various studies to determine...
ergonomic and workplace analyses and recommendations. For example, they were able to demonstrate their capabilities in material handling in supermarkets [6]. In another experimental setup, Davis et al. [7] were even able to incorporate mental stress with the help of the EMG-controlled MMT. They were able to show that mental stress increases disc strain and thus provides a tool to incorporate psychological factors. As the analyses of mental stress in real-life situations can be difficult and EMG measurements are prone to errors as they can be influenced by artefacts [8], like movements [9], a tool that simulates the movement in combination with mental stress would overcome these problems. A possible solution could be the use of machine learning tools to estimate muscle activity (even under the influence of mental stress) and provide input data for MMT.

Similar approaches, based purely on mechanical loading, have been used in several studies. For gait analysis Heller et al. [10] used machine learning techniques to reconstructed muscle activity dependent on kinematic data like walking speed as well as joint kinematics. With moderate static loading, Nussbaum et al. [11,12] used experimental data to train and predict lumbar muscle recruitment. Input data of thoracic angle and external load predicted measured activity of erector spinae, rectus abdominus, oblique externus, and latissimus dorsi. Luh Chang et al. [13] and Wang et al. [14] used arm muscle activity in combination with kinematics to train neural network to predict joint torque. All of these studies are purely based on biomechanical loading and neglect the influence of mental stress.

Therefore, the aim of this study was to investigate the feasibility of using machine learning models based on experimental data to predict and incorporate a submaximal subject-specific stress-induced muscle activity (MA) response which might be the basis for a more holistic description of the influence of workplace environments on the human body.

2. Methods
41 subjects performed 25-degree trunk flexion/extension cycles with and without exposure to cognitive stress, each lasting 120 seconds. The paced auditory serial addition task (PASAT) [15] was used to model cognitive stress (COS) and provides data for the stressed muscle activity. Prior to the stress test, a baseline measurement was taken, which also lasted 120 seconds, providing non-stress data. During the extension, the subject had to push against 5 % maximum load determined by the maximum voluntary contraction test (MVC) beforehand. A dynamometer with an additional TP 500 adapter (CON-TREX® WS; PHYSIOMED AG) set and recorded the motion, moment, and angle (Figure 1). Electromyography (EMG) sensors (Trigno IM sensors from Delsys) recorded six main lower back muscles (m. multifidi, m. erector spinae pars iliacostalis, m. erector spinae pars longissimus - left and right side each).

During the postprocessing a mean cycle for each subject for baseline and stress is computed. Therefore, the data was split at the upper dead centre, aligned to map a full 100 % cycle representing the mean cycle time. The data from each subject was used to compute a mean cycle for torque, speed, and MVC normalized EMG data.
The subject was attached at the leg, knee, hip and thorax to the device.

The records are divided into baseline and stress data. Furthermore, it is assumed that the subject was maximally stressed under the influence of the COS and absolutely relaxed during the baseline measurement. Based on this assumption, the stress level feature is set to zero for the baseline data and one otherwise. A single training data set contains the kinematic variables: Position, velocity, acceleration, and torque. The subject-specific data weight and height are added to the time series data and provide a sample for training and results in 28297 data points.

2.1. Gradient Boosting
The library XGBoost (2020d) for Python is the basis for this implementation. First, the data is collected, then stored in a pandas DataFrame and written to disk as a Python pickle object. This approach allows the data to be loaded into memory during the computational steps and saves time for device access and file parsing. The training includes the following parameters and their corresponding values, and only values without defaults are listed here. All other values are documented at the official XGBoost documentation [16]:

- eta: 0.1
- gamma: 0
- max_depth: 9
- min_child_weight: 6
- subsample: 0.7

Using these hyperparameters, a machine learning model is trained for each mean EMG channel in the low back as an outcome variable and as a function of subject height, weight, stress level, thoracic flexion angle, thoracic velocity, thoracic acceleration, and applied torque. In the case of the stress level feature, baseline data are coded with a value of zero and one for the stress test. This approach assumes that the subject was in a 100% arousal state and allows the prediction of submaximal muscle activity for a lower stress level.

2.2. Neural Network
Similar to the gradient tree model, a third-party Python library is used for the neural network (NN) implementation. TensorFlow [17], as an open source multi language library, supports various types of
devices and was therefore selected for this implementation. The pre-collected data is used for grid search and Bayesian optimization to perform hyperparameter optimization. The following parameters are part of the optimization: Learning Rate; Epochs; Batch size; Number of neurons in layer 1 - 3. The following table (Table 1) shows the outer boundaries of the search space.

Table 1. Parameter boundaries used for hyperparameter grid search. Bayesian optimization uses the optimal parameters as starting parameters.

| Boundary | Learning rate | Epochs | Batch Size | Layer 1 | Layer 2 | Layer 3 |
|----------|---------------|--------|------------|---------|---------|---------|
| -        | 0.01          | 10     | 50         | 14      | 14      | 14      |
| +        | 0.001         | 50     | 500        | 28      | 28      | 28      |

The result of the grid search is used as the initial value for the Bayesian optimization package (2020a). The parameter space was further extended to include learning rate; neurons in input layer; number of hidden layers; number of neurons in hidden layer; activation function: batch size; epoch size. Tanh, Relu, and linear function served as possible activation functions for best data fit. Predefined parameters and ranges are adopted by Montavon et al. [18].

During the training process, Bayesian optimization results are used to generate the NN. First, all input variables are mapped to a range between zero and one. This is necessary to avoid any bias due to the different scales of variables. Then, the data is fed into the Tensorflow training routine with a validation sample size of five percent. The mean absolute error (MAE) was chosen to quantify the model quality. The training process was stopped if this value did not change over 10 iterations.

The quality of the two muscle activity regression models is evaluated with the training data of their specific trial using MAE and person correlation. To predict the submaximal load response with the models, the baseline data was used again, but with a load value of 0.5.

3. Results
Table 2 and table 3 show the differences between prediction and training data. In general, XGBoost shows better results than the TensorFlow neural network (TNN). This is true for both baseline and stress data. The XGBoost models are able to model the data with an MAE below 2.53% m. multifidi muscle activity. In contrast, the TNN error is much larger with up to 68.6% muscle activity in the m. erector spinae pars iliocostalis (Table 2). The similarity analysis with the Pearson correlation coefficient r (Table 3) shows a strong correlation with over 0.78 for the TNN regression model. In this case, the correlation becomes smaller following the increase of MAE. The XGBoost model shows a strong correlation for all muscle activity models above 0.99. In the case of the 50% stress level prediction, both models show different results. In general, the MAE for the XGBoost prediction is lower and closer to the baseline data. Whereas it has its maximum values with 39% and 48% muscle activity in the m. multifidi, the TNN model has its maximum in the m. erector spinae pars iliocostalis with 68% and 113% muscle activity to the baseline and maximum stress. The Pearson correlation coefficient r is smaller for the TNN model, ranging from 0.75 to 0.91 compared to the baseline data and 0.74 to 0.89 for the stress correlation. The XGBoost model varies from 0.85 to 0.95 and 0.77 to 0.92, respectively.
Table 2. MAE values for the TNN and XGBoost training evaluation data.

| Muscle Type                          | Baseline | Stress | 50% level to baseline | 50% level to stress |
|--------------------------------------|----------|--------|------------------------|---------------------|
|                                      | TNN      | XGBoost| TNN                    | XGBoost             |
| erector spinae pars longissimus      |          |        |                        |                     |
| right                               | 13.55    | 1.01   | 14.11 (4.28)           | 1.08 (0.48)         |
| left                                | 19.06    | 1.12   | 19.59 (5.99)           | 1.33 (0.63)         |
| m. erector spinae pars iliocostalis |          |        |                        |                     |
| right                               | 17.62    | 1.00   | 20.95 (7.7)            | 0.99 (0.69)         |
| left                                | 49.34    | 2.00   | 68.60 (82.34)          | 1.70 (1.81)         |
| m. multifidi                        |          |        |                        |                     |
| right                               | 35       | 2.21   | 45.28 (16.35)          | 2.53 (1.34)         |
| left                                | 34.48    | 2.11   | 32.10 (10.2)           | 1.7 (3.97)          |
| mean                                | 28.17    | 1.58   | 33.44 (12.85)          | 1.56 (1.81)         |

Table 3. Pearson r values for the TNN and XGBoost training evaluation data.

| Muscle Type                          | Baseline | Stress | 50% level to baseline | 50% level to stress |
|--------------------------------------|----------|--------|------------------------|---------------------|
|                                      | TNN      | XGBoost| TNN                    | XGBoost             |
| erector spinae pars longissimus      |          |        |                        |                     |
| right                               | 0.97     | 0.999  | 0.98 (0.03)            | 0.99 (0.008)        |
| left                                | 0.96     | 0.999  | 0.97 (0.04)            | 0.99 (0.003)        |
| m. erector spinae pars iliocostalis |          |        |                        |                     |
| right                               | 0.95     | 0.999  | 0.95 (0.04)            | 0.99 (0.001)        |
| left                                | 0.86     | 0.999  | 0.87 (0.02)            | 0.99 (0.002)        |
| m. multifidi                        |          |        |                        |                     |
| right                               | 0.82     | 0.996  | 0.78 (0.19)            | 0.99 (0.004)        |
| left                                | 0.86     | 0.999  | 0.87 (0.11)            | 0.99 (0.001)        |
| mean                                | 0.90     | 0.999  | 0.90 (0.11)            | 0.96 (0.14)         |

The response of the XGBoost for a random subject and a selected muscle (erector spinae pars longissimus) is given in Figure 2. At the 40% position in the cycle, the prediction and target values increase as the subject moves backwards. The curve for the 50% stress level follows the baseline except in the middle of the backward movement where it moves closer to the stress data. Overall, the predictions are following the experimental data quite well. Results of the TNN model are presented in
Figure 3. The baseline and COS prediction do not follow the target values very closely. The prediction of the load level at 50% reflects this trend.

Figure 2. Result of XGBoost model prediction and target muscle activity of m. erector spinae pars longissimus. The yellow line indicates the thoracic angle from 5° to 30°.

Figure 3. TensorFlow model prediction result and target muscle activity of the m. erector spinae pars longissimus. The yellow line indicates the thoracic angle. The yellow line indicates the thoracic angle from 5° to 30°.
4. Discussion
Machine learning models of lumbar muscle recruitment were developed to predict muscle activity for thoracic movements in response to cognitive stress. Both the XGBoost and TNN models are able to predict the activation pattern as a function of thoracic flexion angle, velocity, torque, and stress level, but with different levels of precision.

In particular, the gradient boost approach shows greater accuracy for the relatively small amount of sample data. For the person correlation as a measure of similarity, this trend is repeated with a higher r-value for the gradient-boost method. The average value of r = 0.9 for TNN found in this study is enhanced by the gradient-boost model which was able to increase the correlation to r = 0.996. With an MAE below 2.53 % in the case of the gradient-boost model, the prediction is quite accurate. Despite a good correlation for the maximum and minimum stress data, this trend is not as accurate for the 50% stress level. In this case, one would expect an approximately equal MAE value when comparing the prediction to the baseline and stress curves, but this is not the case. This situation may be caused by a nonlinear transformation from the baseline to the loading data or by missing features such as lateral torque in the model. In addition, the stress-related kinetic changes are influencing the muscle activity, which may be represented by a non-uniformly distributed MAE value. The accessibility to such models allows numerical studies to identify kinetic parameters causing additional stress-related loads. The example in Figure 3 shows the underestimation and overestimation of muscle activity during backward movement. This leads to a less accurate estimate of submaximals. This trend is smaller in the case of the gradient boost method but is still present.

The relatively small sample size and the unknown actual stress level of the subject directly affect the models and cause the lower accuracy of the TNN model. The use of the data for training and analysis is a direct result of the small sample size. Nevertheless, this approach was taken to establish and identify a procedure if more data were available to train a general model. Prediction of the submaximal stress pattern using kinetic data from the baseline study may not reflect the correct trajectory of motion and may result in a less accurate prediction. In addition, to provide training data for submaximal loads, there is a limitation in accurately determining the stress level. Using an individual scaled value that is independent of personal assessment, submaximal measurements are possible. This option would provide the missing data point and allow for much more accurate training. Therefore, a study design that performs submaximal measurement is necessary to capture a range of stress responses.

The results of this study indicate that the combination of machine learning algorithms and experimental data provides a useful basis for the analysis of stress-induced muscle loads. Combining these methods with musculoskeletal modeling opens up the possibility to study a wide variety of movements and tasks. This could enable more realistic assessments of workplace environments. The future potential of these methods is considerably high. The process of using only very limited input data, e.g., collected by wearables, as control variables for a trained algorithm opens up the possibility of not only evaluating workplaces, but also creating direct feedback for potentially hazardous situations.

5. Conclusion
Despite a number of limitations in this study, the results clearly demonstrate that the combined musculoskeletal and mental load can be considered in ergonomic analyses and that the demonstrated methodology could serve as a basis for novel ergonomic workplace designs by applying the technology directly to workers in, for example, wearable-based setups.
References

[1] Ross P 1994 AAOHN Journal 42 171.
[2] Heidarimoghadam R, Mohammadfam I, Babamiri M, Soltanian A R, Khotanlou Hand Sohrabi M S 2020 International journal of occupational safety and ergonomics : JOSE 1.
[3] McAtamney Land Nigel Corlett E 1993 Applied Ergonomics 24 91.
[4] Nelson GS W H 1981 Work practices guide for manual lifting (U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control, National Institute for Occupational Safety and Health).
[5] Waters T R, Putz-Anderson V, Garg Aand Fine L J 1993 Ergonomics 36 749.
[6] Larsen F G, Svenningsen F P, Andersen M S, Zee M deand Skals S 2020 Ann Biomed Eng 48 805.
[7] Davis K G, Marras W S, Heaney C A, Waters T Rand Gupta P 2002 Spine 27 2645.
[8] Konrad P 2005 The ABC of EMG. A Practical Introduction to Kinesiological Electromyography (USA: Noraxon INC.).
[9] Luca C J de, Gilmore L D, Kuznetsov Mand Roy S H 2010 Journal of Biomechanics 43 1573.
[10] Heller B W, Veltink P H, Rijkhoff N J, Rutten W Land Andrews B J 1993 Biol. Cybern. 69 327.
[11] Nussbaum M A, Chaffin D Band Martin B J 1995 Functional Tissue Engineering 28 1015.
[12] Nussbaum M A, Martin B Jand Chaffin D B 1997 Functional Tissue Engineering 30 251..
[13] Luh J-J, Chang G-C, Cheng C-K, Lai J-Sand Kuo T-S 1999 J Electromyogr Kinesiol 9 173.
[14] Wang L, Buchanan T S 2002 IEEE Trans Neural Syst Rehabil Eng 10 30.
[15] Lejuez C W, Kahler C Wand Brown R A 2003 The Behavior Therapist.
[16] 2020 XGBoost Documentation — xgboost 1.2.0-SNAPSHOT documentation. Available at https://xgboost.readthedocs.io/en/latest/#!.
[17] 2020 TensorFlow. Available at https://www.tensorflow.org/.
[18] Montavon G, Orr G Band Müller K-R 2012 Neural networks: tricks of the trade (Berlin: Springer).