Smart Industry Operations 2022-2023: Group Assignment

Operations Management – Industry 4.0/5.0 - Occupational Stress xAI with Decision Trees / Bias Management

Overview

The aim of this assignment is to apply in practice xAI and bias management methods on an operations management case and specifically on occupational stress in workplaces. Occupational stress is a prime concern for workers in work environments. Using the Occupational Health & Safety definition, occupational stress is "the feeling of excessive and prolonged pressure and demands that exceed the worker's perceived resources, capabilities and skills to cope"[1]. The management of occupational stress is important for Operations Management. Stress affects negatively both both individuals (e.g. wellbeing) but also the organisation as whole (e.g. reduced work satisfaction/performance, workers attrition). It specifically leads to reduced performance, quality of work and productivity, alongside increased risk of human errors and labour accidents.

With the increased use of low cost and wearable sensors, individuals and organisations can work together to detect, understand, prevent, or mitigate occupational stress. Measurements such as Heartrate and Heartrate Variability (HRV), combined with Inertial Measurement Unit (IMU), are commonly used physiological measures that indirectly indicate a worker's state[2]. Such studies have shown that younger people may exhibit an overall higher mean in HRV. Equally, a study on stress prediction by Li et al. [32], using Ecological Momentary Assessments (EMA) and an ensemble of regressions to model stress levels of candidates during the COVID-19 pandemic, shows an unequal accuracy in their models for people of different age groups, as well as different lifestyles (living alone vs. living with friends and family). These studies show an underlying need to look into the evaluation of the fairness of models that use measurements to detect stress, as well as looking into bias mitigating techniques to ensure equality and, eventually, equity for all employees in their workplace[3].In their state-of-the-art review paper, Schmidt et al identify some commonly used open-source datasets that have been used recurrently by researchers in the field[4]. These datasets include the WESAD dataset, provided by researchers at Robert Bosch GmbH[5]. This dataset will be used in the present assignment. The labelling of each participant's state in terms of stress was determined through a self-assessment. Biases in the data may arise from such self-assessment, but the data design, and samples may introduce additional biases. The data are from 15 participants with measurements taken over 90 minutes of desk work. The samples are labelled on emotion, and we will retain only emotions: 'baseline' and 'stress' as our output labels. We will ignore samples which are marked 'not defined' or 'amusement'. Some personal attributes of participants included Age, Gender, Weight & Height and Lifestyle. Age and Gender are considered sensitive attributes, even if they contribute to better detecting stress.

---

[1] J. K. Hesselink and A. Jain, "Interventions to prevent and manage psychosocial risks and work-related stress - OSHWiki." http://oshwiki.eu/wiki/Interventions_to_prevent_and_manage_psychosocial_risks_and_work-related_stress

[2] A. M. Nardolillo, A. Baghdadi, and L. A. Cavuoto, "Heart rate variability during a simulated assembly task; influence of age and gender," in *Proceedings of the Human Factors and Ergonomics Society*, 2017, vol. 2017-October, pp. 1853–1857. doi: 10.1177/1541931213601943.

[3] H. Li *et al.*, "Stress prediction using micro-EMA and machine learning during COVID-19 social isolation," *Smart Health*, vol. 23, Mar. 2022, doi: 10.1016/J.SMHL.2021.100242.

[4] P. Schmidt, A. Reiss, R. Duerichen, and K. van Laerhoven, "Wearable affect and stress recognition: A review," Nov. 2018.

[5] P. Schmidt, A. Reiss, R. Duerichen, C. Marberger, and K. van Laerhoven, "Introducing WESAD, a Multimodal Dataset for Wearable Stress and Affect Detection," in *Proceedings of the 20th ACM International Conference on Multimodal Interaction*, Oct. 2018, pp. 400–408. doi: 10.1145/3242969.3242985.

Data Description

"may2014.csv"

The dataset is a processed version of the WESAD dataset.

"wesad_readme.PDF"

This is the full description of the WESAD dataset

PLEASE note that exactly the same data set is also included in Practicals 4 regarding occupational stress (a different data set is employed for decision trees). Similar analysis has been performed in this practical, so the focus of this assessment is to capitalise on the content of Practicals 4 regarding the handling of decision trees and their explainability performance, and apply them on the problem case. So much of the focus lies with your analysis and not so much with data preparation and you can use much of the pre-processing from the practicals with minor adjustments.

Report and Questions to be Answered

The deliverable of the assignment is a complete report compiled as a Jupyter notebook. The organization and the content of your report should be in line with the following guidelines.

1. *Introduction and Exploratory Analysis (15% of assignment mark)*

The introduction should provide a brief overview of the problem, the methodology used to address the problem, and an outline of the report.

The exploratory analysis is where you provide your insights from performing an exploratory analysis of the data. You will include in this as a minimum:

a. Your observations regarding the data distribution from different age groups (SPECIFICALLY up to 27 years old, and 28 years or older), and gender.

b. Your observations from correlation analysis between attributes. Additionally comment on whether any of the two sensitive attributes can be inferred by other attributes.

2. *Stress Model Development with Decision Trees (20% of assignment mark)*

In this step, you will develop different decision tree models to classify data records according to stress. Use only emotions 'baseline' (no stress) and 'stress' as the two categories of classes. Ignore samples which are marked 'not defined' or 'amusement'.

Develop and experiment with different decision trees with the following settings:

max_depth: 3, 4, 5

max_leaf_nodes: between 3 to 10

min_samples_leaf: 1

Select and report 4 of them that produce DIFFERENT outcomes. Do not plot choices which produce similar outcomes. Report your observations on how the max_depth and max_leaf_nodes choices impact on the performance. Don't forget to report performance on both the training and test sets, appropriate for classification tasks. Feel free to use code from the practicals, and especially Practicals 4 but make sure you make appropriate adjustments for this assignment. State your choice or choices regarding the model to adopt, according to the achieved performance and justify it with relevant arguments. You will include in this as a minimum:

a- The cross-validation report outcomes: which would be your decision tree choice according to it?

b- The classification report, including precision, recall, and f1-score, and visualise the Precision-recall curve. Comment on your findings and justify your model choices.

3. *Explainability of Decision Trees (25% of assignment mark)*

In this step, you will seek to offer insights, capitalising on the explainability of the decision trees. Focusing on your selected 4 decision trees (chosen to show different and not identical results, according to the choices made for the maximum depth and number of leaf nodes)

Specifically:

a) Visualise the decision tree structure graphically and comment on how the decision tree makes decisions and point to the visualisation providing examples.
b) Produce text results to make the decision tree transparent. How many nodes each tree has? Which are the decision tree rules?

4. *Bias Management (35% of assignment mark)*

In this step, you will perform the Bias Management analysis that you have practiced in the practicals, but this time for your 4 different choices of Decision Trees. You will perform this twice, first for the AGE sensitive attribute; then for the Gender sensitive attribute.

Include in your analysis:

a). Your insights regarding performance per sensitive category group

b). Your insights regarding performance using fairness metrics (equalised odds, demographic parity difference, false negative rate difference). You can use visualisation aids to plot disparities to support your arguments.

c). Your findings from de-biasing the Decision Tree Models, following (as in the practical 4):

- Reweighting of samples

- Fairness evaluations according to the aforementioned fairness criteria

- Fairlearn postprocessing methods to report false positives / negatives per sensitive group

- Fairlearn postprocessing methods optimising the decision threshold per sensitive group

5. *Overall Quality, Conclusion and Recommendations (5% of assignment mark)*

Provide a summary of your analysis and the choices you made along the way. Highlight the main results you obtained and make recommendations for further analysis.

A general remark on grading: Make sure your code is well-documented and your report is readable, such that what is done is clear and motivated. We will evaluate your code based on what is explained and documented in your report.