

# Stress Detection Using Natural Language Processing (NLP)

Degree programme : BSc in Computer Science  
Thesis advisor : Dr. Souhir Ben Souissi  
Expert : Alain Joray

Healthcare workers are particularly susceptible to stress and burnout. This thesis investigates the feasibility of employing medical progress notes for the construction of NLP-driven stress detection models, aiming to uncover workplace stress among nurses through text.

## Introduction

Nowadays, mental health struggles have become a growing concern for society due to their high impact on individuals, health care systems, and companies. Stress, especially if it becomes overwhelming and prolonged, heavily increases the risk for mental illness as well as medical problems. It is, therefore, a major contributing factor to this concerning structural trend. Early detection and effective management of stress are crucial when it comes to limiting the impact stress has on individuals and society. Most current methods to assess stress used by hospitals or other organizations rely on clinical procedures, such as questionnaires or sensors measuring physical signals, that have inherent limitations such as subjective response bias or intrusiveness. These challenges have boosted the research on more economical, scalable, and unobtrusive methods such as natural language processing (NLP) to assess and detect stress.

## Goal

Health care workers are both systematically important to our society and highly vulnerable to workplace stress. However, due to their demanding schedules, they often lack the time to engage in standard stress assessment procedures. Thus, there is significant potential in detecting workplace stress in health care workers through NLP, utilizing data already generated in their daily tasks. BFH has received two German datasets of patient progress notes written by psychiatric nurses. This bachelor thesis aims to contribute to BFH's research on NLP-driven stress detection by analyzing these unlabeled progress note datasets and through experiments determining whether and under what conditions they can be used to build a stress detection model for nurses. Additionally, the thesis examines the effectiveness of social media data for stress detection, comparing it with progress notes and showcasing the wider applicability of text data in developing stress detection models.

## Results

The thesis' experiments on social media data aligned with existing research, indicating promising potential for NLP-based stress detection models that utilize text data where individuals openly express their opinions and emotions, such as from social media. However, the experiments involving patient progress notes suggested that these unlabeled entries, by themselves, are likely insufficient to serve as a robust base for developing effective machine learning models aimed at detecting workplace stress in nurses.

## Further Research

This thesis's findings do not conclusively dismiss the potential for hidden stress indicators in nurses' progress notes. For future research, it is proposed to collect additional data, like physiological stress markers (e.g., heart rate) from wearable devices that nurses might use while documenting their notes. This supplementary data could help in generating accurate ground truth labels, thereby enabling a more in-depth investigation into the correlation between the writing style of the notes and the nurses' stress levels.



Sarah Mühlemann  
Data Engineering



Word Cloud Showing the Most Used Terms in the Progress Notes (Excluding Stopwords)