

Fibromyalgia Prediction with Machine Learning

project number: 231004

Workshop: Medical Data

names: Vitaly Factor, Mai Shafiki

hosts: Adi Shribman, Dorit Shweiki, Yonatan Bilu

Unraveling the Puzzle of Fibromyalgia Prediction: Shedding Light on the Core of the Issue

Fibromyalgia is a chronic pain disorder, that affects 2% to 6% of the global population, primarily women.

Fibromyalgia has many symptoms that tend to vary from person to person. The main symptom is widespread pain. Other common symptoms include fatigue, sleep disturbances, morning stiffness, headaches, and cognitive difficulties. patients may also experience anxiety and depression.

Fibromyalgia is not just a medical condition; it's a puzzle that affects millions worldwide. Understanding its multifaceted nature, varied symptoms, and diagnostic challenges is crucial to providing better care and support for those living with fibromyalgia.

Challenges In Diagnosis

1. Overlapping Symptoms: Fibromyalgia shares symptoms with other conditions such as arthritis, lupus, and chronic fatigue syndrome, making it difficult to differentiate.
2. Subjectivity of Pain: Pain perception is subjective, and patients may find it challenging to describe their pain accurately, leading to diagnostic confusion.
3. Lack of Objective Tests: The absence of specific biomarkers or imaging abnormalities contributes to the challenge of definitive diagnosis.
4. Misunderstanding and Stigma: Some healthcare professionals and the general public may not fully understand fibromyalgia, leading to skepticism and delayed diagnosis.

Machine Learning-Based Solution

Data Collection: The first step in building the fibromyalgia prediction model involved sourcing data from the UK Biobank database, a comprehensive repository of health-related information.

Feature Selection: From the extensive dataset, a careful selection of relevant features was made. These chosen features were considered to be highly informative for the fibromyalgia prediction task, helping to reduce noise in the data.

Data Preprocessing and Cleaning: To ensure the data's quality and suitability for analysis, thorough preprocessing and cleaning procedures were applied. This included handling missing data, and formatting the data appropriately.

Data Balancing: The UK Biobank database contains less than 1% Fibromyalgia positive Participants. The training process requires a balanced dataset. Downsampling techniques were applied, random selection, and also a clustering selection.

Machine Learning-Based Solution

Machine Learning Model Selection: A machine learning model for classification purposes was selected. Common models utilized in this context include decision trees, support vector machines, logistic regression, and neural networks.

Model Training: The selected machine learning model underwent a training process using a portion of the dataset. During this phase, the model learned patterns in the data that are associated with fibromyalgia.

Model Testing: The model's performance was evaluated by testing it on a separate dataset (typically referred to as a test set). This step gauged the model's ability to make accurate predictions on new, unseen data, providing insights into its accuracy and reliability.

Alternative Machine Learning Approaches for Fibromyalgia Prediction

The research paper titled *"Altered Microbiome Composition in Individuals with Fibromyalgia"* investigates the microbial communities in the digestive tracts of individuals with fibromyalgia.

The research paper highlights significant differences in the gut microbiome composition of individuals with fibromyalgia compared to those without the condition. These findings suggest a potential connection between the gut microbiome and fibromyalgia symptoms, opening the door to further research.

Another notable study, *"POS0257 PAINWAVES: THE POTENTIAL OF MACHINE LEARNING TO DIFFERENTIATE CHRONIC PAIN COHORTS USING ELECTROENCEPHALOGRAPHY"*, investigates the potential of electroencephalography (EEG) data and machine learning to differentiate chronic pain cohorts. While our primary focus remains on Fibromyalgia, exploring studies like this broadens our perspective on how advanced techniques can contribute to a deeper understanding of chronic pain conditions and, ultimately, lead to more accurate predictions and improved patient care

Analysis & Conclusions

Limited Sample Size and Data Completeness

One of the significant challenges we encountered during our machine learning model development to predict fibromyalgia using UK Biobank data was the limitation in sample size and data completeness. Although the UK Biobank is a valuable resource, we found that a substantial portion of the dataset lacked complete responses to the features we initially selected as important for predicting fibromyalgia.

Impact on Model Performance

The limited sample size, coupled with the incomplete data, had a noticeable impact on the performance of our machine learning model. With fewer samples and missing values, the model struggled to generalize effectively and make accurate predictions. As a result, we did not achieve the desired level of accuracy and reliability in predicting fibromyalgia using the UK Biobank data.

Feature Importance Reevaluation

Initially, we believed that certain features were crucial for predicting fibromyalgia based on existing literature and medical knowledge. However, the high prevalence of incomplete responses forced us to reevaluate the importance of these features. It is possible that our assumptions about which features are critical in predicting fibromyalgia may not hold when we have a limited number of complete responses to work with. This highlights the importance of feature engineering and the need to adapt to the dataset's limitations.

In addition, some of these features rely on subjective responses, making their predictability uncertain. We're uncertain about how to deal with this data because it's considered an integral part of the fibromyalgia puzzle.

Leveraging Medical Imaging for Enhanced Fibromyalgia Diagnosis

While we primarily focused on traditional data sources, the potential of medical images, such as brain scans, should be explored for enhancing predictive models. Further research and collaboration with experts in medical imaging may provide valuable insights for improving fibromyalgia diagnosis.