

Springboard Data Science Course

Data Science Capstone Project 1

Orthopedic Biomechanical Features

Michelle Ide - 6/3/2020

Today's medical professionals utilize technology to assist with the interpretation of radiological results. Orthopedic biomechanical measurements, compared to known healthy ranges can indicate disease states. These types of measurements fit well with machine learning algorithms used to classify results. This project proposes to assist professionals in predicting disease states based on abnormalities in biomechanical measurements. The overall goal is to research models that correctly classify as normal/abnormal with an additional level of classification of specific disease states. In medicine, accuracy determines the usefulness of any product, including algorithms, therefore the focus of this project will be determining the model with the highest predictive probability value.

Data

Data for this project was obtained from UCI Machine Learning Repository, <http://archive.ics.uci.edu/ml/datasets/Vertebral+Column#>. (Dr. Henrique da Mota during medical residence period in the Group of Applied Research in Orthopaedics (GARA) of the Centre Medicao-Chirurgical de Radaptation des Massues, Lyon, France).

Included are 2 CSV files, both contain the identical biometric data for each patient,

- 1) Pelvic Incidence
- 2) Pelvic tilt
- 3) Lumbar lordosis angle
- 4) Sacral slope
- 5) Pelvic radius
- 6) Grade of spondylolisthesis

And unique classification columns 1) normal, abnormal and 2) Normal, Disk Hernia, Spondylolisthesis.

There are 310 patients, 100 with normal categories, 210 abnormal

Approach

Data Cleansing:

There are no blank fields or empty values, the only cleaning necessary is to search for outliers using box plot or standard deviation.

Modeling:

In search of the best model various classification algorithms will be tested including KNN classification, Logistic Regression, and Discriminant Analysis to provide improved accuracy on reported results.

Deliverables

The projects' results will be reported to include a) findings b) code with documentation, and c) visualizations to GitHub and, as an option for some users, a website created using Anvil software.