Team 9: Urodynamic Data and Machine Learning

Undergraduate Researchers: Mashal Ali, Nathaniel Choe
Project Manager: Zekun Cao
Faculty Leads: Wilkins Aquino, Jonathan Routh

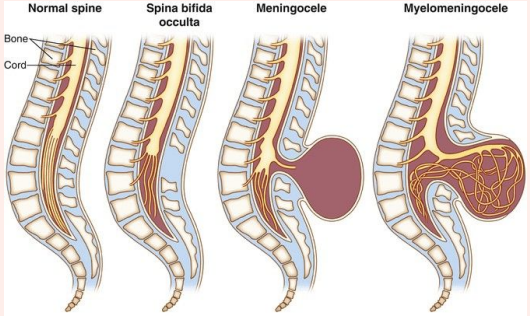
Introduction:

Spina Bifida is a neurogenic bladder dysfunction. It is a neural tube disorder in which the neural tube, the embryonic structure that forms into the baby's brain, spine, and neighboring tissue, and backbone do not fully close, yielding developmental issues as well as infection, renal scarring, and chronic kidney disease. This makes patients more prone to recurrent urinary tract infections (UTI's) and therefore at risk for debilitating bladder and kidney complications. Currently, there is a lack of uniformity in replicating interpretation of urodynamic data, which is pressure data, thus creating the need for a data-driven approach.

Objectives:

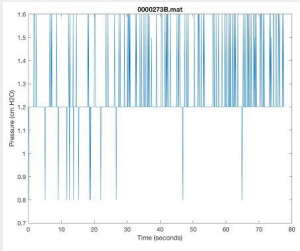
The long-term goal is to assist in better managing long-term health complications for the pediatric spina bifida population. The goals are,

- ➔ develop a machine learning pipeline using urodynamic detrusor pressure data of pediatric spina bifida patients from the Duke University Hospital
- ➔ create a set of visualization tools for the data to supplement clinical interpretations



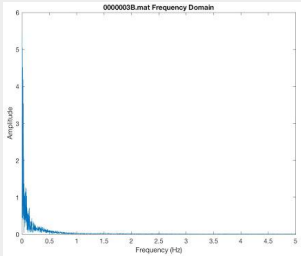
Pipeline:

Clean and Interpret



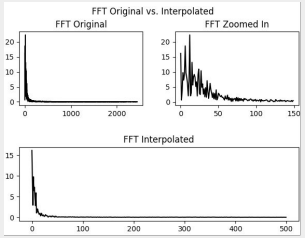
The raw time series data is converted into CSV files. Empty and erroneous files with excessive noise and inhumane pressure increases from movement are removed.

Fourier Transform



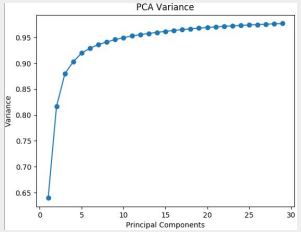
Fast Fourier Transform (FFT) is applied to discount large variations in the pressure values. Further low-pass filtering can also be applied.

Interpolation



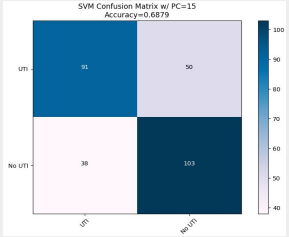
The transformed data is interpolated to the same size to ensure uniformity of the data. This allows for easier dimension reduction.

Dimension Reduction



The data is dimensionally reduced using Principal Component Analysis (PCA); the dataset, reduced to 15 dimensions, had a variance of 96.2%.

Classification

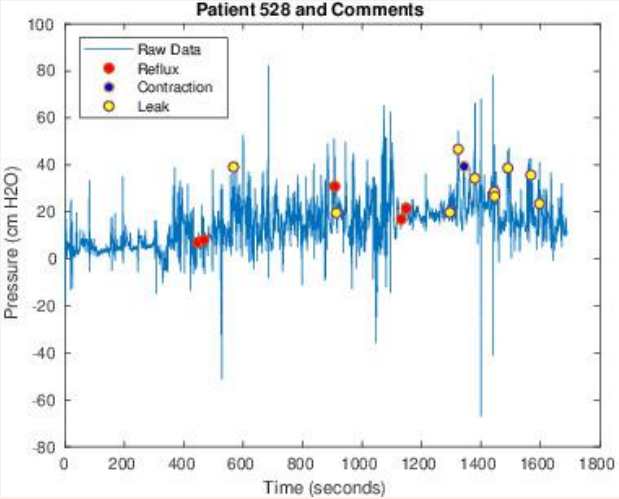


The dataset is now used to train and test a series of classifiers. Data oversampler is used to account for a potentially unbalanced dataset.

Data and Visualization:

757 unique detrusor pressure time-series graphs, collected by Laborie equipment, along with comments, created by the urodynamics staff, were used to obtain empirical results.

Physicians view specific data features and relations like maximum pressure, area under the curve (AUC), and the percentage of data within a certain range of pressure values.



Conclusions:

The pipeline and visualization tools allow urologists to quickly visualize the data in different forms, as well as easily train and test a variety of classification problems. This system can provide information that augment physician diagnoses. In the future, a user interface (i.e: web page) should be built to provide easy access to all the said tools. The user interface would supplement physician recommendations for possible reconstructive surgery to avoid future complications for the pediatric spina bifida population. In addition, upon retrieval and training/testing of labeled data, improvements and/or additions to the pipeline should be made.