A Scheme For Using Machine Learning To Characterize Titanium Dioxide Nanotubes

Jamar Alexander, Alexander Wiley, Eddie Red, Dorian Bohler

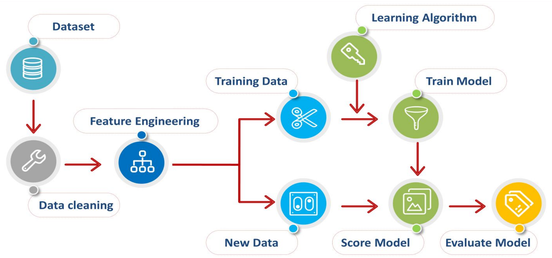
Introduction

Titanium Dioxide Nanotubes (TNTs) are nanostructures with high potential to be used in energy storage and sensors. Students in the R.E.D. Lab have studied methods to reduce defects in TNTs which shall be in Ultrafast Electron Diffraction experiments at SLAC National Lab.

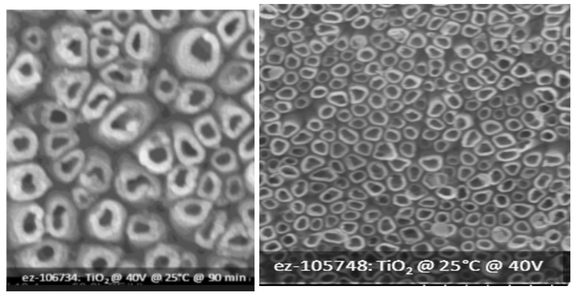
Problem

TNT samples were grown and there is a need to reduce the number of defects by optimizing growth parameters such as the anodization voltage, time, and the components of the solution.. Currently visual inspection methods are used to determine the defects or quality of the TNTs.

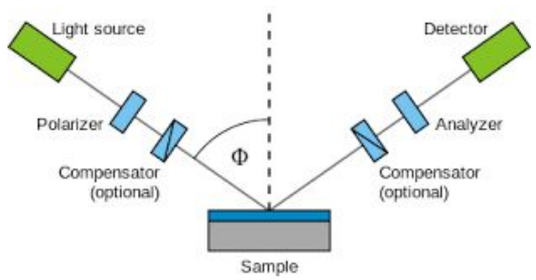
We propose a quantitative method for characterizing the TNTs which employs ellipsometry and machine learning. This method may allow for the determination of TNT length, wall thickness, uniformity, and etc.



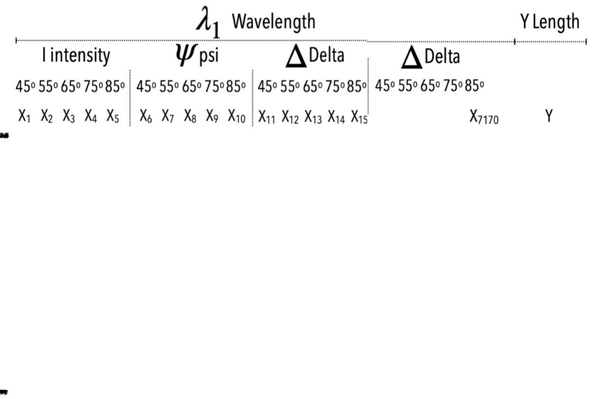
The machine learning process used in materials research.



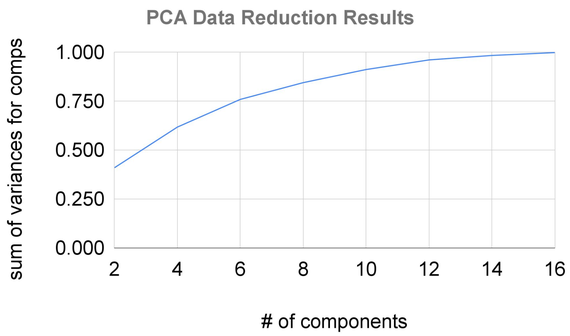
(Left) SEM images of highly un-ordered TNTs grown at Morehouse college (Right) SEM images of highly ordered NTs.



Performing ellipsometry on TNT's would provide the data needed to train an ML model. Models could be developed to predict the physical characteristics of TNTs.



A sample ellipsometry data set from 18 TNTs which results 7171 independent features. This type of dataset is NOT often compatible with training ML models.



Dimensionality reduction is one of the key steps in the feature engineering process. We used Principal Component Analysis (PCA) to reduce the number of training features. The explained variance was used to ensure minimal loss of information from the original data.

..

RESULTS

Using PCA we reduced the number of features from 7171 to 16 while maintaining <99% of the information in the available data set.

NEXT STEPS

Typically <1K samples are needed to train an ML models. We shall investigate several techniques to transform,, simulate, and label the data. Next we would compare the performance of several well known ML algorithms to provide a robust model to predict the TNTs physical parameters.

ACKNOWLEDGEMENTS

US Department of Energy
SLAC National Laboratory
Morehouse College
Dr. Eddie Red & Mr. Dorian Bohler