Motivation and Objectives:

The scope of this project is to be able to predict the future state of a TEM (Transmission Electron Microscope) sample based on the current beam parameters. Since the imaging beam is made of high-energy electrons, there is a natural sample degradation with time as the sample is imaged. This can be detrimental to understanding the original state of the sample, as it is often very difficult to recover the original pristine condition. As such, it would be highly beneficial to operators to be able to predict the state of a sample at some point in the future based on a past sample of images and spectra.

Since this is a multi-term project, the objective of the 1st term (and this class) is to determine the viability of LSTM networks on spectra gathered from the TEM in a time-series forecasting model. If the method is shown to be successful, the 2nd term will focus on refining the model and performance and creating a GUI by which the operators can interact with the model.

User Cases:

@Microscope technician: The microscope technician has no experience in machine learning or python or Jupyter notebooks, but are experts in operating and analyzing data from a TEM. They want to be able to use our developed software to analyze what's happening in the microscope. This will be useful for classifying the data that they obtain or making predictions on reactions happening within the microscope. The technician will want to be able to interact with the raw data from the microscope, and also adjust operating parameters based on the results from the time series predictions.

@Microscope software: some reactions happen faster than operators can realistically control in real time. The time series forecasting model will be able to interact with the microscope software and, based on the predictions of how the reaction is proceeding, inform microscope operating parameters to yield the desired results. This is particularly important for the case of sample damage. After significant exposure to the electron beam, the sample can be damaged; the microscope software would ideally be able to predict how long until significant damage happens under the current operating conditions, and provide a warning to the operator should there be a cause for concern. The software understands code perfectly.

@PNNL ML team: This is a team of experts in machine learning and interpretation of analytical data from the microscope. They will want to be able to see "under the hood" of the software, retrain models and update model hyperparameters, add to the training database, and help operators understand the capabilities of the models. They are also very interested in the contribution to general scientific knowledge from these models, and as such are interested in confidence intervals, how far in the future it is valid to predict the state, etc.

Components:

- A GUI for the operator to interact with (this is a goal for 2nd term)
    - A means of storing and recalling previous analyses
    - No need to look at the actual code, but interact with buttons and visualizations
    - Ability to choose high-level parameters, such as time horizon of forecast or amount of previous data fed into model.

- An LSTM network model used to read incoming data and make time series predictions. This will include receiving spectral data, converting to a proper format for LSTM networks, and then converting the model output back into intuitive visualizations/representations for the operator.
- Software to interpret raw ASCII microscope data (open-source options include HyperSpy and Gatan)
- A means of reading microscope data real-time (potential goal for 2nd term); this includes a repository of incoming data.
- Ability to constantly update the training data repository and update model hyperparamters as more exhaustive datasets are captured.
- The ability to read in microscope operating parameters as well as the raw data, and correlate the two.
- Exhaustive and thorough documentation!

Process Diagram: