The folder contains the materials related to the final project of the course, Introduction to Machine Learning, I passed at Tufts. In this project, we regenerated and assed application of a range of Machine Learning techniques to improve the results from the paper *Computational methods for ubiquitination site prediction using physicochemical properties of protein sequences*.

The focus of the project was on prediction of protein degradation through ubiquitination using physicochemical properties (PCPs) of amino acids in the protein sequences. The project consisted of two main parts:

* Working with data files to generate the training data (preprocessing and engineering steps).
* Apply various machine learning techniques on the training data to develop the best model for prediction of ubiquitination sites.