

Jacob Gursky  
Professor Brown  
Software Design  
4th November 2018

## 90-Second Presentation

One of the most popular machine learning algorithms of all times is linear regression, an algorithm that not only is often quite powerful, but also yields readily interpretable results, as opposed to black box algorithms such as neural networks. Linear regression is the process of deriving a set of coefficients for a given set of predictor variables that describes a response, such as predicting someone's lifespan given various characteristics of their lifestyle. However, given the advent of big data, there is a significant need to make these algorithms faster and more robust to large data sets. The most common implementations of linear regression are present in Python and R, though these languages are often slower than compiled languages such as C++. Therefore, a C++ implementation of linear regression will allow for a scalable implementation of the popular algorithm that enables analysis of large datasets. This project would be relatively straightforward, all that would be necessary would be to import data from a file, and calculate the coefficients of the model using gradient descent, and algorithm for which there is abundant documentation. This project would also be a great addition to the portfolio of any aspiring data scientists for machine learning enthusiasts.