## **Medical Diagnosis using Support Vector Machines**

In this one hour long project-based course, you will learn the basics of support vector machines using Python and scikit-learn. The dataset we are going to use comes from the National Institute of Diabetes and Digestive and Kidney Diseases, and contains anonymized diagnostic measurements for a set of female patients. We will train a support vector machine to predict whether a new patient has diabetes based on such measurements. By the end of this project, you will have created a machine learning model using industry standard tools, and solved a real world medical diagnosis problem.

**Key Concepts**

* By the end of task 1, you will be able to load a dataset from file, and extract its features and labels.
* By the end of task 2, you will be able to split a dataset into training and testing subsets, and normalize the values.
* By the end of task 3, you will be able to create a support vector machine and train it.
* By the end of task 4, you will be able to make a medical diagnosis for a new patient using an SVM.
* By the end of task 5, you will be able to evaluate the accuracy of the SVM classifier.