A machine learning project can be broken down into 6 top-level tasks:

1. Define Problem: Investigate and characterize the problem in order to better understand

the goals of the project.

2. Analyze Data: Use descriptive statistics and visualization to better understand the data

you have available.

3. Prepare Data: Use data transforms in order to better expose the structure of the

prediction problem to modeling algorithms.

4. Evaluate Algorithms: Design a test harness to evaluate a number of standard algorithms

on the data and select the top few to investigate further.

5. Improve Results: Use algorithm tuning and ensemble methods to get the most out of

well-performing algorithms on your data.

6. Present Results: Finalize the model, make predictions and present results.

What we will be covering in this course: (Project Prospective - Pipeline)

. 1: Python Ecosystem for Machine Learning.

. 2: Python and SciPy Crash Course.

. 3: Load Datasets from CSV.

. 4: Understand Data With Descriptive Statistics. (Analyze Data)

. 5: Understand Data With Visualization. (Analyze Data)

. 6: Pre-Process Data. (Prepare Data)

7: Feature Selection. (Prepare Data)

. 8: Resampling Methods. (Evaluate Algorithms)

. 9: Algorithm Evaluation Metrics. (Evaluate Algorithms)

. 10: Spot-Check Classication Algorithms. (Evaluate Algorithms)

. 11: Spot-Check Regression Algorithms. (Evaluate Algorithms)

. 12: Model Selection. (Evaluate Algorithms)

. 13: Pipelines. (Evaluate Algorithms)

. 14: Ensemble Methods. (Improve Results)

. 15: Algorithm Parameter Tuning. (Improve Results)

. 16: Model Finalization. (Present Results)

We will cover 3 basic projects:

**Hello World Project (Iris owers dataset) :** This is a quick pass through the project steps without much tuning or optimizing on a dataset that is widely used as the hello world of

machine learning.

**Regression (Boston House Price dataset) :** Work through each step of the project process with a regression problem.

**Binary Classication (Sonar dataset) :** Work through each step of the project process using all of the methods on a binary classication problem.

How To Install SciPy

There are many ways to install SciPy. For example two popular ways are to use package

management on your platform (e.g. yum on RedHat or macports on OS X) or use a Python

package management tool like pip. The SciPy documentation is excellent and covers how-

to instructions for many dierent platforms on the page Installing the SciPy Stack5. When

installing SciPy, ensure that you install the following packages as a minimum:

. scipy

. numpy

. matplotlib

. pandas

Once installed, you can conrm that the installation was successful. Open the Python

interactive environment by typing python at the command line, then type in and run the

following Python code to print the versions of the installed libraries.