Glossary

**Training Data –** A set of exemplar input and output data from which a computer can learn.

**Model –** A function that can be used to predict outputs for given inputs.

**Supervised Machine Learning** – The computer is given exemplar input data and the respective output data (this is training data). The goal of the computer is to learn a general rule that maps inputs to outputs i.e. to learn a model.

**Classification algorithm** – The input data is divided into two or more classes; the computer learns from the *training data* and produces a *model* that assigns new inputs a class.

**Data Transformation –** Applying the same mathematical function to each and every point in a dataset. For example, square rooting every data point or taking the natural logarithm would be a data transformation.