Introduction :

One-class SVM is an unsupervised algorithm that learns a decision function for novelty detection(also known as outlier detection): classifying new data as similar or different to the training set. In novelty detection, the training data may not be polluted by outliers, and we are interested in detecting anomalies in new observations. Since SVM needs data representing both the classes (in a binary classification system), it can’t be used in the given case.

One Class SVM:

The One-Class SVM has been introduced by Schölkopf for the purpose of novelty detection. It requires the choice of a kernel and a scalar parameter to define a frontier. The RBF kernel is usually chosen although there exists no exact formula or algorithm to set its bandwidth parameter. The \nu parameter, also known as the margin of the One-Class SVM, corresponds to the probability of finding a new, but regular, observation outside the frontier.

Description of various parameters of One Class SVM used (as implemented in scikit-learn):

1. Nu : An upper bound on the fraction of training errors and a lower bound of the fraction of support vectors. Should be in the interval (0, 1]. By default 0.5 will be taken. We have chosen nu=0.1, since we expect minimal outliers in our training data.
2. Kernel: Specifies the kernel type to be used in the algorithm. It must be one of ‘linear’, ‘poly’, ‘rbf’, ‘sigmoid’, ‘precomputed’ or a callable.We have chosen the default RBF kernel.(why?)
3. Gamma : Kernel coefficient for ‘rbf’, ‘poly’ and ‘sigmoid’.

How Does One Class SVM Work In Our Case?