Project 5

Kishan Sarpangala

**Question**:

Train a Random Forest classifier on the MNIST dataset (or a portion of it) and plot each pixel’s importance (like Fig. 7-6 of Jeron).

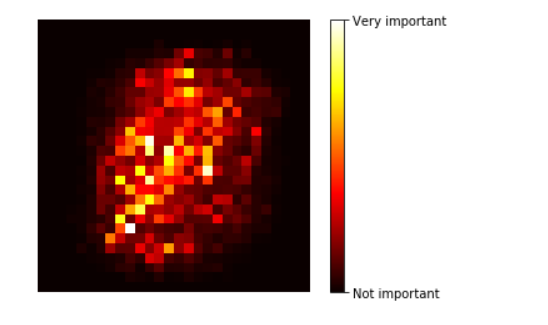
**Answer**:

**What is random forest classifier**?

Random forest are an [ensemble learning](https://en.wikipedia.org/wiki/Ensemble_learning) method for [classification](https://en.wikipedia.org/wiki/Statistical_classification), [regression](https://en.wikipedia.org/wiki/Regression_analysis) and other tasks, that operate by constructing a multitude of [decision trees](https://en.wikipedia.org/wiki/Decision_tree_learning) at training time and outputting the class that is the [mode](https://en.wikipedia.org/wiki/Mode_(statistics)) of the classes (classification) or mean prediction (regression) of the individual trees. Random decision forests correct for decision trees' habit of [overfitting](https://en.wikipedia.org/wiki/Overfitting) to their [training set](https://en.wikipedia.org/wiki/Test_set). [1]

**Pixel’s Importance**:

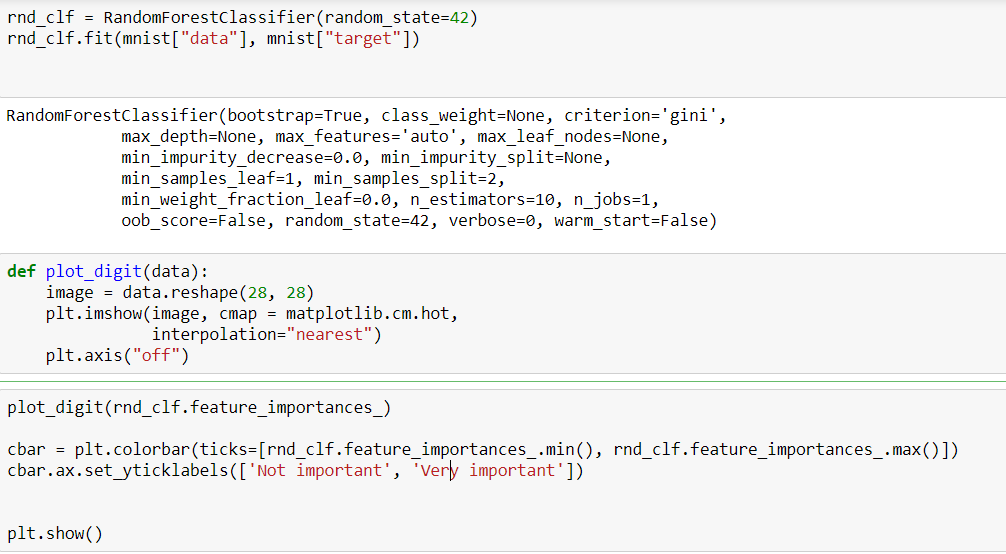
Final image:

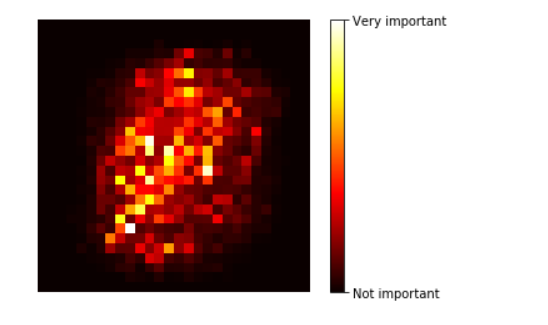


**Output Results**:

* Trained the MNIST dataset in the below code.
* Later tested another set of data and got 95% success rate.
* Clf.score calculated to be 95%. Used Random forest classifier here







**MNIST pixel importance**

**(according to a Random Forest classifier)**