To make pattern recognition problem easy to solve, the input images of digits are typically translated and scaled so that each digit is contained within a fixed sized block (box). This transformation of input variables is known as Preprocessing. Note that new data must also be preprocessed using the same steps as training data.

Pre-processing might also be performed to speed up the computation. Instead of feeding large raw pixel values, the aim is to find useful features which are easy and fast to compute, yet that also preserve useful information for the task. These features will then be used as input to the model. Therefore, preprocessing is also called as Feature Extraction.

Applications in which the training data comprises examples of the input vectors along with their corresponding target vectors are known as Supervised Learning problems. When the aim is to assign each input vector to one of a finite number of discrete categories, then it is called as classification problems. If the desired output consists of one or more continuous variables, then the task is Regression. Example: Digit recognition is a classification problem and predicting yield in a chemical factory is a regression problem.

When training data consists of a set of input vectors $\mathbf{x}$ without any corresponding target values, then it is called Unsupervised learning. The Goal here is may be to discover groups of similar examples (Clustering), or to determine the distribution of data within the input space (density Estimation), or project the data from high-dimensional space to low-dimensions i.e. Dimensionality Reduction.

Finally, the technique of Reinforcement learning is concerned with the problem of finding suitable actions to take in a given situation in order to maximize a reward. Treatment of this technique lies beyond the scope of this book.