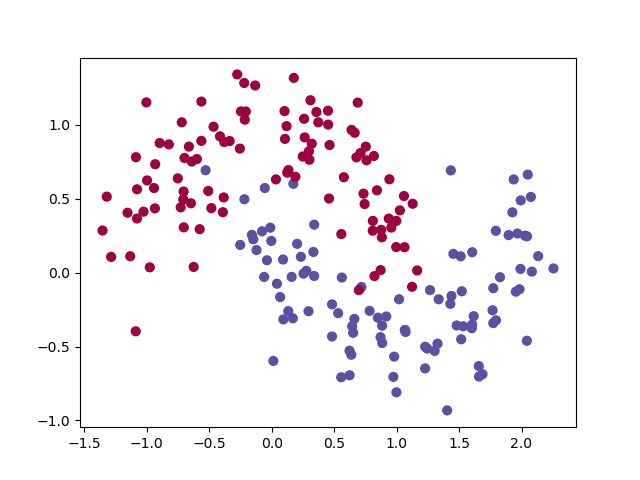
**Assignment1**

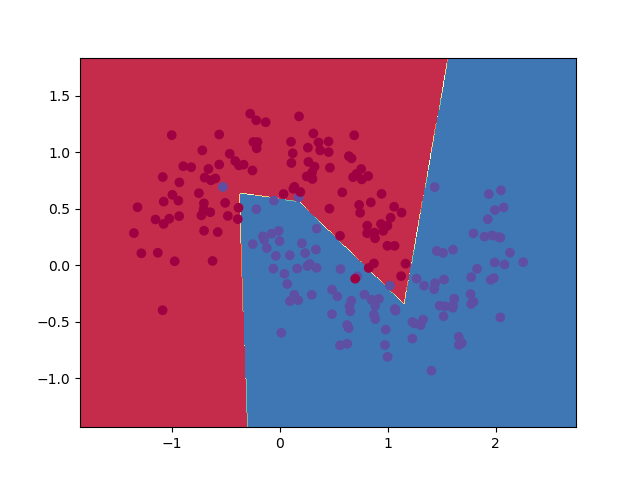
**Author: Ruiqi Kuang**

*TASK1: visualize Make-Moons dataset*

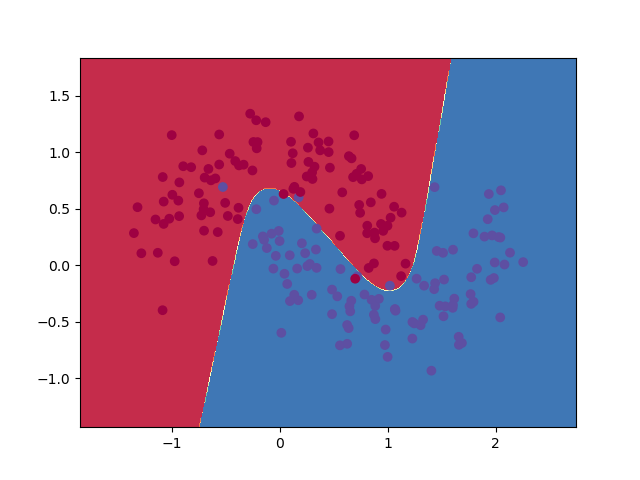


*Task2: train with different activation function*

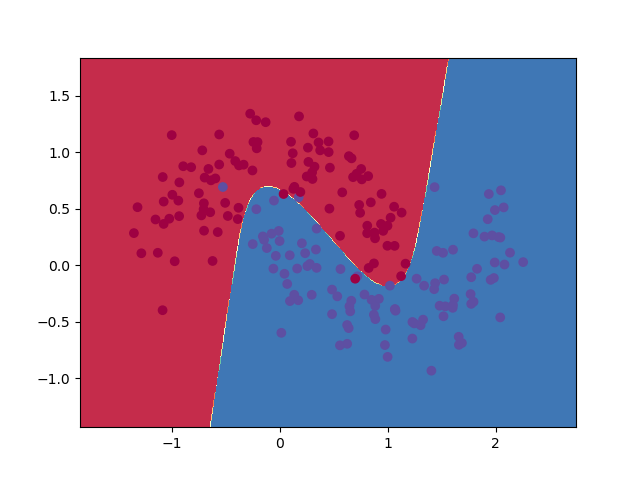
*Relu:*

**

*Tanh:*

**

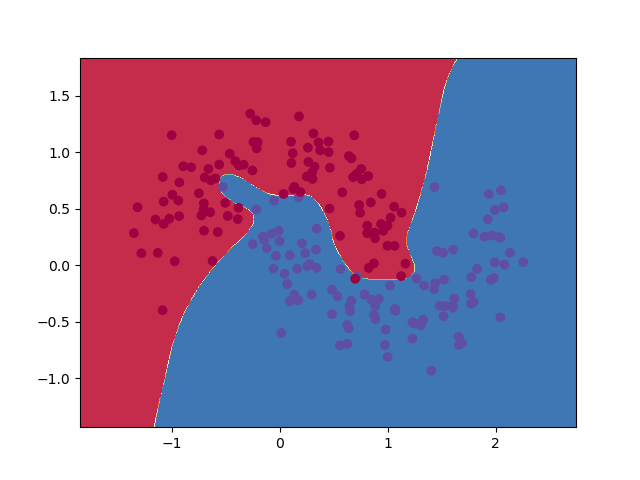
*Sigmoid:*

**

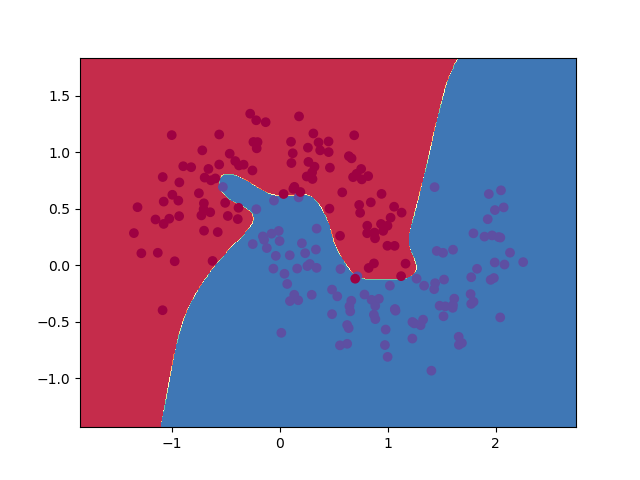
The boundary relu created is several straight lines, while tanh and sigmoid create smooth curves. That may be cause by the fact that relu is a piece-wise function where each part is a proportional function.

*Task3: train with different the number of hidden units*

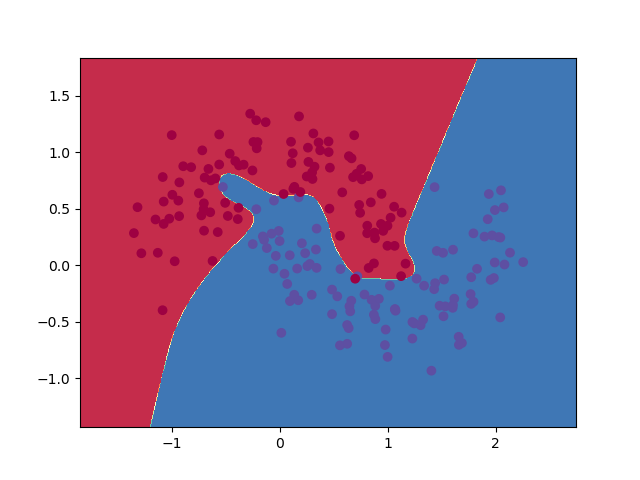
nn\_hidden\_dim=10



nn\_hidden\_dim=20



nn\_hidden\_dim=30



The loss decrease while the dimension of the hidden layer increase. The decision boundary doesn’t change dramatically while dimension increase.