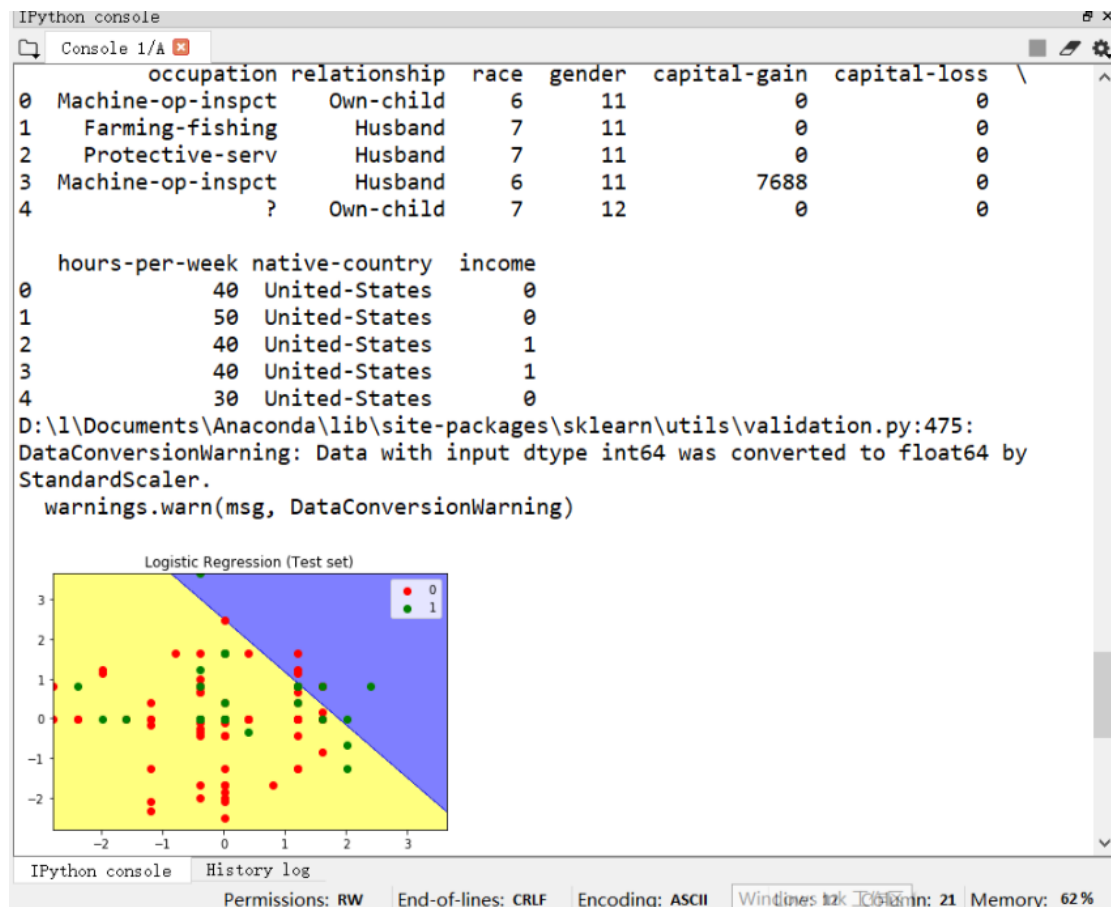


HW1 報告

姓名:劉元坤 學號:U0621006

這次的作業並沒有做出要求的結果，因為有一些 numpy 和 pandas 的內容並不瞭解，不知道他們提供了什麼功能，也不知道我需要什麼功能。雖然在學，但是仍舊有無從下手的感覺。希望老師或者助教可以講解一下這些庫的具體應用方法。



目前程序可以對兩個特徵進行學習，再多加一個特徵就會報錯，我在看 numpy 和 pandas 的文檔，爭取解決。從上圖可以看出訓練的結果並不理想，與實際相差很多，我想很可能是因為兩個特徵確實是有點少，無法使劃分更明確。我也更換了別的特徵進行訓練，發現其它特徵訓練出來的結果相差更多，所以我認為之後在增加其它特徵的時候，教育年限與每週工作時長是兩個非常重要的不可省略的特徵。對於種族與性別，從我之前嘗試的結果來看，我認為會對訓練產生一定影響，但是影響不會太大。資產收益和資產損失由於有太多的 0 值，所以不太容易判斷，但是直觀來看會對年收入產生影響，之後的訓練會增加這兩個特徵。國籍的

話大部分來自美國，所以我認為這個不會對收入產生太大影響。年齡也是一個不可忽略的因素，太年輕或者年齡太大都不利於收入。而教育程度、教育時間是兩個相關的數據，之前講過，教育時間很重要，所以教育程度在一定程度上也能影響年收入。婚姻狀況暫時還未統計，無法得知重要程度。職業對於收入的影響會比較大。