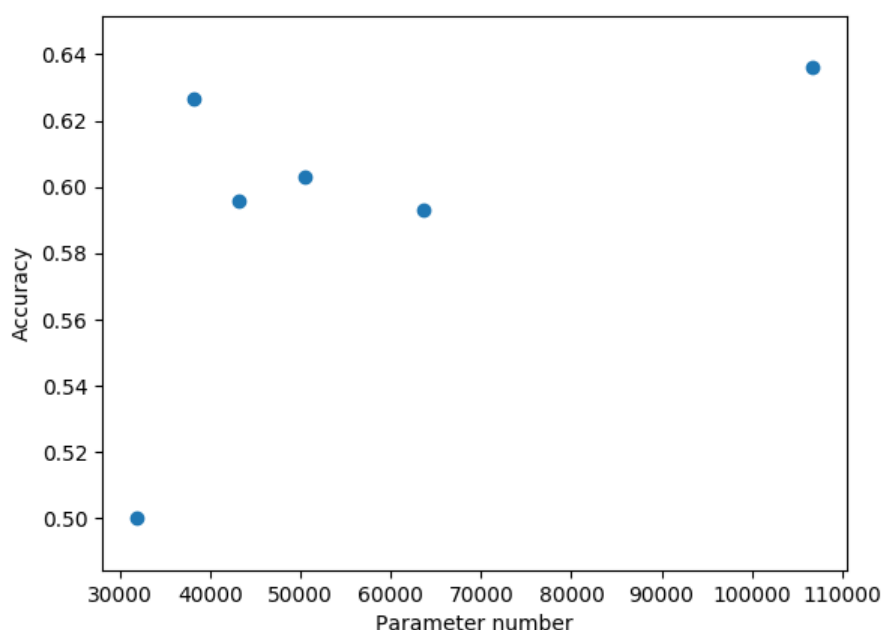


1. 請比較你本次作業的架構，參數量、結果和原 HW3 作業架構、參數量、結果做比較。(1%)

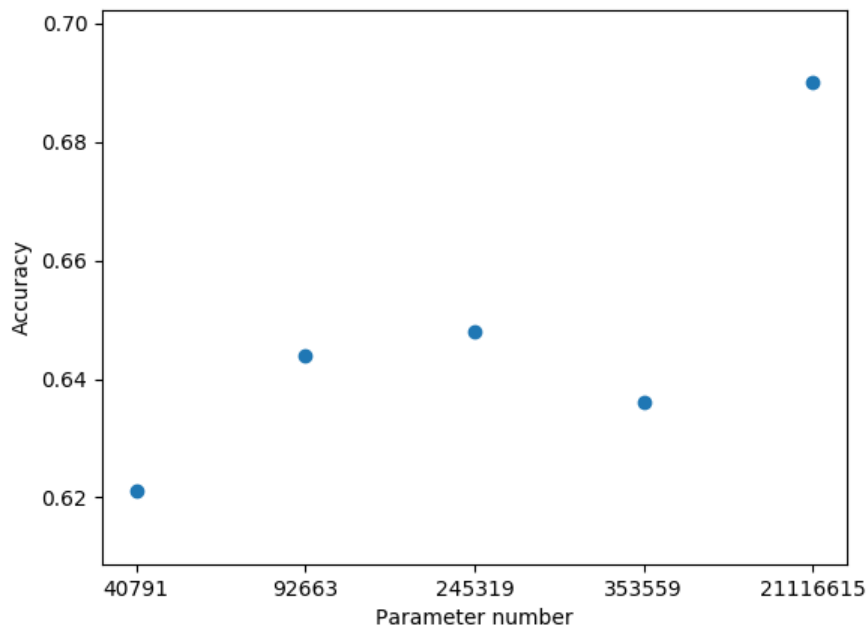
在 HW3 中，我使用 16 層 convolution layers，kernel 的數量分別是 64, 64, 128, 128, 256, 256, 256, 256, 512, 512, 512, 512, 512, 512, 512, 512，其中使用了 4 層的 Maxpooling layer，而 full-connected 的部份共三層 neuron 數分別為 2048, 512, 64, 7。總共的參數量是 21116615，是個相當可觀的數字，model size = 84526368，在 kaggle 上的 accuracy = 0.69。

這次 HW8 中，我 implement 了 MobileNet 的架構，使用 13 層 convolution layers，kernel 數量分別為：32, 32, 32, 32, 32, 32, 32, 32, 32, 32, 64, 64, 64, 128，其中使用了 4 層 Max-pooling layer，full-connected 的部份共一層，neuron 數分為 512, 7。總共的參數量是 38087，model size = 187799，約是 hw3 的 0.2% 左右，然而在 kaggle 上仍然有 0.63 的準確率。在參數量上有極大的縮減，然而在準確率上卻只有小量的退步。

2. 請使用 MobileNet 的架構，畫出參數量-acc 的散布圖（橫軸為參數量，縱軸為 accuracy，且至少 3 個點，參數量選擇時儘量不要離的太近，結果選擇只要大致收斂，不用 train 到最好沒關係。)(1%)



3. 請使用一般 CNN 的架構，畫出參數量-acc 的散布圖（橫軸為參數量，縱軸為 accuracy，且至少 3 個點，參數量選擇時儘量不要離的太近，結果選擇只要大致收斂，不用 train 到最好沒關係。）(1%)



4. 請你比較題 2 和題 3 的結果，並請針對當參數量相當少的時候，如果兩者參數量相當，兩者的差異，以及你認為為什麼會造成這個原因。(2%)

在我所作的實驗中，MobileNet 在 parameter number = 38087 (架構如第一題所述), 訓練了 80 個 epoch 後，達到在 validation accuracy= 0.627; 而單純的 CNN 在 parameter number = 40791, (架構：8 層 CNN，kernel 數量分別為 16, 16, 16, 16, 32, 32, 32, 其中有 4 層 max-pooling layer，full-connected layer 共 1 層，neuron number 分別為 128, 7)，在訓練了 40 個 epoch 後，達到 validation accuracy= 0.621，其實沒有太大差異，有點令我意外，我原本預測由於 MobileNet 可以在一定的參數限制下疊比較多層，因而能得到較 high level 的資訊而表現更好，但實驗出來卻是這樣（當然也可能是因為不同 dataset 的差異）如果因為這樣被扣分我也認了 XD 助教辛苦囉