Przy moich pomiarach dla różnych liczebności i rozmiarów warstw ukrytych, dowiedziałem się że najlepsze wyniki występują dla dwóch warstw ukrytych o liczebności (15, 8) oraz (16, 7). W najbardziej sprzyjających warunkach metryki wskazywały na:  
Accuracy: 0.7433333333333334  
Precision: 0.725  
F-Score: 0.6470588235294118  
Przy trzech i więcej warstw ukrytych, wyniki nie były aż takie wysokie. Dlatego dalsze pomiary badające najlepszą liczbę epok, optymalnego współczynnika uczenia oraz wpływu normalizacji danych na wyniki nie były robione dla wszystkich badanych wcześniej liczb i wymiarowości warstw ukrytych.

Przy badaniu wpływu normalizacji danych na wyniki metryk, zaobserwowałem że normalizacja wnosi minimalny pozytywny wpływ na wyniki metryk. Zazwyczaj wyniki metryk były poprawiane o kilka punktów procentowych (od 0 do 3) na każdą metrykę.

Przy badaniu liczby epok, przyjąłem zbiór:  
[10, 100, 200, 500, 1000, 2000, 5000, 10000]  
Im większa liczba epok zadana do uczenia modelu, tym lepsze były wyniki, jednakże przy liczbie epok wynoszącej 10000 model zbytnio przyuczał się do zadanych danych i nie był już w stanie poprawnie przewidywać wyników dla zbiory testowego. Niestety mój komputer uniemożliwiał mi szkolenia dla liczby epok większej niż 10 000 epok, a i to zajmowało dużo czasu, dlatego większość moich pomiarów była przeprowadzana na 1000 lub 5000 epok dla możliwości większej ilości pomiarów.

Przy pomiarach współczynnika uczenia, przyjąłem zbiór możliwych współczynników:  
[0.01, 0.1, 1, 2, 5, 10, 100]  
najlepsze wyniki metryk były osiągane dla współczynników 2 oraz 10, na drugim miejscu przy współczynniku 5.