F1 score evaluation, Best F1 score at threshold = 0.38, F1 score = 0.786, precision = 76.28%, recall = 81.07%

