Use Adam as optimizer and ReLU as activation function. Adjust DNN model structure by making it deeper and wider, using batch normalization to increase accuracy(so batch size shouldn't be too small) and using dropout to prevent overfitting. After getting a good prediction, use that model to predict several times with different random seeds. Taking the predictions to do voting ensemble. Then, the final prediction is generated by majority decision.