# The data set I choose is the Credit Approval Data Set（CRX）which concerns credit card application. I train it for a classification, and chose the attribute a16 as the prediction field.

#

#

# 1. How did you encode the input variables?

#

# 1>a2,s3,a8,a11,a14,a15 are encoded using the zscore function because these attributes are numeric values and could be made more comprehensible.

# 2>a1,a4,a5,a6,a7,a9,a10,a12,a13 are encoded using the encode\_text\_dummy function, because these attributes are text field and are non-targeted attributes.

# 3>a2,a14 have some missing values, so they are encoded using missing\_median function for filling the missing spaces with the related median value.

# 4>a16 is the chosen the target field, so it is encoded using encode\_text\_index function for encoding targeted text field.

#

#

# 2. How did you choose how to evaluate the accuracy of your model?

#

# Log Loss is chosen here for evaluating the accuracy of the model. Neural network can predict a probability for each of the target classes, which may give high probabilities to predictions that are more close relative, which needs Log loss to penalize confidence in wrong answers.

#

#

# 3. Why is your model either classification or regression?

#

# My model is for classification.

#

#

# 4. Were there any missing data? How did you handle it?

# :

# Yes. In this set, a2&a14 have missing values.

# missing\_median function is used to convert all missing values in the specified column to the related median.

#

#

# • Mention any other challenges or difficulties you had with the dataset.

# :

# 1>When the csv data set is first seen, it looks intricate since there is no direct explainations on website. What is the meaning of each column is complicated.

# 2>Since the raw data was not preprocessed, it was difficult to make compare to get meaningful information. And the data has to be processed first since the input to neural network must be completely numeric.

# 3>I have to try to increase the accuracy of my model, like trying to increase steps of training, or increase the number of the hidden layer's parameters to train.