I’ve found it hard to change the way features were selected. I’ve decided to change the model.

I’ve tried Multinomial Naive Bayes, which, in my mind should perform much poorer than the Logistic Regression and Multi-Level Perceptron, which should perform better than the Logistic Regression.

This actually happened, however Multi-Level Perceptron is not the best tool in my mind for this task, because of the following behavior:

Smaller architecture (less nodes in hidden layers) were performing poorer than logistic regression. Bigger initial random state were also decreasing the quality of the model. Bigger architecture took much longer to train, however, the results were the same or slightly better than the Logistic Regression provided, so the game was not worth it.

I believe that the key is in the way we select features. However, as I’ve noted above, I was unable to make it in the proper way. I’ve heard same thoughts when I was talking to my colleagues.