I pasted NewsDesk, SectionName, SubsectionName together to produce Section variable. Then I merged some small sections with bigger ones based on my judgement - final Section had 20 levels and it was far most important predictor. Then within each Section, I made clusters of articles based on their vocabulary.

Then I extracted some features from Headlines and Snippet, like number of words, number of exclamation marks etc.

I also added feature (last20) expressing how many articles were published recently, because in EDA I found out that fewer articles published was associated with higher popularity (it does make sense). However it doesn't perform well on CV-set (nor test set :)) so I didn't include it in the model (I should have).

This alone would have finished 15th:

modelBayesGLM2 <- bayesglm(data=train, family="binomial", Popular ~ section\*I(log(WordCount + 1)) + section:clusterid + section:weekend + section:poly(ts,5) + (head.question>0) + section:(snip.pipe>0) + section:I(log(last20 + 1)))

Then I added to my model most frequent words, just because it performed well on test set (but not CV :)) Finally I ensembled it with some RF (I forgot the exact formula).

After the competition I played with the models little bit (I tested them on test set :)) and I found a solution that would have finished 2nd.