* .
* .
* Finally here my results for Level 1.
  + I only include the models of word2vec and doc2vec without additional knowledge, because the results are worse than without additional information.
  + What we can see is that count and tfid in combination with LR have pretty much good results and outperform the dense vectorization technques. But BERT outpferoms all
  + Depends on classifier
* .
* . In the training data set exmployer could also give alternatives. The horizontal line clearly indicates that the ids have a huge range of alternatives, which And it important to keep this mind and also to think about ways to maybe predict more than one class.

So much to my presentation. If you do not further questions I would like ask two questions for the discussion.