

"мы должны были выбирать признаки в данном случае по тому на сколько они коррелируют с целевой переменной" - можно так посмотреть, а для классификации даже лучше будет взаимную информацию посчитать

"корреляция почти у всех очень низкая" - бывает и такое, значит выбираем наиболее коррелирующие в любом случае

"я думал, что ассурасу более важная статистика, чем roc" - нет "более важной", есть разные ситуации и разные метрики для этого

"если смотреть в комплексе, то результат то ужасный" - почему же, нет, разные ситуации бывают, плюс мы не знаем баланс классов - если он 50/50, то ассурасу = 60 хотя бы лучше константного предсказания, что уже хоть что-то :)