**Lessons Learned from Previous Kaggle Contests**

1. Title of the interview.
2. URL: <http://blog.kaggle.com/2015/08/21/avito-winners-interview-2nd-place-changsheng-gu-aka-gzs_iceberg/>
3. Title: *Avito Winner’s Interview: 2nd place, Changsheng Gu (aka Gzs\_iceberg)*
4. What model(s) did they use (supervised or not, specific model or models, etc.)?
5. For preprocessing, hash tricks and negative down sampling.
6. For supervised learning methods, FFM, FM and XGBoost.
7. Tools used: [XGBoost](https://github.com/dmlc/xgboost), [pypy](http://pypy.org/) and [lasagne](https://github.com/Lasagne/Lasagne)
8. What influenced their modeling choice and what insights they had from the model?
9. The interview does not contain such information.
10. The winner was trying to find a good approach do model ensembling
11. The advice from the 2nd place winner is “…I encourage people reinventing the tools you used if you have enough time, it helps you understand why and how it works…”
12. Other observations.

The task was addressed in stages:

1. Design data pipeline and validation set
2. Feature extraction
3. Parameters tuning
4. Modeling ensembling

**Exercise on Principle Component Analysis**

File used: women\_tnf\_1984.csv