

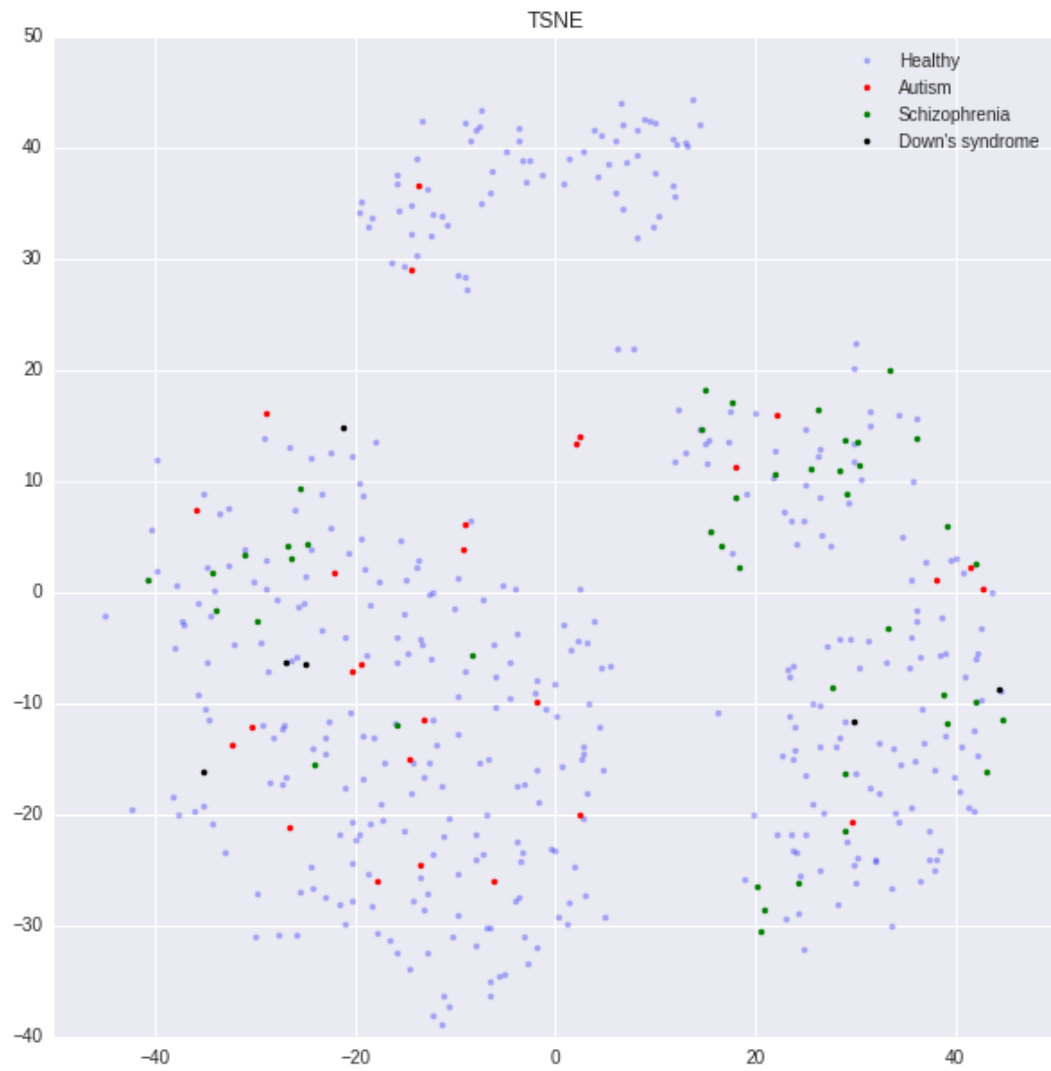
Initial analysis

Sergey Korolev, HSE

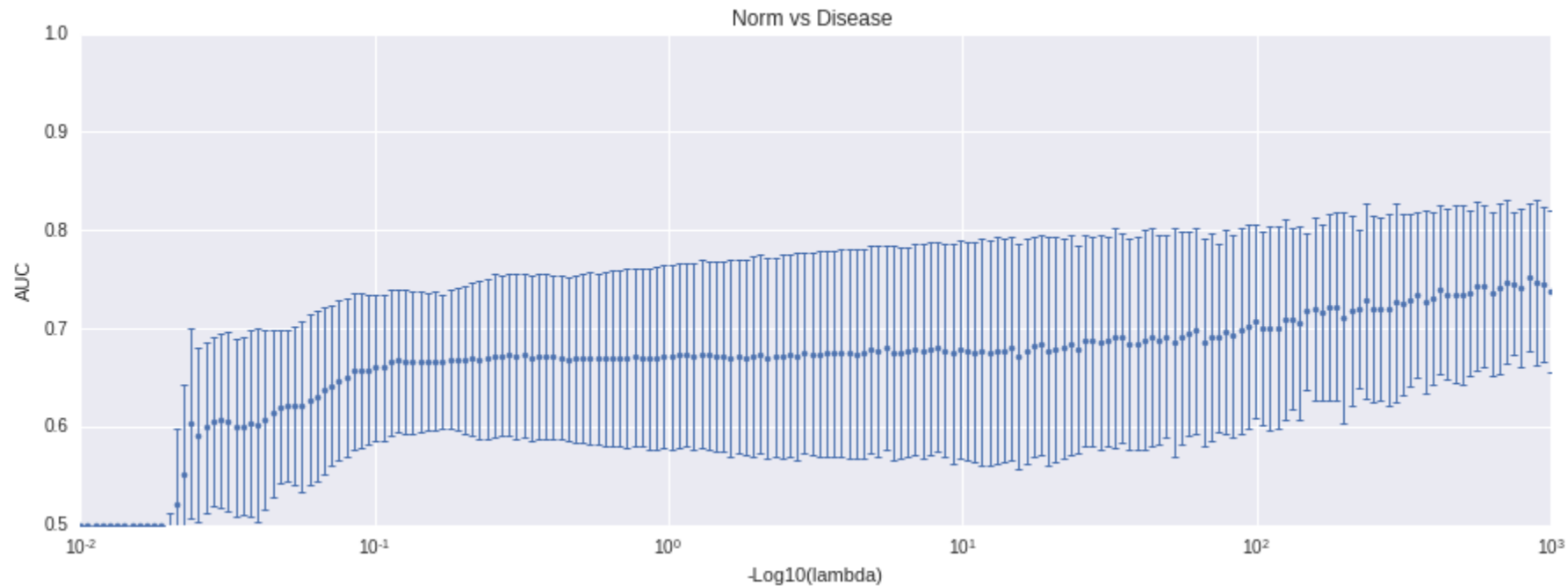
PCA



t-SNE



L1 regularization



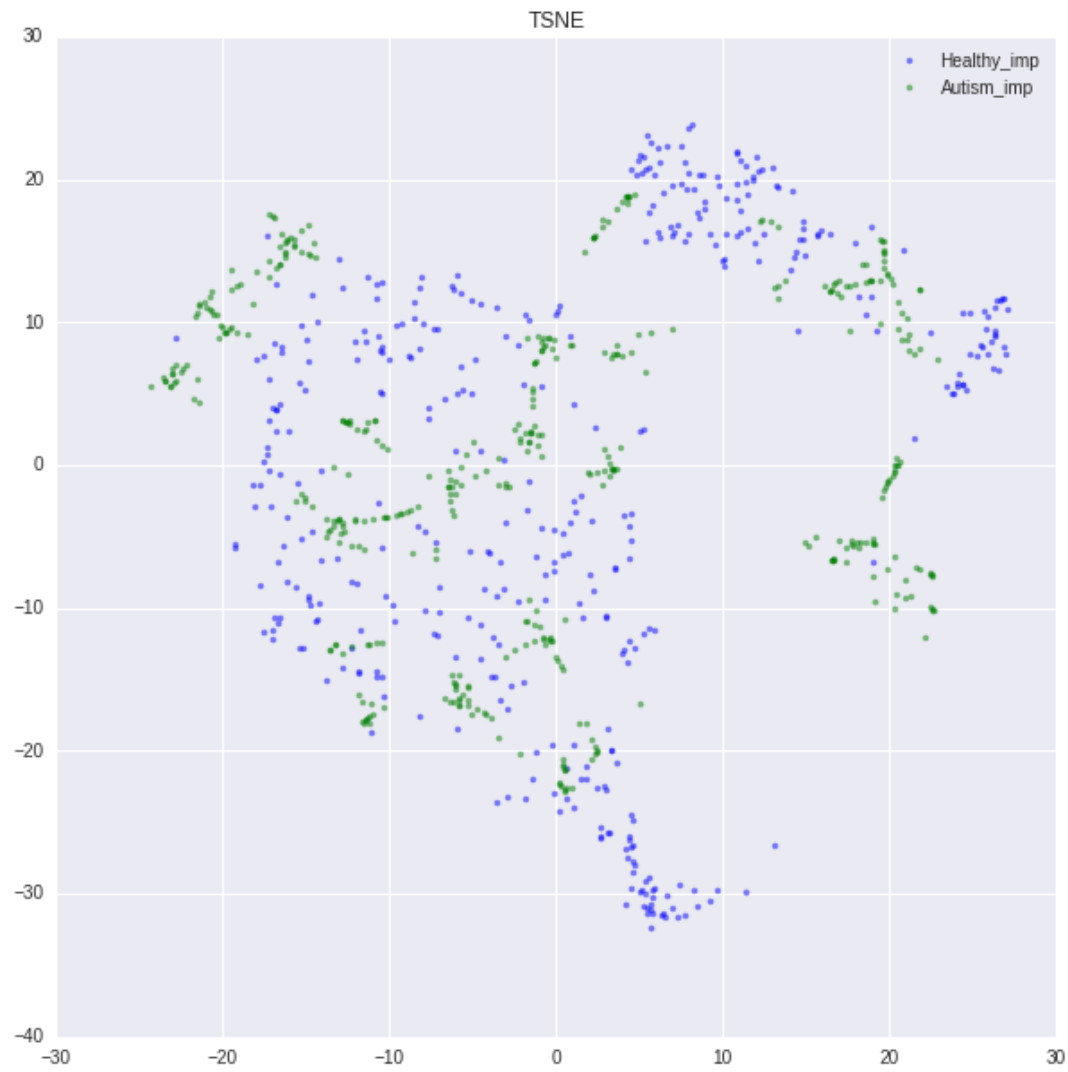
Unbalanced dataset

Weighting

Bootstrap undersampling

Oversampling

SMOTE



Results

Mean \pm std	LogReg with L1	LogReg L1 with SMOTE	LogReg L1 with Balanced weights	LogReg with L1 after age residuals norm
Norm vs Disease	0.75202 \pm 0.07529	0.76842 \pm 0.06774	0.71989 \pm 0.09888	0.78090 \pm 0.07673
Norm vs Autism	0.77851 \pm 0.02485	0.77422 \pm 0.02005	0.67446 \pm 0.07825	0.77474 \pm 0.04460
Norm vs Schizophrenia	0.88057 \pm 0.05056	0.87960 \pm 0.05258	0.81330 \pm 0.06050	0.83598 \pm 0.07026
Autism vs Schizophrenia	0.91963 \pm 0.07608	0.92352 \pm 0.05979	0.91722 \pm 0.04730	0.91963 \pm 0.07608