MATH6380P Final Project
Nexperia Image Classification

CAO Yang

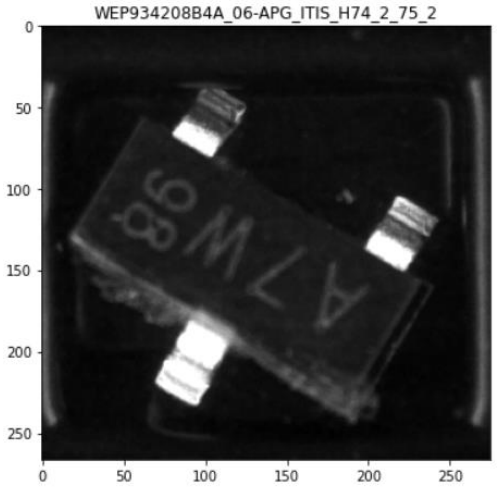
WU Jiamin

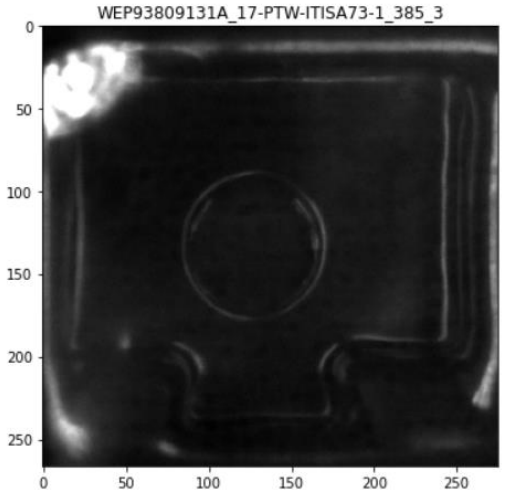
Introduction

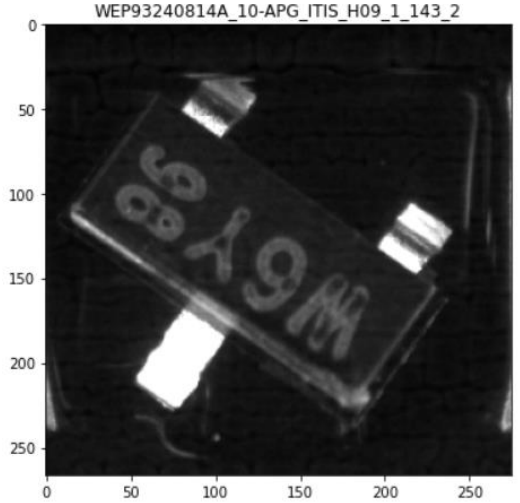
- Dataset
- Feature selection
  - Scattering net
- Image classification
  - LDA
  - Random forest
  - SVM
  - Logistic regression
- Visualization of the selected features
  - PCA
  - MDS
  - T-SNE
- Results and analysis

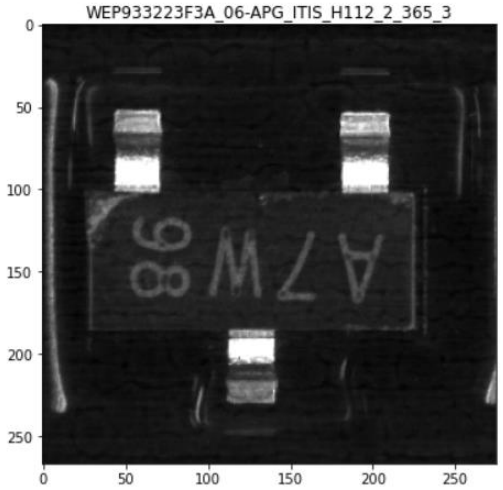
Dataset

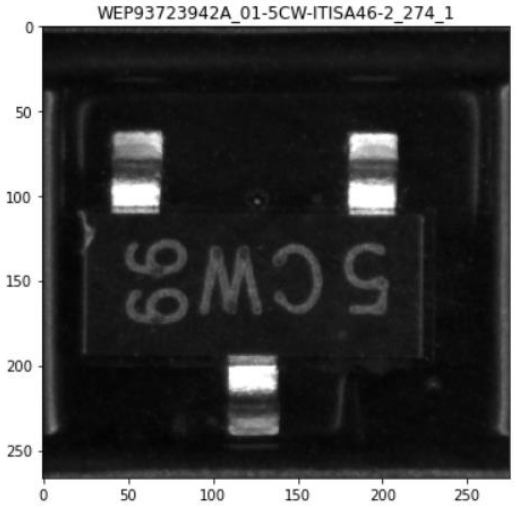
Nexperia image dataset

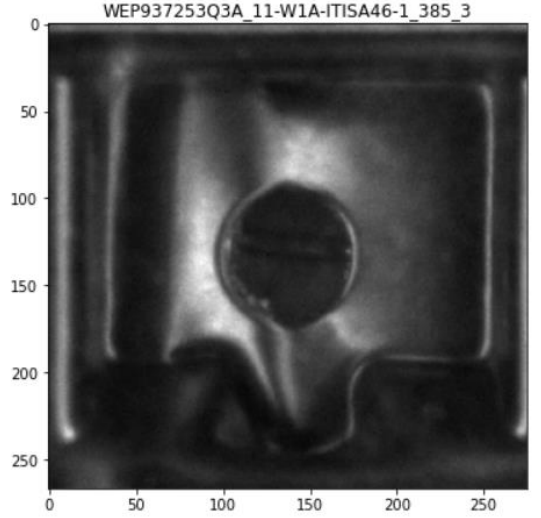






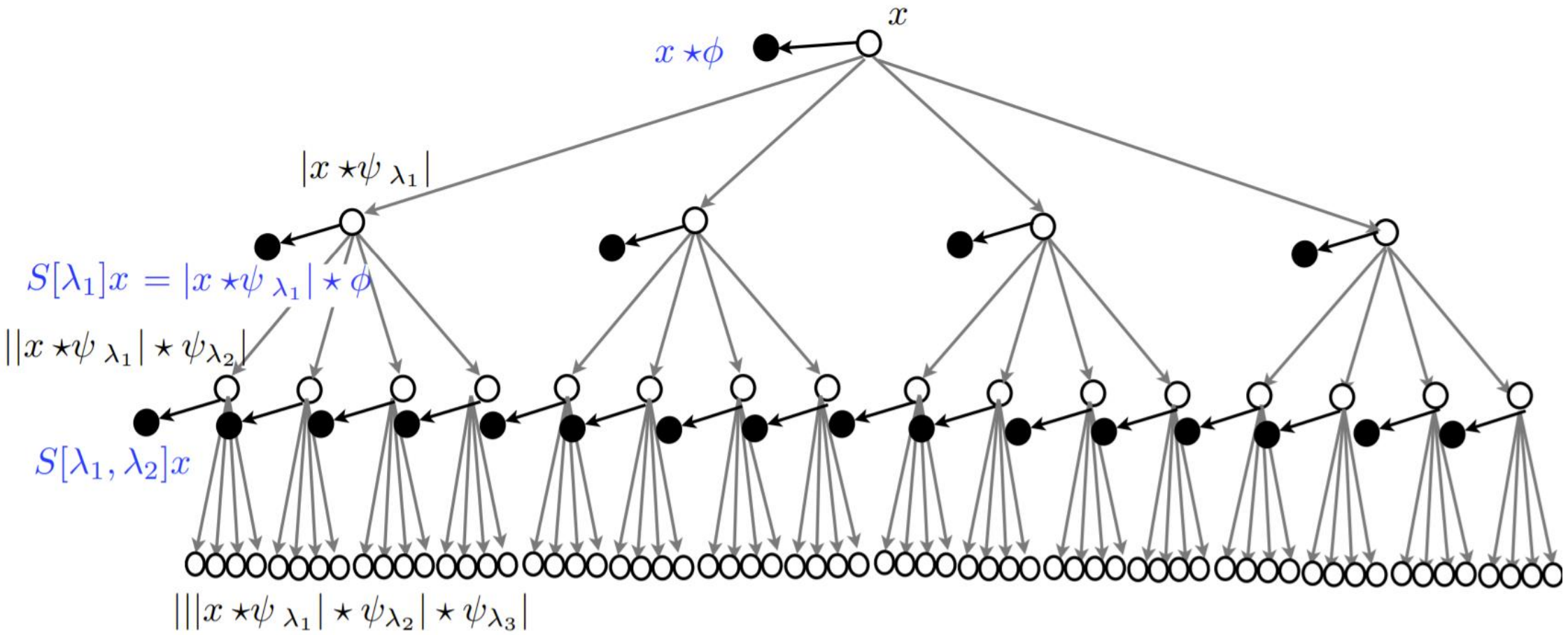






Feature selection

Scattering Net



Y. LeCun et. al.

Scattering Net

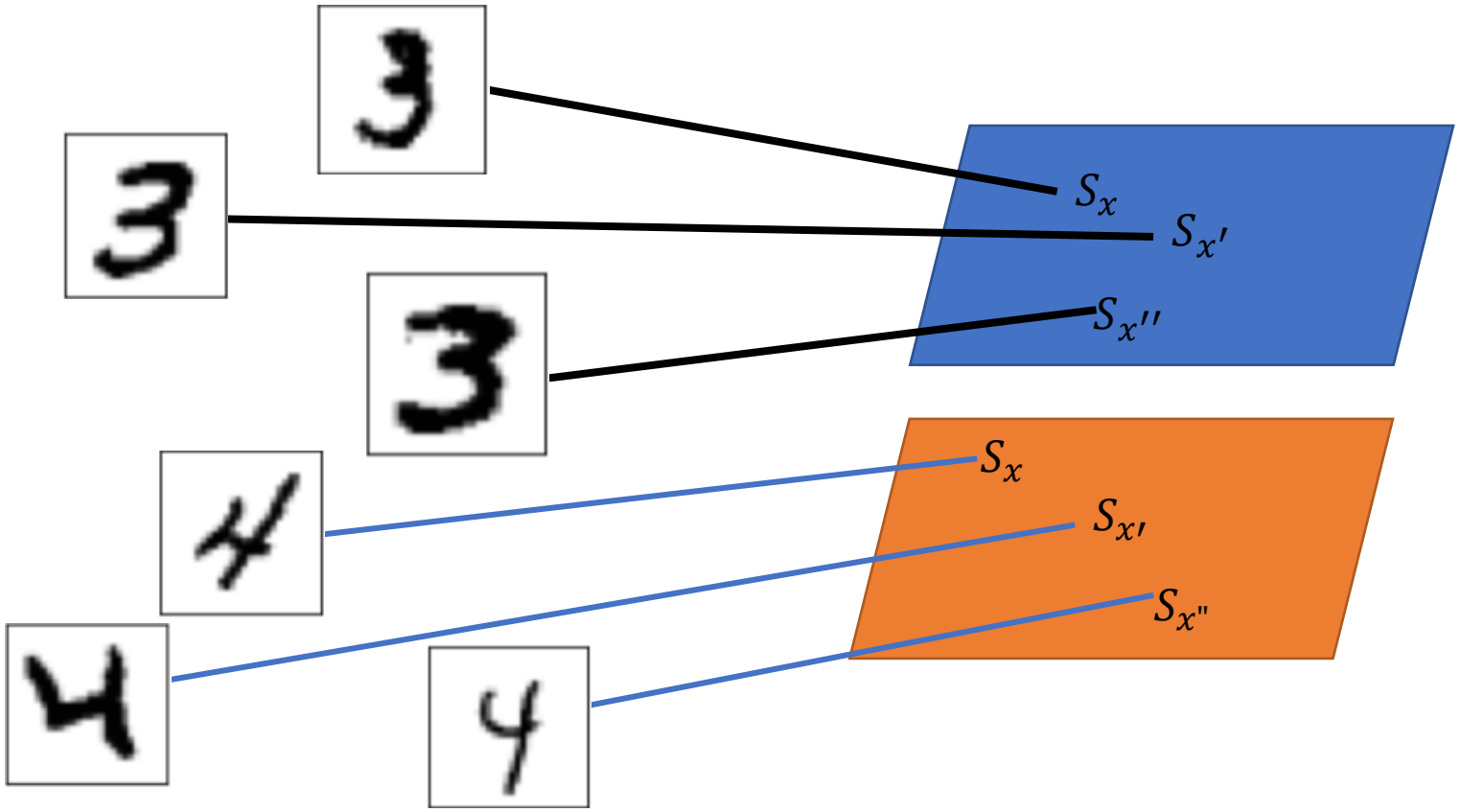
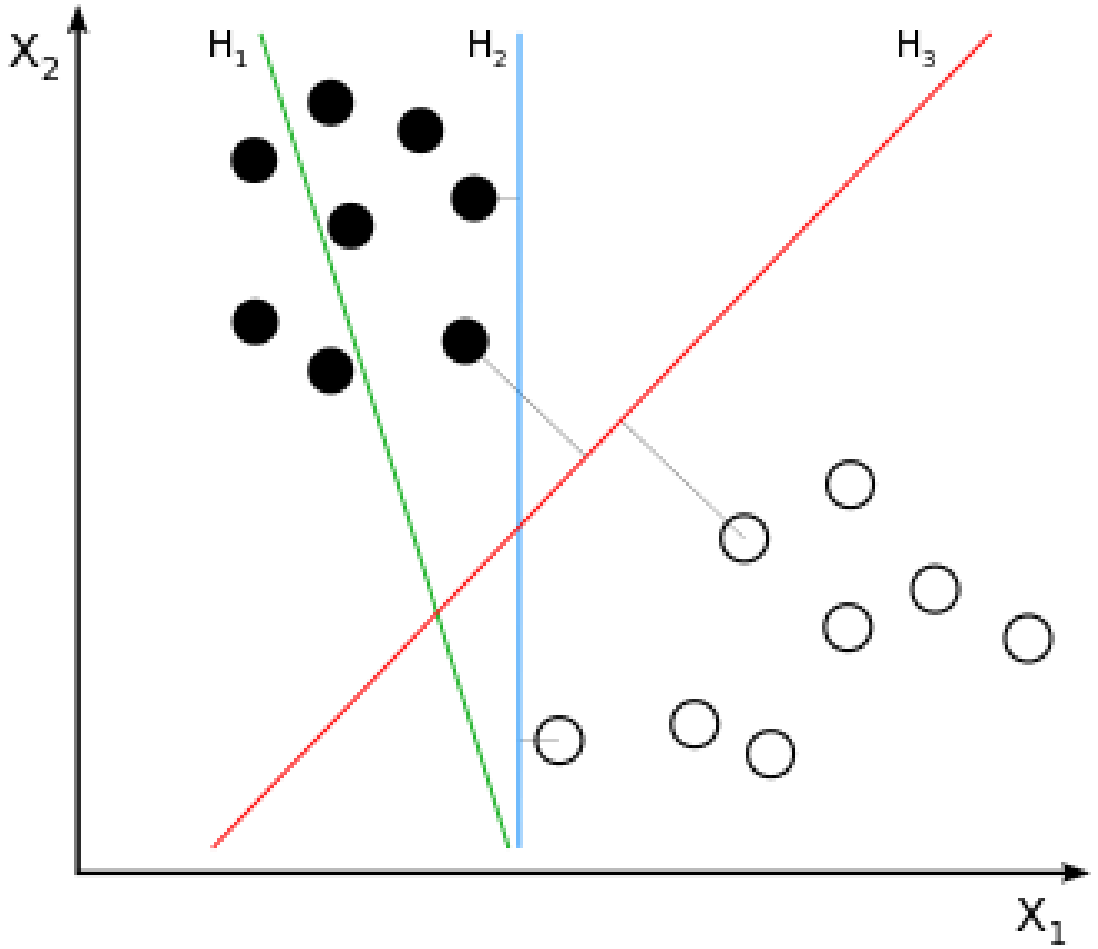


Image classification

Linear discriminant analysis

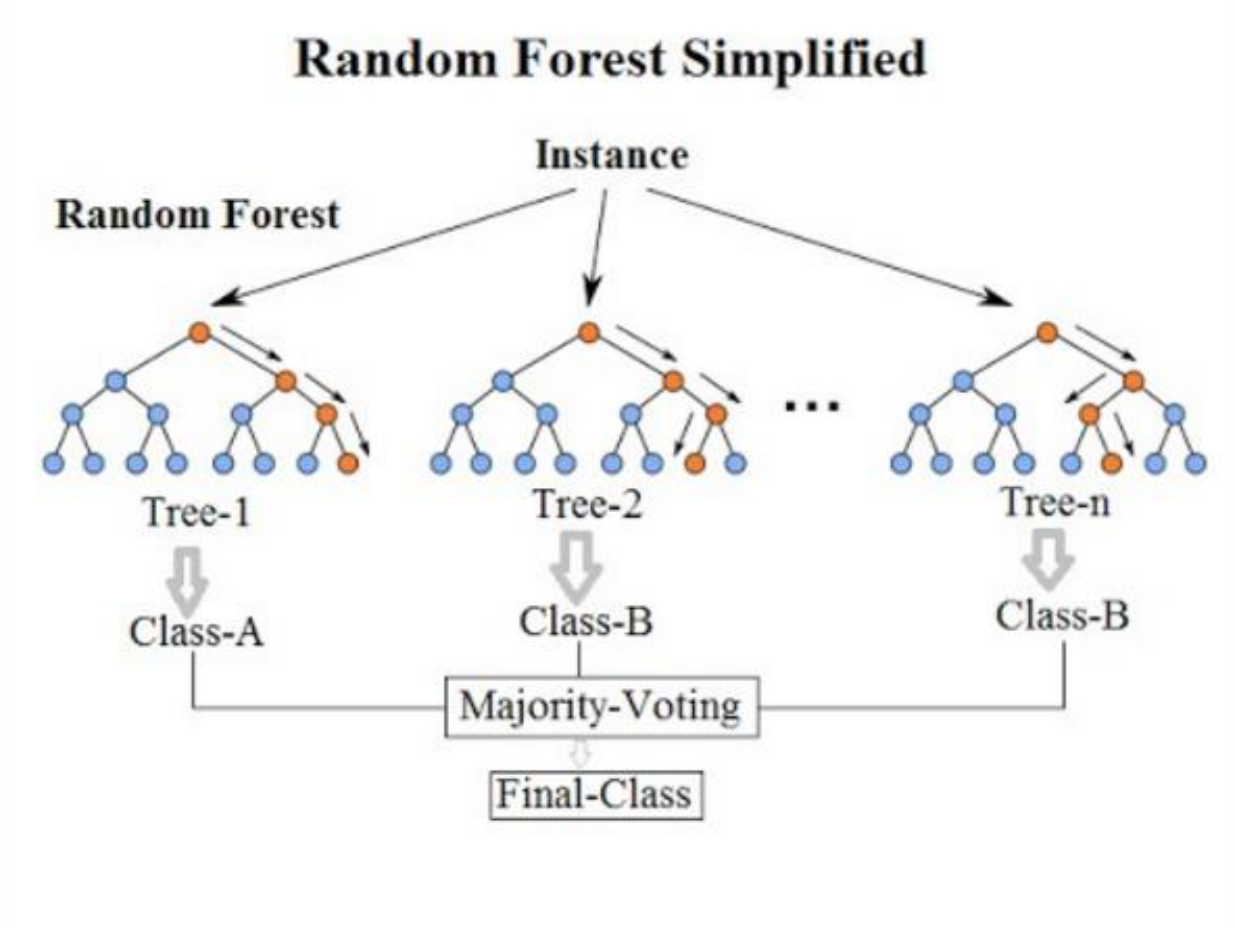
- Assume sample in each class follows normal distribution. Specially, LDA would require the covariance matrix in each class to be the same.
- Prediction power can decrease due to multicollinearity: high correlation with predictor variables
- Can be very sensitive to outliers.

Support vector machine



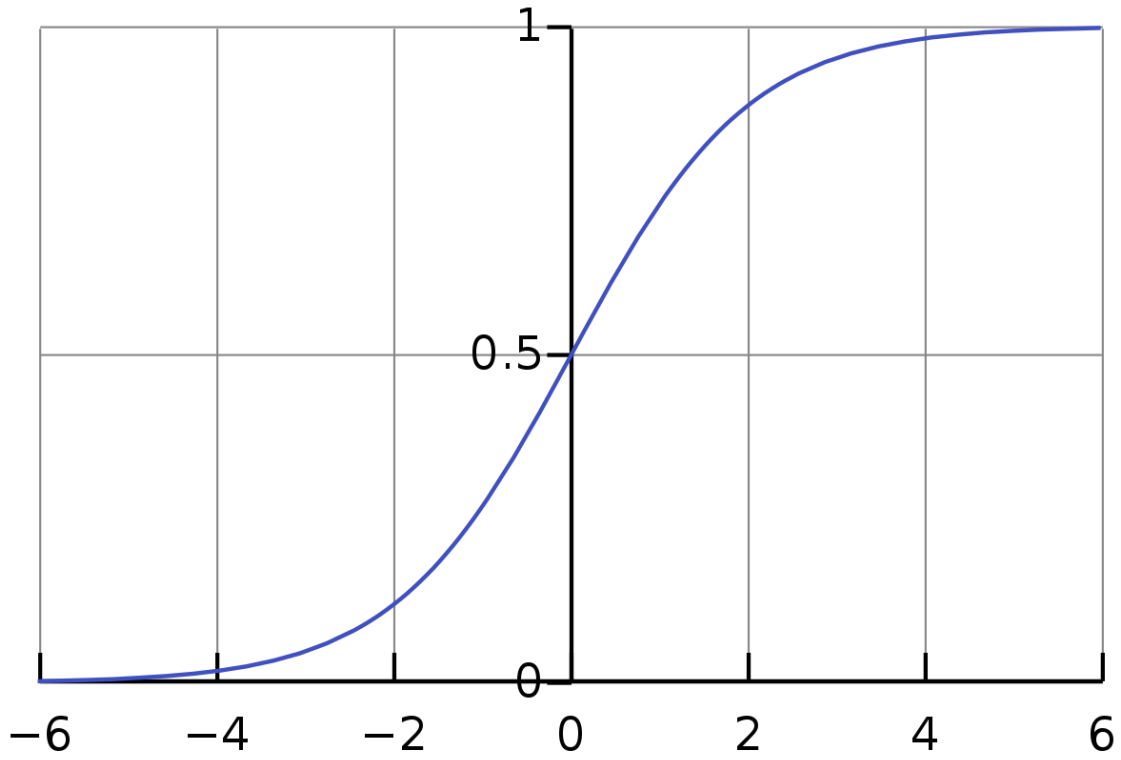
ZackWeinberg

Random forest



Venkata Jagannath

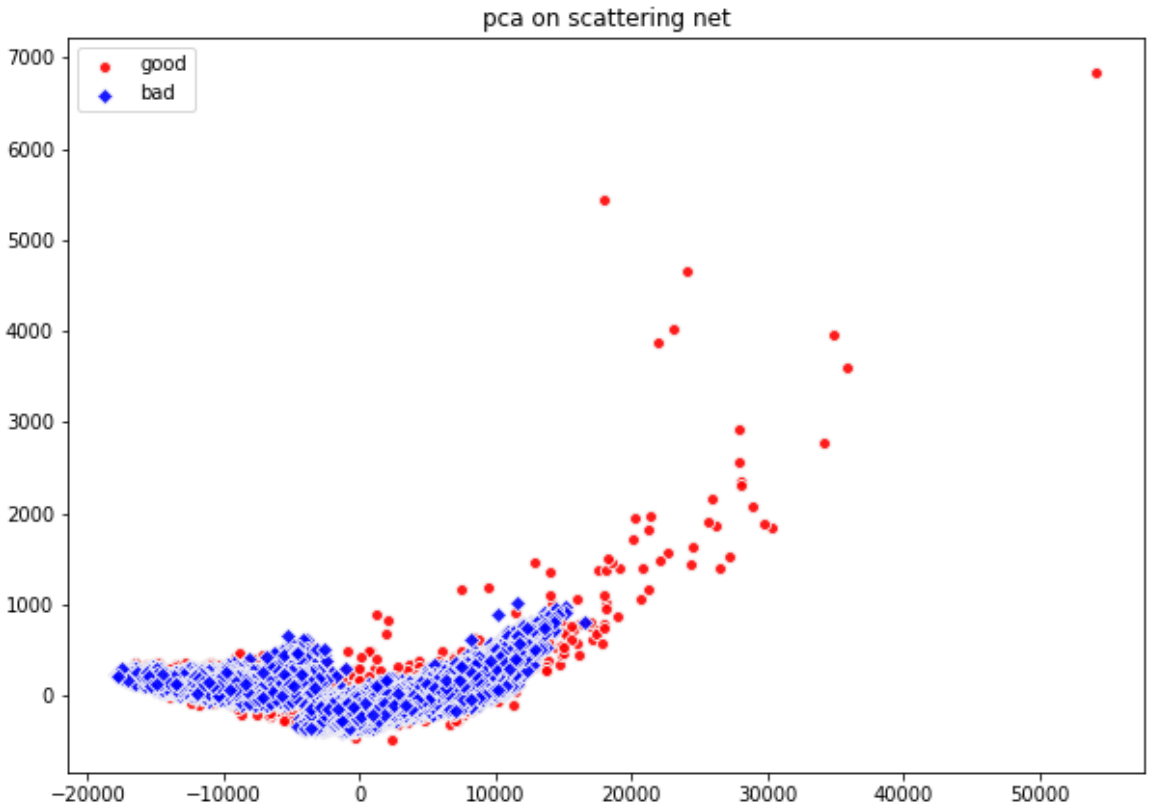
Logistic regression

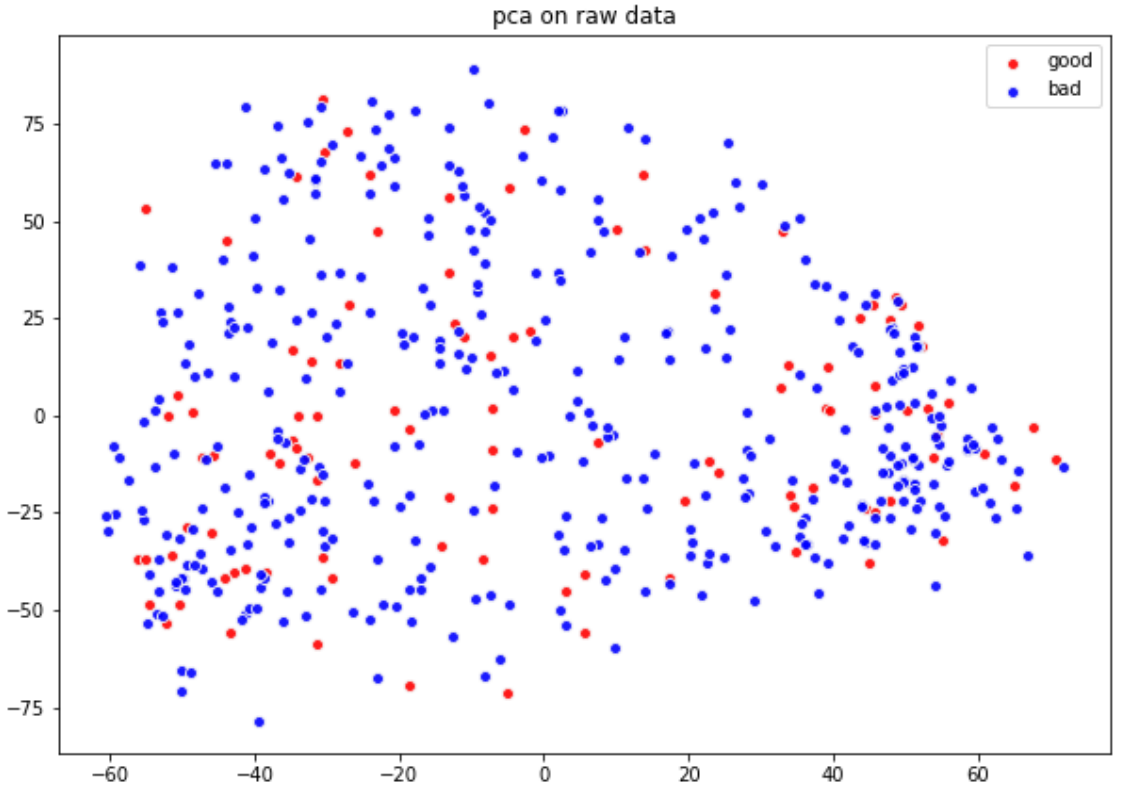


By Qef

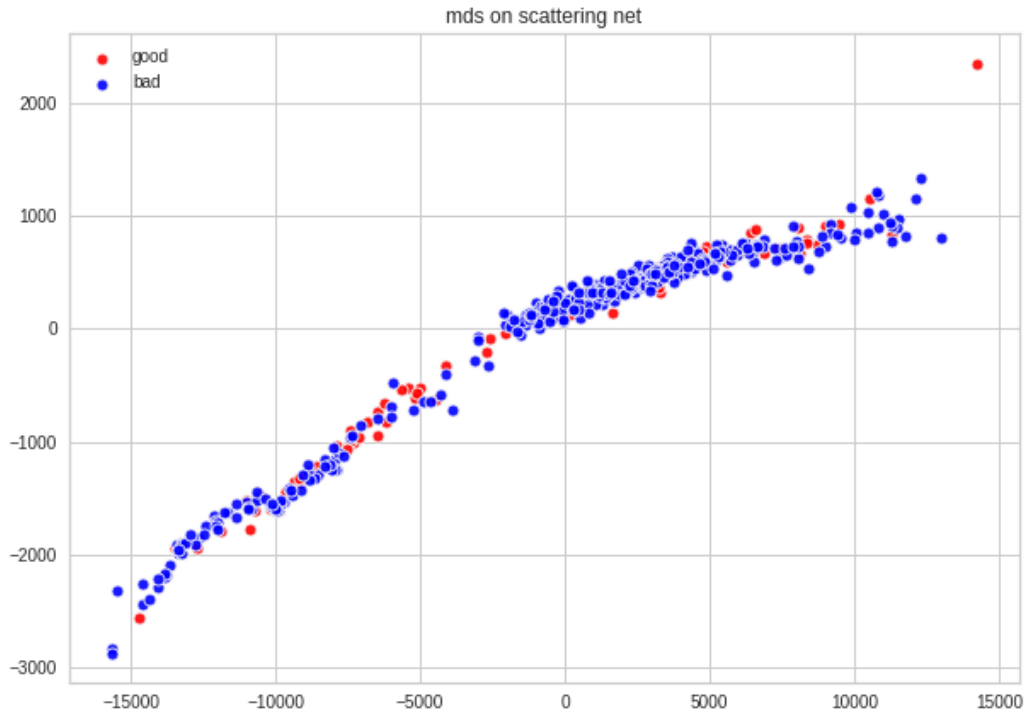
Visualization

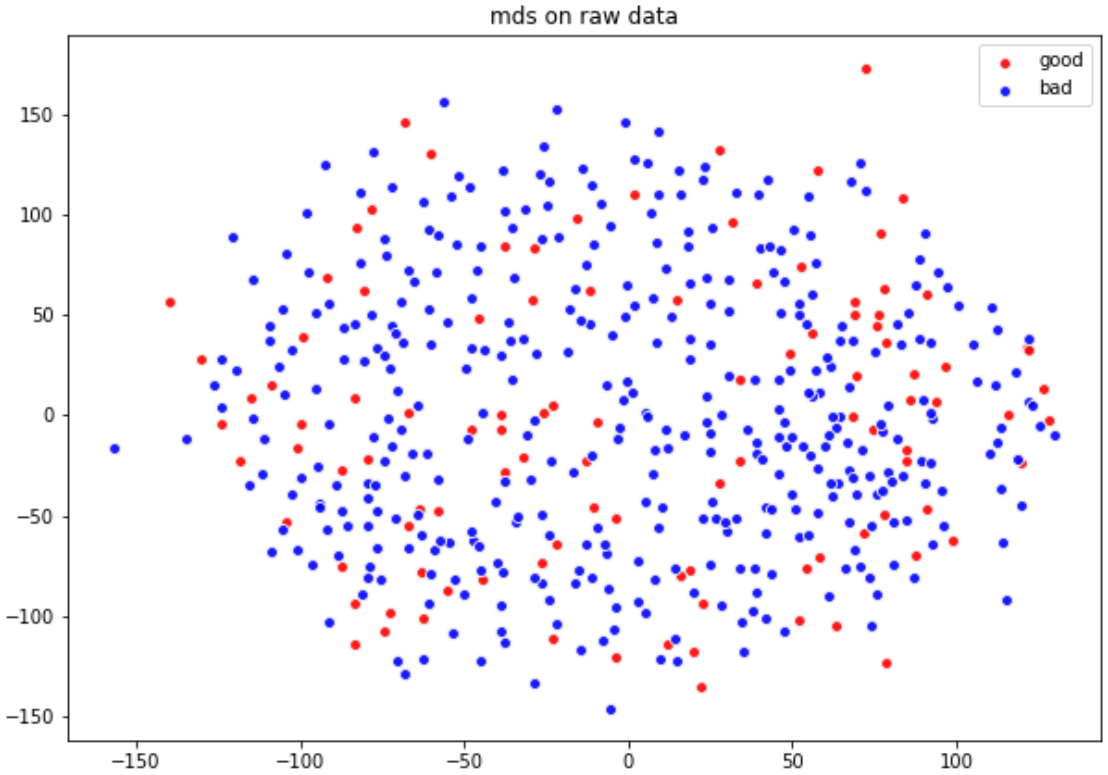
Principal component analysis



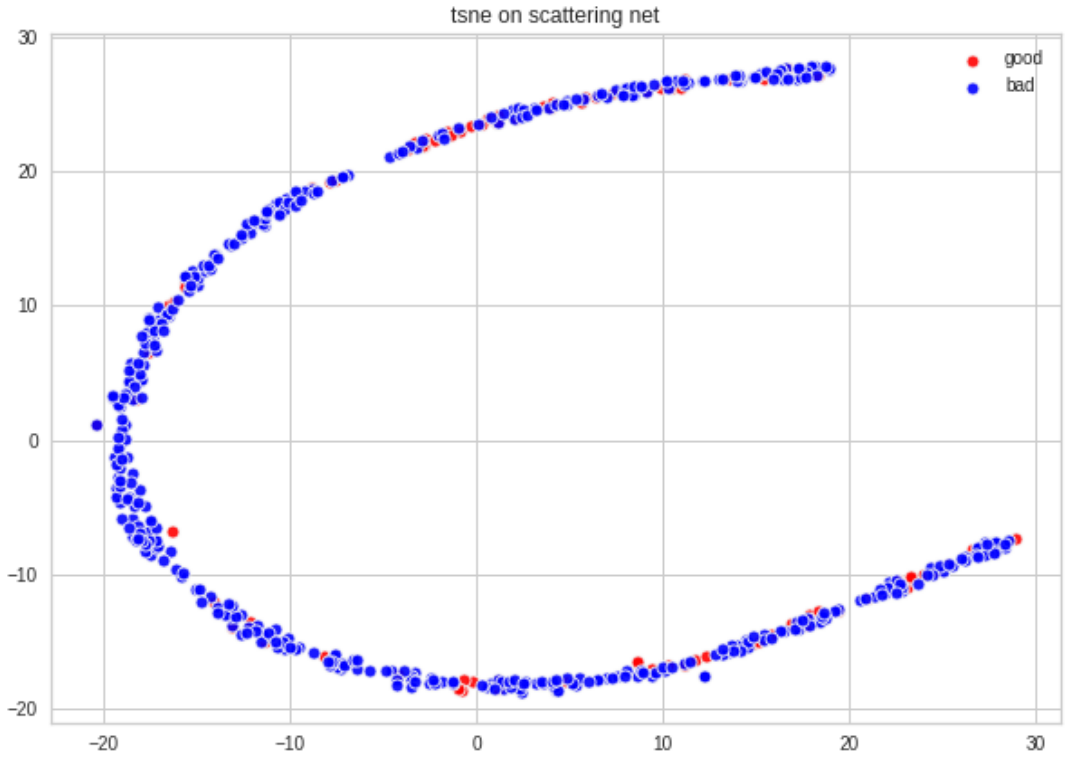


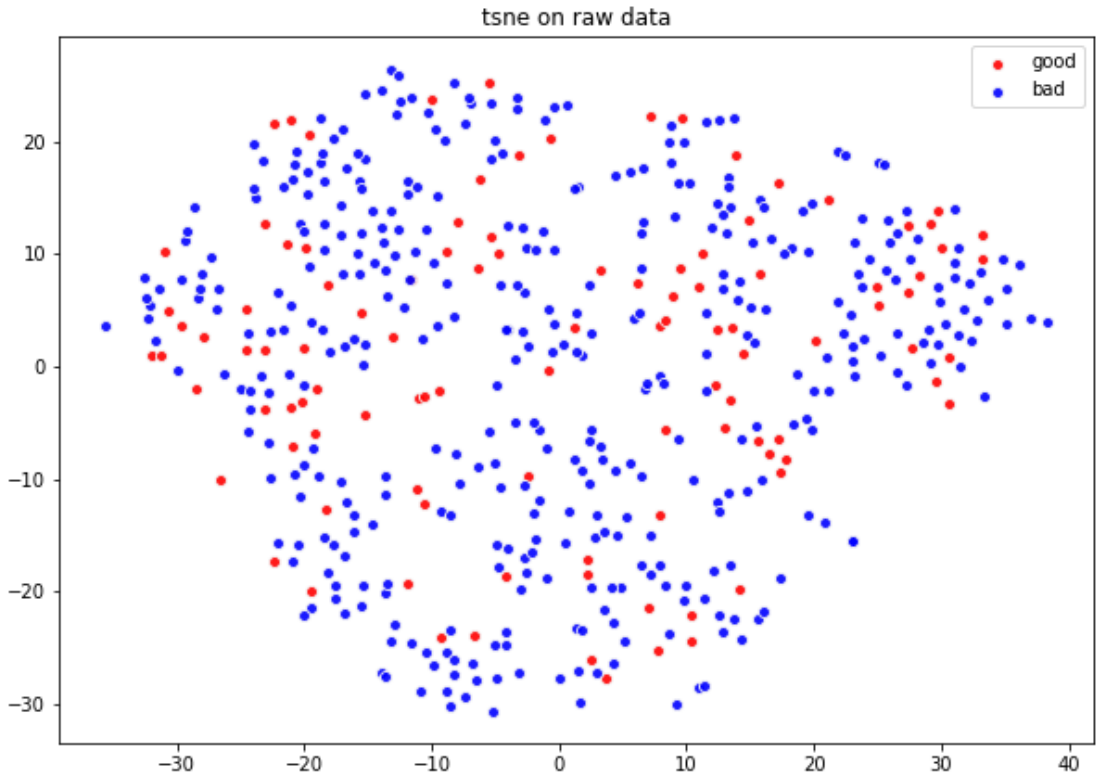
Multidimensional scaling





t-SNE


tsne on scattering net
good
bad
30
20
10
0
−10
−20
−20
−10
0
10
20
30


tsne on raw data
good
bad
20
10
0
−10
−20
−30
−30
−20
−10
0
10
20
30
40

Results and Analysis

Training error

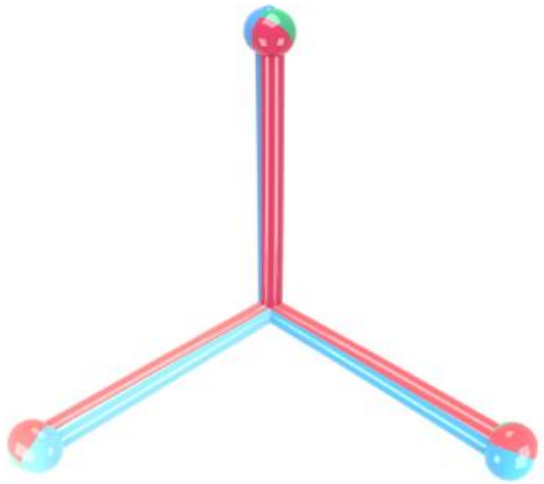
	rf	lr	svm	lda
Raw	0	0.0014800197335964604	0.1351751356684756	0.05000000000000044
scattering	5.803998955278278e-05	0.20987260222293158	0.2042717432310862	0.18961664586900373

Testing error

	rf	lr	svm	lda
Raw	0.5908093994778067 8	0.4981723237597912	0.6895561357702349	0.6503916449086162
scattering	0.1903133159268929 6	0.3919060052219321	0.6	0.4407310704960835

Statistics

	contraction of within class variation (NC1)	Equal-norms of class-means	equal-angularity with	closeness to maximal-angle equiangularity
Scattering net	131.630125519334446	0.32714084	0.9999992996454241	1.0000006705522537



Thank you