



Analysis by classification of RADAR signals returned from the ionosphere.



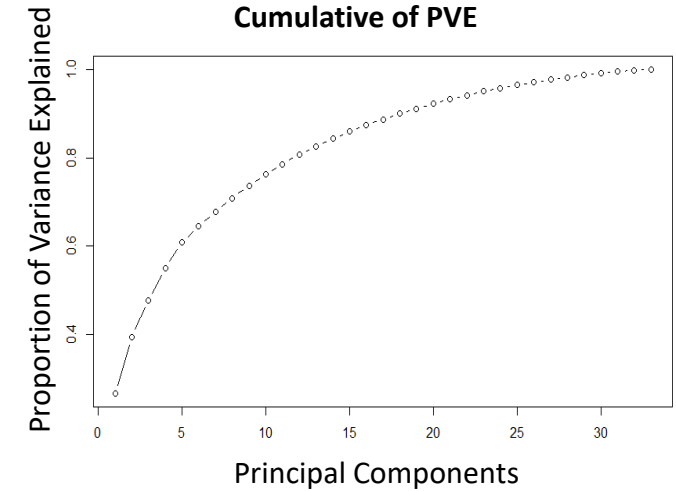
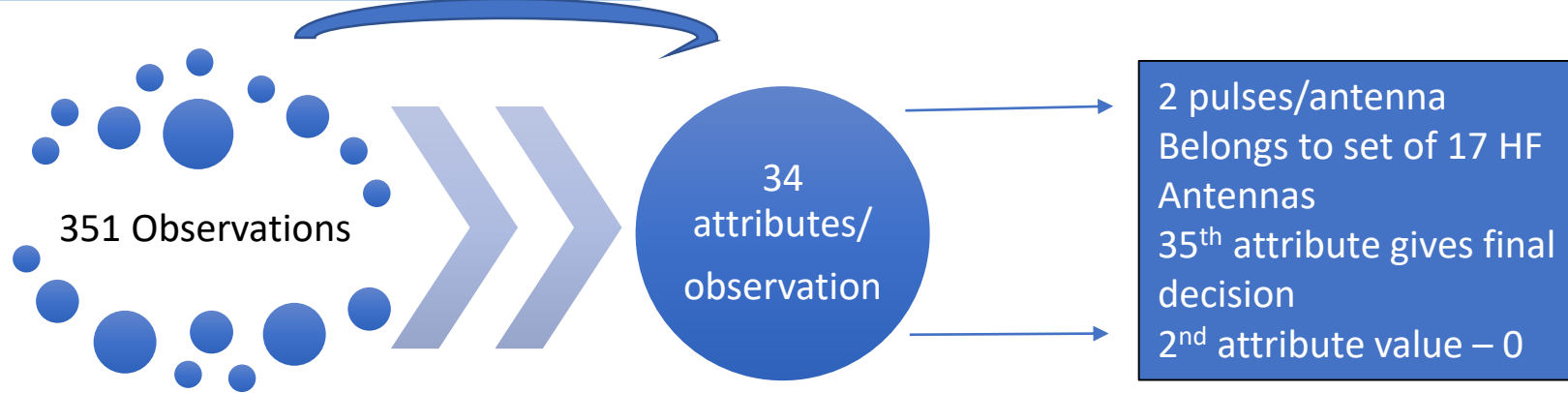
- **Instructor: Dr. Xiaoning Qian**
ECEN 765 – FALL 2017

Submitted by
- **SAMBANDH BHUSAN DHAL**
UIN - 726006080



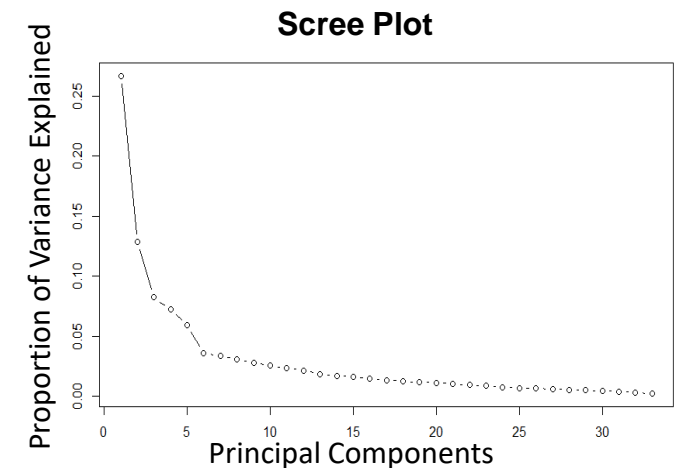
DATASET ANALYSIS:

- Primarily the pulses received by antennas in RADAR system at Goosebay Labrador
<https://archive.ics.uci.edu/ml/datasets/ionosphere>



PCA APPLICATION ON DATASET:

- PCA reduces the dimensionality of the dataset
- In our case, we get 34 Eigen vectors
- We consider 5 principal components as it explains 70% of the variance in our data





Application of different classifiers on the dataset:

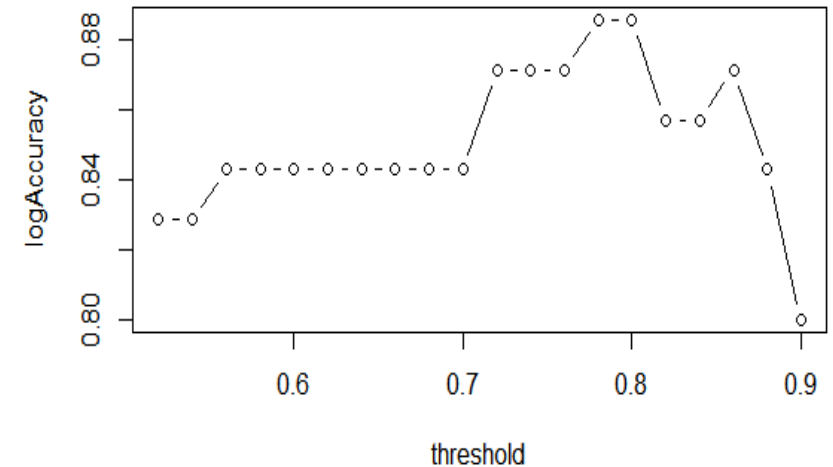
CLASSIFIERS	ACCURACY
KNN(N=3)	0.8714
Logistic Regression (threshold=0.8)	0.8858
LDA	0.80
QDA	0.8858
Bagging (decision trees = 100)	0.9143
Random Forest	0.9429
Support Vector Classifier	0.80
SVM (Polynomial kernel)	0.8000
SVM (Radial kernel)	0.9572

PROCEDURAL CRITERIA

First 20% as testing dataset

rest 80% as training dataset

Log Accuracy vs threshold graph



★ Note: To ensure that there is no overfitting in our model, we go for cross validation of our data



Application of the classification algorithms after cross-validation:

- We divide the data into 5 non-overlapping test data sets and in each case, we use the remaining data as the training set
- We eventually average to find the least error rate

Runs	SVM poly	SVM Radial	SVC	Random Forest	Bagging	QDA	LDA	Logistic	KNN
1	0.2000	0.0428	0.2000	0.0571	0.0857	0.1142	0.2000	0.1142	0.1286
2	0.2285	0.1571	0.2428	0.1714	0.2000	0.1142	0.2285	0.2142	0.2000
3	0.1885	0.0857	0.1428	0.0857	0.1571	0.1428	0.1428	0.1285	0.1143
4	0.1285	0.0571	0.1142	0.0571	0.0857	0.0428	0.0857	0.1428	0.1000
5	0.0422	0.0281	0.0422	0.0140	0.0281	0.05633	0.0563	0.0845	0.0423
MEAN	0.1570	0.0742	0.1485	0.0771	0.1113	0.0941	0.1427	0.1369	0.1170

Future Prospects:

Feed forward neural network in multi layer and single layer networks can be used to achieve higher test accuracy

Conclusion:

After 5 fold cross validation on the data set , Radial SVM provided the best test accuracy (92.58%) followed by Random forest with an accuracy of 92.29%.