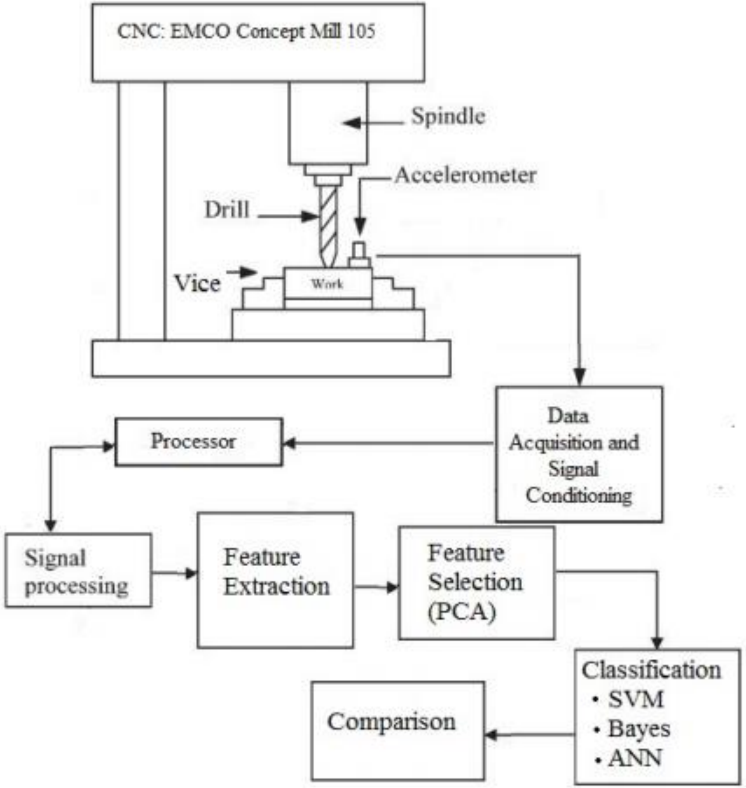
UMassAmherst
The Commonwealth's Flagship Campus

Drill bit fault detection and classification

Nino Figliola & Bhavesh Parkhe

December 10, 2018

Data flow chart



Data flow chart

Preprocessing

- Digital signal filtering
- Feature Extraction
- Feature Selection
- Feature Reduction
- Feature combination

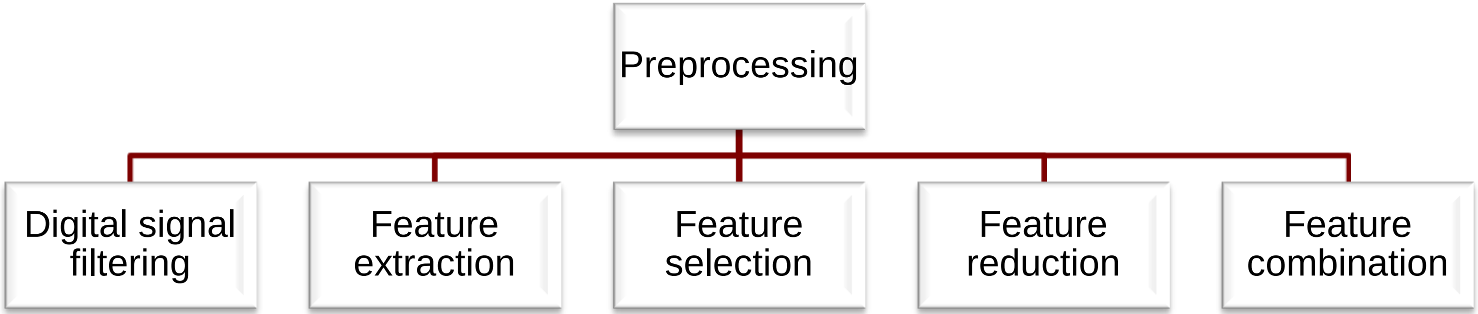
Modelling

- kNN
- ANN
- Decision Tree
- SVM
- Naïve Bayes

Training and Testing

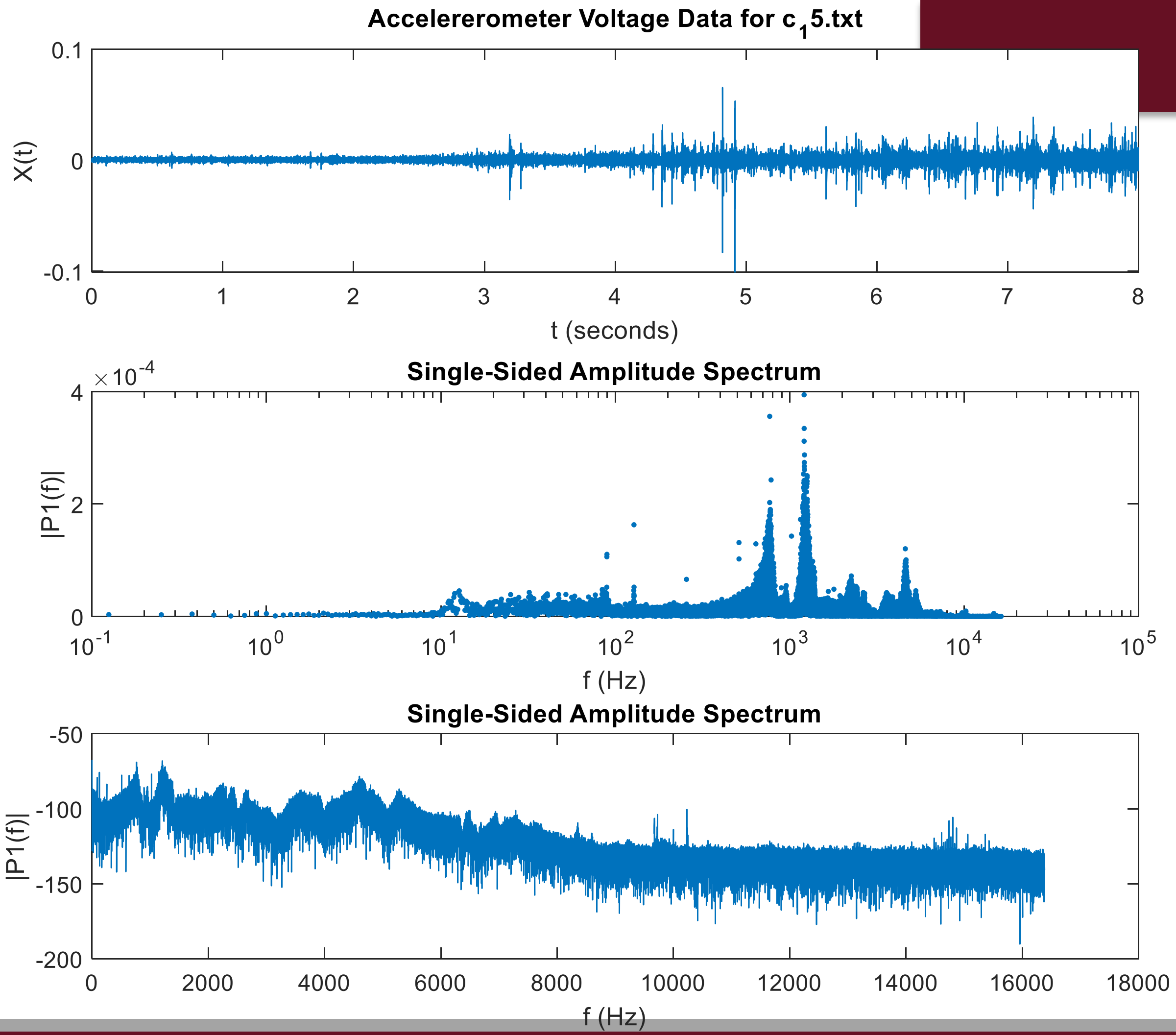
- Fault detection
- Fault Classification
- Individual fault detection

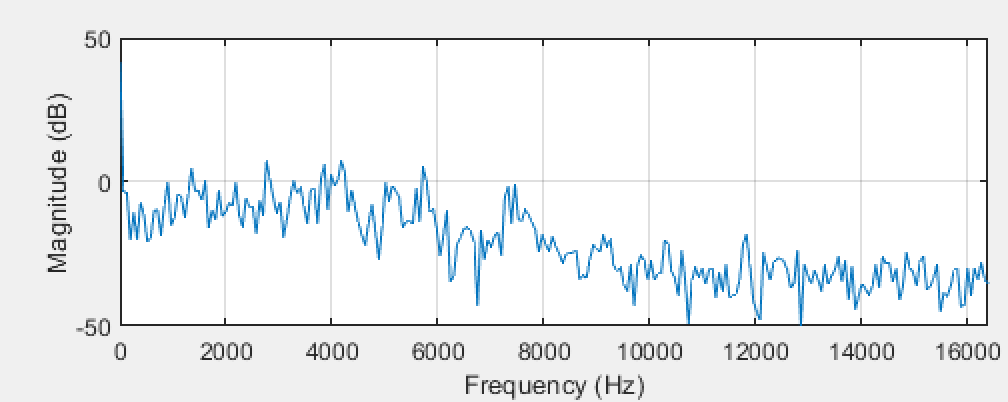
Data preprocessing methods

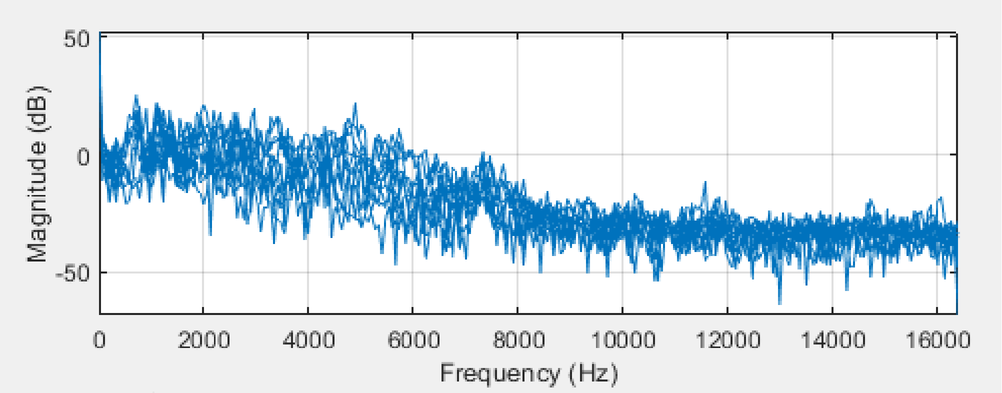


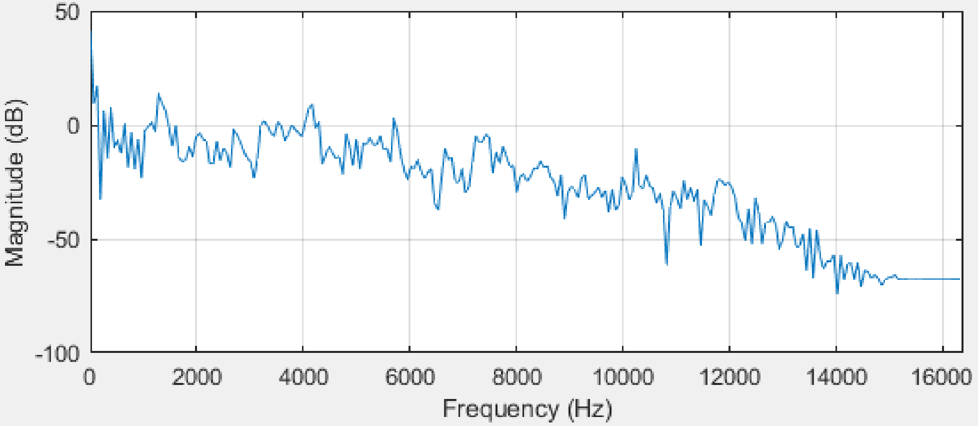
Digital signal filtering

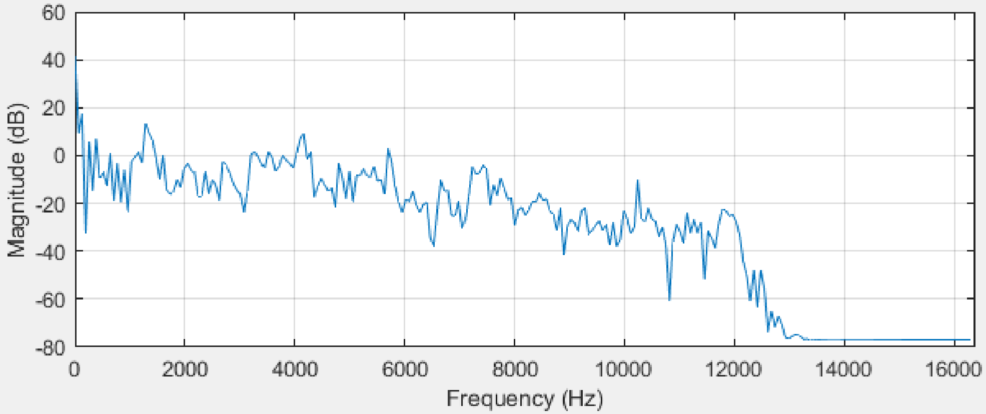
- Remove process noise which doesn’t add value to the data
- Cutoff range of frequencies with low amplitudes
- Lowpass butterworth filter of order 20 used for digital filtering


Acceleromerometer Voltage Data for $c_1$5.txt
X(t)
t (seconds)
Single-Sided Amplitude Spectrum
|P1(f)|
f (Hz)
Single-Sided Amplitude Spectrum
|P1(f)|
f (Hz)


50
0
-50
Magnitude (dB)
0
2000
4000
6000
8000
10000
12000
14000
16000
Frequency (Hz)


50
0
-50
Magnitude (dB)
0
2000
4000
6000
8000
10000
12000
14000
16000
Frequency (Hz)

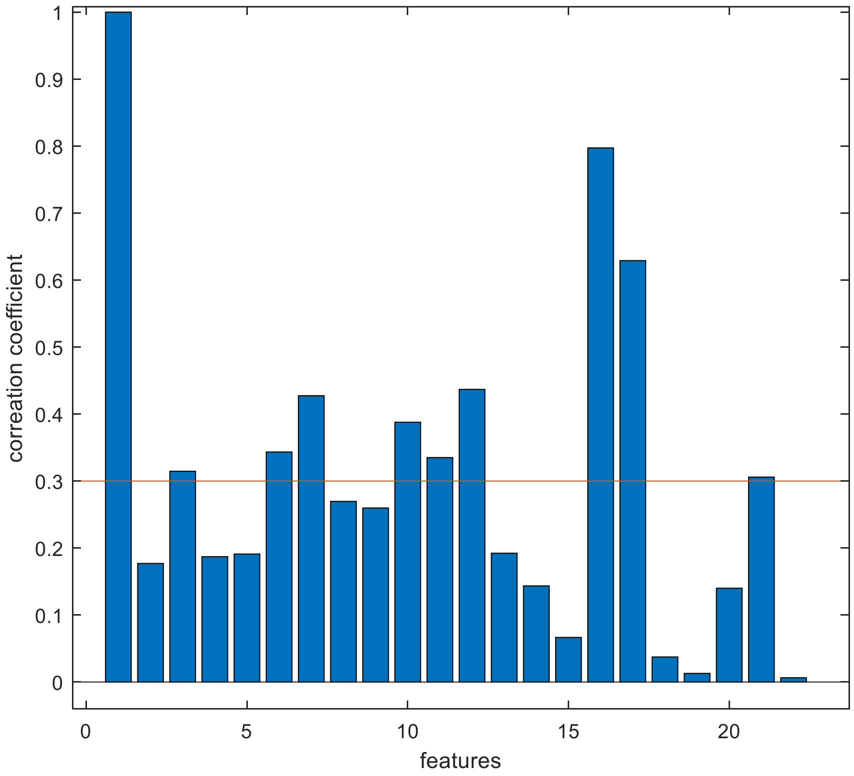

Magnitude (dB)
50
0
-50
-100
0
2000
4000
6000
8000
10000
12000
14000
16000
Frequency (Hz)


Magnitude (dB)
60
40
20
0
-20
-40
-60
-80
0
2000
4000
6000
8000
10000
12000
14000
16000
Frequency (Hz)

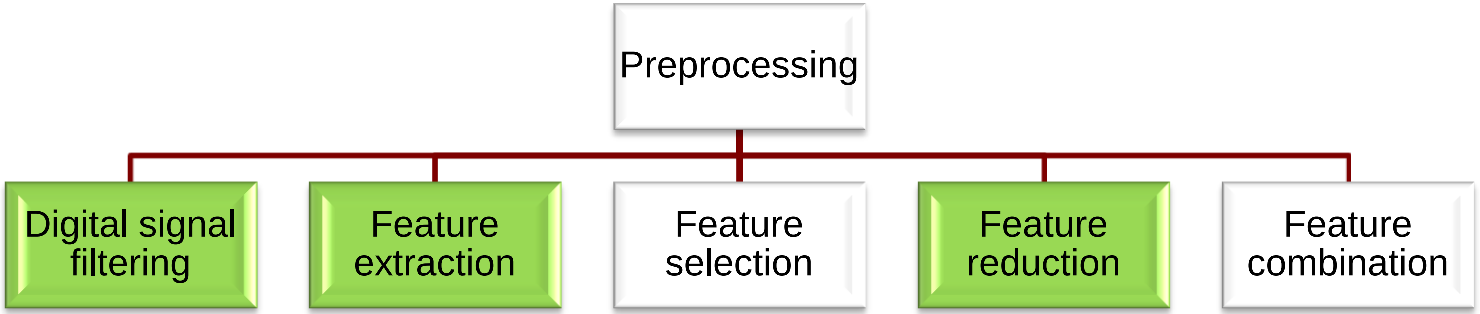
Feature extraction

- Time domain
  - Mean
  - Peak
  - Root mean square
  - Variance
  - Kurtosis
  - Crest factor
  - Shape factor
  - Skewness
- Frequency domain
  - 256 points divided into 8 bins
- Morlet waves
  - Standard deviation
  - Wavelet entropy
  - Kurtosis
  - Skewness
  - Variance

Feature reduction



- Preprocessing
  - Digital signal filtering
  - Feature extraction
  - Feature selection
  - Feature reduction
  - Feature combination

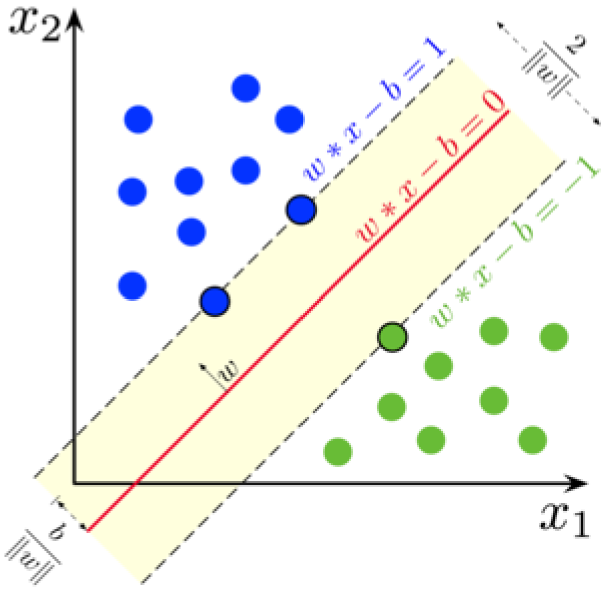


- Steady state: 12kHz
- Penetration: 20kHz

- Time domain (8)
- Frequency domain (8)
- Morlet waves (5)

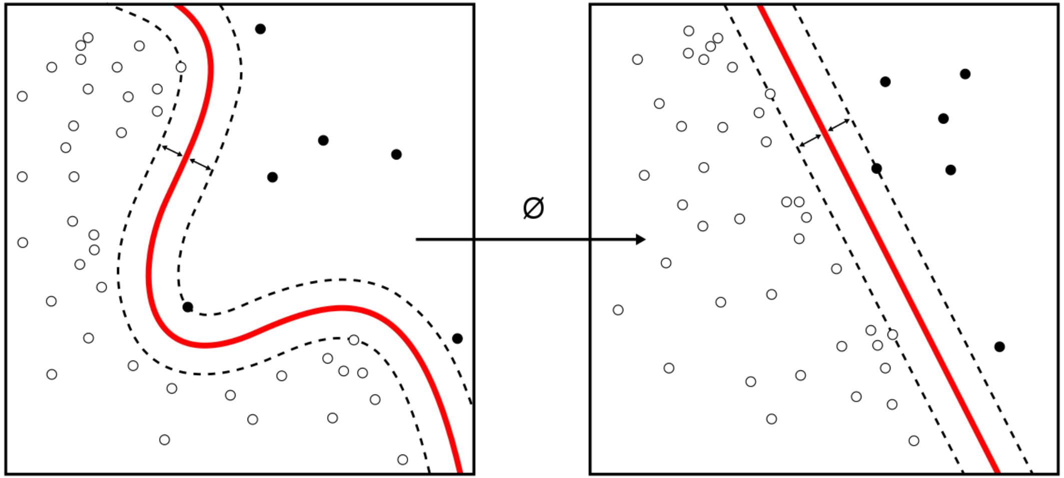
- Fault correlation > 30%

Support vector machine (SVM)



Support vector machine (SVM)	
Kernel Function	Linear
Box Constraint	0.1
Coding Matrix	Onevsone & onevsall
Standardization	False

Naive Bayes classifier



$$p(C_k \mid \mathbf{x}) = \frac{p(C_k)\, p(\mathbf{x} \mid C_k)}{p(\mathbf{x})}$$

$$\text{posterior} = \frac{\text{prior} \times \text{likelihood}}{\text{evidence}}$$

k-nearest neighbor classifier (kNN)

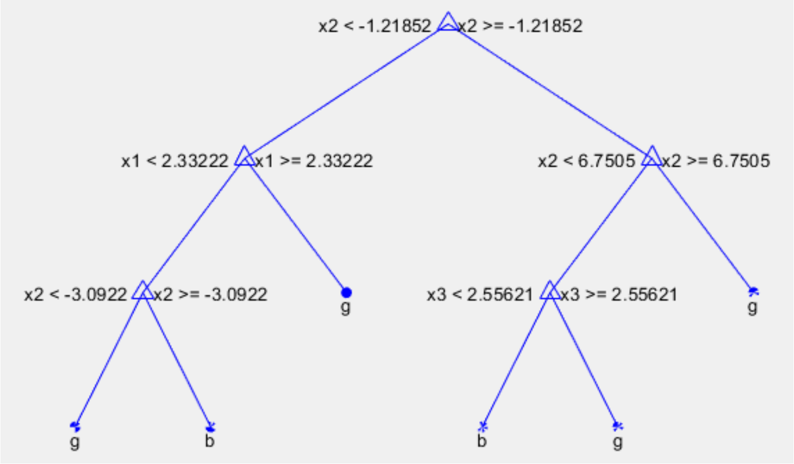
- **kNN**: classification method in which a new object is classified based upon its distance to the nearest training samples of known classification.
  - Strengths: ability to classify nonlinear, multimodal, unlabeled samples based upon similarity to training samples

Design of kNN classifier	
Number of nearest neighbors (k)	3
Distance measurement metric	Euclidean
Equal or weighted voting	Weighted
Distance weight	Squared-inverse

Decision tree classifier

- **Decision tree**: network of binary decision-branches that originate from an initial root node. Branches lead to class labels (“leaves”) at endpoints.
  - Strengths: simple and intuitive, can classify data of multiple types.
- Problem of deep decision trees overfitting data.
  - Can address this by limiting size of the decision tree.
  - Simple, shallow, generalized decision trees can be more robust.

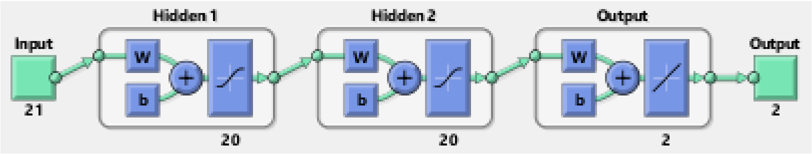
Design of decision tree classifier	
Split predictor technique	CART
Maximum number of splits	5



Neural network classifier (ANN)

- **Artificial neural network**: receives input data and implements weight, biases, and transfer functions to deliver an output.
  - Strengths: recognizing patterns in noisy, complex data and determining their nonlinear relationships.
- ANN candidates: recurrent, feedforward

Design of ANN classifier	
Neural network type	Feedforward
Number of layers	3 (*2 hidden*)
Neurons per hidden layer	20
Transfer function for hidden layers	Tangent-sigmoid
Transfer function for output layer	Pure-linear
Training method	Levenberg-Marquardt
Performance metric	Mean squared error



Classifier training and testing

- 5-fold cross validation for establishing training and testing sets.
- Accuracy determined by comparing the output of the classifier to the known label
- Results from each fold were combined to determine overall accuracy for the classifier

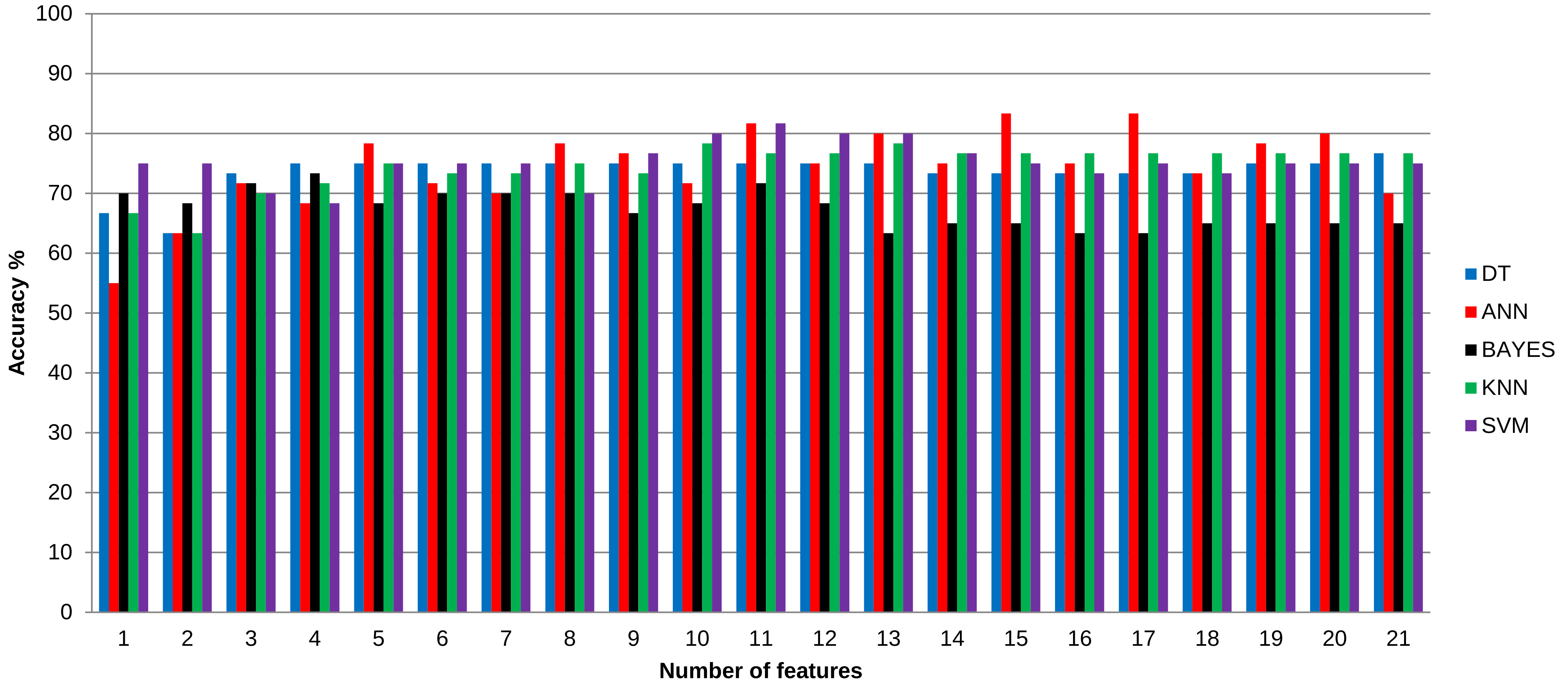
Classifier training and testing

Modes for training

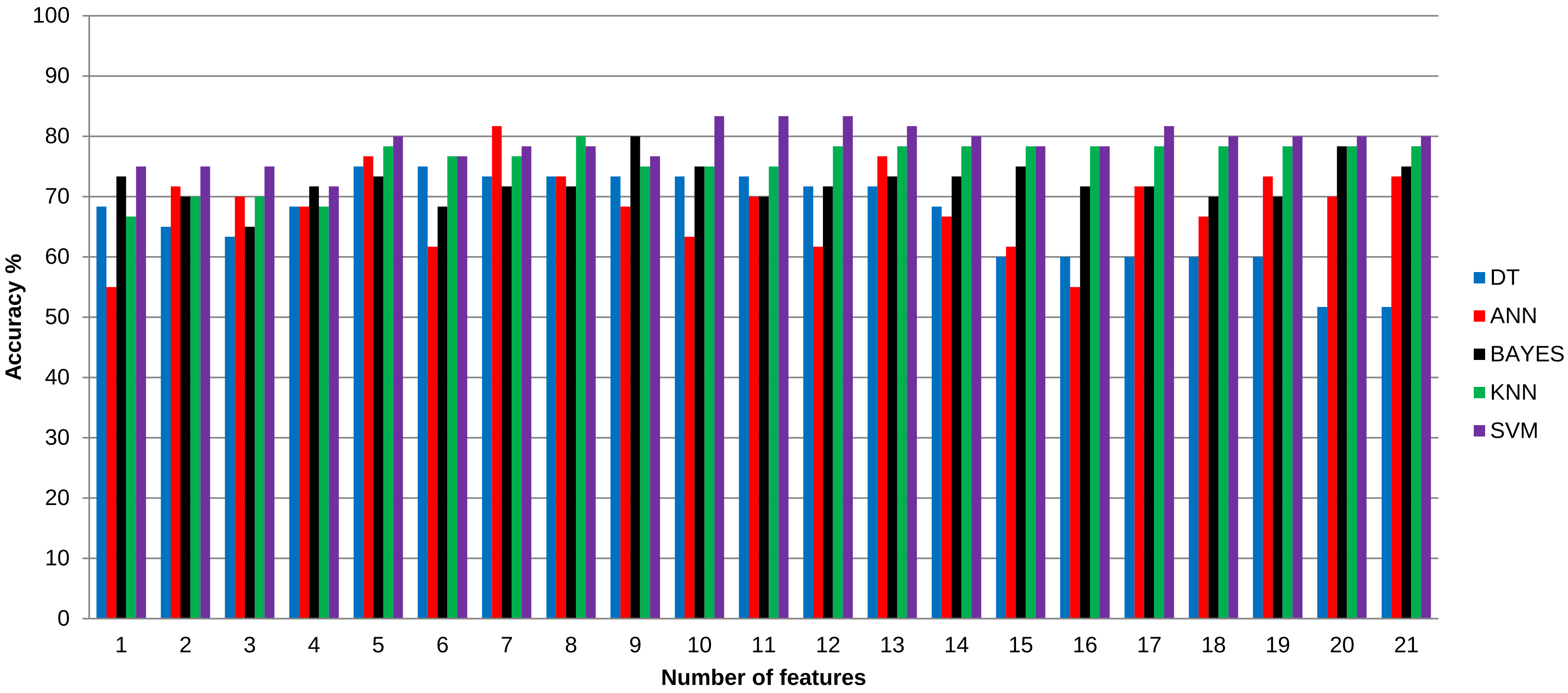
- Fault detection (binary)
- Fault classification (multiclass)
- Individual fault detection (binary)

Fault detection results

Fault detection (penetration)



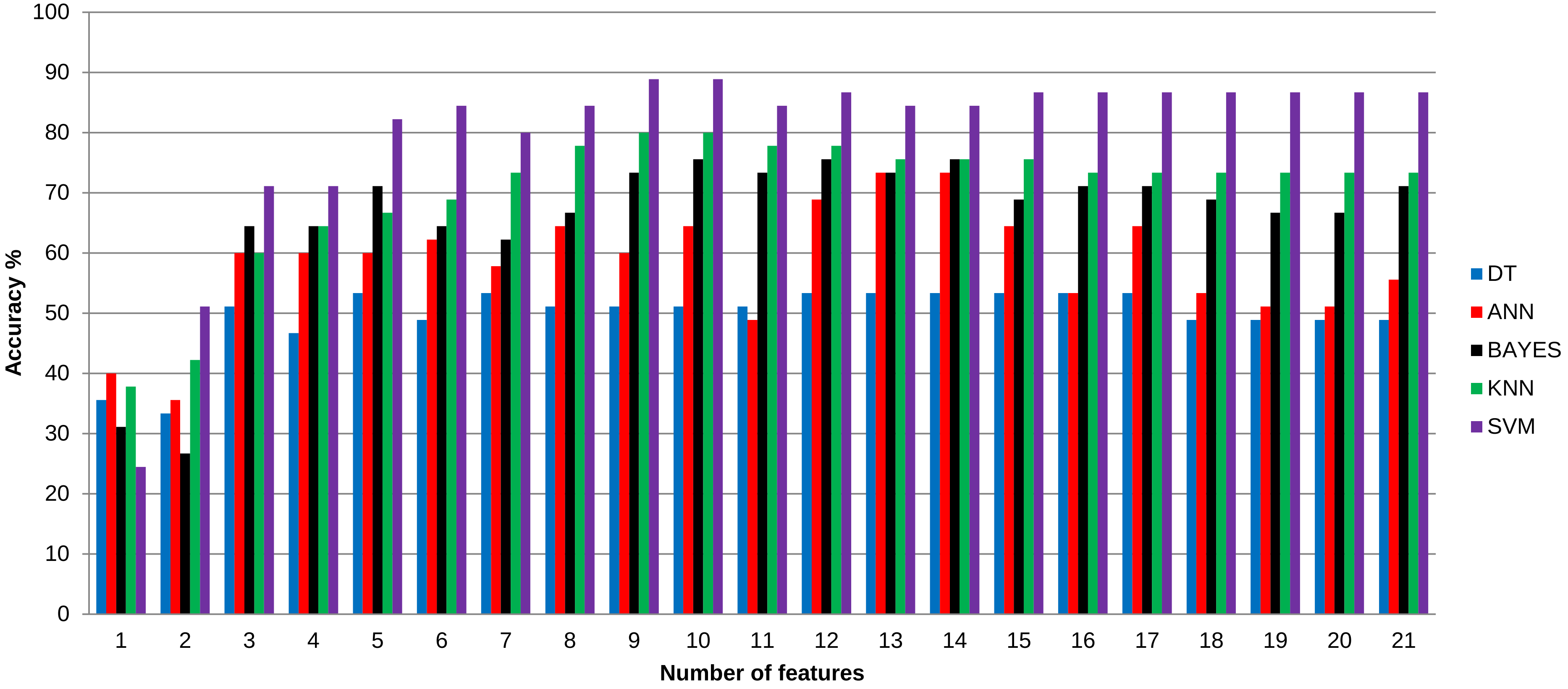
Fault detection (steady-state)



Fault classification results

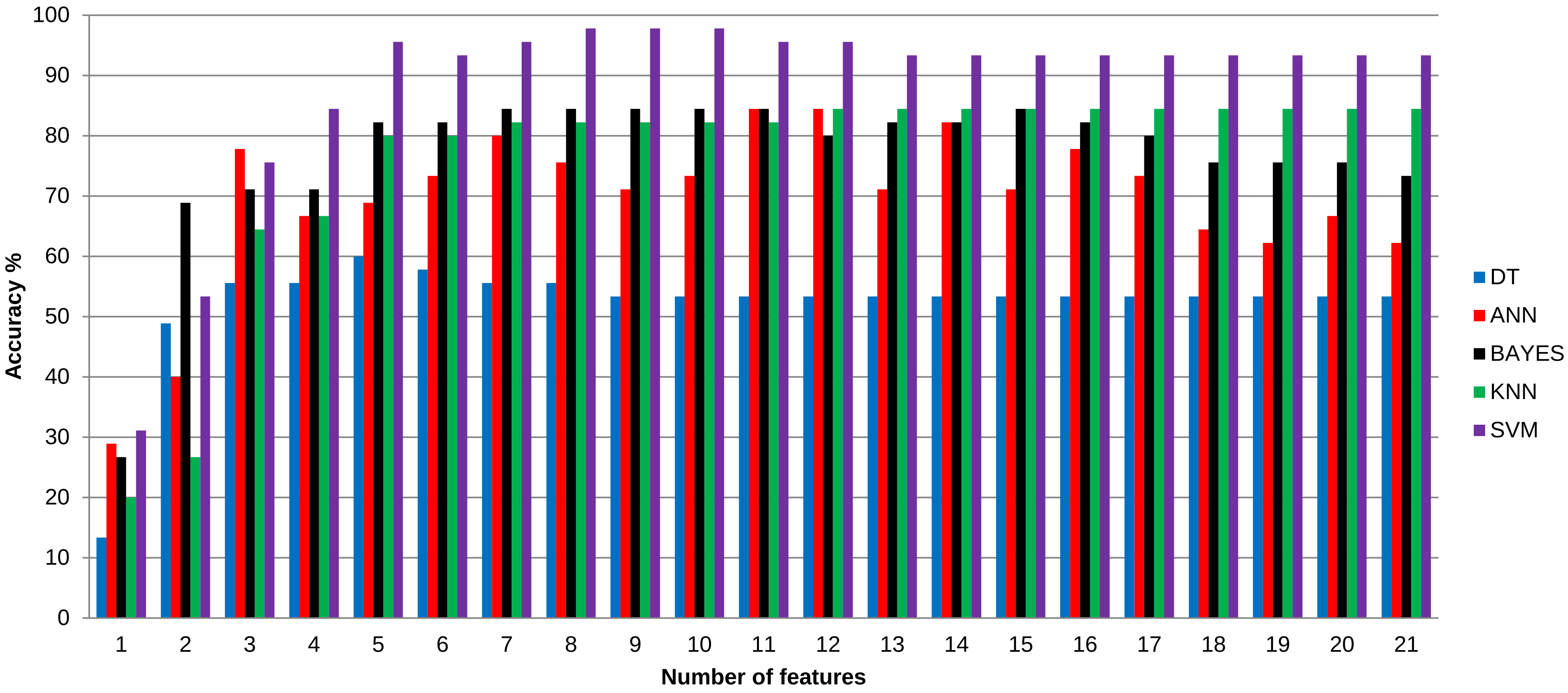
Fault classification (penetration)

UMassAmherst



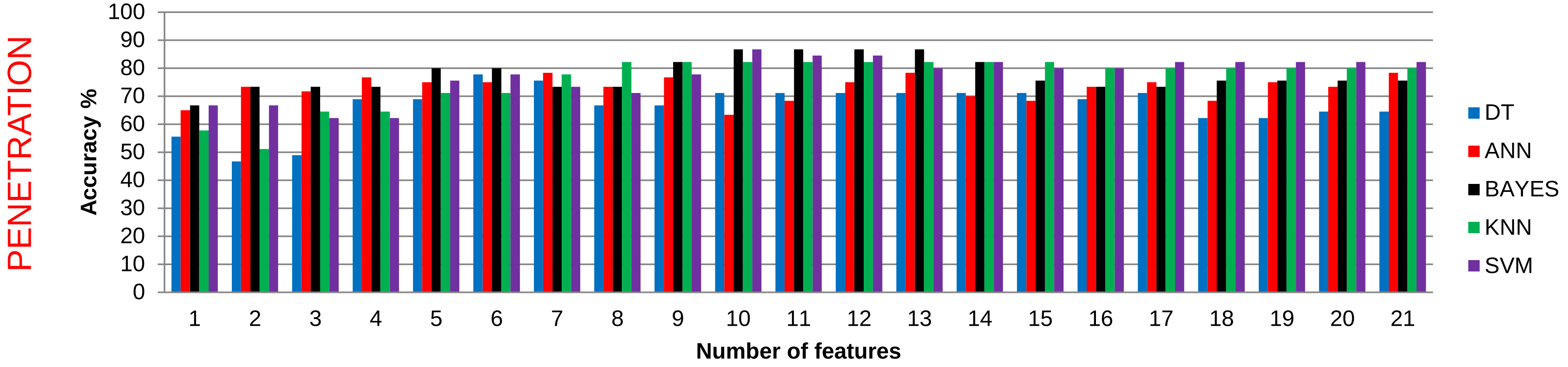
Fault classification (steady-state)

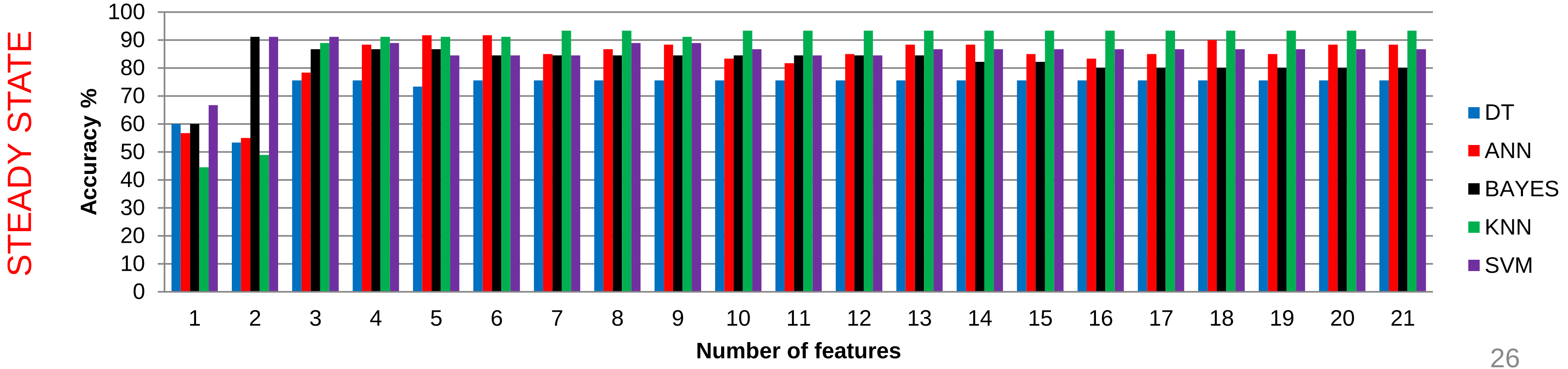
UMassAmherst



Individual wear type detection results

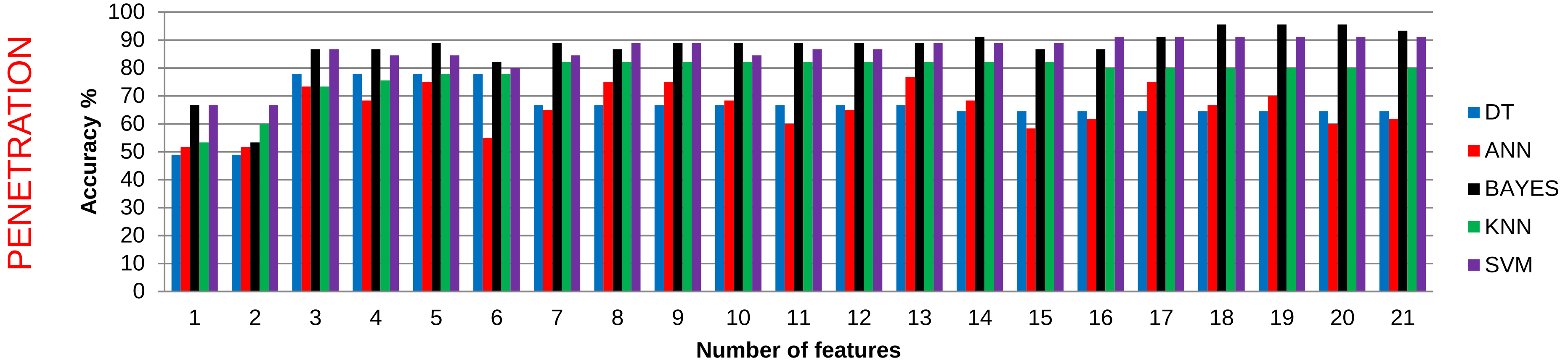
Flank wear detection

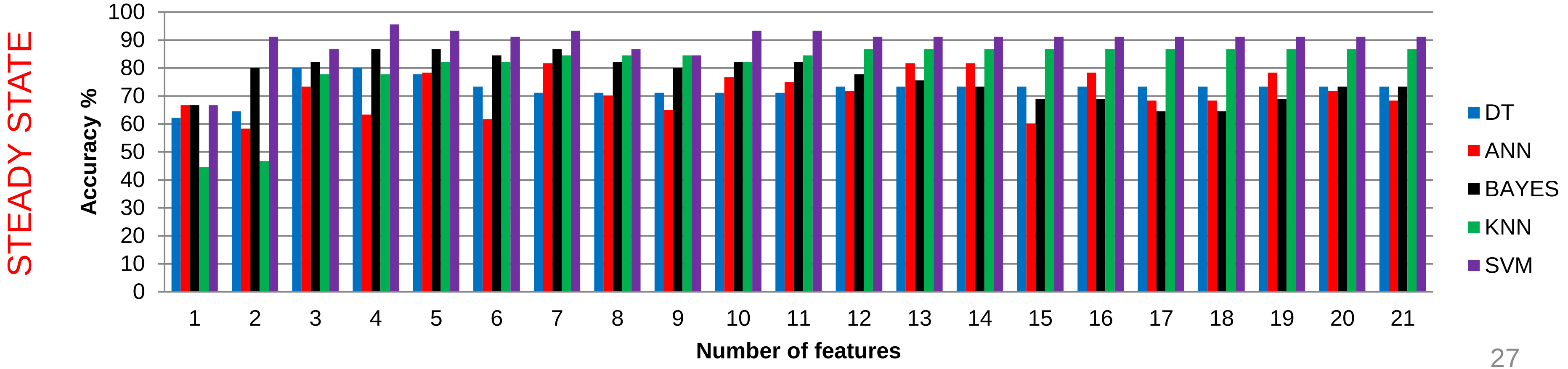




Chisel wear detection

UMassAmherst

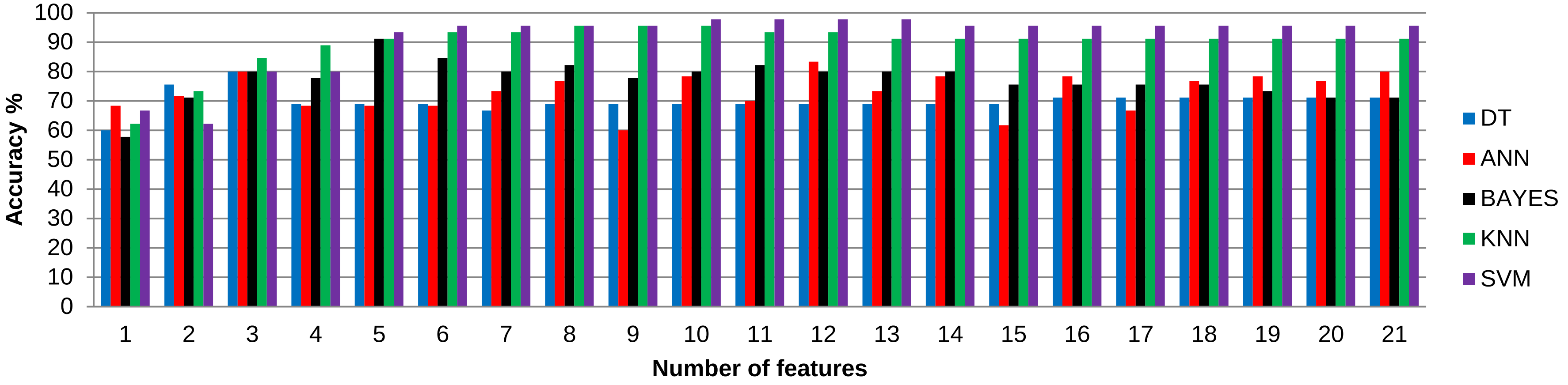




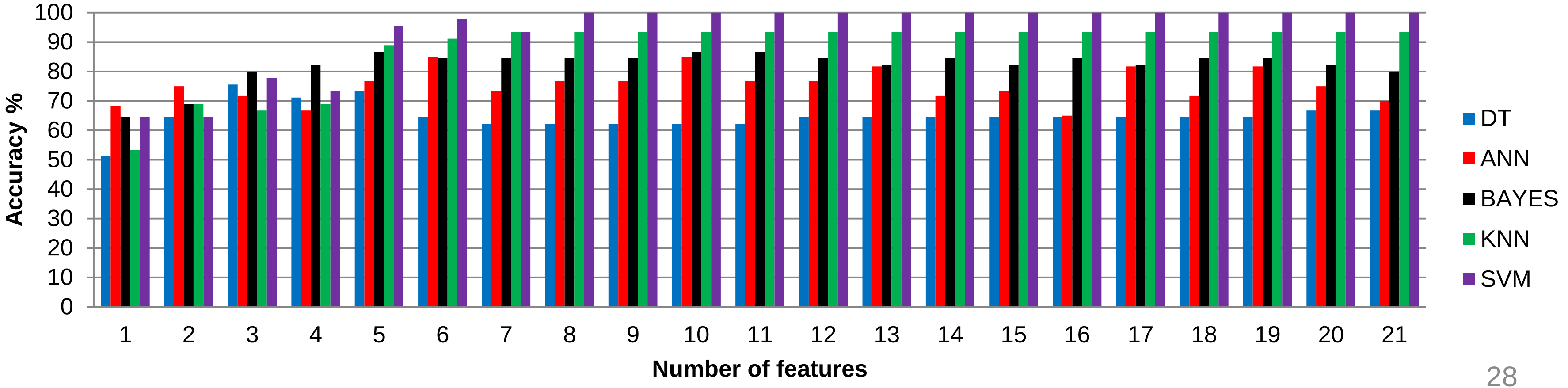
Outer corner wear detection

UMassAmherst

PENETRATION



STEADY STATE

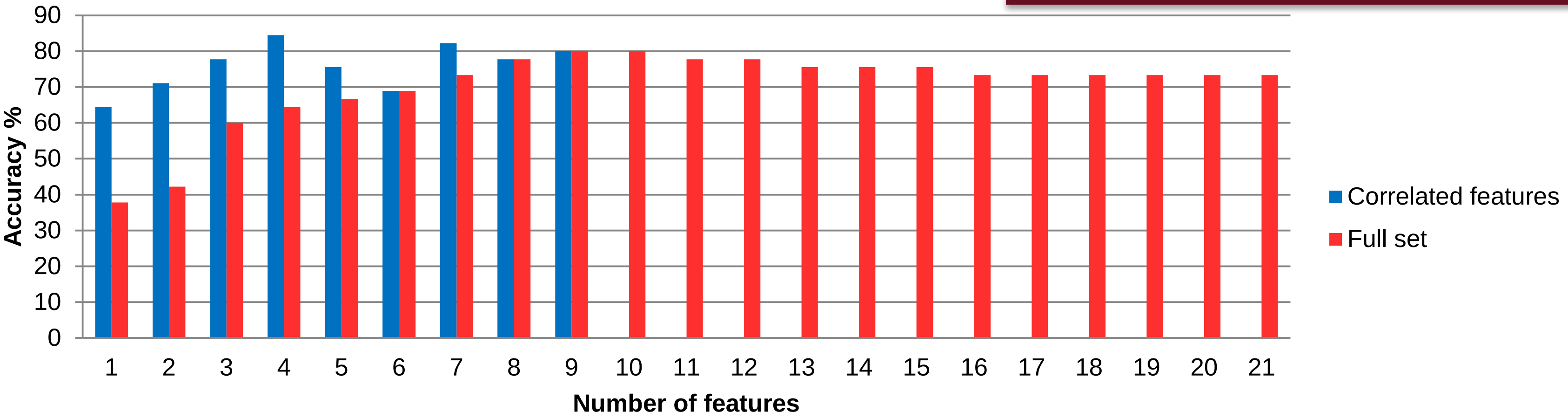


Classification with fault correlation

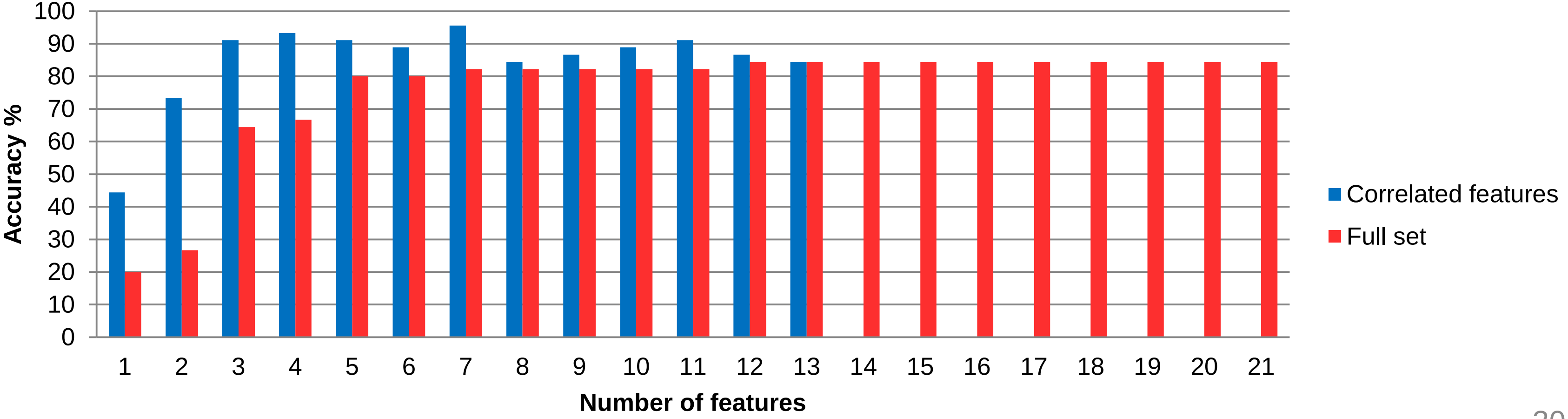
kNN

UMassAmherst

PENETRATION

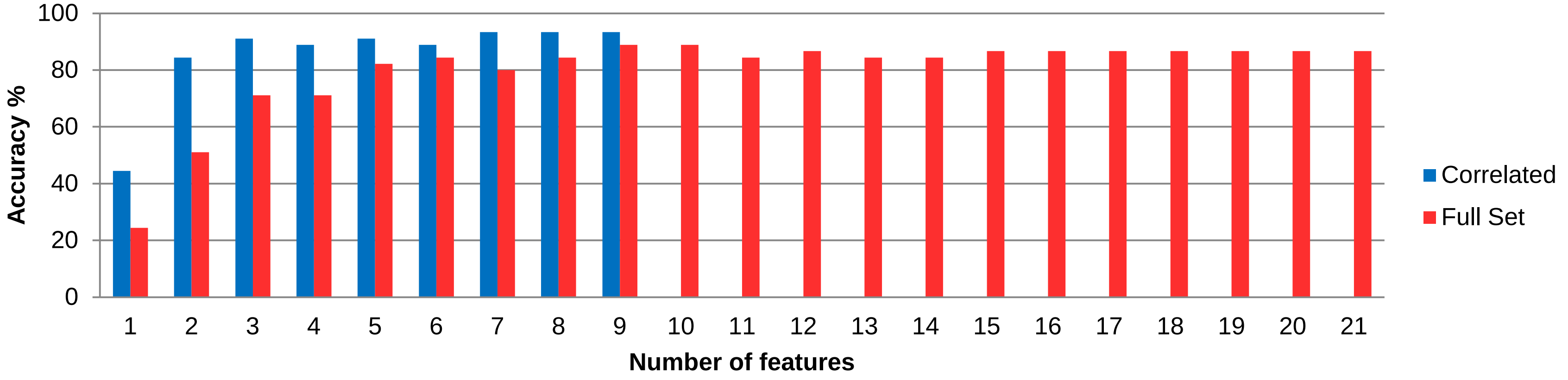


STEADY STATE

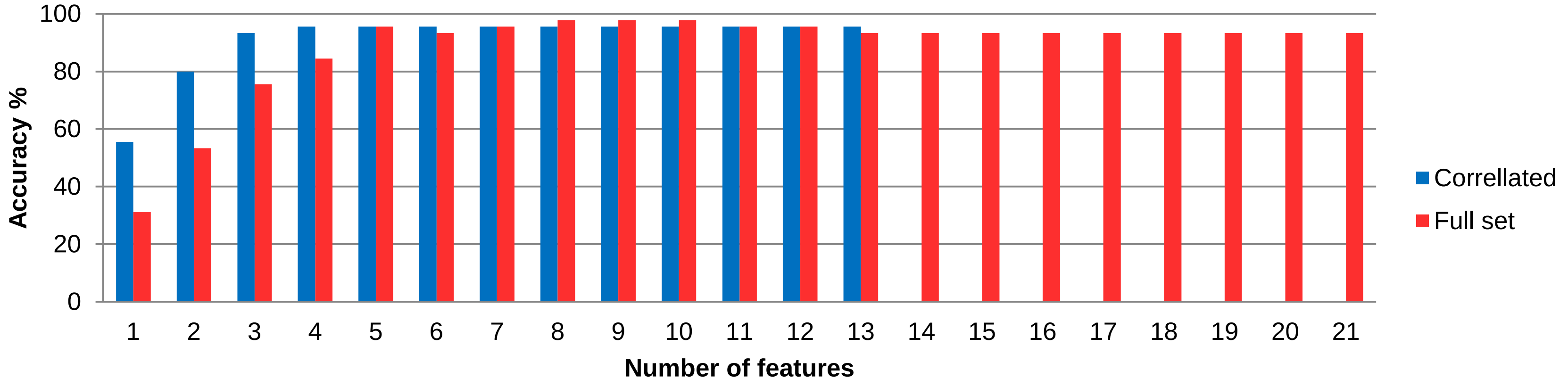


SVM

PENETRATION



STEADY STATE



Conclusions and future work

Conclusions

- SVM & kNN performed best overall out of the five classifiers.
  - **SVM**: 75% for detection, 93% for classification
  - **kNN**: 78% for detection, 84% for classification
- Impact of penetration vs. steady-state data
  - Minimal impact on fault detection accuracy.
  - Higher classification accuracies were derived from steady-state data.
  - Higher individual wear type detection accuracies were derived from steady-state data.

UMassAmherst

Conclusions

- Decision tree performed poorly
  - Excessive pruning = overly generalized
  - Simple classifier excessively sensitive to changes in training data.
- ANN performed poorly
  - Overfitting (able to classify training data without error, but not testing data)
    - Increase the number of training samples by splitting 8-second samples into increments.

UMassAmherst

Future work

- Re-train ANN with more training data.
- Implement ensemble learning techniques for improved accuracy.
- Requiring additional data: detect and classify <u>multiple</u> types of wear affecting the same drill bit.
- Optimization of parameters using ROC curves
- Optimization of the model using detailed analysis of the frequency spectrum

UMassAmherst
The Commonwealth's Flagship Campus