IoT Device and Attack Prediction

Busher Bridi, Jamie Cohen,
& Marcelo Rodrigues

BerkeleyMICS

Overview

- **IoT devices have become increasingly prevalent in today's interconnected world**
  - IoT devices play a vital role in many industries, enhancing efficiency, and ease of use.
- **Clustering IoT devices can provide insight into device management and security**
  - **Problem Statement 1:** Determine whether IoT devices can be clustered based on their features
- **It is crucial to keep IoT devices safe from cyber attacks, especially mission critical ones. Analyzing IoT traffic can detect and possibly mitigate possible threats.**
  - **Problem Statement 2:** Determine whether malicious network activity can be identified and predicted

BerkeleyMICS

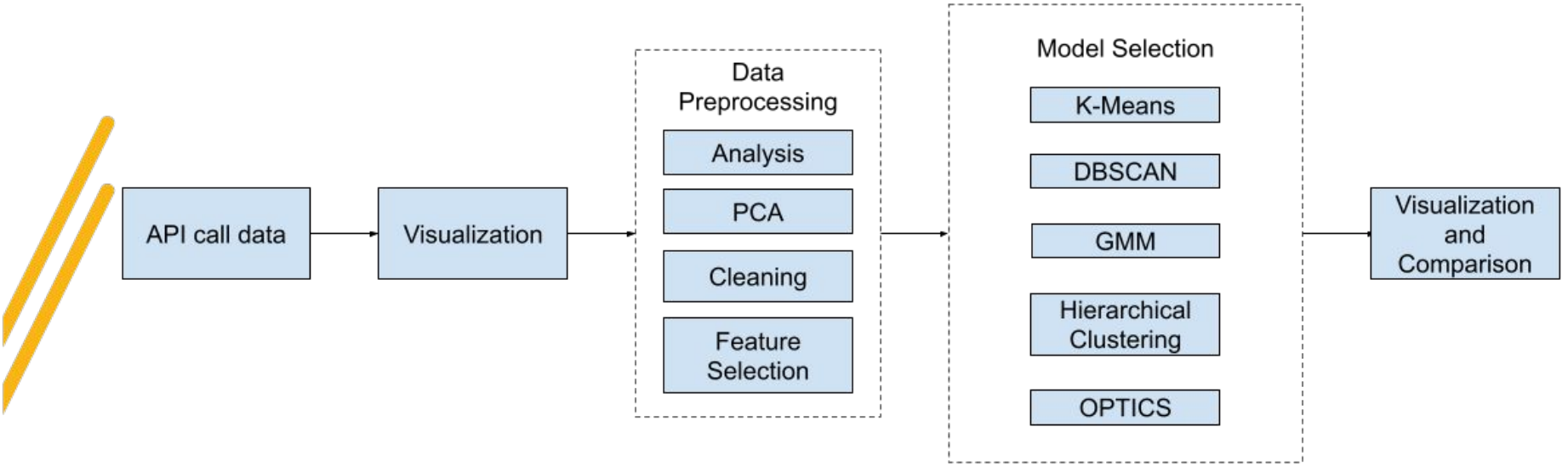
Objective

- **Objective 1:** Create a model that will show which features can be used to identify different types of IoT devices
  - **Success:** An equal amount of distinct cluster to the amount of unique IoT Devices in the datasets

- **Objective 2:** Create a model that can accurately predict type of IoT attack
  - **Success:** 50% accuracy

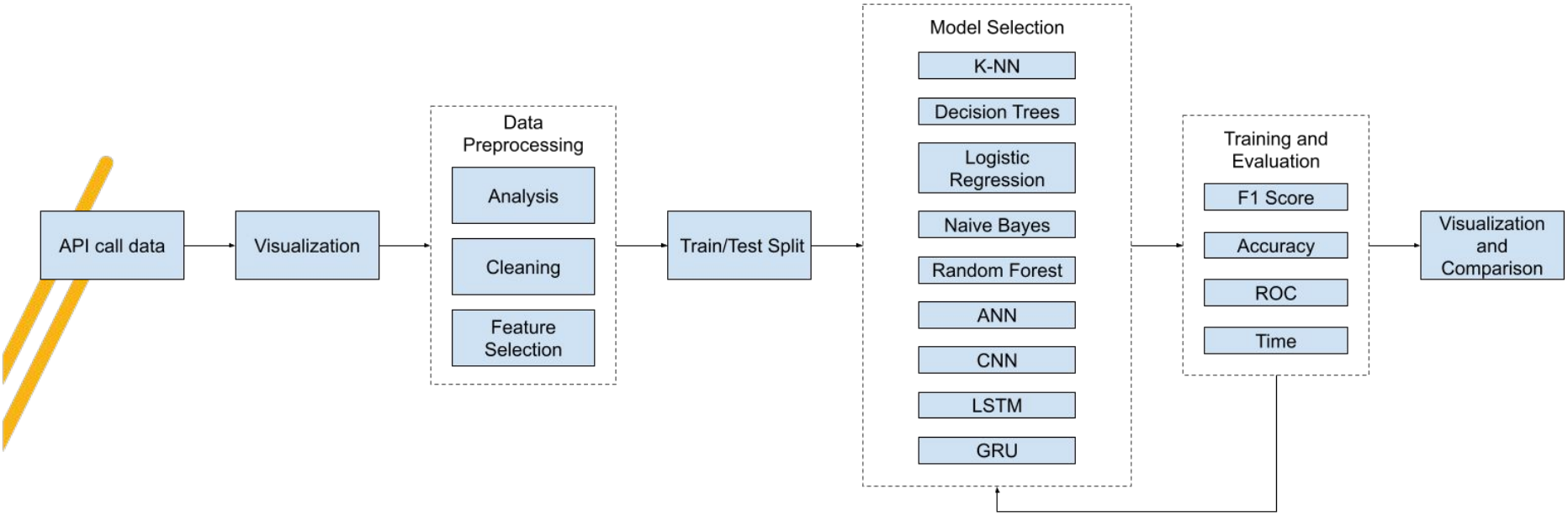
BerkeleyMICS

Methodology - Device Identification



BerkeleyMICS

Methodology - Attack Prediction



BerkeleyMICS

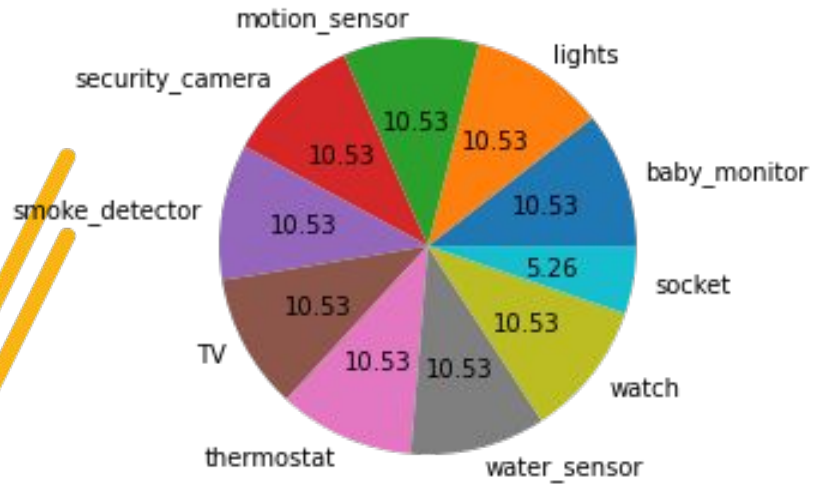
Datasets

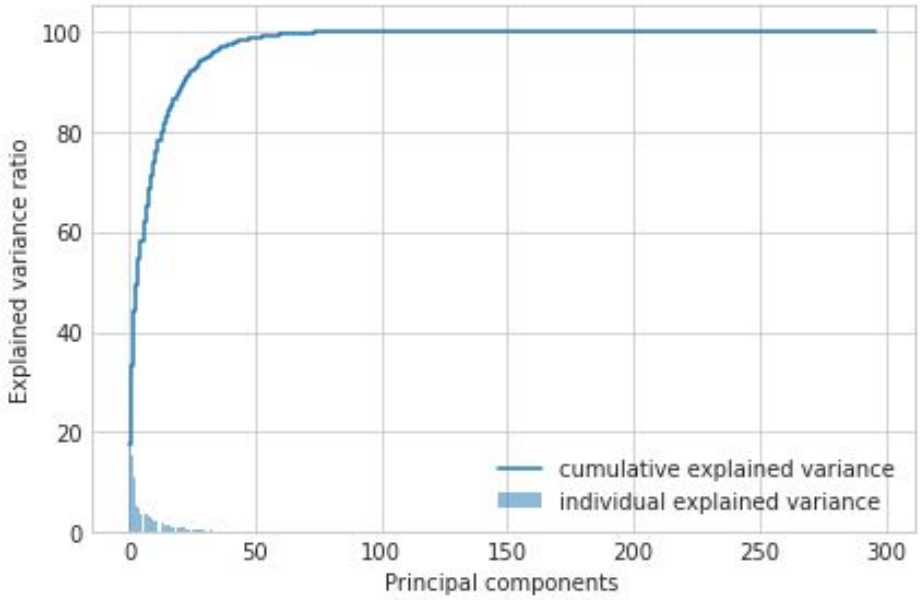
- **Dataset 1: IoT Device Identification**
  - 297 Features
  - 10 Types of IoT Devices within dataset
  - 13,701 data points

- **Dataset 2: IoT Attack Prediction**
  - 16 Features
  - Labeled with attack type
    - Possible types of attack include: SHA, DFA, SFA, SYA, and VNA
  - 10,845 data points

BerkeleyMICS

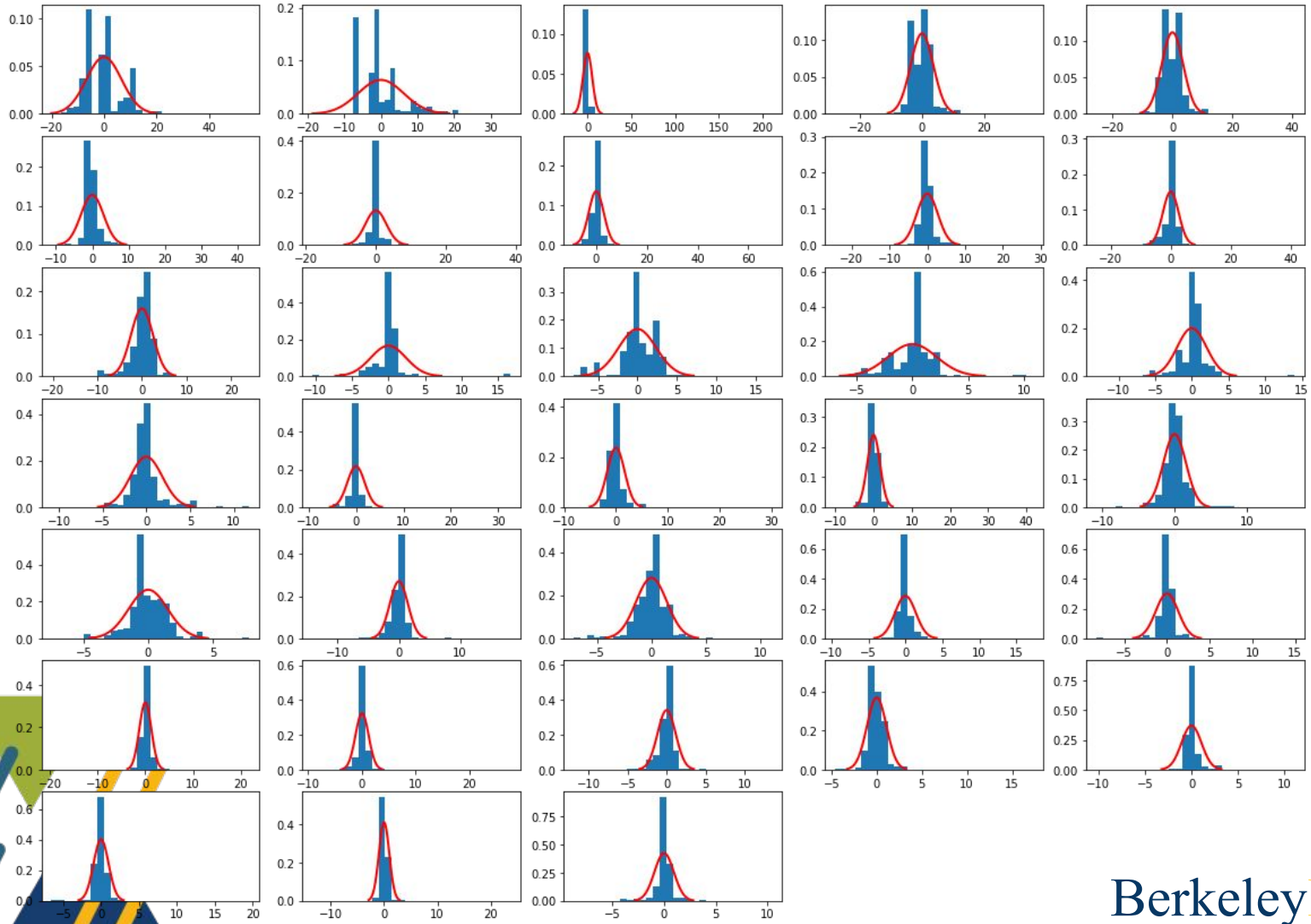
Visualization - Device





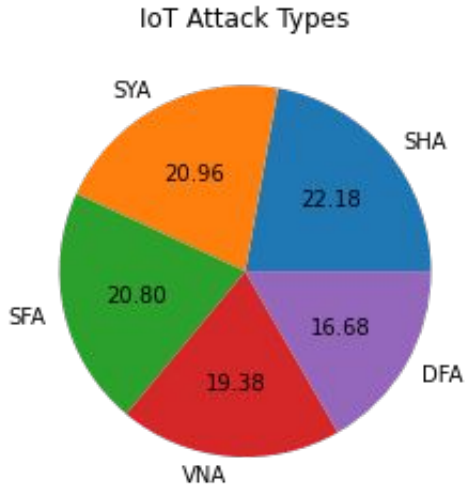
BerkeleyMICS

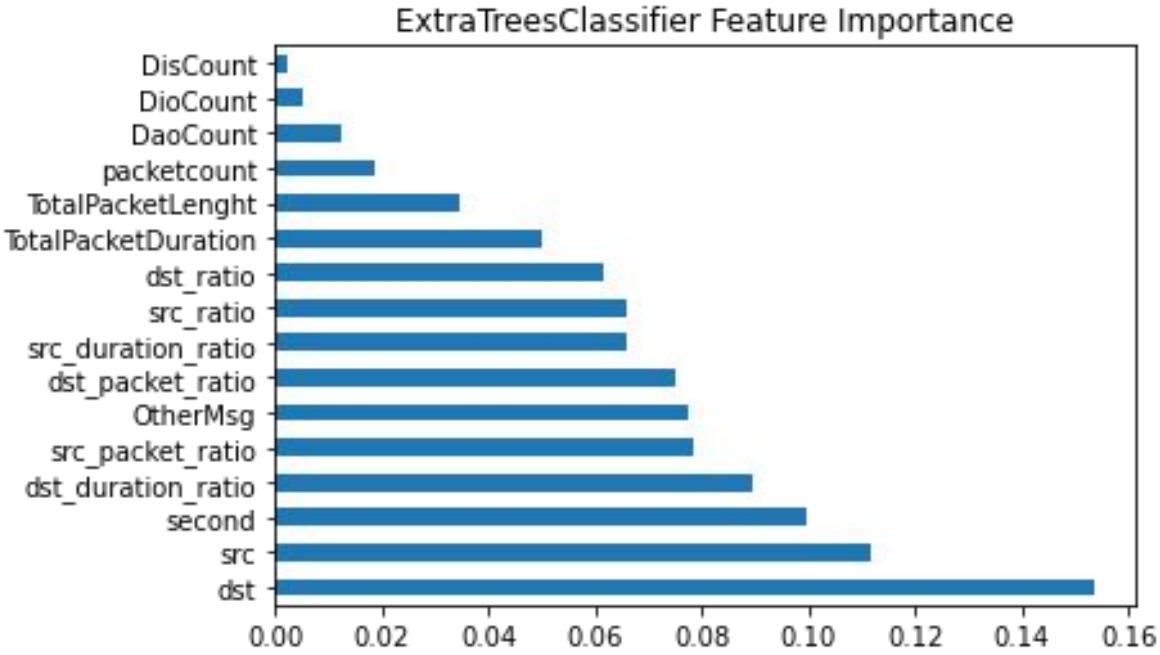
Visualization - Device



BerkeleyMICS

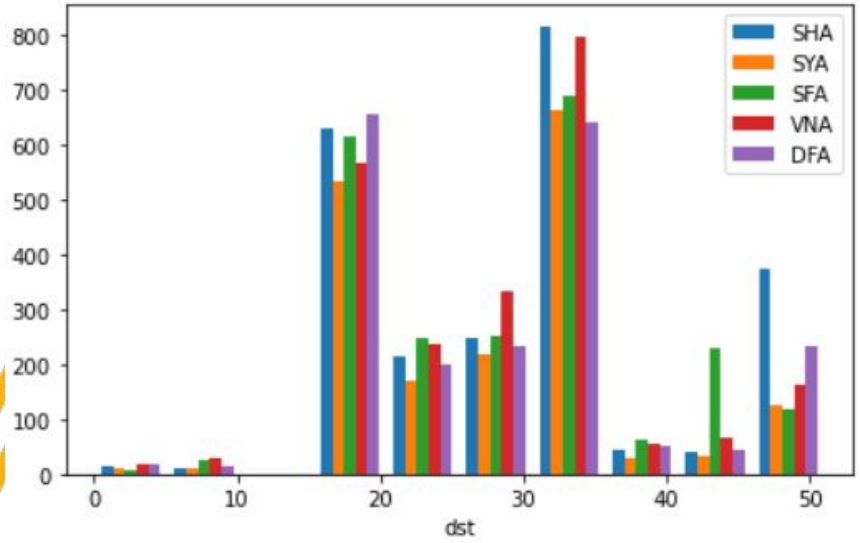
Visualization - Prediction

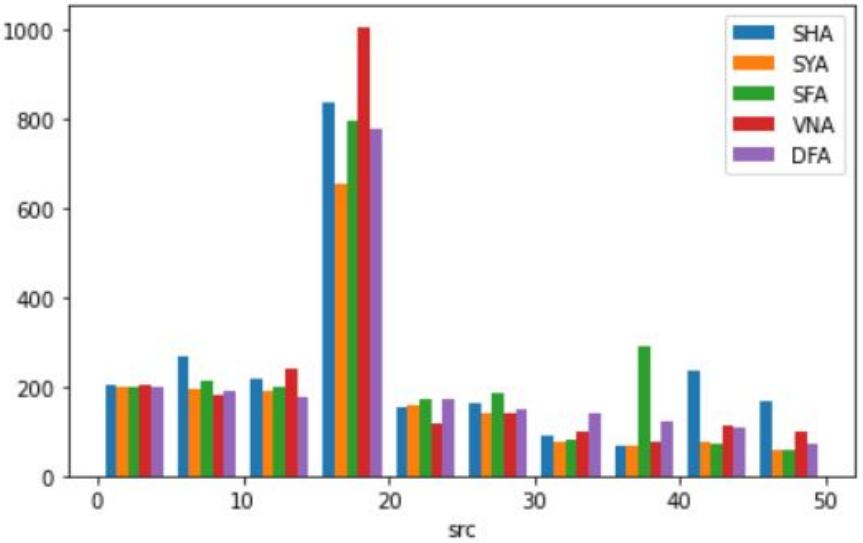


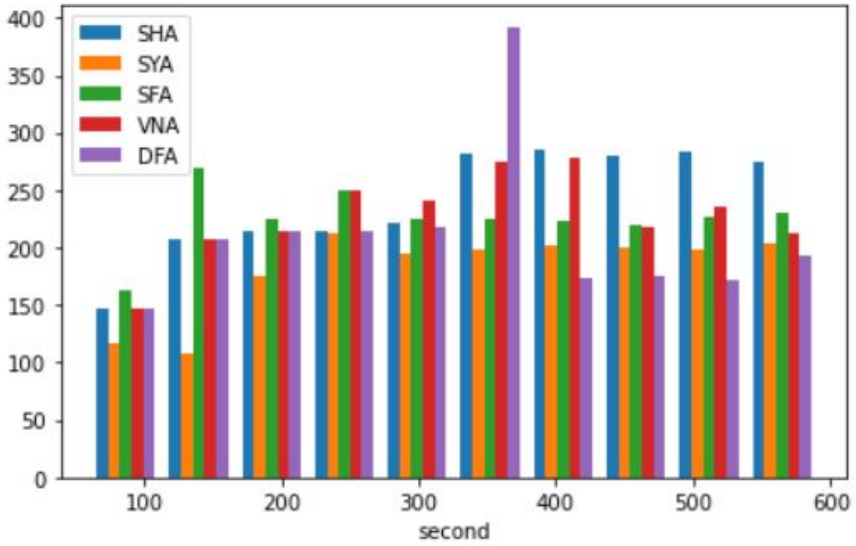


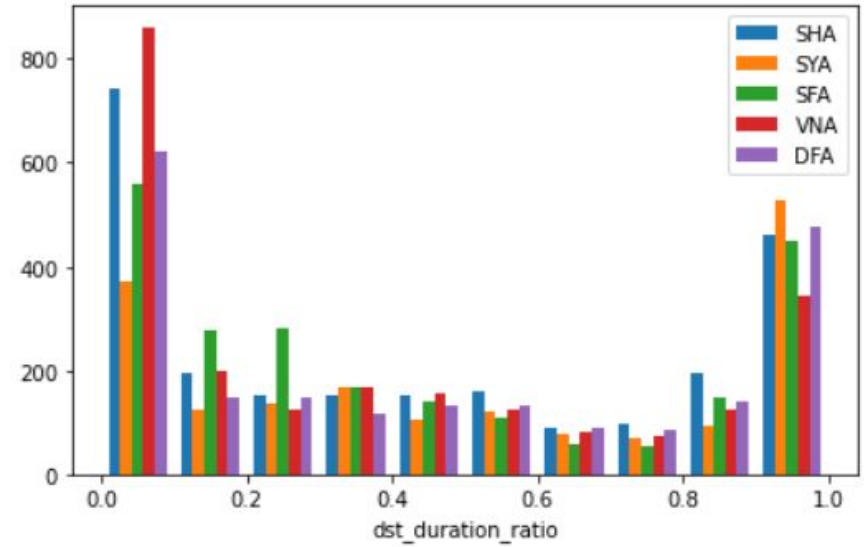
BerkeleyMICS

Visualization - Prediction





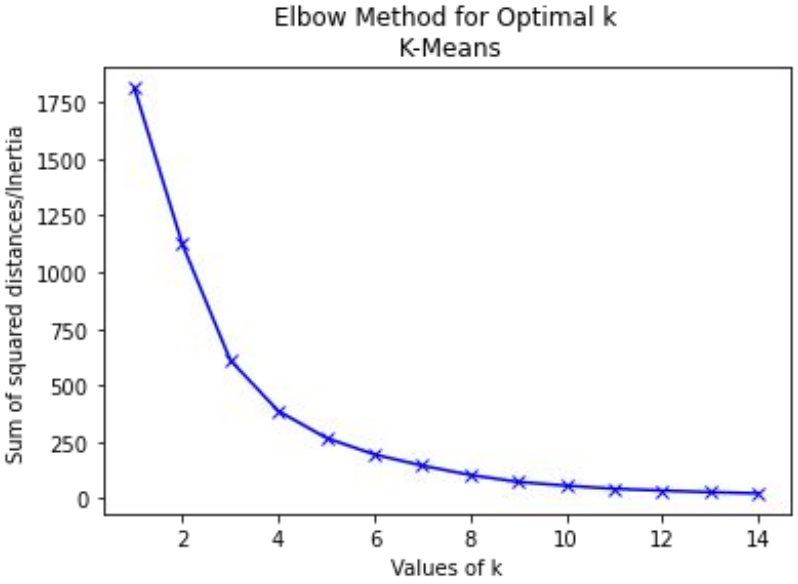


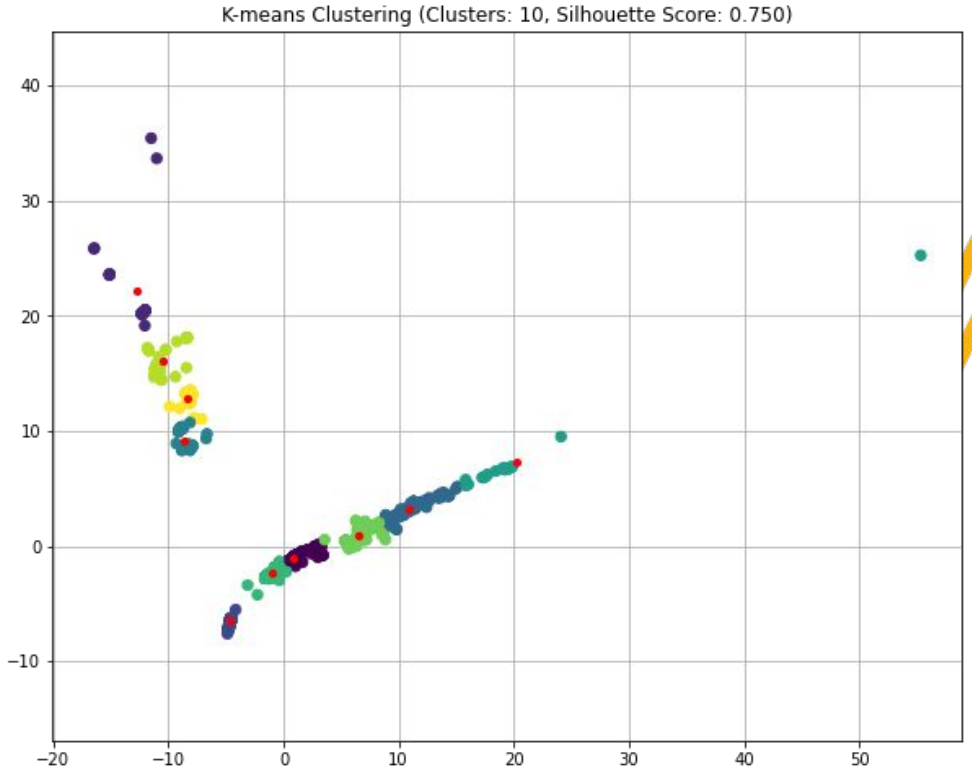


Device Identification

BerkeleyMICS

K-Means



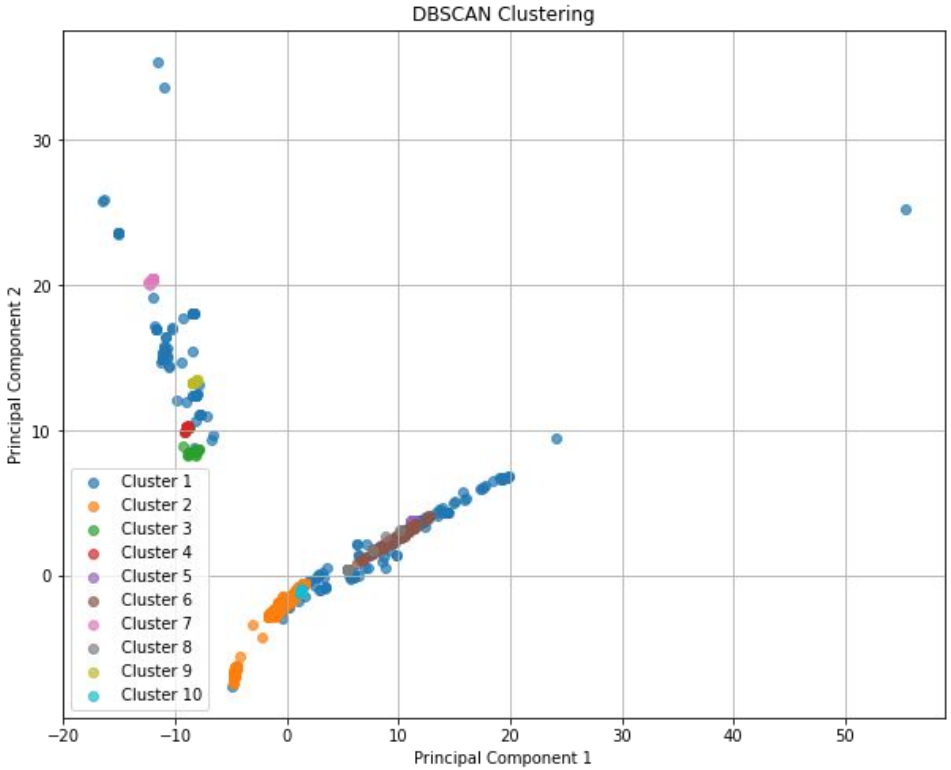


Average mean clusters BIC: 6.84

BerkeleyMICS

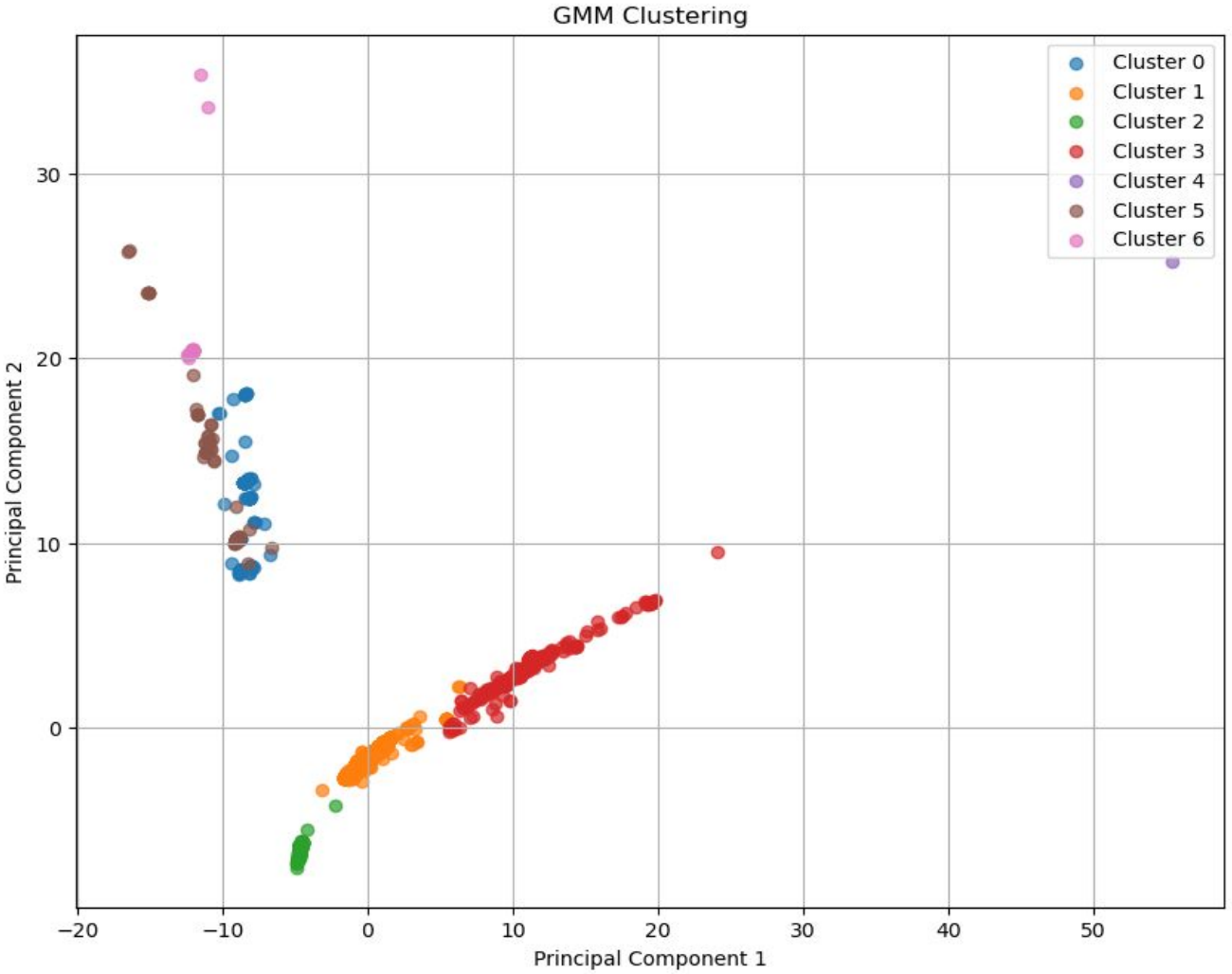
DBSCAN

- Epsilon = 6, Minimum_Samples = 20
- Silhouette Coefficient: 0.418



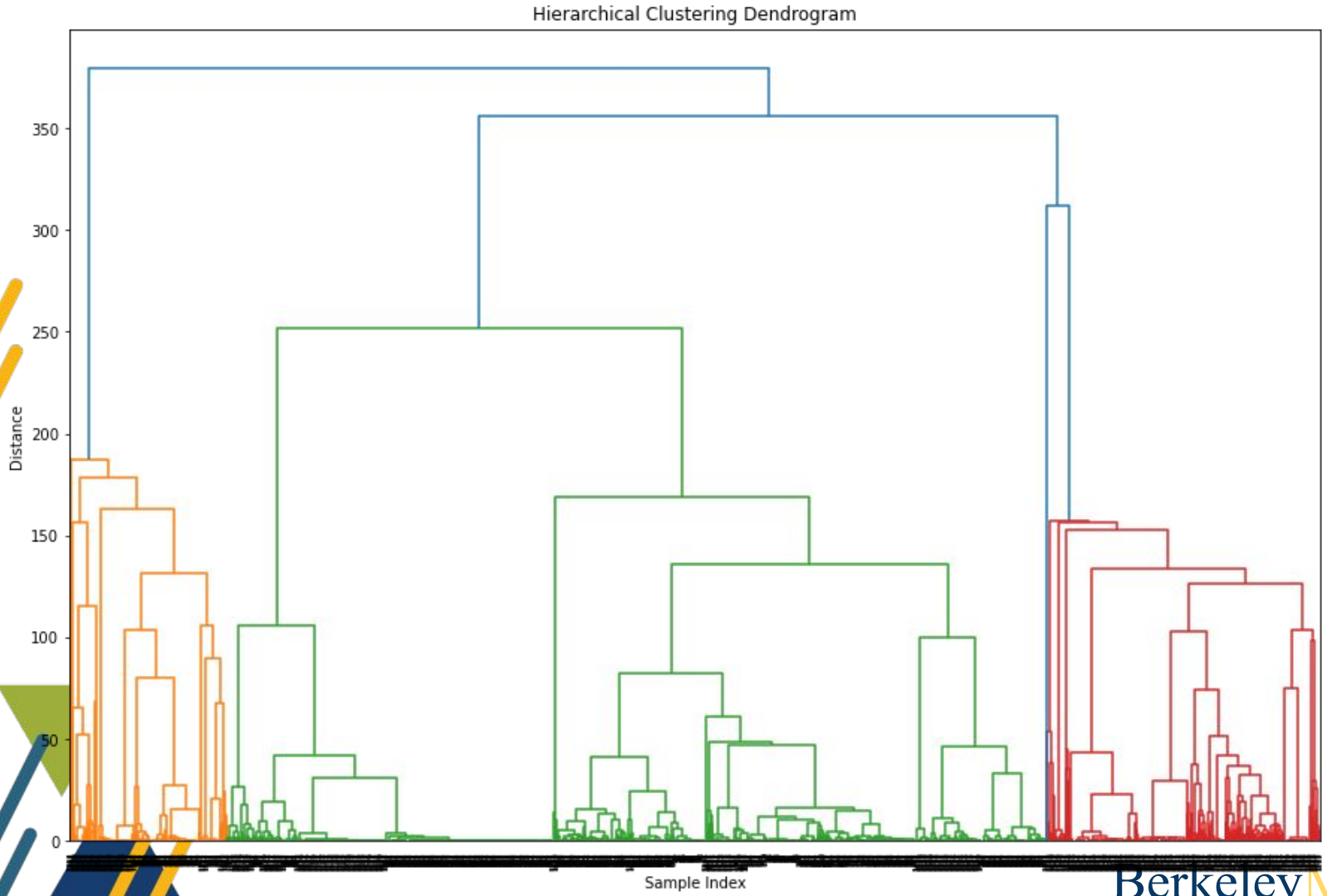
BerkeleyMICS

Gaussian Mixture Model



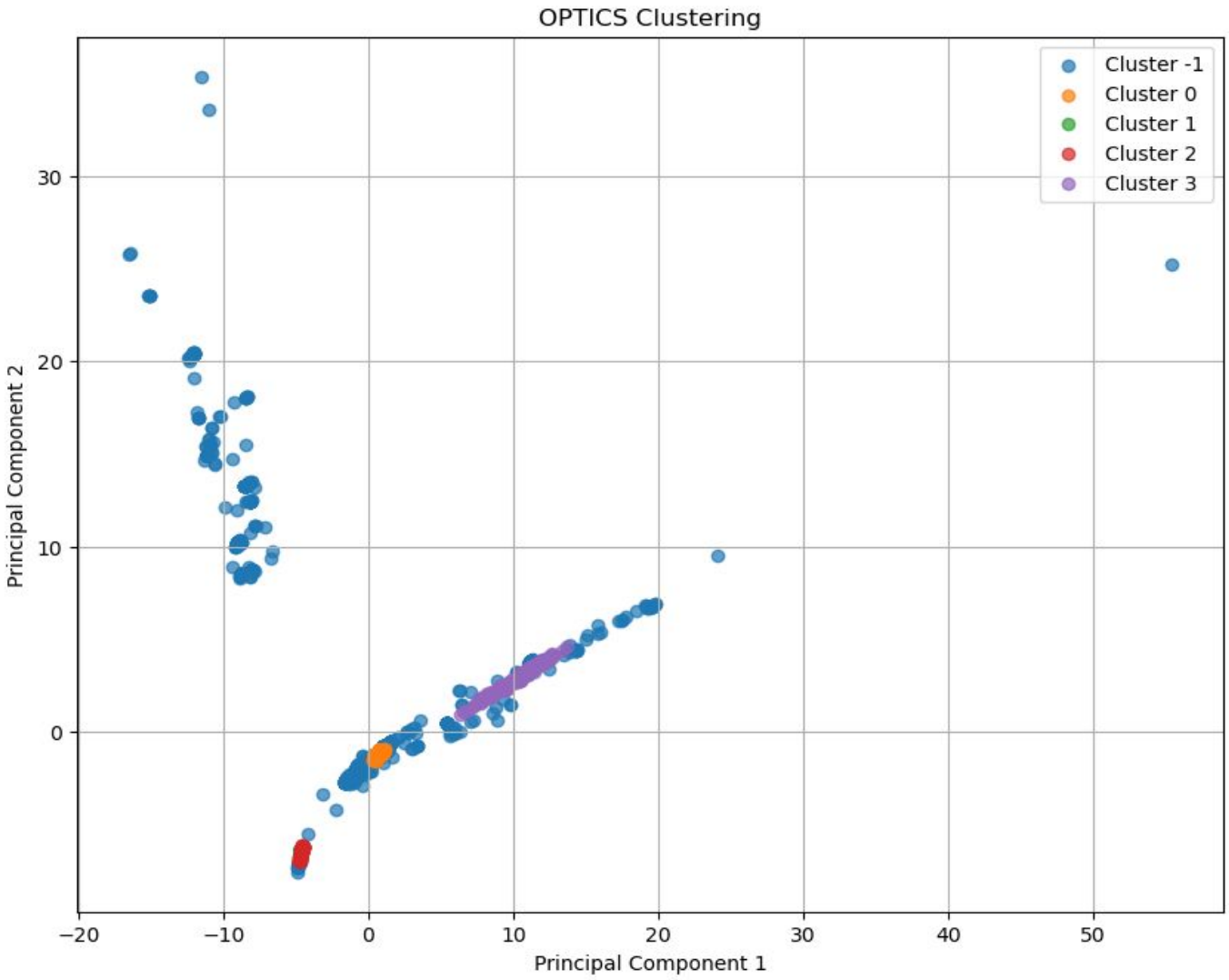
BerkeleyMICS

Hierarchical Clustering



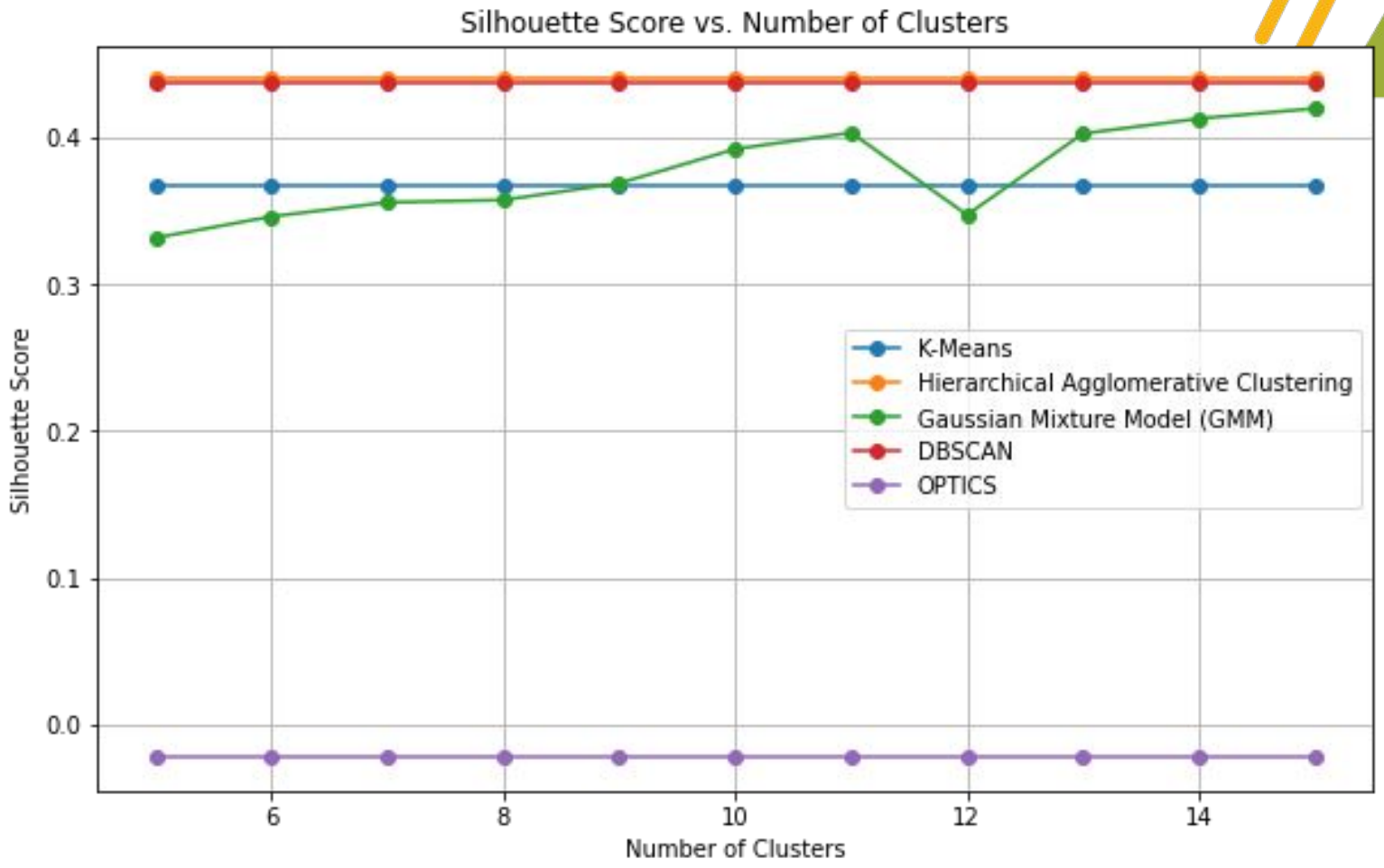
BerkeleyMICS

Optics



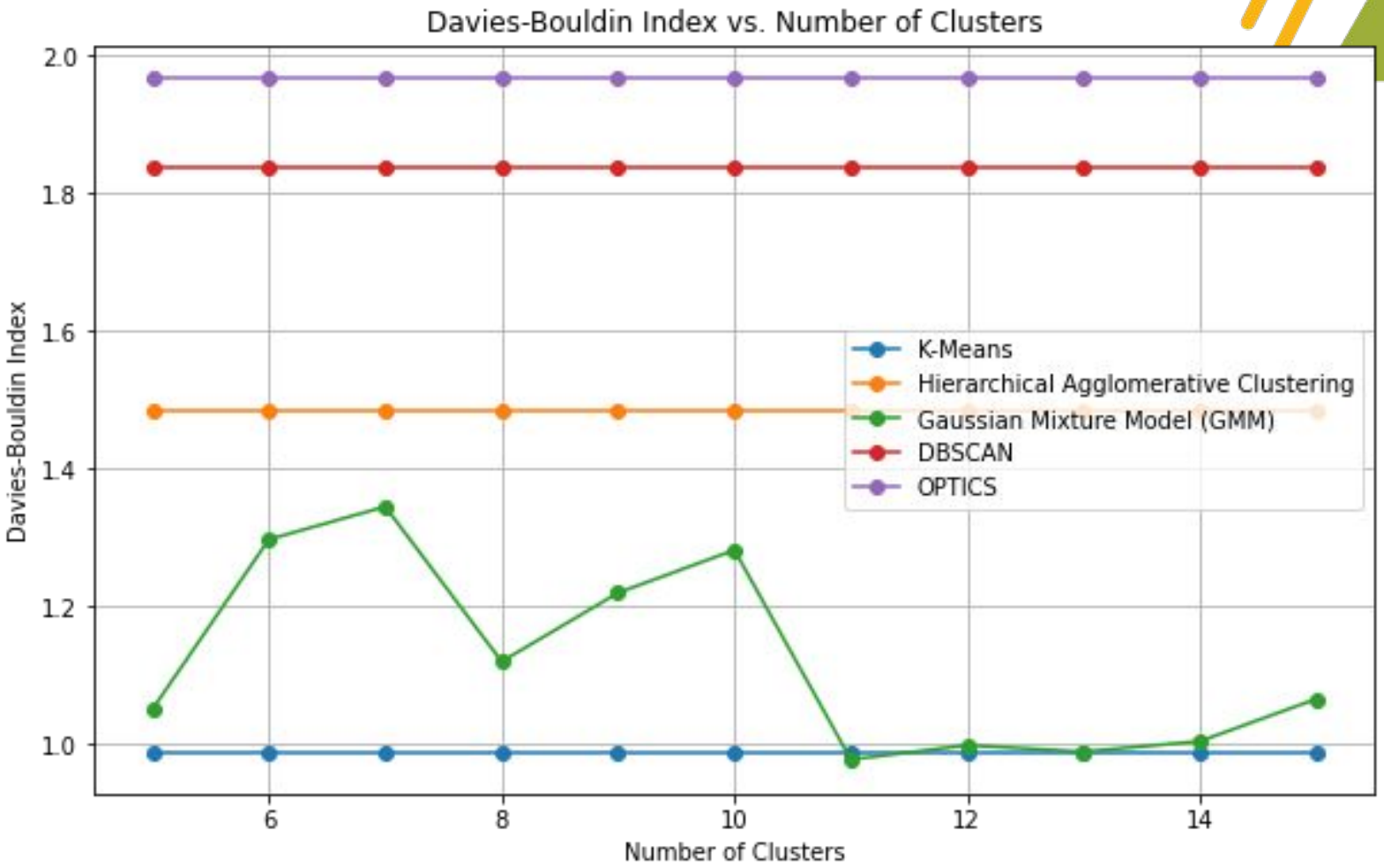
BerkeleyMICS

Comparison



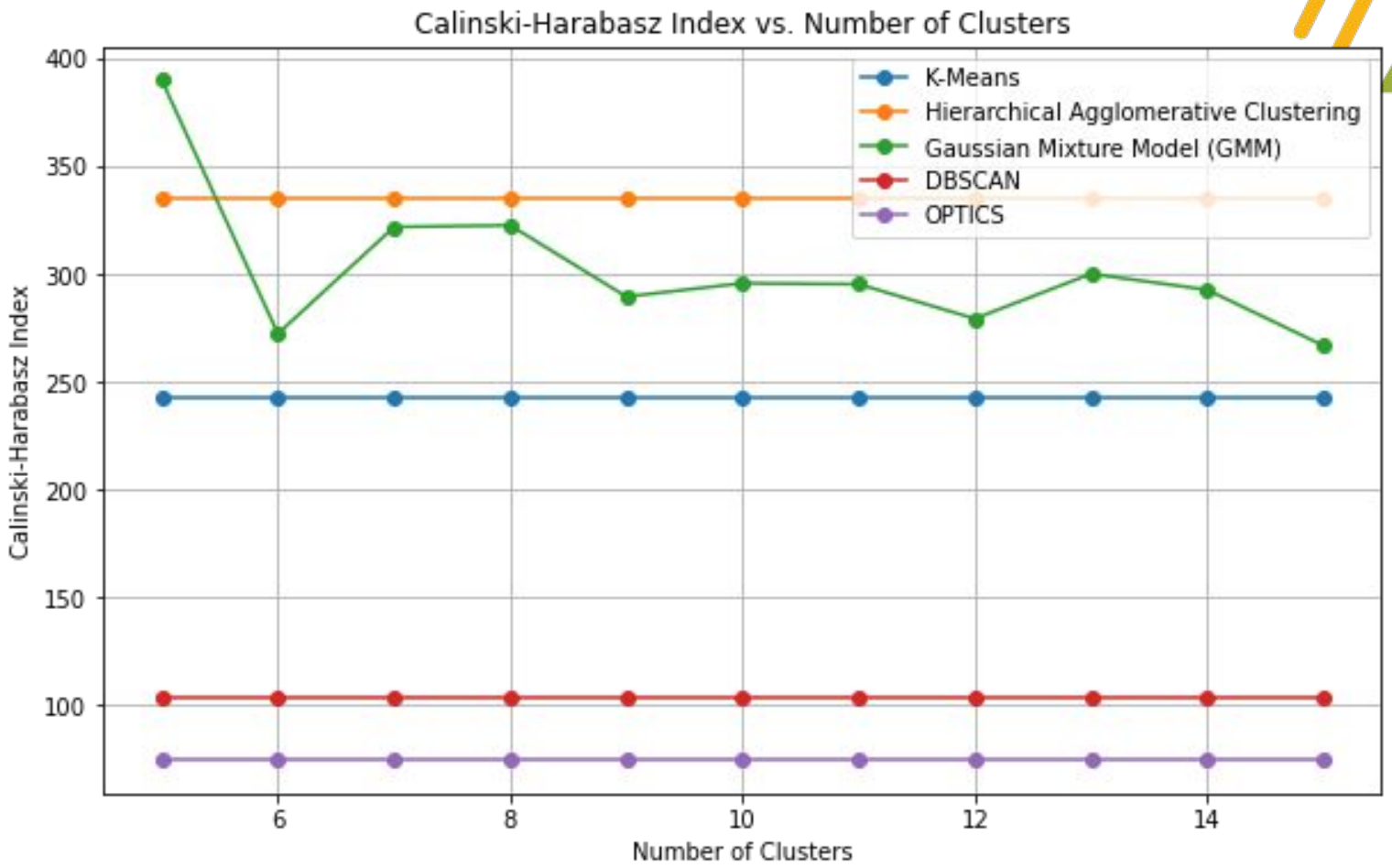
BerkeleyMICS

Comparison



BerkeleyMICS

Comparison

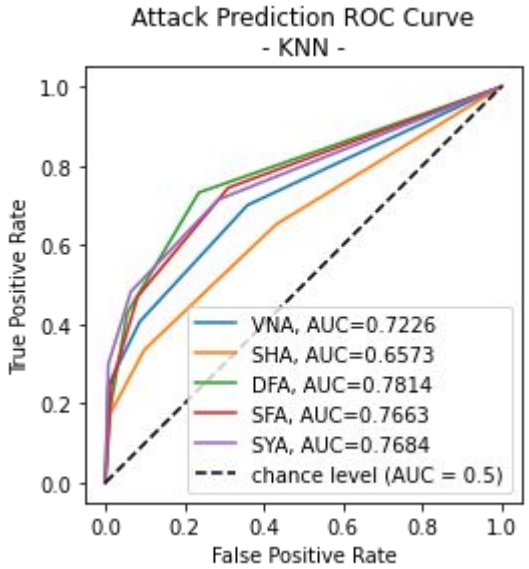


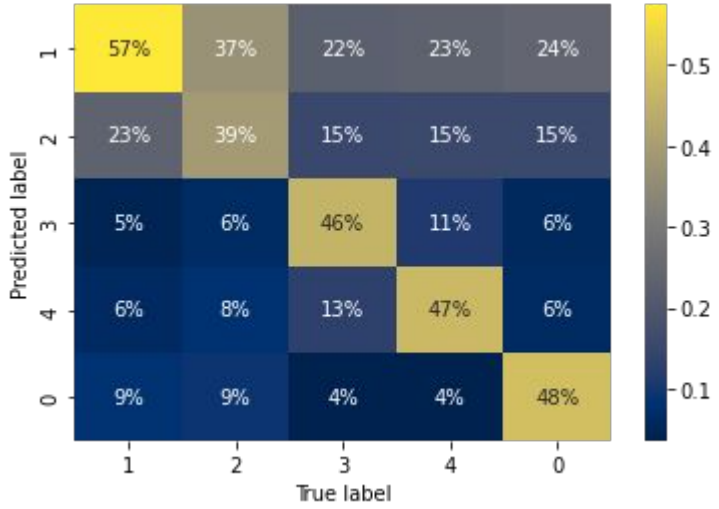
BerkeleyMICS

Attack Prediction

K-Nearest Neighbors

	precision	recall	f1-score	support
0	0.33	0.57	0.42	622
1	0.39	0.39	0.39	727
2	0.57	0.46	0.51	542
3	0.60	0.47	0.53	662
4	0.67	0.48	0.56	701
accuracy			0.47	3254
macro avg	0.51	0.48	0.48	3254
weighted avg	0.51	0.47	0.48	3254

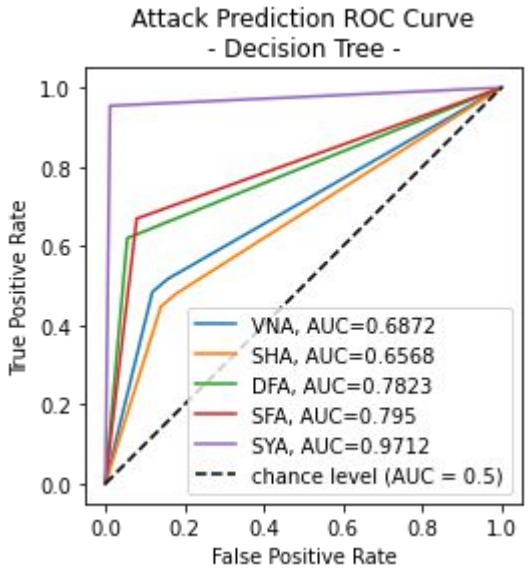


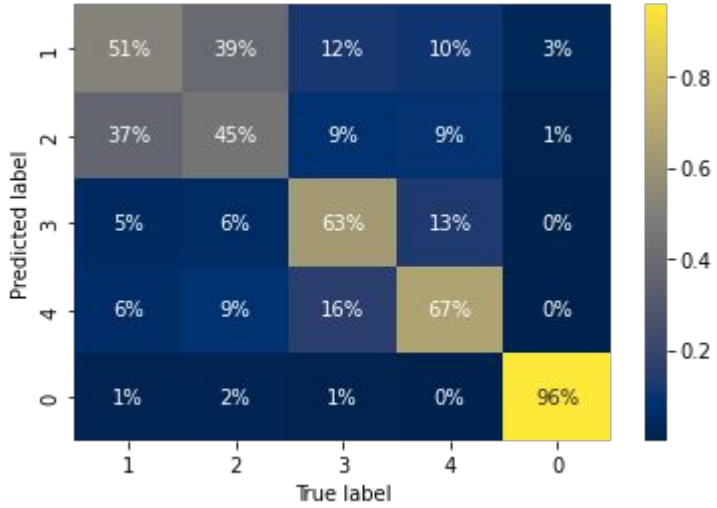


BerkeleyMICS

Decision Trees

	precision	recall	f1-score	support
0	0.42	0.51	0.46	622
1	0.49	0.45	0.47	727
2	0.68	0.63	0.66	542
3	0.70	0.67	0.69	662
4	0.96	0.96	0.96	701
accuracy			0.65	3254
macro avg	0.65	0.64	0.65	3254
weighted avg	0.65	0.65	0.65	3254

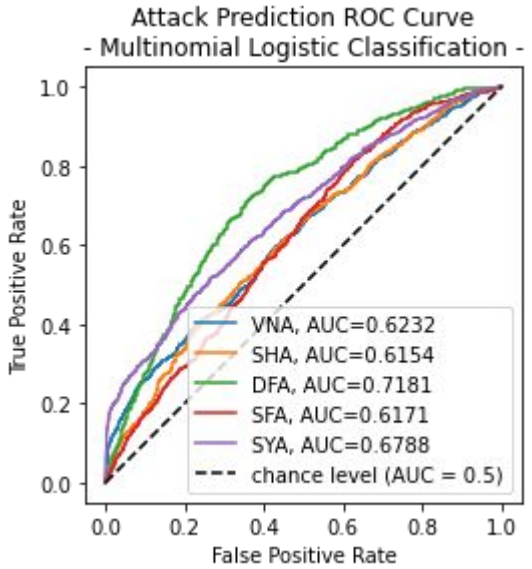


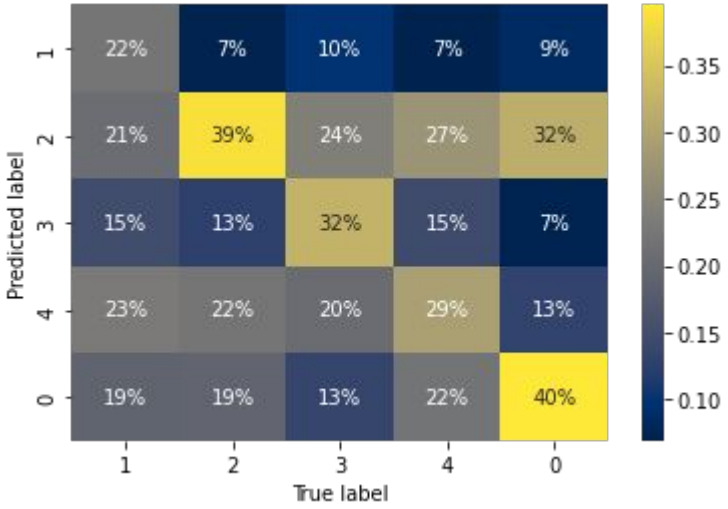


BerkeleyMICS

Logistic Regression

	precision	recall	f1-score	support
0	0.39	0.22	0.28	622
1	0.30	0.39	0.34	727
2	0.34	0.32	0.33	542
3	0.28	0.29	0.28	662
4	0.37	0.40	0.38	701
accuracy			0.33	3254
macro avg	0.34	0.32	0.32	3254
weighted avg	0.33	0.33	0.32	3254

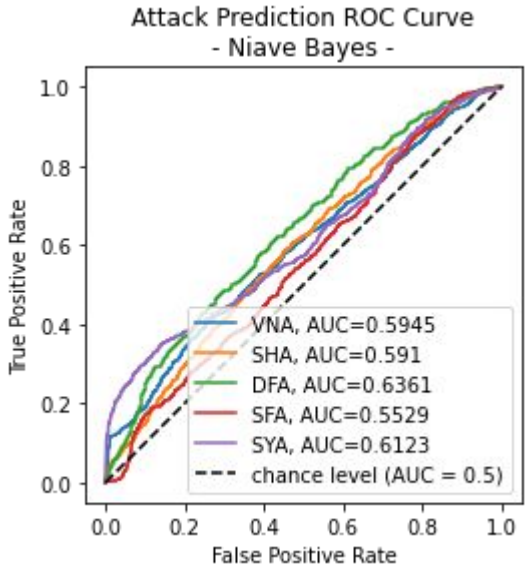


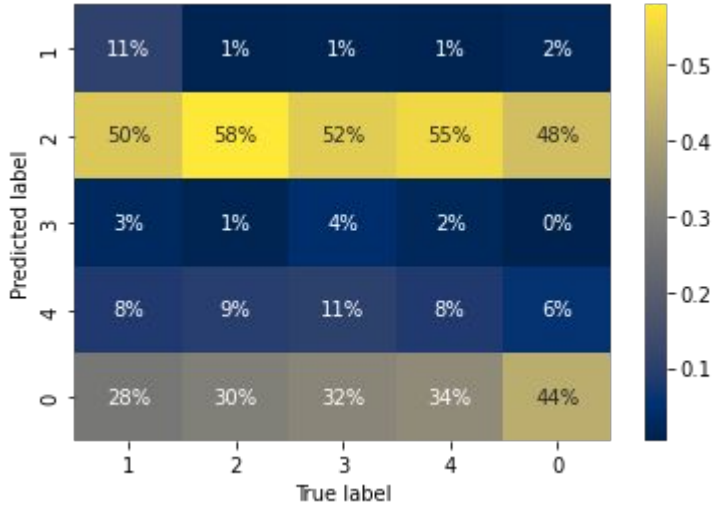


BerkeleyMICS

Naive Bayes

	precision	recall	f1-score	support
0	0.67	0.11	0.19	622
1	0.25	0.58	0.34	727
2	0.35	0.04	0.08	542
3	0.20	0.08	0.12	662
4	0.28	0.44	0.34	701
accuracy			0.27	3254
macro avg	0.35	0.25	0.21	3254
weighted avg	0.34	0.27	0.22	3254

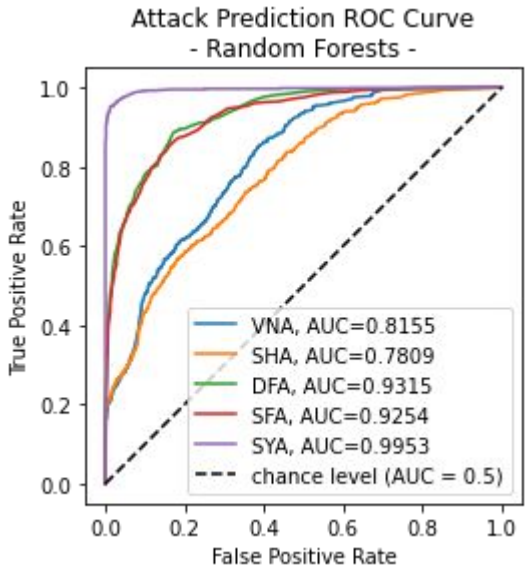


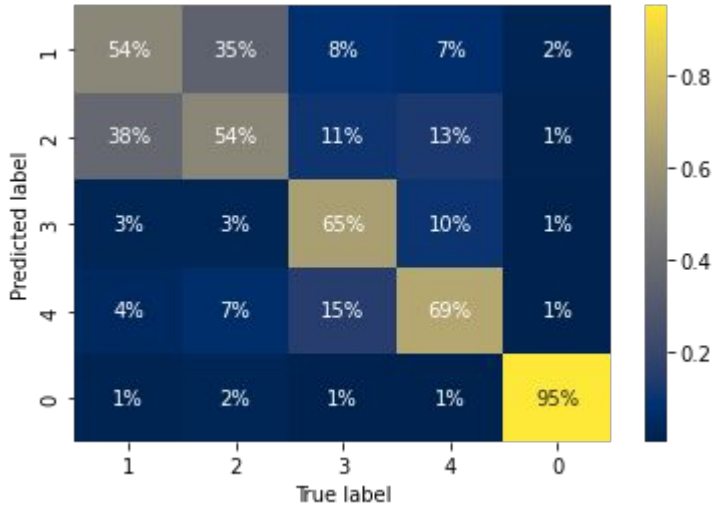


BerkeleyMICS

Random Forests

	precision	recall	f1-score	support
0	0.48	0.54	0.51	622
1	0.50	0.54	0.52	727
2	0.76	0.65	0.70	542
3	0.74	0.69	0.71	662
4	0.95	0.95	0.95	701
accuracy			0.68	3254
macro avg	0.69	0.67	0.68	3254
weighted avg	0.69	0.68	0.68	3254



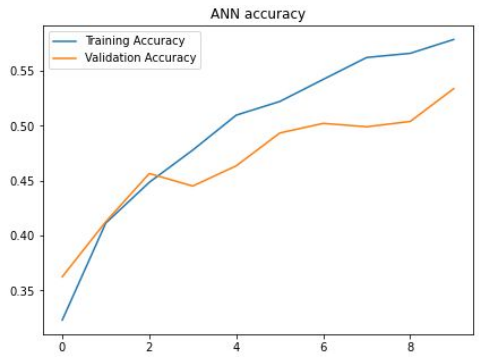


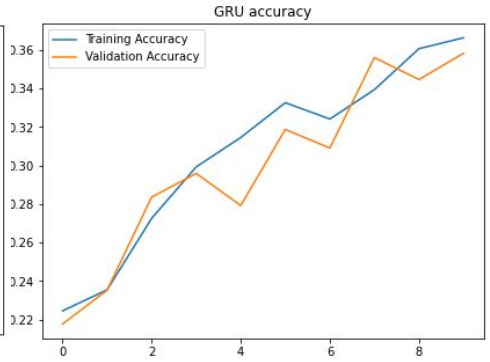
BerkeleyMICS

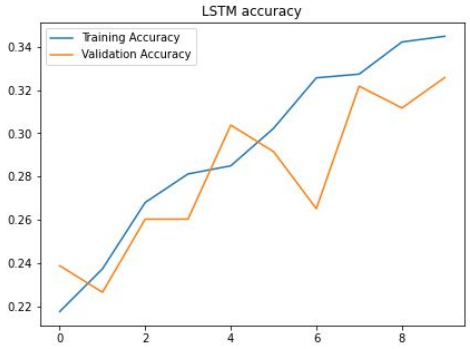
Same Conditions Neural Networks

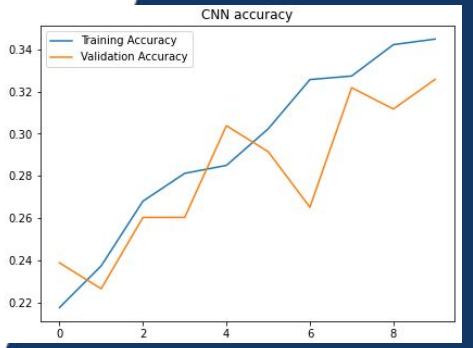
Hidden Layers = [256,256,256,64,64]
Learning rate = 0.001
Optimizer = Adam
Batch size = 25
Epochs = 10

	Accuracy	F1-Score	AUC	Time-To-Compute
ANN	53%	54%	74%	14.163 seconds
GRU	36%	41%	63%	405.860 seconds
LSTM	32%	37%	59%	462.704 seconds
CNN	42%	46%	69%	8.793 seconds









BerkeleyMICS

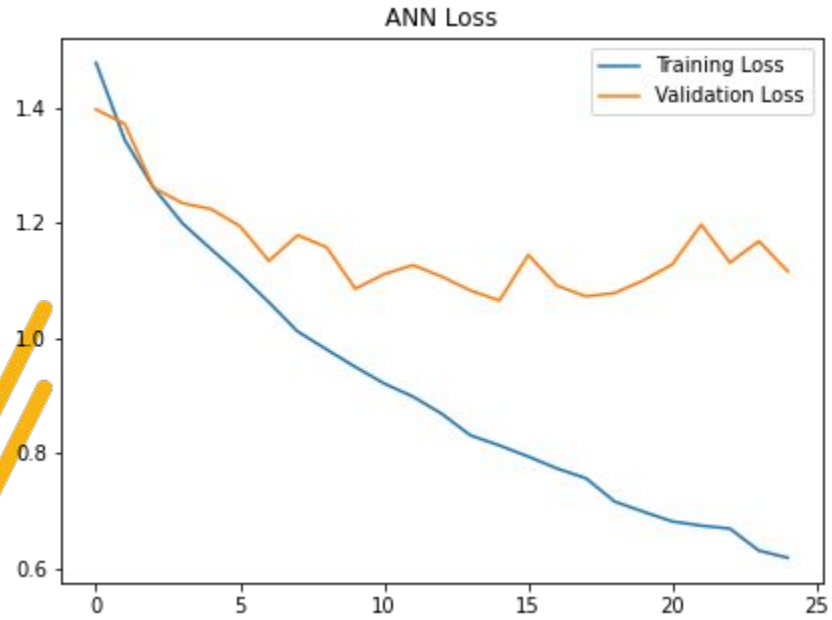
Comparison

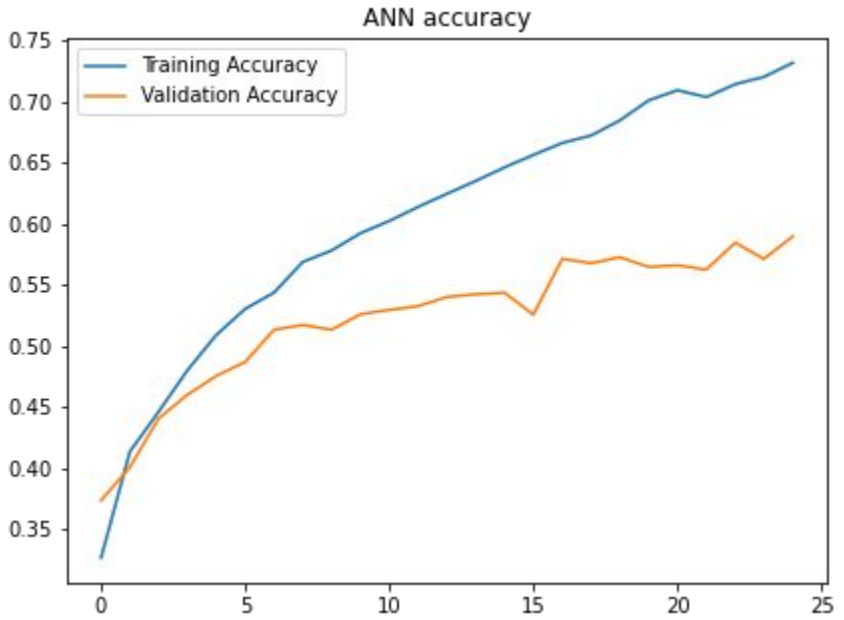
	Accuracy	F1-Score	AUC	Time-To-Compute
K-NN	47%	48%	74%	0.400 seconds
Decision Trees	65%	65%	78%	0.338 seconds
Logistic Regression	33%	32%	65%	0.508 seconds
Naive Bayes	27%	21%	60%	0.261 seconds
Random Forest	68%	68%	89%	2.923 seconds
ANN	53%	54%	74%	14.163 seconds
GRU	36%	41%	63%	405.860 seconds
LSTM	32%	37%	59%	462.704 seconds
CNN	42%	46%	69%	8.793 seconds

BerkeleyMICS

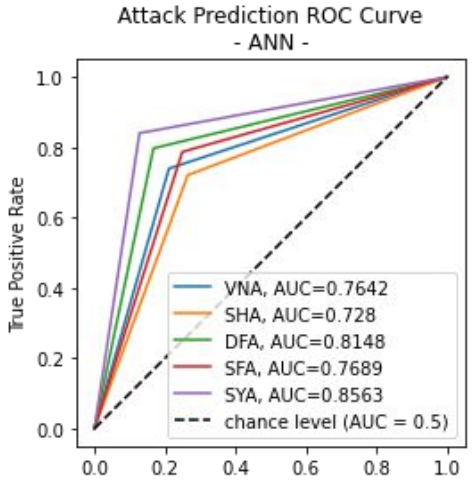
Artificial Neural Network (ANN)

Hidden Layers = [256,256,256,64,64]
Learning rate = 0.001
Optimizer = Adam
Batch size = 25
Epochs = 10



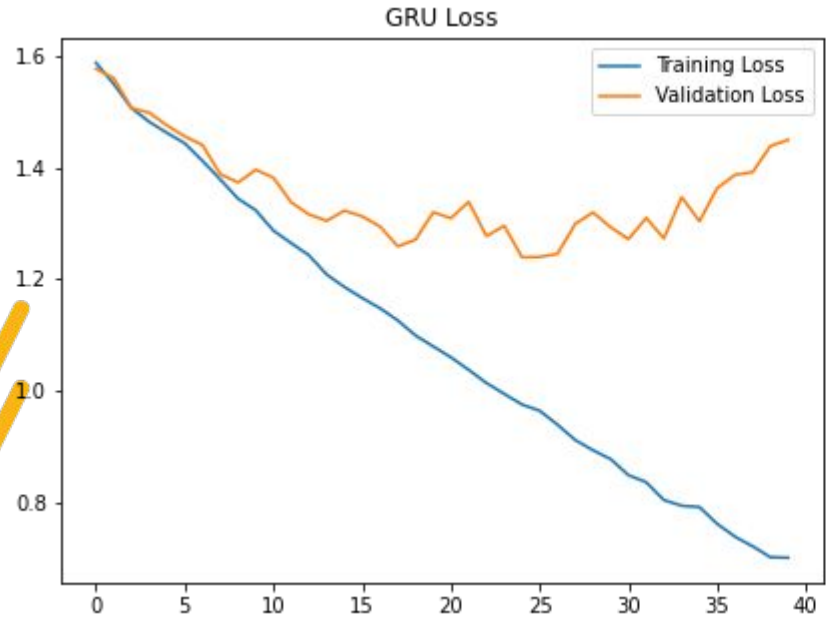


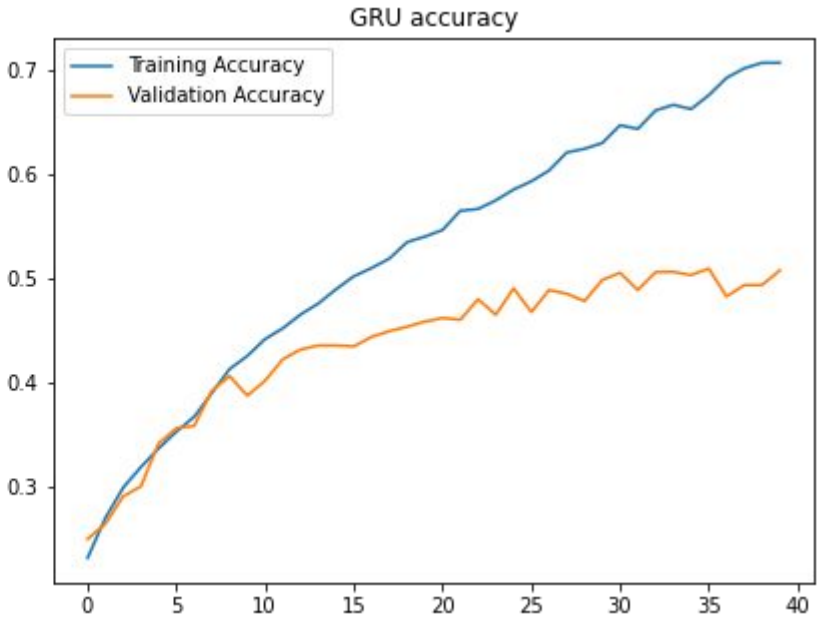
	precision	recall	f1-score	support
VNA	0.45	0.74	0.56	622
SHA	0.44	0.72	0.55	727
DFA	0.49	0.80	0.61	542
SFA	0.45	0.79	0.57	662
SYA	0.64	0.84	0.73	701
micro avg	0.49	0.78	0.60	3254
macro avg	0.49	0.78	0.60	3254
weighted avg	0.50	0.78	0.60	3254
samples avg	0.59	0.78	0.65	3254



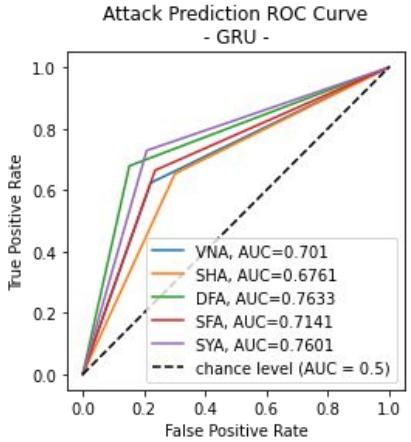
Gated Recurrent Unit (GRU)

Hidden Layers = [64,64]
Learning rate = 0.001
Optimizer = Adam
Batch size = 5
Epochs = 40



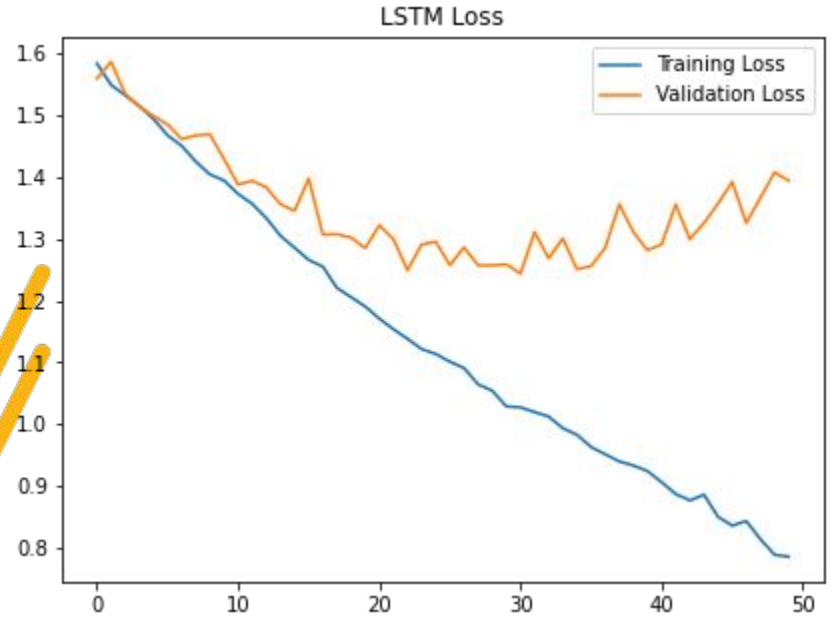


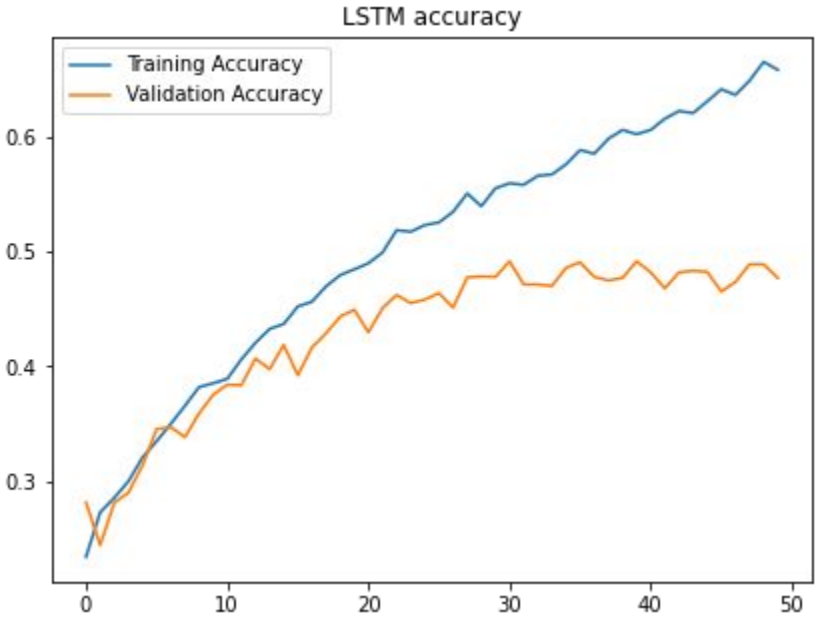
	precision	recall	f1-score	support
VNA	0.40	0.62	0.49	622
SHA	0.38	0.65	0.48	727
DFA	0.47	0.68	0.56	542
SFA	0.42	0.66	0.51	662
SYA	0.49	0.73	0.59	701
micro avg	0.43	0.67	0.52	3254
macro avg	0.43	0.67	0.53	3254
weighted avg	0.43	0.67	0.52	3254
samples avg	0.50	0.67	0.55	3254



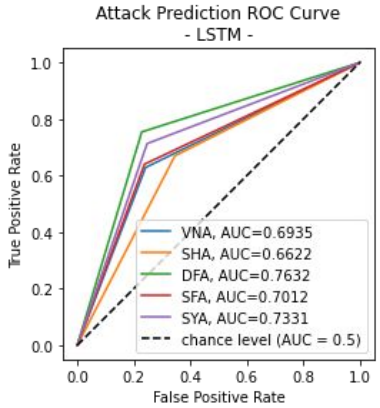
Long Short-Term Memory (LSTM)

Hidden Layers = [64,64]
Learning rate = 0.001
Optimizer = Adam
Batch size = 20
Epochs = 50
Dropout Rate = 0.2



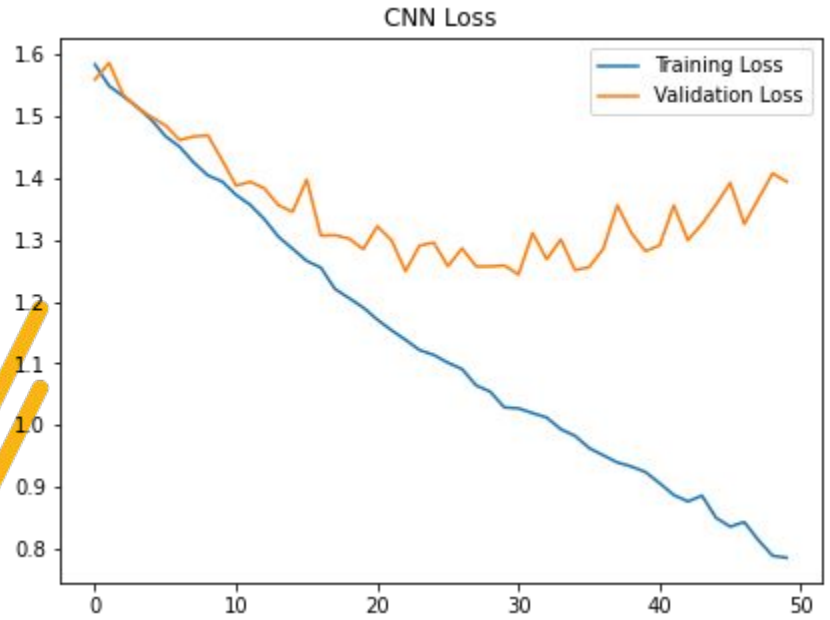


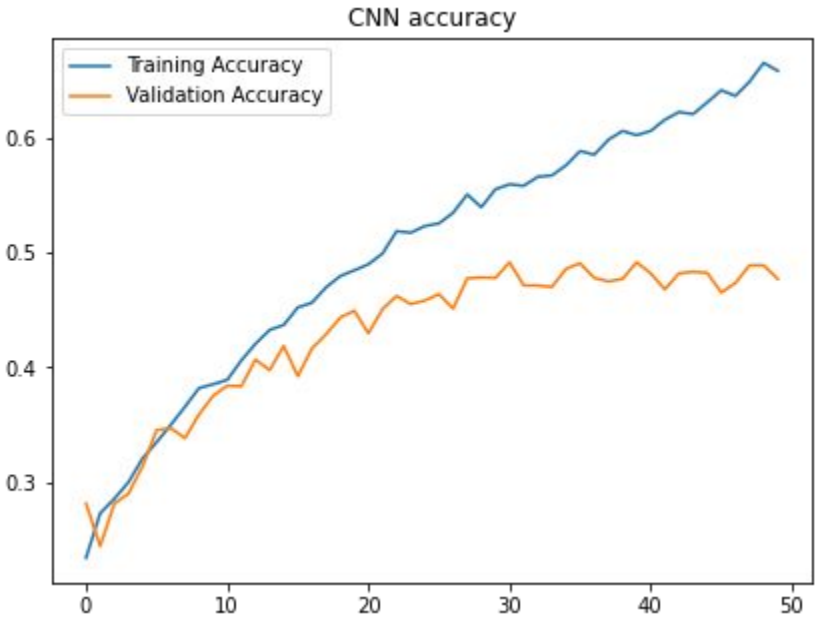
	precision	recall	f1-score	support
VNA	0.38	0.63	0.47	622
SHA	0.36	0.67	0.47	727
DFA	0.40	0.75	0.52	542
SFA	0.41	0.64	0.50	662
SYA	0.44	0.71	0.55	701
micro avg	0.40	0.68	0.50	3254
macro avg	0.40	0.68	0.50	3254
weighted avg	0.40	0.68	0.50	3254
samples avg	0.47	0.68	0.53	3254



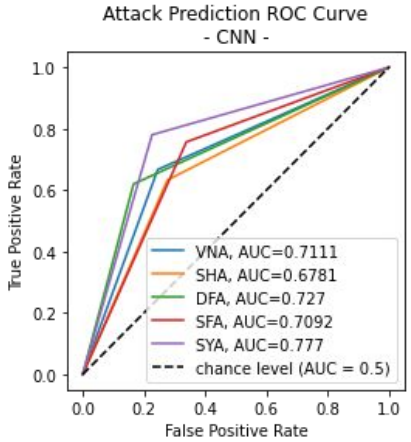
Convolutional Neural Network (CNN)

Hidden Layers = [64,64]
Learning rate = 0.001
Optimizer = Adam
Batch size = 5
Epochs = 50





	precision	recall	f1-score	support
VNA	0.39	0.67	0.49	622
SHA	0.40	0.63	0.49	727
DFA	0.43	0.62	0.51	542
SFA	0.36	0.76	0.49	662
SYA	0.49	0.78	0.60	701
micro avg	0.41	0.69	0.52	3254
macro avg	0.41	0.69	0.52	3254
weighted avg	0.41	0.69	0.52	3254
samples avg	0.49	0.69	0.55	3254



Comparison

	Accuracy	F1-Score	AUC	Time-To-Compute
ANN	61%	60%	79%	25.077 seconds
GRU	51%	52%	72%	831.946 seconds
LSTM	49%	50%	71%	318.898 seconds
CNN	50%	52%	72%	47.961 seconds

BerkeleyMICS

Fairness and Bias

- IoT Device
  - Sample size bias
  - Device similarities
- IoT Prevention

Limitations

- Lack of data
  - Similarities in labels/devices
- Computational Power
  - Ineffectiveness of certain models tested
- Time and Experience

BerkeleyMICS

Future Work

- Fine tuning clustering models
- Optimization of neural networks
- Hybrid models
- Expand datasets

BerkeleyMICS

Questions?

BerkeleyMICS