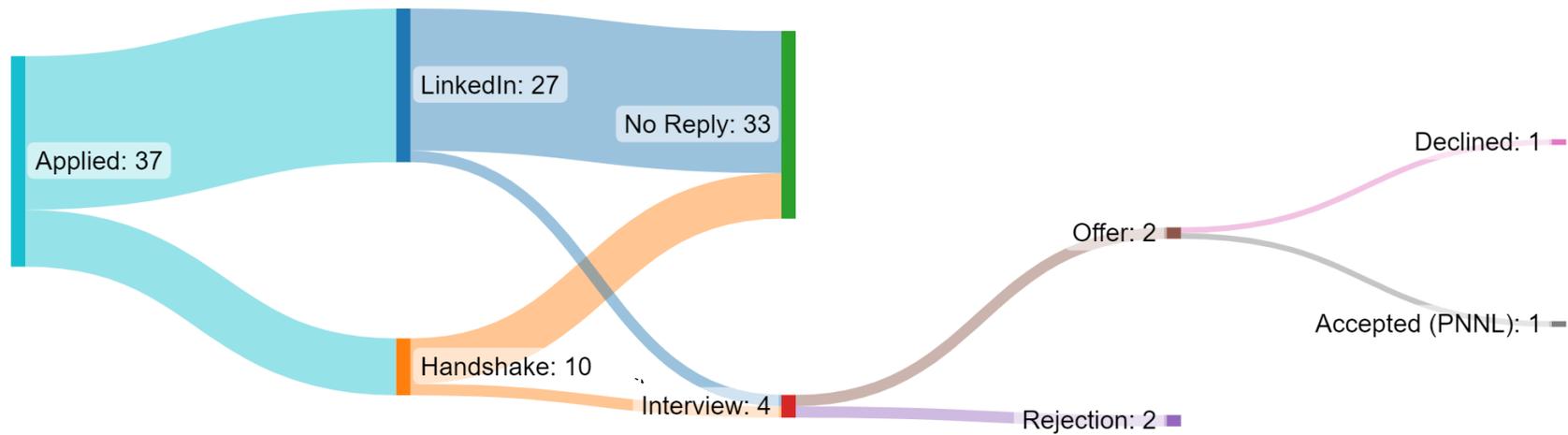


# INTERNSHIP REPORT

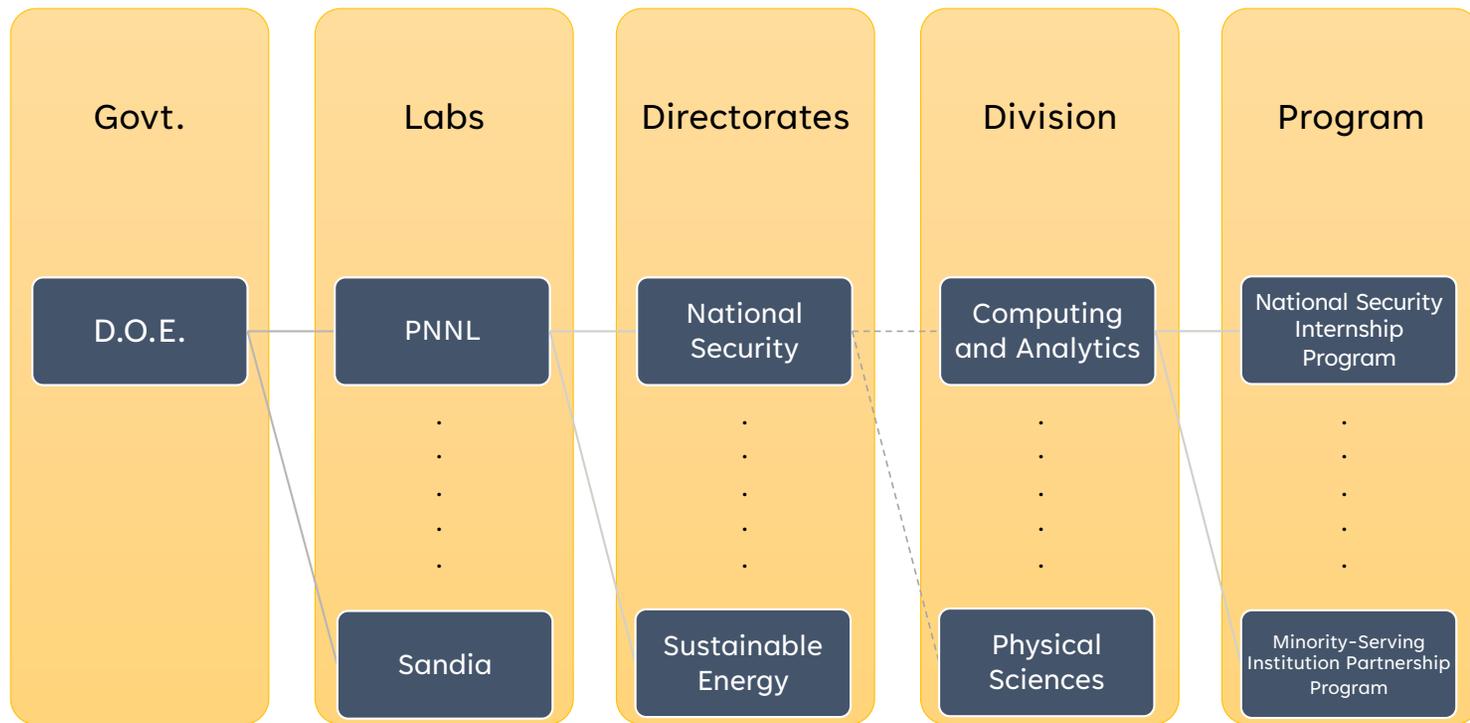
John Cooper

# INTERNSHIP SEARCH PROCESS



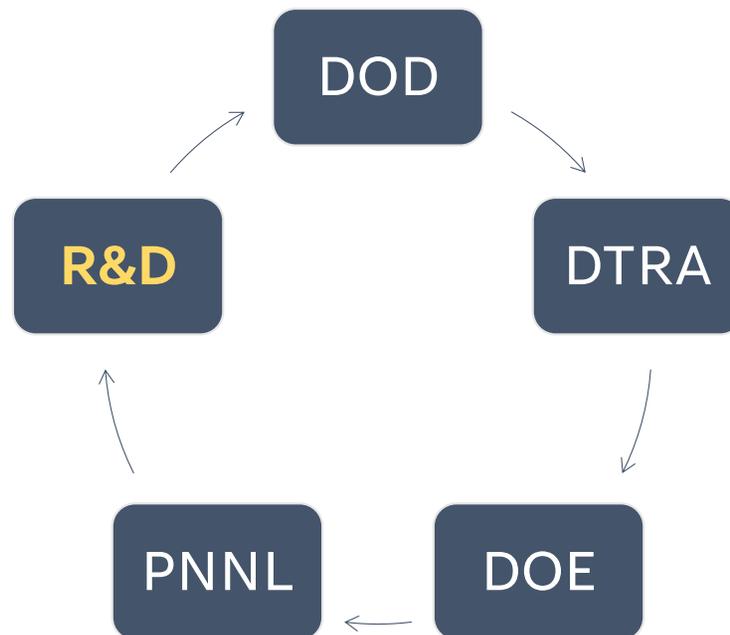
- PNNL listing was on Handshake
- All applications were cold-call resume submissions
- Advice:
  - Apply to organizations that you truly want to work for
  - Government organizations are in competition with industry
  - Network

# EMPLOYER ORGANIZATION



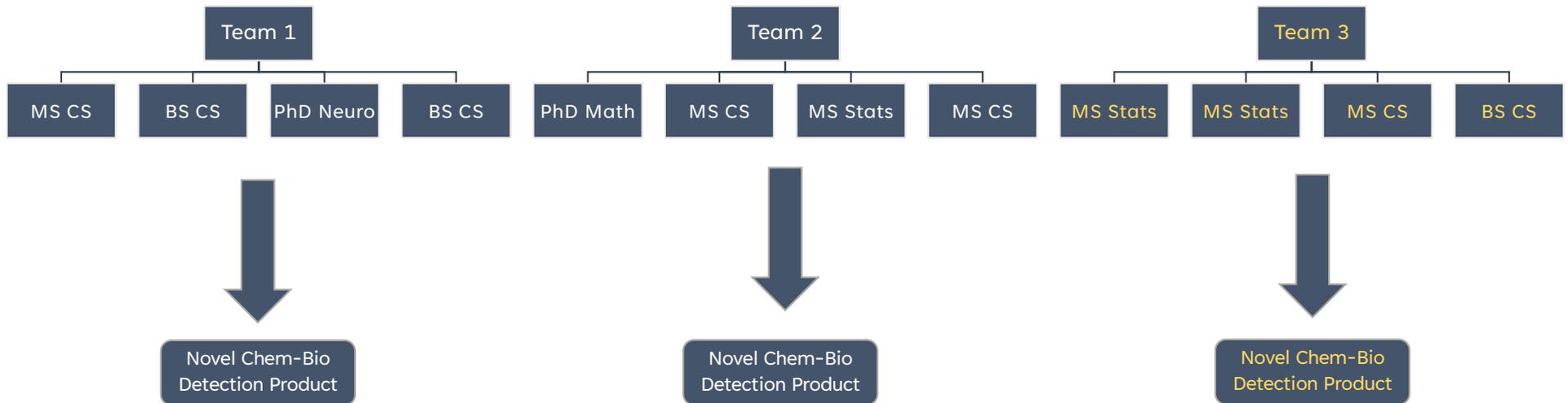
## EMPLOYER ORGANIZATION

- While Pacific Northwest National Laboratory (PNNL) is a DOE laboratory, it can also receive funding for projects from other government organizations such as the Department of Defense (DOD).
- The National Security Internship Program is funded by a subdivision of the DOD called the Defense Threat Reduction Agency (DTRA).



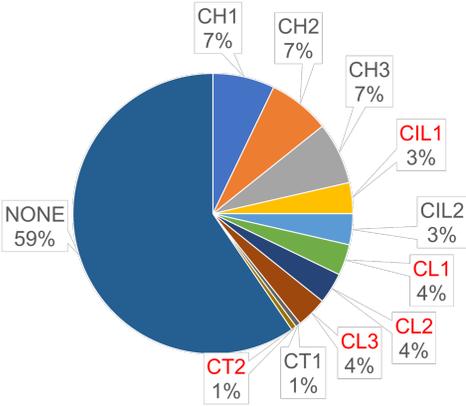
# EMPLOYER ORGANIZATION

- The internship consisted of an R&D competition between three teams, each comprised of a mix of data scientists and developers



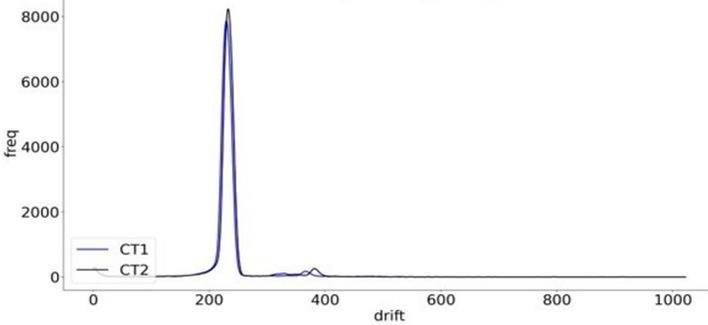
# PROJECT DESCRIPTION

### Imbalanced Data



Small Data (N = 1680)

### Dangerous Co-drift



# PROJECT DESCRIPTION

Small Data



Imbalanced Data



Chemical Co-drift



# PROJECT DESCRIPTION

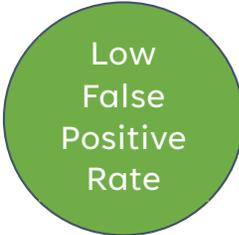
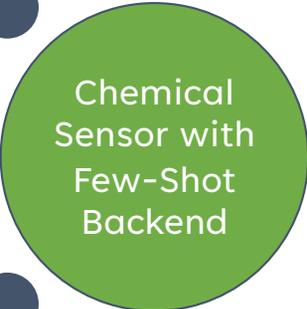
Small Data



Imbalanced Data

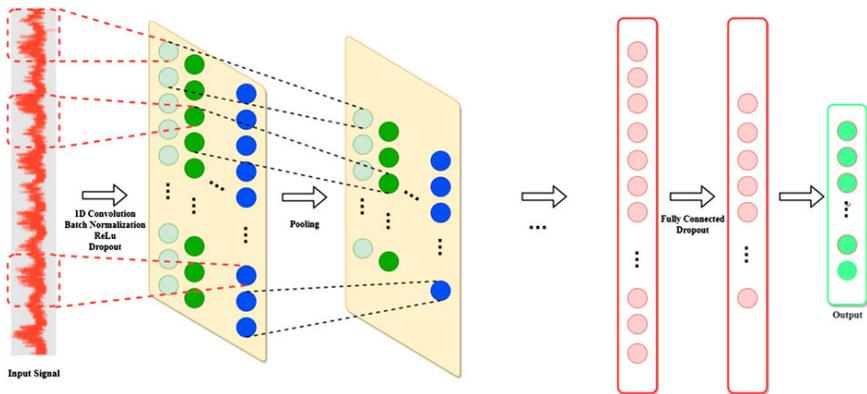


Chemical Co-drift



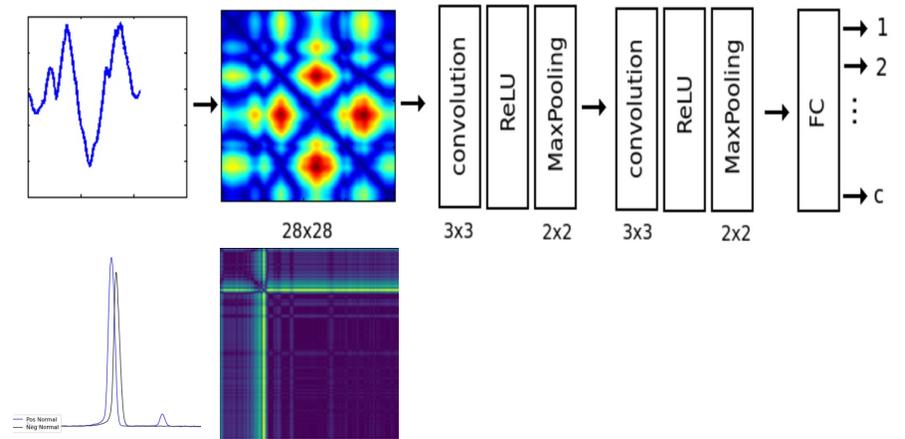
# TECHNIQUES

## Basic 1D CNN



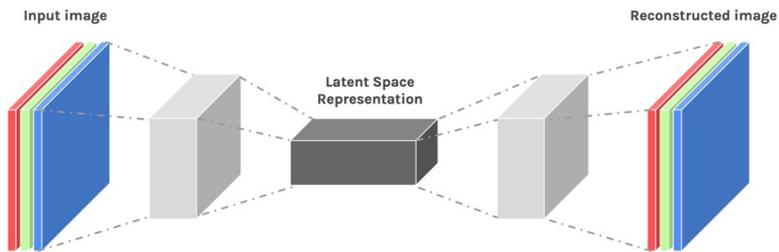
[This Photo](#) by Unknown Author is licensed under [CC BY](#)

## 2D CNN + Recurrence Plots

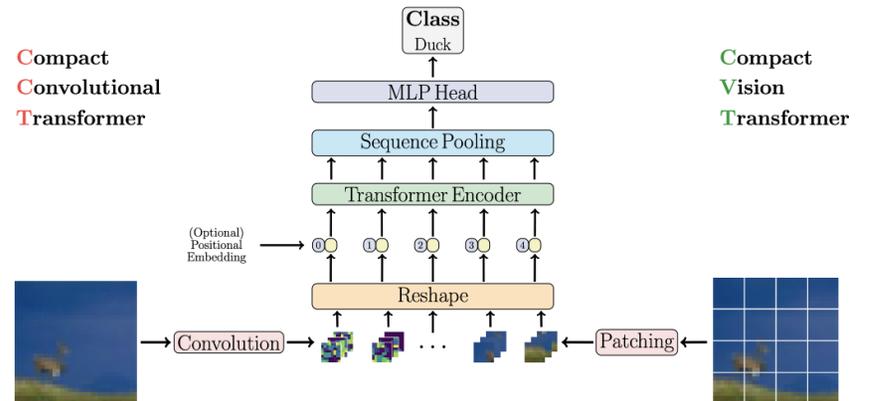


# TECHNIQUES

## Autoencoder for Anomaly Detection



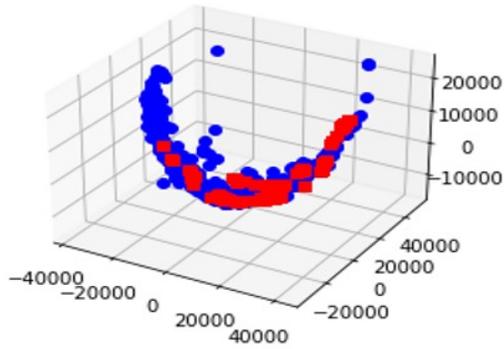
## Compact Convolutional Transformer



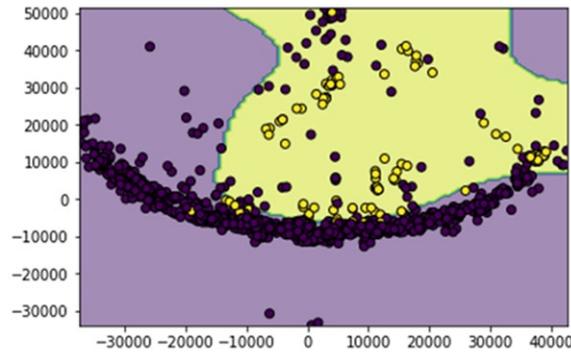
# TECHNIQUES



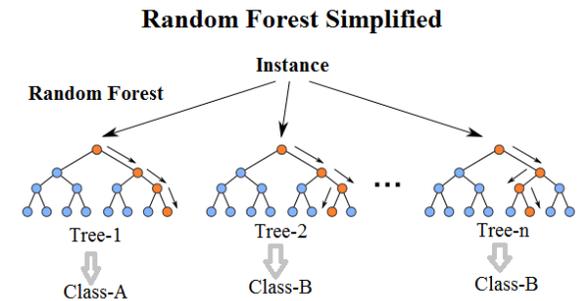
Principal Component Analysis



Support Vector Machines



Random Forests



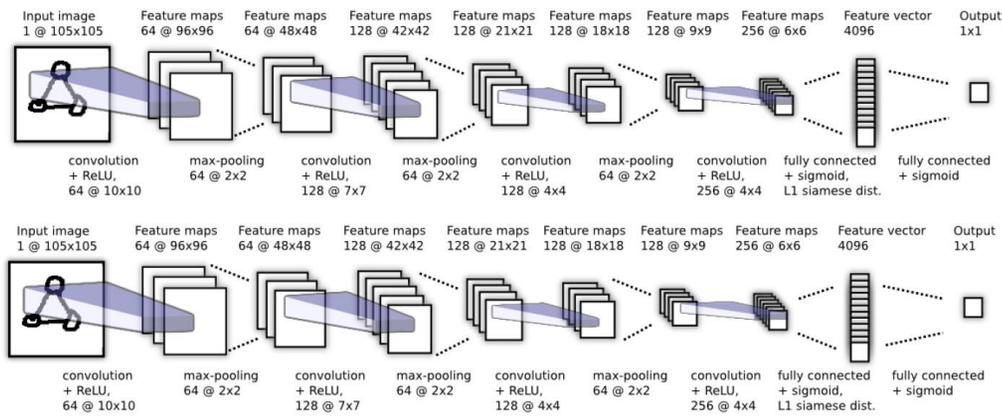
[This Photo](#) by Unknown Author is licensed under [CC BY](#).



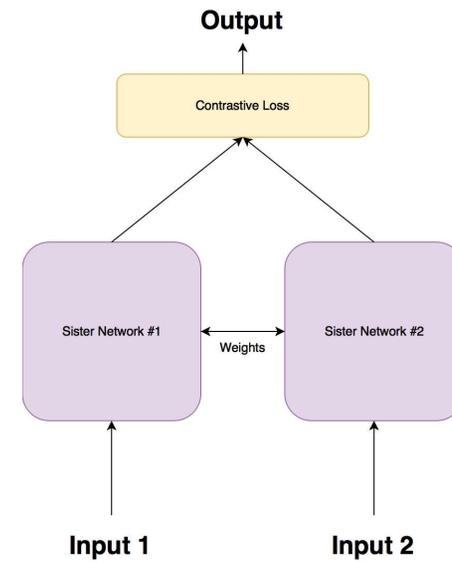
= knew before internship

# TECHNIQUES

## Twin Networks



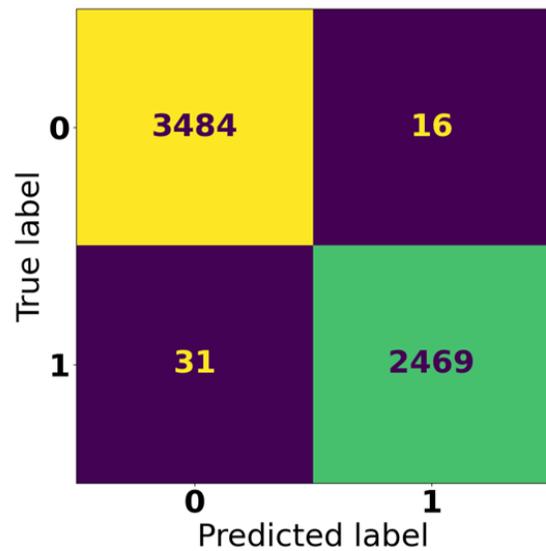
[This Photo](#) by Unknown Author is licensed under [CC BY-NC](#).



[This Photo](#) by Unknown Author is licensed under [CC BY-NC](#).

# TECHNIQUES

## Twin Networks



$$N_{\text{train}} = 1680$$

$$N_{\text{test}} = 6000$$

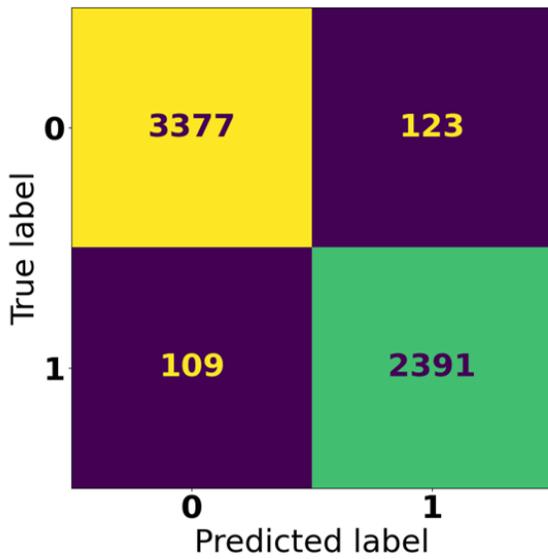
$$FP = 0.26\%$$

$$FN = 0.52\%$$

CH1	470	0	0	0	0	9	0	0	4	0	17
CH2	0	444	0	0	6	4	0	0	1	0	45
CH3	0	0	439	0	0	0	0	1	51	0	9
CIL1	0	0	0	493	0	0	0	7	0	0	0
CIL2	0	0	0	0	465	2	0	0	0	0	33
CL1	15	0	0	0	2	477	0	0	5	1	0
CL2	0	7	0	0	0	0	493	0	0	0	0
CL3	0	0	0	0	0	0	0	498	0	0	2
CT1	0	0	32	0	0	0	0	0	465	0	3
CT2	0	0	0	0	0	0	0	0	0	500	0
NONE	1	0	1	0	9	0	0	0	0	0	989
	CH1	CH2	CH3	CIL1	CIL2	CL1	CL2	CL3	CT1	CT2	NONE

# TECHNIQUES

## Twin Networks



$$N_{\text{train}} = 110$$

$$N_{\text{test}} = 6000$$

$$FP = 2.1\%$$

$$FN = 1.8\%$$

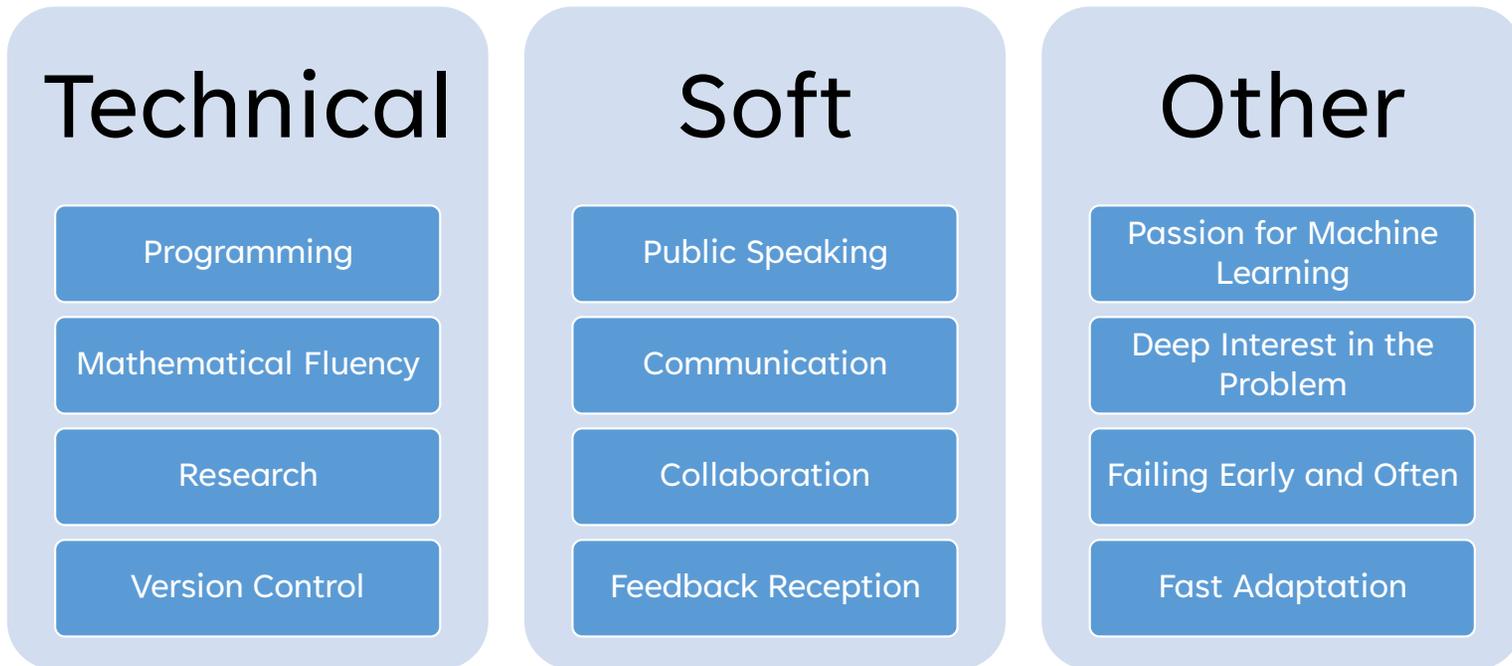
CH1	307	12	0	0	0	32	0	0	0	3	146
CH2	1	331	0	0	0	12	0	0	0	0	156
CH3	5	9	329	0	0	0	9	2	90	0	56
CIL1	0	4	0	474	0	1	0	1	0	0	20
CIL2	35	26	17	0	280	2	0	0	5	60	75
CL1	32	0	0	0	0	415	7	0	8	32	6
CL2	0	4	0	0	0	0	493	0	0	0	3
CL3	0	0	4	33	0	0	0	435	0	0	28
CT1	1	1	1	0	0	0	0	0	385	0	112
CT2	0	0	0	0	0	0	0	0	0	500	0
NONE	21	50	69	0	67	1	0	0	13	2	777
	CH1	CH2	CH3	CIL1	CIL2	CL1	CL2	CL3	CT1	CT2	NONE

## SOFTWARE

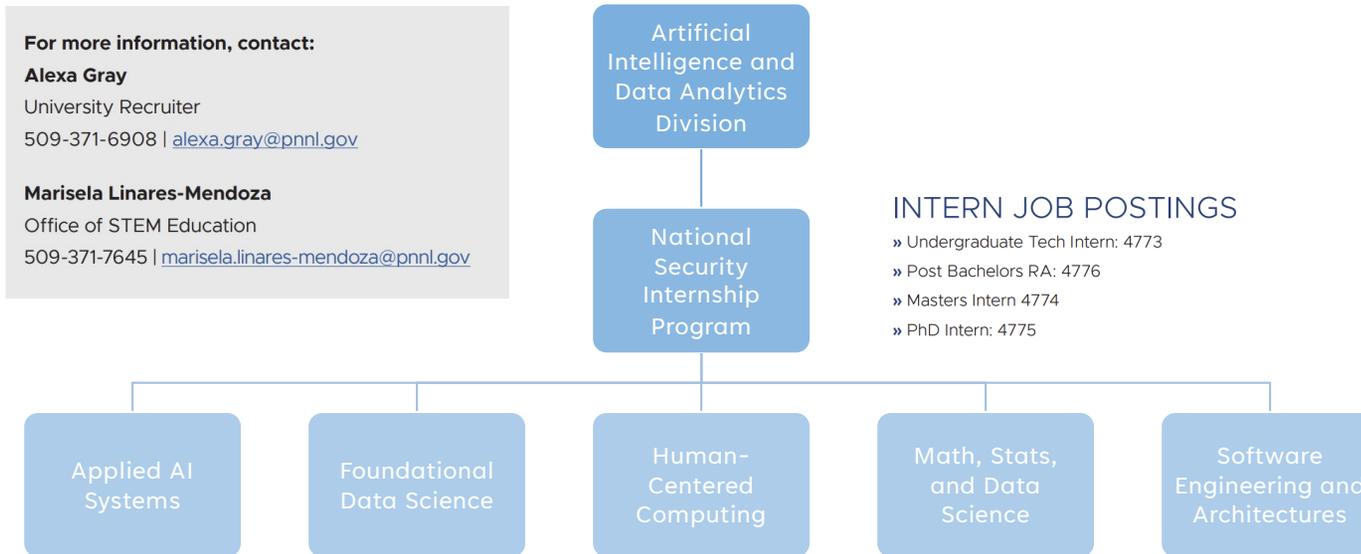


 = knew before internship

# SKILLS FOR SUCCESS



# OPPORTUNITIES AT PNNL



<https://www.pnnl.gov/national-security-internship>  
<https://careers.pnnl.gov/>

# REFERENCES

- [1] N. Shah A. Hassani, S. Walton. 2021. Escaping the Big Data Paradigm with Compact Transformers. <https://arxiv.org/abs/2104.05704>
- [2] S. Cho A. Jinwon. 2015. Variational Autoencoder Based Anomaly Detection Using Reconstruction Probability. <http://dm.snu.ac.kr/static/docs/TR/ SNUDM-TR-2015-03.pdf>
- [3] Defense Threat Reduction Agency. 2022. History. <https://www.dtra.mil/About/ DTRA-History/>
- [4] R. Salakhutdinov G. Koch, R. Zemel. 2015. Siamese Neural Networks for One-shot Image Recognition. <https://www.cs.cmu.edu/~rsalakhu/papers/oneshot1.pdf>
- [5] Y. Gavet J. Deboyle, N. Hatami. 2018. Classification of Time-Series Images Using Deep Convolutional Neural Networks. <https://arxiv.org/abs/1710.00886>
- [6] Pacific Northwest National Laboratory. 2022. Computing and Analytics Division. <https://www.pnnl.gov/computing-and-analytics-research>
- [7] Pacific Northwest National Laboratory. 2022. DOE Capabilities. <https://www.pnnl.gov/doe-capabilities>
- [8] Pacific Northwest National Laboratory. 2022. National Security Internship. <https://www.pnnl.gov/national-security-internship>
- [9] Pacific Northwest National Laboratory. 2022. Sponsors. <https://www.pnnl.gov/sponsors>
- [10] B. Ince B. Hauck K. O'Donnell R. Dereje C. Harden V. McHugh M. Wade P. Riley, S. Deshpande. 2021. Random Forest and Long Short-Term Memory Based Machine Learning Models for Classification of Ion Mobility Spectrometry Spectra. <https://spie.org/Publications/Proceedings/Paper/10.1117/12.2585829?SSO=1>
- [11] O. Abdelljaber O. Avci M. Gabbouj S. Kiranyaz, T. Ince. 2019. 1-D Convolutional Neural Networks for Signal Processing Applications. <https://ieeexplore.ieee.org/document/8682194>
- [12] B.C. Song S.H Lee. 2021. Vision Transformers for Small-Size Datasets. <https://arxiv.org/abs/2112.13492>
- [13] W. Bi N. Sebe B. Lepri M. De Nadai Y. Liu, E. Sangineto. 2021. Efficient Training of Visual Transformers with Small Datasets. <https://arxiv.org/abs/2106.03746>



THANK YOU

John Cooper