

How to reduce the data requirement

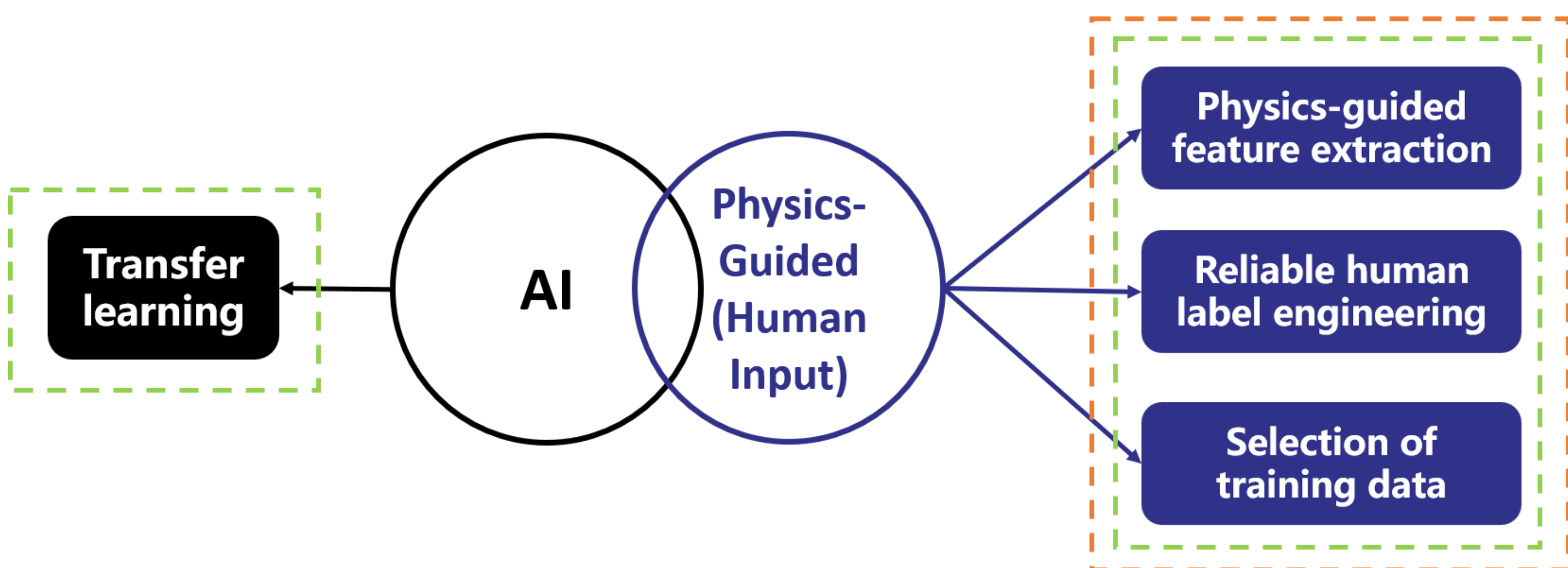
Performance of the target domain

$$\varepsilon_t(h) \leq \varepsilon_s(h) + d_{\mathcal{H}}(\mathcal{D}_s, \mathcal{D}_t) + \min\{\mathbb{E}_{\mathcal{D}_s}[l_d(f_s, f_t)], \mathbb{E}_{\mathcal{D}_t}[l_d(f_s, f_t)]\}$$

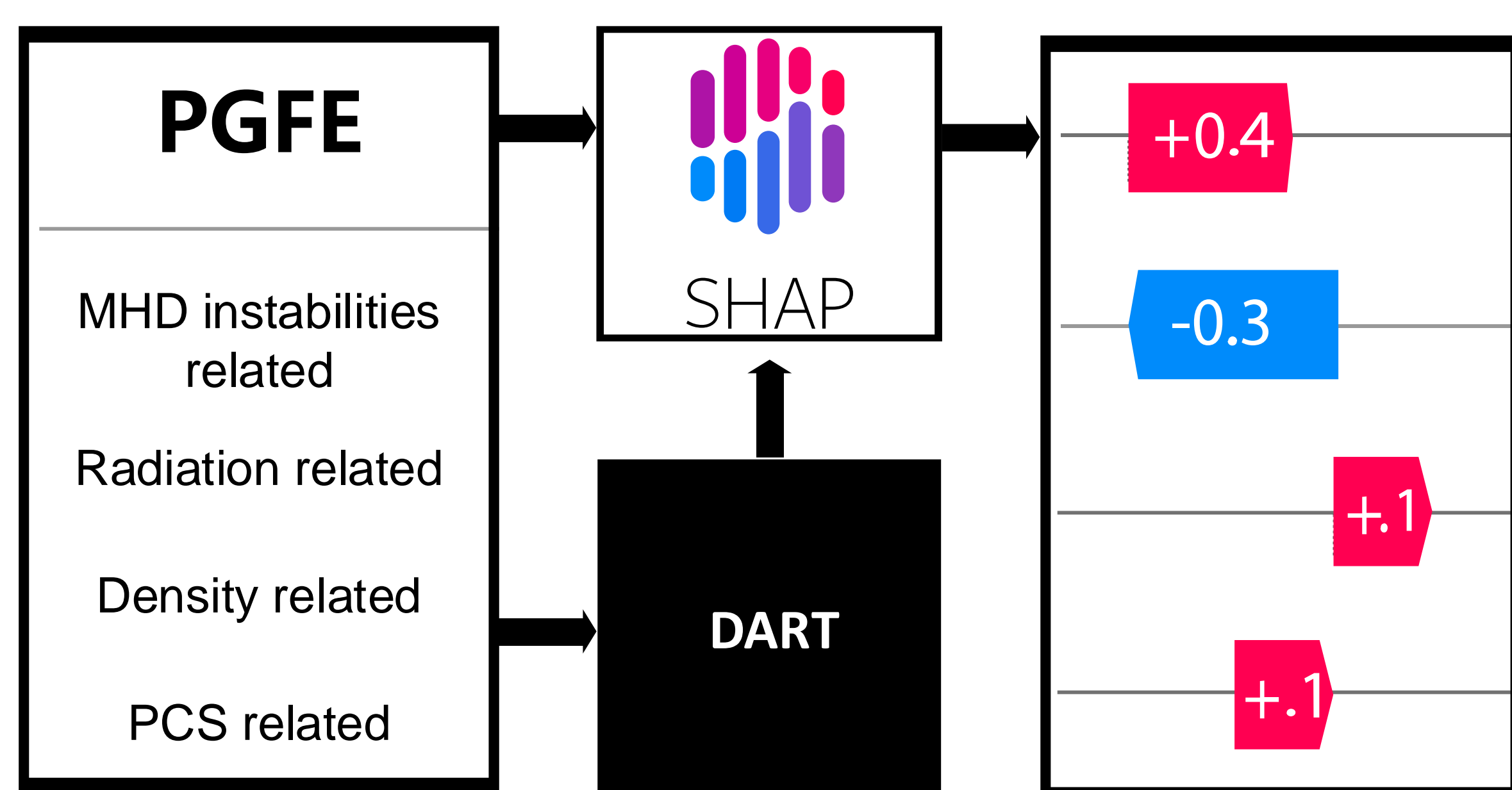
Reduce the distance of source and target domain

Improve the performance of source domain

Intersection of AI and physics

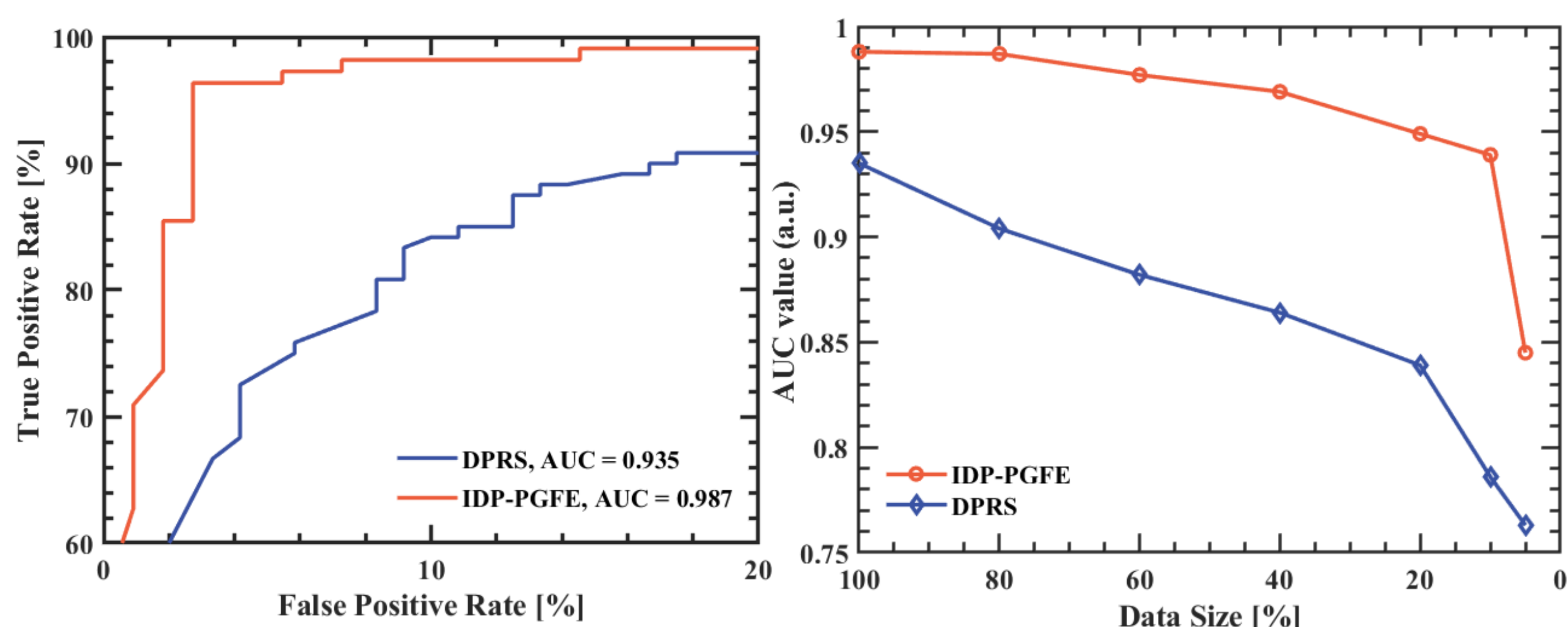


20 disruptive shots + 120 nondisruptive shots: IDP-PGFE



CS. Shen, W. Zheng, YH. Ding et al. Nuclear Fusion, 63 046024, 2023

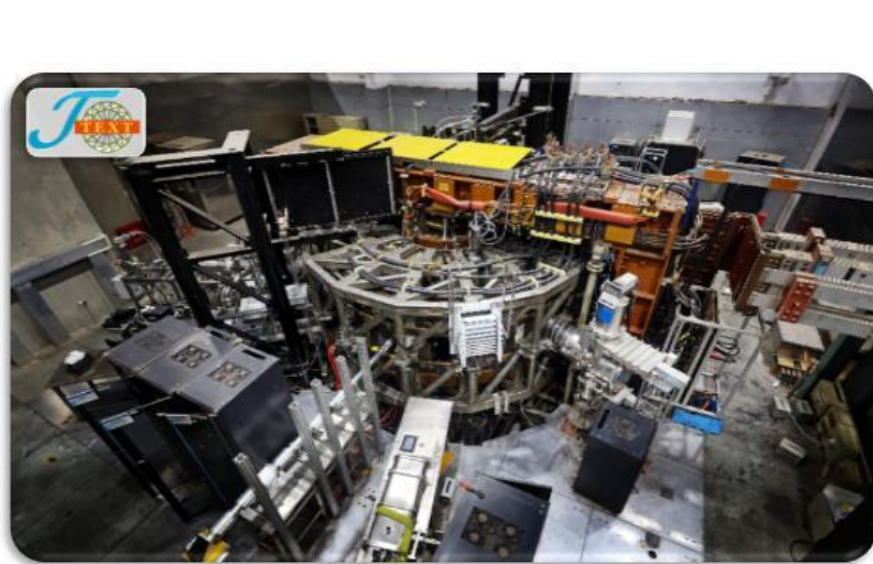
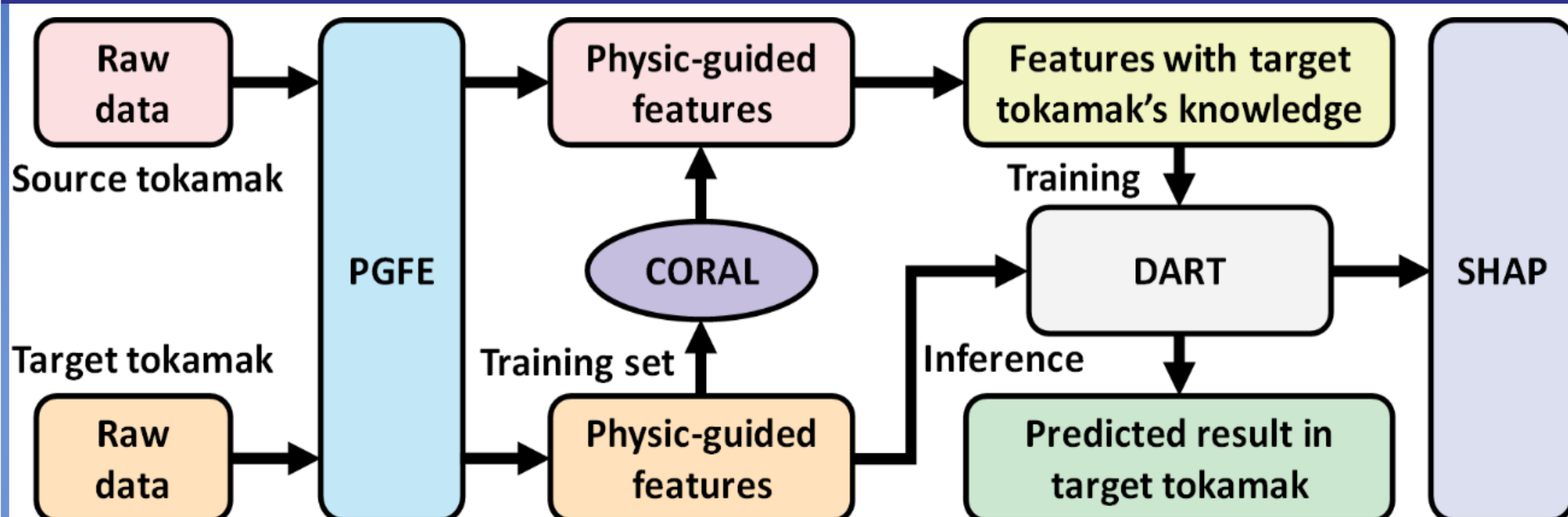
TPR~97.27%, FPR~5.45%, AUC ~0.98 > 20 dis and 120 non, AUC ~0.93



PGFE could Improve the performance of source domain

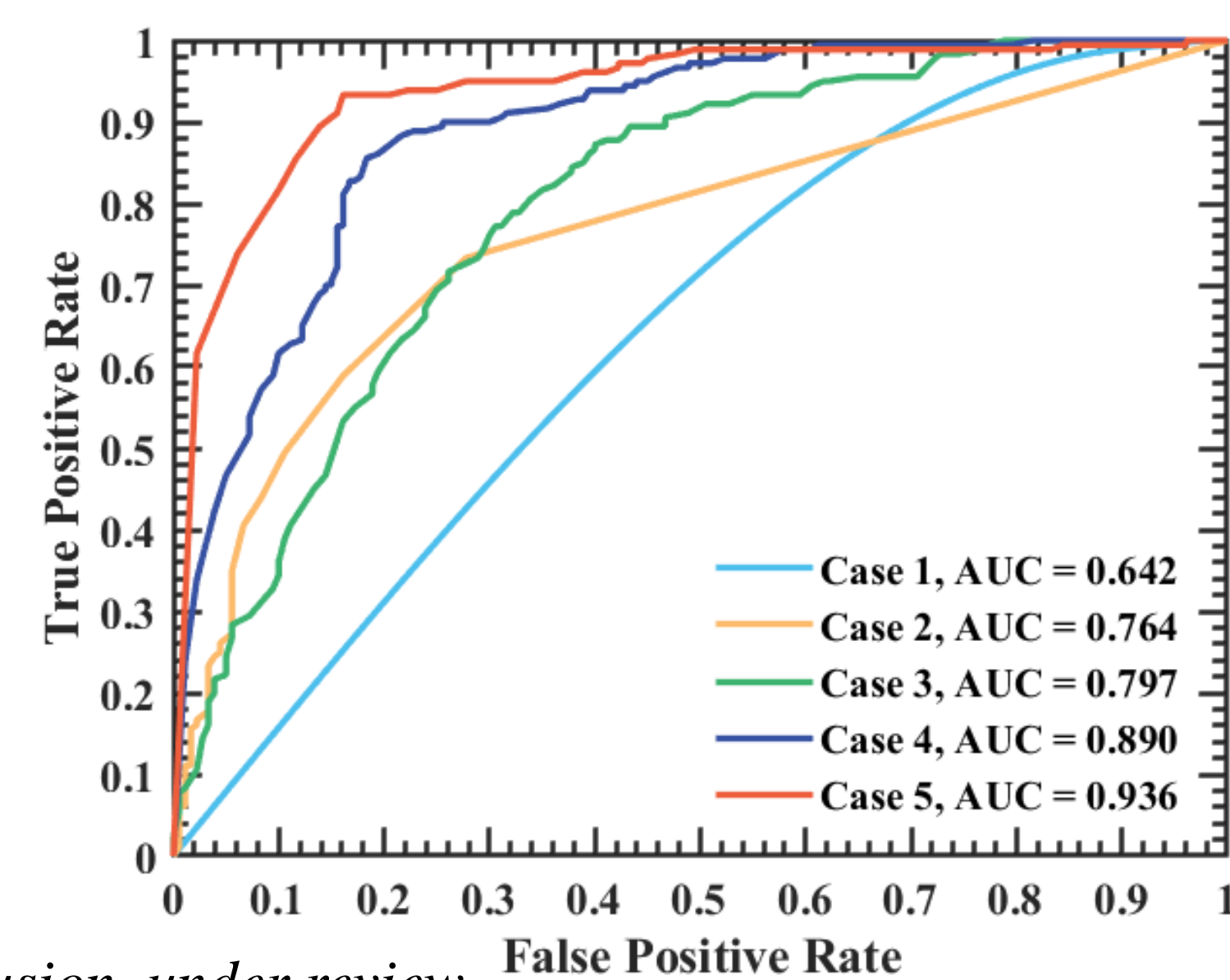
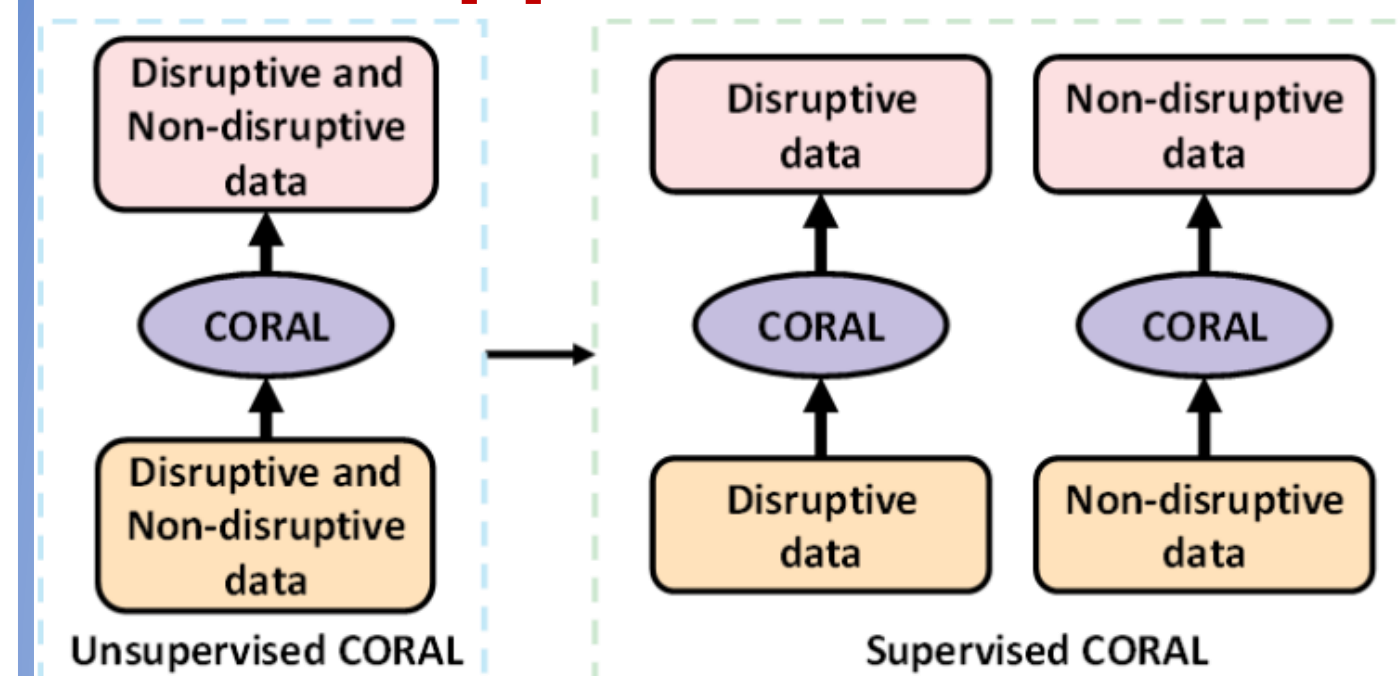
PGFE could Reduce the distance of source and target domain

10 disruptive shots + 100 nondisruptive shots: CORAL



Compare to J-TEXT, EAST just like the future tokamak.

Improved CORAL as an SDA approach from the UDA approach



CS. Shen, W. Zheng, BH. Guo et al. Nuclear Fusion, under review.

Case No.	EAST Data	Data Strategy	CORAL Strategy	AUC
1	None	/	/	0.642
2	110 (10)	Mixing	/	0.764
3	110 (10)	CORAL	Unsupervised	0.797
4	110 (10)	CORAL	Supervised	0.890
5	1896 (355)	Full data	/	0.936

TPR = 90 %, FPR = 25.56%, AUC = 0.89 > Supervised CORAL reaches the BEST
CORAL and PGFE could Reduce the distance of source and target domain

Zero-shot: more physics guided (human input)

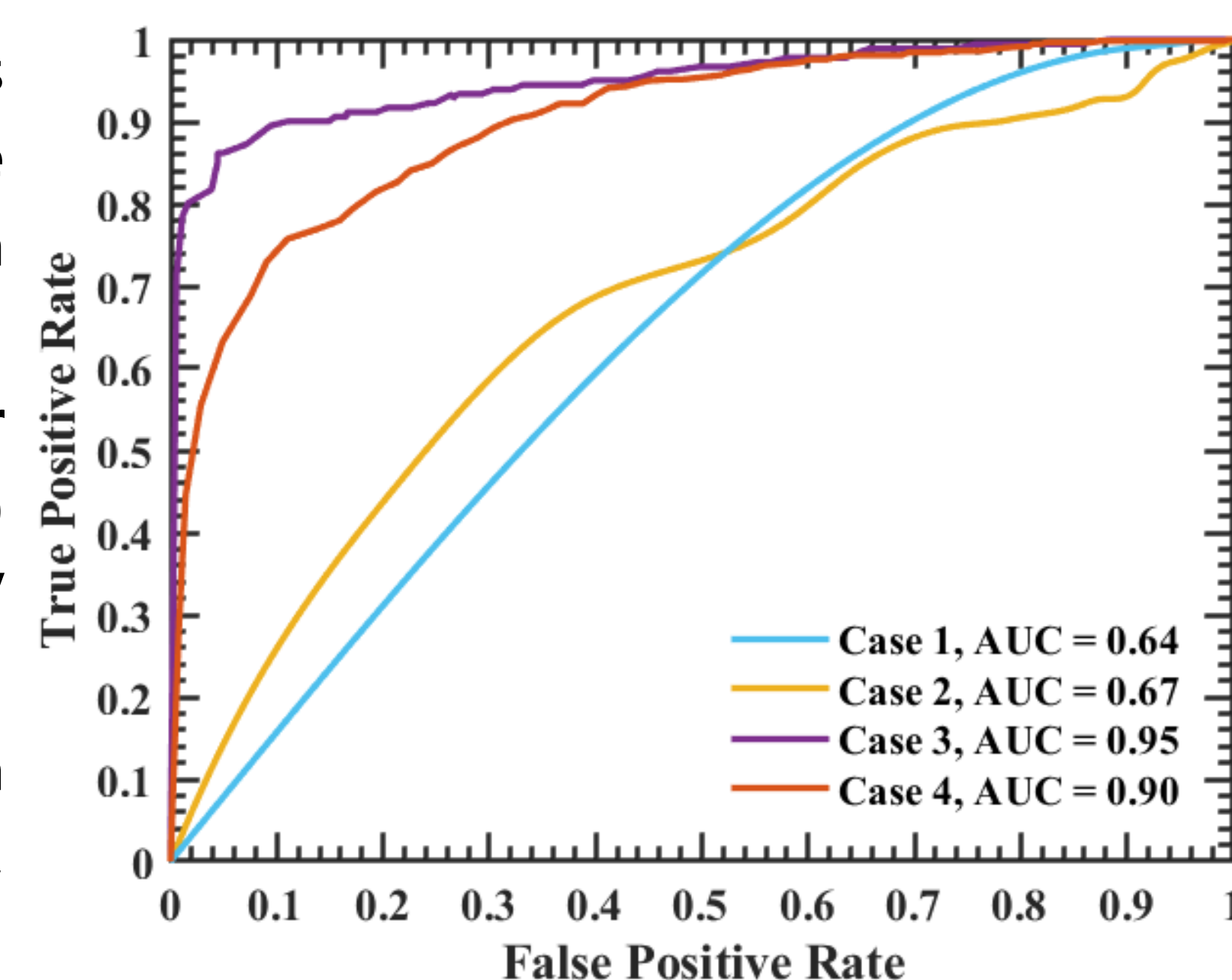
Improved data quality

More accurate true negative samples for J-TEXT have been labeled through the anomaly detection model.

XK. Ai, W. Zheng, CS. Shen et al. Nuclear Engineering and Technology, under review.

PGFE-Upgrade

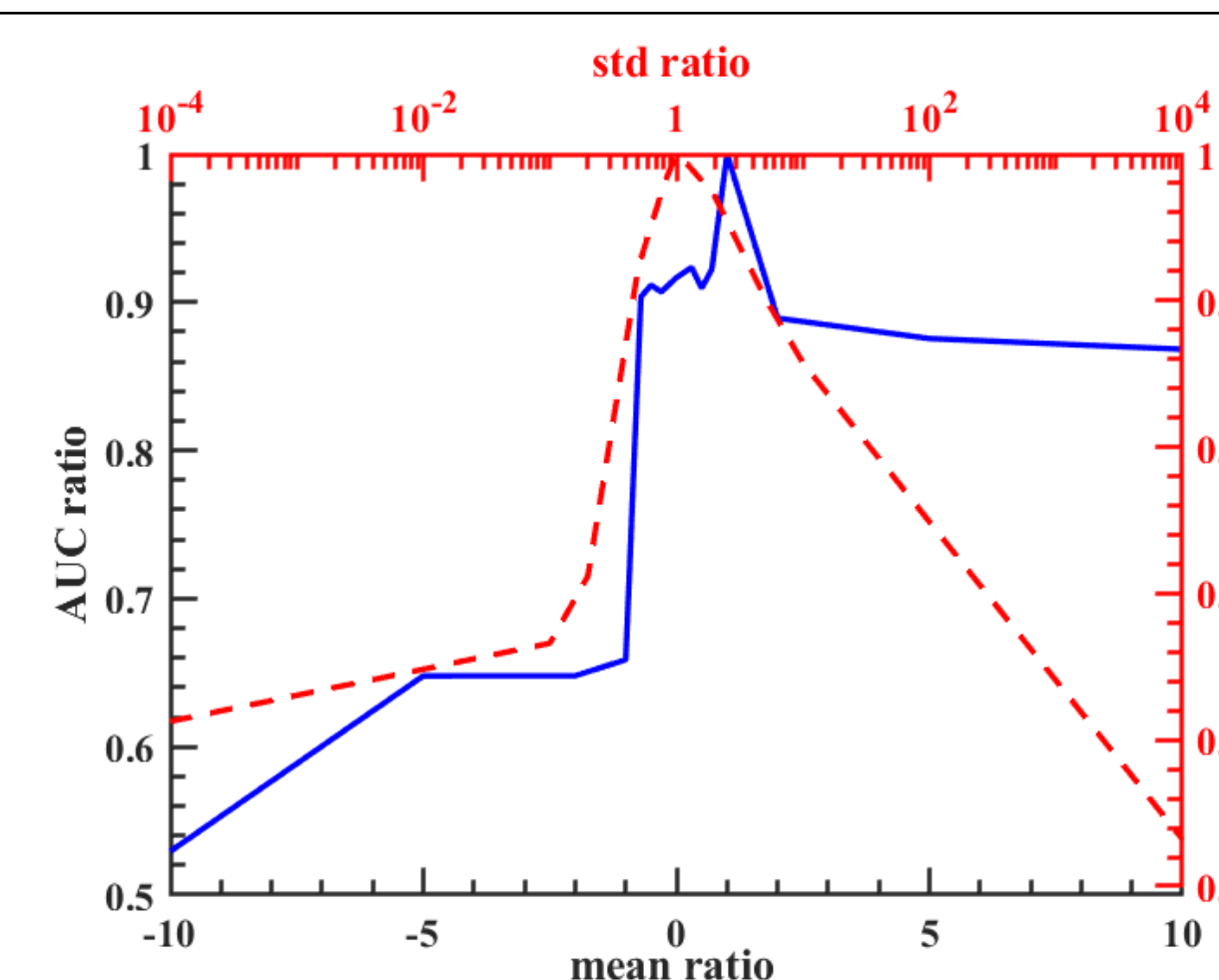
- Extract MHD instabilities related features from the whole Mirnov array based on SVD rather than FFT and CSD.
- Selected channels with similar chord core distances (r/a) between the SXR and XUV arrays in J-TEXT and EAST.
- Added scaled features such as XUV_ratio, ne0/nG, dr/a, dZ/a, etc.



Case No.	EAST Data	Transfer Strategy	Improved data quality	PGFE-U	AUC
1	None	/	N	N	0.64
2	None	/	Y	Y	0.67
3	1896 (355)	/	Y	Y	0.95
4	None	Estimated normalized parameters	Y	Y	0.90

Sensitivity analysis of the normalized parameters

- Performance declines by less than 90%
- Estimated mean ratio between -1~2.
- Estimated std ratio between 0.5~5.



Summary

- Improve the performance of source domain and reduce the distance of source and target domain are two main approaches to make the cross machine disruption prediction get higher performance and fewer data
- The cross machine performance could reach AUC=0.9 zero-shot performance by CORAL and physics guided approach.