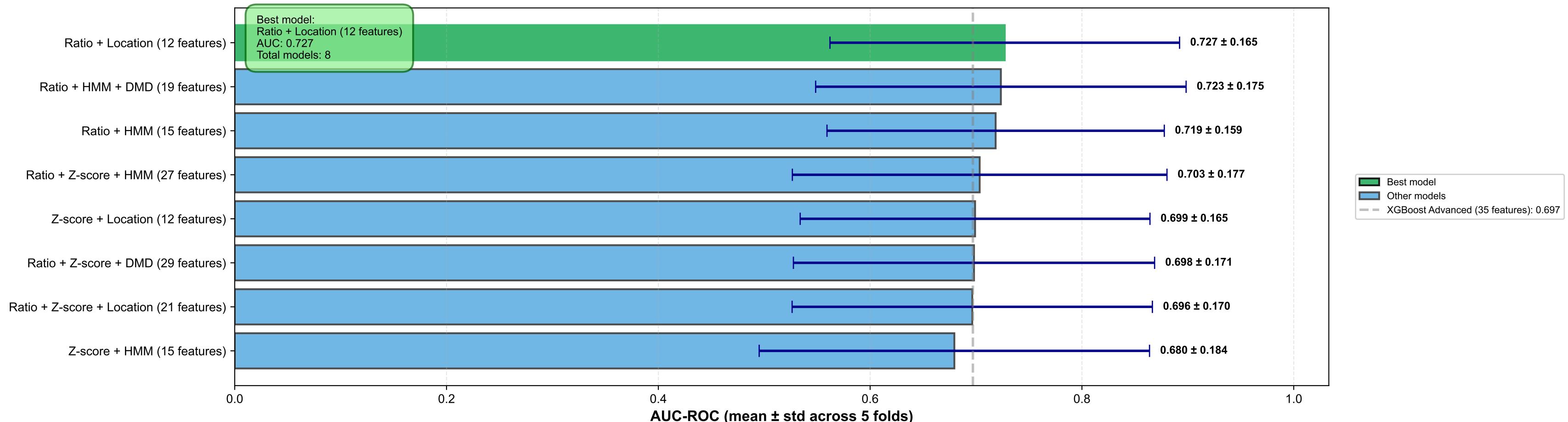


Ablation Study Performance with Z-Score Threshold Sensitivity Analysis

Ablation Study: Model Performance by Feature Set



B. Z-Score Threshold Sensitivity Analysis (h=8)

Threshold	Precision	Recall	F1-Score	TP	FP	FN	TN
1.0sigma	0.299	0.366	0.329	5,148	12,061	8,924	27,920
2.0sigma	0.229	0.110	0.148	1,545	5,194	12,527	34,787
3.0sigma	0.150	0.007	0.013	98	554	13,974	39,427

2sigma threshold optimal for h=8 horizon: Balances precision (0.229) and recall (0.110)
 Higher thresholds (3sigma) increase precision but sacrifice recall
 Lower thresholds (1sigma) increase recall but reduce precision