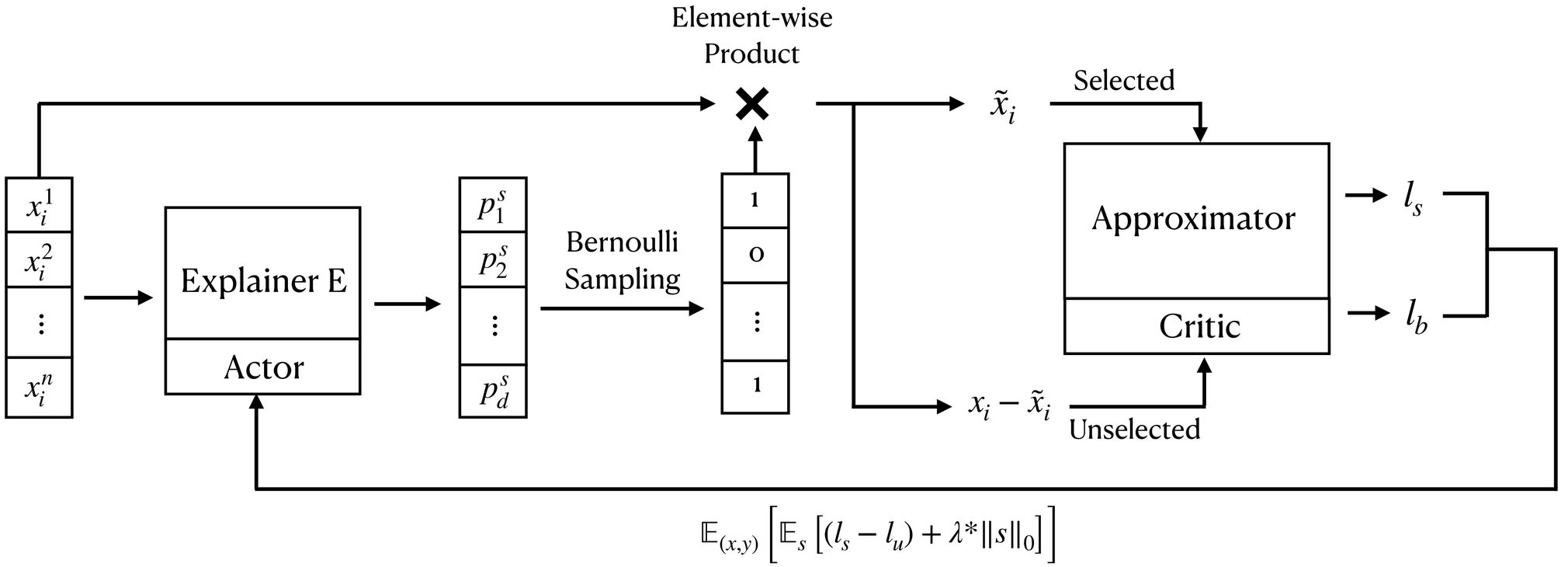
D-ACES Framework Proposal and Analysis

Proposal of D-ACES



- **Input to Actor:** Only feature values
- **Critic**: use both selected and unselected feature values to train
- **Loss function:** Direct Replacement Version Loss

Dual-Feature Actor-Critic
Explainable Feature Selection

(D-ACES)

D-ACES Framework Proposal and Analysis

Comparison with State-of-the-Art IFS Methods (Syn A)

Table 6.1: Comparison of Various *IFS* Methods: Best Results on Synthetic Dataset A

Methods	Mean(%)	Synethetic Datasets											
		A.1		A.2		A.3		A.4		A.5		A.6	
D-ACES	TPR	**100**		**100**		89.6		**96.7**		**83.1**		**73.2**	
	FDR	**0**		**0**		0		**13.5**		**18.9**		**7.5**	
INVASE	TPR	**100**	◇	**100**	◇	**100**	▲10.4	100	▲3.3	72.7	▼10.4	**72.0**	▼1.2
	FDR	**0**	◇	**0**	◇	**0**	◇	39.3	▲25.8	23.3	▲4.4	**4.6**	▼2.9
L2X	TPR	**100**	◇	**99.8**	▼0.2	80.5	▼9.6	82.7	▼14	75.9	▼7.2	78.3	▲5.1
	FDR	**0**	◇	**0.3**	▲0.3	19.5	▲19.5	20.2	▲6.7	26.5	▲7.6	21.8	▲14.3
LIME	TPR	39.0	▼61	**100**	◇	**100**	▲10.4	40.8	▼55.9	57.1	▼26	55.9	▼17.3
	FDR	61.0	▲61	**0**	◇	**0.1**	▲0.1	49.4	▲35.9	37.0	▲18.1	44.1	▲36.6
SHAP	TPR	60.6	▼39.4	88.0	▼12	94.0	▲4.4	64.5	▼32.2	66.9	▼16.2	65.2	▼8
	FDR	39.4	▲39.4	12.0	▲12	6.0	▲6	34.0	▲20.5	32.8	▲13.9	34.8	▲27.3