A Implementation Details

A.1 Datasets

Table 3: Dataset overview

	# Instances	# Features	Task
Adult	48 842	15	Binary classification
Churn Modeling	10 000	14	Binary classification

Data Details. ADULT contains demographic information such as age, education, and capital gain. The target variable is wether a person's income is >50k or ≤ 50. CHURN contains details of a bank's customers, the target variable is whether the customer left the bank or continues being a customer.

Data preprocessing. We normalize the input features of ADULT. For CHURN, we drop the feature columns "RowNumber", "CustomerId", and "Surname", encode the categorical columns and normalize the numerical ones to [0, 1] by min-max normalization.

A.2 Parameter settings

For all the models, we use the Optuna library [4] for hyperparameter tuning, the tuning is done on the validation set of each split (5 splits in total). For each experiment, the number of iterations is set to 100. Random seed = 0

MLP. We use a fixed architecture for both datasets [100, 64, 32]. We tune on the learning rate (*lr*) in the range of $[1e-2, 5e-3]$, dropout rate (*Dropout*) in the range of [0.0, 0.5] and Lambda (λ) in the range of $[1e-7, 1e-2]$ for L1 regularization.

XNNTab Follows the same setting details as MLP for the neural network component, the parameter settings for the sparse autoencoder are the $l1$ coefficient $\alpha = 1e-3$ for ADULT and CHURN.

Random Forest. We tuned on number of estimators (*n_estimators*) from the list [100, 200, 300], maximum depth (*max_depth*) in the range of [5, 10, 15, 20], the minimum samples split (*min_samples_split*) in the range of [2, 10], and the minimum samples leaf (*min_samples_leaf*) in the range of [1, 10].

XGBoost. We set the booster to "gbtree", and "early_stopping_rounds" to 12 and tuned on learning rate (*learning_rate*) in the range of [0.001, 0.4], maximum depth ($max/_d epth$) in the range of [3, 10], subsample in the range of [0.5, 1], and lambda (λ) in the range of [0.1, 10]

Decision Trees. We tuned on maximum depth (*max_depth*) in the range of [5, 10, 15, 20], the minimum samples split (*min_samples_split*) in the range of [2, 10], and the minimum samples leaf (*min_samples_leaf*) in the range of [1, 10].

Logistic Regression. We tuned on maximum of iterations (*max_iter*) in the range of [100, 200].

For our rule-based classifier:

Skope-rules. we set *precision_min* = 1, *recall_min* = 0.2, and the default settings for the rest of the parameters.

B Qualitative Results

B.1 ADULT

[4] https://optuna.org

Table 4: Dictionary features for the ADULT dataset. $|T_j|$ - size of the training subset that strongly activate feature j. Coverage of the rule reported as number of samples and percentage of samples. Table sorted by $|T_j|$.

j	|T_j|	Description	Coverage
40	15294	marital_status_Married is False and educational_num < 13 and capital_gain <= 8028.0	11625/0.76
9	11968	marital_status_Married is False and age <= 37.5 and educational_num < 12	7271/0.61
8	11634	marital_status_Married is False and age <= 34.5 and educational_num < 12	6517/0.56
54	10889	marital_status_Married is False and age <= 31.5 and educational_num < 13	5953/0.55
37	10755	marital_status_Married is False and age <= 31.5 and educational_num < 12	5735/0.53
49	9304	marital_status_Married is False and age <= 29.5 and educational_num < 11	4920/0.52
11	9030	marital_status_Married is False and age <= 26.5 and educational_num < 13	4387/0.49
31	8385	marital_status_Married is False and age <= 26.5 and educational_num < 11	4098/0.49
38	8313	marital_status_Married is False and age <= 25.5 and educational_num < 13	4053/0.48
44	7457	marital_status_Married is False and relationship_Not_in_family is False and age <= 25.5	3330/0.45
45	3975	marital_status_Married is False and age <= 22.5 and hours_per_week <= 32.5	1516/0.38
5	1758	age <= 20.5 and hours_per_week <= 24.5	754/0.43
34	1658	occupation_Other_service is False and capital_gain > 14682.0	547/0.33
26	1358	marital_status_Widowedand is False and capital_gain > 14682.0	546/0.40
50	982	age <= 72.0 and capital_gain > 15022.0	532/0.55
63	978	gender is Female and age <= 18.5 and hours_per_week <= 24.5	236/0.24
12	521	occupation_Handlers_cleaners is False and educational_num > 10 and capital_gain > 15022.0	404/0.79
61	480	capital_gain > 70654.5	152/0.32
7	290	occupation_Farming_fishing is False and capital_gain > 19266.0 and hours_per_week > 27.5	209/0.73
18	226	relationship_Other_relative is False and capital_gain > 26532.0	185/0.82
22	225	occupation_Farming_fishing is False and capital_gain > 26532.0 and hours_per_week > 18.0	183/0.82
3	212	capital_gain > 34569.0	153/0.73
58	154	capital_gain > 37702.5	153/0.99
42	153	capital_gain > 37702.5	153/1.00
36	152	capital_gain > 70654.5	152/1.00
4	152	capital_gain > 70654.5	152/1.00
33	152	capital_gain > 67047.0	152/1.00
53	152	capital_gain > 70654.5	152/1.00
21	152	capital_gain > 70654.5	152/1.00
60	152	capital_gain > 70654.5	152/1.00
16	152	capital_gain > 70654.5	152/1.00
13	152	capital_gain > 70654.5	152/1.00
64	152	capital_gain > 70654.5	152/1.00

B.2 CHURN

Table 5: Dictionary features for the CHURN dataset. $|T_j|$ - size of the training subset that strongly activate feature j. Coverage of the rule reported as number of samples and percentage of samples. Table sorted by $|T_j|$.

j	$\lvert T_j \rvert$	**Description**	**Coverage**
6	4955	`Age < 45 and NumOfProducts < 3`	4560 / 0.91
47	4631	`Age ≤ 55.5 and NumOfProducts ≥ 2`	2898 / 0.62
2	4628	`CreditScore > 379.0 and Age < 40`	3561 / 0.76
3	4457	`CreditScore > 409.0 and NumOfProducts ≥ 2`	2939 /0.66
30	4419	`Age < 45 and Balance > 76177.5`	2594 /0.57
43	3966	`Age ≤ 42`	2238 / 0.62
9	2999	`Age < 46.5 and NumOfProducts ≥ 2`	2533 / 0.84
12	2574	`Age < 35 and NumOfProducts ≥ 2`	1193 / 0.46
18	1805	`Geography_Spain == False and Age ≥ 45`	1081 / 0.59
36	425	`Age > 59.5 and IsActiveMember == 1`	255 / 0.60
7	424	`NumOfProducts ≥ 3`	196 / 0.46
24	363	`CreditScore > 361.5 and NumOfProducts ≥ 3`	211 / 0.57
1	350	`CreditScore > 379.0 and NumOfProducts ≥ 3`	205 / 0.59
38	303	`Age > 64.0 and IsActiveMember == 1`	172 / 0.57
28	270	`Balance ≤ 244642.5 and NumOfProducts ≥ 3`	215 /0.79
40	226	`NumOfProducts ≥ 3`	196 / 0.87
11	218	`Age > 62.5 and IsActiveMember == True`	164 /0.74
35	150	`Age ≥ 38 and Balance ≥ 75720`	98 /0.65
18	146	`Geography_Germany == True and IsActiveMember == False`	54 /0.29
4	127	`Age > 61.5 and NumOfProducts < 2 and IsActiveMember == True`	101 / 0.79
16	112	`Age > 66.0 and NumOfProducts < 2 and IsActiveMember == True`	60 / 0.55
5	53	`Age > 77`	12 / 0.23