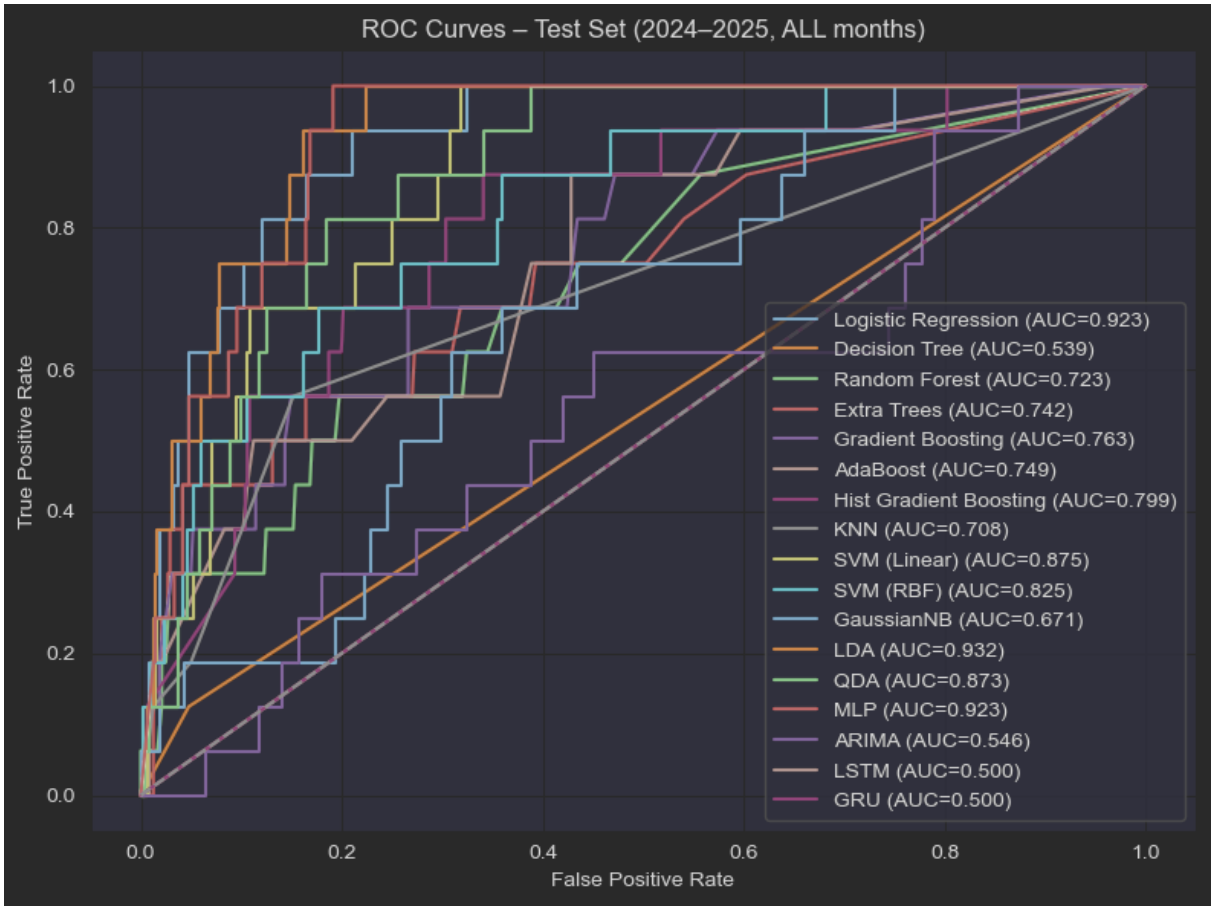
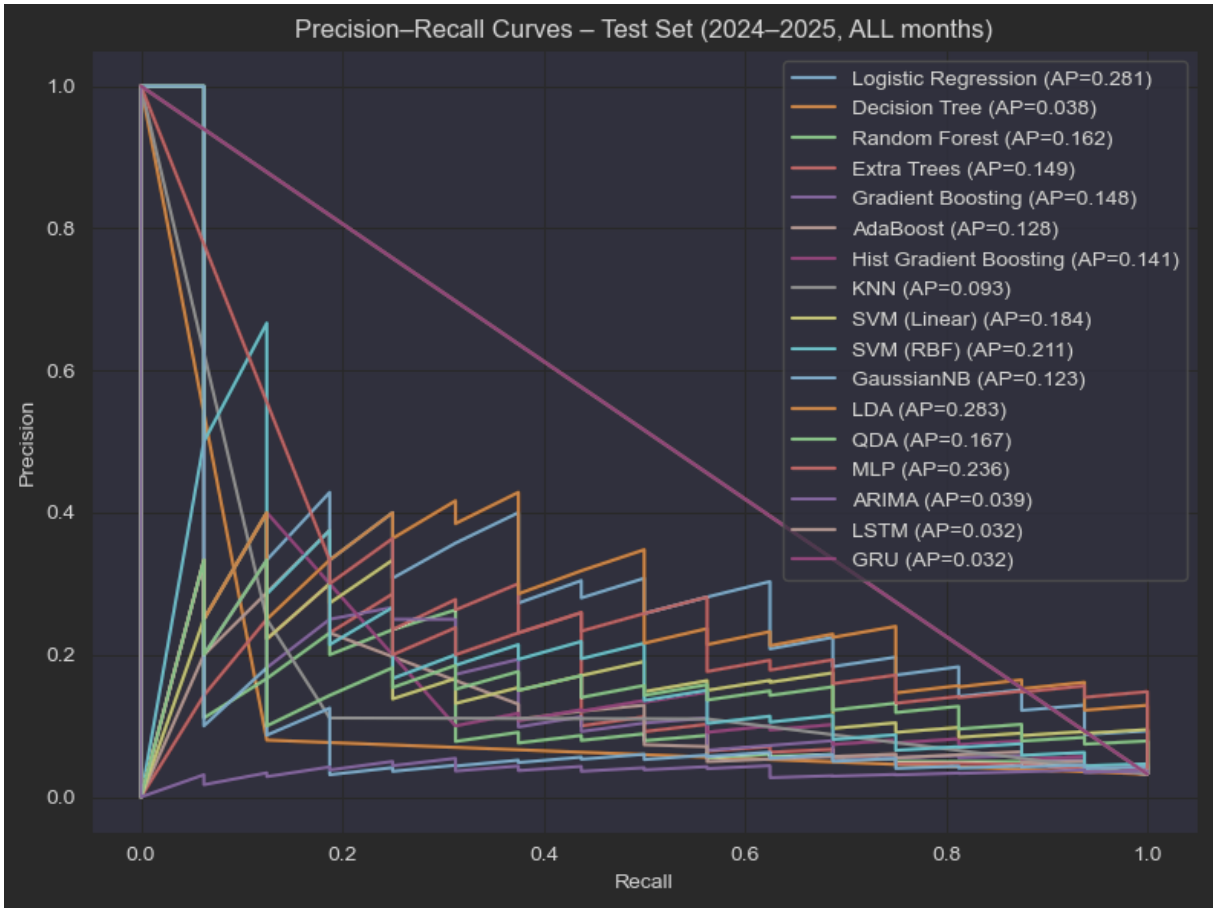
Below are raw outputs of supervised learning models on ECA

	Model	precision_at_4	recall_at_4
0	Logistic Regression	0.5000	0.5000
3	Extra Trees	0.5000	0.5000
4	Gradient Boosting	0.4375	0.4375
9	SVM (RBF)	0.4375	0.4375
2	Random Forest	0.3750	0.3750
7	KNN	0.3750	0.3750
13	MLP	0.3750	0.3750
11	LDA	0.3750	0.3750
8	SVM (Linear)	0.3750	0.3750
10	GaussianNB	0.3125	0.3125
5	AdaBoost	0.2500	0.2500
12	QDA	0.1875	0.1875
6	Hist Gradient Boosting	0.1875	0.1875
1	Decision Tree	0.1250	0.1250
15	LSTM	0.1250	0.1250
16	GRU	0.1250	0.1250
14	ARIMA	0.0625	0.0625

	model	Precision	Accuracy	Recall	AUC	F1score	WEI	WEI_share
0	Logistic Regression	0.4000	0.9536	0.1250	0.9382	0.1905	0.0643	1.0
1	Decision Tree	0.1053	0.9153	0.1250	0.5382	0.1143	0.0847	1.0
2	Random Forest	0.3125	0.9399	0.3125	0.7376	0.3125	0.0904	1.0
3	Extra Trees	0.8750	0.9727	0.4375	0.7618	0.5833	0.0749	1.0
4	Gradient Boosting	0.3000	0.9344	0.3750	0.7876	0.3333	0.0746	1.0
5	AdaBoost	0.3529	0.9426	0.3750	0.7786	0.3636	0.3500	1.0
6	Hist Gradient Boosting	0.0000	0.9563	0.0000	0.8292	0.0000	0.0618	1.0
7	KNN	1.0000	0.9617	0.1250	0.7294	0.2222	0.0617	1.0
8	SVM (Linear)	0.0000	0.9508	0.0000	0.8793	0.0000	0.0680	1.0
9	SVM (RBF)	0.6667	0.9590	0.1250	0.8468	0.2105	0.0821	1.0
10	GaussianNB	0.1034	0.8934	0.1875	0.6839	0.1333	0.1747	1.0
11	LDA	0.3333	0.9481	0.1875	0.9362	0.2400	0.0630	1.0
12	QDA	0.1429	0.9153	0.1875	0.8614	0.1622	0.0971	1.0
13	MLP	0.3333	0.9454	0.2500	0.9261	0.2857	0.0537	1.0
14	ARIMA	0.0559	0.5219	0.6250	0.5488	0.1026	0.4877	1.0
15	LSTM	0.0000	0.9563	0.0000	0.5000	0.0000	0.0832	1.0
16	GRU	0.0000	0.9563	0.0000	0.5000	0.0000	0.0740	1.0


ROC Curves – Test Set (2024–2025, ALL months)
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
Logistic Regression (AUC=0.923)
Decision Tree (AUC=0.539)
Random Forest (AUC=0.723)
Extra Trees (AUC=0.742)
Gradient Boosting (AUC=0.763)
AdaBoost (AUC=0.749)
Hist Gradient Boosting (AUC=0.799)
KNN (AUC=0.708)
SVM (Linear) (AUC=0.875)
SVM (RBF) (AUC=0.825)
GaussianNB (AUC=0.671)
LDA (AUC=0.932)
QDA (AUC=0.873)
MLP (AUC=0.923)
ARIMA (AUC=0.546)
LSTM (AUC=0.500)
GRU (AUC=0.500)


Precision–Recall Curves – Test Set (2024–2025, ALL months)
Precision
Recall
0.0
0.2
0.4
0.6
0.8
1.0
Logistic Regression (AP=0.281)
Decision Tree (AP=0.038)
Random Forest (AP=0.162)
Extra Trees (AP=0.149)
Gradient Boosting (AP=0.148)
AdaBoost (AP=0.128)
Hist Gradient Boosting (AP=0.141)
KNN (AP=0.093)
SVM (Linear) (AP=0.184)
SVM (RBF) (AP=0.211)
GaussianNB (AP=0.123)
LDA (AP=0.283)
QDA (AP=0.167)
MLP (AP=0.236)
ARIMA (AP=0.039)
LSTM (AP=0.032)
GRU (AP=0.032)

From the results and plots, we can see that the extra tree model can perform best in predicting. So, we choose to evaluate that model.

2024 results — Extra Trees

4 rows 4 rows × 4 cols

	Extra Trees_prob	actual_peak	is_CP	timestamp
901	0.64250	22485.78	1	2024-06-20
900	0.53500	23851.82	1	2024-06-19
887	0.46875	18764.47	0	2024-06-06
904	0.37500	19076.84	0	2024-06-23

Hits: 2/4 | True CP in month: 4 | Precision@4: 50.00%
Recall@4: 50.00%

4 rows 4 rows × 4 cols

	Extra Trees_prob	actual_peak	is_CP	timestamp
942	0.80250	23355.86	1	2024-07-31
941	0.71500	22371.08	1	2024-07-30
940	0.61875	22244.97	1	2024-07-29
926	0.58625	21637.19	0	2024-07-15

Hits: 3/4 | True CP in month: 4 | Precision@4: 75.00%
Recall@4: 75.00%

4 rows 4 rows × 4 cols

	Extra Trees_prob	actual_peak	is_CP	timestamp
944	0.84125	22103.13	1	2024-08-02
943	0.79500	23179.24	1	2024-08-01
947	0.49375	20269.79	0	2024-08-05
945	0.40875	21899.00	0	2024-08-03

Hits: 2/4 | True CP in month: 4 | Precision@4: 50.00%
Recall@4: 50.00%

4 rows 4 rows × 4 cols

	Extra Trees_prob	actual_peak	is_CP	timestamp
992	0.20500	20886.45	1	2024-09-19
994	0.12250	18923.97	0	2024-09-21
979	0.06625	18005.31	0	2024-09-06
983	0.05250	17733.50	0	2024-09-10

Hits: 1/4 | True CP in month: 4 | Precision@4: 25.00%
Recall@4: 25.00%

For 2024’s results, the model can achieve an average precision of 50% and an average recall of 50%. While in the previous algorithm, it cannot be evaluated on ECA, so almost every metric is 0.

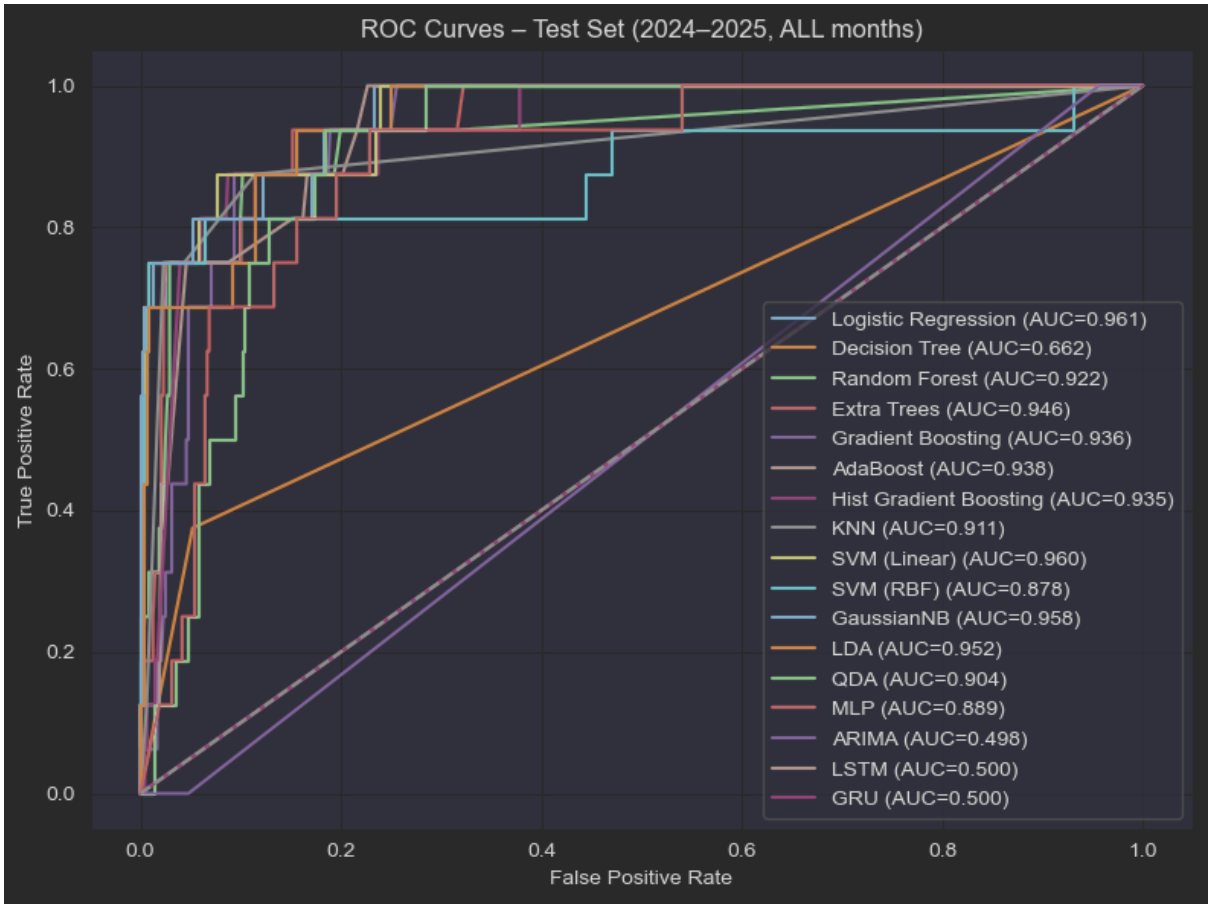
	Original	Limited to June-September	Monte Carlo
Accuracy	0	0	0.5
Recall	0	0	0
Precision	0	0	0

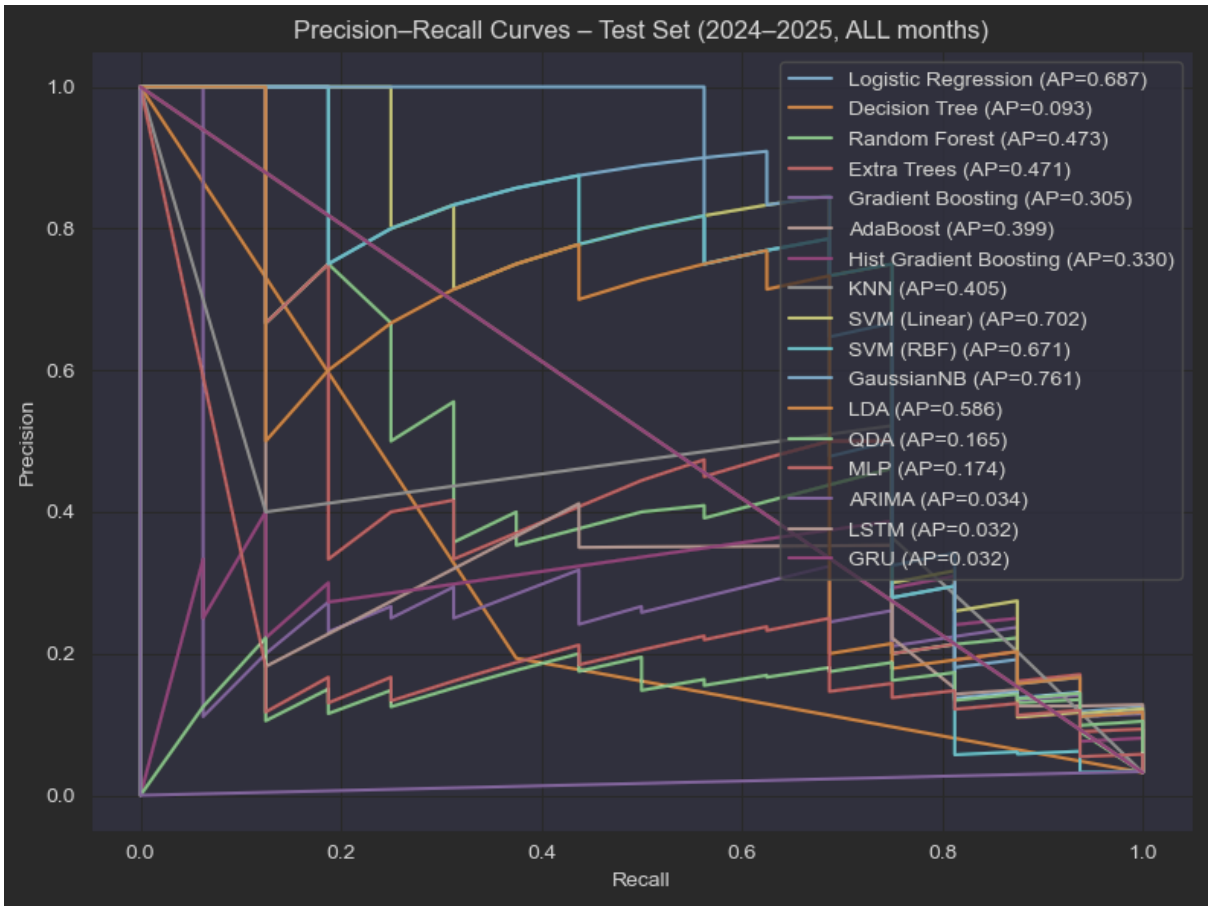
This showed that the supervised model can perform much better than the old ones.

Next, below are raw outputs of supervised learning models on RTO/TESLA

	model	Precision	Accuracy	Recall	AUC	F1score	WEI	WEI_share
0	Logistic Regression	0.5714	0.9645	0.7500	0.9680	0.6486	0.0564	1.0
8	SVM (Linear)	0.6667	0.9727	0.7500	0.9673	0.7059	0.0468	1.0
10	GaussianNB	0.2121	0.8525	0.8750	0.9646	0.3415	0.1460	1.0
11	LDA	0.7143	0.9727	0.6250	0.9573	0.6667	0.0631	1.0
5	AdaBoost	0.4615	0.9508	0.7500	0.9488	0.5714	0.3115	1.0
3	Extra Trees	0.5217	0.9590	0.7500	0.9477	0.6154	0.0604	1.0
6	Hist Gradient Boosting	0.4800	0.9536	0.7500	0.9429	0.5854	0.0539	1.0
4	Gradient Boosting	0.4074	0.9426	0.6875	0.9422	0.5116	0.0682	1.0
2	Random Forest	0.4615	0.9508	0.7500	0.9247	0.5714	0.0611	1.0
7	KNN	0.6316	0.9699	0.7500	0.9143	0.6857	0.0514	1.0
12	QDA	0.1379	0.8989	0.2500	0.8988	0.1778	0.1013	1.0
13	MLP	0.1333	0.9262	0.1250	0.8845	0.1290	0.0734	1.0
9	SVM (RBF)	0.7059	0.9754	0.7500	0.8811	0.7273	0.0581	1.0
1	Decision Tree	0.2727	0.9290	0.3750	0.6646	0.3158	0.0710	1.0
15	LSTM	0.0000	0.9563	0.0000	0.5000	0.0000	0.0833	1.0
16	GRU	0.0000	0.9563	0.0000	0.5000	0.0000	0.0794	1.0
14	ARIMA	0.0000	0.9508	0.0000	0.4971	0.0000	0.0895	1.0

	Model	precision_at_4	recall_at_4
0	Logistic Regression	0.7500	0.7500
8	SVM (Linear)	0.7500	0.7500
10	GaussianNB	0.7500	0.7500
11	LDA	0.6875	0.6875
9	SVM (RBF)	0.6250	0.6250
6	Hist Gradient Boosting	0.5625	0.5625
2	Random Forest	0.5625	0.5625
5	AdaBoost	0.5000	0.5000
3	Extra Trees	0.3750	0.3750
4	Gradient Boosting	0.3750	0.3750
7	KNN	0.3750	0.3750
1	Decision Tree	0.1875	0.1875
12	QDA	0.1875	0.1875
13	MLP	0.1875	0.1875
14	ARIMA	0.1250	0.1250
15	LSTM	0.1250	0.1250
16	GRU	0.1250	0.1250


ROC Curves – Test Set (2024–2025, ALL months)
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
Logistic Regression (AUC=0.961)
Decision Tree (AUC=0.662)
Random Forest (AUC=0.922)
Extra Trees (AUC=0.946)
Gradient Boosting (AUC=0.936)
AdaBoost (AUC=0.938)
Hist Gradient Boosting (AUC=0.935)
KNN (AUC=0.911)
SVM (Linear) (AUC=0.960)
SVM (RBF) (AUC=0.878)
GaussianNB (AUC=0.958)
LDA (AUC=0.952)
QDA (AUC=0.904)
MLP (AUC=0.889)
ARIMA (AUC=0.498)
LSTM (AUC=0.500)
GRU (AUC=0.500)


Precision–Recall Curves – Test Set (2024–2025, ALL months)
Precision
Recall
0.0
0.2
0.4
0.6
0.8
1.0
Logistic Regression (AP=0.687)
Decision Tree (AP=0.093)
Random Forest (AP=0.473)
Extra Trees (AP=0.471)
Gradient Boosting (AP=0.305)
AdaBoost (AP=0.399)
Hist Gradient Boosting (AP=0.330)
KNN (AP=0.405)
SVM (Linear) (AP=0.702)
SVM (RBF) (AP=0.671)
GaussianNB (AP=0.761)
LDA (AP=0.586)
QDA (AP=0.165)
MLP (AP=0.174)
ARIMA (AP=0.034)
LSTM (AP=0.032)
GRU (AP=0.032)

From the results and plots, we can see that the SVM (Linear) model performs best in predicting. So, we choose to evaluate that model.

2024 results – SVM (Linear)

4 rows | 4 rows × 4 cols

	SVM (Linear)_prob	actual_peak	is_CP	timestamp
900	0.991512	23851.82	1	2024-06-19
899	0.976383	23097.18	1	2024-06-18
901	0.894208	22485.78	1	2024-06-20
898	0.814411	21647.62	1	2024-06-17

Hits: 4/4 | True CP in month: 4 | Precision@4: 100.00%
Recall@4: 100.00%

4 rows | 4 rows × 4 cols

	SVM (Linear)_prob	actual_peak	is_CP	timestamp
920	0.953414	21836.21	0	2024-07-09
940	0.953346	22244.97	1	2024-07-29
926	0.948989	21637.19	0	2024-07-15
942	0.868142	23355.86	1	2024-07-31

Hits: 2/4 | True CP in month: 4 | Precision@4: 50.00%
Recall@4: 50.00%

4 rows | 4 rows × 4 cols

	SVM (Linear)_prob	actual_peak	is_CP	timestamp
943	0.977480	23179.24	1	2024-08-01
969	0.956275	22749.83	1	2024-08-27
944	0.918841	22103.13	1	2024-08-02
968	0.876354	22466.42	1	2024-08-26

Hits: 4/4 | True CP in month: 4 | Precision@4: 100.00%
Recall@4: 100.00%

4 rows | 4 rows × 4 cols

	SVM (Linear)_prob	actual_peak	is_CP	timestamp
991	0.264484	20180.01	0	2024-09-18
989	0.198206	21546.47	1	2024-09-16
992	0.142151	20886.45	1	2024-09-19
985	0.124292	19582.95	0	2024-09-12

Hits: 2/4 | True CP in month: 4 | Precision@4: 50.00%
Recall@4: 50.00%

For 2024's results, the model can achieve an average precision of 75% and an average recall of 75%. For June and August, it can even achieve 100% accuracy.

	Original	Limited to June-September	Monte Carlo
Accuracy	0.333333333	0.333333333	0.428571429
Recall	0.6	0.6	0.6
Precision	0.428571429	0.428571429	0.6

While in the previous algorithm, both Recall and Precision are lower than the new model. And the new model is even more accurate.

Though the models used for the datasets are different, we can find that the logistic regression model can achieve almost the second-best on both sides when evaluating normal power load, while maintaining the same precision and recall value when predicting 4CP. So, I think we can simply conclude that logistic regression is a better model in general evaluation.