

Multi site model evaluation:

- Data sources of : watch/view/ purchase data.
- Watch contribution to performance is ~ + 6.15% in precision of class1.

All features model results:

```
Label counts:
0      162165
1      34666
Name: vi_label, dtype: int64
Metrics:
Accuracy: 0.7811523591304216
ROC AUC: 0.7550880170242418
Precision (micro): 0.7811523591304216
Recall (micro): 0.7811523591304216
Precision (macro): 0.6781388317920642
Recall (macro): 0.7550880170242418
```

	precision	recall	f1-score	support
0	0.93	0.80	0.86	162165
1	0.43	0.71	0.54	34666
accuracy			0.78	196831
macro avg	0.68	0.76	0.70	196831
weighted avg	0.84	0.78	0.80	196831

Feature importance of ranker:

- 1) f_max_LEAF_cat_watch_propensity_cosine_sim_in_last_90_days 0.004316
- 2) f_num_views_from_LEAF_cat_in_last_30_days 0.003276
- 3) f_var_time_since_watch_from_LEAF_cat 0.002978
- 4) f_num_views_from_LEAF_cat_in_last_3_days 0.002136
- 5) f_max_LEAF_cat_view_propensity_cosine_sim_in_last_30_days 0.001699
- 6) f_num_purchases_from_LEAF_cat_in_last_360_days 0.001606
- 7) f_max_LEAF_cat_view_propensity_jaccard_sim_in_last_30_days 0.001586
- 8) f_var_time_since_purchase_from_LEAF_cat 0.001406
- 9) f_num_watchs_from_LEAF_cat_in_last_90_days 0.001254
- 10) f_p75_META_cat_purchase_propensity_jaccard_sim_in_last_1_days 0.001226
- 11) f_poisson_META_cat_purchase_propensity_jaccard_sim_in_last_3_days 0.001193
- 12) f_num_watchs_from_LEAF_cat_in_last_360_days 0.001096
- 13) f_mean_time_since_view_from_L2_cat 0.001096
- 14) f_num_views_from_LEAF_cat_in_last_14_days 0.001085
- 15) f_poisson_META_cat_purchase_propensity_cosine_sim_in_last_3_days 0.001081
- 16) f_min_time_since_view_from_L2_cat 0.001059
- 17) f_num_views_from_LEAF_cat_in_last_5_days 0.001051
- 18) f_p75_META_cat_purchase_propensity_cosine_sim_in_last_1_days 0.001037
- 19) f_p75_META_cat_purchase_propensity_jaccard_sim_in_last_3_days 0.001014
- 20) f_num_views_from_L2_cat_in_last_30_days 0.001005
- 21) f_p75_META_cat_purchase_propensity_cosine_sim_in_last_3_days 0.000997
- 22) f_num_views_from_L2_cat_in_last_60_days 0.000980

23) f_p50_META_cat_purchase_propensity_jaccard_sim_in_last_1_days 0.000968
 24) f_num_views_from_L2_cat_in_last_360_days 0.000965
 25) f_poisson_META_cat_purchase_propensity_cosine_sim_in_last_2_days 0.000956
 26) f_poisson_META_cat_purchase_propensity_jaccard_sim_in_last_7_days 0.000954
 27) f_min_time_since_view_from_LEAF_cat 0.000953
 28) f_max_time_since_view_from_L2_cat 0.000947
 29) f_num_views_from_L2_cat_in_last_180_days 0.000946
 30) f_num_views_from_L2_cat_in_last_90_days 0.000946
 31) f_poisson_L2_cat_purchase_propensity_jaccard_sim_in_last_1_days 0.000943
 32) f_poisson_L2_cat_purchase_propensity_cosine_sim_in_last_2_days 0.000939
 33) f_poisson_META_cat_purchase_propensity_cosine_sim_in_last_1_days 0.000936
 34) f_var_time_since_view_from_LEAF_cat 0.000934
 35) f_poisson_META_cat_watch_propensity_jaccard_sim_in_last_1_days 0.000921
 36) f_p50_META_cat_purchase_propensity_cosine_sim_in_last_3_days 0.000907
 37) f_p75_L2_cat_purchase_propensity_jaccard_sim_in_last_1_days 0.000892
 38) f_p50_META_cat_purchase_propensity_jaccard_sim_in_last_3_days 0.000890
 39) f_p75_META_cat_purchase_propensity_jaccard_sim_in_last_2_days 0.000888
 40) f_max_L2_cat_view_propensity_cosine_sim_in_last_360_days 0.000885
 41) f_p75_L2_cat_purchase_propensity_jaccard_sim_in_last_2_days 0.000882

Top 30 features model multi site:

```
Label counts:
0      162165
1      34666
Name: vi_label, dtype: int64
Metrics:
Accuracy: 0.8040755775258979
ROC AUC: 0.7491886626844316
Precision (micro): 0.8040755775258979
Recall (micro): 0.8040755775258979
Precision (macro): 0.6908957436714542
Recall (macro): 0.7491886626844316
```

	precision	recall	f1-score	support
0	0.92	0.83	0.88	162165
1	0.46	0.66	0.54	34666
accuracy			0.80	196831
macro avg	0.69	0.75	0.71	196831
weighted avg	0.84	0.80	0.82	196831

Feature importance of ranker:

1) f_var_time_since_watch_from_LEAF_cat 0.059936
 2) f_num_views_from_LEAF_cat_in_last_30_days 0.045752
 3) f_num_views_from_LEAF_cat_in_last_14_days 0.042673
 4) f_var_time_since_purchase_from_LEAF_cat 0.039961

5) f_max_LEAF_cat_view_propensity_cosine_sim_in_last_30_days 0.038901
6) f_max_LEAF_cat_watch_propensity_cosine_sim_in_last_90_days 0.037873
7) f_max_LEAF_cat_view_propensity_jaccard_sim_in_last_30_days 0.036934
8) f_min_time_since_view_from_LEAF_cat 0.036657
9) f_num_purchases_from_LEAF_cat_in_last_360_days 0.036528
10) f_num_watchs_from_LEAF_cat_in_last_360_days 0.036416
11) f_num_views_from_L2_cat_in_last_60_days 0.035760
12) f_num_views_from_L2_cat_in_last_180_days 0.034645
13) f_p50_META_cat_purchase_propensity_jaccard_sim_in_last_1_days 0.034284
14) f_poisson_META_cat_purchase_propensity_jaccard_sim_in_last_7_days 0.034155
15) f_p75_META_cat_purchase_propensity_cosine_sim_in_last_3_days 0.032665
16) f_p75_META_cat_purchase_propensity_jaccard_sim_in_last_3_days 0.032284
17) f_p75_META_cat_purchase_propensity_cosine_sim_in_last_1_days 0.031887
18) f_poisson_META_cat_purchase_propensity_cosine_sim_in_last_2_days 0.031763
19) f_num_views_from_L2_cat_in_last_90_days 0.031724
20) f_num_watchs_from_LEAF_cat_in_last_90_days 0.031718
21) f_poisson_META_cat_purchase_propensity_cosine_sim_in_last_3_days 0.028863
22) f_num_views_from_LEAF_cat_in_last_5_days 0.028321
23) f_poisson_META_cat_purchase_propensity_jaccard_sim_in_last_3_days 0.028168
24) f_p75_META_cat_purchase_propensity_jaccard_sim_in_last_1_days 0.027940
25) f_num_views_from_L2_cat_in_last_360_days 0.027508
26) f_num_views_from_LEAF_cat_in_last_3_days 0.027177
27) f_max_time_since_view_from_L2_cat 0.025265
28) f_min_time_since_view_from_L2_cat 0.024617
29) f_mean_time_since_view_from_L2_cat 0.024134
30) f_num_views_from_L2_cat_in_last_30_days 0.015490

file_name = "xgb_model_ms_top30_ol.pkl"