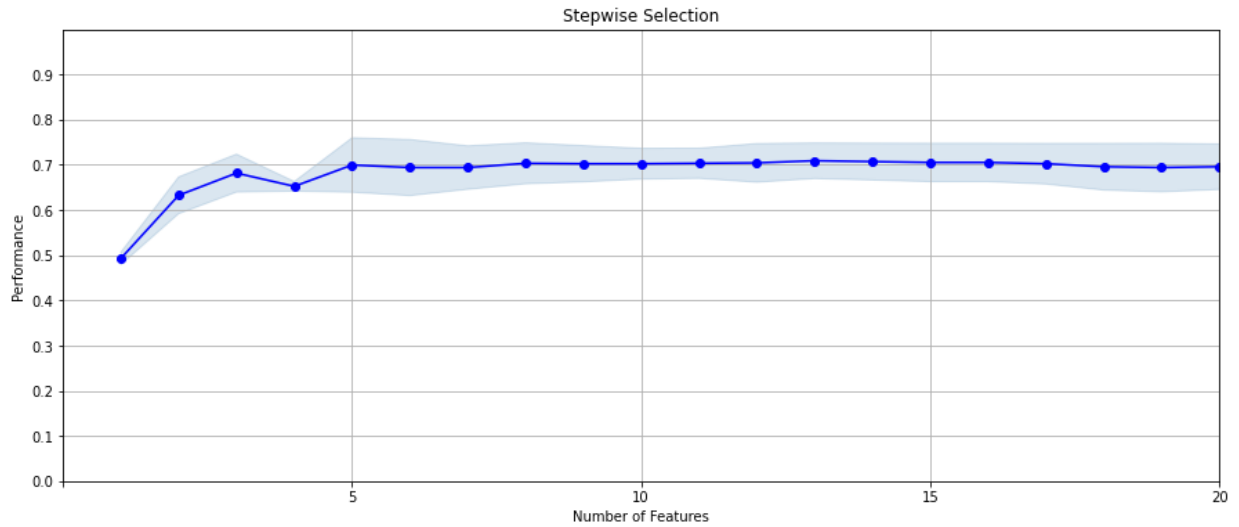# BACKWARD SELECTION

Santhoshini Jayachandran (8218065467)
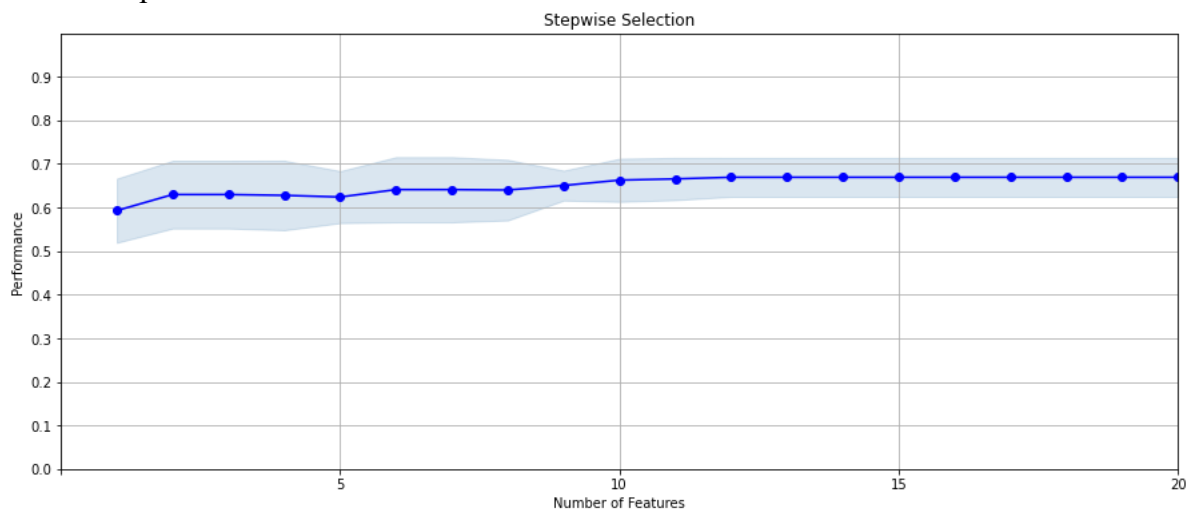
LGBM
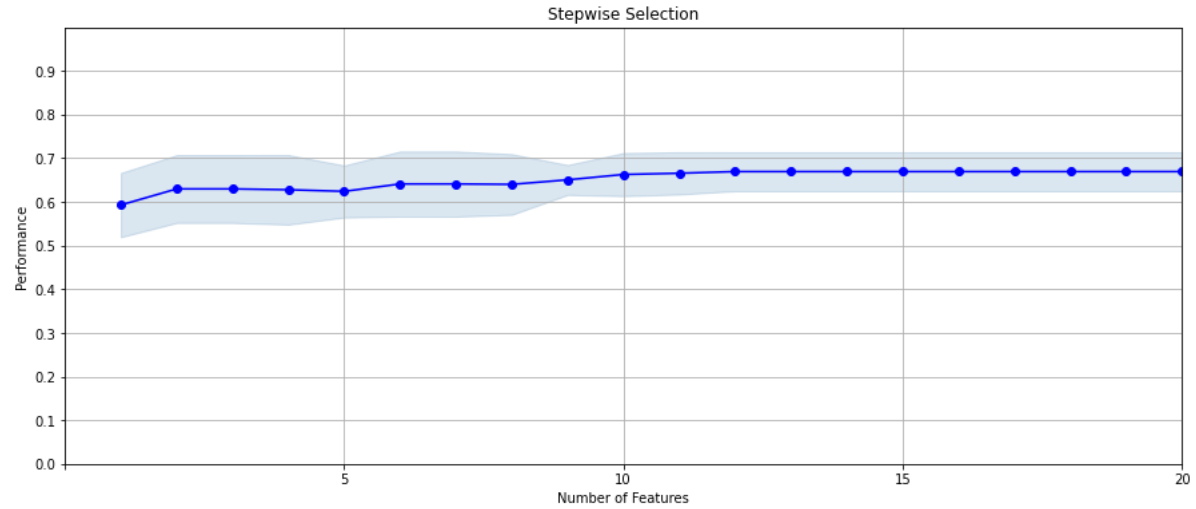1. Num_filter= 150, Num_wrapper= 20

FORWARD SELECTION

LGBM

1. Num_filter= 150, Num_wrapper= 20
First attempt:

Second attempt: LGBM has no stochastic nature so same result is obtained for both the run.

Stepwise Selection

First Attempt			Second Attempt	
add variables in this order	**variable name**		**add variables in this order**	**variable name**
1.0	card_merch_total_14		1.0	card_merch_total_14
2.0	Card_Merchnum_desc_max_60		2.0	Card_Merchnum_desc_max_60
3.0	card_merch_total_30		3.0	card_merch_total_30
4.0	Merchnum_desc_total_7		4.0	Merchnum_desc_total_7
5.0	card_zip3_total_30		5.0	card_zip3_total_30
6.0	Cardnum_avg_1		6.0	Cardnum_avg_1
7.0	Card_Merchnum_desc_total_30		7.0	Card_Merchnum_desc_total_30
8.0	Card_Merchdesc_total_3		8.0	Card_Merchdesc_total_3
9.0	card_zip3_max_30		9.0	card_zip3_max_30
10.0	Card_Merchdesc_total_7		10.0	Card_Merchdesc_total_7
11.0	Cardnum_max_0		11.0	Cardnum_max_0
12.0	card_merch_max_60		12.0	card_merch_max_60
13.0	Card_Merchnum_desc_total_14		13.0	Card_Merchnum_desc_total_14
14.0	Card_Merchdesc_total_30		14.0	Card_Merchdesc_total_30
15.0	card_merch_max_30		15.0	card_merch_max_30
16.0	Card_Merchdesc_max_14		16.0	Card_Merchdesc_max_14
17.0	Card_Merchdesc_max_30		17.0	Card_Merchdesc_max_30
18.0	Card_Merchnum_desc_max_14		18.0	Card_Merchnum_desc_max_14
19.0	card_merch_total_60		19.0	card_merch_total_60
20.0	Card_Merchnum_desc_max_30		20.0	Card_Merchnum_desc_max_30

2. Num_filter= 300, Num_wrapper= 25

Stepwise Selection
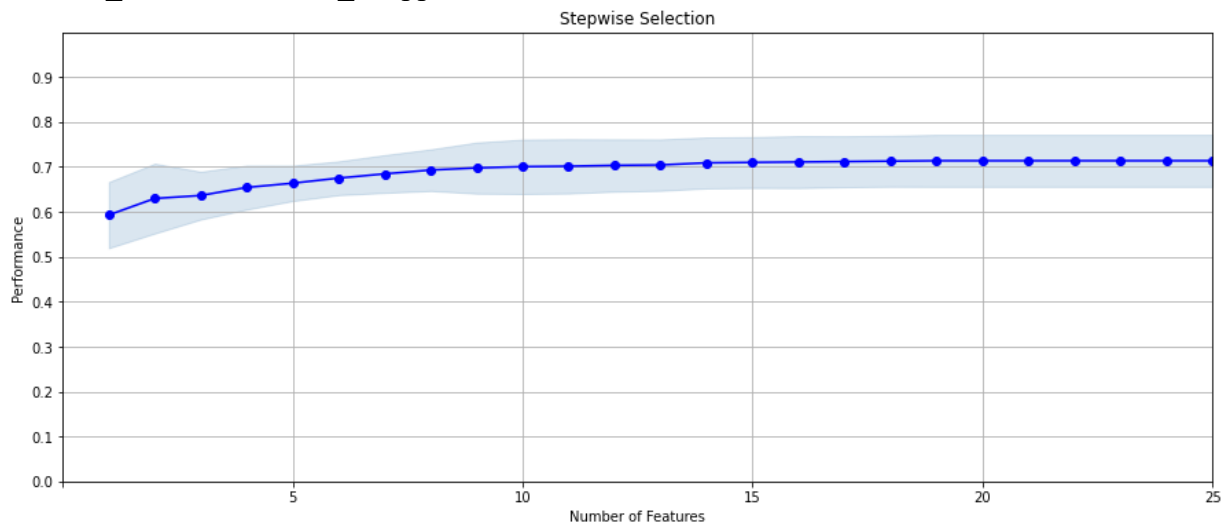
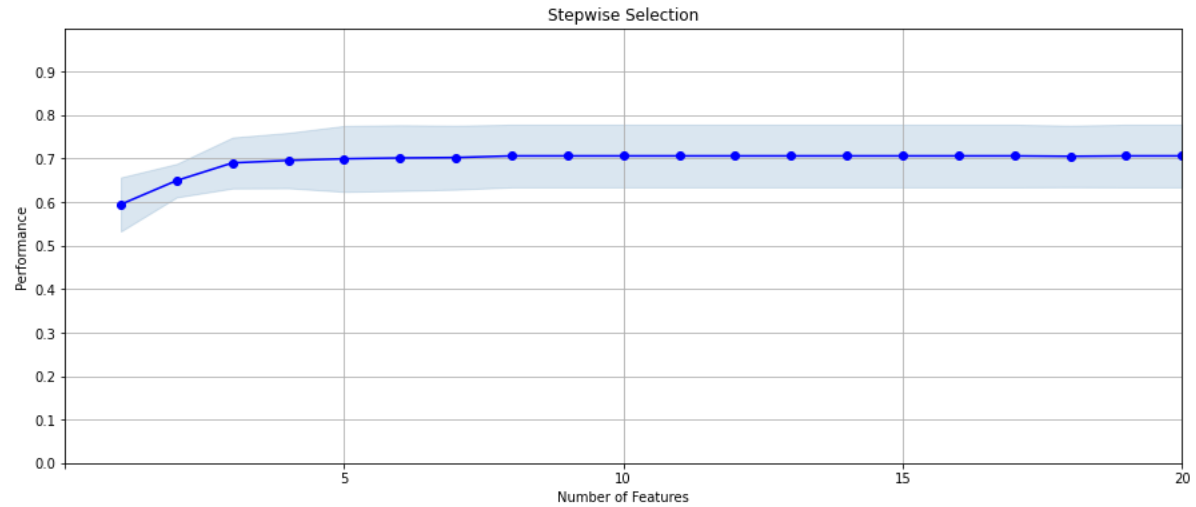| wrapper order	variable	filter score
1	card_merch_total_14	0.6283541638475630
2	Card_Merchnum_desc_max_60	0.5838934744987920
3	card_zip3_variability_max_7	0.37546760473758800
4	card_zip3_max_30	0.6120288967015240
5	Cardnum_max_0	0.5279133209093690
6	card_merch_total_7	0.6355457332057540
7	Merchnum_desc_actual/med_60	0.3573414788170500
8	zip3_avg_3	0.35808927794516900
9	card_merch_avg_30	0.5202343648028800
10	zip3_actual/max_30	0.38912697294419200
11	card_merch_avg_3	0.522054963493437
12	Merchnum_desc_total_0	0.4989755944314700
13	Card_Merchdesc_total_30	0.6055993197979510
14	card_zip3_avg_7	0.5320855448130660
15	card_merch_total_60	0.5974170205935560
16	Merchnum_desc_total_14	0.48862569793182100
17	Merchnum_variability_avg_30	0.3639662164361260
18	Card_Merchdesc_max_60	0.5841780640663190
19	zip3_variability_med_7	0.4093761055417740
20	Card_Merchnum_desc_total_7	0.6160712214586610
21	card_merch_total_30	0.6145933821936370
22	Card_Merchnum_desc_total_30	0.6050525053043720
23	card_merch_max_30	0.602176073672655
24	Card_Merchdesc_max_30	0.5980533431761530

25	Card_Merchnum_desc_max_14	0.597943461275128

3. Num_filter = 350, Num_wrapper =20

Stepwise Selection
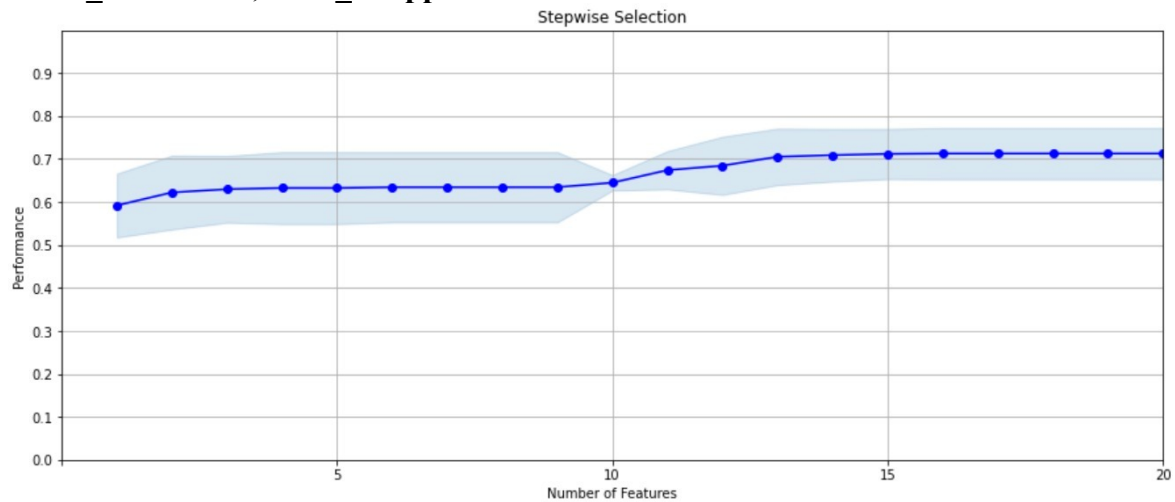
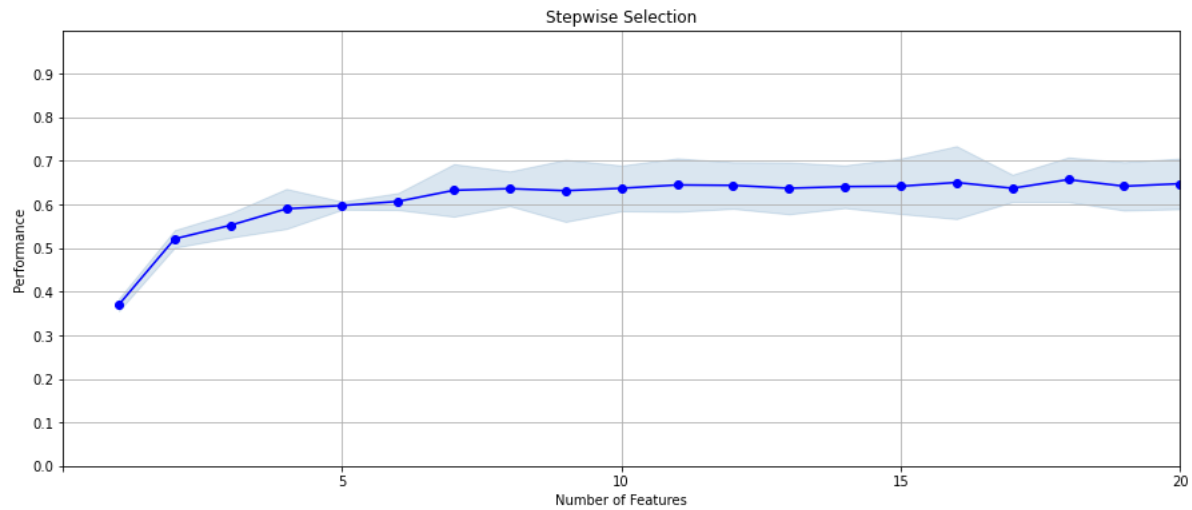| wrapper order	variable	filter score
1	card_merch_state_total_14	0.6280604719303840
2	card_merch_state_max_60	0.5889723677029930
3	merchnum_desc_total_7	0.5138036290582740
4	merch_zip_med_7	0.3801237091526120
5	card_merch_total_14	0.6280394939363000
6	merchnum_desc_avg_7	0.45386521125741400
7	card_merch_max_60	0.5889409007118670
8	merchnum_desc_total_14	0.49097523326925200
9	merch_zip_avg_7	0.4539830861901790
10	card_state_max_14	0.6002801106964570
11	card_state_total_1	0.6075726874123780
12	merchnum_desc_avg_60	0.3537856592966560
13	merchnum_desc_total_1	0.5220379473490380
14	merchnum_desc_total_0	0.49911195139301700
15	card_merch_state_total_30	0.6150971313811360
16	card_zip_avg_1	0.5112480181094560
17	card_merch_total_30	0.6150551753929670
18	card_merch_desc_total_30	0.6052308182540880
19	card_merch_state_total_60	0.5983825629803850
20	card_merch_total_60	0.5983196289981330

## 4. Num_filter = 300, Num_wrapper =30, leaves= 4

Stepwise Selection

wrapper order	variable	filter score
1	card_zip_total_14	0.6504807278694390
2	card_zip3_max_30	0.6120288967015240
3	card_merch_total_0	0.5442033374978320
4	Card_Merchdesc_avg_60	0.5167818207991600
5	merch_zip_total_14	0.43945106116261900
6	Card_Merchnum_desc_avg_3	0.517344017174143
7	Merchnum_desc_avg_0	0.5080877793755260
8	Card_Merchnum_desc_avg_0	0.50682272115248
9	Card_Merchnum_desc_avg_7	0.5172485068828520
10	Card_Merchdesc_avg_14	0.5164177565269440
11	Card_Merchdesc_avg_7	0.5159277747671020
12	Card_Merchnum_desc_avg_60	0.5138123253183330
13	Card_Merchdesc_avg_1	0.5125174640810290
14	Card_Merchnum_desc_avg_1	0.5089920022338490
15	Card_Merchdesc_avg_0	0.5068017431583960
16	Card_Merchdesc_avg_3	0.5163421149118260
17	Merchnum_desc_avg_1	0.5048878032892510
18	Merchnum_avg_1	0.5058845353377280
19	card_merch_avg_30	0.5202343648028800
20	card_merch_avg_3	0.522054963493437

RF

1. Num_filter= 150, Num_wrapper= 20

Second attempt: Num_filter= 150, Num_wrapper= 20
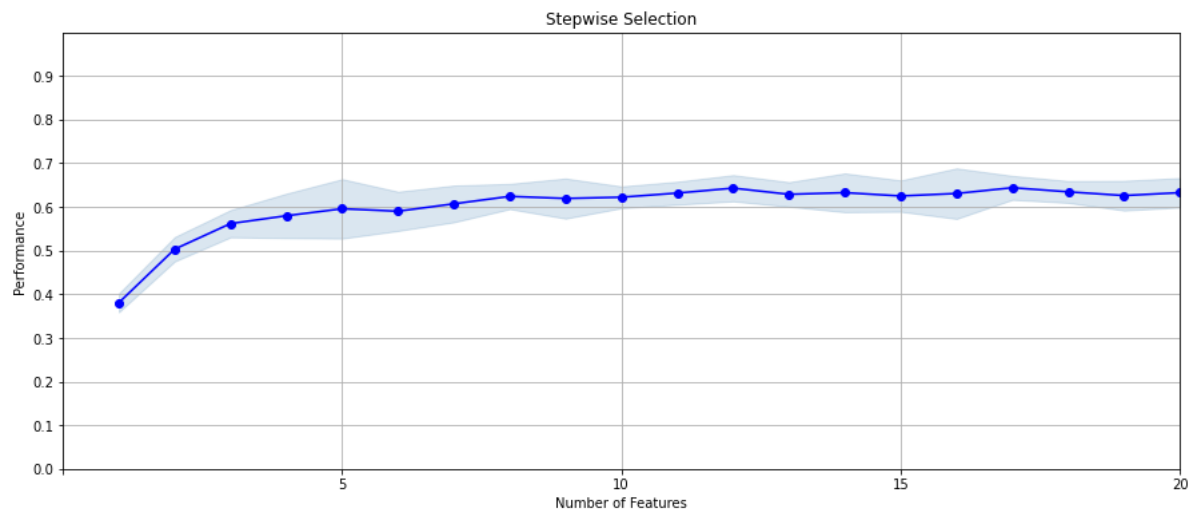
Second attempt: LGBM has stochastic nature, so we get the different results.

First Attempt		Second attempt	
add variables in this order	**variable name**	**add variables in this order**	**variable name**
1.0	card_zip3_total_7	1.0	card_zip3_total_3
2.0	Card_Merchdesc_max_60	2.0	Card_Merchdesc_max_30
3.0	Card_Merchnum_desc_total_14	3.0	Card_Merchdesc_total_30
4.0	Merchnum_desc_total_1	4.0	Merchnum_total_0
5.0	Cardnum_total_3	5.0	card_zip_total_30

6.0	card_zip_avg_7		6.0	card_merch_total_0
7.0	zip3_actual/avg_60		7.0	card_zip_max_60
8.0	Card_Merchnum_desc_total_30		8.0	Cardnum_total_14
9.0	card_zip3_max_3		9.0	card_zip3_max_14
10.0	Cardnum_max_1		10.0	card_zip3_avg_3
11.0	Card_Merchdesc_avg_7		11.0	Card_Merchdesc_avg_1
12.0	card_merch_max_14		12.0	card_merch_total_14
13.0	Card_Merchnum_desc_max_60		13.0	card_zip3_max_7
14.0	card_merch_med_30		14.0	Card_Merchdesc_total_0
15.0	Card_Merchnum_desc_max_0		15.0	Cardnum_avg_3
16.0	Merchnum_desc_total_3		16.0	card_merch_max_0
17.0	card_zip_total_3		17.0	Card_Merchdesc_max_3
18.0	card_merch_total_60		18.0	Card_Merchdesc_total_14
19.0	Card_Merchnum_desc_max_3		19.0	card_zip3_avg_0
20.0	card_merch_max_30		20.0	Card_Merchnum_desc_total_7

2. Num_filter= 300, Num_wrapper= 25

Stepwise Selection
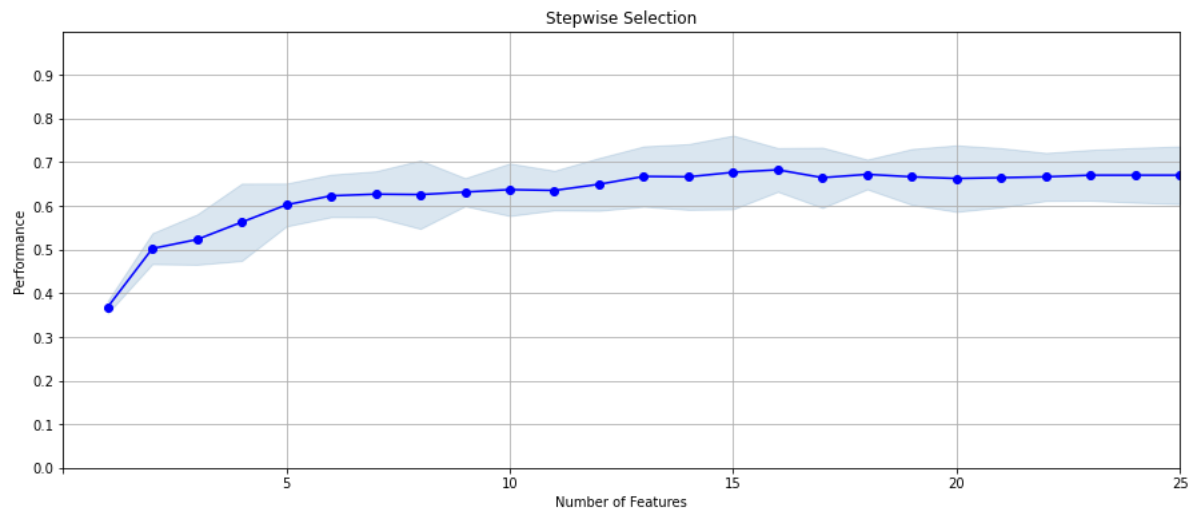
wrapper order	variable	filter score
1	card_zip3_total_3	0.6567020897627960
2	card_zip3_actual/toal_3	0.35693097176189700
3	merch_zip_actual/med_60	0.3738031524477570
4	card_merch_max_60	0.5892238263025260
5	card_zip_total_30	0.635873110749883
6	Cardnum_total_7	0.5620477957149860
7	Merchnum_max_30	0.46005170887353900
8	card_zip3_med_0	0.4873459907695640

9	card_zip_max_0	0.5392450429980540
10	zip3_avg_0	0.4669014993639940
11	Cardnum_max_60	0.4798792412368490
12	Card_Merchnum_desc_total_0	0.5335843125490540
13	Merchnum_avg_1	0.5058845353377280
14	Merchnum_desc_total_1	0.5215869204762280
15	card_merch_avg_3	0.522054963493437
16	Merchnum_total_1	0.5225005725368030
17	card_zip_total_60	0.6169900932561970
18	merch_zip_avg_30	0.4099350234861420
19	zip3_actual/med_14	0.39330699178714500
20	zip3_actual/avg_60	0.5110021060994840
21	card_merch_avg_1	0.5127691504869180
22	card_zip3_total_30	0.6283605424255570
23	card_zip3_total_7	0.6741055544914690
24	card_zip_avg_30	0.5354486797740110
25	card_merch_avg_0	0.510177258899607

Given all these trials I think I'll go with FS LGBM, num_filter=300, num_wrapper=25. It has good set of variables as well and performance is greater than .70.