

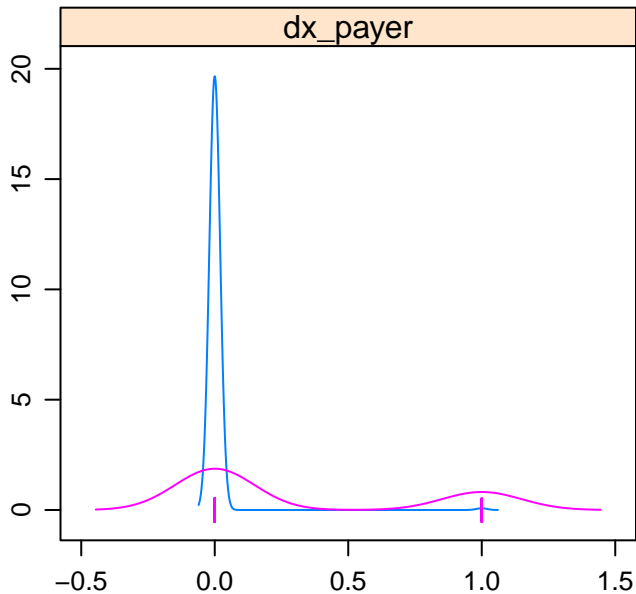
FALSE



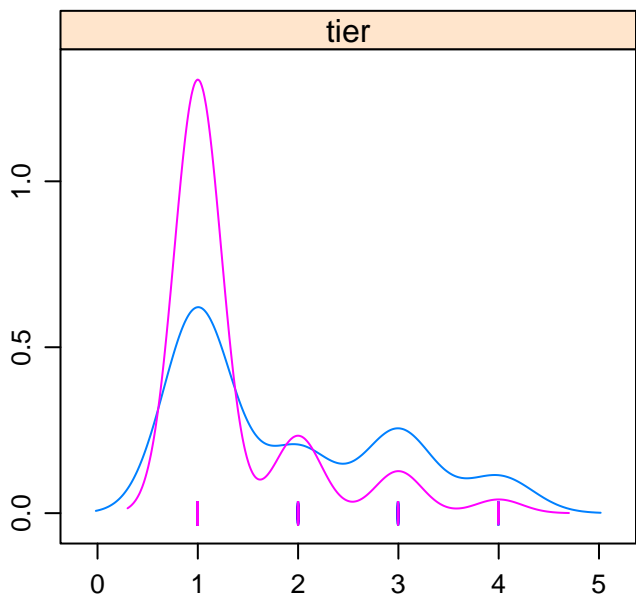
TRUE



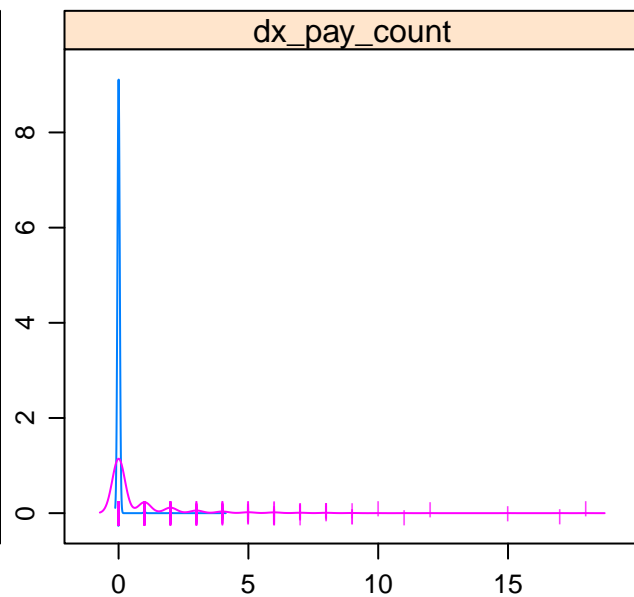
dx_payer



tier

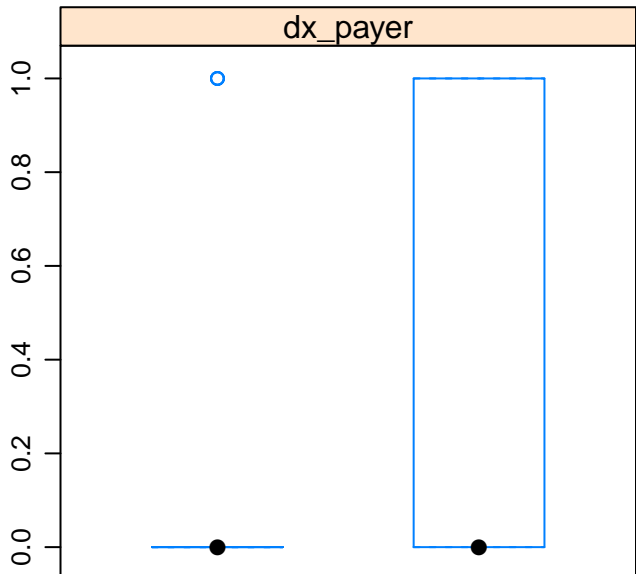


dx_pay_count

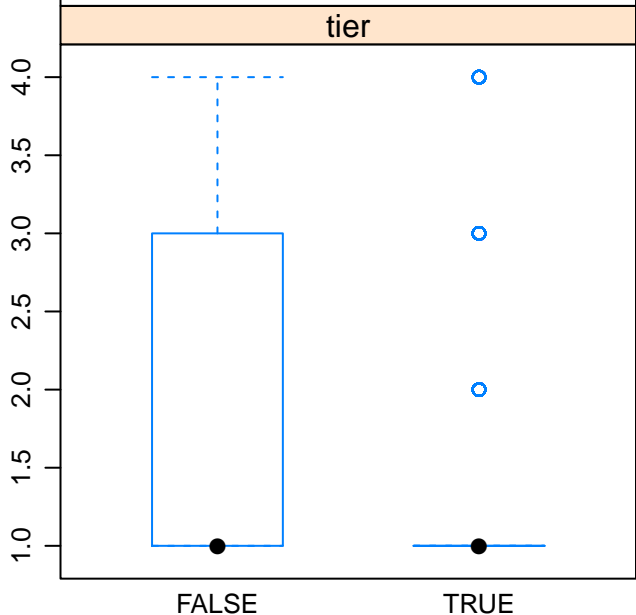


Feature

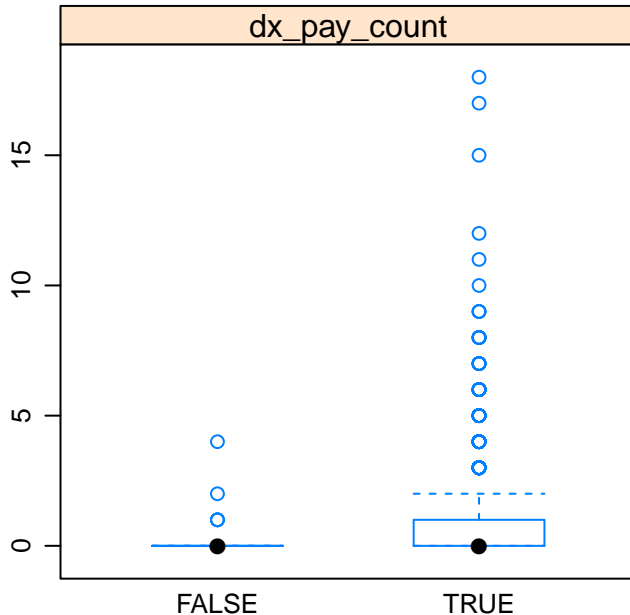
dx_payer



tier



dx_pay_count

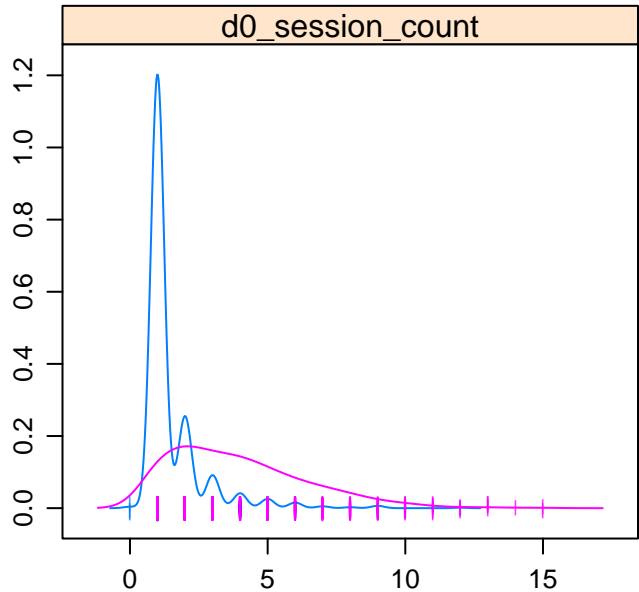
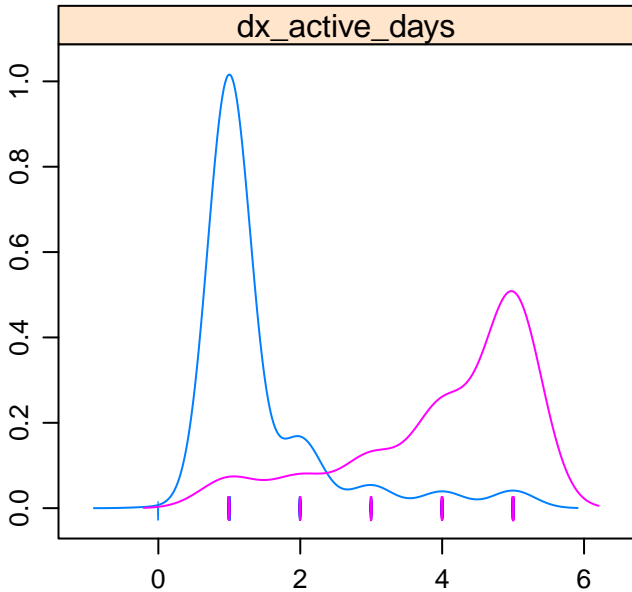
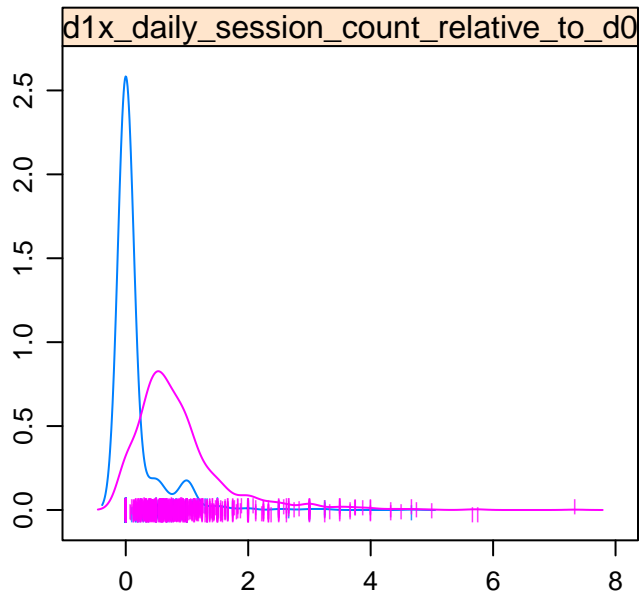
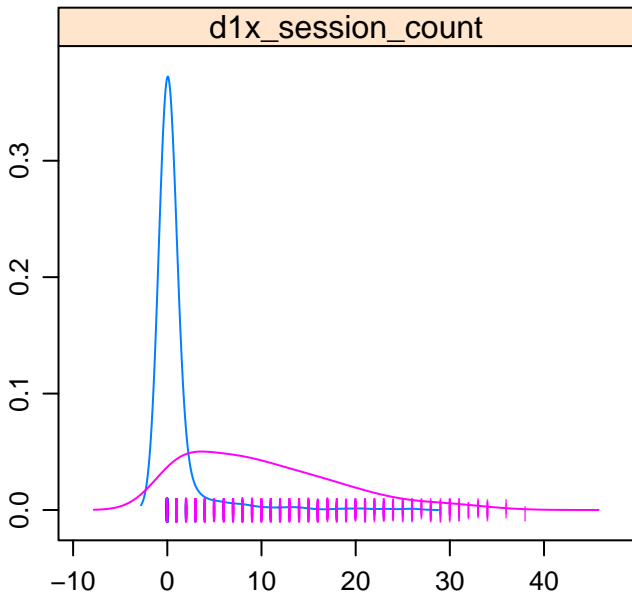


Feature

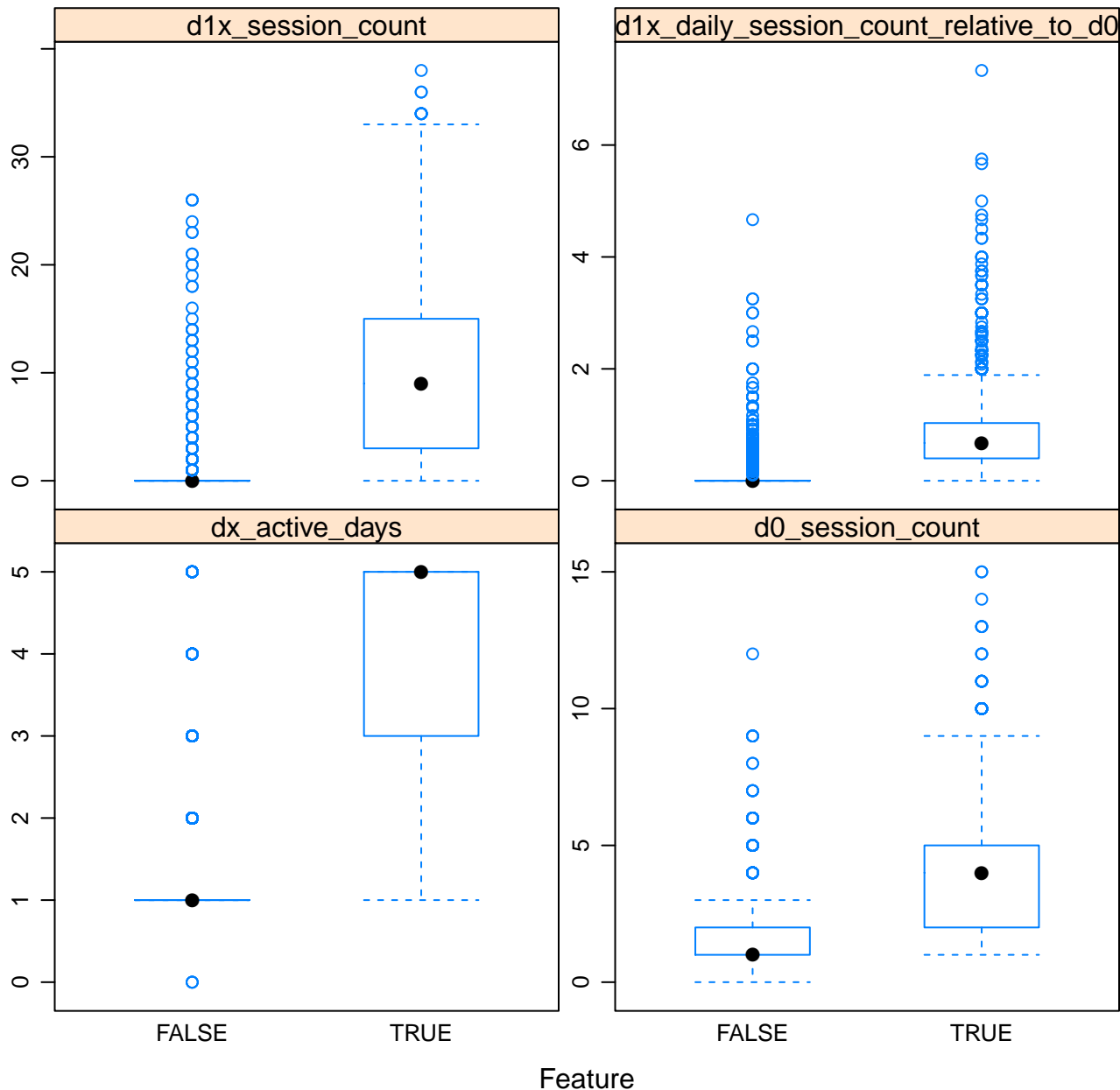
FALSE



TRUE



Feature



Selection By Filter

Outer resampling method: Cross-Validated (10 fold, repeated 5 times)

Resampling performance:

Accuracy	Kappa	AccuracySD	KappaSD
0.8863	0.7727	0.02107	0.04215

Using the training set, 7 variables were selected:

tier, dx_pay_count, dx_payer, dx_active_days, d0_session_count...

During resampling, the top 5 selected variables (out of a possible 7):

d0_session_count (100%), dlx_daily_session_count_relative_to_d0 (100%), dlx_sess

On average, 7 variables were selected (min = 7, max = 7)

	variables	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1:	dx_active_days	0.9063	0.9100	0.9135	0.9162	0.9226	0.9295
2:	dlx_session_coun	0.8965	0.9117	0.9190	0.9151	0.9203	0.9263
3:	dlx_daily_sessio	0.8684	0.8746	0.8755	0.8764	0.8771	0.8874
4:	d0_session_count	0.8274	0.8302	0.8346	0.8377	0.8457	0.8513
5:	dx_pay_count	0.6294	0.6409	0.6440	0.6500	0.6641	0.6718
6:	dx_payer	0.6315	0.6363	0.6487	0.6499	0.6624	0.6714
7:	tier	0.6129	0.6228	0.6241	0.6336	0.6421	0.6700

	variables	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1:	dx_active_days+tier	0.9116	0.9260	0.9285	0.9303	0.9328	0.9540
2:	d0_session_count+dlx_session_coun	0.9144	0.9240	0.9262	0.9249	0.9285	0.9300
3:	dx_active_days+dx_payer	0.9074	0.9157	0.9254	0.9239	0.9309	0.9398
4:	dx_active_days+dx_pay_count	0.9092	0.9219	0.9244	0.9238	0.9277	0.9347
5:	d0_session_count+dx_active_days	0.9069	0.9120	0.9165	0.9228	0.9244	0.9593
6:	dlx_session_count+tier	0.8993	0.9190	0.9252	0.9221	0.9293	0.9350
7:	dlx_session_coun+dx_pay_count	0.9064	0.9122	0.9125	0.9203	0.9295	0.9434
8:	dlx_session_coun+dx_payer	0.9062	0.9119	0.9195	0.9203	0.9282	0.9363
9:	dlx_session_coun+dx_active_days	0.9051	0.9194	0.9219	0.9202	0.9250	0.9275
10:	dlx_daily_sessio+dx_active_days	0.8642	0.9173	0.9230	0.9185	0.9306	0.9521
11:	d0_session_count+dlx_daily_sessio	0.8924	0.9000	0.9202	0.9162	0.9329	0.9338
12:	dlx_daily_sessio+dlx_session_coun	0.8981	0.9099	0.9127	0.9108	0.9155	0.9159
13:	dlx_daily_sessio+dx_pay_count	0.8737	0.8852	0.9029	0.8974	0.9074	0.9169
14:	dlx_daily_sessio+dx_payer	0.8703	0.8823	0.9013	0.8968	0.9068	0.9235
15:	dlx_daily_sessio+tier	0.8647	0.8717	0.8840	0.8875	0.8957	0.9249
16:	d0_session_count+dx_pay_count	0.8376	0.8530	0.8650	0.8640	0.8775	0.8855
17:	d0_session_count+dx_payer	0.8259	0.8599	0.8656	0.8640	0.8765	0.8891
18:	d0_session_count+tier	0.8366	0.8492	0.8664	0.8626	0.8774	0.8817
19:	dx_pay_count+tier	0.7211	0.7295	0.7422	0.7418	0.7499	0.7673
20:	dx_payer+tier	0.7233	0.7327	0.7402	0.7413	0.7527	0.7572
21:	dx_pay_count+dx_payer	0.6265	0.6439	0.6501	0.6500	0.6546	0.6752

	variables	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
--	-----------	------	---------	--------	------	---------	------

	variables	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1:	dx_active_days+dx_payer+tier	0.9167	0.9306	0.9412	0.9375	0.9454	0.9518
2:	dx_active_days+dx_pay_count+tier	0.9247	0.9265	0.9346	0.9359	0.9431	0.9516
3:	dlx_daily_sessio+dx_active_days+tier	0.9268	0.9269	0.9323	0.9334	0.9389	0.9427
4:	d0_session_count+dx_active_days+tier	0.9120	0.9277	0.9385	0.9332	0.9405	0.9448
5:	dlx_session_count+dx_active_days+tier	0.9171	0.9232	0.9312	0.9328	0.9408	0.9528
6:	dlx_daily_sessio+dx_active_days+dx_payer	0.9076	0.9115	0.9258	0.9261	0.9406	0.9452
7:	dlx_daily_sessio+dx_active_days+dx_pay_count	0.9083	0.9203	0.9248	0.9252	0.9326	0.9394
8:	d0_session_count+dlx_daily_sessio+dx_active_days	0.9047	0.9236	0.9258	0.9245	0.9302	0.9363
9:	dx_active_days+dx_pay_count+dx_payer	0.9123	0.9169	0.9239	0.9238	0.9294	0.9369
10:	dlx_session_count+dx_active_days+dx_pay_count	0.9135	0.9163	0.9208	0.9233	0.9273	0.9405
11:	dlx_session_coun+dx_active_days+dx_payer	0.9100	0.9166	0.9244	0.9233	0.9278	0.9380
12:	d0_session_count+dlx_session_count+dx_active_days	0.9134	0.9163	0.9227	0.9232	0.9279	0.9364
13:	d0_session_count+dx_active_days+dx_payer	0.8888	0.9143	0.9227	0.9220	0.9353	0.9471
14:	d0_session_count+dx_active_days+dx_pay_count	0.8886	0.9121	0.9211	0.9218	0.9337	0.9529
15:	dlx_daily_sessio+dlx_session_count+dx_active_days	0.9029	0.9147	0.9210	0.9191	0.9268	0.9283

	variables	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1:	dlx_session_coun+dx_active_days+dx_payer+tier	0.9256	0.9302	0.9407	0.9385	0.9460	0.9492
2:	dlx_daily_sessio+dx_active_days+dx_payer+tier	0.9322	0.9373	0.9393	0.9384	0.9404	0.9426
3:	d0_session_count+dx_active_days+dx_pay_count+tier	0.9149	0.9294	0.9375	0.9381	0.9475	0.9612
4:	dlx_daily_sessio+dx_active_days+dx_pay_count+tier	0.9244	0.9348	0.9365	0.9380	0.9408	0.9543
5:	dx_active_days+dx_pay_count+dx_payer+tier	0.9256	0.9334	0.9384	0.9377	0.9431	0.9473
6:	dlx_session_coun+dx_active_days+dx_pay_count+tier	0.9300	0.9336	0.9362	0.9376	0.9396	0.9494
7:	d0_session_count+dlx_daily_sessio+dx_active_days+tier	0.9294	0.9303	0.9340	0.9361	0.9423	0.9450
8:	d0_session_count+dx_active_days+dx_payer+tier	0.9212	0.9256	0.9374	0.9359	0.9429	0.9528
9:	d0_session_count+dlx_session_coun+dx_active_days+tier	0.9027	0.9149	0.9422	0.9348	0.9513	0.9615
10:	dlx_daily_sessio+dlx_session_coun+dx_active_days+tier	0.9090	0.9277	0.9329	0.9326	0.9405	0.9517

	variables	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1:	dlx_session_count+dx_active_days+dx_pay_count+tier	0.9313	0.9343	0.9369	0.9374	0.9410	0.9434
2:	dlx_daily_sessio+dx_active_days+dx_pay_count+tier	0.9265	0.9313	0.9376	0.9374	0.9422	0.9500
3:	dx_active_days+dx_pay_count+dx_payer+tier	0.9228	0.9283	0.9364	0.9367	0.9446	0.9517
4:	d0_session_count+dx_active_days+dx_pay_count+tier	0.9143	0.9306	0.9346	0.9364	0.9432	0.9597

	variables	Min.	1st Qu.	Median
1:	d0_session_count+dlx_daily_sessio+dx_active_days+dx_payer+tier	0.9274	0.9379	0.9436
2:	d0_session_count+dlx_daily_sessio+dx_active_days+dx_pay_count+tier	0.9352	0.9368	0.9412
3:	dlx_session_coun+dx_active_days+dx_pay_count+dx_payer+tier	0.9266	0.9342	0.9391
4:	dlx_daily_sessio+dx_active_days+dx_pay_count+dx_payer+tier	0.9293	0.9342	0.9395
5:	dlx_daily_sessio+dlx_session_coun+dx_active_days+dx_pay_count+tier	0.9144	0.9221	0.9369
6:	dlx_daily_sessio+dlx_session_coun+dx_active_days+dx_payer+tier	0.9300	0.9311	0.9340
7:	d0_session_count+dlx_session_coun+dx_active_days+dx_pay_count+tier	0.9266	0.9341	0.9397
8:	d0_session_count+dlx_session_coun+dx_active_days+dx_payer+tier	0.9281	0.9301	0.9329
9:	d0_session_count+dx_active_days+dx_pay_count+dx_payer+tier	0.9231	0.9269	0.9303
10:	d0_session_count+dlx_daily_sessio+dlx_session_coun+dx_active_days+tier	0.9204	0.9295	0.9372
11:	d0_session_count+dlx_daily_sessio+dlx_session_coun+dx_payer+tier	0.9190	0.9261	0.9326
12:	d0_session_count+dlx_daily_sessio+dlx_session_coun+dx_pay_count+tier	0.9218	0.9288	0.9331
13:	d0_session_count+dlx_daily_sessio+dx_pay_count+dx_payer+tier	0.9190	0.9221	0.9279
14:	d0_session_count+dlx_session_coun+dx_pay_count+dx_payer+tier	0.9134	0.9230	0.9306
15:	d0_session_count+dlx_daily_sessio+dx_active_days+dx_pay_count+dx_payer	0.9176	0.9258	0.9313
16:	d0_session_count+dlx_daily_sessio+dlx_session_coun+dx_active_days+dx_payer	0.8857	0.9167	0.9459
17:	d0_session_count+dlx_daily_sessio+dlx_session_coun+dx_active_days+dx_pay_count	0.9127	0.9243	0.9342
18:	dlx_daily_sessio+dlx_session_coun+dx_pay_count+dx_payer+tier	0.9093	0.9176	0.9332
19:	d0_session_count+dlx_daily_sessio+dlx_session_coun+dx_pay_count+dx_payer	0.9047	0.9157	0.9225
20:	dlx_daily_sessio+dlx_session_coun+dx_active_days+dx_pay_count+dx_payer	0.9067	0.9209	0.9249
21:	d0_session_count+dlx_session_coun+dx_active_days+dx_pay_count+dx_payer	0.8973	0.9045	0.9169

	variables	Min.	1st Qu.	Median
	Mean	3rd Qu.	Max.	
1:	0.9414	0.9455	0.9515	
2:	0.9411	0.9447	0.9477	
3:	0.9394	0.9427	0.9550	
4:	0.9382	0.9425	0.9449	
5:	0.9379	0.9547	0.9616	
6:	0.9374	0.9413	0.9525	
7:	0.9372	0.9409	0.9440	
8:	0.9364	0.9424	0.9501	
9:	0.9359	0.9453	0.9557	
10:	0.9355	0.9391	0.9515	
11:	0.9343	0.9438	0.9500	
12:	0.9338	0.9361	0.9505	
13:	0.9321	0.9411	0.9525	
14:	0.9311	0.9388	0.9498	
15:	0.9310	0.9354	0.9453	
16:	0.9304	0.9478	0.9495	
17:	0.9301	0.9374	0.9403	
18:	0.9273	0.9363	0.9383	
19:	0.9242	0.9366	0.9410	
20:	0.9241	0.9318	0.9346	
21:	0.9194	0.9363	0.9420	
	Mean	3rd Qu.	Max.	

	variables	Min.	1st Qu.
1:	d0_session_count+dlx_daily_sessio+dlx_session_coun+dx_active_days+dx_payer+tier	0.9348	0.9365
2:	d0_session_count+dlx_daily_sessio+dx_active_days+dx_pay_count+dx_payer+tier	0.9230	0.9323
3:	d0_session_count+dlx_daily_sessio+dlx_session_coun+dx_active_days+dx_pay_count+tier	0.9227	0.9369
4:	dlx_daily_sessio+dlx_session_coun+dx_active_days+dx_pay_count+dx_payer+tier	0.9281	0.9308
5:	d0_session_count+dlx_session_coun+dx_active_days+dx_pay_count+dx_payer+tier	0.9199	0.9258
6:	d0_session_count+dlx_daily_sessio+dlx_session_coun+dx_pay_count+dx_payer+tier	0.9176	0.9263
7:	d0_session_count+dlx_daily_sessio+dlx_session_coun+dx_active_days+dx_pay_count+dx_payer	0.9106	0.9262

	Median	Mean	3rd Qu.	Max.
1:	0.9402	0.9418	0.9447	0.9543
2:	0.9427	0.9416	0.9530	0.9561
3:	0.9393	0.9405	0.9480	0.9548
4:	0.9374	0.9380	0.9389	0.9568
5:	0.9354	0.9360	0.9467	0.9524
6:	0.9361	0.9336	0.9418	0.9452
7:	0.9339	0.9302	0.9370	0.9411

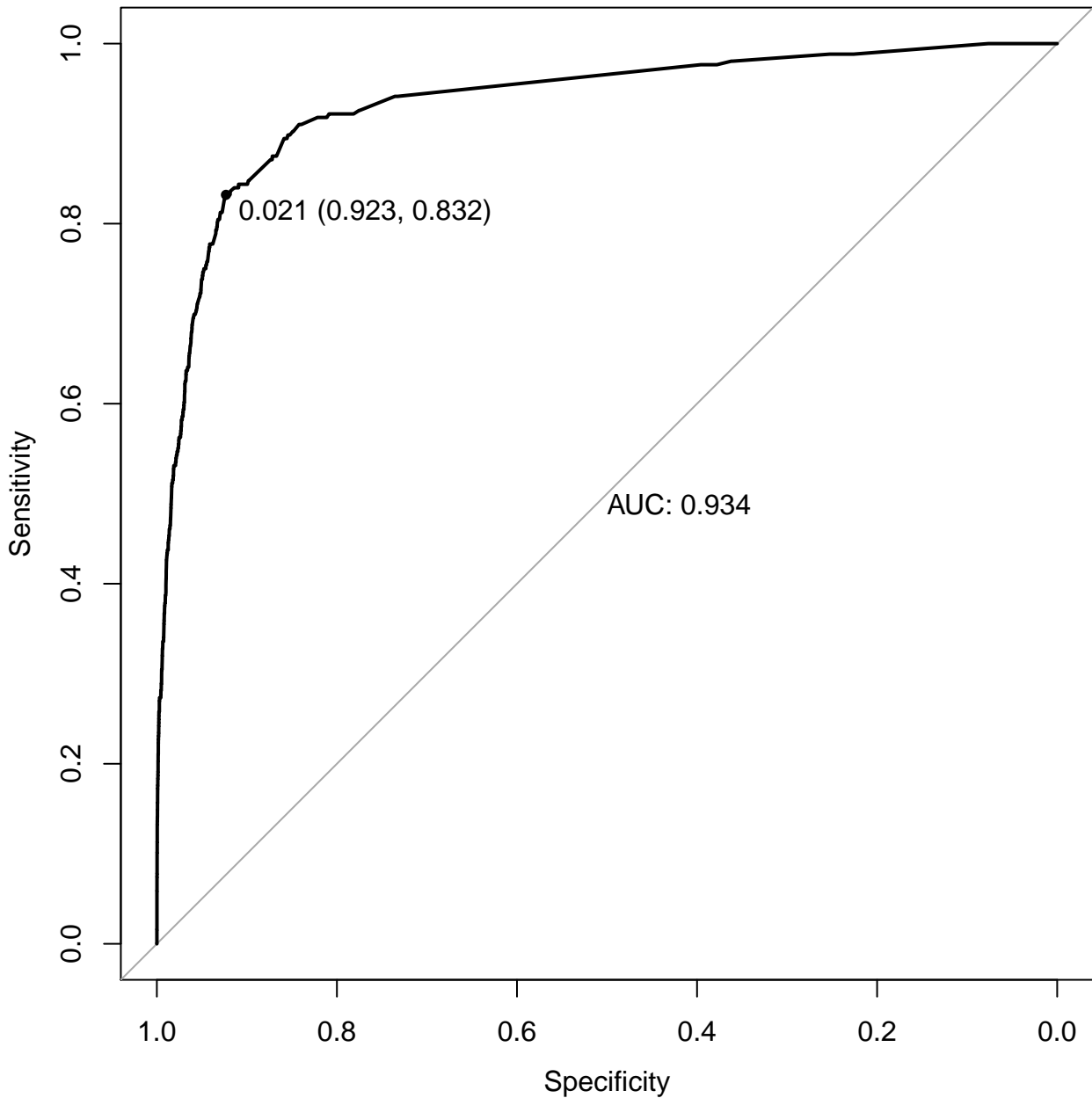
	variables	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1:	d0_session_count+dx_active_days+dx_pay_count+tier	0.9285	0.932	0.9363	0.9358	0.9376	0.9448

	variables	Min.	1st Qu.	Median	Mean	3rd Qu.
1:	d0_session_count+dlx_daily_sessio+dx_active_days+dx_pay_count+tier	0.9297	0.9349	0.9412	0.9419	0.9459
	Max.					
1:	0.959					

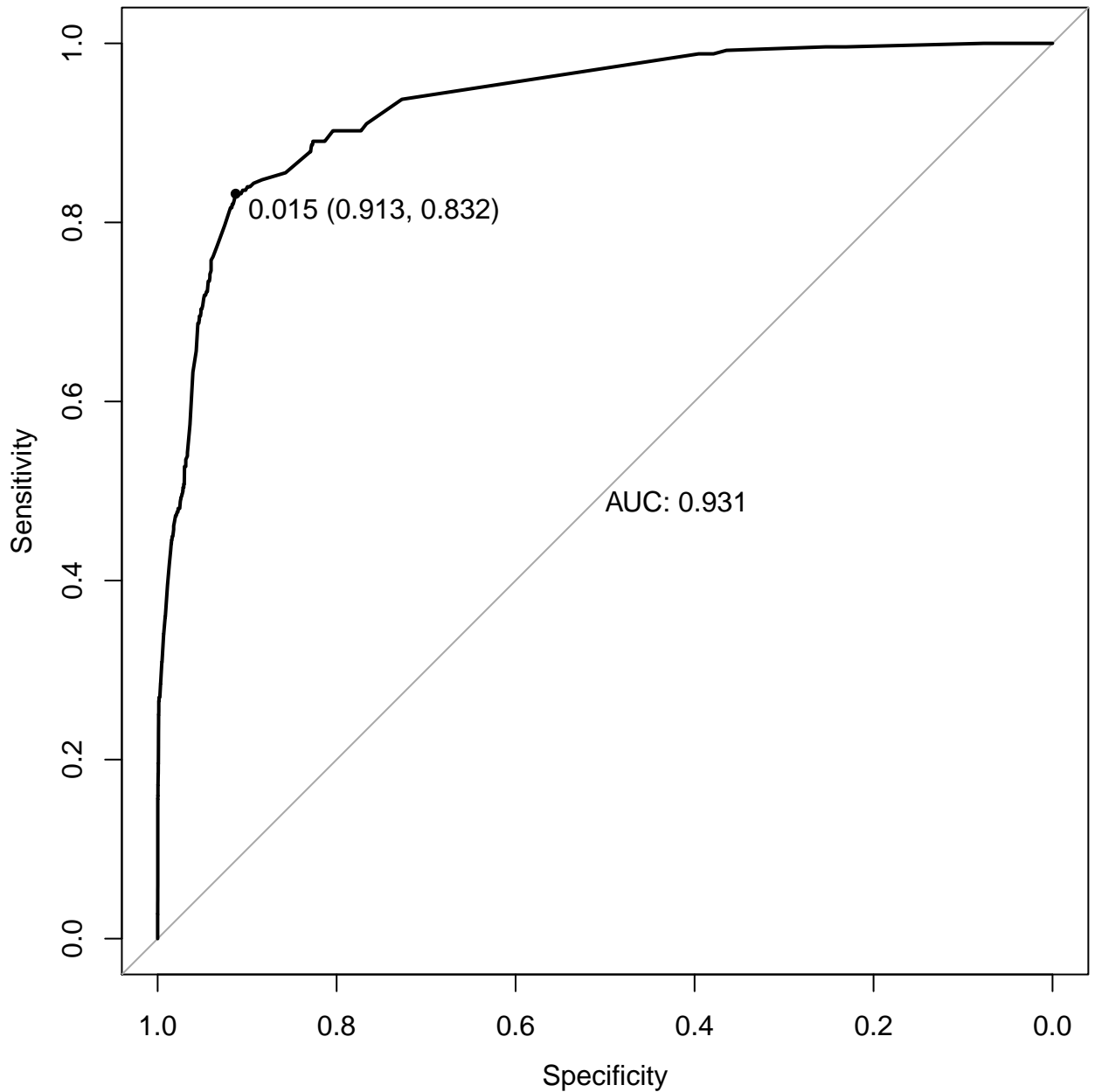
```

                                variables  Min. 1st Qu.
1: d0_session_count+dlx_daily_sessio+dlx_session_coun+dx_active_days+dx_pay_count+dx_payer+tier 0.9324 0.9361
   Median   Mean 3rd Qu.   Max.
1: 0.9444 0.9424 0.9474 0.9511
```


ROC curve of full model (7 variables)



ROC curve of simplified model (4 variables)



```
[1] "I will choose the simplified model with 4 variables for now:"
[2] " - tier"
[3] " - dx_pay_count"
[4] " - dx_active_days"
[5] " - d0_session_count"
[6] "The difference in AUC is insignificant compared to full v2 model."
[7] "For more detailed comments, see code:"
[8] " -> 7_logit_v2_feature_selection"
[9] "This was calculated on DA - Google play - marketing"
[10] "But we got very similar results on iOS too."
[11] " - tier is not as good a predictor on iOS"
[12] " - dlx_daily_session_count_relative_to_d0 looks more promising on iOS"
```