CRIC Autoplier Survival Analysis

Setup

Download Data

Sanity Check Data

	patientid	mu_egfr_rate	sigma_egfr_rate	acearb_v3y0	african_ancestry_v3y0	age_integer_v3y0	albuminuria_v3y0	apol1_risk_alleles_v3y0
0	1010001	-0.068201	0.188699	1	6.6	70	1	0
1	1010007	-0.105450	0.199372	1	82.2	69	3	1
2	1010008	-0.012396	0.218453	1	80.7	68	2	1
3	1010013	-0.054251	0.244502	1	84.3	61	3	2
4	1010015	-0.290542	0.361062	1	10.4	49	4	0

5 rows × 131 columns

Formatting

Partition Data

Omics to pathway transformation

Autoplier

Autoplier is an unsupervised dimension reduction method that transforms the longitudinal metabolomics data into the pathway representation.

Setup

Estimate number of latent variables

```
Number of Latent Variables is 42
```

Estimate regularization constant

```
targetsparsity is 0.9964349376114082


L1 regularization value is 1e-07
```

Train autoplier

```
2022-03-25 16:35:55.100740: I tensorflow/core/platform/cpu_feature_guard.cc:151] This TensorFlow binary is optimized with
oneAPI Deep Neural Network Library (oneDNN) to use the following CPU instructions in performance-critical operations:  AV
X2 FMA
To enable them in other operations, rebuild TensorFlow with the appropriate compiler flags.
Trained Model Sparsity is 0.9738955823293173
```

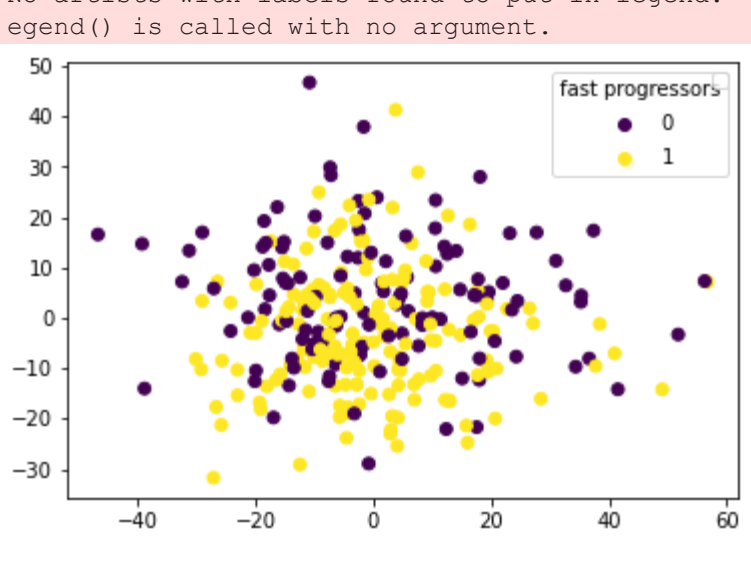
Association of Latent variables with eGFR decline

We define a logistic regression classifier with optional adjustment for clinical variables and return the strongest predictive pathways for the positive and negative cases.

Full cohort analysis

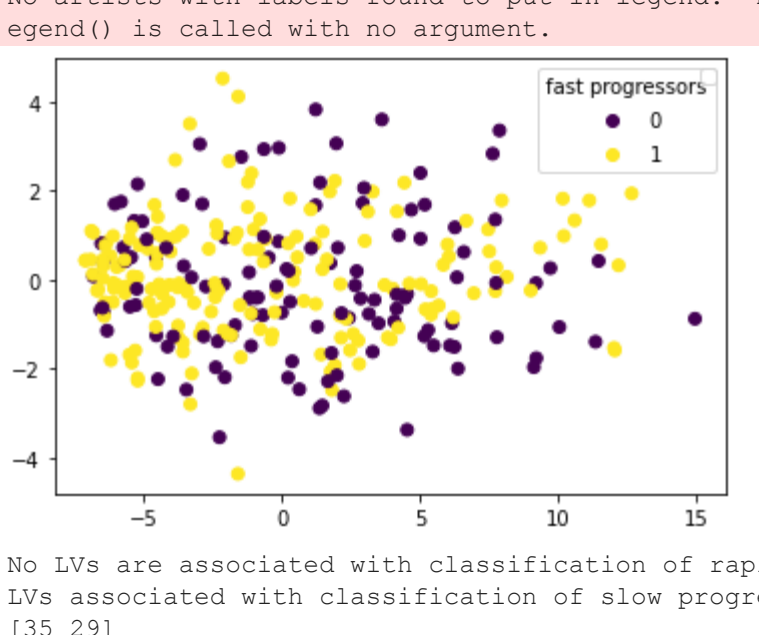
benchmark model

```
No artists with labels found to put in legend.  Note that artists whose label start with an underscore are ignored when l
egend() is called with no argument.
```

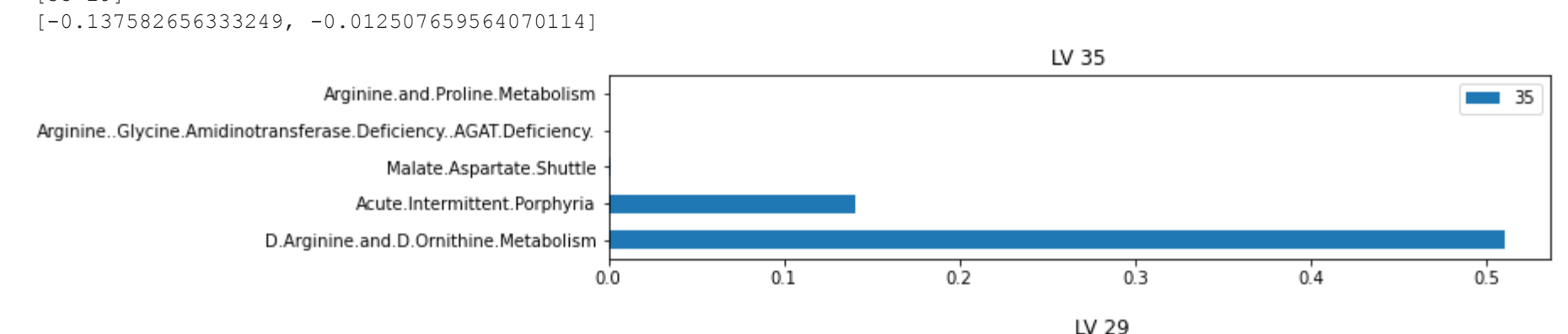


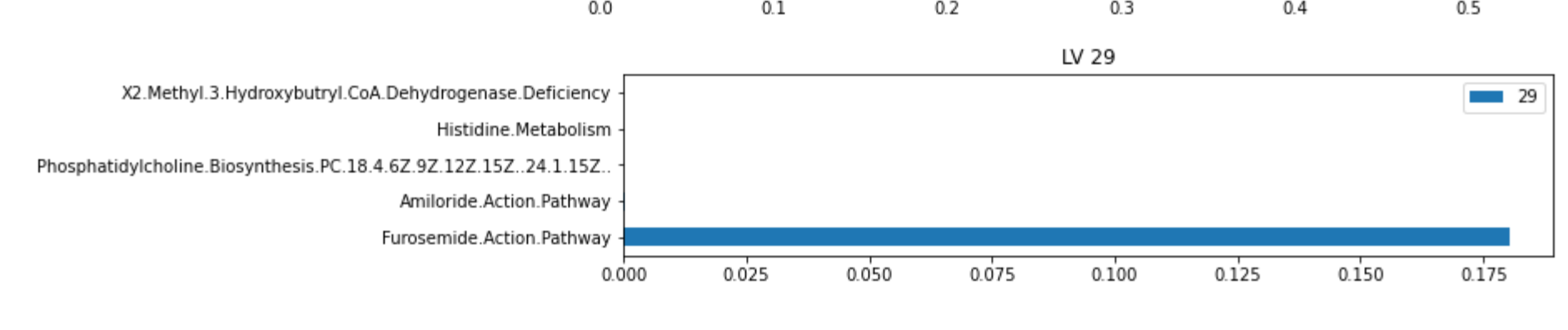
Top latent variables for predicting rapid eGFR decline unadjusted

```
No artists with labels found to put in legend.  Note that artists whose label start with an underscore are ignored when l
egend() is called with no argument.
```



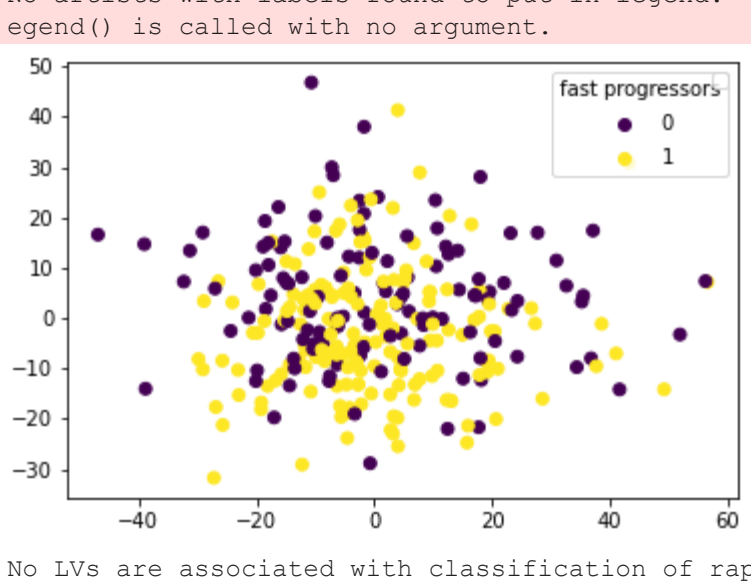

```
No LVs are associated with classification of rapid progressors!
LVs associated with classification of slow progressors :
[35 29]
[-0.137582656333249, -0.012507659564070114]
```





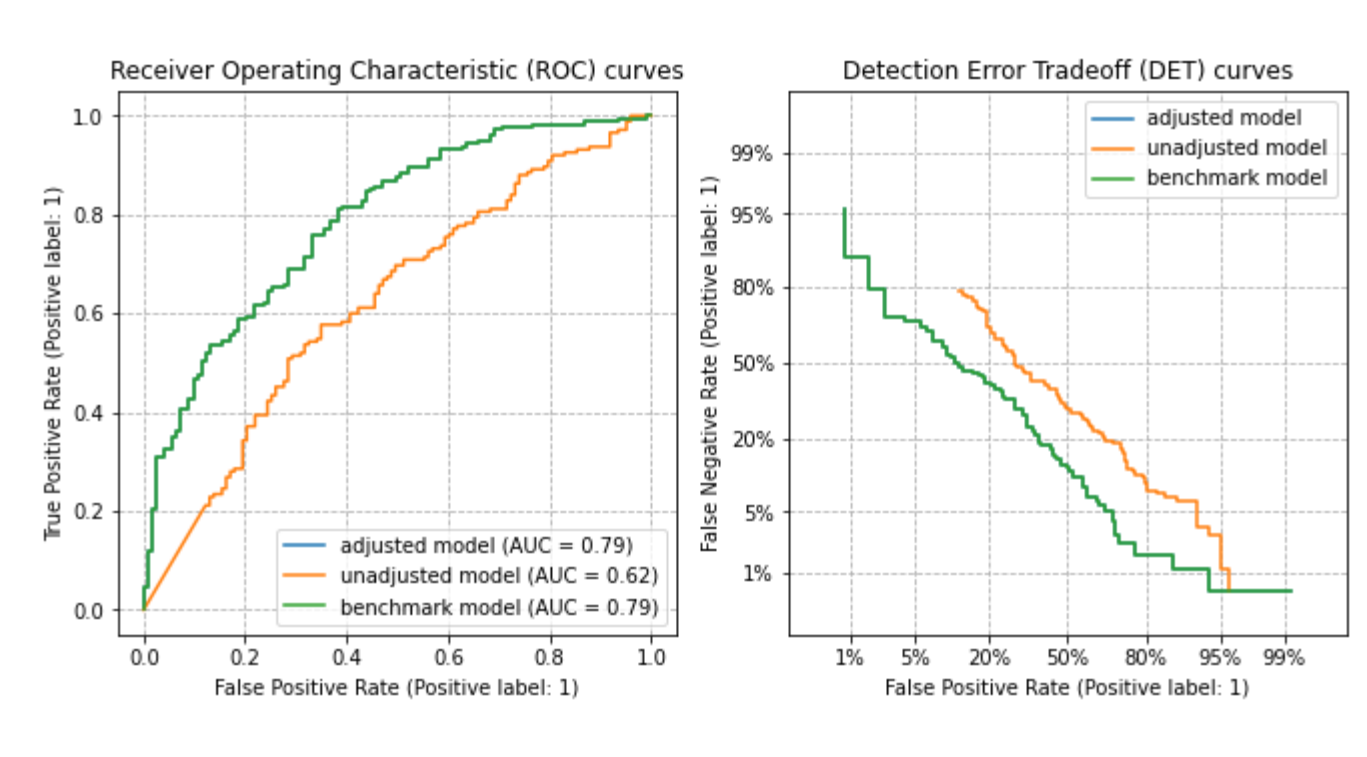
Top latent variables for predicting rapid eGFR decline with adjustment for clinical variables

```
No artists with labels found to put in legend.  Note that artists whose label start with an underscore are ignored when l
egend() is called with no argument.
```




```
No LVs are associated with classification of rapid progressors!
No LVs are associated with classification of slow progressors!
```

Comparison of ROC curves between adjusted and unadjusted models



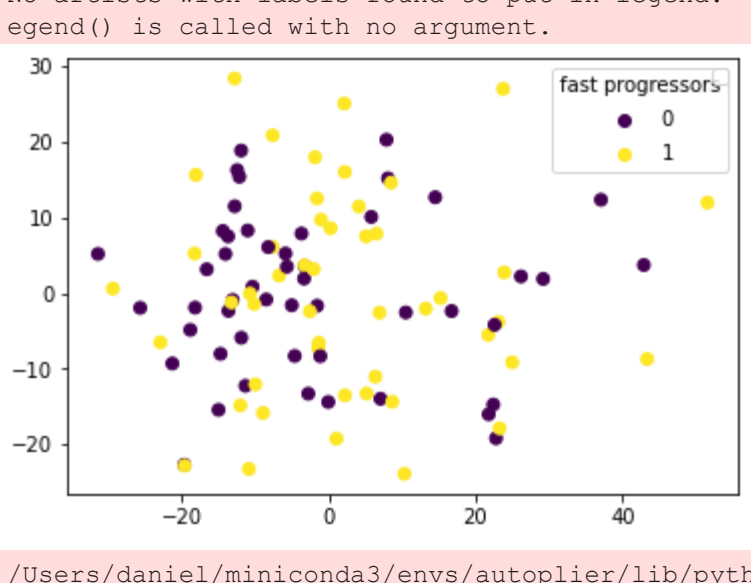
Normoalbuminuria group analysis

```
/var/folders/v3/d9c1pzwx01d7sf9x7_4wzy4c0000gq/T/ipykernel_46964/186333728.py:12: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-v
iew-versus-a-copy
  clinred['log2acr'] = np.log2(clinred['urine_albumin_v3y0']+1) - np.log2(clinred['urine_creatinine_v3y0']+1)
```

benchmark model

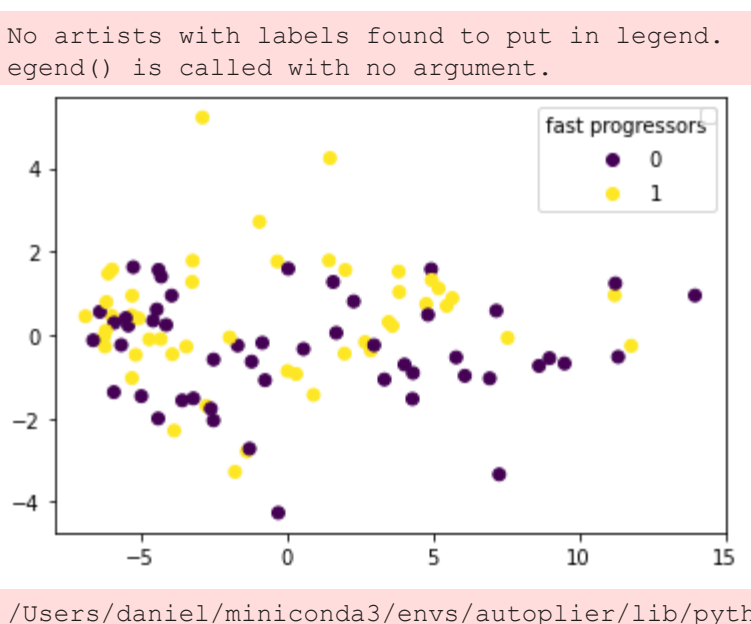
```
No artists with labels found to put in legend.  Note that artists whose label start with an underscore are ignored when l
egend() is called with no argument.
```



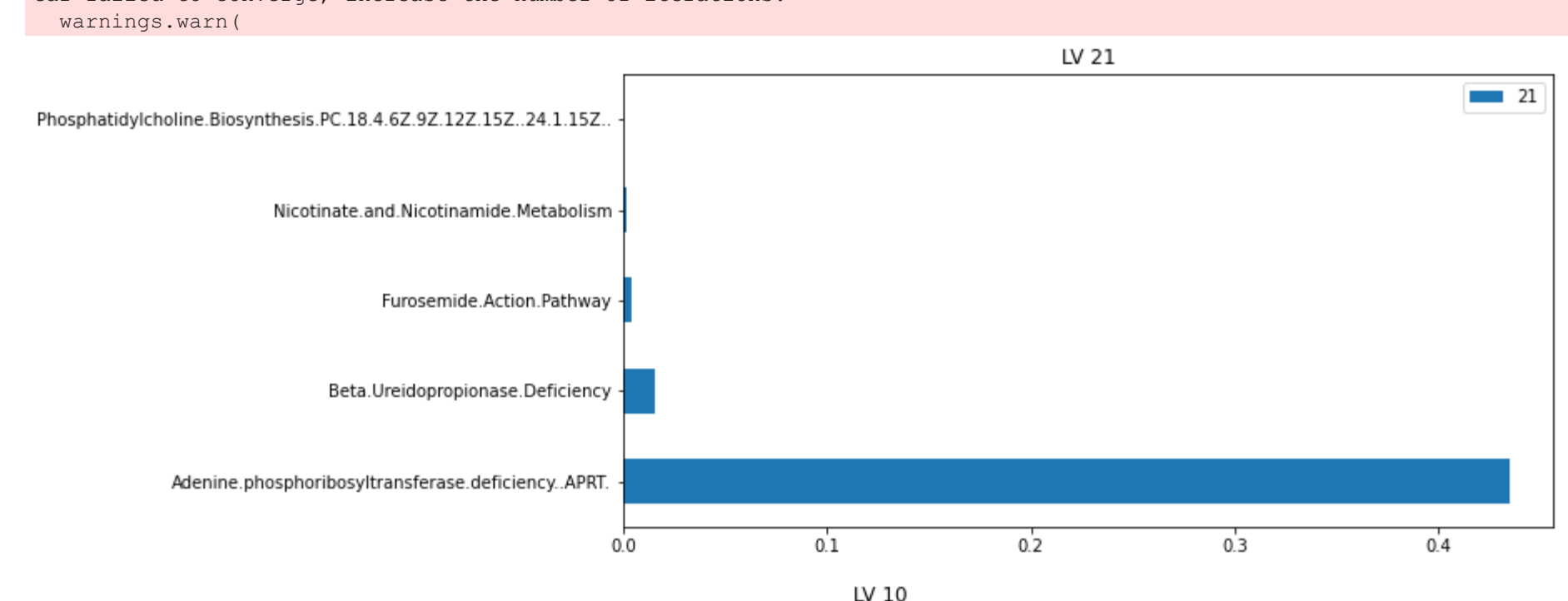

```
/Users/daniel/miniconda3/envs/autoplier/lib/python3.9/site-packages/sklearn/svm/_base.py:1206: ConvergenceWarning: Liblin
ear failed to converge, increase the number of iterations.
  warnings.warn(
/Users/daniel/miniconda3/envs/autoplier/lib/python3.9/site-packages/sklearn/svm/_base.py:1206: ConvergenceWarning: Liblin
ear failed to converge, increase the number of iterations.
  warnings.warn(
/Users/daniel/miniconda3/envs/autoplier/lib/python3.9/site-packages/sklearn/svm/_base.py:1206: ConvergenceWarning: Liblin
ear failed to converge, increase the number of iterations.
  warnings.warn(
/Users/daniel/miniconda3/envs/autoplier/lib/python3.9/site-packages/sklearn/svm/_base.py:1206: ConvergenceWarning: Liblin
ear failed to converge, increase the number of iterations.
  warnings.warn(
/Users/daniel/miniconda3/envs/autoplier/lib/python3.9/site-packages/sklearn/svm/_base.py:1206: ConvergenceWarning: Liblin
ear failed to converge, increase the number of iterations.
  warnings.warn(
/Users/daniel/miniconda3/envs/autoplier/lib/python3.9/site-packages/sklearn/svm/_base.py:1206: ConvergenceWarning: Liblin
ear failed to converge, increase the number of iterations.
  warnings.warn(
/Users/daniel/miniconda3/envs/autoplier/lib/python3.9/site-packages/sklearn/svm/_base.py:1206: ConvergenceWarning: Liblin
ear failed to converge, increase the number of iterations.
  warnings.warn(
/Users/daniel/miniconda3/envs/autoplier/lib/python3.9/site-packages/sklearn/svm/_base.py:1206: ConvergenceWarning: Liblin
ear failed to converge, increase the number of iterations.
  warnings.warn(
/Users/daniel/miniconda3/envs/autoplier/lib/python3.9/site-packages/sklearn/svm/_base.py:1206: ConvergenceWarning: Liblin
ear failed to converge, increase the number of iterations.
  warnings.warn(
/Users/daniel/miniconda3/envs/autoplier/lib/python3.9/site-packages/sklearn/svm/_base.py:1206: ConvergenceWarning: Liblin
ear failed to converge, increase the number of iterations.
  warnings.warn(
/Users/daniel/miniconda3/envs/autoplier/lib/python3.9/site-packages/sklearn/svm/_base.py:1206: ConvergenceWarning: Liblin
ear failed to converge, increase the number of iterations.
  warnings.warn(
/Users/daniel/miniconda3/envs/autoplier/lib/python3.9/site-packages/sklearn/svm/_base.py:1206: ConvergenceWarning: Liblin
ear failed to converge, increase the number of iterations.
  warnings.warn(
/Users/daniel/miniconda3/envs/autoplier/lib/python3.9/site-packages/sklearn/svm/_base.py:1206: ConvergenceWarning: Liblin
ear failed to converge, increase the number of iterations.
  warnings.warn(
/Users/daniel/miniconda3/envs/autoplier/lib/python3.9/site-packages/sklearn/svm/_base.py:1206: ConvergenceWarning: Liblin
ear failed to converge, increase the number of iterations.
  warnings.warn(
```

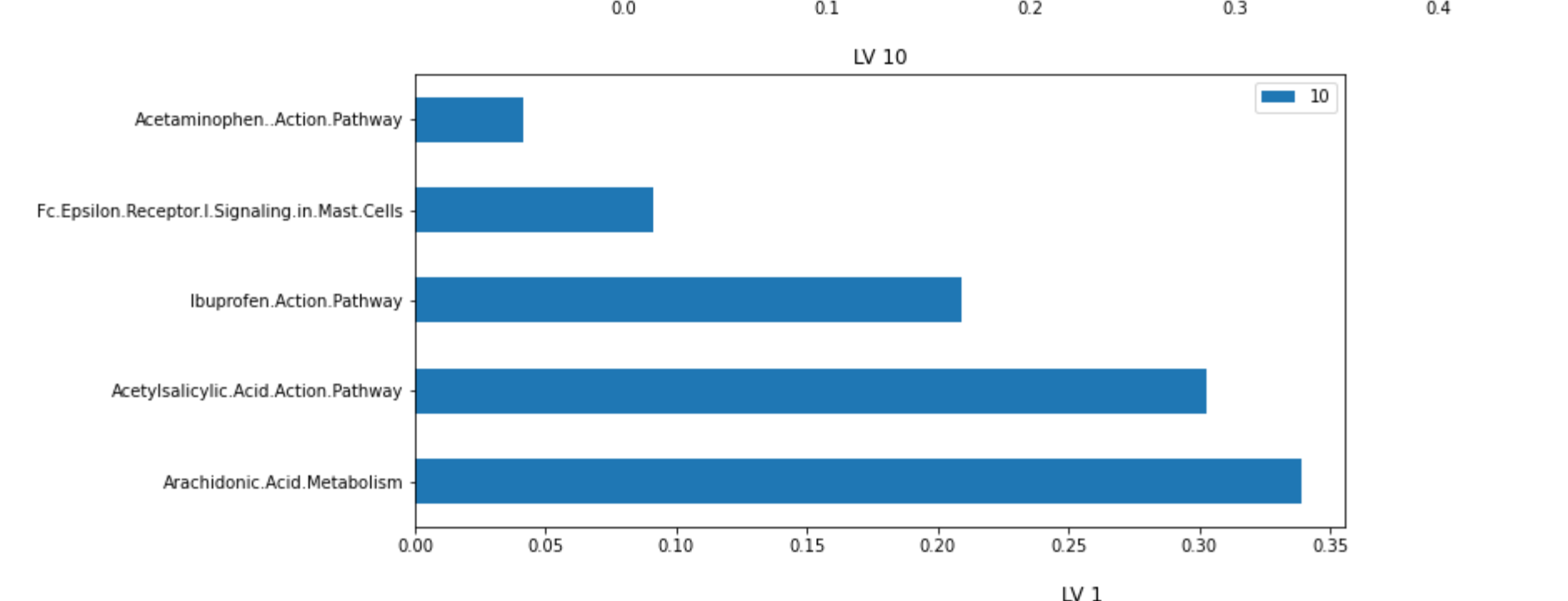
Top latent variables for predicting rapid eGFR decline unadjusted

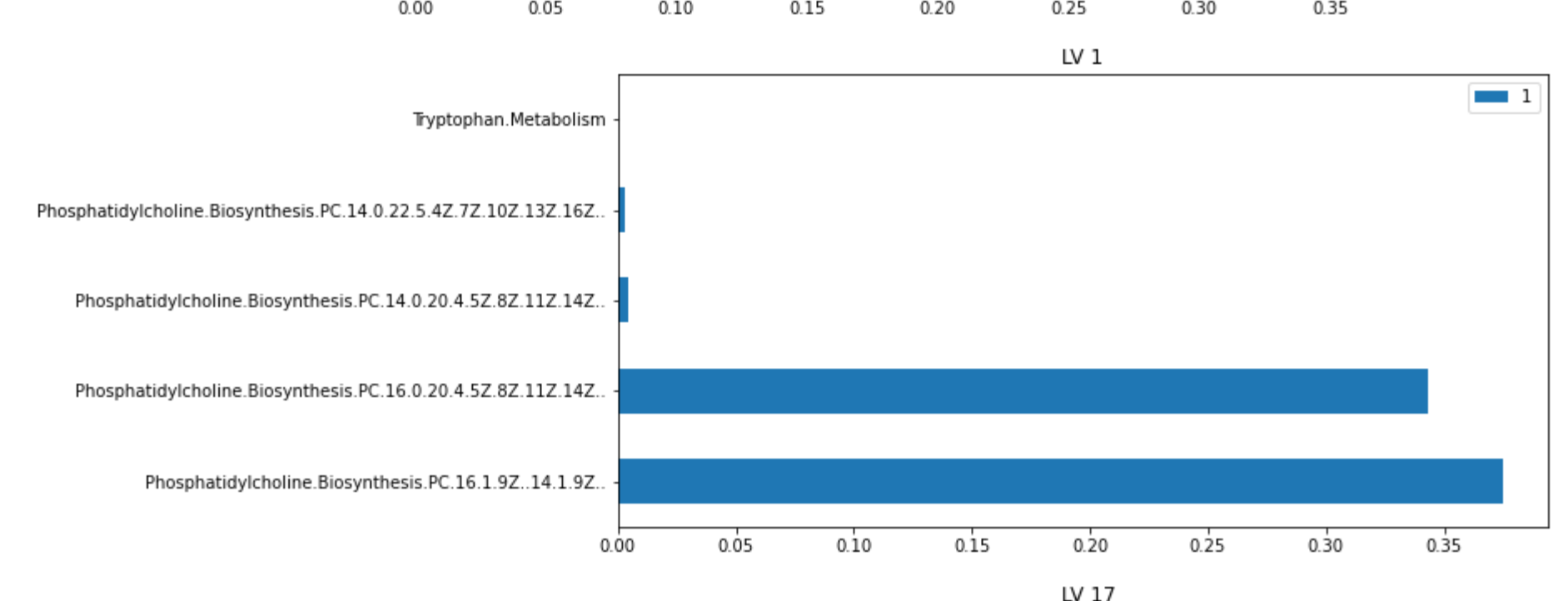
```
No artists with labels found to put in legend.  Note that artists whose label start with an underscore are ignored when l
egend() is called with no argument.
```

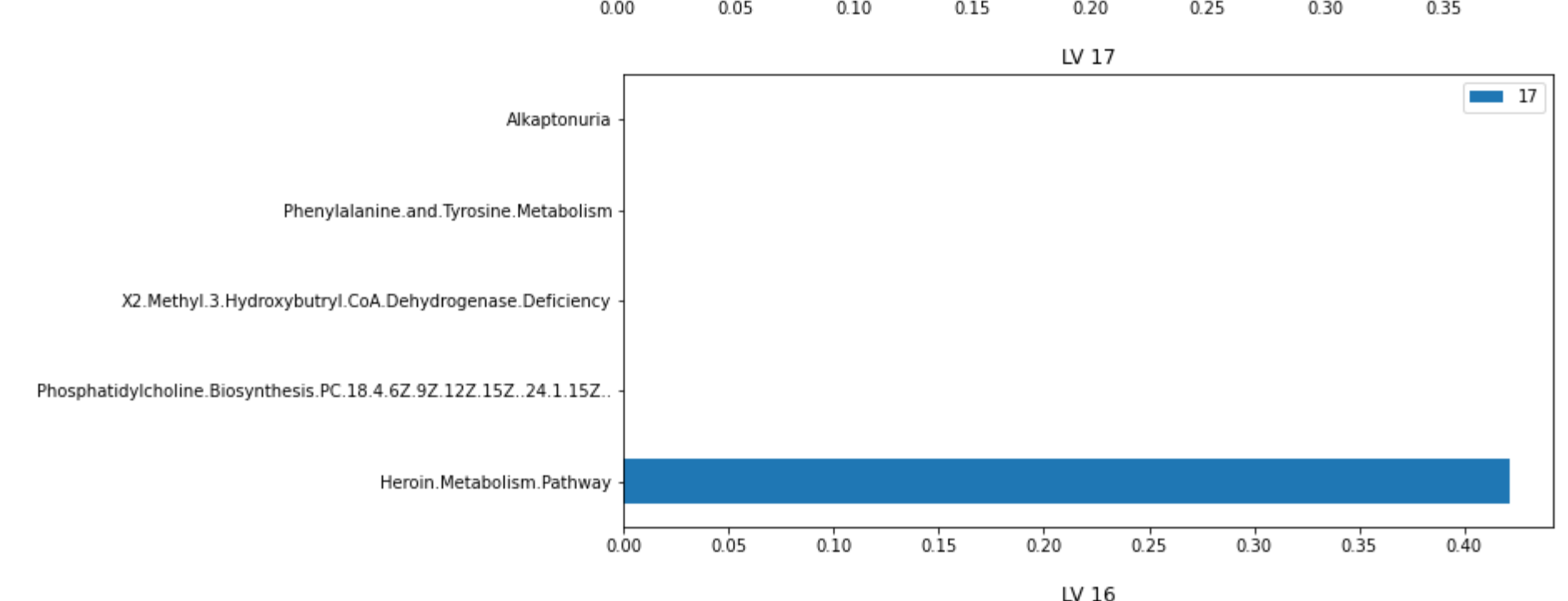


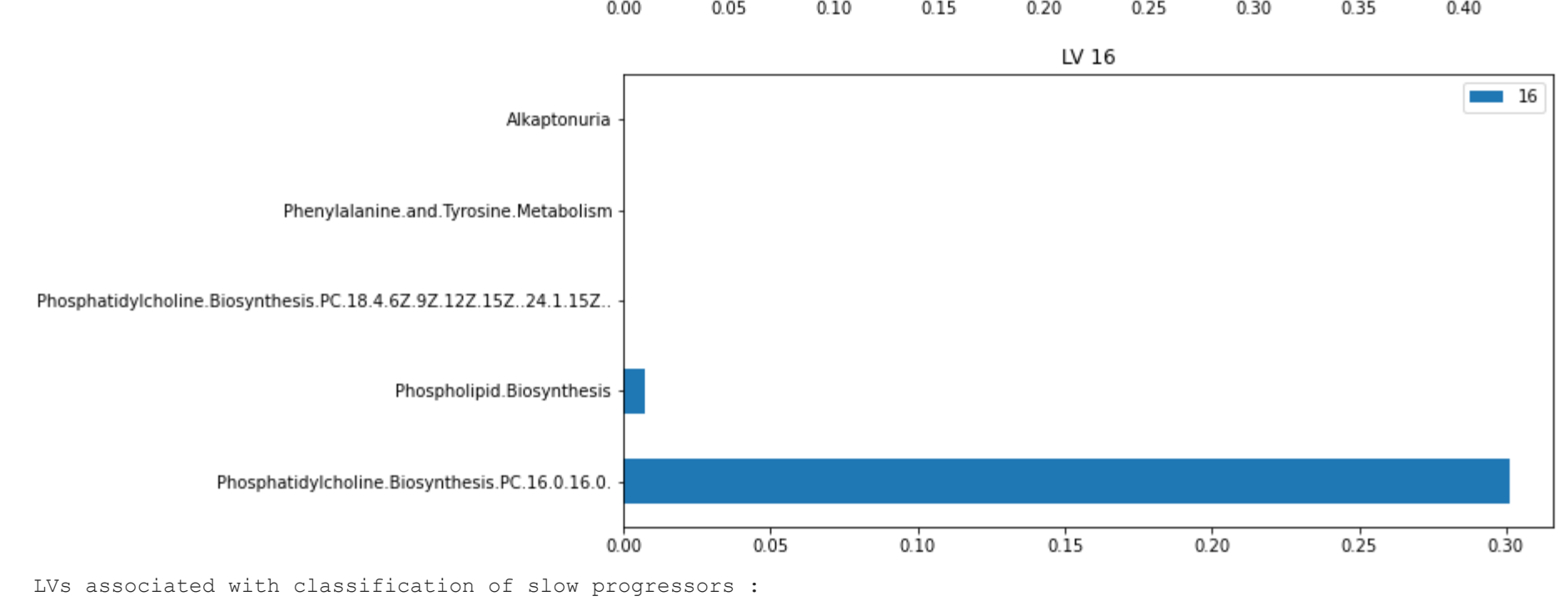

```
/Users/daniel/miniconda3/envs/autoplier/lib/python3.9/site-packages/sklearn/svm/_base.py:1206: ConvergenceWarning: Liblin
ear failed to converge, increase the number of iterations.
  warnings.warn(
/Users/daniel/miniconda3/envs/autoplier/lib/python3.9/site-packages/sklearn/svm/_base.py:1206: ConvergenceWarning: Liblin
ear failed to converge, increase the number of iterations.
  warnings.warn(
/Users/daniel/miniconda3/envs/autoplier/lib/python3.9/site-packages/sklearn/svm/_base.py:1206: ConvergenceWarning: Liblin
ear failed to converge, increase the number of iterations.
  warnings.warn(
/Users/daniel/miniconda3/envs/autoplier/lib/python3.9/site-packages/sklearn/svm/_base.py:1206: ConvergenceWarning: Liblin
ear failed to converge, increase the number of iterations.
  warnings.warn(
/Users/daniel/miniconda3/envs/autoplier/lib/python3.9/site-packages/sklearn/svm/_base.py:1206: ConvergenceWarning: Liblin
ear failed to converge, increase the number of iterations.
  warnings.warn(
/Users/daniel/miniconda3/envs/autoplier/lib/python3.9/site-packages/sklearn/svm/_base.py:1206: ConvergenceWarning: Liblin
ear failed to converge, increase the number of iterations.
  warnings.warn(
/Users/daniel/miniconda3/envs/autoplier/lib/python3.9/site-packages/sklearn/svm/_base.py:1206: ConvergenceWarning: Liblin
ear failed to converge, increase the number of iterations.
  warnings.warn(
LVs associated with classification of rapid progressors :
[21 10  1 17 16]
[3.0537751998562066, 3.109404178273325, 3.699788940346257, 5.252856810332323, 6.059349369849345]
/Users/daniel/miniconda3/envs/autoplier/lib/python3.9/site-packages/sklearn/svm/_base.py:1206: ConvergenceWarning: Liblin
ear failed to converge, increase the number of iterations.
  warnings.warn(
```



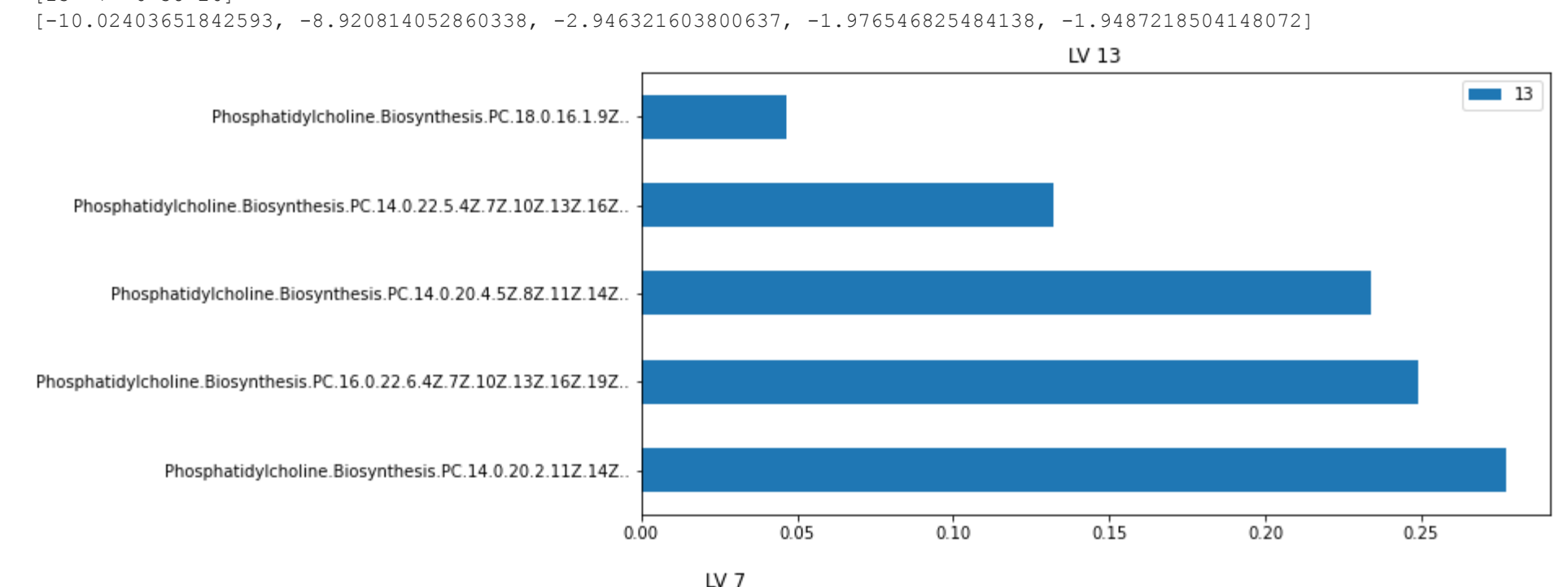


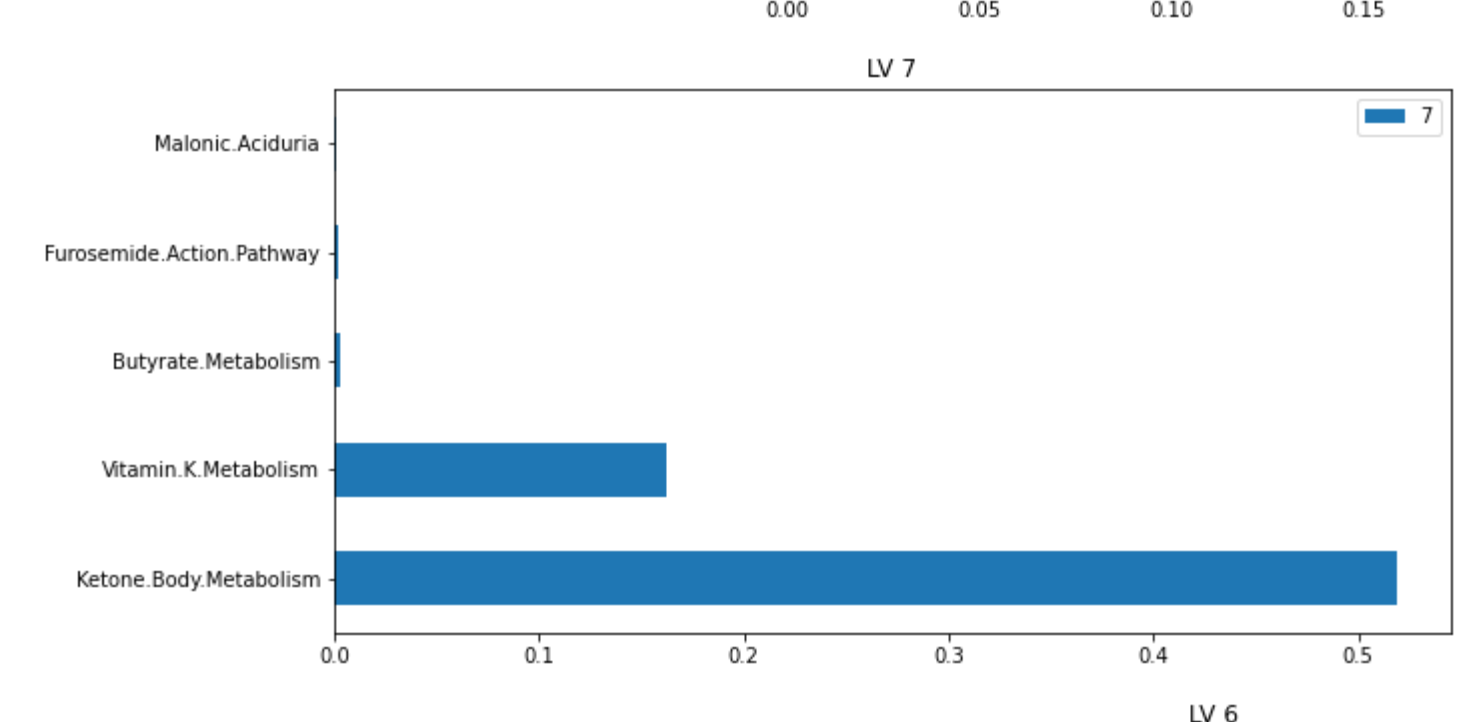


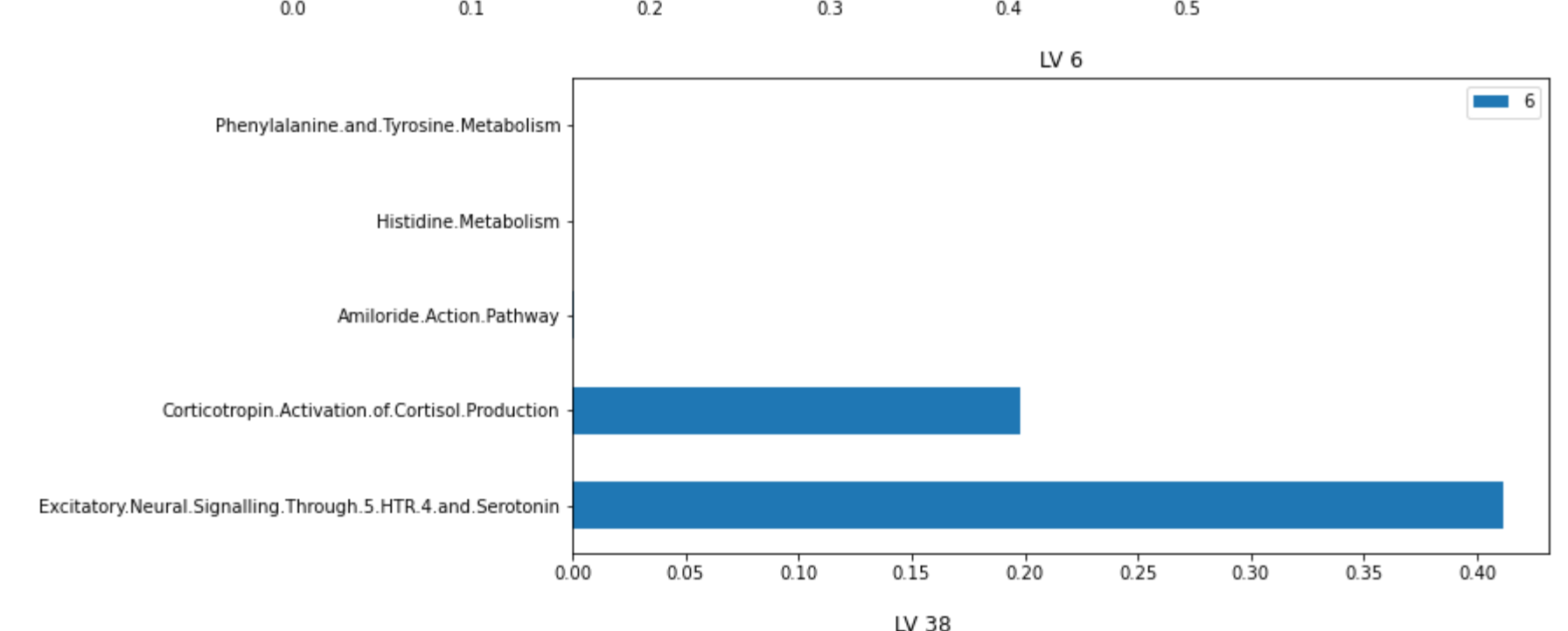


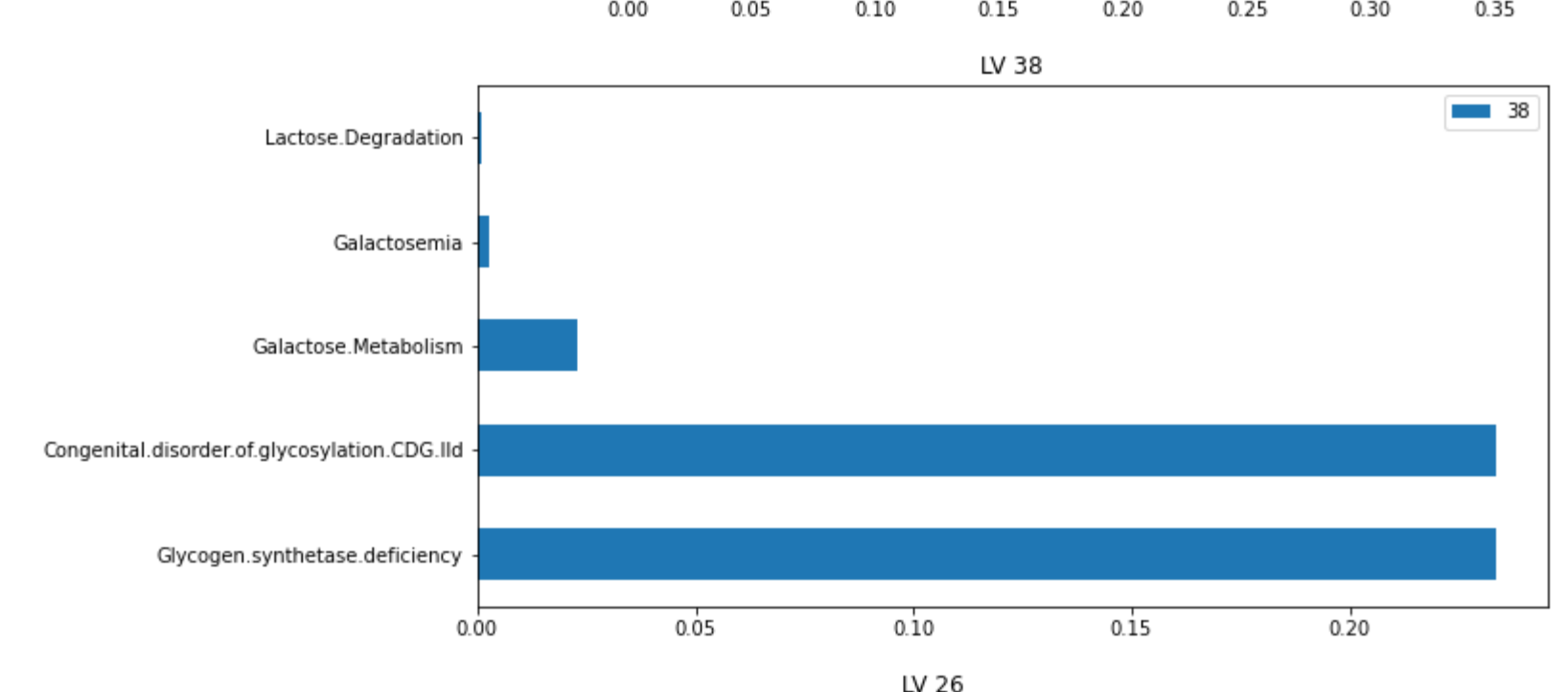


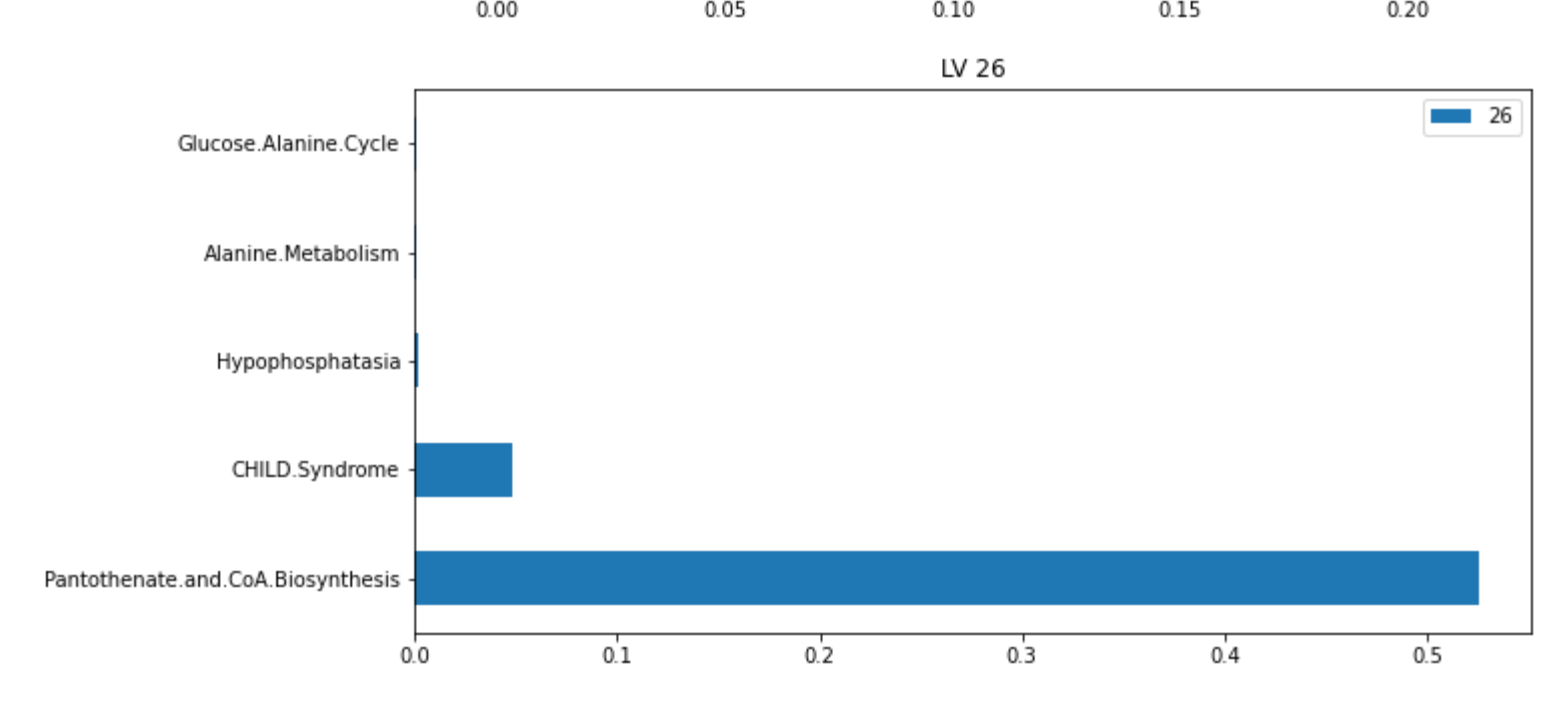

```
LVs associated with classification of slow progressors :
[13  7  6 38 26]
[-10.02403651842593, -8.920814052860338, -2.946321603800637, -1.976546825484138, -1.9487218504148072]
```





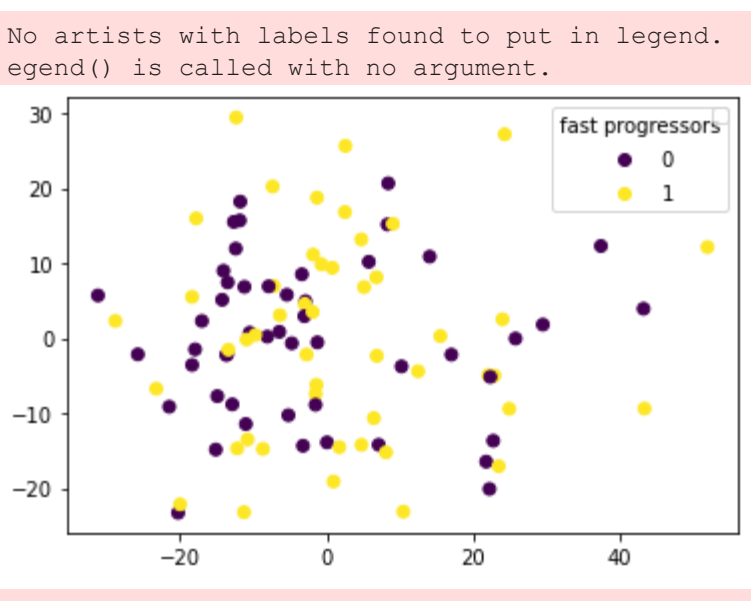




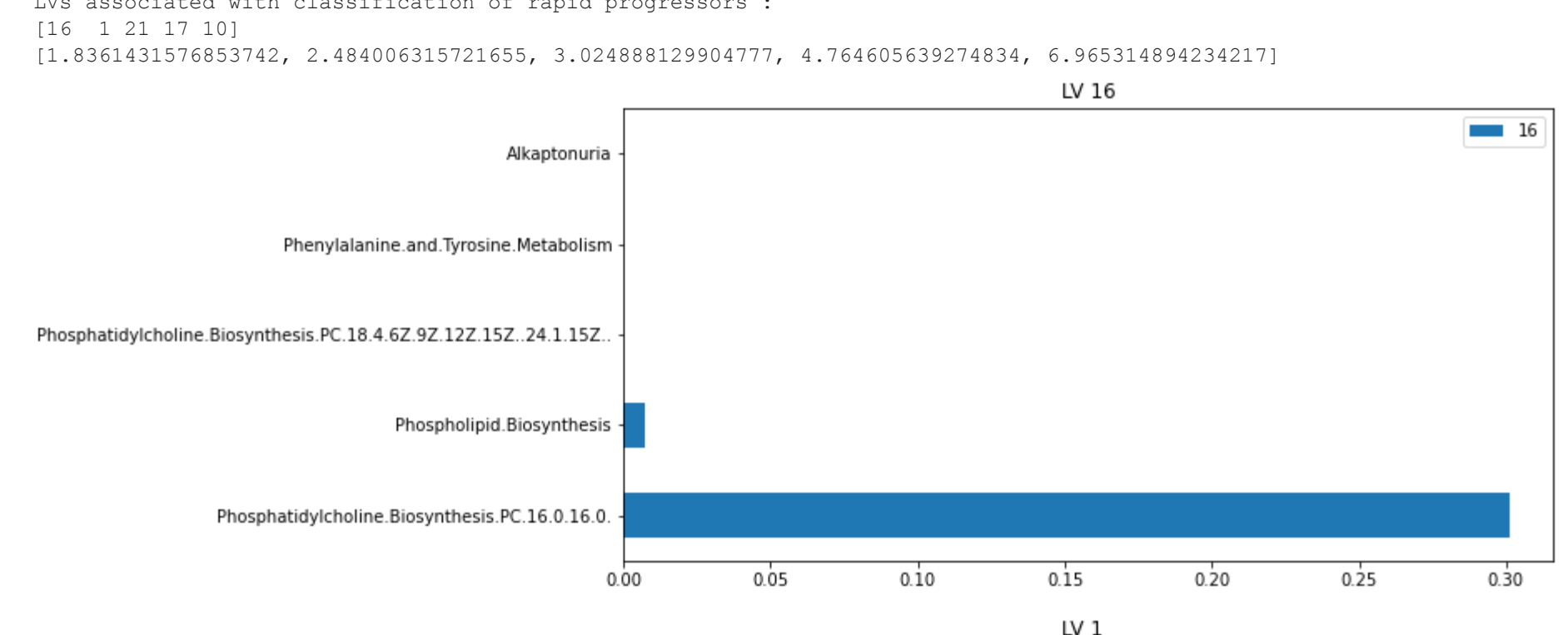


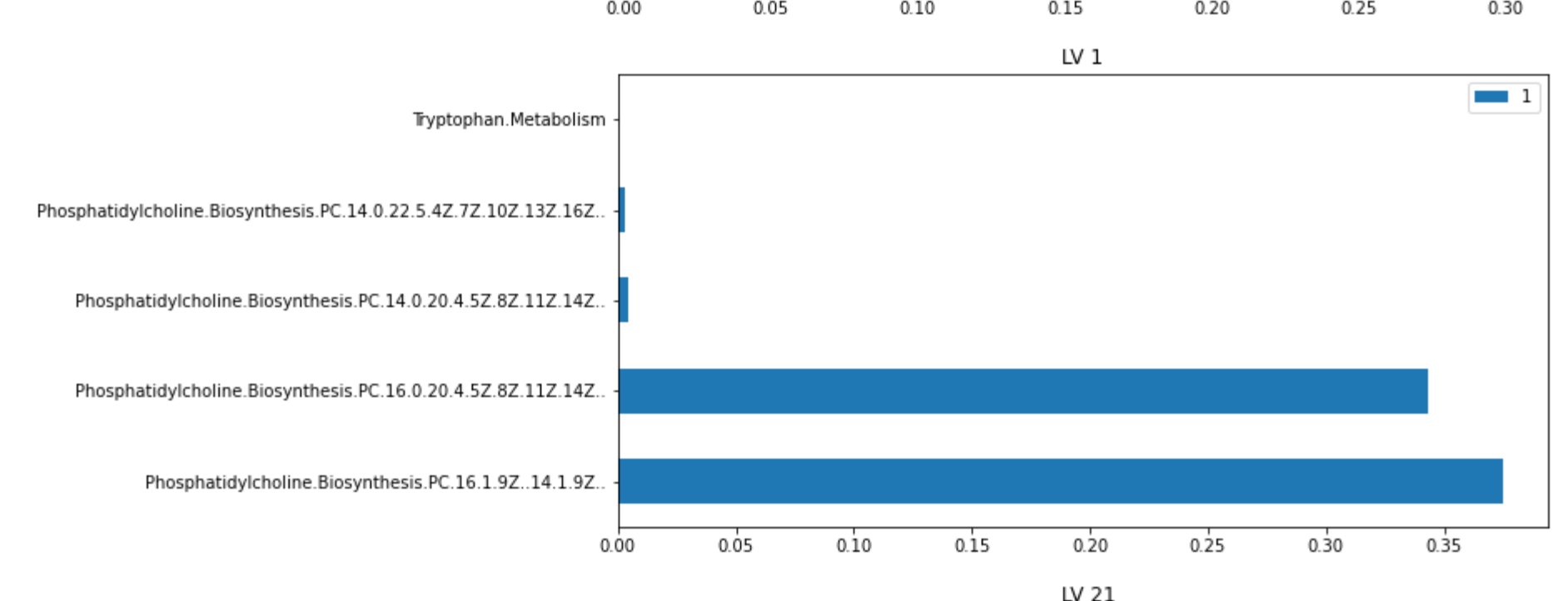
Top latent variables for predicting rapid eGFR decline with adjustment for clinical variables

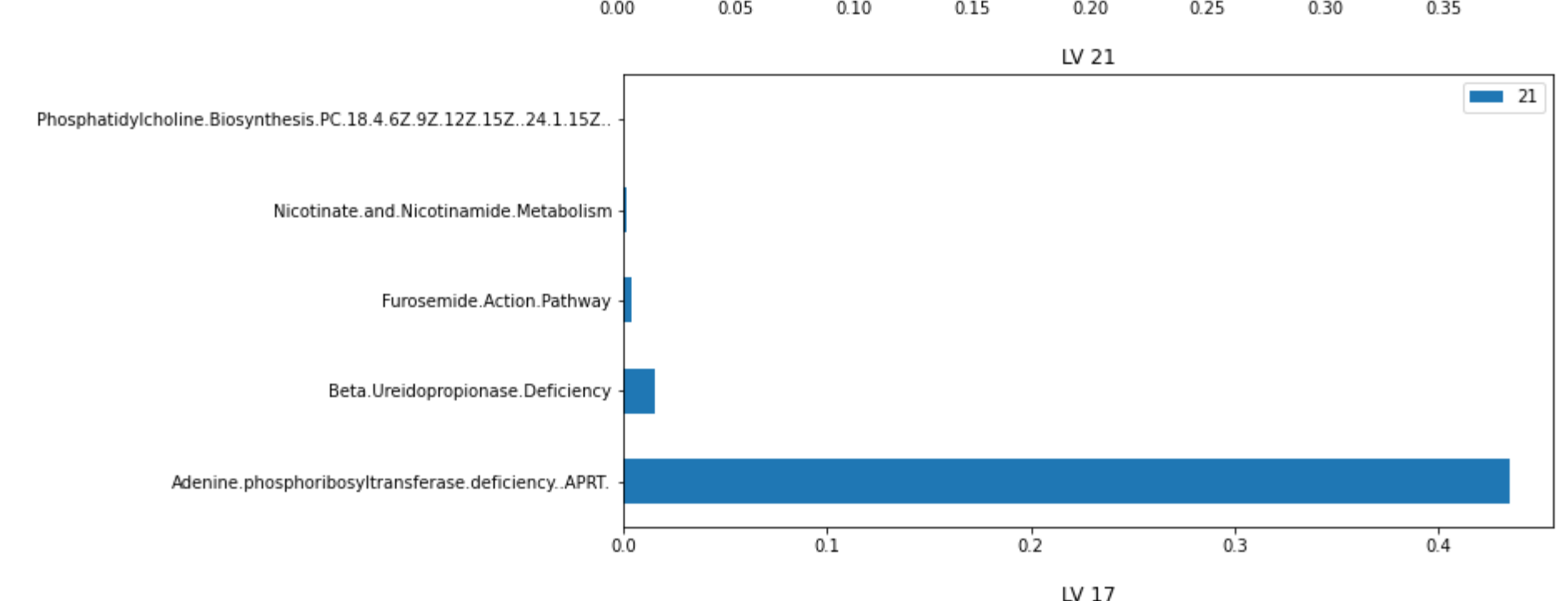
```
No artists with labels found to put in legend.  Note that artists whose label start with an underscore are ignored when l
egend() is called with no argument.
```

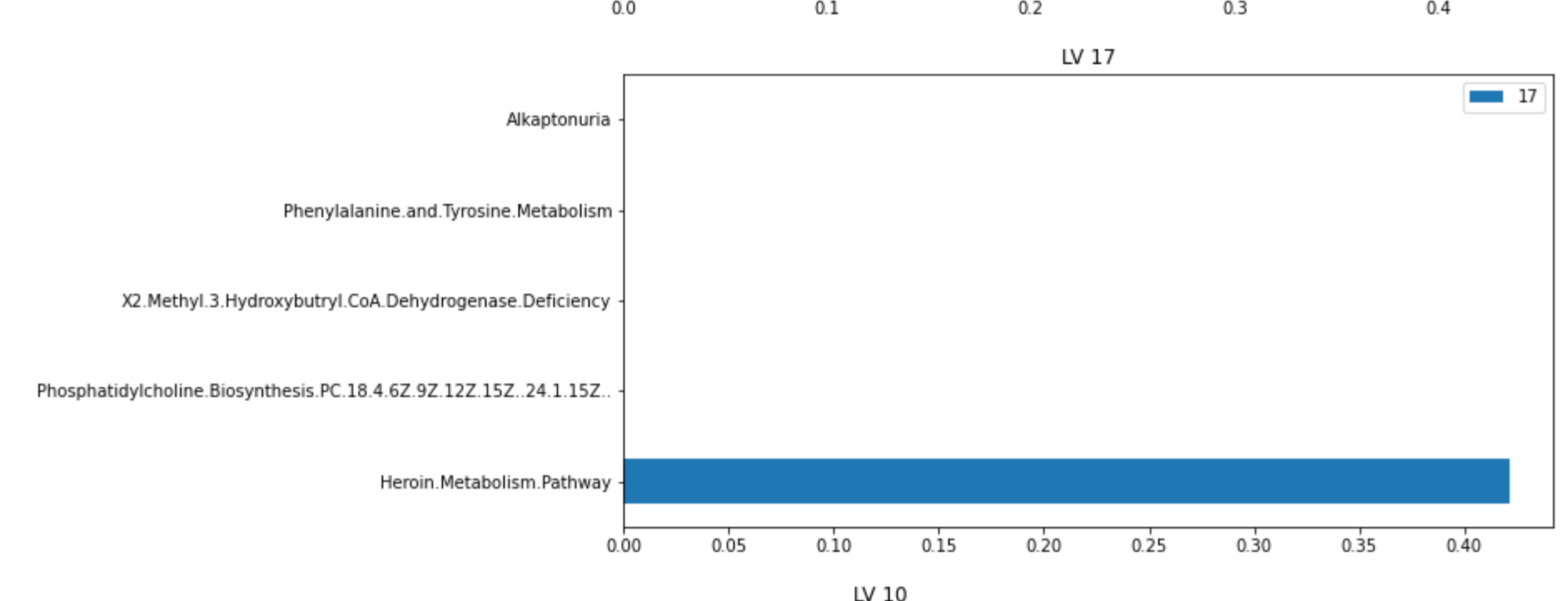


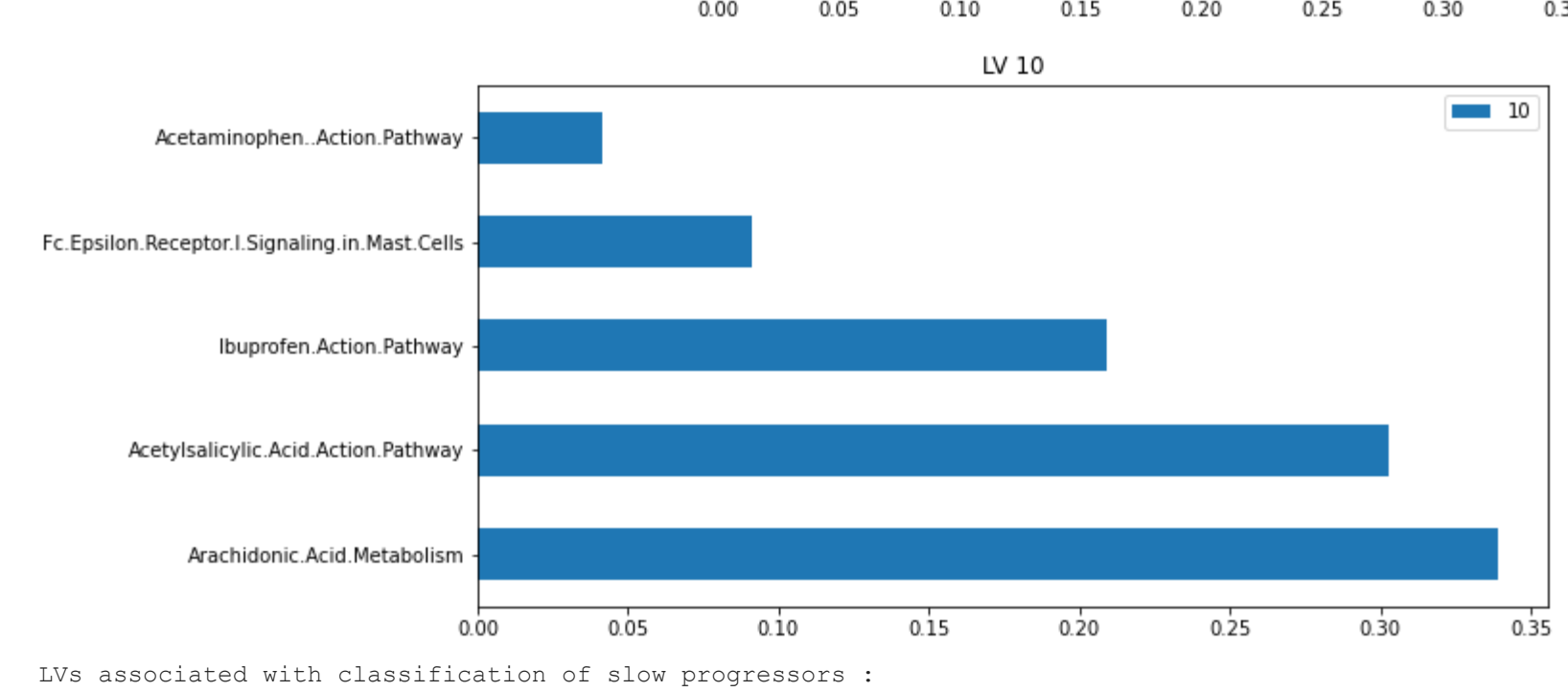

```
/Users/daniel/miniconda3/envs/autoplier/lib/python3.9/site-packages/sklearn/svm/_base.py:1206: ConvergenceWarning: Liblin
ear failed to converge, increase the number of iterations.
  warnings.warn(
LVs associated with classification of rapid progressors :
[16  1 21 17 10]
[1.8361431576853742, 2.484006315721655, 3.024888129904777, 4.764605639274834, 6.965314894234217]
```



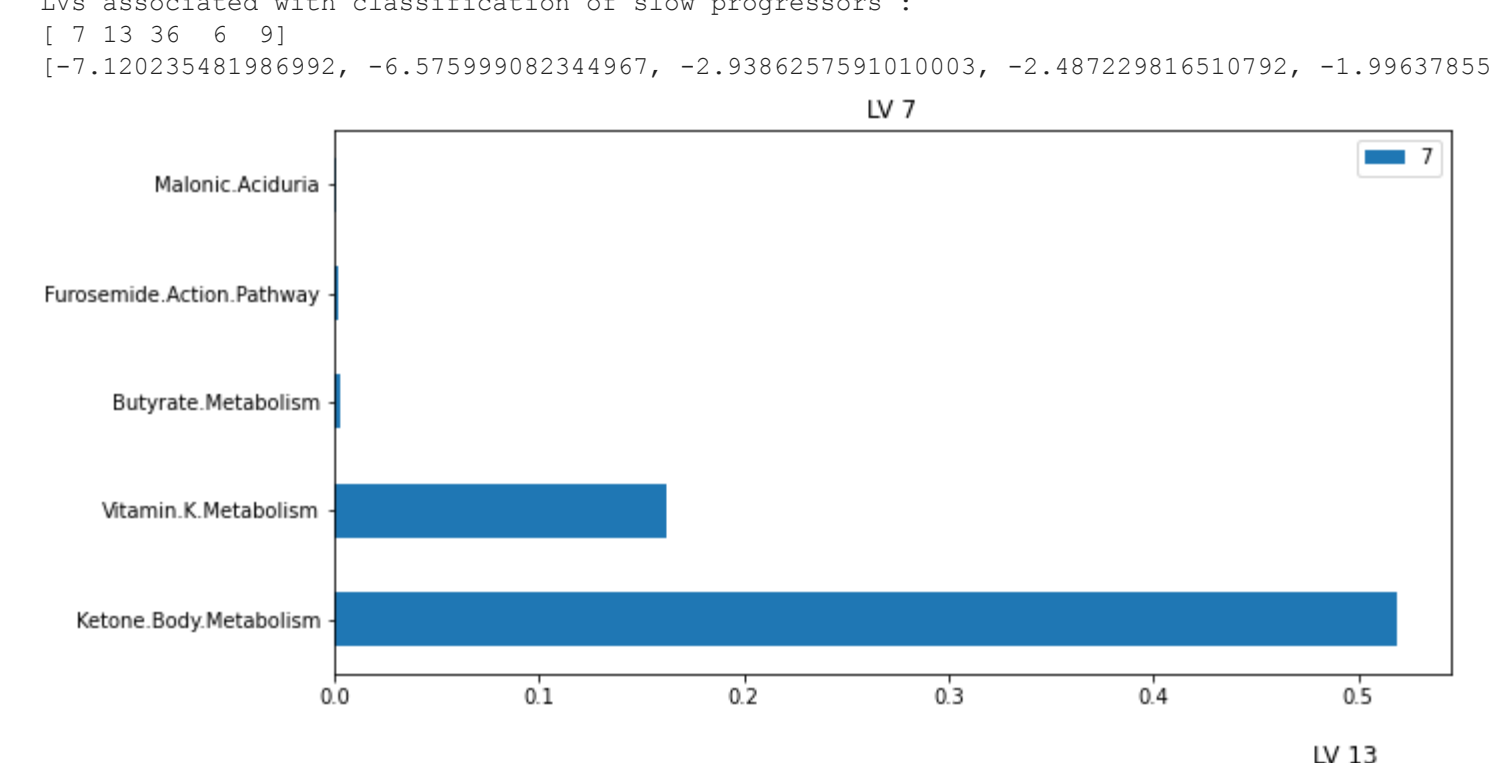


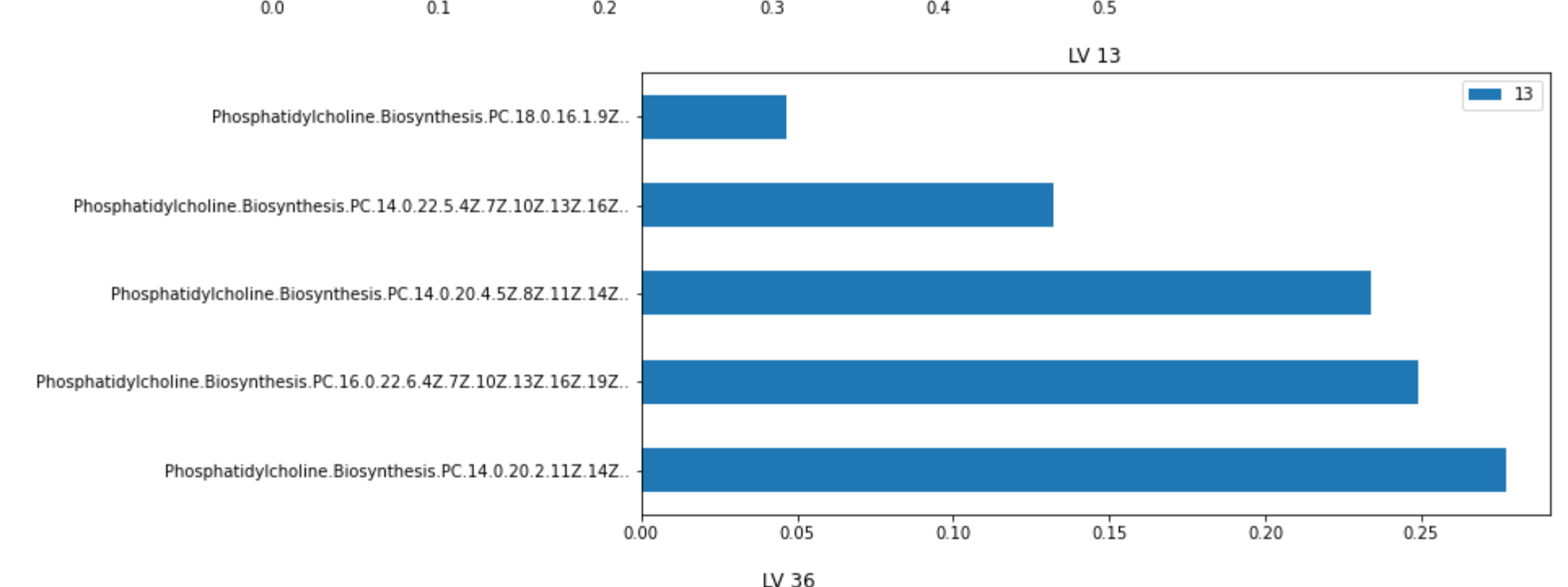


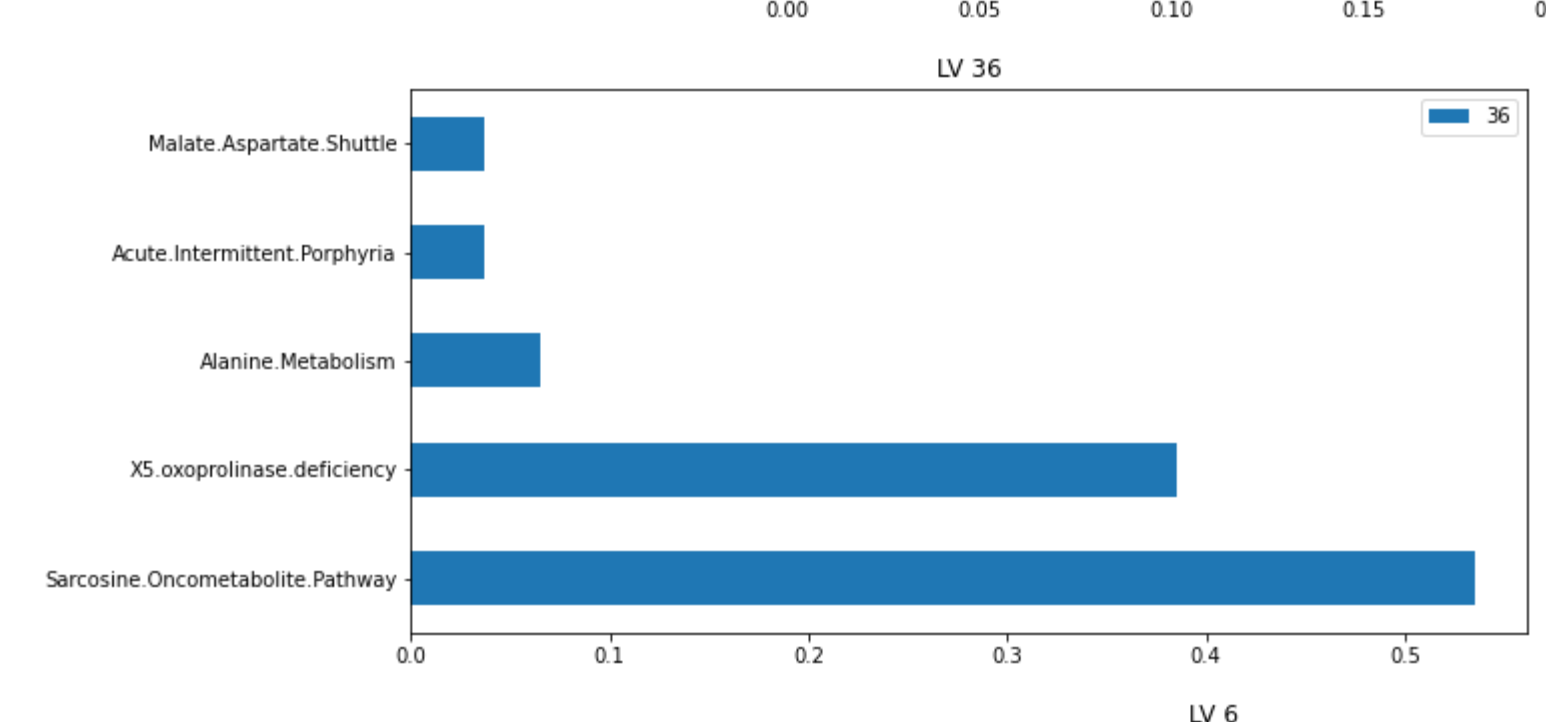


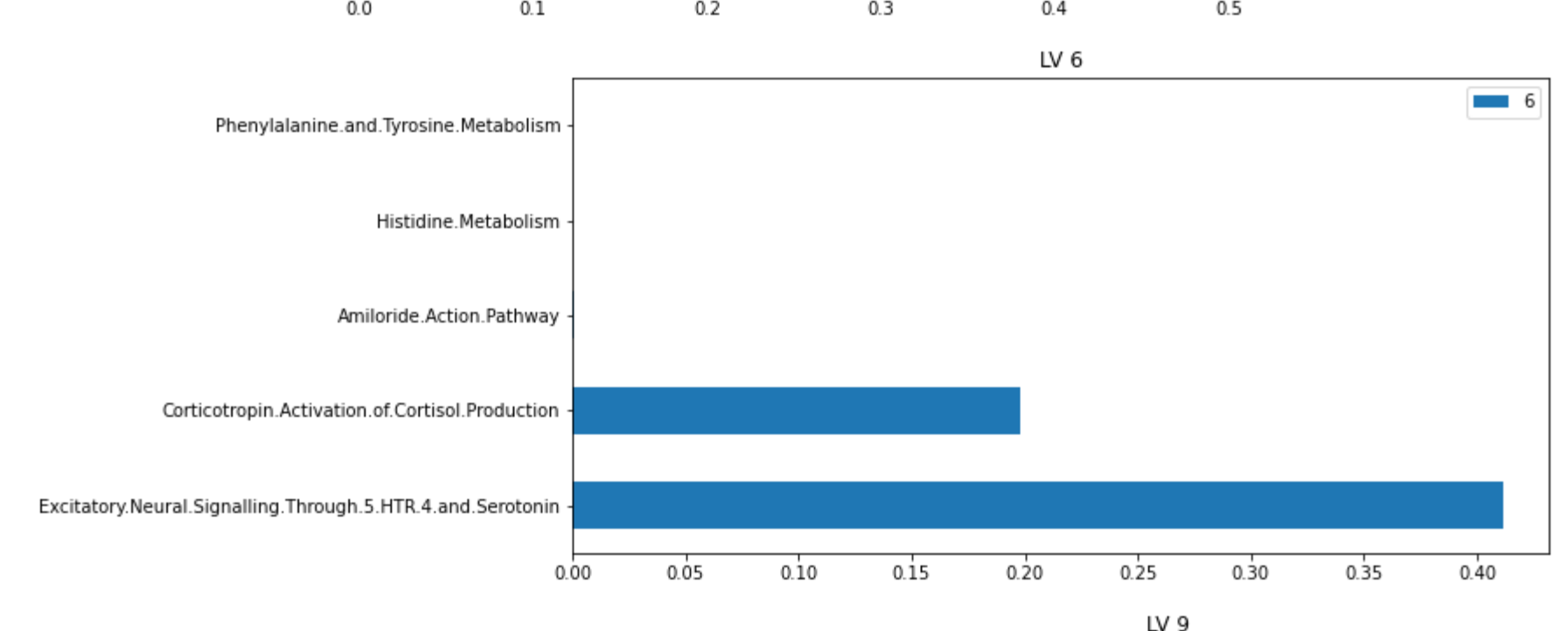


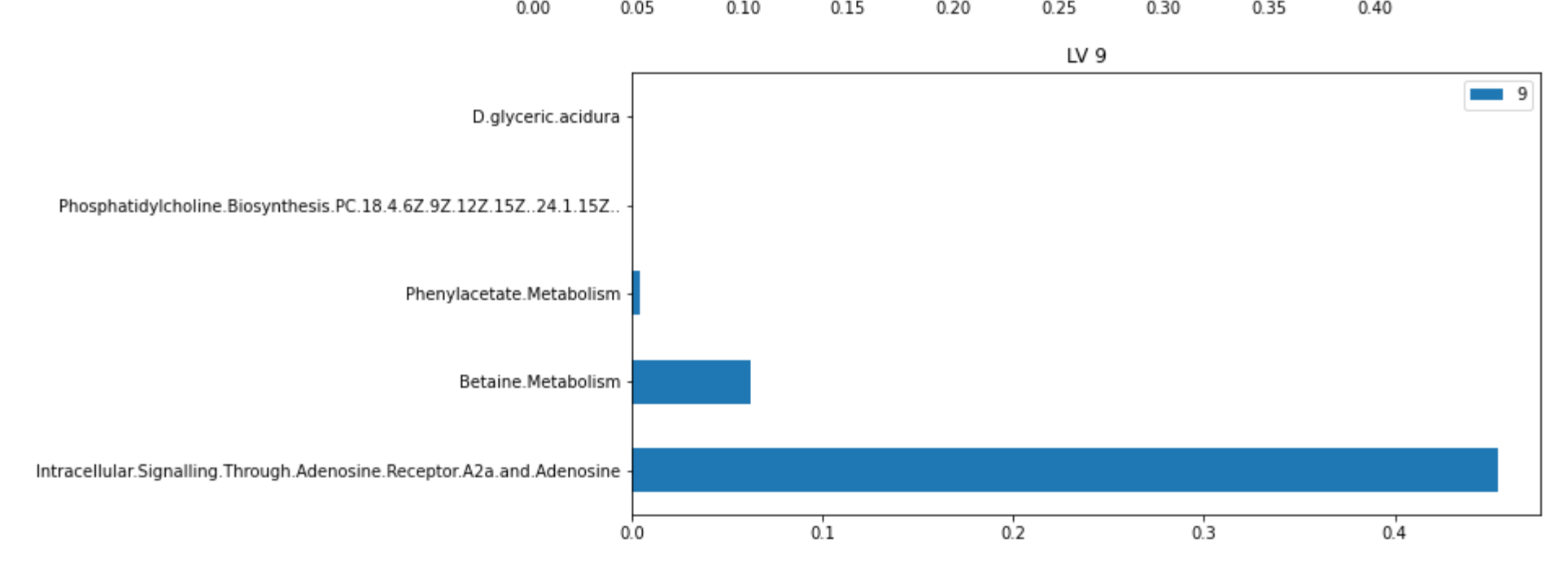

```
LVs associated with classification of slow progressors :
[ 7 13 36  6  9]
[-7.120235481986992, -6.575999082344967, -2.9386257591010003, -2.487229816510792, -1.9963785549379622]
```



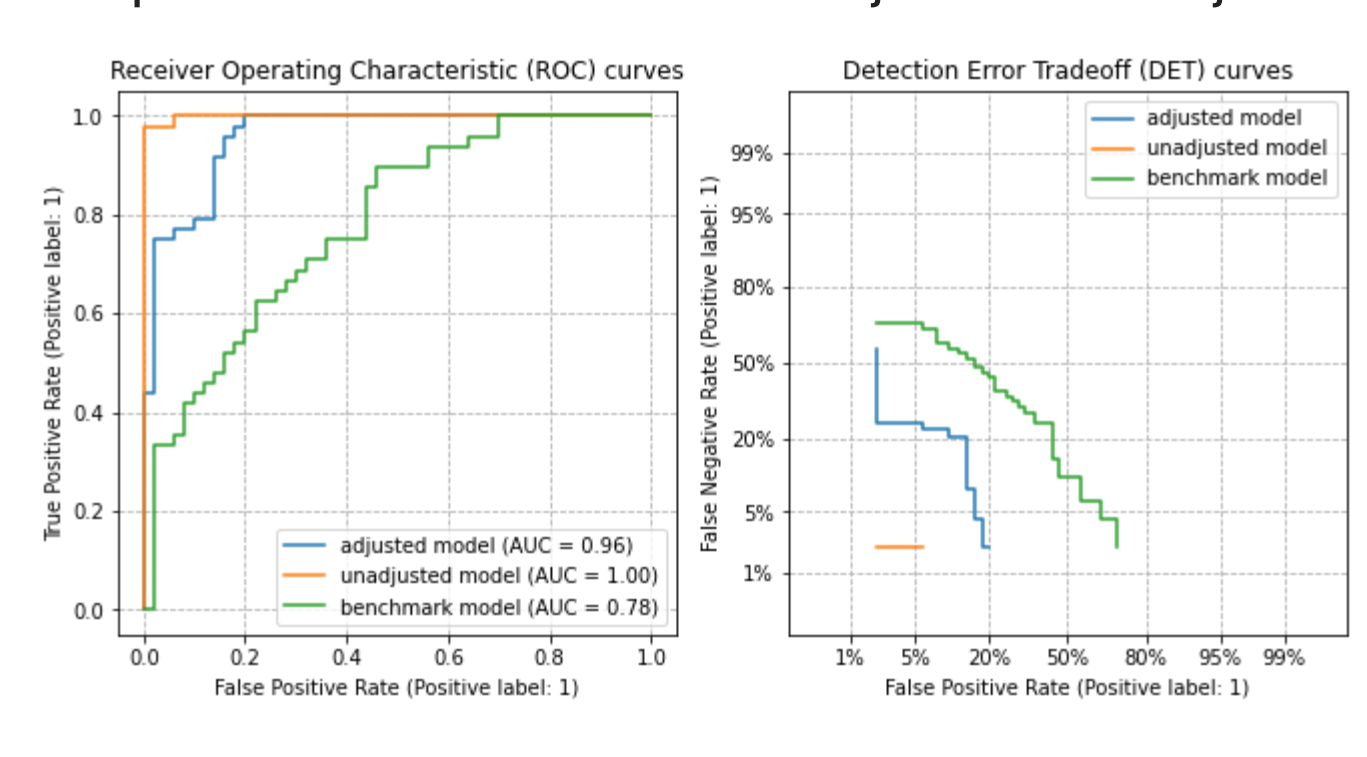








Comparison of ROC curves between adjusted and unadjusted models



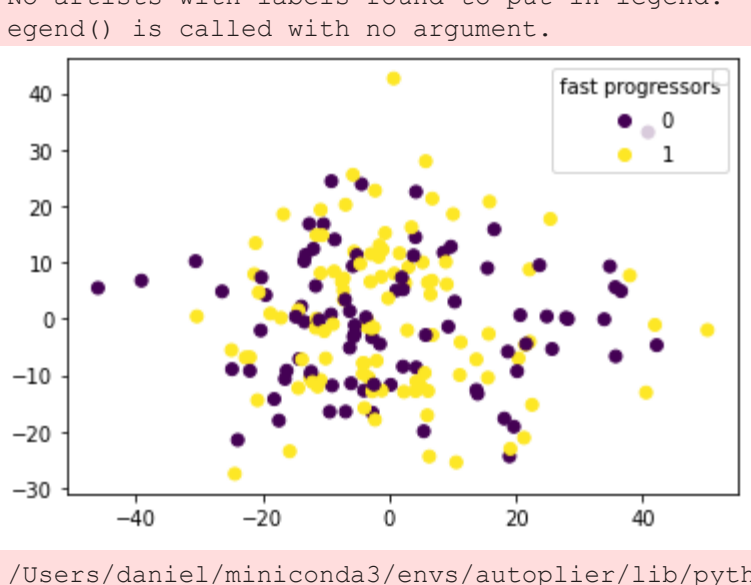
Normo and Micro albuminuria group analysis

```
/var/folders/v3/d9c1pzwx01d7sf9x7_4wzy4c0000gq/T/ipykernel_46964/3313650863.py:12: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-v
iew-versus-a-copy
  clinred['log2acr'] = np.log2(clinred['urine_albumin_v3y0']+1) - np.log2(clinred['urine_creatinine_v3y0']+1)
```

benchmark model

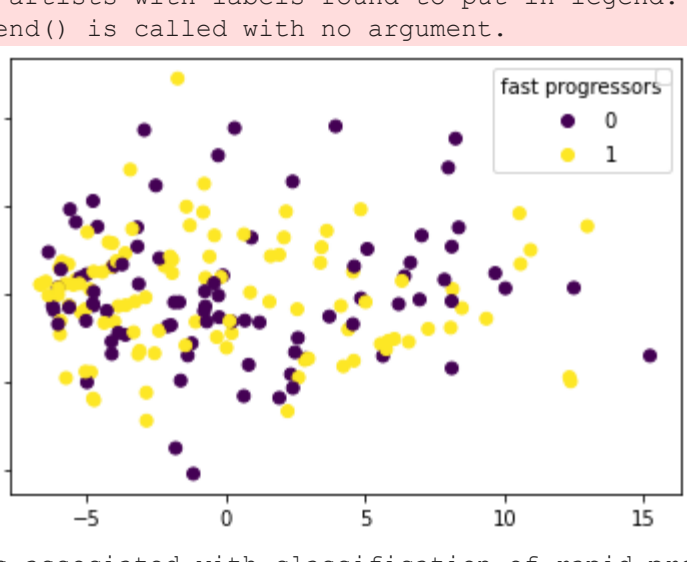
```
No artists with labels found to put in legend.  Note that artists whose label start with an underscore are ignored when l
egend() is called with no argument.
```



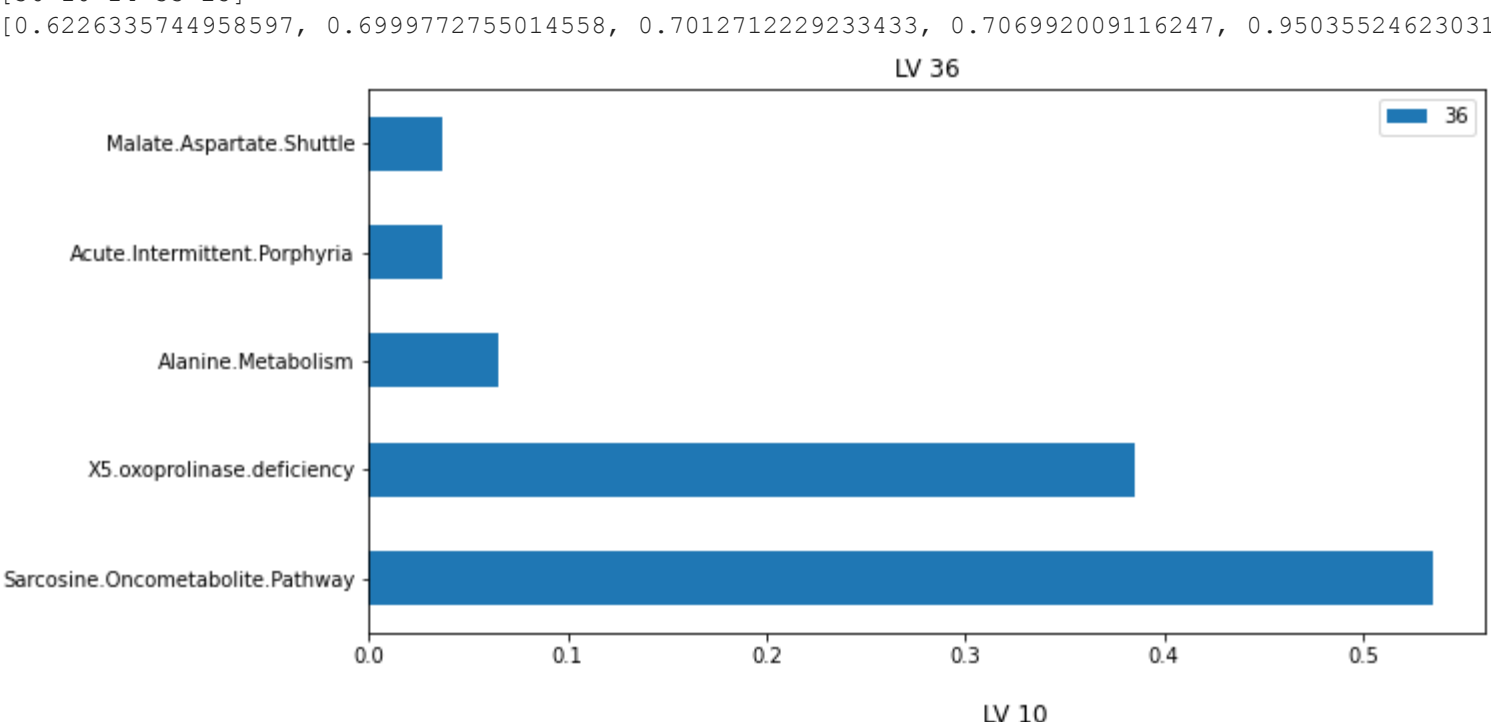

```
/Users/daniel/miniconda3/envs/autoplier/lib/python3.9/site-packages/sklearn/svm/_base.py:1206: ConvergenceWarning: Liblin
ear failed to converge, increase the number of iterations.
  warnings.warn(
/Users/daniel/miniconda3/envs/autoplier/lib/python3.9/site-packages/sklearn/svm/_base.py:1206: ConvergenceWarning: Liblin
ear failed to converge, increase the number of iterations.
  warnings.warn(
/Users/daniel/miniconda3/envs/autoplier/lib/python3.9/site-packages/sklearn/svm/_base.py:1206: ConvergenceWarning: Liblin
ear failed to converge, increase the number of iterations.
  warnings.warn(
```

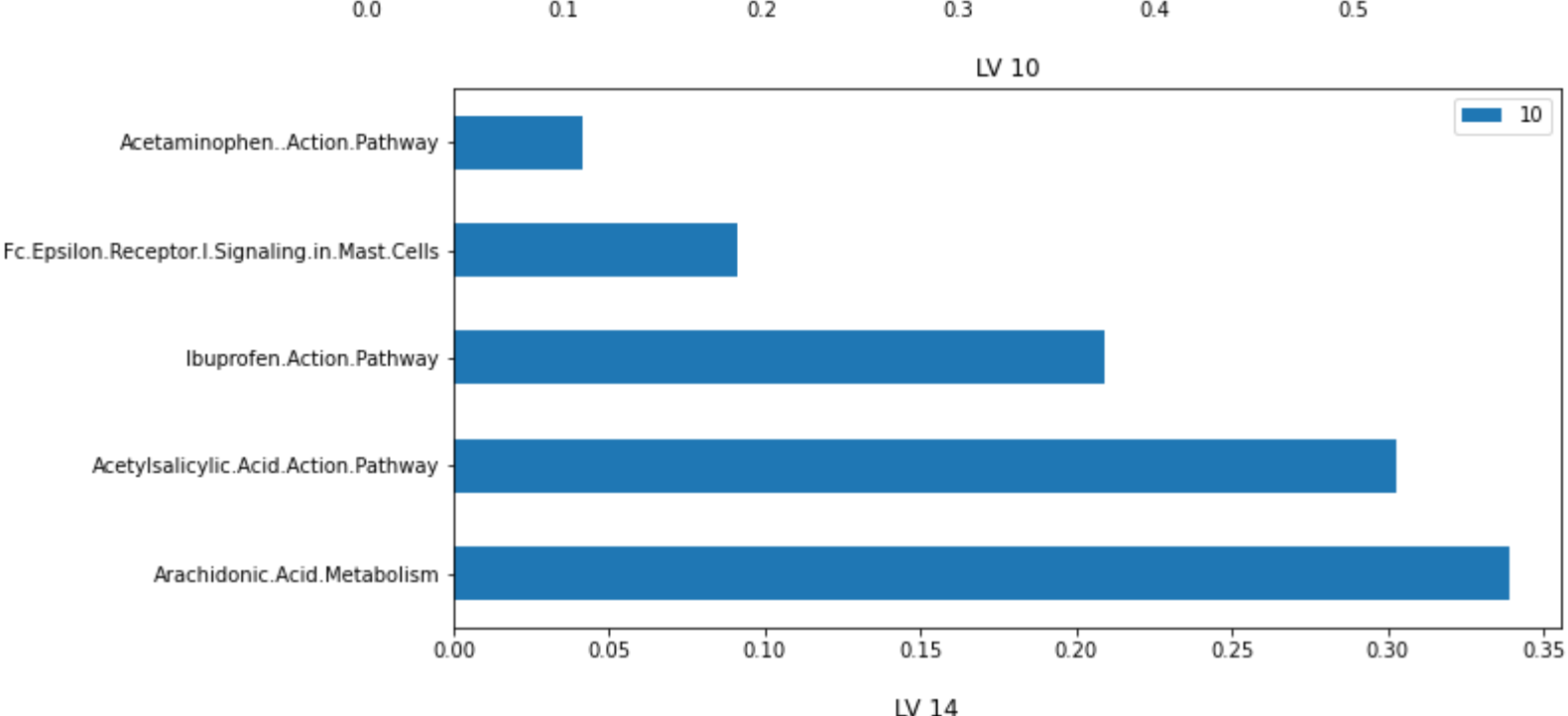
Top latent variables for predicting rapid eGFR decline unadjusted

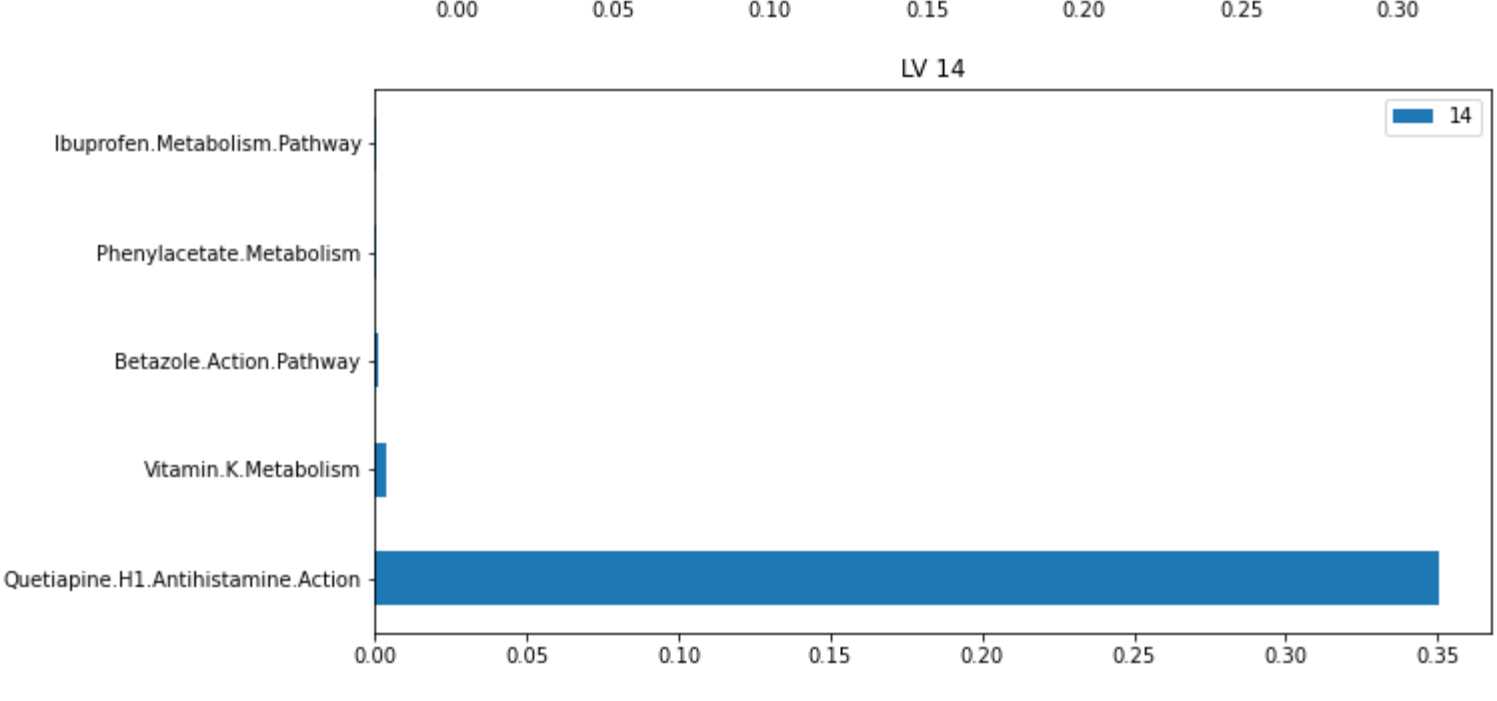
```
No artists with labels found to put in legend.  Note that artists whose label start with an underscore are ignored when l
egend() is called with no argument.
```

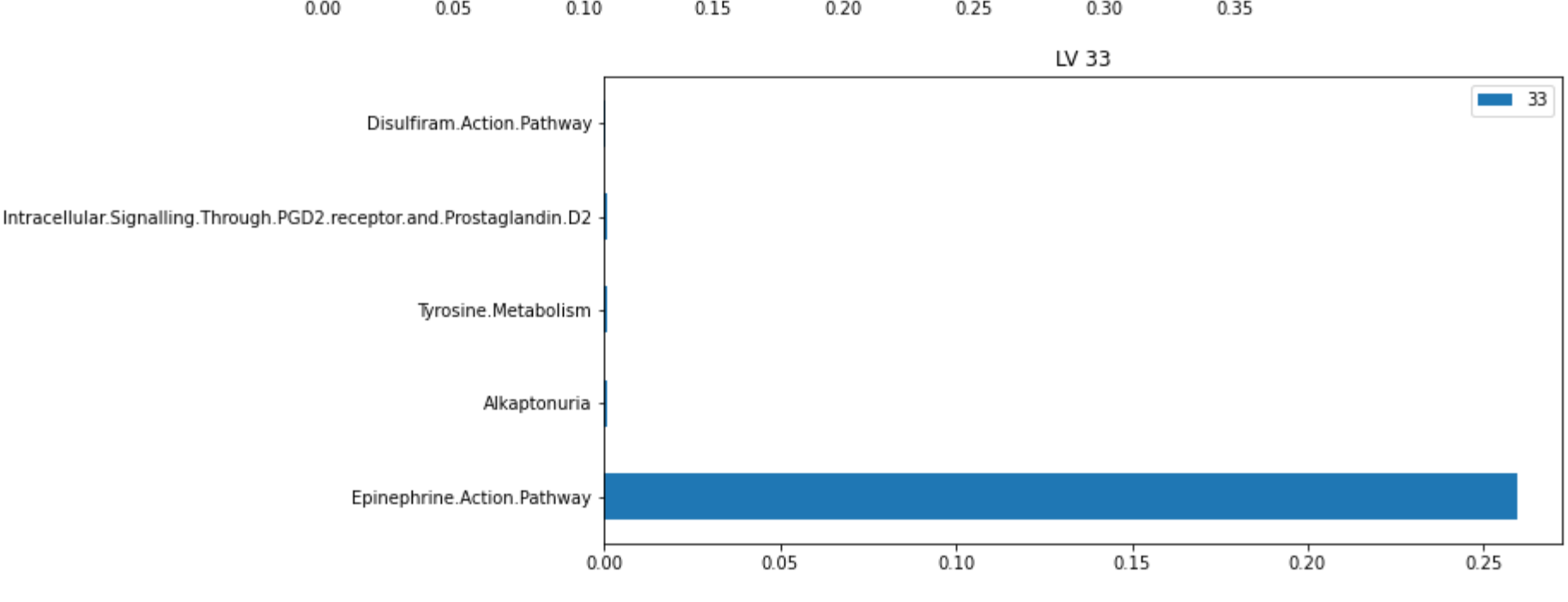


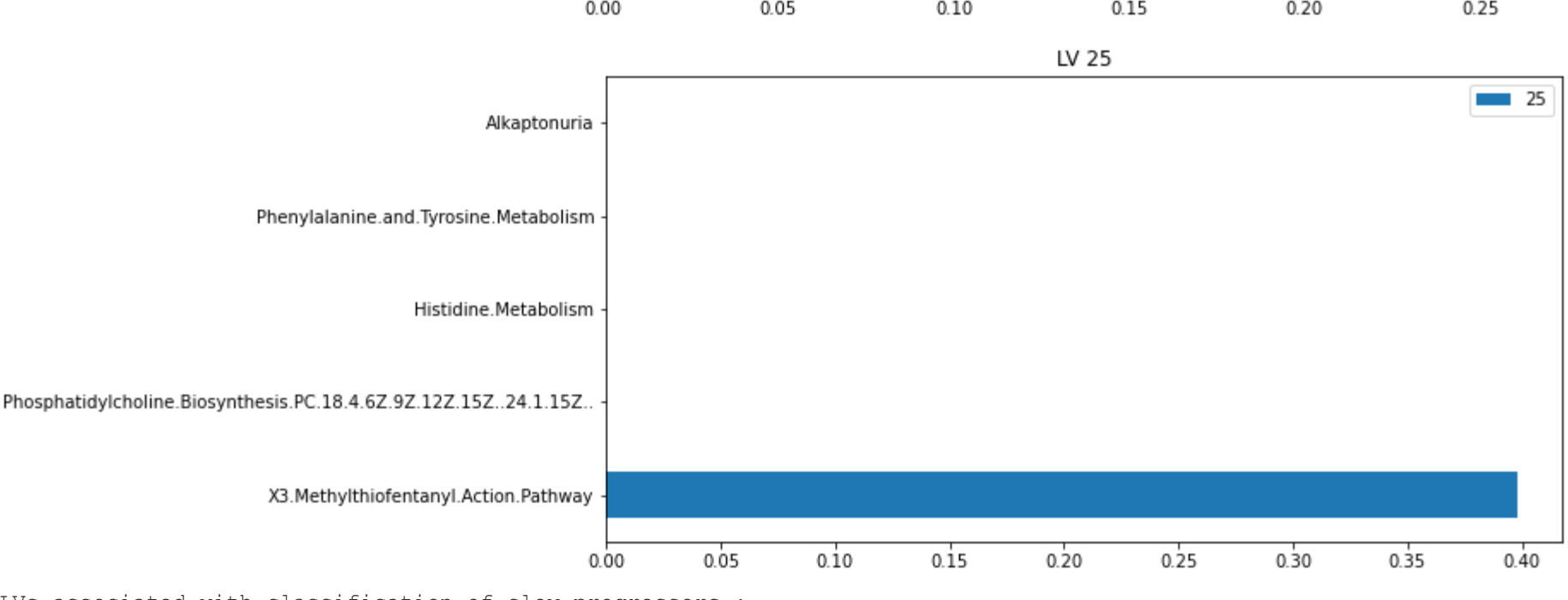

```
LVs associated with classification of rapid progressors :
[36 10 14 33 25]
[0.6226335744958597, 0.6999772755014558, 0.7012712229233433, 0.706992009116247, 0.9503552462303162]
```



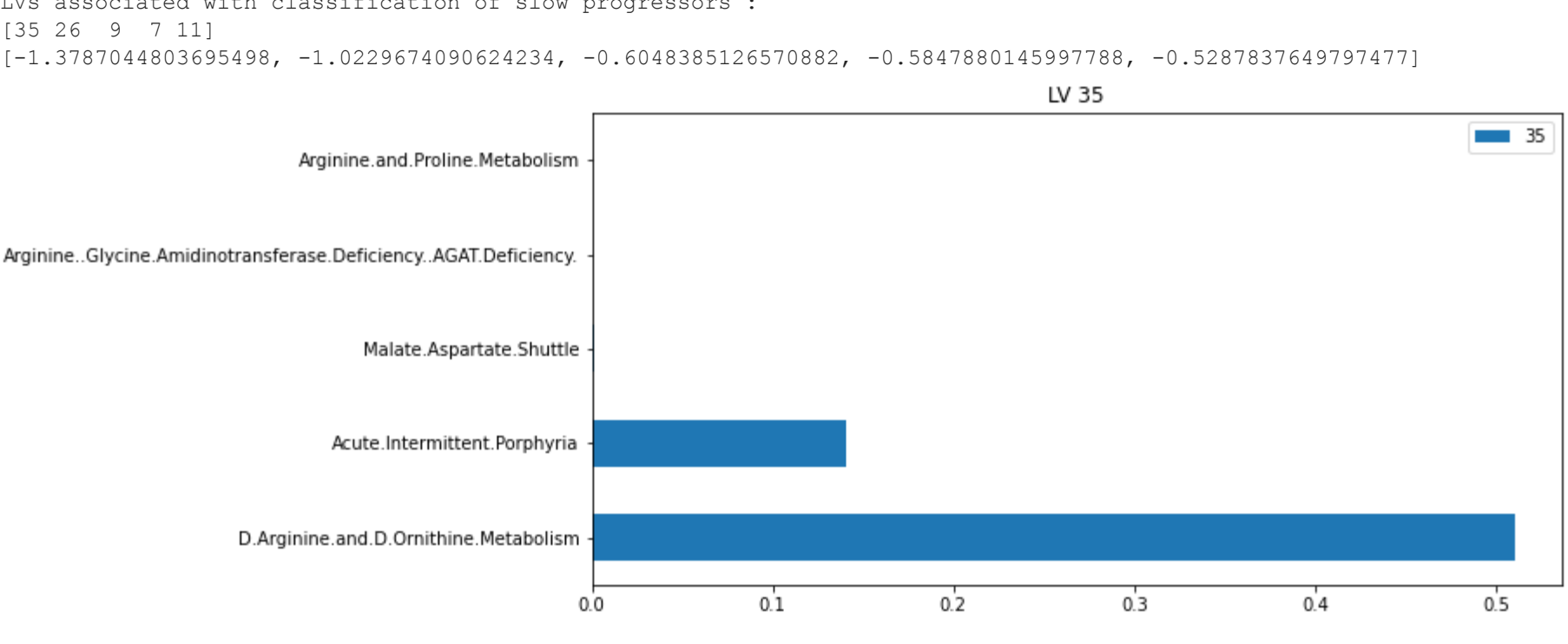


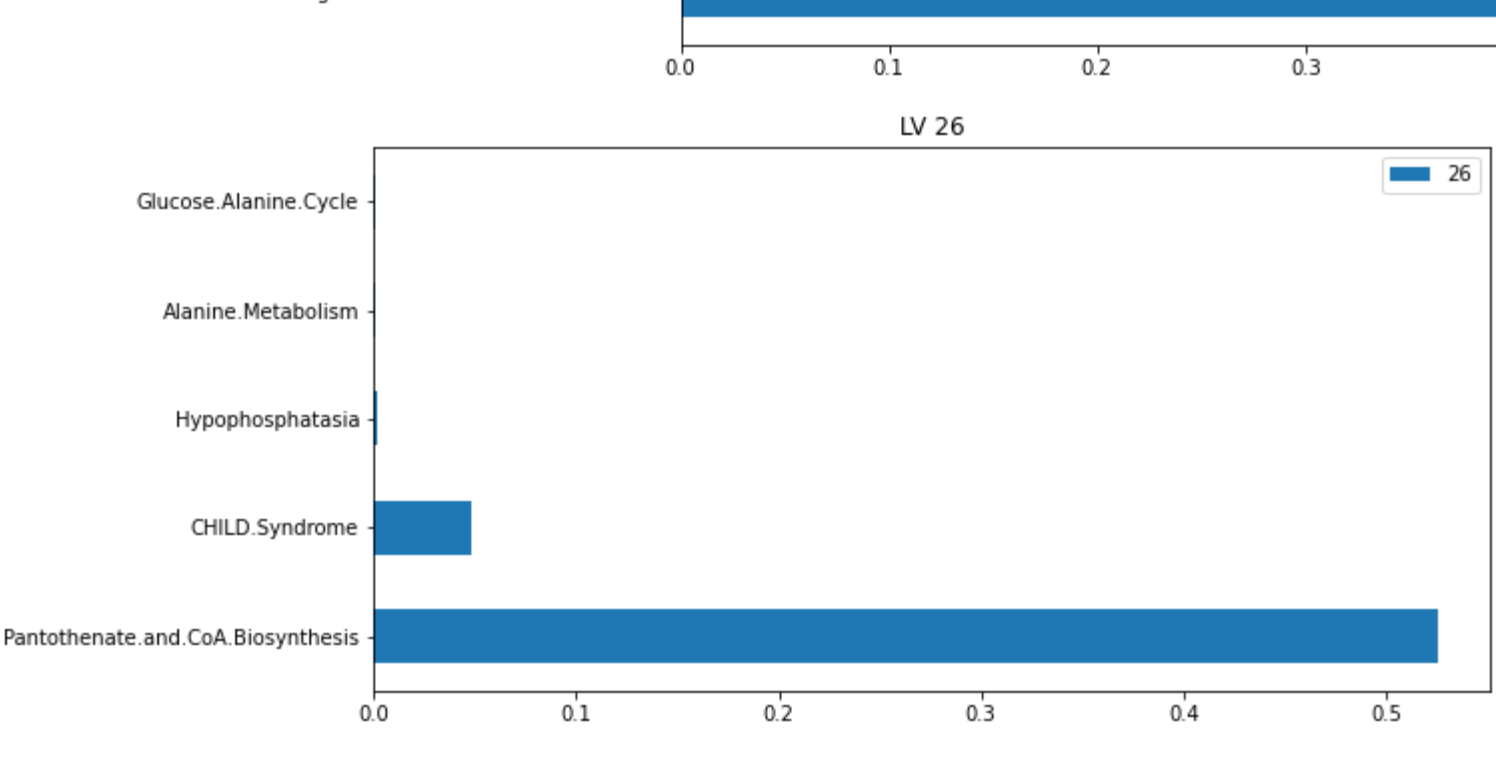


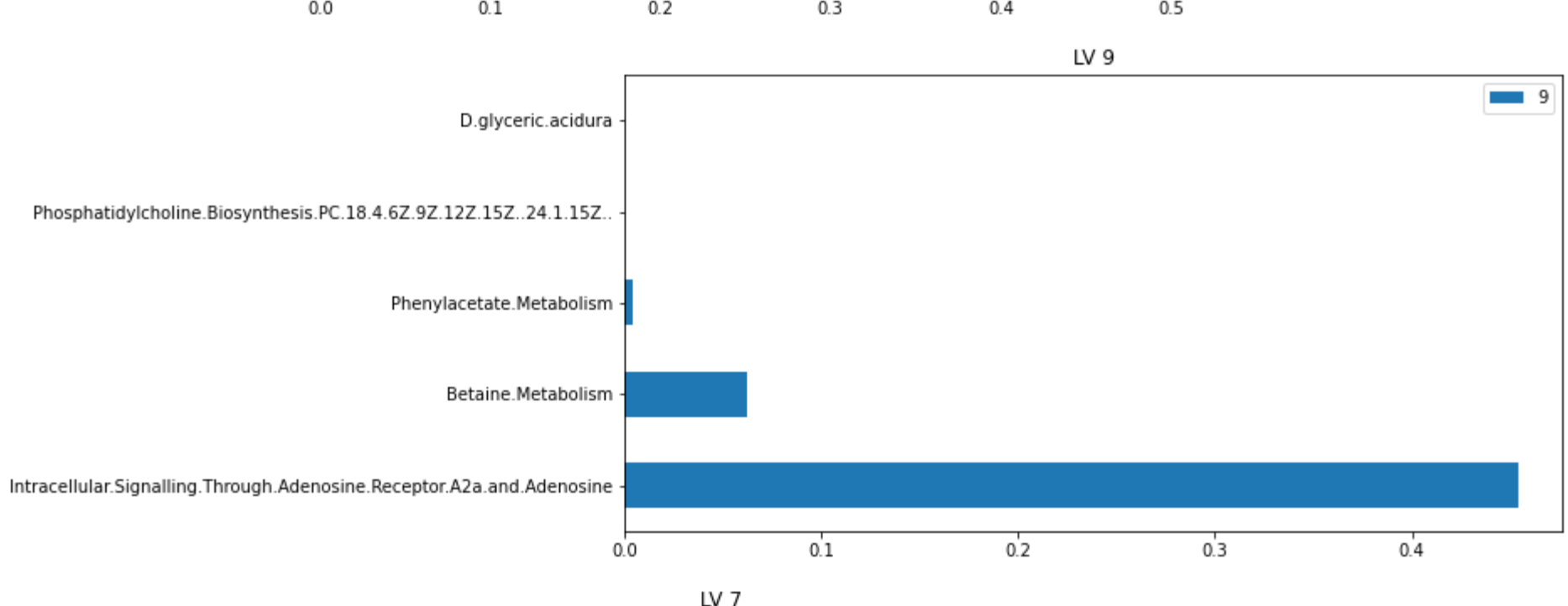


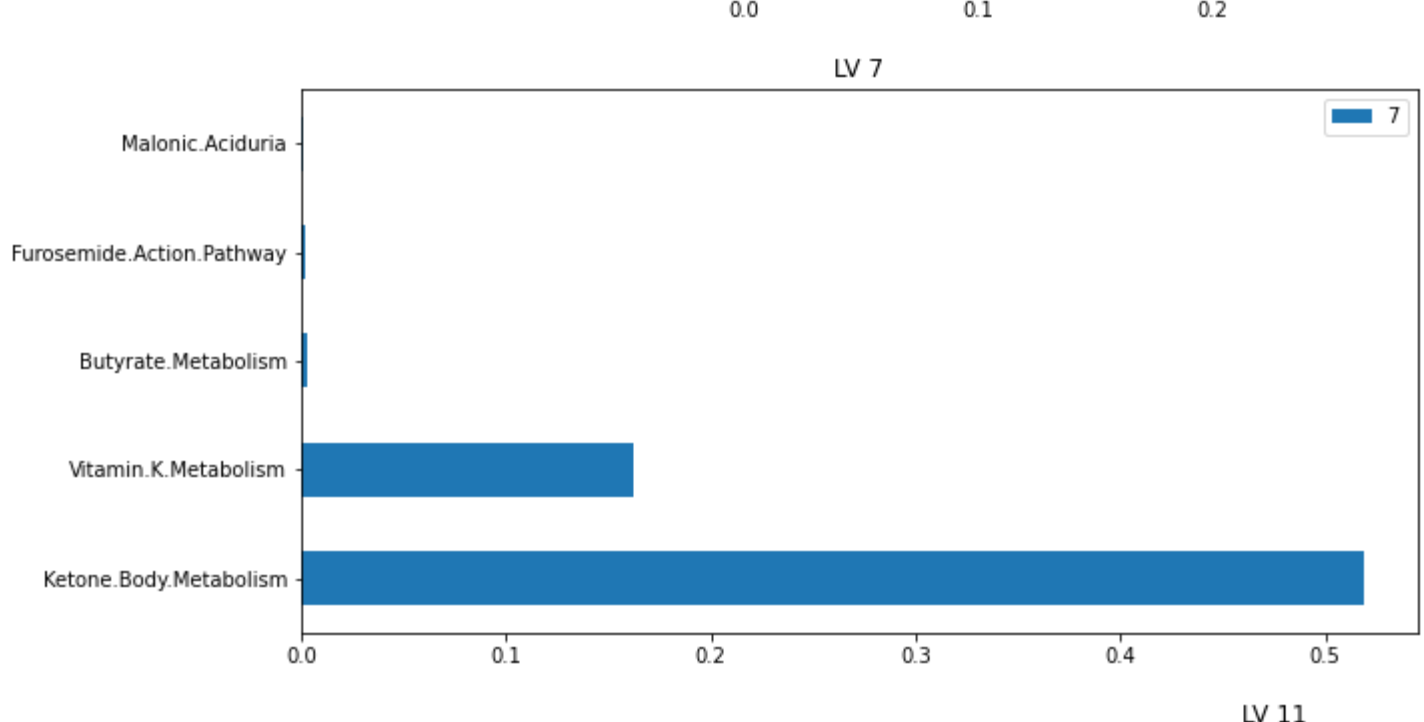


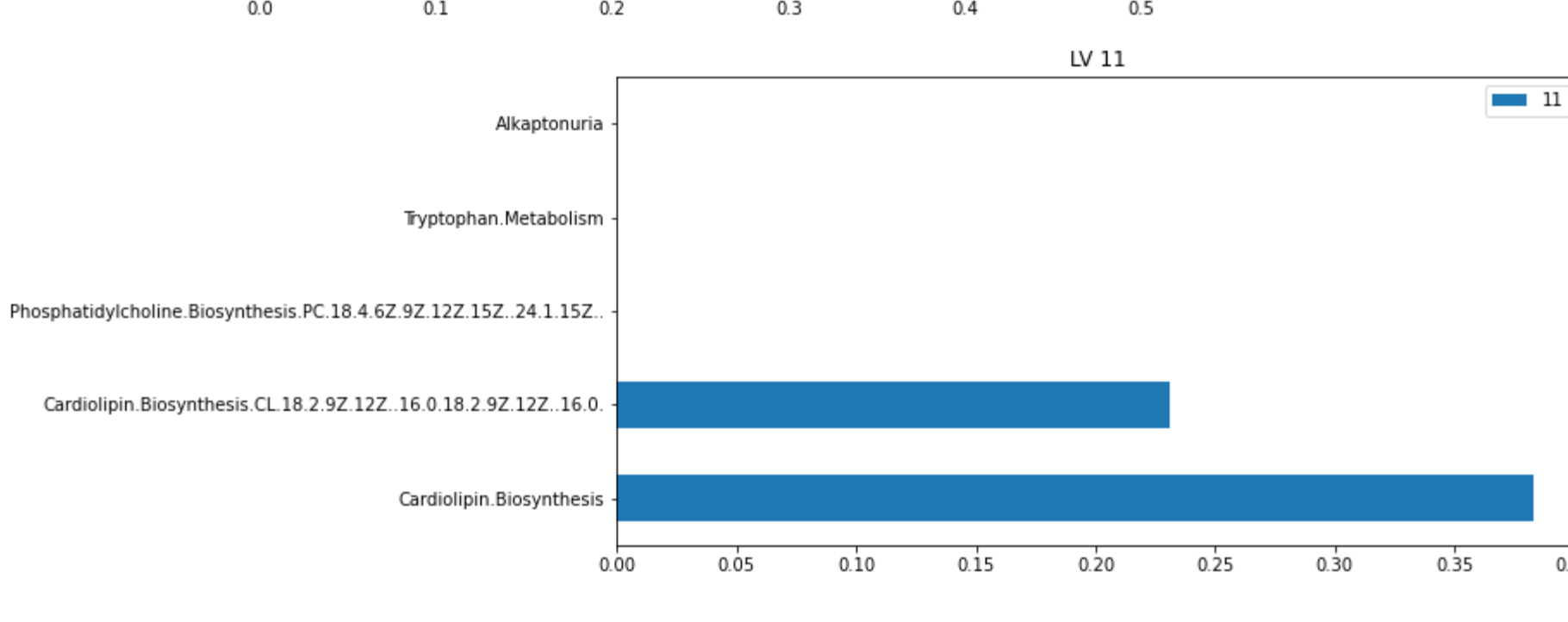

```
LVs associated with classification of slow progressors :
[35 26  9  7 11]
[-1.3787044803695498, -1.0229674090624234, -0.6048385126570882, -0.5847880145997788, -0.5287837649797477]
```





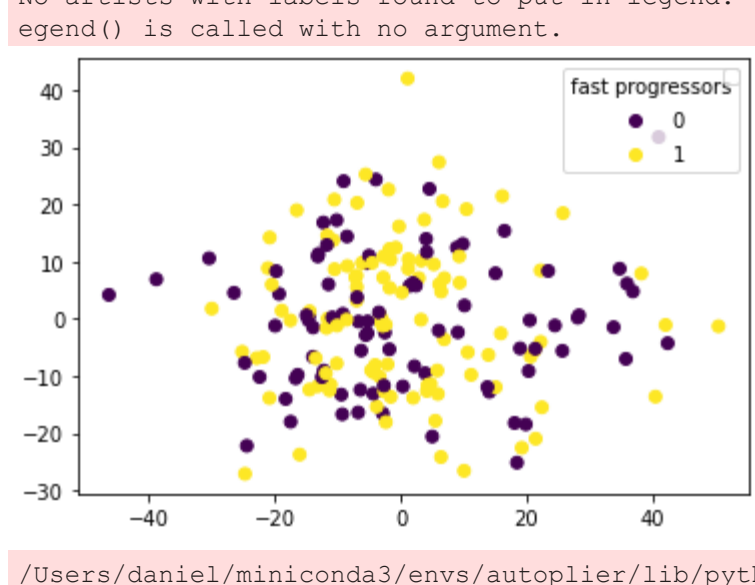






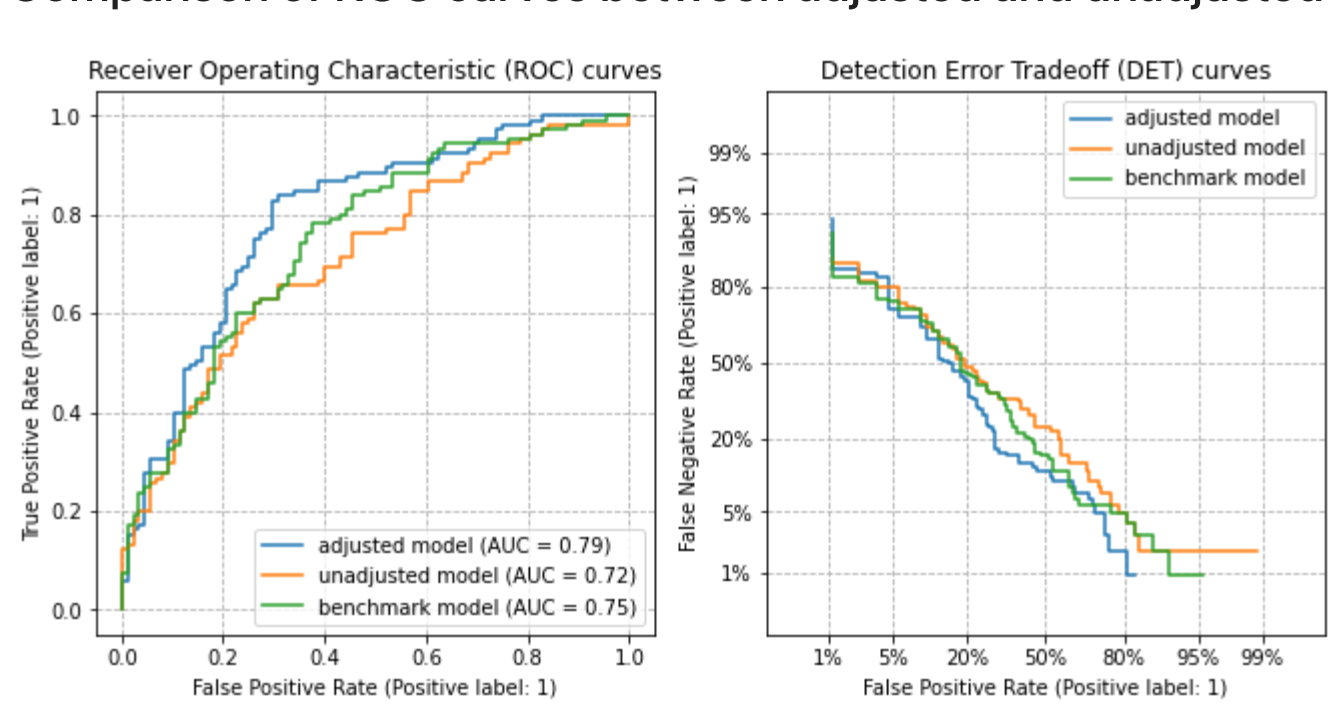
Top latent variables for predicting rapid eGFR decline with adjustment for clinical variables

```
No artists with labels found to put in legend.  Note that artists whose label start with an underscore are ignored when l
egend() is called with no argument.
```




```
/Users/daniel/miniconda3/envs/autoplier/lib/python3.9/site-packages/sklearn/svm/_base.py:1206: ConvergenceWarning: Liblin
ear failed to converge, increase the number of iterations.
  warnings.warn(
/Users/daniel/miniconda3/envs/autoplier/lib/python3.9/site-packages/sklearn/svm/_base.py:1206: ConvergenceWarning: Liblin
ear failed to converge, increase the number of iterations.
  warnings.warn(
/Users/daniel/miniconda3/envs/autoplier/lib/python3.9/site-packages/sklearn/svm/_base.py:1206: ConvergenceWarning: Liblin
ear failed to converge, increase the number of iterations.
  warnings.warn(
No LVs are associated with classification of rapid progressors!
No LVs are associated with classification of slow progressors!
```

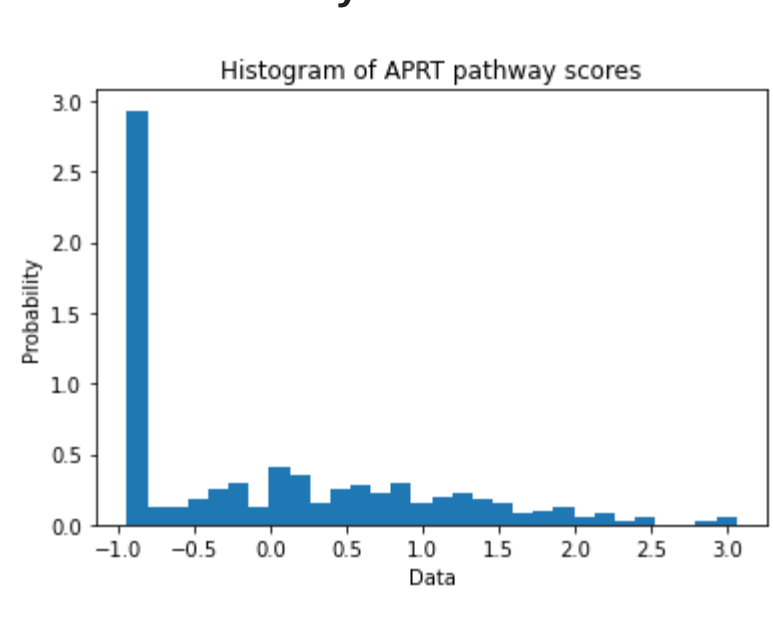
Comparison of ROC curves between adjusted and unadjusted models

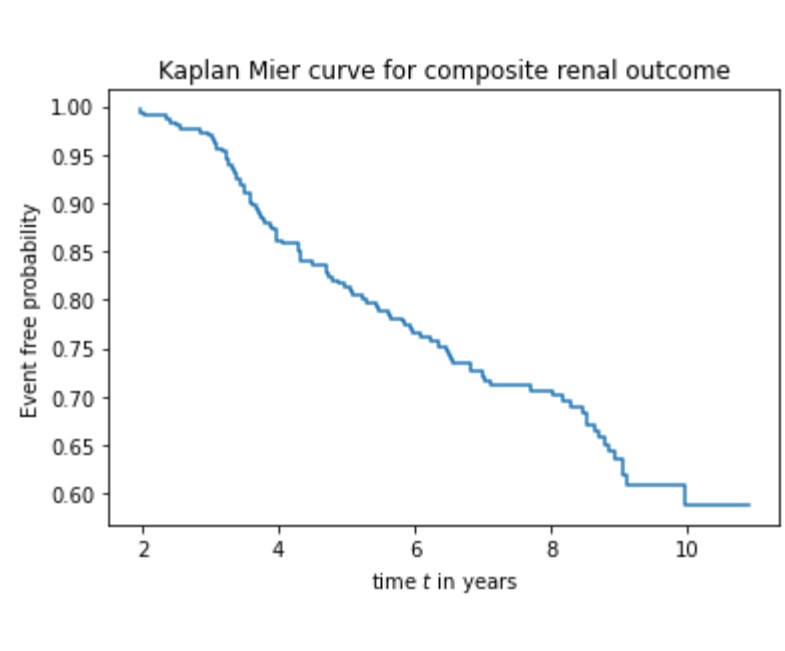


Association of APRT pathway scores with composite renal outcome

We consider the composite event: Dialysis, Kidney Transplant, 50% eGFR decline, or eGFR<15 ml/min/1.73m^2. Two patient groups are split along the median APRT pathway score and the significance between the two survival curves is calculated. Finally, the prognostic strength of the binned High/Low category is quantified by C-statistic in a Cox proportional hazard model with adjustment for clinical variables.

Full cohort analysis

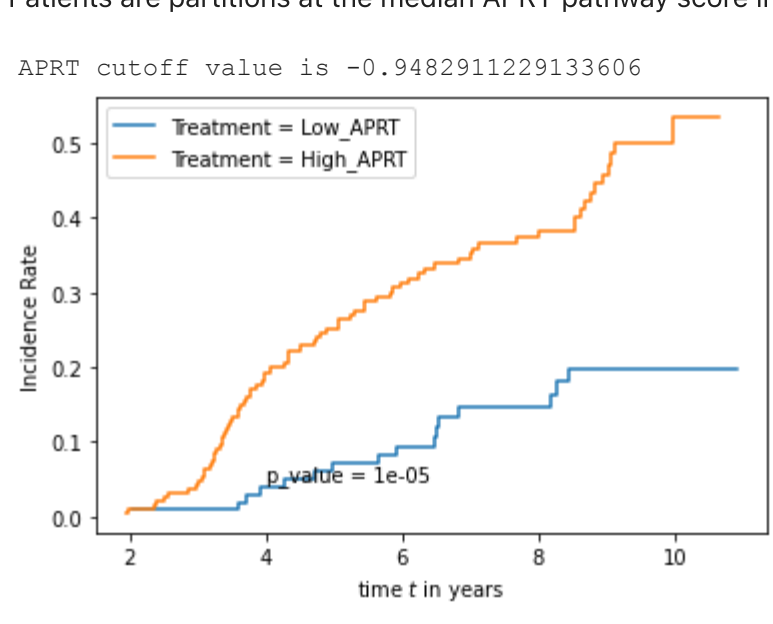




Is there any difference in event probability between groups with high or low APRT pathway?

Patients are partitions at the median APRT pathway score into high_APRT and low_APRT groups and come survival curves

```
APRT cutoff value is -0.9482911229133606
```



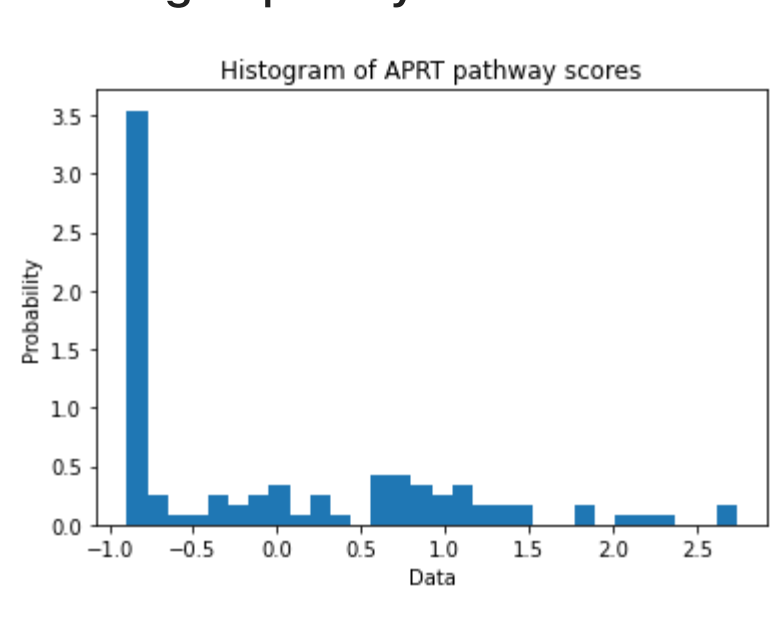
group	counts	observed	expected	statistic
High_APRT	191	77	55.900141	21.099859
Low_APRT	107	16	37.099859	-21.099859

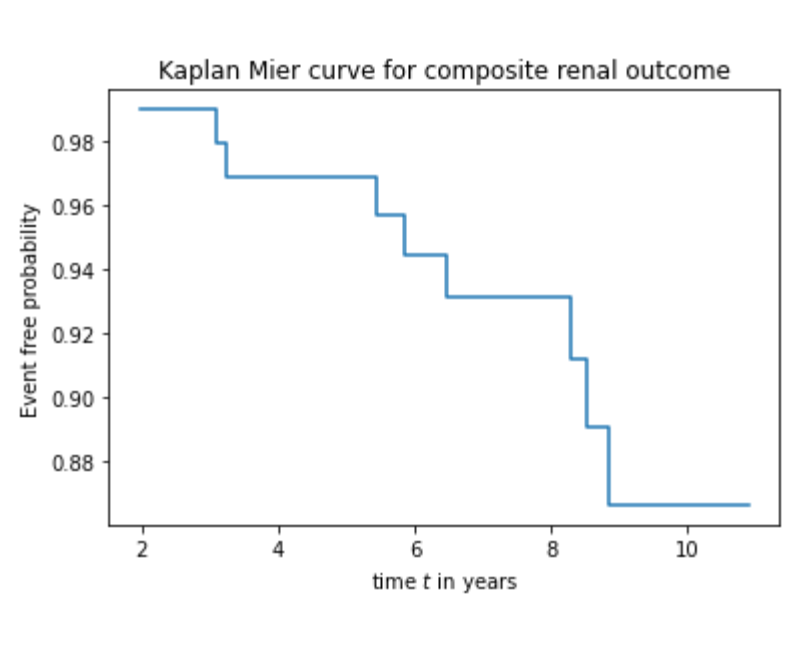
Cox Proportional Hazard

```
Concordance statistic: 0.8613754989253914
```

	log hazard ratio	hazard ratio
race_ethnicity_cat2_v3y0=2	0.570778	1.769644e+00
log2acr	0.298844	1.348300e+00
hemoglobin_a1c_v3y0	0.253634	1.288700e+00
map_v3y0	0.026513	1.026868e+00
waist_v3y0	0.001877	1.001879e+00
age_integer_v3y0	-0.022382	9.778665e-01
egfr_cric_v3y0	-0.074537	9.281731e-01
sex_v3y0=2	-0.385734	6.799511e-01
group=Low_APRT	-0.781791	4.575858e-01
race_ethnicity_cat2_v3y0=4	-14.492598	5.080947e-07
race_ethnicity_cat2_v3y0=3	-15.515763	1.826375e-07

Normo group analysis

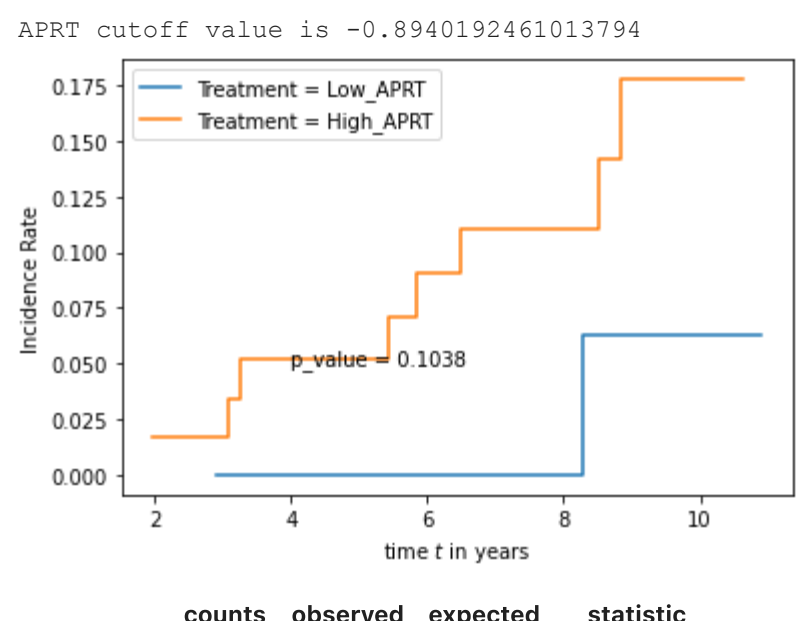




Is there any difference in event probability between groups with high or low APRT pathway?

Patients are partitions at the median APRT pathway score into high_APRT and low_APRT groups and come survival curves

```
APRT cutoff value is -0.8940192461013794
```



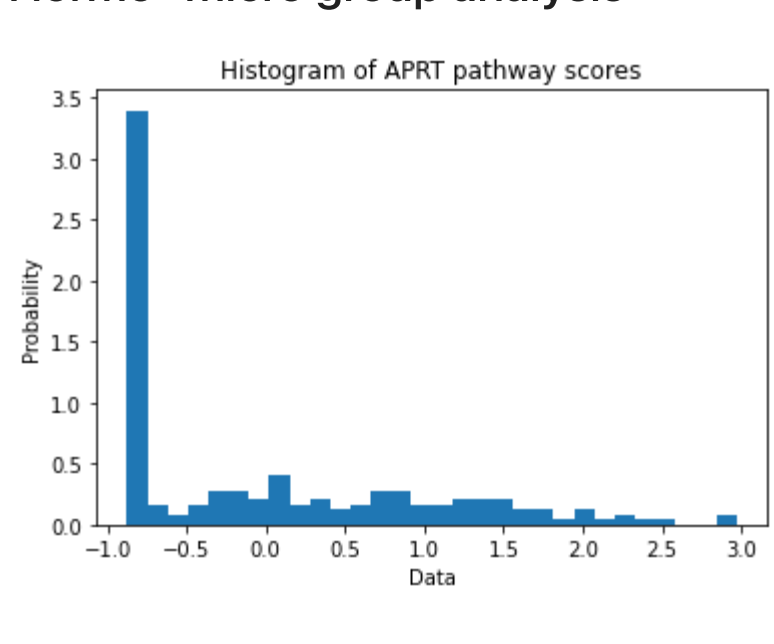
group	counts	observed	expected	statistic
High_APRT	58	8	5.643266	2.356734
Low_APRT	40	1	3.356734	-2.356734

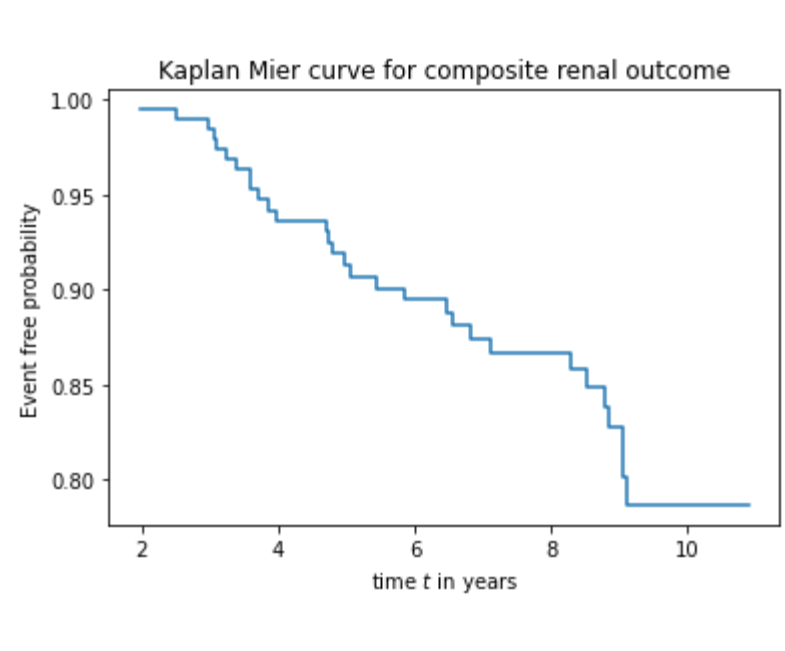
Cox Proportional Hazard

```
Concordance statistic: 0.9178515007898894
```

	log hazard ratio	hazard ratio
hemoglobin_a1c_v3y0	1.399658	4.053815e+00
map_v3y0	0.073505	1.076273e+00
age_integer_v3y0	0.063644	1.065712e+00
waist_v3y0	0.047713	1.048869e+00
egfr_cric_v3y0	-0.232334	7.926814e-01
log2acr	-0.834036	4.342928e-01
race_ethnicity_cat2_v3y0=2	-1.089853	3.362660e-01
sex_v3y0=2	-1.688430	1.848095e-01
group=Low_APRT	-2.680214	6.854847e-02
race_ethnicity_cat2_v3y0=4	-15.682497	1.545889e-07

Normo+micro group analysis

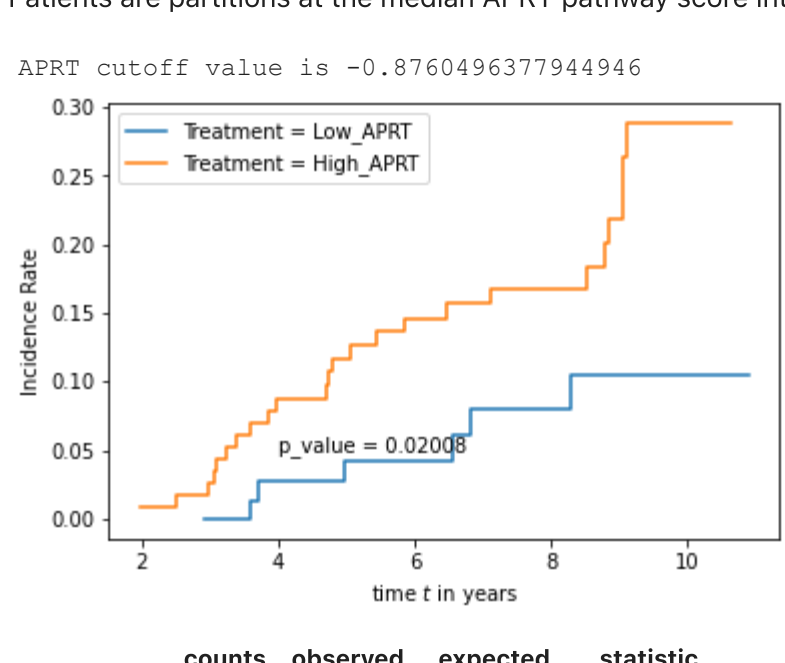




Is there any difference in event probability between groups with high or low APRT pathway?

Patients are partitions at the median APRT pathway score into high_APRT and low_APRT groups and come survival curves

```
APRT cutoff value is -0.8760496377944946
```



group	counts	observed	expected	statistic
High_APRT	115	24	17.745071	6.254929
Low_APRT	78	6	12.254929	-6.254929

Cox Proportional Hazard

```
Concordance statistic: 0.870426474015381
```

	log hazard ratio	hazard ratio
hemoglobin_a1c_v3y0	0.727909	2.070746e+00
sex_v3y0=2	0.645626	1.907181e+00
log2acr	0.117556	1.124744e+00
age_integer_v3y0	0.079374	1.082609e+00
map_v3y0	0.030089	1.030546e+00
waist_v3y0	0.007595	1.007624e+00
egfr_cric_v3y0	-0.121025	8.860122e-01
race_ethnicity_cat2_v3y0=2	-0.382402	6.822206e-01
group=Low_APRT	-0.746379	4.740799e-01
race_ethnicity_cat2_v3y0=3	-12.580324	3.439021e-06
race_ethnicity_cat2_v3y0=4	-14.611482	4.511426e-07