MLCB

Friday, May 31, 2024 4:56 PM

Lucid
https://lucid.app/lucidchart/4de75c75-94b5-4029-92ac-32daf85c10a7/edit?viewport_loc=-1855%2C-2042%2C5625%2C3378%2C0_0&invitationId=inv_a05796dc-bfac-4f99-b4de-8a2c9b06b2fe

M3GPRfigure

Mention Figure 1,
Relate MTMO1 to actual drug formulation situation
shortening

Diagram: Where is the task defined

VAR1, VAR2, Realization1

??-TBD

Union of all the variables (missing variable scenario)

Make Data1, Data2, …, Data n more general,

M3GPR-O can work for MTMO5
M3GPR-O (output as tasks)

MTMO4 will be more relevant for drug formulation

Measured by single instrument, each instrument can generate multiple outputs (often use 4-5 instruments)

Task represents modality or could be flexible

Molecule--> task, assays-->output

MTMO4: new excipient

MTMO2 would be more common than MTMO1 (T2O3 may have less data, usually the formulation is already locked)

Reduce the number of categorical variable will significantly improve the performance of simple GP on Tm2(generated correlated output variable), but seem to make the performance of original output variables worse

Corr 0.5, corr between task2-out1 and task2-out2 and corr between task1-out1 and task1-out2 are often different, so M3GPR-O performs worse

Corr 0.9
After optimization

```
Correlation matrix between tasks
[[ 1.     -0.4346  0.9894 -0.4809]
 [-0.4346  1.     -0.542   0.9978]
 [ 0.9894 -0.542   1.     -0.5848]
 [-0.4809  0.9978 -0.5848  1.    ]]

          Tm_1      NP_1      Tm_2      NP_2
Tm_1  1.000000 -0.409118  0.906527 -0.379267
NP_1 -0.409118  1.000000 -0.426765  0.893893
Tm_2  0.906527 -0.426765  1.000000 -0.206267
NP_2 -0.379267  0.893893 -0.206267  1.000000

MAE      Tm_1   NP_1  Tm_2   NP_2
```

rf 0.45 5.44 0.76 5.98

```
SimpleGP      0.26   2.86  0.65   3.78
```

M3GPR-O 0.23 4.21 0.42 4.49
M3GPR-R 0.23 2.16 0.26 3.20

Avg_STD
rf NA NA NA NA
SimpleGP 0.87 11.41 1.23 11.30
M3GPR-O 0.51 8.06 0.96 10.84
M3GPR-R 1.00 10.91 0.94 10.31

4 cate
MAE
RF 0.59 5.78 0.91 6.14
SimpleGP 0.51 5.93 0.94 5.47
M3GPR-O 0.34 4.02 0.47 4.01
M3GPR-R 0.64 3.80 0.61 5.19
M3GPR-RL 0.34 2.22 0.41 4.26
M3GPR-V 0.61 5.55 0.48 6.65

MAPE

RF	0.01	2.26	0.02	2.63
SimpleGP	0.01	2.45	0.02	2.63
M3GPR-O	0.01	1.0	0.01	1.32
M3GPR-R	0.01	0.85	0.01	0.9
M3GPR-V	0.01	1.07	0.01	1.13

Avg_STD
SimpleGP 0.96 13.45 1.37 14.37

SimpleGP 0.96 12.45 1.37 14.37
M3GPR-O 0.55 7.92 1.03 13.64
M3GPR-R 1.26 11.99 1.18 11.30
M3GPR-RL 0.64 6.15 0.69 6.67
M3GPR-V 0.50 4.66 0.55 5.18

Corr 0.5

```
          Tm_1      NP_1      Tm_2      NP_2
Tm_1  1.000000 -0.409118  0.527813 -0.196244
NP_1 -0.409118  1.000000 -0.328386  0.412322
Tm_2  0.527813 -0.328386  1.000000  0.616769
NP_2 -0.196244  0.412322  0.616769  1.000000

Correlation matrix between tasks
[[ 1.     -0.3229  0.8133 -0.4226]
 [-0.3229  1.     -0.8074  0.9835]
 [ 0.8133 -0.8074  1.     -0.8627]
 [-0.4226  0.9835 -0.8627  1.    ]]
```

MAE	Task1-Out1	Task1-Out2	Task2-Out1	Task2-Out2
RFR	0.45	5.44	0.51	5.59
Simple GPR	0.26	2.86	0.68	5.57
M3GPR-O	**0.20**	5.51	0.69	6.27
M3GPR-R	0.22	**1.67**	0.61	5.68
M3GPR-RL	0.50	3.42	**0.41**	**4.03**
M3GPR-V	0.49	3.45	0.62	4.90
M3GPR-O (4-cate)	0.36	4.11	0.67	6.08
M3GPR-R	0.56	4.97	0.59	6.83

MAPE (0.5)

RF	0.01	2.26	0.01	1.43
Simple GPR	0.01	2.45	0.01	1.14
M3GPR-O (2-cate)	0.0	2.45	0.01	1.44
M3GPR-O (4-cate)	0.01	1.02	0.01	1.36
M3GPR-RL (2-cate)	0.01	1.42	0.01	0.62
M3GPR-V (4-cate)	0.02	1.73	0.01	1.49
M3GPR-RL (4-cate)	0.01	1.16	0.01	1.0

MAPE(0.9)

RF	0.01	2.26	0.02	2.63
SimpleGP	0.01	2.45	0.02	2.63
M3GPR-O	0.01	1.0	0.01	1.32
M3GPR-R	0.01	0.85	0.01	0.9
M3GPR-V	0.01	1.07	0.01	1.13

0.5

Uncertainty	Task1-Out1	Task1-Out2	Task2-Out1	Task2-Out2
RFR	NA	NA	NA	NA
Simple GPR	0.87	11.41	1.74	15.64
M3GPR-O	**0.44**	7.97	1.70	16.13
M3GPR-R	1.75	15.54	1.35	12.75
M3GPR-RL	1.18	11.02	1.29	11.92
M3GPR-V	0.91	**7.86**	**0.95**	**8.25**

MAPE(Cor0.5)	T1-O1	T1-O2	T2-O1	T2-O2		MAPE(Cor0.9)	T1-O1	T1-O2	T2-O1	T2-O2
RFR	0.01	2.26	0.01	1.43		RFR	0.01	2.26	0.02	2.63
Simple GPR	0.01	2.45	0.01	1.14	-	Simple GPR	0.01	2.45	0.02	2.63
M3GPR-O	0.01	**1.02**	0.01	1.36		M3GPR-O	0.01	1.00	0.01	1.32
M3GPR-R	0.01	1.16	0.01	**1.00**		M3GPR-R	0.01	**0.85**	0.01	**0.90**
M3GPR-V	0.02	1.73	0.01	1.49		M3GPR-V	0.01	1.07	0.01	1.13