Study the effect of pH levels by Buffer type

Wednesday, July 3, 2024 11:23 AM

Buffer Type label correspondence with actual types
0: Ace (Acetate)
1: His (Histidine)
2: Cit (Citrate)
3: Phos (Phosphate)

Additive_label Correspondence
0: Arginine, solid shape
1: Glycine, hollow shape

Acetate

Buffer_Type_label	0.0
Sugar_Salt_label	0.0
Additive_label	1.0
pH	4.5
task_ind	0.0

Histidine

Buffer_Type_label	1.0
Sugar_Salt_label	0.0
Additive_label	1.0
pH	6.5
task_ind	1.0

Cit (Citrate)

Buffer_Type_label	2.0
Sugar_Salt_label	0.0
Additive_label	0.0
pH	5.0
task_ind	1.0

Phos (Phosphate)

Buffer_Type_label	3.0
Sugar_Salt_label	0.0
Additive_label	0.0
pH	7.0
task_ind	1.0

Model accuracy affect pattern

Use original scale for embedding 3

pH level and target: linear? Model accuracy, use another model

Acetate: 3.7-5.7

Histidine: 5-7
Citrate: 3.7-5.7
Phosphate: 6.2-8.2

Citrate, Phosphate: why choose Arginine

Citrate
```
np.mean(new_train_mean[df_X_train_buffer['Additive_label']==1])
```
91.14

```
np.mean(new_train_mean[df_X_train_buffer['Additive_label']==0])
```
91.61

Phosphate
```
np.mean(new_train_mean[df_X_train_buffer['Additive_label']==1])
```
90.82

```
np.mean(new_train_mean[df_X_train_buffer['Additive_label']==0])
```
90.56

MK-6204 dataset
Ignore glycine
If excipients column has Control: didn't add any excipient

NR_CE_SDS: device

Target properties: add up to 100
NT: non-tested

Keep Sugar and Additive as the same with optimal training data solution

```
===============Ace==========
==Among training data==
Buffer_Type_label    0.0
Sugar_Salt_label     0.0
Additive_label       1.0
pH                   4.5
task_ind             0.0
Name: 1, dtype: float64

==After simulation==
Buffer_Type_label   0.0
Sugar_Salt_label    0.0
Additive_label      1.0
pH                  3.8
task_ind            0.0
Name: 16, dtype: float64

===============His==========
==Among training data==
```

==Among training data==
Buffer_Type_label   1.0
Sugar_Salt_label   0.0
Additive_label   1.0
pH   6.5
task_ind   1.0
Name: 35, dtype: float64

==After simulation==
Buffer_Type_label   1.0
Sugar_Salt_label   0.0
Additive_label   1.0
pH   5.0
task_ind   2.0

Name: 50, dtype: float64
===============Cit==========
==Among training data==
Buffer_Type_label   2.0
Sugar_Salt_label   0.0
Additive_label   0.0
pH   5.0
task_ind   1.0
Name: 36, dtype: float64

==After simulation==
Buffer_Type_label   2.0
Sugar_Salt_label   0.0
Additive_label   0.0
pH   3.7
task_ind   1.0
Name: 38, dtype: float64

===============Phos==========
==Among training data==
Buffer_Type_label   3.0
Sugar_Salt_label   0.0
Additive_label   0.0
pH   7.0
task_ind   1.0
Name: 46, dtype: float64

==After simulation==
Buffer_Type_label   3.0
Sugar_Salt_label   0.0
Additive_label   0.0
pH   6.2
task_ind   1.0
Name: 38, dtype: float64