

C:\Users\adnan\anaconda3\lib\site-packages\sklearn\base.py:493: FutureWarning: The feature names should match those that were passed during fit. Starting version 1.2, an error will be raised.

Feature names unseen at fit time:

- molecule_id
- pXC50

Feature names must be in the same order as they were in fit.

```
warnings.warn(message, FutureWarning)
```

ValueError Traceback (most recent call last)

<ipython-input-26-02f2f41f5531> in <module>

```
41     read_and_fit_dataset(dataset)
```

```
42
```

```
--> 43 main_02()
```

<ipython-input-26-02f2f41f5531> in main_02()

```
39     datasets = ['CHEMBL203', 'CHEMBL204', 'CHEMBL205', 'CHEMBL228', 'CHEMBL233',  
'CHEMBL251', 'CHEMBL253', 'CHEMBL260', 'CHEMBL267', 'CHEMBL339']
```

```
40     for dataset in datasets:
```

```
--> 41         read_and_fit_dataset(dataset)
```

```
42
```

```
43 main_02()
```

<ipython-input-26-02f2f41f5531> in read_and_fit_dataset(dataset)

```
17
```

```
18     #predicting new sets of data using the models
```

```
--> 19     df_pred_train = base_model.predict(x_train)
```

```
20     df_pred_test = base_model.predict(x_test)
```

```
21
```

~\anaconda3\lib\site-packages\sklearn\ensemble_forest.py in predict(self, X)

```
969     check_is_fitted(self)
```

```
970     # Check data
```

```
--> 971     X = self._validate_X_predict(X)
```

```
972
```

```
973     # Assign chunk of trees to jobs
```

~\anaconda3\lib\site-packages\sklearn\ensemble_forest.py in _validate_X_predict(self, X)

```
577     Validate X whenever one tries to predict, apply, predict_proba."""
```

```
578     check_is_fitted(self)
```

```
--> 579     X = self._validate_data(X, dtype=DTYPE, accept_sparse="csr", reset=False)
```

```
580     if issparse(X) and (X.indices.dtype != np.intc or X.indptr.dtype != np.intc):
```

```
581         raise ValueError("No support for np.int64 index based sparse matrices")
```

```
~\anaconda3\lib\site-packages\sklearn\base.py in _validate_data(self, X, y, reset,
validate_separately, **check_params)
```

```
564     raise ValueError("Validation should be done on X, y or both.")
565     elif not no_val_X and no_val_y:
--> 566         X = check_array(X, **check_params)
567         out = X
568     elif no_val_X and not no_val_y:
```

```
~\anaconda3\lib\site-packages\sklearn\utils\validation.py in check_array(array, accept_sparse,
accept_large_sparse, dtype, order, copy, force_all_finite, ensure_2d, allow_nd,
ensure_min_samples, ensure_min_features, estimator)
```

```
744         array = array.astype(dtype, casting="unsafe", copy=False)
745     else:
--> 746         array = np.asarray(array, order=order, dtype=dtype)
747     except ComplexWarning as complex_warning:
748         raise ValueError(
```

```
~\anaconda3\lib\site-packages\numpy\core\_asarray.py in asarray(a, dtype, order)
```

```
83
84     """
--> 85     return array(a, dtype, copy=False, order=order)
86
87
```

ValueError: could not convert string to float: 'ChEMBL109428'