Machine Learning
Seoul National University

Homework #(**3**)
JARRY Guillaume

---

1 Q1

3) To choose the best model, the metric we used to classify the different parameters is the squared distance to the original (ie, how well the mean result of the gaussian process fit the curve).

We first ran a visual exploration of the parameter space to restrain the value to an interval that was tolerable for the computation time of approximately one hour. Here is the code used below:

```
# TODO:
# Find the best hyperparameters for GP regression
p_interval = np.arange(2, 3, 0.1)
l_interval = np.arange(3, 4, 0.1)
accuracy_list = np.empty((len(p_interval), len(l_interval)))


for i, p in enumerate(p_interval):
    for j, l in enumerate(l_interval):
        kernel = ExpSineSquared(length_scale=l, periodicity=p)
        gpr = GPRegressor(kernel=kernel, noise_level=0.5)
        gpr.fit(X_train, y_train)
        y_pred_mean, y_pred_std = gpr.predict(X_test)
        accuracy_list[i][j] = accuracy(y_pred_mean, true_f, X_test)

index = np.unravel_index(np.argmin(accuracy_list), (accuracy_list.shape))

length_scale = 3 + 0.1*index[0]
periodicity = 2 + 0.1*index[1]

kernel = ExpSineSquared(length_scale=l, periodicity=p)
gpr = GPRegressor(kernel=kernel, noise_level=0.5)
gpr.fit(X_train, y_train)
y_pred_mean, y_pred_std = gpr.predict(X_test)
print(accuracy(y_pred_mean, true_f, X_test))
```

And here is the plot obtained.

Machine Learning
Seoul National University

Homework #(**3**)
JARRY Guillaume