ML Implementing Challenge

AI Engineering Hiring

Challenge

You are presented with a multivariate **regression** problem and a blind test.

Data description

In the file *training_data.csv,* you will find the training data set of **800 samples**. Each sample has **20 features**, named *[feature_0, feature_1, … , feature_19].*

As an AI engineer, you will be in charge of training a model to predict a target variable that you will find in the *target* column of the *training_data* file*.*

Additionally, the file *blind_test_data.csv* contains **200 samples** with the same 20 features as above; you must propose a solution that allows a user to predict the missing *target* value for these samples using the model you created using the training data. Make a special effort to develop or propose a solution that allows these bulk operations to be carried out regularly.

Tasks

Using the programming language and libraries of your choice, your tasks are the following:

1. Train a model using the training data set.

2. Create a project that can be deployed easily. The PoC needs to use the model you created to provide new predictions, (use the provided blind test dataset to test your solution).
3. Describe briefly how you can evolve the solution..

Submission

You will submit by email to **ds_interviewers@wizeline.com**:

1. Your code for the challenge (if needed you can create a zip file), with brief comments explaining its purpose.
2. Your 200 predictions for the blind test in csv format with a single column ***target_pred***.

