Name: Pulkit Arora
Group: 3CO10
Roll no : 102103267
Kaggle Team Name: 102103267_Pulkit_Arora
Kaggle ID: https://www.kaggle.com/pulkittarora

Code using Random Forest Regression

```
#Import Libraries
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestRegressor
from sklearn.metrics import mean_squared_error as mse, r2_score as r2

# Load training dataset
train_data = pd.read_parquet('/kaggle/input/2024ucs654labeval1007/Lab
Eval/train_data.parquet').sample(frac=0.6, random_state=42)

# Load test dataset
test_data = pd.read_parquet('/kaggle/input/2024ucs654labeval1007/Lab
Eval/test_data.parquet')

# Reset index for both dataframes
test_data.reset_index(inplace=True)
train_data.reset_index(inplace=True)

# Preprocess data
test_ids = test_data['id']
test_data = test_data.drop('data_type', axis=1)
train_data = train_data.drop(['id', 'era', 'data_type', 'T0', 'T1', 'T2', 'T3', 'T4', 'T5', 'T6', 'T8',
'T9'], axis=1)

# Convert numeric columns to categorical
num_cols = train_data.select_dtypes(include=['int8']).columns
train_data[num_cols] = train_data[num_cols].astype('category')

# Select features and target variable
features = train_data.drop('T7', axis=1)
target = train_data['T7']

# Split data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(features, target, test_size=0.2,
random_state=42)

# Initialize Random Forest Regressor
rf_model = RandomForestRegressor(random_state=42, n_jobs=-1)

# Train the model
rf_model.fit(X_train, y_train)

# Make predictions on the testing set
y_pred_test = rf_model.predict(X_test)

# Evaluate the model
mean_error = mse(y_test, y_pred_test)
```

```
r_squared = r2(y_test, y_pred_test)

print("Mean Squared Error:", mean_error)
print("R-squared:", r_squared)

# Make predictions on the test data
pred_test = rf_model.predict(test_data)

# Create submission dataframe
submission = pd.DataFrame({'ID': test_ids, 'Target': pred_test})

# Save submission file
submission.to_csv('submission.csv', index=False)

submission.head()
```