ReadMe

# Introduction

The program utilizes both R and Python programming languages and the operating system used is Windows 10.

**The Python code uses XGBoost and randomForest libraries. The R code uses forecast, imputeTS, dplyr and prophet libraries**.

The files used and a brief description is provided below

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Type** | **Description** | **Folder Name** |
| XGB | Python Notebook | Does the XGBoost predictions | Present in PRWPython |
| RF | Python Notebook | Does the RandomForest predictions | Present in PRWPython |
| packagesInstall.R | R file | Installs the packages forecast, imputeTS , dplyr and prophet | Present in PRWR |
| config.R | R file | Sets the configuration for the working directory | Present in PRWR |
| Arima.R | R file | Implements the Arima model | Present in PRWR |
| ETS.R | R file | Implements the ETS model | Present in PRWR |
| Prophet.R | R file | Implements the Prophet model | Present in PRWR |
| Blend.R | R file | Blends the results of the Arima, ETS,Prophet,XGB and RandomForest | Present in PRWR |

# Installations

1. Install Anaconda for Windows
   1. The installer link is in <https://www.continuum.io/downloads>
   2. Choose the Python 3.6 version ( 64 bit installer )
2. Install RStudio
   1. The installer link is in <https://www.rstudio.com/products/rstudio/download/>
   2. The version to choose is **RStudio Desktop** ( Open Source license )
3. Install XGBoost for Anaconda in Windows
   1. The instructions are provided in the link <https://www.ibm.com/developerworks/community/blogs/jfp/entry/Installing_XGBoost_For_Anaconda_on_Windows?lang=en>

# Instructions

1. Extract the **PRW** zip file
2. Copy the extracted folder **PRWPython** into the startup directory of **Jupyter Notebook**
3. Invoke the Windows command line
4. Type **jupyter notebook** in the command line
5. Navigate to the folder PRWPython directory in **Jupyter Notebook**
6. Open the notebook **XGB**
7. Execute all the cells in **XGB** 
   1. This will generate the file **“predictionsXGBInterpolationSeed.csv” in** the PRWPython folder
8. Open the notebook **RF**
9. Execute all the cells in **RF**
   1. This will generate the file **“predictionsRFInterpolationWithSeed.csv” in** the PRWPython folder
10. Open the R project **PRWR** from the folder **PRWR**
11. Open the **packagesInstall.R** file
12. Do Cntrl+A to select everything in the file and click the Run button in the RStudio editor
    1. This will install the **forecast, imputeTS,prophet and dplyr packages**
13. Open the **config.R** file
14. Change the DATA\_DIR variable to “**<JupyterNotebook Startup Directory>/ PRWPython**”
15. Open the **Arima.R** file
16. Do Cntrl+A to select everything in the file and click the Run button in the RStudio editor
    1. This will generate the file **“PredictionsAutoArimaStineApprox.csv”** in the ROutput folder
17. Open the **ETS.R** file
18. Do Cntrl+A to select everything in the file and click the Run button in the RStudio editor
    1. This will generate the file **“PredictionsETSStine.csv”** in the ROutput folder
19. Open the **Prophet.R** file
20. Do Cntrl+A to select everything in the file and click the Run button in the RStudio editor
    1. This will generate the file **“ProphetScaled.csv”** in the ROutput folder
21. Open the **Blend.R** file
22. Do Cntrl+A to select everything in the file and click the Run button in the RStudio editor
23. Navigate to the folder “**<JupyterNotebook Startup Directory>/ PRWPython/ROutput**”
24. The file “**ArimaApproxETSXGBRFProphetScaled.csv**” is the FINAL submission file