

Rubric

Submission format: GitHub repository (submitted by link to canvas)

Individual Assignment

General Description: Submit to Canvas assignments a link to your GitHub repository for this project.

Why am I doing this? The goal of this assignment is to gain experience in developing a forecasting machine learning model and analyzing the differences between multiple models using time series data. This assignment challenges you to go through the entire data science pipeline in a real world scenario and allows you to provide recommendations based on your results and conclusions. This is your opportunity to think critically and not only build machine learning models, but also practice and learn how to convey your results, analyses, inferences, and thoughts in a concise and easy-to-understand manner to an unknowledgeable audience.

- Course Learning Objective: Research and model decision making
- Course Learning Objective: Machine learning programming techniques and execution
- Course Learning Objective: Analyses of results to provide practical recommendations to a real-world problem

What am I going to do? You will first carefully read through the hook document which provides background information about the topic at hand. You will be presented with a deliverable at the end of the page which describes the analysis goal that you are attempting to achieve. There are two articles attached which provide information about the current state of the housing market. Be sure to read through these articles thoroughly as you will need to link your results back to what you have read and provide recommendations to potential home buyers on when you think is the best time and season to purchase a home. You will perform any data processing and cleaning on the raw data and do any exploratory data analysis you think will be helpful in understanding the data. Then, you will develop multiple forecasting models based on what you researched and learned are the most fitting for this question and discuss your accuracy and error metric in context. Discuss your reasoning for the models created and any tuning efforts along the process.

Tips for success:

- Take time in researching. The first tutorial on Google is not necessarily the best for answering this specific research question. Think about what metric you are predicting and what parameters you have to consider to develop the most accurate model.
- Have an audience in mind before you start. Who would benefit the most from this study? What would they like to know from your results?

How will I know I have Succeeded? You will meet expectations on this case study when you follow the criteria in the rubric below.

Spec Details	
Formatting	<ul style="list-style-type: none"> • One Github Repository (submitted via link on Canvas) • The top level page should contain <ul style="list-style-type: none"> ◦ A README.md file (which auto displays) ◦ A LICENSE.md file (use MIT as default) ◦ A SRC folder ◦ A DATA folder ◦ A FIGURES folder
README.md	<ul style="list-style-type: none"> • Goal: This file serves as an orientation to everyone who comes to your repository, it should enable them to get their bearings. • SRC section <ul style="list-style-type: none"> ◦ Installing/Building your code ◦ Usage of your code • DATA section <ul style="list-style-type: none"> ◦ Data Dictionary ◦ Link to data ◦ Relevant notes about use of data • FIGURES section <ul style="list-style-type: none"> ◦ Table of contents describing all figures produced and summarizing their takeaways • REFERENCES section <ul style="list-style-type: none"> ◦ All references should be listed at the end of the Readme.md file (Use IEEE Documentation style) ◦ Include any acknowledgements
LICENSE.md	<ul style="list-style-type: none"> • Goal: This file explains to a visitor the terms under which they may use and cite your repository. • Usually, the MIT license is appropriate.
SRC folder	<ul style="list-style-type: none"> • Goal: This folder contains all the source code for your project. • Include ALL code files you produce
DATA folder	<ul style="list-style-type: none"> • Goal: This folder contains all of the data for this project • If your data fits in GitHub, place all of it here; otherwise, upload a .txt file explaining the process to obtain the dataset.
Figures folder	<ul style="list-style-type: none"> • Goal: This folder contains all of the figures generated by your project • Include with every figure relevant notes about the figure

Acknowledgements: Special thanks to Professor Alonzi for providing similar rubric outlines.