

MScFE 632 Machine Learning in Finance

Group Work Project # 2

[See grading rubric here.](#)

Scenario

The team did so well on the previous project that the portfolio strategists want more! The good news is that you learned more: Your toolkit now includes boosting (among other forms of ensemble learning), Support Vector Machines, and Neural Networks. Admittedly, these are more technical and advanced than the previous methods. Nevertheless, as the head strategist (or Spider-Man's Uncle Ben) said, "With great power comes great responsibility." Your next responsibility will be to expand the list of best practices for time series. These challenges are more difficult than the first set and include:

- Category 5: Linear Discriminant Analysis
- Category 6: Support Vector Machines
- Category 7: Neural Networks

Recall that each of these sections requires:

- **Advantages:** The benefits of using this methodology
- **Basics:** Definition and classification (e.g., this is a linear regression that...)
- **Computation:** A Jupyter notebook that illustrates the method
- **Disadvantages:** Difficulties or known issues with this methodology
- **Equations:** Equations that summarize how the model works
- **Features:** Features of the model (e.g., works well with missing values)
- **Guide:** List of inputs and outputs
- **Hyperparameters:** List of hyperparameters that need tuning
- **Illustration:** Visuals (figures, flowcharts, graphs) that show HOW the model works; can cite existing references
- **Journal:** One reference to a journal article (no Wikipedia page or blog) that illustrates the use of that technique as it is applied to finance. Do not include a required reading from this course.

- **Keywords:** Create “tags” that identify this model

Tasks

Step 1

As a group, all three members read through the feedback received from instructors for GWP1. All group members will incorporate the suggested corrections on the three items and work together to improve the previous submission.

Note: If there are only two students in the group, then the pair decides on only two items. Each member does one individually.

Step 2

As a group, for the new categories, all three members work together on writing the basics and keywords.

- Category 5: Linear Discriminant Analysis
- Category 6: Support Vector Machines
- Category 7: Neural Networks

Step 3

Individually, each student is responsible for one topic—from the new categories—by writing its:

- **Advantages:** The benefits of using this methodology
- **Computation:** A Jupyter notebook that illustrates the method
- **Disadvantages:** Difficulties or known issues with this methodology
- **Equations:** Equations that summarize how the model works
- **Features:** Features of the model (e.g., works well with missing values)
- **Guide:** List of inputs and outputs
- **Hyperparameters:** List of hyperparameters that need tuning
- **Illustration:** Visuals (figures, flowcharts, graphs) that show HOW the model works; can cite existing references
- **Journal:** One reference to a journal article (no Wikipedia page or blog) that illustrates the use of that technique as it is applied to finance. Do not include a required reading from this course.

For example:

- Team Member A works on Model 1
- Team Member B works on Model 2
- Team Member C works on Model 3

Step 4

As a group, the team members work together on the section called “**Technical Section**” where they describe how models in general have their hyperparameters tuned. This section should include examples of hyperparameters used from the individual reports.

Step 5

As a group, the team members work together on a section called “**Marketing Alpha**,” which integrates the advantages and features from the individual reports to show how ML techniques can do well! The result is 1-2 pages. The emphasis is on the new models (not those from GWP1).

Step 6

As a group, the team members work together on a section called “**Learn More**,” which combines the references to the journal articles from the individual reports and other ML websites that emphasize the strengths of ML algorithms (as opposed to websites that teach how to do it or provide code samples). Be sure the references are in proper MLA format.

Step 7

As a group, the team members work together on a section called “**Comparing Models**,” which has rows that identify features of models (e.g., “Handles Missing Data”) and columns that are specific models (e.g., “Classification Trees”). In the box corresponding to that row and column, they will give a summary (e.g., handles well or handles poorly; not applicable; etc.).

Submission Requirements and Format

One team member submits the following on behalf of the entire group:

- 1) A **zipped folder** including:
 - a) A single executable Jupyter notebook* that addresses all the computations

- b) A duplicate version of the Jupyter notebook code and output in PDF or HTML format
 - i) In order to include the output of the code, you must RUN the code before downloading the PDF.
- 2) **One PDF document** with all sections EXCEPT ANY CODE. This PDF should just contain text, formulas, and graphs, but no Python or other code. In fact, this document should contain all the sections EXCEPT any parts or sections that have code. Please be sure that code only appears in the two files above.
 - a) Use the available Report Template and fill out the required information on the first page.

****Use Google Colab or GitHub to collaborate** in completing the executable Python program.*

*The PDF file with the answers to the questions must be uploaded **separately** from the zipped folder that includes any other types of files. This allows Turnitin to generate a similarity report.*

Rubric

Your instructor will evaluate your group submission for GWP2 using the following rubric:

Quantitative Analysis (Open-Ended Questions)	Technical and Non-Technical Reports	Writing and Formatting
40 Points	30 Points	20 Points
<p>The group is able to apply results, formulas, and their knowledge of theory to real-life finance scenarios by doing the following:</p> <ul style="list-style-type: none"> • Providing all the necessary information to support their arguments. • Presenting arguments that reflect group discussion and research. • Using authoritative references to support a position and provide updated information. • Concluding with practical takeaways for more insightful financial decision-making. 	<p>Technical Reports contain 3 parts:</p> <ol style="list-style-type: none"> 1) code for each question (be sure to explicitly state the question number), 2) the corresponding output of that code, and 3) interpretations and/or recommended courses of action that reasonably follow from those results. <p>Note: Technical reports will include the technicalities of models, such as names, methods of estimation, parameter values, etc., and exclude generalities about the work done. It should NOT include names of Python code that were used.</p>	<p>A submission that looks professional should:</p> <ul style="list-style-type: none"> • Include the axes, labels, and scales in graphs. • Be free of significant grammatical errors or typos. • Be an organized, well-structured, and easy-to-read document. • Include proper citations and a bibliography in MLA format.
	<p>Non-Technical Reports contain 3 parts:</p> <ol style="list-style-type: none"> 1) clear explanation of results; 2) the recommended course of action that follows; and 3) the identification of factors that impact each portfolio. <p>Note: AVOID all references to model names, algorithms, and unnecessary details. Instead, focus on the investment decision.</p>	