

# MScFE 632 Machine Learning in Finance

## Group Work Project # 3

[See grading rubric here.](#)

### Scenario

The team did so well on the previous two projects that the portfolio strategists want more! Your group has already presented a half-dozen ideas. The strategists came back with three big questions:

1. How do we know the models have the right parameters?
2. How well can the models be expected to work for predicting future cases?
3. How can the models be used together?

Responding to these tasks will surely give the strategists confidence in your ability to use these models profitably. Let's consider each of these questions as issues.

The first question asks, "How do we know the models have the right parameters?" This question requires that we know how to optimize hyperparameters. This means we'll have to explain which hyperparameters are used, how we optimize them, and how we know they are optimal.

The second question asks, "How well can the models be expected to work for predicting future cases?" This question requires that our models work well both within-sample—on training data—as well as out-of-sample—on testing data. Bias measures the accuracy on training data, and variance measures the accuracy on testing data. As such, this question is motivating us to discuss the process of optimizing the bias-variance tradeoff.

The third question asks, "How can the models be used together?" This question requires us to think about the various ways models can be combined. Some methods exist that combine the same type of model—homogeneous ensemble methods. Other methods exist that allow us to combine different types of models—heterogeneous ensemble methods. We want to show that, when combined, multiple models can sometimes outperform individual models. This will motivate us to show how models can be used together.

So, in summary, the three topics to be addressed are:

- Issue 1: Optimizing Hyperparameters
- Issue 2: Optimizing the Bias-Variance Tradeoff
- Issue 3: Applying Ensemble Learning—Bagging, Boosting, or Stacking

However, each topic will have two sections: a technical and a non-technical section. Be careful not to mix the two up. For example, if you decide to include equations and code, those should go into the technical section. Points will be deducted if those are in the non-technical section.

Within each section, you will select the relevant areas. You may use definitions, equations, etc. Be sure that the non-technical section is very specific. It should answer each of the three questions posed above.

## Tasks

### Step 1

**As a group**, all three members read through the feedback from GWP2. All group members will incorporate the suggested corrections on the three items and work together to improve the questions submitted from GWP2.

**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**, all three members plan a strategy as to what goes in the technical and non technical section of each question. The deliverable from this section will be an outline.

- Issue 1: Optimizing Hyperparameters
  - Technical
  - Non-technical
- Issue 2: Optimizing the Bias-Variance Tradeoff
  - Technical
  - Non-technical

- Issue 3: Applying Ensemble Learning—Bagging, Boosting, or Stacking
  - Technical
  - Non-technical

**Note:** Groups of two: Pick Issue 1 **OR** 2 and Issue 3.

### Step 3

**Individually**, each student writes 3–5 pages addressing their issue. Be sure that the title of each section (technical, non-technical) is in a large font and in bold face. Within each section, use bullets (e.g., description) that are from the outline in the previous step.

### Step 4

**Individually**, each student reviews another student’s part. Here, be sure to check accuracy, clarity, spelling, etc.

	<b>Wrote</b>	<b>Reviewed</b>
Student A	Issue 1	Issue 3
Student B	Issue 2	Issue 1
Student C	Issue 3	Issue 2

**Note:** Groups of two: Each student reviews the other’s part.

Each section should indicate who was the author and who was the reviewer.

### Step 5

Revisit GWP1 and update the marketing material that persuades potential investors that the ML tools will add value.

## 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

---

MScFE | © 2022 - WorldQuant University – All rights reserved. Group Work Project 3 | PAGE 3

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 GWP3 using the following rubric:

--	--	--

<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"> <li>• Providing all the necessary information to support their arguments.</li> <li>• Presenting arguments that reflect group discussion and research.</li> <li>• Using authoritative references to support a position and provide updated information.</li> <li>• Concluding with practical takeaways for more insightful financial decision-making.</li> </ul>	<p>Technical Reports contain 3 parts:</p> <ol style="list-style-type: none"> <li>1) <b>code</b> for each question (be sure to explicitly state the question number),</li> <li>2) the corresponding output of that <b>code</b>, and</li> <li>3) <b>interpretations</b> and/or <b>recommended courses of action</b> that reasonably follow from those results.</li> </ol> <p><b>Note:</b> 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>Non-Technical Reports contain 3 parts:</p> <ol style="list-style-type: none"> <li>1) clear explanation of results;</li> <li>2) the recommended course of action that follows; and</li> <li>3) the identification of factors that impact each portfolio.</li> </ol> <p><b>Note:</b> AVOID all references to model names, algorithms, and unnecessary details. Instead, focus on the investment decision.</p>	<p>A submission that looks professional should:</p> <ul style="list-style-type: none"> <li>• Include the axes, labels, and scales in graphs.</li> <li>• Be free of significant grammatical errors or typos. <ul style="list-style-type: none"> <li>• Be an organized, well-structured, and easy-to-read document.</li> </ul> </li> <li>• Include proper citations and a bibliography in MLA format.</li> </ul>
--	--	--