Monitoring Reliability and Robustness of Agents for Dynamic Pricing in different (Re-)Commerce Markets

Monitoring der Zuverlässigkeit und Robustheit von Agenten für dynamische Bepreisung in unterschiedlichen (Re-)Commerce Märkten

Nikkel Mollenhauer

Universitätsbachelorarbeit
zur Erlangung des akademischen Grades

Bachelor of Science
(B. Sc.)

im Studiengang IT-Systems Engineering
eingereicht am 30. Juni 2022 am Fachgebiet
Enterprise Platform and Integration Concepts
der Digital-Engineering-Fakultät der Universität Potsdam

Gutachter	Dr. Rainer Schlosser
Betreuer	Johannes Huegle Alexander Kastius

Abstract

Zusammenfassung

Contents

1 Demo Chapter

This is where you can write some meta information about your chapter. For example, this chapter is based on one of my publications [NC25], and I just blindly copied everything without adjusting it. Just a heads-up warning.

Sadly, if you cite your own publications, they will appear in the bibliography. Thus, make sure to cite your papers with yourself as one of the authors.

This chapter shows off some of the basic formats of this thesis. Many packages are included in order for you to be able to start immediately without having to manually add all of the important things. The features deemed most important are now presented.

Here is just some filler text.[15] The following citations use the command textcite: Name and Co-Author [NC25]; Name et al. [Nam+30]. The first reference has a short list of authors, the second one a long list.

Also note that you can include to-do notes if necessary. Delete this chapter!

We now state a theorem and restate it later on again. Have a look at the source code in order to see how the theorem is written. Many macros are used, and all of them can be used without using math mode explicitly. Note that we can refer to inequality (1.1) as an inequality through the magic of an option in its label.

▶ **Theorem 1.1 (Variable Drift).** Let $(\mathcal{F}_t)_{t \in \mathbf{N}}$ be a filtration, $(X_t)_{t \in \mathbf{N}}$ be a random process over $\mathbf{R}_0^+$ adapted to $\mathcal{F}$, $x_{\min} > 0$, and let $T = \inf\{t \mid X_t < x_{\min}\}$. Additionally, let D denote the smallest real interval that contains at least all values $x \geq x_{\min}$ that, for all $t \leq T$, any X_t can take. Furthermore, suppose that

1. $X_0 \geq x_{\min}$ and that
2. there is a monotonically increasing function $h \colon D \to \mathbf{R}^+$ such that, for all $t < T$, we have $X_t - \mathrm{E}[X_{t+1} \mid \mathcal{F}_t] \geq h(X_t)$.

Then

$$\mathrm{E}[T \mid \mathcal{F}_0] \leq \frac{x_{\min}}{h(x_{\min})} + \int_{x_{\min}}^{X_0} \frac{1}{h(z)} \mathrm{d}z\,. \tag{1.1}$$

◀

15 Here is a footnote with a strange number (if that floats your boat). Note how the footnote mark is *above* the period at the end of the sentence.

(a) This is the caption of the subfigure that displays the logo of the HPI.

(b) This is the caption of the subfigure that displays the logo of the UP.

Figure 1.1: These are the two logos featured on the title page. Figure 1.1 (a) belongs to the HPI, whereas Figure 1.1 (b) belongs to the UP.

Please shift your attention to Figure 1.1. This reference was created using the package cleveref, which knows in what environment the label is defined in. This way, you can easily change a theorem into a lemma, and the name of the reference will be adjusted automatically. A wrapfigure like **??** is referenced just like a normal figure.

Of course, you can also use tables in a fancy style. See, for example, Table 1.1. This document already contains packages in order to also handle larger tables. Hence, it is possible to use tables spanning multiple pages or to rotate a page into landscape in order to fit in a wider table.

Before we continue, consider the following obvious theorem. We conjecture that it also holds for $n = 2$.

▶ **Theorem 1.2.** Let $a, b, c, n \in \mathbb{N}^+$ with $n > 2$. Then

$$a^n + b^n \neq c^n \, . \qquad ◀$$

Since the proof is straightforward, it is omitted. Nonetheless, we present a proof in order to show off the proof environment.

Proof of Theorem 1.2. Unfortunately, there is too little space in this PDF for the proof. ■

You can have very expressive and fancy enumerations from the package enumitem. Again, we can easily reference an item like item (i).

(i) The labels of the items can be nicely chosen.

(ii) Note how the labels are left-aligned. This does not look good but should demonstrate what is easily possible.

Table 1.1: This is a nicely formatted table. Thus, the caption is *above* the content. If not, the data could not be interpreted meaningfully. As a rule of thumb, never use vertical lines[1], and use horizontal lines sparingly. If you think that a table is illegible and thus needs vertical lines, then your spacing between columns is wrong and should be increased. Always use some whitespace first before you use some additional lines.

Text	**Number**
This is some text. Thus, it is left-aligned.	**0**
Numbers are right-aligned.	**1**
The numbers are formatted in bold using the package array.	**2**

We can even interrupt this enumeration and easily resume it immediately.

(iii) We continue where we left off.

Recall that Theorem 1.1 was as follows:

▶ **Theorem 1.1 (Variable Drift).** Let $(\mathcal{F}_t)_{t\in\mathbf{N}}$ be a filtration, $(X_t)_{t\in\mathbf{N}}$ be a random process over $\mathbf{R}_0^+$ adapted to $\mathcal{F}$, $x_{\min} > 0$, and let $T = \inf\{t \mid X_t < x_{\min}\}$. Additionally, let D denote the smallest real interval that contains at least all values $x \geq x_{\min}$ that, for all $t \leq T$, any X_t can take. Furthermore, suppose that

1. $X_0 \geq x_{\min}$ and that
2. there is a monotonically increasing function $h\colon D \to \mathbf{R}^+$ such that, for all $t < T$, we have $X_t - \mathrm{E}[X_{t+1} \mid \mathcal{F}_t] \geq h(X_t)$.

Then

$$\mathrm{E}[T \mid \mathcal{F}_0] \leq \frac{x_{\min}}{h(x_{\min})} + \int_{x_{\min}}^{X_0} \frac{1}{h(z)} \mathrm{d}z\ . \tag{1.1}$$

◀

Note that the reference above still refers to the first occurrence of the theorem. However, the theorem is repeated without any noise. That is, it is identical to the other occurrence.

From the next page on, other than a warp figure and some filler text, there is not much more to see. Thank you very much for taking your time and reading so far. I hope you got an impression of what this template is capable of. Have fun using it, and create a great thesis!

1 Except you know what you are doing.

2 Introduction

This chapter will introduce the underlying concepts of the Reinforcement Learning approach and prepare us for the introduction of measures to monitor and evaluate agents trained in such a way.

Demo Citation: [KLM96]

Short description of Reinforcement learning:
- Agent interacts with the environment only on the basis of a state, which contains information about the environment that the agent can use to decide on one of a number of specified actions it can take.

2.1 An Overview of Reinforcement-Learning

Agents to be trained for real-world use Training in an isolated environment Need to make sure they are "good" What we want to offer with our framework Determining the grade of an agent using monitoring

3 Related Work

3.1 Approaches to evaluating RL-agents

3.1.1 ...on the fly (while training)

3.1.2 ...after training has finished

4 What makes a good agent?

4.1 Good agent = high profit, few outliers

4.2 Overview of market components

4.2.1 Focus on how agents make profit etc.

4.3 How realistic the market is

4.3.1 Restrictions for evaluation arising from this

5 Different approaches

5.1 During vs. After training

5.2 Tensorboard? (Not built by us)

5.3 Macro

5.3.1 Agent-monitoring

5.3.2 Live-monitoring

5.4 Micro

5.4.1 Exampleprinter

5.5 Static

5.5.1 Policyanalyzer

6 Our workflow

6.1 Training continuously saves models

6.1.1 Automatic monitoring at certain intervals

6.1.2 -> Can we discard agents prematurely due to results from this?

6.1.3 First analysis if available with finished training

6.2 Manual invocation of monitoring functionalities

6.2.1 When is this necessary/a good idea? Why?

7 Interpreting the results

7.1 Graphs and diagrams are available...

7.1.1 ...comparing with other agents/models

7.1.2 ...which hyperparameters influence the results in what ways?

7.1.3 ...can we augment e.g. Grid-Search with our analysis?

7.1.4 -> Would need to make results "machine-readable" again

8 Conclusions & Outlook

Bibliography

[KLM96] Leslie Pack Kaelbling, Michael L. Littman and Andrew W. Moore. **Reinforcement Learning: A Survey**. *Journal of Artificial Intelligence Research* 4 (1996), 237–285 (see page 5).

[Nam+30] My Name, First Co-Author, Second Co-Author, Third Co-Author and Fourth Co-Author. **Dear Lord! How Did This Get Accepted?** *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik* 42:1 (2030), 2–1024 (see page 1).

[NC25] My Name and A Co-Author. **Useless Stuff That No One Cares About**. In: *Proceedings of the coolest Annual ACM Symposium on Theory of Computing (STOC'XX)*. ACM Press, 2025, 42–1337 (see page 1).

Declaration of Authorship

I hereby declare that this thesis is my own unaided work. All direct or indirect sources used are acknowledged as references.

Potsdam, 4th April 2022

Nikkel Mollenhauer