



---

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

Monitoring der Zuverlässigkeit und Robustheit von  
Agenten zur dynamischen 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**  
**Betreuer**

Dr. Rainer Schlosser  
Johannes Huegle  
Alexander Kastius



# Abstract

---



# Zusammenfassung

---



# Contents

---

<b>Abstract</b>	<b>iii</b>
<b>Zusammenfassung</b>	<b>v</b>
<b>Contents</b>	<b>vii</b>
<b>1 Demo Chapter</b>	<b>1</b>
<b>2 Introduction</b>	<b>5</b>
2.1 Introduction to the Recommence platform . . . . .	5
2.1.1 The Circular Economy model . . . . .	5
2.1.2 Using the simulated marketplace to train agents . . . . .	5
2.2 Reliability and Robustness . . . . .	6
<b>3 Related Work</b>	<b>7</b>
3.1 Approaches to evaluating RL-agents . . . . .	7
3.1.1 ...on the fly (while training) . . . . .	7
3.1.2 ...after training has finished . . . . .	7
<b>4 What makes a good agent?</b>	<b>9</b>
4.1 Good agent = high profit, few outliers . . . . .	9
4.2 Overview of market components . . . . .	9
4.2.1 Focus on how agents make profit etc. . . . .	9
4.3 How realistic the market is . . . . .	9
4.3.1 Restrictions for evaluation arising from this . . . . .	9
<b>5 Different approaches</b>	<b>11</b>
5.1 During vs. After training . . . . .	11
5.2 Tensorboard? (Not built by us) . . . . .	11
5.3 Macro . . . . .	11
5.3.1 Agent-monitoring . . . . .	11
5.3.2 Live-monitoring . . . . .	11
5.4 Micro . . . . .	11
5.4.1 Exampleprinter . . . . .	11

5.5	Static . . . . .	11
5.5.1	Policyanalyzer . . . . .	11
<b>6</b>	<b>Our workflow</b>	<b>13</b>
6.1	Training continuously saves models . . . . .	13
6.1.1	Automatic monitoring at certain intervals . . . . .	13
6.1.2	-> Can we discard agents prematurely due to results from this? . . . . .	13
6.1.3	First analysis if available with finished training . . . . .	13
6.2	Manual invocation of monitoring functionalities . . . . .	13
6.2.1	When is this necessary/a good idea? Why? . . . . .	13
<b>7</b>	<b>Interpreting the results</b>	<b>15</b>
7.1	Graphs and diagrams are available... . . . .	15
7.1.1	...comparing with other agents/models . . . . .	15
7.1.2	...which hyperparameters influence the results in what ways? . . . . .	15
7.1.3	...can we augment e.g. Grid-Search with our analysis? . . . . .	15
7.1.4	-> Would need to make results "machine-readable" again . . . . .	15
<b>8</b>	<b>Conclusions &amp; Outlook</b>	<b>17</b>
	<b>Bibliography</b>	<b>19</b>
	<b>Declaration of Authorship</b>	<b>21</b>



*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.<sup>15</sup> 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.

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.

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

► **Theorem 1.1 (Variable Drift).** Let  $(\mathcal{F}_t)_{t \in \mathbb{N}}$  be a filtration,  $(X_t)_{t \in \mathbb{N}}$  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  $h: D \rightarrow \mathbf{R}^+$  such that, for all  $t < T$ , we have  $X_t - \mathbf{E}[X_{t+1} \mid \mathcal{F}_t] \geq h(X_t)$ .

Then

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

<sup>15</sup> 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

$$a^n + b^n \neq c^n .$$



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<sup>1</sup>, 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.	<b>0</b>
Numbers are right-aligned.	<b>1</b>
The numbers are formatted in bold using the package array.	<b>2</b>

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  $(\mathcal{F}_t)_{t \in \mathbb{N}}$  be a filtration,  $(X_t)_{t \in \mathbb{N}}$  be a random process over  $\mathbb{R}_0^+$  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  $h: D \rightarrow \mathbb{R}^+$  such that, for all  $t < T$ , we have  $X_t - \mathbb{E}[X_{t+1} \mid \mathcal{F}_t] \geq h(X_t)$ .

Then

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

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!

<sup>1</sup> Except you know what you are doing.



*This thesis builds upon the bachelors project "Online Marketplace Simulation: A Testbed for Self-Learning Agents" of the Enterprise Platform and Integration Concepts research group at the Hasso-Plattner-Institute. Therefore, the project will be referenced and all examples and experiments, unless otherwise mentioned, will have been conducted using its framework.*

## 2.1 Introduction to the Recommerce platform

### 2.1.1 The Circular Economy model

The main goal of the aforementioned bachelors project was to develop a realistic online marketplace that simulates a Circular Economy. A market is most commonly referred to as being a "Circular Economy" if it includes the three activities of reduce, reuse and recycle [KRH17]. This means that while in a classical Linear Economy market each product is being sold once at its *new price* and after use being thrown away, in a Circular Economy, recycling and thereby waste reduction is a major focus. In our project, we modelled this by adding two additional price channels, *re-buy price* and *used price*, to the pre-existing *new price* of a product.

The *re-buy price* is defined as the price a vendor is willing to pay a customer to buy back a used product, while the *used price* is defined as the price the vendor sets for products they previously bought back and now want to sell alongside new products (whose price is defined by the *new price*).

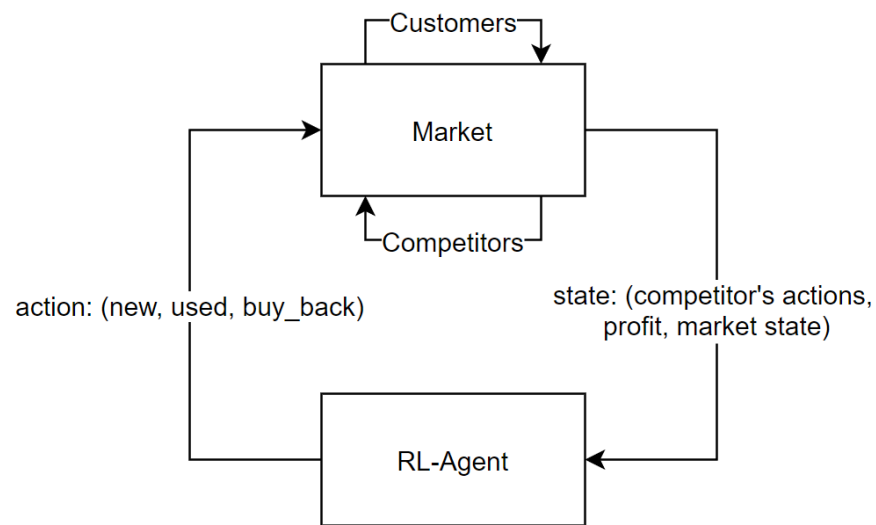
From now on, we will refer to the Circular Economy market simply as *the market*.

### 2.1.2 Using the simulated marketplace to train agents

After the initial market was modelled the goal was to train agents using different reinforcement-learning algorithms to dynamically set prices on this marketplace, both in monopolistic scenarios as well as in competition with rule-based vendors which set prices following a strict set of pre-defined rules. These rules can range from simply undercutting the lowest competitor's price to more advanced techniques such as price-fixing and -gouging.

Is the following sentence a footnote?

find out if we have such competitors



**Figure 2.1:** The standard reinforcement-learning model in the context of our market.

Reinforcement-learning agents are trained through a process of trial-and-error. They interact with the market through an observable state and an action which influences the following state. [Figure 2.1](#) illustrates the RL-model in the context of our market. The goal of the agent is to maximize the so-called reinforcement signal, which in our case is the profit the agent made during the last cycle, since we want to train agents to maximize profits on real markets. By observing which prices lead to which reinforcement signal, the agents get more profitable over the course of training.

Create a *nicer* diagram with the three prices

## 2.2 Reliability and Robustness

Term definitions, Ambiguity

### **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**



### **6.1 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

---

- [KRH17] Julian Kirchherr, Denise Reike and Marko Hekkert. **Conceptualizing the circular economy: An analysis of 114 definitions**. *Resources, Conservation and Recycling* 127 (2017), 221–232. ISSN: 0921-3449. DOI: <https://doi.org/10.1016/j.resconrec.2017.09.005>. URL: <https://www.sciencedirect.com/science/article/pii/S0921344917302835> (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, 5th April 2022

---

Nikkel Mollenhauer