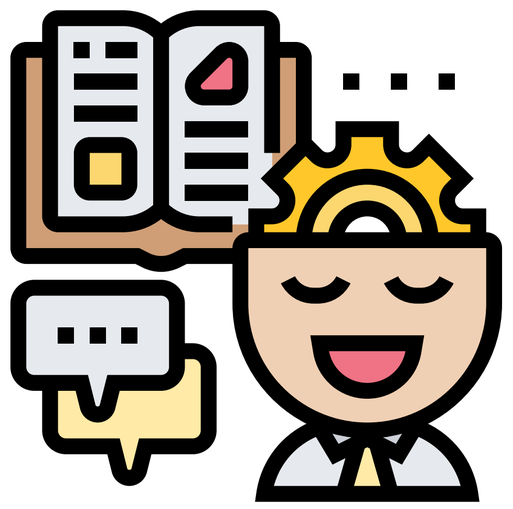
Implementing Reinforcement Learning on Next Best Action:
the data requirements and preparation



This master's thesis undertakes a rigorous examination of the applicability of Reinforcement Learning (RL) techniques for Next Best Action (NBA) in marketing, by embarking on the data requirements and preparation process. It thus conducts analyses on the limitations and requirements of marketing data needed to train RL models.

Author
Khanh Ha Dao

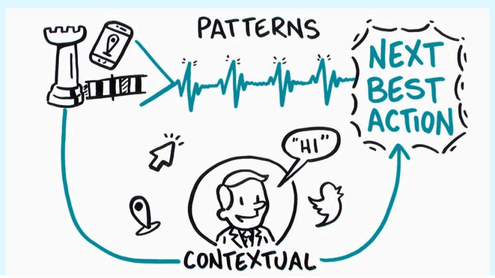
Education institution
University of Hasselt
Business and Information Systems Engineering

Promotor
Prof.dr. Koenraad Vanhoof

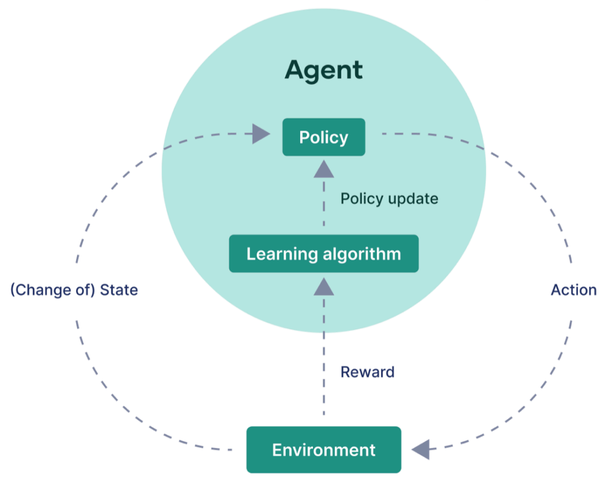
Assistent
Leen Jooken

Introduction

The concept '**Next Best Action**' has emerged as a powerful strategy to enhance customer engagement and optimize marketing campaigns.



Reinforcement Learning



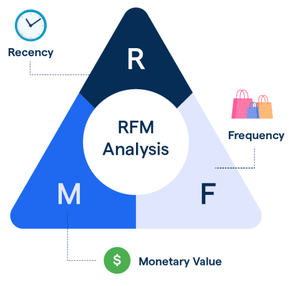
Objective

What are the **requirements** regarding marketing data, which is used as input in RL, in application to the field of NBA?
Are RL models **suitable** to implement in the context of NBA?

Methodology

- **Exploratory literature review**: theoretical knowledge about RL, NBA and data requirements using scientific databases and online libraries
- **Empirical experiment**: RFM-analysis to obtain required input data

RFM-analysis



Experiment

The data preparation analysis is conducted to obtain the key inputs for an RL model:

- **State space**: defining loyalty levels
- **Action space**: combining *Sub-category* and *Discount*
- **State transition probability matrices**: calculating probabilities of state transitions
- **Rewards**: assigning scores to each state transition

State transition probability matrix

	Need attention	Promising	Potential loyalists	Loyalists
Need attention	0.23	0.06	0.19	0.52
Promising	0.13	0.07	0.13	0.67
Potential loyalists	0.21	0.11	0.05	0.63
Loyalists	0.29	0.11	0.11	0.49

Reward matrix

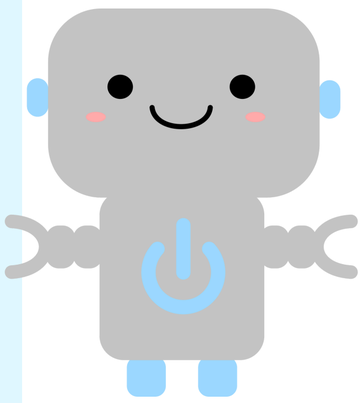
	Need attention	Promising	Potential loyalists	Loyalists
Need attention	0	1	2	3
Promising	-1	0	1	2
Potential loyalists	-2	-1	0	1
Loyalists	-3	-2	-1	1

Results

Results obtained are **5 state transition probability matrices** and a **reward matrix**. Each cell of the probability matrix represents the probability that a certain customer state (row) will transition to another customer state (column) after engaging with a certain marketing action. Each cell of the reward matrix represents the corresponding amount of reward/punishment of each state transition.

Discussion

When working with raw marketing data, **data analysis** and **transformation** have to be conducted to be able to obtain the necessary inputs for the RL algorithms. Findings from this experiment suggest that it is **feasible** to use RL to support NBA-systems. The data preparation analysis is however quite extensive and requires **exhaustive procedures**.



UHASSELT