Data Intelligence Application 2018/19 Project



Advertising

Antoniazzi Matteo (895712)

Bonali Luca (896641)

Chittò Pietro (899045)

Lamparelli Andrea (894005)

Ravelli Leonardo (894222)

Index

1. Introduction

The product we decided to use for this project is the AMAZON ECHO PLUS. We’re basing our project on this version of the product w.r.t. other ones (like Echo Dot, Echo Sub etc.) because, even if the Echo Dot is the more sold product for its small price, it lacks some features like the hub for connecting more smart devices to make a more advanced “smart home”.

1.1 Product description

The second generation of Amazon Echo Plus was released in September 2018.

It’s a smart speaker developed by Amazon. Echo devices connect to the voice-controlled intelligent personal assistant service Alexa, which responds to the names "Alexa", "Echo", or "Computer". The features of this device include: voice interaction, music playback, making to-do lists, setting alarms, streaming podcasts, in addition to providing weather, traffic and other real-time information. It can also control several smart devices, acting as a home automation hub like, for example, Smart TV, specific appliances, lamps, windows, doors, temperature control etc. It shares design similarities with the first-generation Echo, but also doubles as a smart home hub, connecting to most common wireless protocols to control connected smart devices within a home.

We consider it more customizable than its direct competitor “Google Home” and Alexa's encourages faster and broader development and support from third-parties of its skills market.

2. Classes and environment description

2.1 Features Selection

We describe our possible customers by means of 3 main features, with the following values:

* Age: Young, Adults, Retires
* Home status: Living alone, Living with family
* Welfare: Normal, Richer

We decided to divide w.r.t. the age because younger people are more inclined to accept new technologies to make their lives simpler than older ones. We also thought that a person who lives alone is more worn to have something that can help and fasten the way he approaches some of his daily tasks, especially if young. Of course, we consider that this product is a commodity and so, even if the price is not too large, some class of customer may not be interested in buying our product.

2.2 Class descriptions

In the following tables we show how, using the previous explained features, we’ve created our main class of customers.

For readability, we split the 3D features tensor into 2 tables according to the feature home status. In each cell of the table we reported the probability of a user to belong to that specific class.

Each colour represents one class.

|  |  |  |  |
| --- | --- | --- | --- |
| **FAMILY 0.40** | Young 0.50 | Workers 0.40 | Retires 0.1 |
| Richer 0.55 | 0,11 | 0,088 | 0,022 |
| Normal 0.45 | 0,09 | 0,072 | 0,018 |

|  |  |  |  |
| --- | --- | --- | --- |
| **ALONE 0.60** | Young 0.50 | Workers 0.40 | Retires 0.1 |
| Richer 0.55 | 0,165 | 0,132 | 0,033 |
| Normal 0.45 | 0,135 | 0,108 | 0,027 |

* **Class 1:** this class is characterized by people who live alone, like young student or worker, who have a good economic wellness and are more willing to spend money on this commodity. Young families with economic possibilities are in this class too. We assume that this is the class that we expect to click more on our advertised product.

* **Class 2:** in the second class we have other kinds of families, from younger to older, which are less incline to spend money on a product like that, so they will less probably click on our ad. Considering age , we can include in this class also older families, or retired couples, who have money to spend, and are curious about these new technologies.
* **Class 3:** this class is composed by all those people who, for different reasons, are not so interested in our product, can be for economic reason for example. But advertising the product on those people, may bring some of them to click on the product anyway.

2.3 Phases

We identified 4 different phases in our scenario:

1. Market launch: this is the first phase, when the product enters the market. We assumed that for the first 3 months the demand for all the classes remains approximatively the same, after that, we hypothesized some smooth changes, in particular in the medium-high price range.
   * Class 1: here the demand is overall high for the prices below 1000€, after which decreases. We assumed that the customers evaluate our product basing on the price of the previous model at the same phase.
   * Class 2: in this class the customers have less money than the previous ones, so the demand decreases if the price exceeds 500€.
   * Class 3: the demand is generally low since the customers of this class are not very interested in buying our product.

|  |  |  |
| --- | --- | --- |
| Phase | Period | Duration |
| Market launch | From February to August | 7 months |

1. Competitors’ new products: we assumed that in September the 2 main Samsung’s competitors (Apple Inc. and Huawei) decide to release their new products. This leads to an abrupt change in the demand that decrease drastically for high prices.
   * Class 1: the demand falls for prices above 400€; we assumed that this kind of customers prefers, cost being equal, to buy the new smartphone in the market (we’ve assumed the new iPhone model).
   * Class 2: similar consideration for this type of customers, that prefers to buy a new and cheaper phone (the new Huawei model in this case).
   * Class 3: the demand softly decreases and there is a flattening of the demand of what we previously called rich members. As we already said, these few people consider the last released smartphone as the best in the market.

|  |  |  |
| --- | --- | --- |
| Phase | Period | Duration |
| Competitors’ new product | From September to November | 3 months |

1. Holiday: we hypothesized that during winter holidays, due to the fact that there are several festivities, people increase the demand, because, usually, technologic product are a very popular gift.
   * Class 1: the demand remains more or less the same for lower prices (people take advantage of holiday offers for example) and increase a little bit for medium-higher prices (this kind of customers allows himself to spend a little more for a gift)
   * Class 2: here the demand increases a little bit for lower prices (again for holiday offers) and remains the same for the others.
   * Class 3: also in this class the demand increases a little, especially for very high prices, due to the rich members. However, it decreases for medium prices range (we assumed that retirees – not the rich group – don’t spend too much money for a not well-known product).

|  |  |  |
| --- | --- | --- |
| Phase | Period | Duration |
| Holiday | From December to January | 2 months |

1. New model: Samsung releases the new smartphone model (Samsung Galaxy S11). This is the last phase we decided to consider.
   * Class 1: because this class of customer is predominantly composed by wealthy people, as soon as the new model of a smartphone is released, the demand strongly decreases, also due to the fact that, in general, the price of the new model is close to the previous model release price.
   * Class 2: the demand increases for low prices because usually, as soon as the new model is released, the prices of the previous models decrease. Instead, for higher prices it decreases due to the fact that this kind of customers prefer to keep their money for the new model.
   * Class 3: the demand trend returns as in the competitors’ new product phase. The motivations are very similar of the ones expressed in the aforementioned phase.

|  |  |  |
| --- | --- | --- |
| Phase | Period | Duration |
| New model | From February to April | 3 months |

3. Time horizon and candidates

3.1 Time horizon

The time horizon we decided to consider (as already introduced in the phases description) starts from February 2019 and it ends in April 2020.

3.2 Candidates selection

We considered a price range from 0€ to 1400€ and we identified 19 possible prices/candidates:

We divided the range into intervals of 100€, except for the prices from 500€ to 900€ where we reduced the interval to 50€, because we assumed that in this range a small variation of the price would lead in a significant variation in customers behaviour and so in the demand curve.

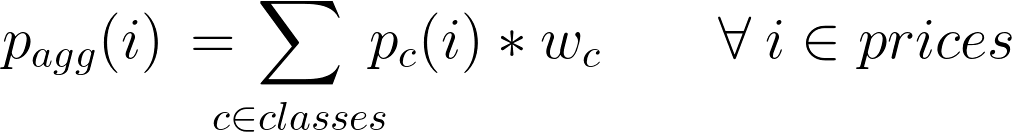
|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 100 | 200 | 300 | 400 |  | 1000 | 1100 | 1200 | 1300 | 1400 |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 500 | 550 | 600 | 650 | 700 | 750 | 800 | 850 | 900 |

Despite we already known that the **highlighted prices** would get us a negative reward (since we supposed a production cost of 350€ for unit) we decided to leave them in our algorithm to have a more complete analysis.

4. Aggregated demand curve

In this section we’ll show the result we obtained considering the aggregated demand curve. This curve is computed as the weighted sum of the demands of the single classes:



Where pagg(i) is the percentage of customers that would buy our product at price *i*. In our experiment we use these percentage as probabilities for the conversion rate.

4.1 K-testing

Because the Sequential A/B testing is not practical and difficult to perform in a non-stationary environment, we decided to consider only the first phase demand curve during all the time of the test.

We perform the sequential A/B testing comparing the candidates 2 by 2 from the lower to the higher.

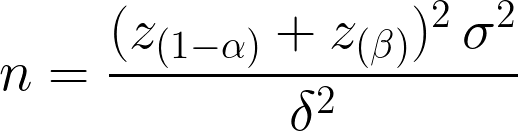
* Hypothesis definition:

|  |  |  |
| --- | --- | --- |
| Hypothesis | | Action performed |
| H0 | u1 = u2 | Select the new price/candidate |
| H1 | u1 > u2 | Keep the old price/candidate |

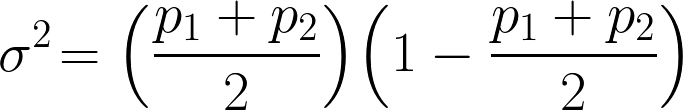
* Accuracy selection:

|  |  |
| --- | --- |
| Parameter | Value |
| Significance level | 0.005 |
| Power level | 0.85 |
| Alternative hypothesis relaxing coefficient | 0.05 |

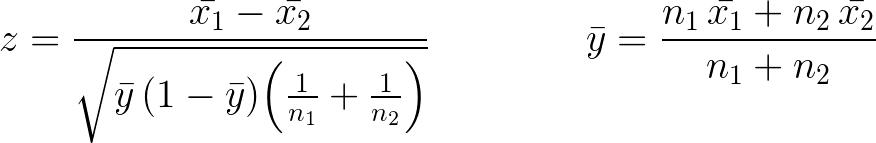
* Number of sample selection. Each time we perform the single A/B testing we compute the minimum number of samples needed to take a good decision. This is done at every comparison because the number of samples is dependent by the standard deviation of the two current candidates:



With:



* Statistic computation: if *z* is greater than *z(1-α)* then the null hypothesis is rejected.



If, for some reason, one between x1 or x2 is 0, we directly use the probability associated to each candidate (respectively p1 and p2).

At the end of the experiment the Sequential A/B testing gave us XXX€ as result. [other considerations to put here]

4.2 UCB1/TS

Results of the UCB1 and Thomson Sampling algorithms applied in the non-stationary environment that we’ve described previously (*See chapter: 2.3 Phases*).

4.2.1 UCB1

4.2.2 TS

4.3 SW-UCB1/SW-TS

Results of the Sliding Window UCB1 and Sliding Window Thomson Sampling algorithms applied in the non-stationary environment that we’ve described previously (*See chapter: 2.3 Phases*).

Here we set the length of the Sliding Window as: SOMETHING

4.3.1 SW-UCB1

4.3.2 SW-TS

5. Disaggregation

5.1 Description

5.2 UCB/TS

5.2.1 UCB1

5.2.2 TS

5.3 SW-UCB1/SW-TS

5.3.1 SW-UCB1

5.3.2 SW-TS

6. Conclusion