

---

# Rapport de stage

## Sujet : Chaînage concurrentiel automatique de produits

---

MATTHIEU AKHAVAN

MENTION DONNÉES ET CONNAISSANCES  
PROMOTION 2021-2022  
MASTER D'INFORMATIQUE

11/04/2022 - 30/09/2022

***Tuteur Universitaire :***  
Mention : CHLOÉ BRAUD  
chloe.braud@irit.fr

***Tuteurs Entreprise :***  
GUILLAUME HALB  
guillaume.halb@mercio.io  
YOAN OZANNEAUX  
yoann.ozanneaux@mercio.io

## Résumé

Vous lisez le rapport de mon stage de 5 mois et deux semaines dans l'entreprise Mercio, à Paris, spécialisé dans le retail et la big data. Apporter ma pierre à l'édifice dans une équipe recherche et développement est très formateur. Identifier et répondre à un besoin via la recherche d'une solution à concevoir m'a introduit au monde du travail en entreprise.

C'est jusqu'à ce jour l'expérience professionnelle la plus complète que j'ai eue au cours de mes études en informatique.

Utiliser en cas concret les compétences longuement acquises au cours de mon parcours universitaire est un formidable enjeu. Mon sujet tourne autour d'une problématique qui allie des enjeux purement autour des données et de l'intelligence artificielle, mais contient aussi un aspect développement logiciel pour l'implémentation de la solution algorithmique trouvée. Le tout, répond à un besoin concret présent dans la grande distribution : l'alignement à la concurrence.

Le défi est de taille : au vu du poids des acteurs de la grande distribution dans l'économie contemporaine, le fait de s'aligner sur les prix de sa concurrence est crucial en tant que vendeur. Ainsi, d'un point de vue concurrentiel, faire correspondre les prix des produits identifiés comme similaires par le consommateur lambda passe d'abord par relier des produits entre eux. La tâche est simple pour des produits identiques, mais qu'en est-il des produits propres à chaque enseigne qui se compare ? Quelle solution algorithmique mettre en œuvre pour répondre à la problématique ?

Des solutions axées machine learning ont été proposées lors de ce stage, certaines ont pu être testées, avec un succès relatif. Les démarches mises en œuvres seront exposées sous un œil critique, actuellement à l'orée du 5<sup>e</sup> mois de stage. Mercio désirant obtenir une solution qui puisse convenir à une bonne partie de sa clientèle, j'ai également été formé à plusieurs outils de développement très utilisés dans le mode du travail, en plus de m'intégrer dans un cycle de développement Agile.

D'un point de vue technique, le développement de l'algorithme s'est fait en Python, le développement logiciel s'est fait en Java et en React. J'ai donc appris des compétences très solides dans le monde du travail dans des domaines clés, variés, mais ciblés de l'informatique d'aujourd'hui et de demain, sans pour autant dévier de mon sujet initial. Ce sont des compétences qu'il n'est possible d'acquérir de manière solide qu'en les appliquant avec un véritable objectif.

Des apports concrets et pérennes (bien que partiels à cette date) ont bien été créés pour l'entreprise d'un point de vue amélioration du logiciel, qui sert d'outil à leurs clients désireux de s'aligner plus facilement à la concurrence.

Dans tous les cas, j'ai eu l'impression d'avoir énormément appris en peu de temps, et je tache de le faire transparaître à travers ce rapport.

## Abstract

You are reading the report of my 5 months and two weeks internship in the Mercio company, located in Paris and specialized in retail and big data. Being part of the workforce in a research and development team is very formative. Identifying and responding to a need, to search of a solution to design in this purpose, deeply introduced me to the world of corporate work.

This is by far the most comprehensive work experience I have ever had during my studies in computer science. To use in a concrete case the skills I acquired over the course of my university journey represents an exciting challenge. My subject is about an issue that combines both big data and artificial intelligence topics. It also contains a software development aspect for the implementation of the algorithmic solution that would be found. It globally answers to a concrete need in the retail industry: alignment with the competition.

The challenge is significant: given the huge weight of the retail sector in the contemporary economy, aligning with the competition is a major challenge. The fact of being aligned with the prices of within its own competition range is crucial for retailers. Thus, from a competitive point of view, matching the prices of products identified as similar by the average consumer starts with linking products together. This is a simple task for identical products, but what about products that are unique to each retailer that are compared? What algorithmic solution should be implemented to address this issue?

Machine learning solutions were proposed during the elaboration of such an algorithm, some of them were tested, with a relative success for each. The chosen approaches will be exposed under a critical eye, currently at the beginning of the 5th month of internship. Mercio wishing to obtain a solution which could suit a good part of its customers, I also was trained in several development tools used in the work industry, in addition to integrating myself in an Agile development cycle.

From a technical point of view, the development of the algorithm was done in Python, the software development was done in Java and React. So I learned very solid skills in central, various, but focused areas of today's and tomorrow's computer science, all without deviating from my initial subject. These are skills that can only be acquired in a solid way through applying them towards a real objective.

Substantial contributions (although partial at this date) have indeed been created for the company from a software improvement point of view. These changes will be useful for Mercio customers to align themselves more easily with the concurrency. Overall, I felt like I learned a lot in a short amount of time, and that's what I try to convey in this report.

---

Remerciements sincères à toute l'équipe de Mercio...

# Table des matières

<b>1</b>	<b>Mise en contexte</b>	<b>5</b>
1.1	Introduction générale . . . . .	5
1.2	Modèle économique de l'entreprise . . . . .	6
1.3	Structure de l'application . . . . .	6
1.4	L'équipe Produit et l'équipe Projet . . . . .	7
1.5	Environnement de travail et matériel . . . . .	7
<b>2</b>	<b>Sujet du stage : Alignement horizontal de produits concurrents</b>	<b>8</b>
2.1	Définition du besoin . . . . .	8
2.2	Contexte de l'application . . . . .	9
2.3	Pertinence d'un lien de chaînage horizontal . . . . .	10
2.4	Répondre à la problématique et au besoin . . . . .	11
2.5	Données initiales à disposition . . . . .	11
<b>3</b>	<b>Démarche proposée</b>	<b>12</b>
3.1	Agilité . . . . .	12
3.2	Planification des travaux . . . . .	12
<b>4</b>	<b>Travail réalisé</b>	<b>13</b>
4.1	Analyse de la tâche . . . . .	13
4.2	Solutions possibles et leurs justifications . . . . .	13
4.3	Résultats expérimentaux, tests . . . . .	13
<b>5</b>	<b>Bilan de production</b>	<b>14</b>
5.1	Objectifs atteints . . . . .	14
5.2	Apport du travail réalisé . . . . .	14
5.3	Perspectives d'évolution des travaux . . . . .	14
5.4	Bilan général de la démarche . . . . .	14
<b>6</b>	<b>Bilan sur le plan personnel</b>	<b>15</b>
6.1	Compétences mises en avant . . . . .	15
6.2	Projet professionnel . . . . .	15
<b>7</b>	<b>Conclusion</b>	<b>16</b>
<b>8</b>	<b>Glossaire</b>	<b>17</b>
<b>9</b>	<b>Annexes</b>	<b>18</b>

# 1 Mise en contexte

## 1.1 Introduction générale

Mercio, de raison sociale HereForRetail, est une société dans le secteur du retail. C'est-à-dire que ses clients sont les maillons finaux de la chaîne de la grande distribution française, qui vendent des produits en rayons, que ce soit dans l'alimentaire, le bricolage ou les jouets. Mercio propose à ses clients un logiciel web qui leur permet d'ajuster leurs prix en rayon selon les stratégies définies en interne. Les stratégies s'expriment alors sous forme d'un ensemble de règles de prix configurables sur l'application Mercio Pricing, dont toutes les données sont emmagasinées dans le nuage. Les règles que ces clients peuvent appliquer à leurs produits ont des priorités entre elles selon la stratégie de l'enseigne et sont paramétrables en profondeur : il y a la très connue règle d'arrondi (prix qui fini par 99 centimes par exemple), l'alignement par rapport à la concurrence, et bien d'autres.

Avec les milliers de magasins que certains clients ont sur le territoire français et les milliers de produits vendus en magasins, le nombre de prix à traiter pour les plus gros clients va jusqu'à plusieurs dizaines de millions de prix. Le fait que la plateforme est capable de traiter autant de prix en peu de temps (une dizaine de minutes la plupart du temps) place Mercio en qualité d'entreprise Big Data dans le vaste secteur de la tech.

Au fil des avancées dans le domaine lucratif des sciences des données, de plus en plus d'entreprises proclament utiliser des techniques d'intelligence artificielle pour obtenir des performances calculatoires jamais égalées par le passé, et ce, avec des technologies de plus en plus à la pointe dans leur domaine. Mercio n'est pas en reste, et c'est ainsi que j'ai répondu à leur appel d'offre pour résoudre un challenge de data science en tant que développeur stagiaire en fin de Bac+5 Informatique.

Mon sujet au sein de l'entreprise est de développer une solution algorithmique et logicielle pour faire une suggestion de correspondance entre chaque produit donné d'une enseigne client et le produit jugé le plus similaire possible chez chaque concurrent dont le client met à disposition les données. Ce sujet, ses enjeux pour l'entreprise, mais aussi pour moi en tant qu'apprenti, seront bien plus développés dans les parties suivantes.

Cependant, avant de parler en profondeur du sujet, prendre connaissances des descriptions qui suivent est capitale pour comprendre la place que j'ai en tant que stagiaire dans cette entreprise de 21 employés, stagiaires compris (au moment où je rédige cette introduction) car la dimension économique, mais également le plan technologique dans lesquels se situe Mercio influencent grandement sa structure et ses cycles de production. Tout autant sa manière de fonctionner que les solutions technologiques choisies, rien n'est laissé au hasard, mais bel et bien conçu pour répondre à des besoins fonctionnels spécifiques et à des moments clés du cycle de vie de l'entreprise.

## 1.2 Modèle économique de l'entreprise

La licence du logiciel de pricing est vendue pour trois ans et tous les différents modules de l'application ne sont pas fournis à tous les clients, et de façon différente. En effet, certains clients ne sont intéressés que par certains modules. Ces clients peuvent aussi négocier des modifications, voire fonctionnalités supplémentaires qui sont facturées en fonction de la masse de travail requise pour ce faire. Ce mode de fonctionnement est appelé B2B pour « Business to Business » en opposition au B2C « Business to Customer ».

Les développeurs, qui représentent le plus gros des employés de Mercio, sont scindés en deux équipes : l'équipe projet qui traite avec les clients directement et appliquent des correctifs et changements selon leurs besoins, et l'équipe Produit qui développe les fonctionnalités standard sur le long-terme. Je travaille actuellement sur mon sujet dans l'équipe produit, qui s'apparente davantage au secteur recherche et développement de beaucoup d'entreprises. Nous reviendrons dessus plus tard.

Les enseignes de retail avec qui échangent les commerciaux de l'entreprise ne sont pas exclusivement des enseignes françaises, mais aussi européennes, notamment en Allemagne. Mercio Pricing est reconnu par beaucoup d'acteurs économiques clés du retail comme une solide solution de pricing. Le pricing est le fait d'attribuer un prix à un produit selon ses besoins stratégiques de vente. Cette solution logicielle se positionne bien sur le marché en termes de qualité de recommandation de prix et de quantité de données traitées par jour, avec des atouts stratégiques, comme le géo-pricing, une méthode d'attribution de prix qui prend en compte la concurrence locale de chaque magasin de l'enseigne cliente de Mercio.

## 1.3 Structure de l'application

L'application Mercio Pricing est une solution web, accessible par les clients dans les horaires prévues à cet effet en semaine. La solution d'un client est déployée sur un serveur distant et accessible par identifiant et mot de passe sécurisés.

Une fois connecté, le client a accès au service de pricing grâce aux données qu'il fournit sur son site de stockage Microsoft Azure dédié.

Pour synthétiser, l'application qu'utilisent les clients est divisée en deux parties : la première est la solution standard qui est le cœur de l'application, nommée activepricing ou appelé la plupart du temps le produit. Les travaux autour du produit rassemblent le gros des avancées qui sont faites pour améliorer le service accordé à tous les clients de manière générale.

La seconde partie est le starter, qui comme son nom l'indique, sert à démarrer un projet par-dessus le produit. Il est la partie personnalisable par les projets et doit être mise à jour suivant les changements majeurs apportés par les nouveautés du produit. Il arrive parfois que des changements du produit apportent des inconvénients au niveau des projets pour des raisons techniques ou simplement par la spécification

des données des clients (appelée simplement 'spec' dans le milieu), mais je reviendrai dessus plus tard.

Le produit consiste globalement en un serveur codé en Java SpringBoot et d'une vue codée en React TypeScript. Le serveur est lui-même scindé en trois unités interdépendantes : ces parties sont appelées en interne " l'application, l'édition et le pricer ".

Je ne parlerai que de l'édition coté serveur et de la vue, car la partie développement logiciel de mon stage ne touche qu'à ces composantes.

## 1.4 L'équipe Produit et l'équipe Projet

Les experts des prix qui utilisent la plateforme de pricing de Mercio chez les clients sont appelés pricers. Fournir aux pricers des outils adaptés à leurs besoins est le nerf de la guerre pour Mercio, qui cherche à gagner des nouveaux clients européens tout en standardisant de plus en plus ses nouveaux contrats tout comme ses anciens. L'objectif est de réduire le plus possible le temps passé à répondre à des besoins spécifiques à travers les progrès réalisés sur le produit.

L'équipe Produit est composé de 7 personnes incluant :

- Le manager de l'équipe produit Yoann Ozanneaux et tuteur officiel pour les stages.
- Mon tuteur d'entreprise Guillaume Halb, responsable de mes travaux au sein de l'équipe Produit.
- Des développeurs qui ont de plusieurs mois à 4 ans d'ancienneté.
- Les deux stagiaires dont je fais partie.

L'équipe Projet, elle, s'occupe des solutions logicielles mises en œuvre pour les différents clients. Chaque projet ayant ses propres spécificités, ils sont attribués à un ou plusieurs membres de l'équipe en fonction de la nature et de la charge que constituent les demandes du client pour des améliorations et du support. Un membre de cette équipe doit veiller à la formation du client à l'utilisation de la plateforme de pricing (bien que la documentation soit très bien fournie), anticiper ses demandes, faire respecter les exigences sur les données que le client fournit à l'entreprise, pré-traiter certaines données, etc.

## 1.5 Environnement de travail et matériel



## 2 Sujet du stage : Alignement horizontal de produits concurrents

### 2.1 Définition du besoin

Une entreprise a naturellement l'obligation de répondre aux besoins de ses clients et de son potentiel marché si elle veut rester compétitive. Ainsi, le besoin d'un acteur de la grande distribution à ne pas vendre trop cher ou trop peu cher ses produits est réel et concerne tout aussi bien l'alimentaire, que le bricolage, ou bien même les jouets pour ne citer que les secteurs des clients de Mercio.

**Besoin concret de l'utilisateur** Les distributeurs ont pour la plupart des Marques Nationales (MN), des Marques De Distributeur (MDD), et des marques Premiers Prix. Le gros de ces ventes s'axe autour des MN et des MDD

Les MN peuvent tout à fait se retrouver dans une enseigne comme dans une autre, comme l'on trouvera aussi bien de l'eau de la marque Cristalline dans tous les magasins en France.

A partir du moment où un produit a le même identifiant produit partout en France, c'est-à-dire le même EAN, il est facile de s'ajuster à la concurrence, dans la mesure où on connaît le prix de ce même produit chez les magasins concurrents. La connaissance de ce prix peut s'acquérir facilement si un distributeur paie l'accès à une base de données telles que Nielsen-TradeDimension, afin d'obtenir quel EAN est vendu à tel prix dans telle enseigne.

Dans le cas où l'utilisateur souhaite relier des produits de sa MDD avec des produits de la MDD correspondante de son concurrent, les produits sont moins faciles à relier, car il faut parcourir le catalogue de son concurrent pour trouver le bon produit.

Effectuer ce type de lien entre produits dans l'optique d'ensuite aligner les prix des produits correspondants d'appelle le "chainage horizontal".

Bien entendu, le but derrière l'alignement à la concurrence est de faire un maximum de profit en vendant ces produits similaires ou identiques que les concurrents proposent aussi.

Si un distributeur est trop cher sur des produits clés par rapport à ses concurrents directs, les consommateurs vont se diriger chez ceux-ci pour acheter ces produits clés. D'une autre manière, si un certain produit n'a pas d'équivalent chez ses concurrents directs, alors on peut se permettre de garder les prix de vente élevés afin de garantir les marges faciles.

Pour résumer, les liens de chaînage horizontal sont donc la base d'une solution sous forme de données à ce besoin en alignement.

Un deuxième besoin rempli par le chaînage horizontal est l'indice prix. Ce dernier

permet à un distributeur de se positionner comme étant plus cher ou moins cher qu'un concurrent sur des références liées entre elles de cette façon.

Par exemple, si sur 10 références de produits, un distributeur E est moins cher de 10% en moyenne que son concurrent F sur les références correspondantes, alors l'indice prix de E sur F est de 0.9. Dans l'idée, plus l'indice prix est bas par rapport à la concurrence, plus on attire les clients qui veulent acheter leurs produits préférés le moins cher possible.

Avoir beaucoup de références liées à sa concurrence permet de renforcer la fiabilité de son indice prix, car plus il y a de liens, plus l'indice prix sera représentatif du catalogue.

**Besoin en outil** Quand le pricer utilise pour la première fois le module du "chaînage horizontal" dans la plateforme Mercio Pricing, les liens peuvent être très longs à tous écrire à la main s'ils ne sont pas déjà exportés d'un fichier rempli en interne. Il y a aussi le cas où on n'a aucun lien à exporter dans le module, et tout est à faire en partant de zéro, en rentrant les références produit une par une en s'aidant de filtres.

L'idée derrière les recommandations automatiques de lien, ou pour faire court le matching automatique, c'est que les pricers puissent rapidement appréhender ce module afin de commencer le plus tôt possible à utiliser l'alignement à la concurrence et d'établir leurs indices prix.

D'un point de vue fonctionnalité, l'utilisateur aurait des liens à valider, avec plusieurs options pour filtrer selon les catégories de produits, afin de créer de bons liens rapidement, sans avoir à taper et à chercher parmi tous les produits de la concurrence.

## 2.2 Contexte de l'application

Dans la plateforme ActivePricing, il y a de nombreuses façon d'appliquer des règles de prix sur les produits. On peut citer, parmi les règles de prix les plus utilisées, la règle d'arrondi et l'alignement à la concurrence. Cette dernière est directement concernée par mon sujet.

Afin de d'appliquer ces règles, l'application dispose de plusieurs modules, tantôt dépendants ou indépendants des autres pour fonctionner. Cependant, les données traitées restent globalement les mêmes à travers toute l'application.

**Module Éditeur de stratégie** Dans l'éditeur de stratégie, l'utilisateur précise des stratégies comme des ensembles de règles. Au sein d'une stratégie, on choisit les priorités et l'ordre entre les règles que l'on veut appliquer, et sur quels produits en

particulier.

Les prix sont calculés pour des centaines de milliers de références produit à travers plusieurs milliers de magasins pour les plus gros clients. Ces calculs se font très vite grâce à la technologie Active Pivot, un logiciel propriétaire axé performance, qui fait de l'agrégation de données et des analyses dans des hypercubes.

Les résultats des prix après application sont transparents grâce à une interface web sur laquelle on retrouve la construction du prix, son impact sur le chiffre d'affaires, etc.

**Module Chaînage horizontal** Comment prépare-t-on les renseignements nécessaires à l'utilisation des règles d'alignement à la concurrence ? L'alignement d'un produit A du client à sa concurrence signifie que pour chaque produit vendu par un concurrent pour lequel on a déclaré une correspondance, on applique un prix qui suit un coefficient qualitatif appliqué à ce triplet : Produit A, enseigne concurrente E, produit concurrent X. Concrètement, les pricers relient les produits de chez eux à des produits de leurs concurrents pour pouvoir appliquer un coefficient de prix et s'ajuster à la concurrence de manière qualitative, quantitative ou arbitrairement une combinaison des deux.

Exemple concret d'un utilisateur qui veut chaîner les produits de son catalogue aux produits du catalogue de ses concurrents :

- Si un produit à lier est vendu par pack de 12 chez le client et par pack de 6 chez son concurrent, le client choisira un lien quantitatif de coefficient égal à 2.0 pour lier ces deux références.
- Si un autre produit pèse 1kg chez le client et son correspondant se trouve uniquement par 3kg dans le catalogue de son concurrent, alors le client choisira un lien volumique de 0.333 pour lier ces deux références.

**Module de Cohérence** On peut aussi créer des liens avec coefficients entre produits d'un même catalogue :

si l'utilisateur décide par exemple que tous ses biscuits non bio doivent être moins cher de 20% que ses biscuits bio, et ce, dans tous les rayons de tous ses magasins, alors il peut créer ce type de lien dans le module de cohérence de l'application.

## 2.3 Pertinence d'un lien de chaînage horizontal

Pour pouvoir définir ce qu'est un lien de chaînage concurrentiel pertinent (ou simplement lien de matching), il faut au préalable exposer quelques exemples concrets définis à la main pour se rendre compte de la difficulté de la conception d'un tel algorithme.

TODO : exemples avec des petites images

## 2.4 Répondre à la problématique et au besoin

Mon travail consiste donc à concevoir un algorithme qui prend en entrée des produits concurrents et des produits clients pour calculer des liens cohérents en sortie pour tous les produits du catalogue des clients. Ensuite, il s'agira d'implémenter ces résultats en tant que recommandations de liens dans la plateforme Mercio Pricing, au sein du module de chaînage horizontal. Ainsi, l'utilisateur pourra utiliser le plus tôt possible la règle d'alignement à la concurrence dans son parcours.

Le client aura accès à un indice prix aussi d'autant plus fiable qu'il aura créé beaucoup liens par le moyen de ces recommandations.

## 2.5 Données initiales à disposition

## 3 Démarche proposée

Concevoir une démarche est l'étape primordiale de tout sujet de recherche et développement. Nous cherchons à savoir quels sont les connaissances à acquérir afin d'appliquer une solution adaptée, quels sont les solutions similaires proposées dans la littérature, quels moyens possède-t-on pour déployer pragmatiquement la solution mise en œuvre.

INSERER IMAGE EXEMPLE MATCHING 1

Ici, un pricer décide de lier un produit A de son enseigne au produit X de son concurrent.

Le produit A et le produit X possèdent plusieurs caractéristiques comme un nom, des catégories, un volume, un packaging, une marque, un prix, et bien d'autres dans la réalité.

Nous sommes ici dans le cas facile où le produit n'est pas exactement identique à celui de son concurrent, mais le produit semble concorder aux attentes de l'alignement concurrentiel : il s'agit de la marque de distributeur de l'enseigne respectivement dans les deux cas. Les noms semblent sémantiquement désigner le même produit, l'ordre de grandeur du volume vendu en rayon est proche, les catégories concordent assez bien, etc. Nous avons tout pour dire qu'il est cohérent d'aligner A sur X pour ce concurrent-ci.

INSERER IMAGE EXEMPLE MATCHING 2

### 3.1 Agilité

### 3.2 Planification des travaux

## 4 Travail réalisé

### 4.1 Analyse de la tâche

### 4.2 Solutions possibles et leurs justifications

### 4.3 Résultats expérimentaux, tests

## 5 Bilan de production

### 5.1 Objectifs atteints

Au moment de l'écriture du rapport, il reste 8 semaines de stage

**Avancée relative** Ma principale phase d'apprentissage avec les technologies à manipuler est terminée, ce qui veut dire que j'avance beaucoup plus vite sur tout ce que je fais à partir de maintenant. Ainsi, je peux réitérer sur tous les points précédemment abordés avec plus d'expertise.

**Travaux nécessaires restants**

### 5.2 Apport du travail réalisé

### 5.3 Perspectives d'évolution des travaux

### 5.4 Bilan général de la démarche

**Point de vue méthodologie scientifique**

**Une démarche axée production**

## 6 Bilan sur le plan personnel

### 6.1 Compétences mises en avant

Compétences utilisées

Compétences acquises

### 6.2 Projet professionnel



## 7 Conclusion

Si c'était à refaire ...

Pour finir

## 8 Glossaire

### Acronymes

MN :

MDD :

### Terminologie

distributeur :

## 9 Annexes