**HYPERLOOP FEATURE ENHANCEMENT PROPOSAL**

**for Machine Learning model generated features**

|  |  |  |  |
| --- | --- | --- | --- |
| **History** | | | |
| **NR** | **Data** | **Autor** | **Ver** |
| **1** | **01.02.2017** | **A.I. DAMIAN** | **Draft 1** |
| **2** | **02.02.2017** | **A.I. DAMIAN** | **Draft 2** |

Contents

[1 Features derived from Machine Learning models for client brand preference 1](#_Toc473794652)

[2 Features derived from Machine Learning models for client product categories preference 2](#_Toc473794653)

[3 Features derived from Machine Learning models for client price range (level) preference 5](#_Toc473794654)

# Features derived from Machine Learning models for client brand preference

Obs.: all 17 proposed features have descriptive names instead of short cryptic codes due to the need for high self-explain ability)

|  |  |  |
| --- | --- | --- |
| Name | Description | Notes |
| BP01\_APIVITA | Client brand preference for APIVITA | Real value score [0.00-1.00] |
| BP02\_OCCITANE | Client brand preference for L’OCCITANE | Real value score [0.00-1.00] |
| BP03\_VICHY | Client brand preference for VICHY | Real value score [0.00-1.00] |
| BP04\_DR\_HART | Client brand preference for DR HART | Real value score [0.00-1.00] |
| BP05\_NUROFEN | Client brand preference for NUROFEN | Real value score [0.00-1.00] |
| BP06\_AVENE | Client brand preference for AVENE | Real value score [0.00-1.00] |
| BP06\_BIODERMA | Client brand preference for BIODERMA | Real value score [0.00-1.00] |
| BP07\_LA\_ROCHE | Client brand preference for LA ROCHE POSAY | Real value score [0.00-1.00] |
| BP08\_ERBOLARIO | Client brand preference for L’ ERBOLARIO | Real value score [0.00-1.00] |
| BP09\_PARASINUS | Client brand preference for PARASINUS | Real value score [0.00-1.00] |
| BP10\_NESTLE | Client brand preference for NESTLE | Real value score [0.00-1.00] |
| BP11\_OXYANCE | Client brand preference for OXYANCE | Real value score [0.00-1.00] |
| BP12\_DUCRAY | Client brand preference for DUCRAY | Real value score [0.00-1.00] |
| BP13\_SCHOLL | Client brand preference for SCHOLL | Real value score [0.00-1.00] |
| BP14\_ASPENTER | Client brand preference for ASPENTER | Real value score [0.00-1.00] |
| BP15\_TRUSSA | Client brand preference for TRUSSA | Real value score [0.00-1.00] |
| BP16\_SREPSILS | Client brand preference for SREPSILS | Real value score [0.00-1.00] |
| BP17\_MELVITA | Client brand preference for MELVITA | Real value score [0.00-1.00] |

Client brand preferences features (BP\_<BRAND>) define the actual inclination (disposition) that a particular client has towards buying a certain brand. The real value score is in range 0.00 up to 1.00 where a value of 0 means that the client has never bought that particular brand, a value of 1.00 means a potential client that buys only that particular brand and a value around 0.5 means the client buys about half of his/hers products from that particular brand.

The brand preference features can be used directly or in a model (ranging from simple statistical models to complex machine learning models) for the following purposes:

1. Customer behavior segmentation
2. Market-Basket based segmentation or individual market basket inference
3. Products/Buy recommendation
4. New product campaigns (for certain brands) or other kind of campaigns

# Features derived from Machine Learning models for client product categories preference

(obs.: all 60 features have descriptive names instead of cryptic codes due to the need for high self-explain ability)

|  |  |  |
| --- | --- | --- |
| Name | Description | Notes |
| CP01\_BABY | Client preference for products tagged as BABY products (category) | Real value score [0.00-1.00] |
| CP02\_COSMETICE | Client preference for products tagged as COSMETICE products (category) | Real value score [0.00-1.00] |
| CP03\_DERMOCOSM | Client preference for products tagged as DERMOCOSMETICE products (category) | Real value score [0.00-1.00] |
| CP04\_PHARMA | Client preference for products tagged as PHARMA products (category) | Real value score [0.00-1.00] |
| CP05\_SPECIALE | Client preference for products tagged as SPECIALE products (category) | Real value score [0.00-1.00] |
| CP06\_OTC | Client preference for products tagged as OTC products (category) | Real value score [0.00-1.00] |
| CP07\_MEDICAT\_ALTERN | Client preference for products tagged as MEDICATIE\_ALTERNATIVA products (category/ subcategory) | Real value score [0.00-1.00] |
| CP08\_BARBATI | Client preference for products tagged as BARBATI products (category/ subcategory) | Real value score [0.00-1.00] |
| CP09\_CADOURI | Client preference for products tagged as CADOURI products (category/ subcategory) | Real value score [0.00-1.00] |
| CP10\_COPII | Client preference for products tagged as COPII products (category/ subcategory) | Real value score [0.00-1.00] |
| CP11\_FRUM\_PRENAT | Client preference for products tagged as FRUMUSETE\_PRE\_NATAL products (category/ subcategory) | Real value score [0.00-1.00] |
| CP12\_FRUMUSETE | Client preference for products tagged as FRUMUSETE products (category/ subcategory) | Real value score [0.00-1.00] |
| CP13\_HOMEOPATIE | Client preference for products tagged as HOMEOPATIE products (category/ subcategory) | Real value score [0.00-1.00] |
| CP14\_PROMO | Client preference for products tagged as PROMO products (category/ subcategory) | Real value score [0.00-1.00] |
| CP15\_R\_APARAT\_RESP | Client preference for products tagged as R\_APARAT\_RESPIRATOR products (category/ subcategory) | Real value score [0.00-1.00] |
| CP16\_S\_ORGANE\_SENZ | Client preference for products tagged as S\_ORGANE\_SENZITIVE products (category/ subcategory) | Real value score [0.00-1.00] |
| CP17\_SANATATE | Client preference for products tagged as SANATATE products (category/ subcategory) | Real value score [0.00-1.00] |
| CP18\_STYLIST | Client preference for products tagged as STYLIST products (category/ subcategory) | Real value score [0.00-1.00] |
| CP19\_INGRIJIRE\_CORP | Client preference for products tagged as INGRIJIRE\_CORP products (category/ subcategory) | Real value score [0.00-1.00] |
| CP20\_INGRIJIRE\_PERS | Client preference for products tagged as INGRIJIRE\_PERSONALA products (category/ subcategory) | Real value score [0.00-1.00] |
| CP21\_SANGE\_ORG\_HP | Client preference for products tagged as B\_SANGE\_SI\_ORGANE\_HEMATOPOETICE products (category/ subcategory) | Real value score [0.00-1.00] |
| CP22\_PENTRU\_EA | Client preference for products tagged as PENTRU\_EA products (category/ subcategory) | Real value score [0.00-1.00] |
| CP23\_SAMPON\_USCAT | Client preference for products tagged as SAMPON\_USCAT products (category/ subcategory) | Real value score [0.00-1.00] |
| CP24\_SIS\_MUSC\_SCHEL | Client preference for products tagged as M\_SISTEMUL\_MUSCULO\_SCHELETIC products (category/ subcategory) | Real value score [0.00-1.00] |
| CP25\_INGRIJIRE\_INTIMA | Client preference for products tagged as INGRIJIRE\_INTIMA products (category/ subcategory) | Real value score [0.00-1.00] |
| CP26\_INGRIJJIRE\_FATA | Client preference for products tagged as INGRIJJIRE\_FATA products (category/ subcategory) | Real value score [0.00-1.00] |
| CP27\_CADOURI\_MIXTE | Client preference for products tagged as CADOURI\_MIXTE products (category/ subcategory) | Real value score [0.00-1.00] |
| CP28\_N06\_PSIHOAN | Client preference for products tagged as N06\_PSIHOANALEPTICE products (category/ subcategory) | Real value score [0.00-1.00] |
| CP29\_ANTIHIST\_UZ\_SIST | Client preference for products tagged as R06\_ANTIHISTAMINICE\_DE\_UZ\_SISTEMIC products (category/ subcategory) | Real value score [0.00-1.00] |
| CP30\_INCALT\_ACCESOR | Client preference for products tagged as INCALTAMINTE\_SI\_ACCESORII products (category/ subcategory) | Real value score [0.00-1.00] |
| CP31\_PROT\_SEX | Client preference for products tagged as PROTECTIE\_SEXUALA products (category/ subcategory) | Real value score [0.00-1.00] |
| CP32\_TRACT\_DIG\_MET | Client preference for products tagged as A\_TRACTUL\_DIGESTIV\_SI\_METABOLISM products (category/ subcategory) | Real value score [0.00-1.00] |
| CP33\_BARBIERIT | Client preference for products tagged as BARBIERIT products (category/ subcategory) | Real value score [0.00-1.00] |
| CP34\_CARDIOVASC | Client preference for products tagged as C\_SISTEM\_CARDIOVASCULAR products (category/ subcategory) | Real value score [0.00-1.00] |
| CP35\_BAIE | Client preference for products tagged as BAIE products (category/ subcategory) | Real value score [0.00-1.00] |
| CP36\_PARFUM | Client preference for products tagged as PARFUMERIE products (category/ subcategory) | Real value score [0.00-1.00] |
| CP37\_INGRIJ\_PICIO | Client preference for products tagged as INGRIJIRE\_PICIOARE products (category/ subcategory) | Real value score [0.00-1.00] |
| CP38\_INGRIJIRE\_SOL | Client preference for products tagged as INGRIJIRE\_SOLARA products (category/ subcategory) | Real value score [0.00-1.00] |
| CP39\_DISP\_MEDICAL | Client preference for products tagged as DISPOZITIVE\_MEDICALE products (category/ subcategory) | Real value score [0.00-1.00] |
| CP40\_JUCARII | Client preference for products tagged as JUCARII products (category/ subcategory) | Real value score [0.00-1.00] |
| CP41\_PENTRU\_COPII | Client preference for products tagged as PENTRU\_COPII products (category/ subcategory) | Real value score [0.00-1.00] |
| CP42\_INGRIJIRE\_PAR | Client preference for products tagged as INGRIJIREA\_PARULUI products (category/ subcategory) | Real value score [0.00-1.00] |
| CP43\_OCHELARI | Client preference for products tagged as OCHELARI products (category/ subcategory) | Real value score [0.00-1.00] |
| CP44\_SIST\_NERVOS | Client preference for products tagged as N\_SISTEMUL\_NERVOS products (category/ subcategory) | Real value score [0.00-1.00] |
| CP45\_MAKE\_UP | Client preference for products tagged as MAKE\_UP products (category/ subcategory) | Real value score [0.00-1.00] |
| CP46\_ANALGEZICE | Client preference for products tagged as N02\_ANALGEZICE products (category/ subcategory) | Real value score [0.00-1.00] |
| CP47\_CASA | Client preference for products tagged as ACCESORII\_CASA products (category/ subcategory) | Real value score [0.00-1.00] |
| CP48\_JUCARII | Client preference for products tagged as JUCARII products (category/ subcategory) | Real value score [0.00-1.00] |
| CP49\_ANTIINFECTI1 | Client preference for products tagged as J\_ANTIINFECTIOASE\_DE\_UZ\_SISTEMIC products (category/ subcategory) | Real value score [0.00-1.00] |
| CP50\_INGR\_DENTARA | Client preference for products tagged as INGRIJIRE\_DENTARA products (category/ subcategory) | Real value score [0.00-1.00] |
| CP51\_DIETETICE1 | Client preference for products tagged as V06\_DIETETICE products (category/ subcategory) | Real value score [0.00-1.00] |
| CP52\_NETERAPEUTICE | Client preference for products tagged as V07\_NETERAPEUTICE products (category/ subcategory) | Real value score [0.00-1.00] |
| CP53\_OFTALMOLOGIC | Client preference for products tagged as S01\_PRODUSE\_OFTALMOLOGICE products (category/ subcategory) | Real value score [0.00-1.00] |
| CP54\_ANESTEZICE | Client preference for products tagged as N01\_ANESTEZICE products (category/ subcategory) | Real value score [0.00-1.00] |
| CP55\_PREP\_NAZALE | Client preference for products tagged as R01\_PREP\_NAZALE products (category/ subcategory) | Real value score [0.00-1.00] |
| CP56\_TOALETA\_ZILNIC | Client preference for products tagged as TOALETA\_ZILNICA products (category/ subcategory) | Real value score [0.00-1.00] |
| CP57\_HRANA | Client preference for products tagged as HRANA products (category/ subcategory) | Real value score [0.00-1.00] |
| CP58\_PSIHOLEPTICE | Client preference for products tagged as N05\_PSIHOLEPTICE products (category/ subcategory) | Real value score [0.00-1.00] |
| CP59\_TINCT\_HOMEO | Client preference for products tagged as XRNT\_TINCTURI\_HOMEOPATE products (category/ subcategory) | Real value score [0.00-1.00] |
| CP60\_ANESTEZICE | Client preference for products tagged as ANESTEZICE products (category/ subcategory) | Real value score [0.00-1.00] |

Client category preferences features (CP\_<CATEG>) define the actual inclination (disposition) that a particular client has towards buying a certain category of products (products tagged/labeled as of a certain category type) – observation: a product has more than one “tag” meaning it has category and one or more subcategories. The real value score is in range 0.00 up to 1.00 where a value of 0 means that the client has never bought that particular category of products, a value of 1.00 means a potential client that buys only that particular category of products and a value around 0.5 means the client buys about half of his/hers products from that particular category of products.

The product segment preference features can be used directly or in a model (ranging from simple statistical models to complex machine learning models) for the following purposes:

1. Customer behavior segmentation including margin-revenue segmentation
2. Market-Basket based segmentation or individual market basket inference
3. Products/Buy recommendation
4. Churn prediction
5. Direct use in new product campaigns (for particular product categories), narrowly-targeted campaigns or other kind of campaigns

# Features derived from Machine Learning models for client price range (level) preference

(obs.: all 15 proposed features have descriptive names instead of short cryptic codes due to the need for high self-explain ability)

|  |  |  |
| --- | --- | --- |
| Name | Description | Notes |
| PR01\_BABY | Price range score for BABY product category | Real value score [0.00-1.00] |
| PR02\_COSMETICE | Price range score for COSMETICE product category | Real value score [0.00-1.00] |
| PR03\_DERMOCOSM | Price range score for DERMOCOSMETICE product category | Real value score [0.00-1.00] |
| PR04\_PHARMA | Price range score for PHARMA product category | Real value score [0.00-1.00] |
| PR05\_SPECIALE | Price range score for SPECIALE product category | Real value score [0.00-1.00] |
| PR06\_OTC | Price range score for OTC product category | Real value score [0.00-1.00] |
| PR07\_MEDICAT\_ALTERN | Price range score for MEDICATIE\_ALTERNATIVA product category | Real value score [0.00-1.00] |
| PR08\_BARBATI | Price range score for BARBATI product category | Real value score [0.00-1.00] |
| PR09\_FEMEI | Price range score for FEMEI product category | Real value score [0.00-1.00] |
| PR11\_COPII | Price range score for COPII product category | Real value score [0.00-1.00] |
| PR12\_FRUMUSETE | Price range score for FRUMUSETE product category | Real value score [0.00-1.00] |
| PR13\_HOMEOPATIE | Price range score for HOMEOPATIE product category | Real value score [0.00-1.00] |
| PR14\_SANATATE | Price range score for SANATATE product category | Real value score [0.00-1.00] |
| PR15\_ ANALGEZICE | Price range score for N02\_ANALGEZICE product category | Real value score [0.00-1.00] |

The client price range features (PR\_<CATEG>) define the actual disposition of a client to buy cheap or expensive products from a certain category of products. For example a value of 0.1 for PR09\_FEMEI would mean that particular client buys very cheap women-related products while a value of 0.9of PR09\_FEMEI would mean that particular client buys some of the most expensive products in that category (women-related products). This particular feature is applicable in various features but has great potential as a predictor in the churn prediction model. The range of models that can use the client price range features contains the following:

1. Behavior bases customer segmentation
2. Client lifecycle analysis and inference
3. Product/buy recommendation
4. Churn prediction
5. Value related segmentation models
6. Targeted campaigns