

Projet 7

Implémentez un modèle de scoring

Khalil Henchi

Soutenance, le 09 février 2022

Encadrement : M. Amine Hadj-Youcef

Jury d'évaluation : M. Souhail Toumdi

Plan

1. **Rappel de la problématique**
2. **Nettoyage de données**
3. **Analyse exploratoire de données**
4. **Modélisation : Machine Learning**
5. **Déploiement : API et Dashboard interactif**
6. **Conclusion et Perspectives**

Présentation de la problématique

Contexte

- **Client**
Prêt à dépenser : Crédits à la consommation
- **Analyse de l'existant**
 - Personnes avec peu ou pas d'historique de prêt

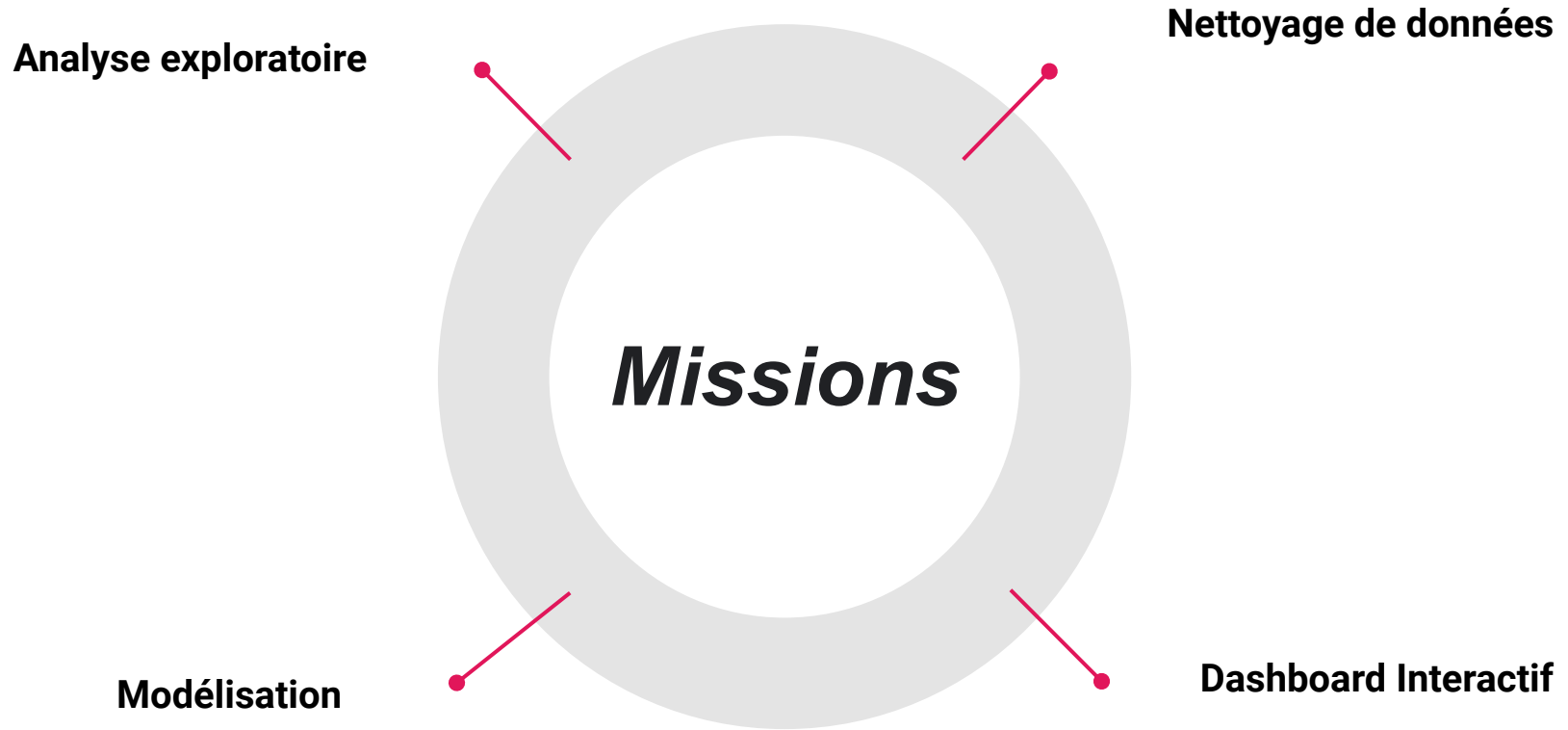


Problématique

- **Attribution manuelle des crédits: peu Fiable et manque de transparence**
- **Est-il possible de réaliser une classification à l'aide des données disponible?**

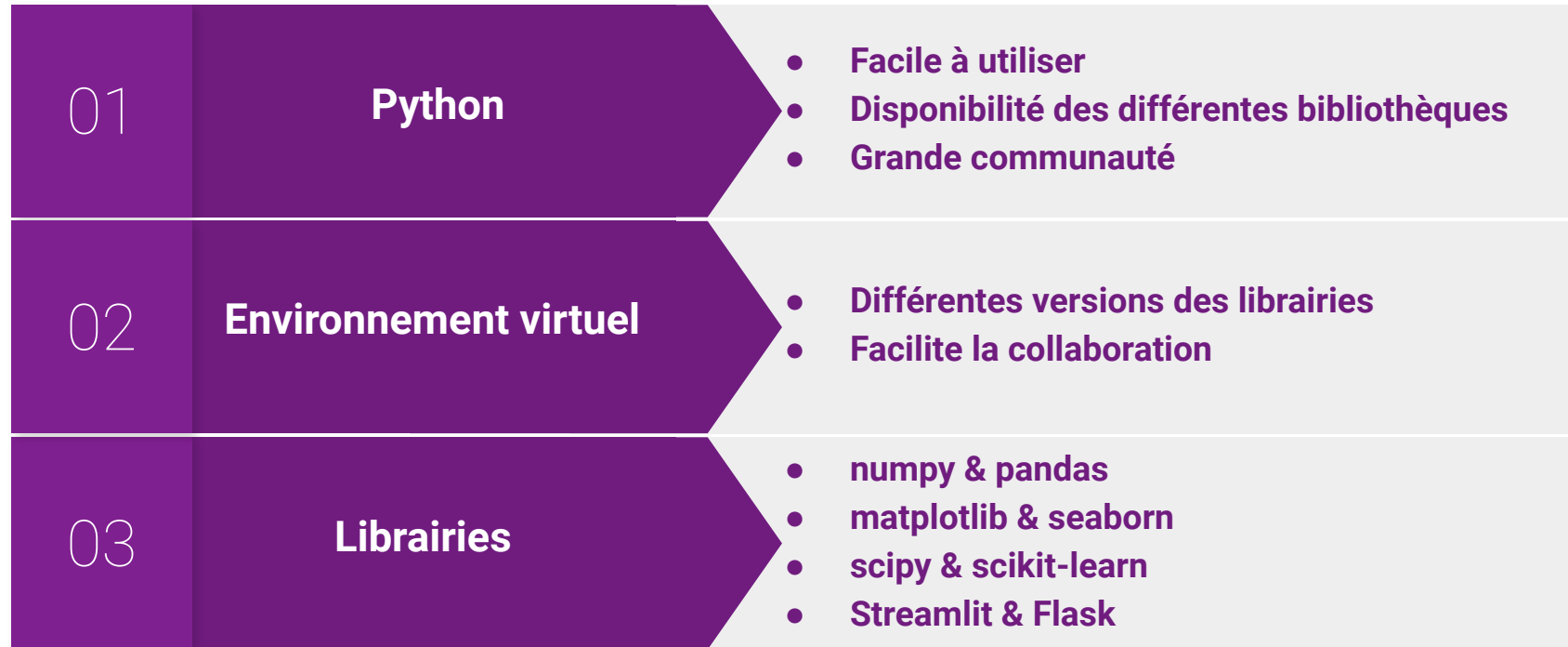
Objectifs

- **Fiabiliser le processus : Développer modèle de scoring : probabilité défaut de paiement**
- **Réaliser un dashboard pour faciliter l'utilisation de l'outil développé**



Travail à réaliser

- Préparer l'environnement du travail
- Nettoyer les données
- Valider la qualité des données
- Décrire et sélectionner les informations pertinentes
- Analyser le jeu
- Extraire des features
- Réduction de dimension
- Développer modèle de scoring
- Développer un dashboard interactif



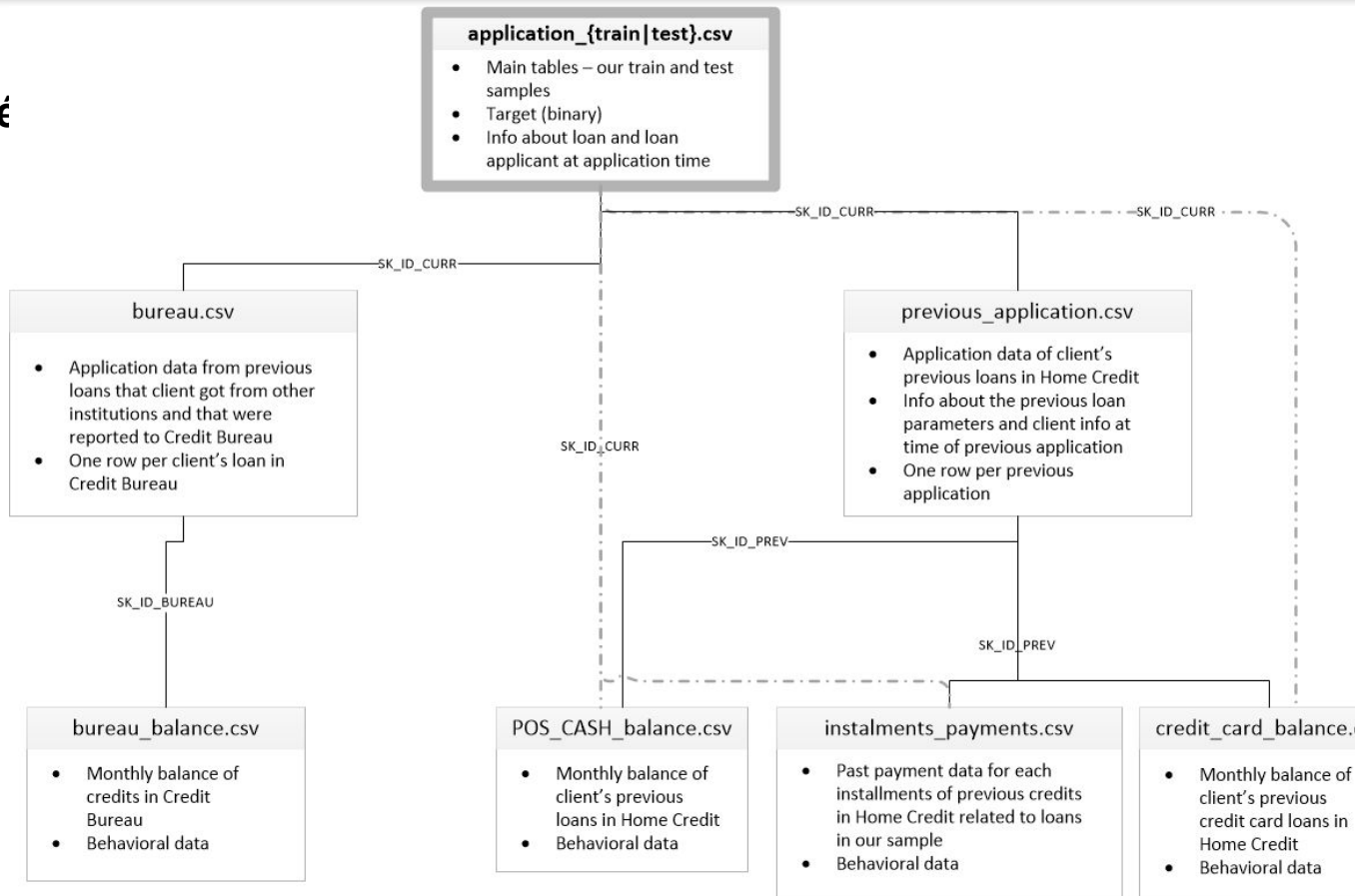
Nettoyage et Analyse Exploratoire

01	Exploration de jeu de données et Identification des variables	<ul style="list-style-type: none">• Quelles informations ?• Quantités de données• Quantités de données manquantes• Informations pertinentes
02	Contrôle des variables	<ul style="list-style-type: none">• Vérifier la pertinence de chaque variable quant aux besoins métiers
03	Suppression des variables non pertinentes	<ul style="list-style-type: none">• Supprimer les variables non pertinentes en se basant sur le contrôle de l'étape 2
03	Suppression des doublons	<ul style="list-style-type: none">• Supprimer les observations dupliquées en se basant sur une variable unique
04	Nettoyage métier	<ul style="list-style-type: none">• Contrôler les variables et ses valeurs en se basant sur des connaissances métier
06	Traitement des valeurs aberrantes	<ul style="list-style-type: none">• Supprimer• Imputation• Traiter séparément
07	Traitement des valeurs manquantes	<ul style="list-style-type: none">• Imputation moyenne / mode / médiane• Supprimer• Prédiction

Présentation de donn  

9 fichiers

- 307 511 clients
- 122 Indicateurs
- informations g  n  rales
- informations sur les pr  ts pr  c  dents
- capacit   ou non de payer cr  dit (solvabilit  ) : Cible



Autres opérations de nettoyage

- Kaggle kernels
- Feature engineering : pourcentage du montant des crédits sur la totalité des revenus, la durée totale du paiement des crédits, ...
- Gestion des données manquantes : suppression lignes et colonnes
- Standardisation, encodage des variables catégorielles

Analyse exploratoire

01	Exploration de jeu de données et Identification des variables	<ul style="list-style-type: none">• Quelles informations ?• Quantités de données• Quantités de données manquantes• Informations pertinentes
02	Analyse univariée	<ul style="list-style-type: none">• Travailler variable par variable• statistiques classiques : moyenne, variance, médiane, etc.• histogramme, boxplot
03	Analyse multivariée	<ul style="list-style-type: none">• Analyse croisée• Diagramme de dispersion• Matrice de corrélation
04	Traitement des valeurs manquantes	<ul style="list-style-type: none">• Imputation moyenne / mode / médiane• Supprimer• Prédiction
05	Traitement des valeurs aberrantes (outliers)	<ul style="list-style-type: none">• Supprimer• Imputation• Traiter séparément
06	Transformation variable	<ul style="list-style-type: none">• Logarithme• Racine carrée / cubique
07	Création de variables	<ul style="list-style-type: none">• Création de nouveaux variables• Combiner

Analyses univariées

Première étape d'une analyse exploratoire

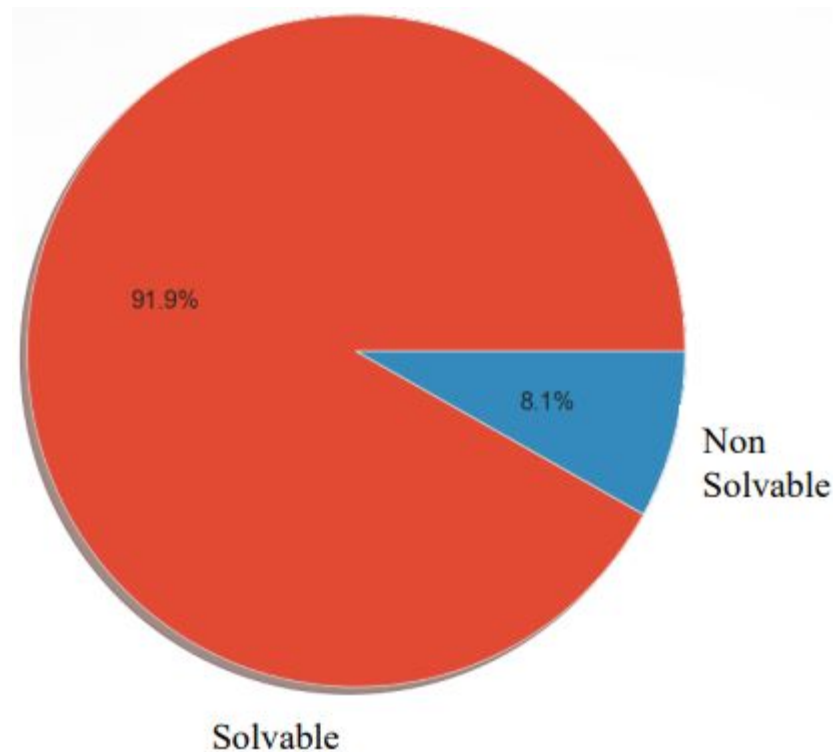
- travailler variable par variable
- numériquement et graphiquement

Variable numérique

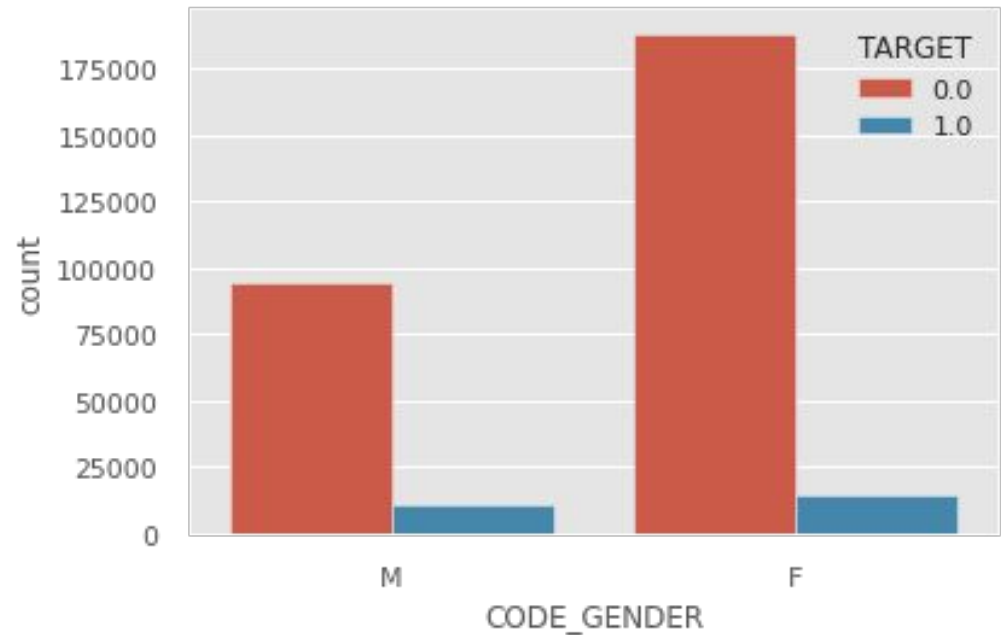
- à valeurs dans R
- statistiques classiques : moyenne, variance, médiane, etc.
- représentations associées :
 - Histogramme
 - Boxplot
 - Diagramme circulaire (Pie chart)

Déséquilibre de la Target

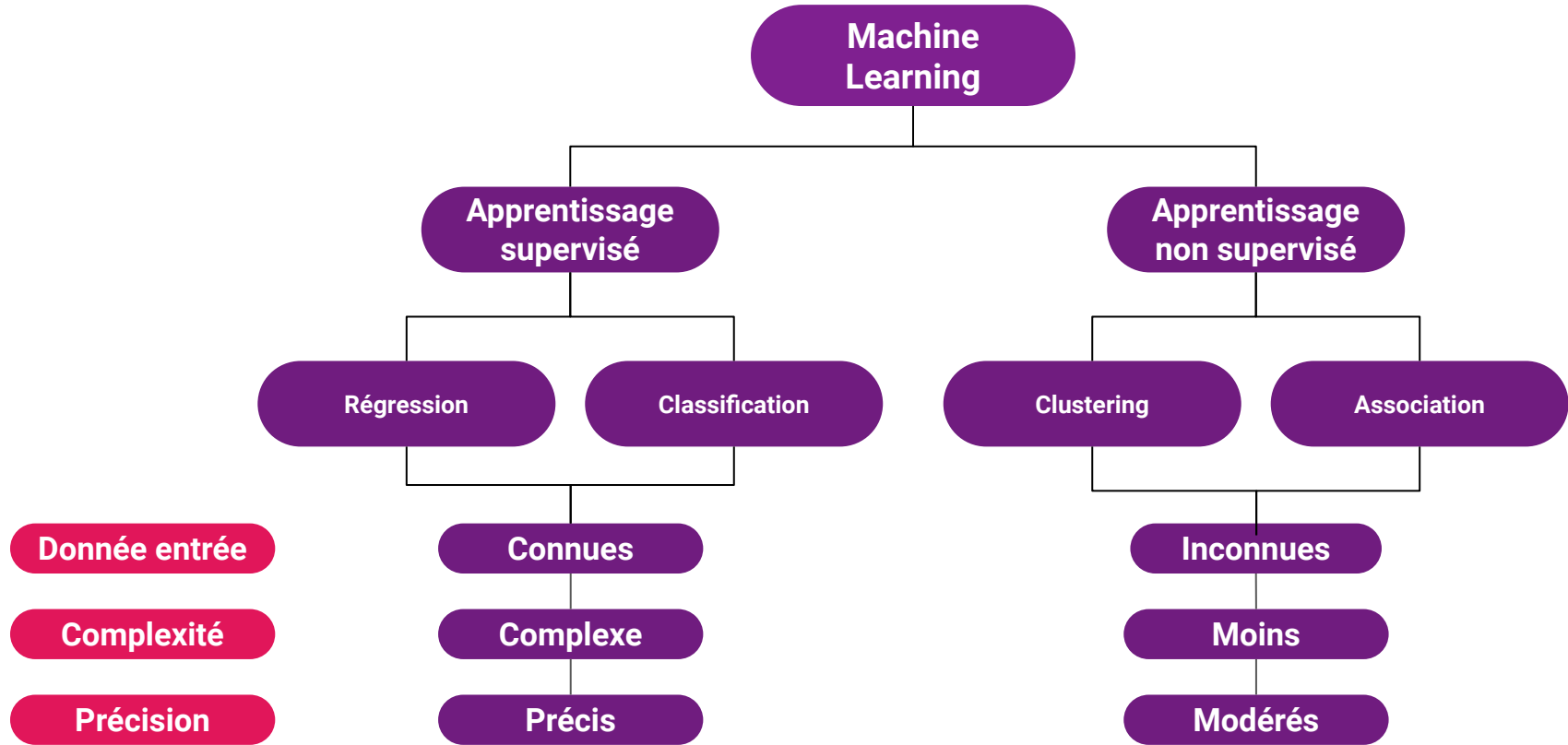
Fort déséquilibre entre les personnes solvables et les personnes non solvables.



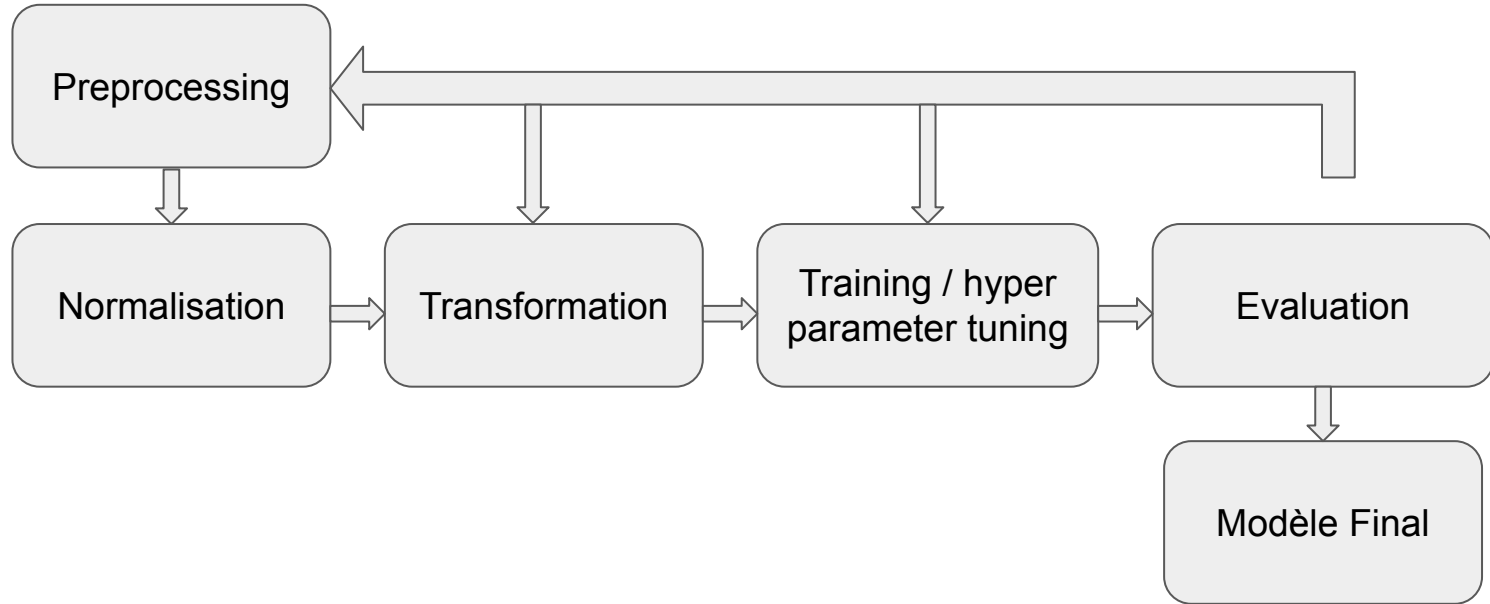
**Genre masculin a tendance à
être plus défaillants**



Modélisation : entraînement de modèles



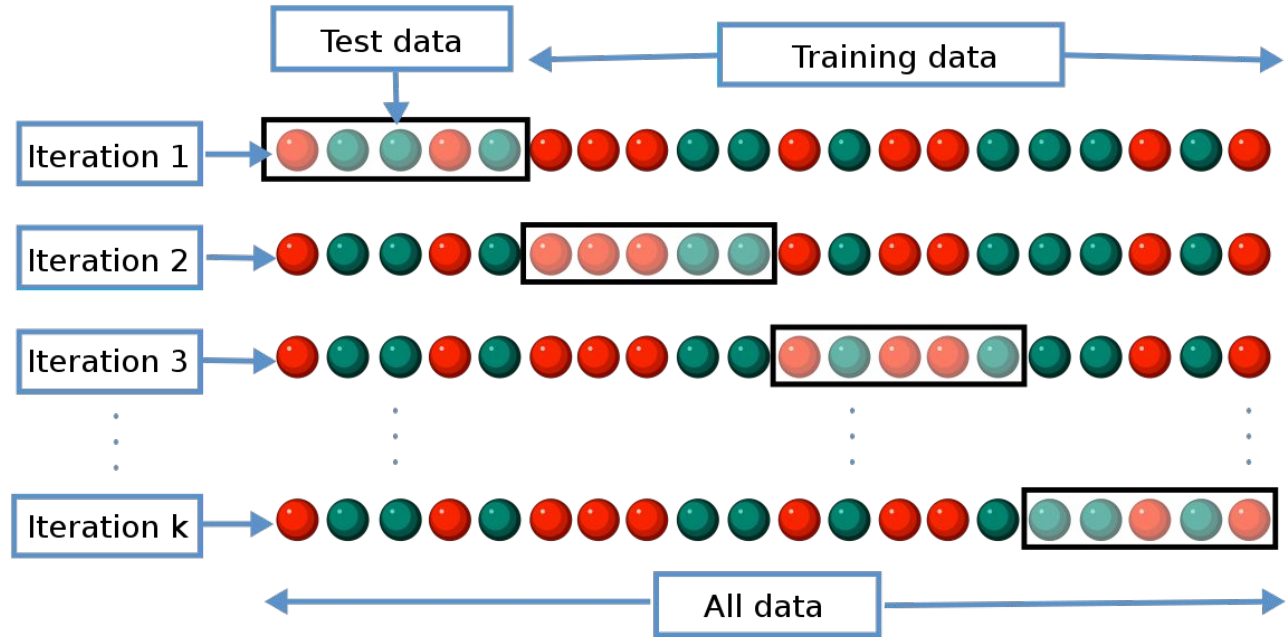
Data project Work Flow



Préparation des données

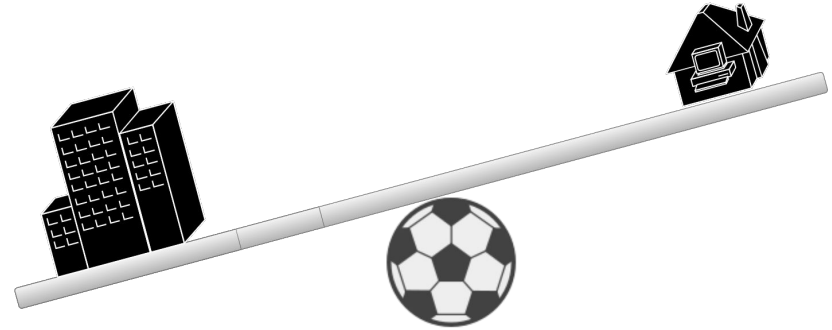
Division du jeu de données

- Train (80%) :
Validation croisée
- Test (20%)



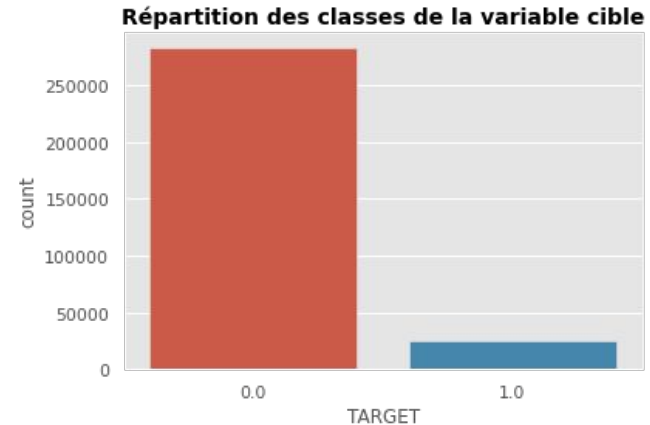
Déséquilibre de la cible

- Impact performance
- Prédiction incorrecte



Solutions

- Class weights
- Over-sampling
- Under-sampling
- SMOTE



Fonction coût à optimiser

Classe 0 = Solvabilité

Classe 1 = Difficulté de paiement

TP : Accorder / Rembourser = gain

TN : Non accorder / Non rembourser = neutre

FN : Non accorder / Rembourser = perte

FP : Accorder / Non rembourser = perte

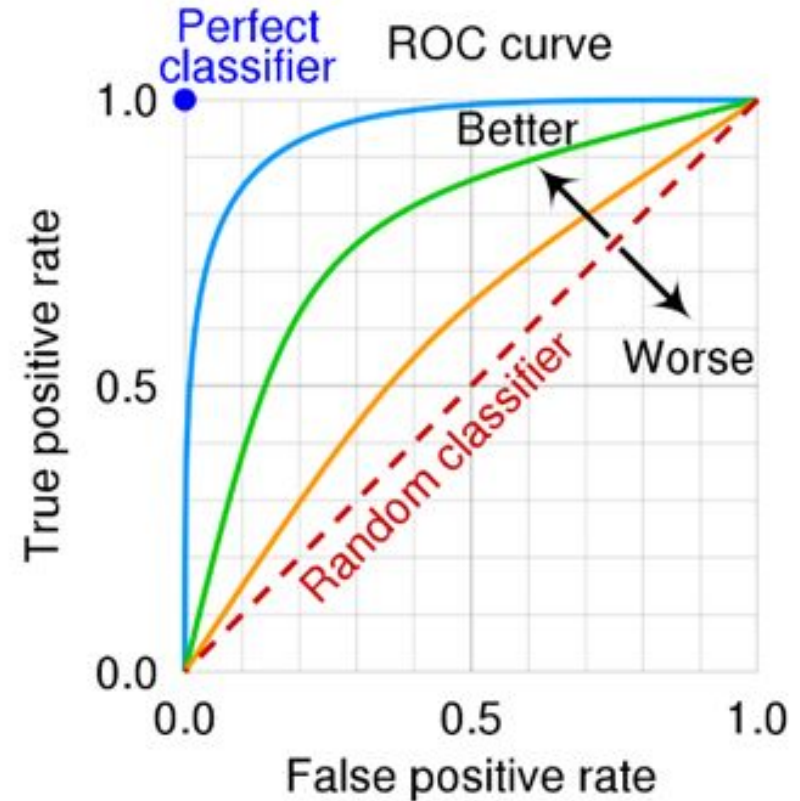
			Prédiction	
			Positive	Negative
			0	1
Réelle	Positive	0	True positive (TP)	False negative (FN)
	Negative	1	False positive (FP)	True negative (TN)

Evaluation des modèles

AUC = Aire sous la courbe ROC :

→ capacité d'un classificateur
à distinguer les classes

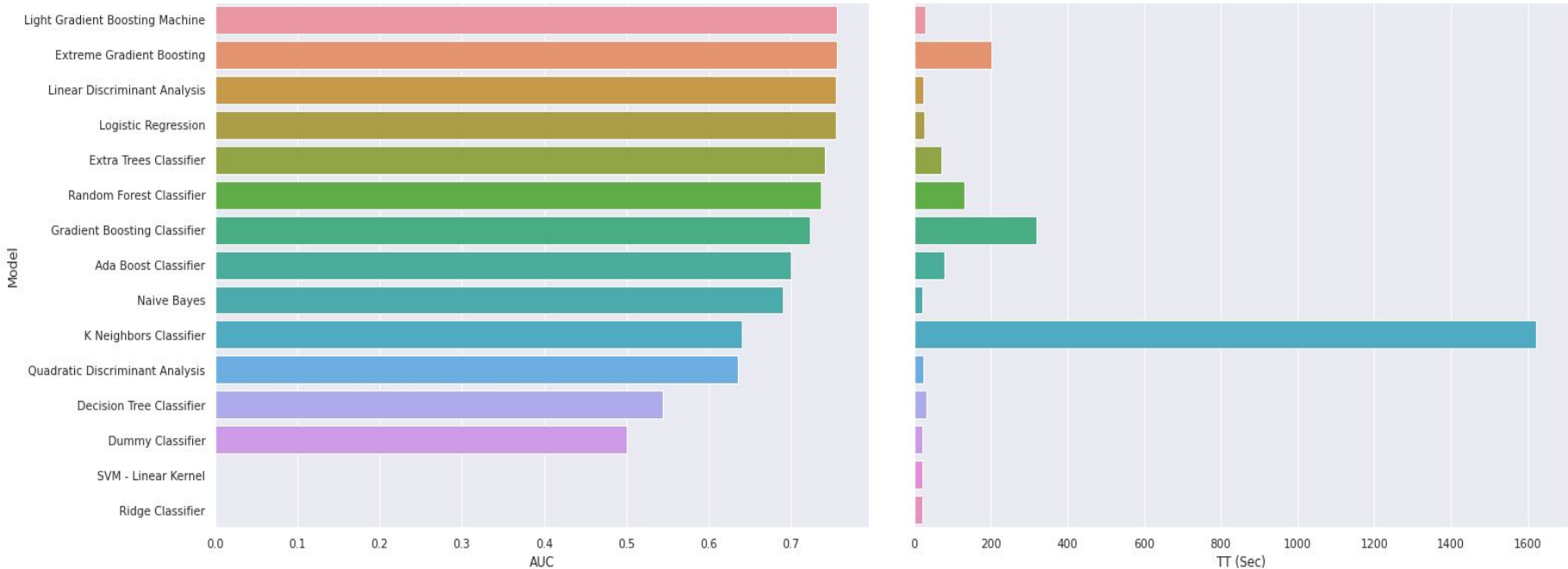
$$F_\beta = (1 + \beta^2) \frac{\text{Precision} \cdot \text{Recall}}{\beta^2 \cdot \text{Precision} + \text{Recall}}$$



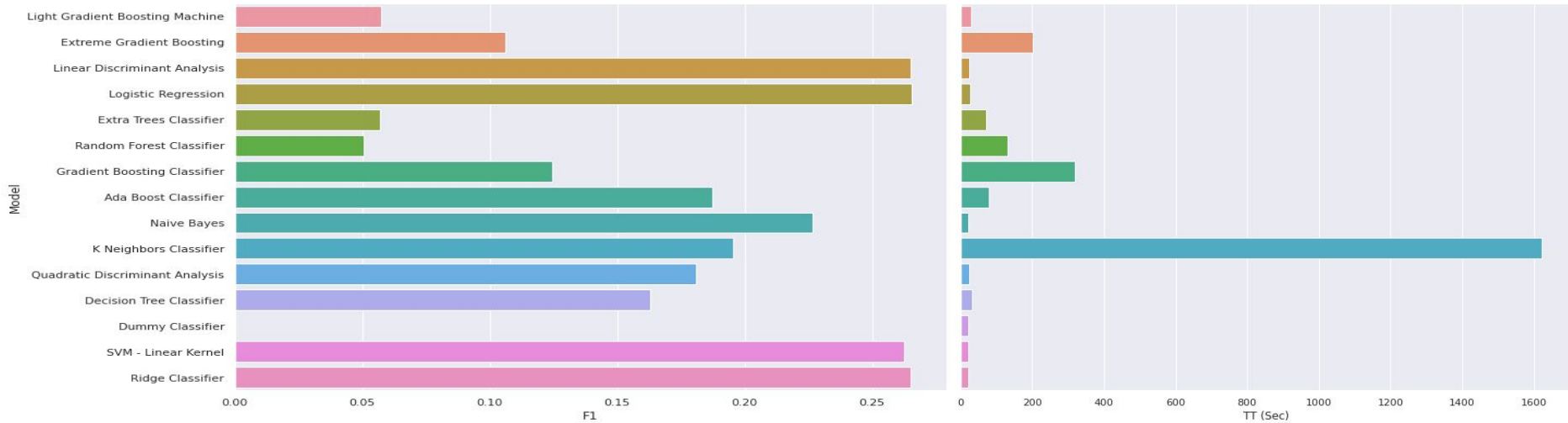
Choix des modèles

	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC	F10Score	TT (Sec)
lightgbm	Light Gradient Boosting Machine	0.9181	0.7562	0.0306	0.4418	0.0573	0.0473	0.0992	0.0309	29.4550
xgboost	Extreme Gradient Boosting	0.9175	0.7560	0.0604	0.4426	0.1063	0.0885	0.1400	0.0610	203.0390
lda	Linear Discriminant Analysis	0.6893	0.7549	0.6900	0.1641	0.2651	0.1541	0.2184	0.6688	24.6900
lr	Logistic Regression	0.6938	0.7544	0.6823	0.1650	0.2657	0.1552	0.2181	0.6617	25.8910
et	Extra Trees Classifier	0.9175	0.7411	0.0307	0.3988	0.0570	0.0459	0.0920	0.0310	69.9840
rf	Random Forest Classifier	0.9175	0.7361	0.0271	0.3869	0.0506	0.0403	0.0844	0.0273	130.2630
gbc	Gradient Boosting Classifier	0.9105	0.7222	0.0783	0.3028	0.1243	0.0941	0.1187	0.0788	319.6870
ada	Ada Boost Classifier	0.8850	0.6999	0.1631	0.2197	0.1871	0.1267	0.1284	0.1635	78.5190
nb	Naive Bayes	0.6545	0.6901	0.6225	0.1384	0.2264	0.1078	0.1587	0.6016	21.6530
knn	K Neighbors Classifier	0.5438	0.6401	0.6830	0.1141	0.1956	0.0655	0.1172	0.6509	1623.1170
qda	Quadratic Discriminant Analysis	0.8379	0.6354	0.2342	0.1590	0.1811	0.1044	0.1082	0.2324	23.7600
dt	Decision Tree Classifier	0.8381	0.5445	0.1939	0.1404	0.1629	0.0758	0.0770	0.1932	33.1900
dummy	Dummy Classifier	0.9188	0.5000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	21.7910
svm	SVM - Linear Kernel	0.6855	0.0000	0.6884	0.1622	0.2625	0.1509	0.2149	0.6670	22.9040
ridge	Ridge Classifier	0.6893	0.0000	0.6901	0.1641	0.2651	0.1541	0.2185	0.6689	21.6560

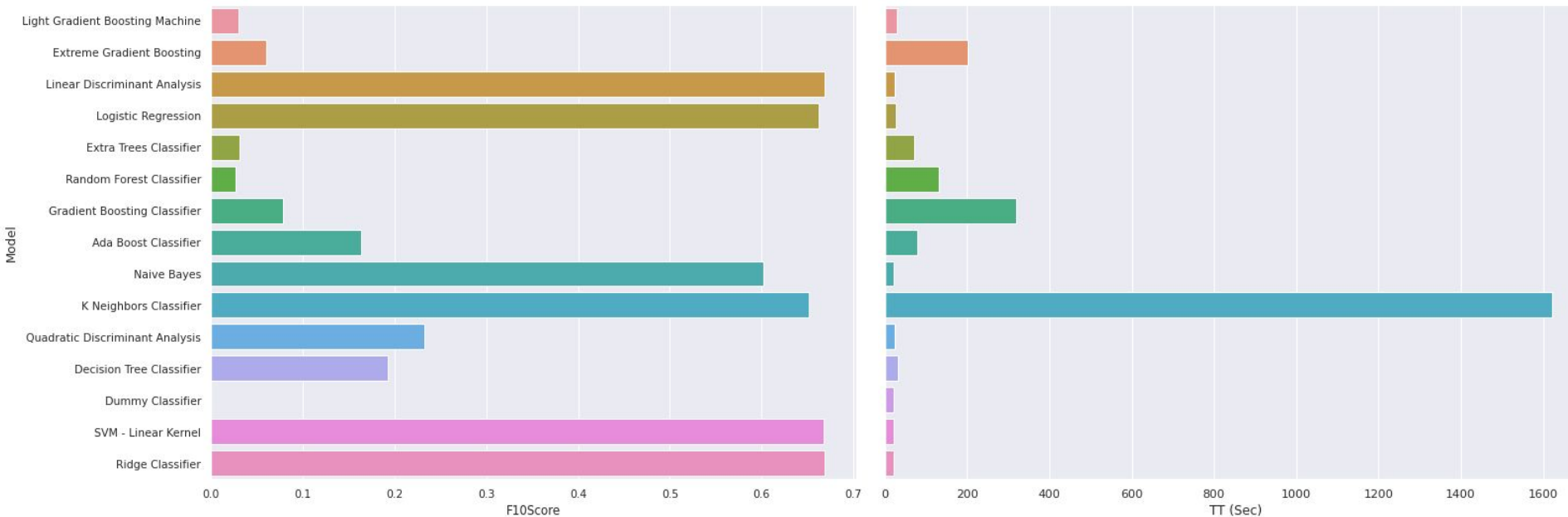
Choix des modèles AUC



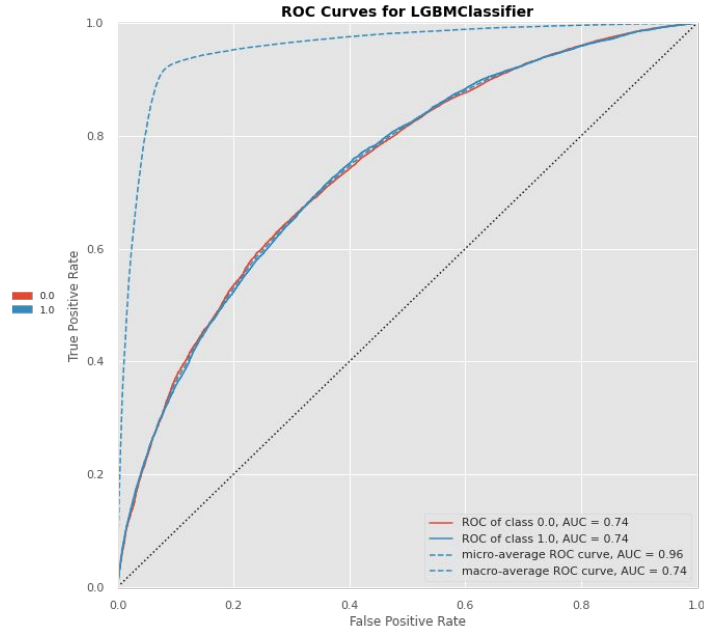
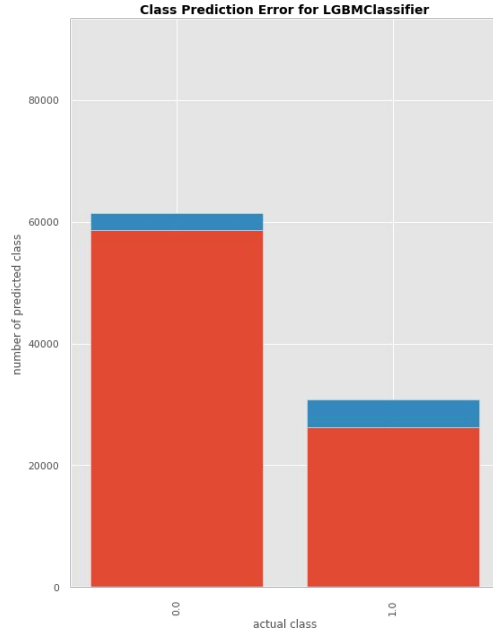
Choix des modèles F1 Score



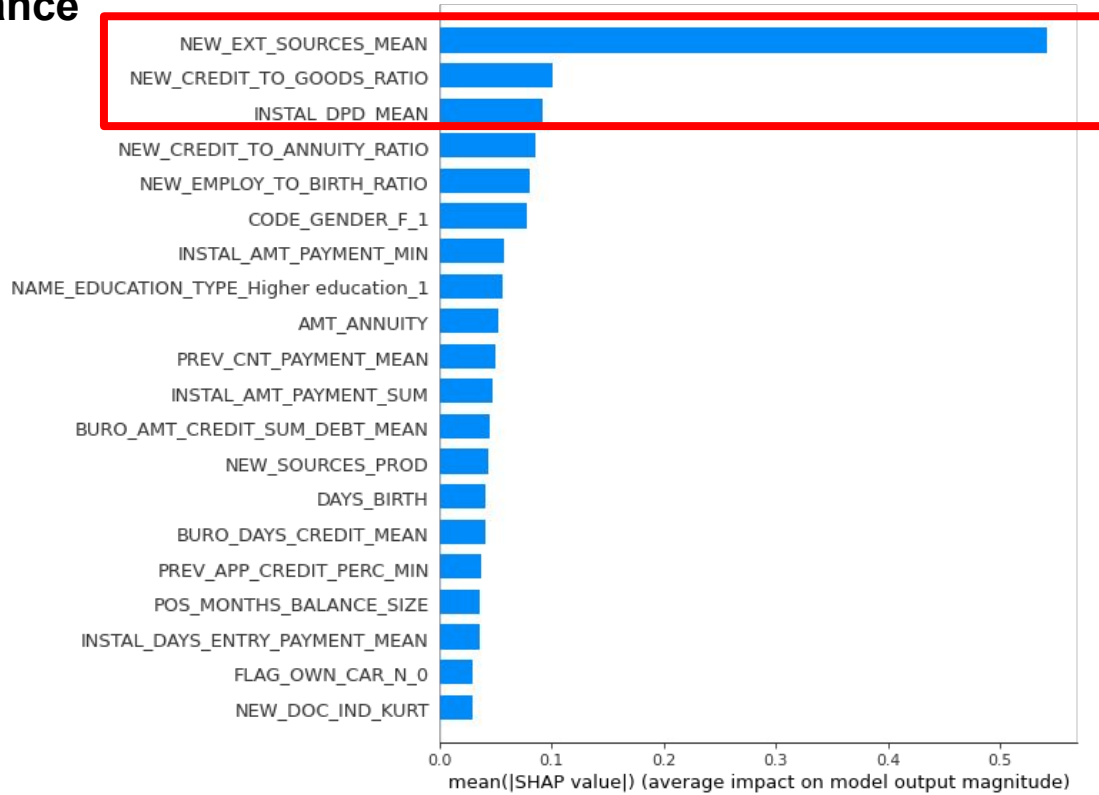
Choix des modèles F10 Score



Modèle Choisi LGBM



Features Importance

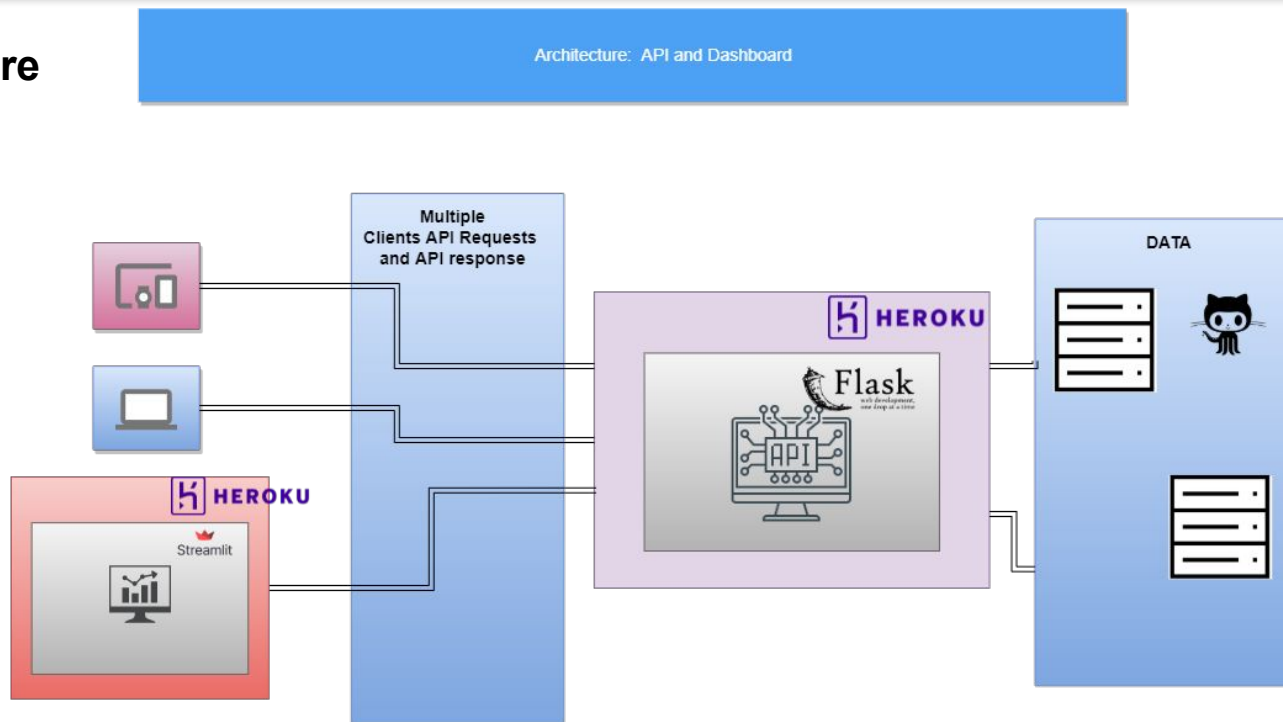


Dashboard Interactif : Architecture et Déploiement

Technologies

01	Flask	Création API
02	Streamlit	Création Dashboard
03	Heroku	Déploiement
04	Git / GitHub	Versioning / Stockage de data

Architecture



API

<https://khalil-henchi-oc-p7-api.herokuapp.com/>

[← → ↺](#) khalil-henchi-oc-p7-api.herokuapp.com

API for Home Credit Default Risk Prediction, created by Khalil Henchi

[← → ×](#) khalil-henchi-oc-p7-api.herokuapp.com/get_data/

```
{
  "X": {
    "SK_ID_CURR":
    {
      "0": 389086, "1": 159006, "2": 111210, "3": 393584, "4": 159312, "5": 446262, "6": 383229, "7": 455506, "8": 252474, "9": 440644, "10": 43506, "11": 182351, "12": 379400, "13": 174071, "14": 429171, "15": 437201, "16": 204888, "17": 158740, "18": 370063, "19": 181700, "20": 43506, "21": 371453, "22": 126010, "23": 273533, "24": 243414, "25": 174798, "26": 453208, "27": 328356, "28": 167258, "29": 275215, "30": 424477, "31": 348192, "32": 397708, "33": 208646, "34": 447974, "35": 126038, "36": 410455, "37": 289540, "38": 331902, "39": 174686, "40": 202827, "41": 5018, "42": 347369, "43": 246188, "44": 327294, "45": 204464, "46": 368021, "47": 378268, "48": 450283, "49": 282254, "50": 337818, "51": 3067, "52": 319104, "53": 194104, "54": 150599, "55": 425789, "56": 150368, "57": 196277, "58": 418384, "59": 257313, "60": 392724, "61": 448914, "62": 407002, "63": 294447, "64": 149762, "65": 407792, "66": 427030, "67": 448493, "68": 141713, "69": 257674, "70": 324692, "71": 357908, "72": 9273, "73": 245781, "74": 124585, "75": 128222, "76": 174556, "77": 107122, "78": 388802, "79": 102212, "80": 207233, "81": 279316, "82": 272615, "83": 438190, "84": 1081, "85": 189992, "86": 1091, "87": 273171, "88": 1101, "89": 273220, "90": 1111, "91": 356868, "92": 420890, "93": 309587, "94": 1141, "95": 196111, "96": 1151, "97": 2846, "98": 1191, "99": 389047, "100": 411392, "101": 274427, "102": 1221, "103": 138394, "104": 1231, "105": 416837, "106": 1241, "107": 181306, "108": 1251, "109": 245374, "110": 1261, "111": 116699, "112": 1271, "113": 329657, "114": 11445473, "115": 171472, "116": 1331, "117": 355776, "118": 1341, "119": 369399, "120": 1351, "121": 403006, "122": 1361, "123": 226928, "124": 1371, "125": 227612, "126": 1381, "127": 166639, "128": 1391, "129": 395642, "130": 1401, "131": 9943, "132": 1441, "133": 328803, "134": 1451, "135": 143639, "136": 1461, "137": 300236, "138": 1471, "139": 412438, "140": 1481, "141": 412779, "142": 1491, "143": 339669, "144": 1501, "145": 263272, "146": 1511, "147": 313004, "148": 1521, "149": 174096, "150": 1561, "151": 102477, "152": 1571, "153": 316998, "154": 1581, "155": 190699, "156": 1591, "157": 123787, "158": 1601, "159": 287715, "160": 1611, "161": 214545, "162": 1621, "163": 108277, "164": 1631, "165": 384023, "166": 1641, "167": 157365, "168": 1691, "169": 404159, "170": 1701, "171": 266192, "172": 1721, "173": 402712, "174": 1731, "175": 236474, "176": 1741, "177": 420956, "178": 1751, "179": 421274, "180": 1761, "181": 405421, "182": 1811, "183": 251438, "184": 1821, "185": 234001, "186": 1831, "187": 138503, "188": 1841, "189": 354551, "190": 1851, "191": 251873, "192": 1861, "193": 207460, "194": 1871, "195": 147914, "196": 1881, "197": 191900, "198": 1891, "199": 397340, "200": 311, "201": 213108, "202": 1941, "203": 408282, "204": 1951, "205": 202905, "206": 1961, "207": 176427, "208": 1971, "209": 331152, "210": 1981, "211": 305498, "212": 1991, "213": 311253, "214": 2001, "215": 383532, "216": 2011, "217": 198037, "218": 2021, "219": 5319, "220": 2061, "221": 298971, "222": 2071, "223": 245352, "224": 2081, "225": 134820, "226": 2091, "227": 167920, "228": 2101, "229": 424379, "230": 2111, "231": 306953, "232": 2121, "233": 292371, "234": 2131, "235": 168182, "236": 2141, "237": 219657, "238": 2181, "239": 422629, "240": 2191, "241": 445229, "242": 2201, "243": 449013, "244": 2211, "245": 243997, "246": 2221, "247": 330704, "248": 2231, "249": 189010, "250": 2241, "251": 400324, "252": 2251, "253": 327810, "254": 2261, "255": 25718, "256": 2271, "257": 162622, "258": 2311, "259": 409096, "260": 2321, "261": 241145, "262": 2331, "263": 191061, "264": 2341, "265": 327691, "266": 2351, "267": 401937, "268": 2361, "269": 356276, "270": 2371, "271": 437960, "272": 2381, "273": 225064, "274": 2391, "275": 3162, "276": 2431, "277": 318845, "278": 2441, "279": 372153, "280": 2451, "281": 247520, "282": 2461, "283": 176823, "284": 2471, "285": 179469, "286": 2481, "287": 109412, "288": 2491, "289": 102779, "290": 2501, "291": 316343, "292": 2511, "293": 390868, "294": 2551, "295": 215132, "296": 2561, "297": 366272, "298": 2571, "299": 296200, "300": 2581, "301": 414925, "302": 2591, "303": 325518, "304": 2601, "305": 314626, "306": 2611, "307": 195639, "308": 2621, "309": 266626, "310": 2631, "311": 363748, "312": 2641, "313": 4705, "314": 2681, "315": 315976, "316": 2691, "317": 376395, "318": 2701, "319": 108945, "320": 2711, "321": 139878, "322": 2721, "323": 351917, "324": 2731, "325": 149758, "326": 2741, "327": 128019, "328": 2751, "329": 144652, "330": 2761, "331": 279957, "332": 2801, "333": 410850, "334": 2811, "335": 455181, "336": 2821, "337": 103860, "338": 2831, "339": 174737, "340": 2841, "341": 222160, "342": 2851, "343": 383798, "344": 2861, "345": 149941, "346": 2871, "347": 341139, "348": 2881, "349": 396633, "350": 2891, "351": 137823, "352": 2931, "353": 350584, "354": 2941, "355": 285443, "356": 2951, "357": 258765, "358": 2961, "359": 341694, "360": 2971, "361": 396967, "362": 2981, "363": 114565, "364": 2991, "365": 446879, "366": 3001, "367": 420426, "368": 3011, "369": 17, "370": 61, "371": 3051, "372": 143304, "373": 3061, "374": 151946, "375": 3071, "376": 236245, "377": 3081, "378": 254423, "379": 3091, "380": 308942, "381": 3101, "382": 221745, "383": 3111, "384": 221997, "385": 3121, "386": 132633, "387": 3131, "388": 406340, "389": 3171, "390": 397763, "391": 3181, "392": 246265, "393": 3191, "394": 208015, "395": 3201, "396": 412591, "397": 3211, "398": 279529, "399": 3221, "400": 313046, "401": 3231, "402": 370175, "403": 3241, "404": 229123, "405": 3251, "406": 249757, "407": 3261, "408": 9354, "409": 3301, "410": 375153, "411": 3311, "412": 265660, "413": 3321, "414": 329770, "415": 3331, "416": 126731, "417": 3341, "418": 188623, "419": 3351, "420": 420933, "421": 3361, "422": 342955, "423": 3371, "424": 342391, "425": 3381, "426": 362844, "427": 3421, "428": 441167, "429": 3431, "430": 260214, "431": 3441, "432": 247949, "433": 3451, "434": 206613, "435": 3461, "436": 127270, "437": 3471, "438": 434397, "439": 3481, "440": 366863, "441": 3491, "442": 131363, "443": 3511, "444": 306190, "445": 3551, "446": 439588, "447": 3561, "448": 441024, "449": 3571, "450": 153148, "451": 3581, "452": 238430, "453": 3591, "454": 234410, "455": 3601, "456": 205514, "457": 3611, "458": 341180, "459": 3621, "460": 144468, "461": 3631, "462": 3161, "463": 3161, "464": 3161, "465": 3161, "466": 3161, "467": 3161, "468": 3161, "469": 3161, "470": 3161, "471": 3161, "472": 3161, "473": 3161, "474": 3161, "475": 3161, "476": 3161, "477": 3161, "478": 3161, "479": 3161, "480": 3161, "481": 3161, "482": 3161, "483": 3161, "484": 3161, "485": 3161, "486": 3161, "487": 3161, "488": 3161, "489": 3161, "490": 3161, "491": 3161, "492": 3161, "493": 3161, "494": 3161, "495": 3161, "496": 3161, "497": 3161, "498": 3161, "499": 3161, "500": 3161, "501": 3161, "502": 3161, "503": 3161, "504": 3161, "505": 3161, "506": 3161, "507": 3161, "508": 3161, "509": 3161, "510": 3161, "511": 3161, "512": 3161, "513": 3161, "514": 3161, "515": 3161, "516": 3161, "517": 3161, "518": 3161, "519": 3161, "520": 3161, "521": 3161, "522": 3161, "523": 3161, "524": 3161, "525": 3161, "526": 3161, "527": 3161, "528": 3161, "529": 3161, "530": 3161, "531": 3161, "532": 3161, "533": 3161, "534": 3161, "535": 3161, "536": 3161, "537": 3161, "538": 3161, "539": 3161, "540": 3161, "541": 3161, "542": 3161, "543": 3161, "544": 3161, "545": 3161, "546": 3161, "547": 3161, "548": 3161, "549": 3161, "550": 3161, "551": 3161, "552": 3161, "553": 3161, "554": 3161, "555": 3161, "556": 3161, "557": 3161, "558": 3161, "559": 3161, "560": 3161, "561": 3161, "562": 3161, "563": 3161, "564": 3161, "565": 3161, "566": 3161, "567": 3161, "568": 3161, "569": 3161, "570": 3161, "571": 3161, "572": 3161, "573": 3161, "574": 3161, "575": 3161, "576": 3161, "577": 3161, "578": 3161, "579": 3161, "580": 3161, "581": 3161, "582": 3161, "583": 3161, "584": 3161, "585": 3161, "586": 3161, "587": 3161, "588": 3161, "589": 3161, "590": 3161, "591": 3161, "592": 3161, "593": 3161, "594": 3161, "595": 3161, "596": 3161, "597": 3161, "598": 3161, "599": 3161, "600": 3161, "601": 3161, "602": 3161, "603": 3161, "604": 3161, "605": 3161, "606": 3161, "607": 3161, "608": 3161, "609": 3161, "610": 3161, "611": 3161, "612": 3161, "613": 3161, "614": 3161, "615": 3161, "616": 3161, "617": 3161, "618": 3161, "619": 3161, "620": 3161, "621": 3161, "622": 3161, "623": 3161, "624": 3161, "625": 3161, "626": 3161, "627": 3161, "628": 3161, "629": 3161, "630": 3161, "631": 3161, "632": 3161, "633": 3161, "634": 3161, "635": 3161, "636": 3161, "637": 3161, "638": 3161, "639": 3161, "640": 3161, "641": 3161, "642": 3161, "643": 3161, "644": 3161, "645": 3161, "646": 3161, "647": 3161, "648": 3161, "649": 3161, "650": 3161, "651": 3161, "652": 3161, "653": 3161, "654": 3161, "655": 3161, "656": 3161, "657": 3161, "658": 3161, "659": 3161, "660": 3161, "661": 3161, "662": 3161, "663": 3161, "664": 3161, "665": 3161, "666": 3161, "667": 3161, "668": 3161, "669": 3161, "670": 3161, "671": 3161, "672": 3161, "673": 3161, "674": 3161, "675": 3161, "676": 3161, "677": 3161, "678": 3161, "679": 3161, "680": 3161, "681": 3161, "682": 3161, "683": 3161, "684": 3161, "685": 3161, "686": 3161, "687": 3161, "688": 3161, "689": 3161, "690": 3161, "691": 3161, "692": 3161, "693": 3161, "694": 3161, "695": 3161, "696": 3161, "697": 3161, "698": 3161, "699": 3161, "700": 3161, "701": 3161, "702": 3161, "703": 3161, "704": 3161, "705": 3161, "706": 3161, "707": 3161, "708": 3161, "709": 3161, "710": 3161, "711": 3161, "712": 3161, "713": 3161, "714": 3161, "715": 3161, "716": 3161, "717": 3161, "718": 3161, "719": 3161, "720": 3161, "721": 3161, "722": 3161, "723": 3161, "724": 3161, "725": 3161, "726": 3161, "727": 3161, "728": 3161, "729": 3161, "730": 3161, "731": 3161, "732": 3161, "733": 3161, "734": 3161, "735": 3161, "736": 3161, "737": 3161, "738": 3161, "739": 3161, "740": 3161, "741": 3161, "742": 3161, "743": 3161, "744": 3161, "745": 3161, "746": 3161, "747": 3161, "748": 3161, "749": 3161, "750": 3161, "751": 3161, "752": 3161, "753": 3161, "754": 3161, "755": 3161, "756": 3161, "757": 3161, "758": 3161, "759": 3161, "760": 3161, "761": 3161, "762": 3161, "763": 3161, "764": 3161, "765": 3161, "766": 3161, "767": 3161, "768": 3161, "769": 3161, "770": 3161, "771": 3161, "772": 3161, "773": 3161, "774": 3161, "775": 3161, "776": 3161, "777": 3161, "778": 3161, "779": 3161, "780": 3161, "781": 3161, "782": 3161, "783": 3161, "784": 3161, "785": 3161, "786": 3161, "787": 3161, "788": 3161, "789": 3161, "790": 3161, "791": 3161, "792": 3161, "793": 3161, "794": 3161, "795": 3161, "796": 3161, "797": 3161, "798": 3161, "799": 3161, "800": 3161, "801": 3161, "802": 3161, "803": 3161, "804": 3161, "805": 3161, "806": 3161, "807": 3161, "808": 3161, "809": 3161, "810": 3161, "811": 3161, "812": 3161, "813": 3161, "814": 3161, "815": 3161, "816": 3161, "817": 3161, "818": 3161, "819": 3161, "820": 3161, "821": 3161, "822": 3161, "823": 3161, "824": 3161, "825": 3161, "826": 3161, "827": 3161, "828": 3161, "829": 3161, "830": 3161, "831": 3161, "832": 3161, "833": 3161, "834": 3161, "835": 3161, "836": 3161, "837": 3161, "838": 3161, "839": 3161, "840": 3161, "841": 3161, "842": 3161, "843": 3161, "844": 3161, "845": 3161, "846": 3161, "847": 3161, "848": 3161, "849": 3161, "850": 3161, "851": 3161, "852": 3161, "853": 3161, "854": 3161, "855": 3161, "856": 3161, "857": 3161, "858": 3161, "859": 3161, "860": 3161, "861": 3161, "862": 3161, "863": 3161, "864": 3161, "865": 3161, "866": 3161, "867": 3161, "868": 3161, "869": 3161, "870": 3161, "871": 3161, "872": 3161, "873": 3161, "874": 3161, "875": 3161, "876": 3161, "877": 3161, "878": 3161, "879": 3161, "880": 3161, "881": 3161, "882": 3161, "883": 3161, "884": 3161, "885": 3161, "886": 3161, "887": 3161, "888": 3161, "889": 3161, "890": 3161, "891": 3161, "892": 3161, "893": 3161, "894": 3161, "895": 3161, "896": 3161, "897": 3161, "898": 3161, "899": 3161, "900": 3161, "901": 3161, "902": 3161, "903": 3161, "904": 3161, "905": 3161, "906": 3161, "907": 3161, "908": 3161, "909": 3161, "910": 3161, "911": 3161, "912": 3161, "913": 3161, "914": 3161, "915": 3161, "916": 3161, "917": 3161, "918": 3161, "919": 3161, "920": 3161, "921": 3161, "922": 3161, "923": 3161, "924": 3161, "925": 3161, "926": 3161, "927": 3161, "928": 3161, "929": 3161, "930": 3161, "931": 3161, "932": 3161, "933": 3161, "934": 3161, "935": 3161, "936": 3161, "937": 3161, "938": 3161, "939": 3161, "940": 3161, "941": 3161, "942": 3161, "943": 3161, "944": 3161, "945": 3161, "946": 3161, "947": 3161, "948": 3161, "949": 3161, "950": 3161, "951": 3161, "952": 3161, "953": 3161, "954": 3161, "955": 3161, "956": 3161, "957": 3161, "958": 3161, "959": 3161, "960": 3161, "961": 3161, "962": 3161, "963": 3161, "964": 3161, "965": 3161, "966": 3161, "967": 3161, "968": 3161, "969": 3161, "970": 3161, "971": 3161, "972": 3161, "973": 3161, "974": 3161, "975": 3161, "976": 3161, "977": 3161, "978": 3161, "979": 3161, "980": 3161, "981": 3161, "982": 3161, "983": 3161, "984": 3161, "985": 3161, "986": 3161, "987": 3161, "988": 3161, "989": 3161, "990": 3161, "991": 3161, "992": 3161, "993": 3161, "994": 3161, "995": 3161, "996": 3161, "997": 3161, "998": 3161, "999": 3161, "1000": 3161, "1001": 3161, "1002": 3161, "1003": 3161, "1004": 3161, "1005": 3161, "1006": 3161, "1007": 3161, "1008": 3161, "1009": 3161, "1010": 3161, "1011": 3161, "1012": 3161, "1013": 3161, "1014": 3161, "1015": 3161, "1016": 3161, "1017": 3161, "1018": 3161, "1019": 3161, "1020": 3161, "1021": 3161, "1022": 3161, "1023": 3161, "1024": 3161, "1025": 3161, "1026": 3161, "1027": 3161, "1028": 3161, "1029": 3161, "1030": 3161, "1031": 3161, "1032": 3161, "1033": 3161, "1034": 3161, "1035": 3161, "1036": 3161, "1037": 3161, "1038": 3161, "1039": 3161, "1040": 3161, "1041": 3161, "1042": 3161, "1043": 3161, "1044": 3161, "1045": 3161, "1
```

Présentation Dashboard Interactif

Fonctionnalités

- ❑ **Classification**
 - ❑ **Prédiction**
- ❑ **Analyse**
 - ❑ **Visualisation**
 - ❑ **Comparaison de clients**
- ❑ **Interprétabilité**
 - ❑ **Globale: variables importantes**
 - ❑ **Locale: SHAP**

<https://khalil-henchi-oc-p7-dashboard.herokuapp.com/>



Home Credit Default Risk Dashboard

Khalil Henchi

Prêt à dépenser

Mission du dashboard

ID Client

Veuillez saisir un ID client :

200038

Probabilité

0.15

Seuil : 0.5

Décision : Crédit



Score : 0

Données client

☐ Afficher les informations descriptives du client

Jeu de Données

☐ Afficher les informations descriptives de l'ensemble des clients

Interprétation des résultats

☐ Afficher l'interprétation des résultats

← → ↺ khalil-henchi-oc-p7-dashboard.herokuapp.com



Données client

☒ Afficher les informations descriptives du client

	AMT_CREDIT	AMT_ANNUITY	AMT_GOODS_PRICE	REGION_POPULATION_RELATIVE	DAYS_BIRTH	DAYS_EMPLOYED	DAYS_REGISTRATION	DAYS_ID_PUBLISH	REGION_RATING_CLIENT	REG_CITY_NOT_LIVE_CITY	EXIT_SOURCE_1	EXIT_SOURCE_2	EXIT_SOURCE_3	YEARS_BEGINEYPUBLISH	LIVINGAREA_MED	DEF_30_CNT_SOCIAL_CREDIT
100039	733,315,0000	39089	675000	0.0102	-11694	-1090	-1007	-1007	2	1	0.5060	0.3217	0.4118	0.9816	0.0749	

Jeu de Données

☒ Afficher les informations descriptives de l'ensemble des clients

Quel jeu voulez-vous analyser ?

Ensemble de clients

	AMT_ANNUITY	AMT_GOODS_PRICE	REGION_POPULATION_RELATIVE	DAYS_BIRTH	DAYS_REGISTRATION	DAYS_ID_PUBLISH	EXIT_SOURCE_1	EXIT_SOURCE_2	DEF_30_CNT_SOCIAL_CREDIT	DEF_60_CNT_SOCIAL_CREDIT	DAYS_LAST_PHONE_CHANGE	NEW_CREDIT_TO_ANNUITY_RATIO	NEW_CREDIT_TO_GOODS_PRICE_RATIO	NEW_DOC_INDO_K	NEW_INCOME_RATIO
369086	0.0937	-0.8590	-0.7433	1.8639	1.0416	1.2869	-1.6024	0.8938	-0.3587	-0.3022	-0.1819	-1.7118	-0.5075	0.3716	0.4
159008	0.4022	0.7042	0.2338	-0.8854	0.4340	0.0781	-0.0035	0.0710	-0.3587	-0.3022	0.1728	0.8600	0.7384	0.3716	0.4
111210	1.8195	0.6527	1.0001	0.8510	0.2627	-0.1096	-1.1394	1.9799	-0.3587	-0.3022	-0.1602	-1.2009	0.0085	0.3716	0.4
393584	1.4868	1.0805	-1.2593	-0.1189	-0.2909	-0.8864	-0.0035	-0.9494	-0.3587	-0.3022	-0.4483	-0.1831	0.1226	0.3716	0.4
159312	2.0895	2.1484	-0.8088	-0.5280	-1.1721	0.8678	-0.0035	0.0188	-0.3587	-0.3022	-1.7347	-0.9279	-0.0886	0.3716	-0.7
446282	1.7422	1.1042	-1.3076	-0.7474	0.8448	0.7363	-0.0035	-1.3726	-0.3587	-0.3022	-0.8487	-0.1824	1.0020	0.3716	-1.4
361229	0.6507	0.5214	0.0797	-1.0233	-0.7792	1.3760	1.0426	-0.7401	-0.3587	-0.3022	1.1229	0.5878	1.8807	0.3716	-1.4
455506	0.2396	0.0125	0.9340	1.3459	0.9634	0.8873	-0.0035	-0.8662	-0.3587	-0.3022	-0.6459	-0.2348	0.4488	0.3716	0.4
252474	0.3083	0.6527	-1.0447	1.4451	0.8087	0.4029	-0.0035	0.8982	-0.3587	-0.3022	1.3836	0.3600	-1.0962	0.3716	0.4
647642	0.1204	-0.5671	-0.6007	0.0708	-0.8087	1.1165	-0.0035	1.0410	-0.3587	-0.3022	-0.1828	-0.1470	1.4704	0.3716	0.7

Analyse Univariée

Quel ensemble voulez-vous analyser ?

BURD_AMT_CREDIT_MAX_OVERDUE_MEAN

Distribution de la variable : BURD_AMT_CREDIT_MAX_OVERDUE_MEAN



Comparaison avec profils de clients

Sélectionnez les variables à comparer :

similaires

SK_ID_CURR x AMT_ANNUITY x AMT_GOODS_PRICE x REGION_POPULATION_RELATIVE x DAYS_BIRTH



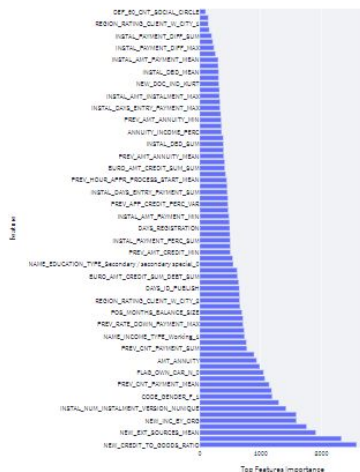
← → ↺ khalil-henchi-oc-p7-dashboard.herokuapp.com

Interprétation des résultats

☒ Afficher l'interprétation des résultats

Importance de Variables

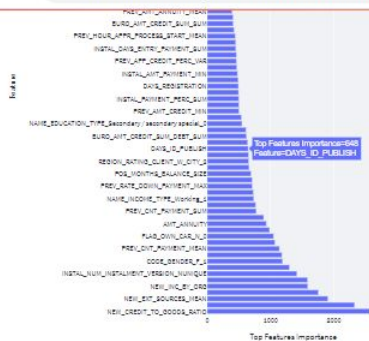
☒ Afficher importance de variables



Analyse SHAP

☐ Analyse SHAP

← → ↺ khalil-henchi-oc-p7-dashboard.herokuapp.com



Analyse SHAP

☒ Analyse SHAP

Veuillez choisir le plot SHAP à afficher :

Force Plot



Conclusion

Livrables

- Jupyter Notebook détaillant le travail de nettoyage et modélisation
- Support de présentation

Compétences acquises

- Prétraiter des données pour obtenir un jeu de données exploitable.
- Mettre en œuvre des techniques de réduction de déséquilibre.
- Déployer un dashboard pour interagir avec le modèle.

Perspectives

- Améliorer le dashboard
- Présenter l'outil aux experts métiers
- Optimiser l'imputation des données

***Merci pour votre
Attention !
A vos Questions***



khalilhenchi@gmail.com



/khalil-henchi

