



République Tunisienne Ministère de l'Enseignement Supérieur et de
la Recherche Scientifique

Licence en Big Data et Analyse de données



Prédiction du churn des clients Tunisie Télécom

Réalisé par:
Fatma Ezzahra Ben abdallah

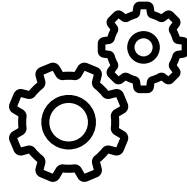


Encadré par:
Mr. Chedi Bechikh
Mr. Hatem Jemai

Année universitaire 2023 - 2024

Plan

Introduction
générale



Méthodologie
du projet



Conclusion et
perspectives

01

02

03

04

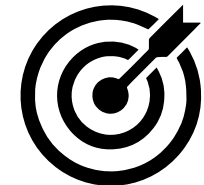
05



Cadre du
projet



Mise en
œuvre

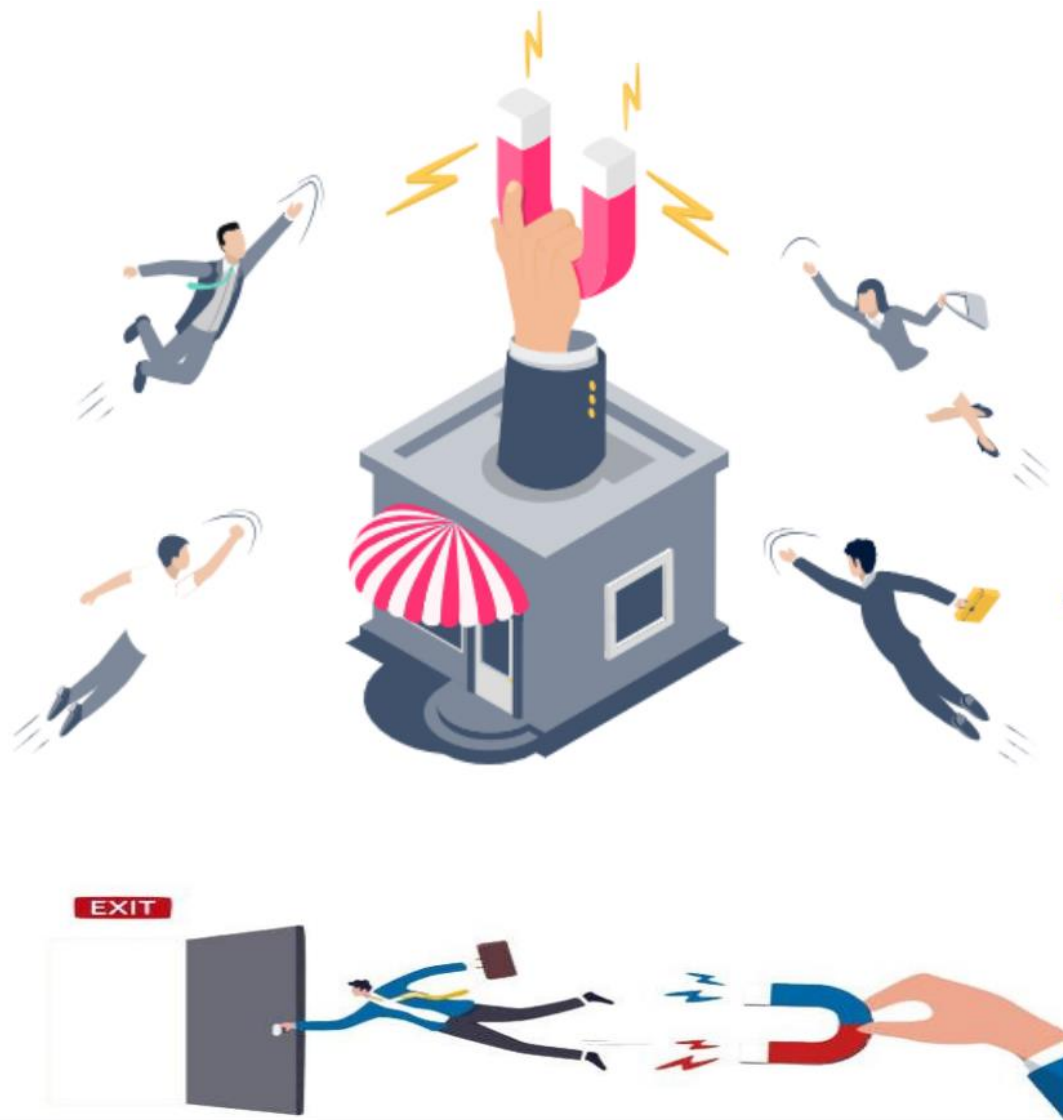




Introduction générale



Impact considérable sur le taux de résiliation des clients



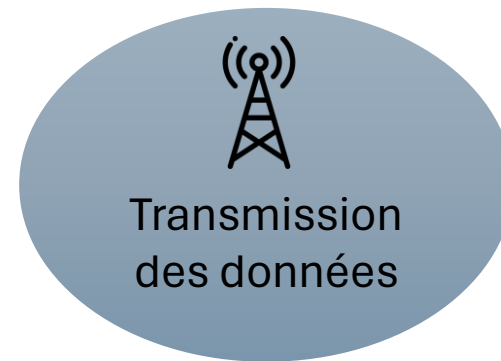
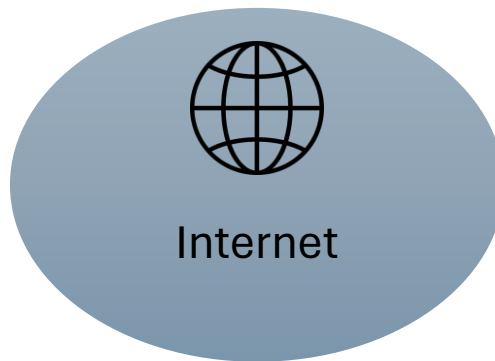
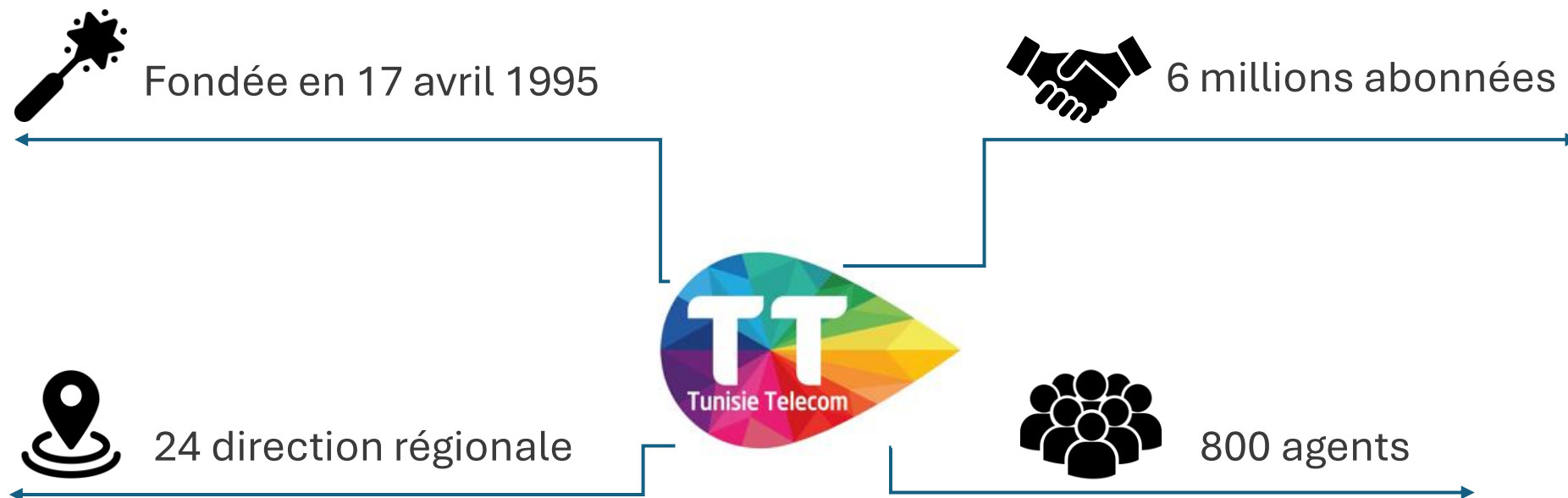


Cadre du projet

Présentation de
l'organisme
d'accueil

Problématique

Solution
proposée



Comment Tunisie Télécom peut-elle réduire le taux désabonnement des clients afin de maintenir sa part de marché ?





Modèle de prédiction du Churn en utilisant différents algorithmes d'apprentissage automatique supervisée

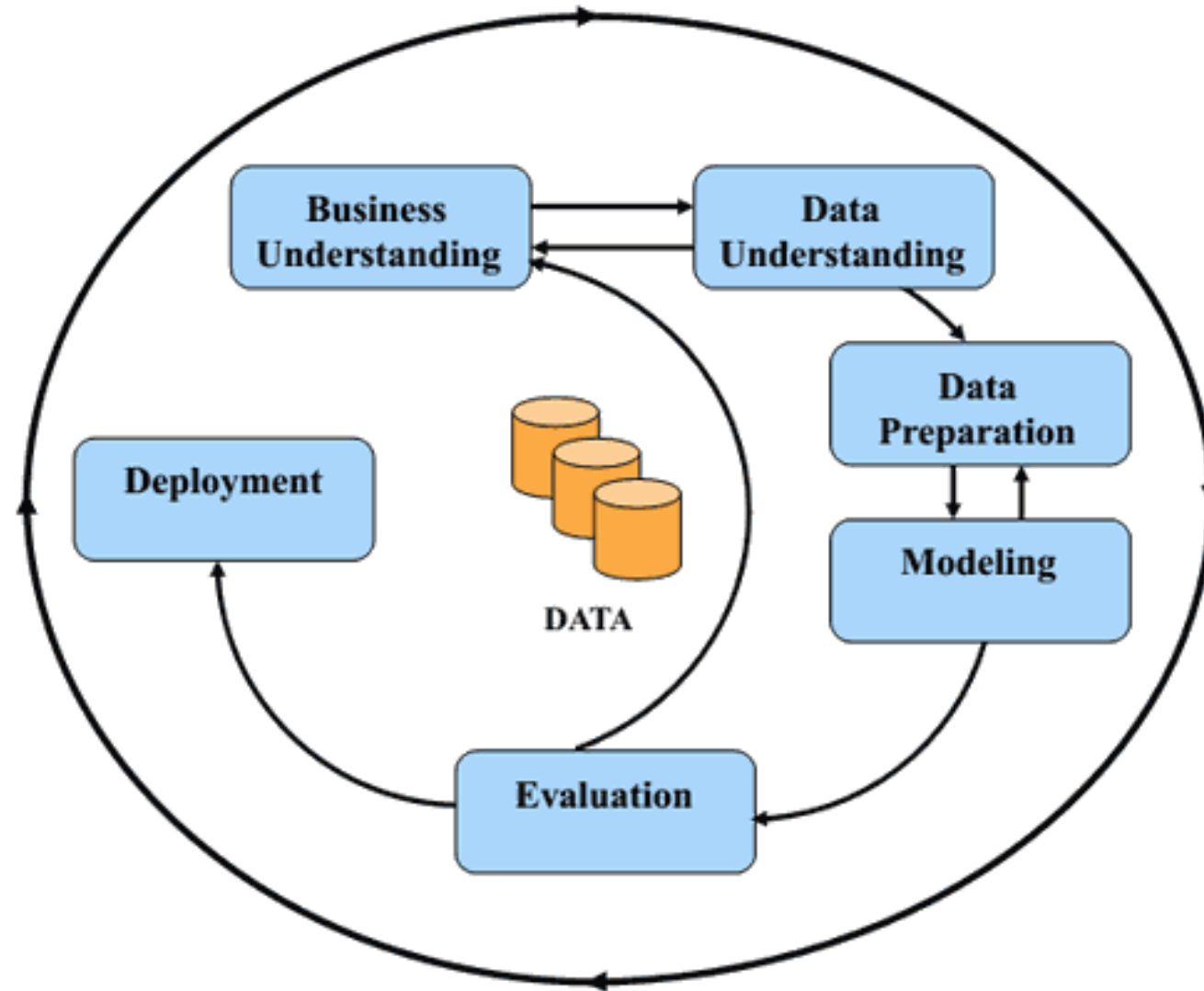


Élaboration des tableaux de bord

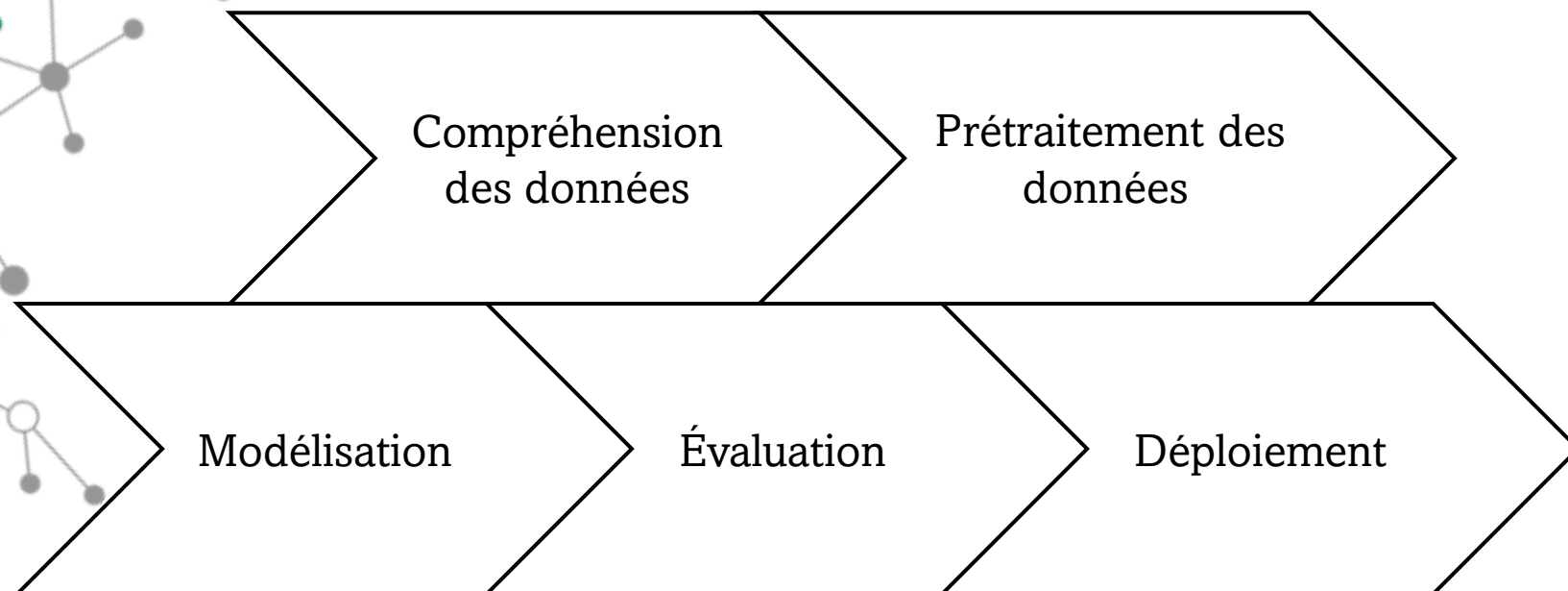


Méthodologie de projet

CRISP-DM : Cross-Industry Standard Process for Data Mining



Mise en œuvre



Compréhension des données

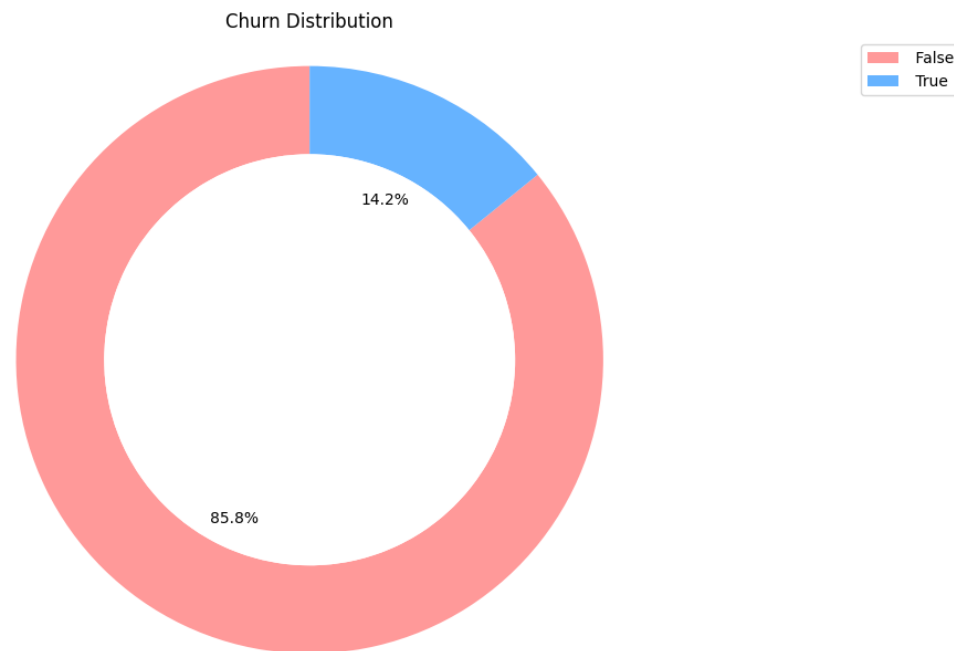
13

23 Variables

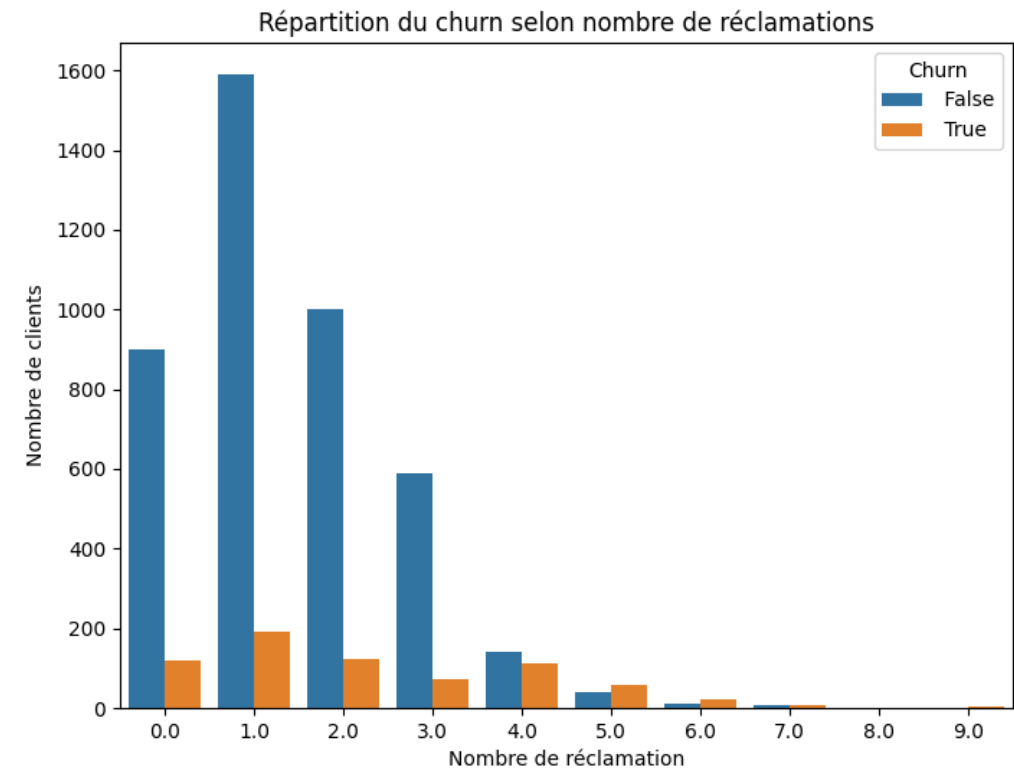
5000 Lignes

id_client	genre	age	marie	num_tel	nb_jours_a	duree_appel_j	nb_appel_j	cout_appel_j	duree_appel_n	nb_appel_nuit	cout_appel_n	duree_appel_n	nb_appel_nuit	cout_appel_n	duree_appel_n	nb_appel_nuit	cout_appel_n	active	nb_ms	nb_rec	churn	offer_type
382-4657	femme	37	yes	98505453	128	265.1	110	45.07	197.4	99	16.78	244.7	91	11.01	10	3	2.7	yes	25	1	False	Hayya
371-7191	homme	46	no	97321658	107	161.6	123	27.47	195.5	103	16.62	254.4	103	11.45	13.7	3	3.7	yes	26	1	False	PRE - 1=11
358-1921	homme	50	no	98653270	137	243.4	114	41.38	121.2	110	10.3	162.6	104	7.32	12.2	5	3.29	no	0	0	False	PRE - 900 bonus
375-9999	homme	78	yes	96303256	84	299.4	71	50.9	61.9	88	5.26	196.9	89	8.86	6.6	7	1.78	no	0	2	False	PRE - AHLA
330-6626	femme	75	yes	96412387	75	166.7	113	28.34	148.3	122	12.61	186.9	121	8.41	10.1	3	2.73	no	0	3	False	PRE - Binetna
391-8027	femme	23	no	98142367	118	223.4	98	37.98	220.6	101	18.75	203.9	118	9.18	6.3	6	1.7	no	0	0	False	PRE - Classic
355-9993	femme	67	yes	98501260	121	218.2	88	37.09	348.5	108	29.62	212.6	118	9.57	7.5	7	2.03	yes	24	3	False	PRE - Club Optimum P
329-9001	homme	52	yes	99606321	147	157	79	26.69	103.1	94	8.76	211.8	96	9.53	7.1	6	1.92	no	0	0	False	PRE - Corporate Optim
335-4719	femme	68	no	99421753	117	184.5	97	31.37	351.6	80	29.89	215.8	90	9.71	8.7	4	2.35	no	0	1	False	PRE - CSS 1000% New
330-8173	femme	43	yes	99203170	141	258.6	84	43.96	222	111	18.87	326.4	97	14.69	11.2	5	3.02	yes	37	0	False	PRE - CSS Mobile 1000
329-6603	homme	47	no	95606231	65	129.1	137	21.95	228.5	83	19.42	208.8	111	9.4	12.7	6	3.43	no	0	4	True	PRE - CSSM 35mil/min
344-9403	femme	25	yes	98741230	74	187.7	127	31.91	163.4	148	13.89	196	94	8.82	9.1	5	2.46	no	0	0	False	PRE - Day Pass
363-1107	femme	58	yes	98632140	168	128.8	96	21.9	104.9	71	8.92	141.1	128	6.35	11.2	2	3.02	no	0	1	False	PRE - Doublé
394-8006	femme	32	no	96321586	95	156.6	88	26.62	247.6	75	21.05	192.3	115	8.65	12.3	5	3.32	no	0	3	False	PRE - Double Reinstal
366-9238	femme	39	no	99201369	62	120.7	70	20.52	307.2	76	26.11	203	99	9.14	13.1	6	3.54	no	0	4	False	PRE - E.M. 1000%
351-7269	femme	58	yes	98512647	161	332.9	67	56.59	317.8	97	27.01	160.6	128	7.23	5.4	9	1.46	no	0	4	True	PRE - Elissa 300%
350-8884	femme	52	yes	93601476	85	196.4	139	33.39	280.9	90	23.88	89.3	75	4.02	13.8	4	3.73	yes	27	1	False	PRE - Employe TT
386-2923	femme	72	no	93602500	93	190.7	114	32.42	218.2	111	18.55	129.6	121	5.83	8.1	3	2.19	no	0	3	False	PRE - ESS 1000% New
356-2992	homme	79	no	98503260	76	189.7	66	32.25	212.8	65	18.09	165.7	108	7.46	10	5	2.7	yes	33	1	False	PRE - ESS Mobile 35mi
373-2782	homme	67	no	90032100	73	224.4	90	38.15	159.5	88	13.56	192.8	74	8.68	13	2	3.51	no	0	1	False	PRE - EST 1000% New
396-5800	femme	79	yes	98416870	147	155.1	117	26.37	239.7	93	20.37	208.8	133	9.4	10.6	4	2.86	no	0	0	False	PRE - Jawhara 35Mil
393-7984	femme	26	yes	98560321	77	62.4	89	10.61	169.9	121	14.44	209.6	64	9.43	5.7	6	1.54	no	0	5	True	PRE - New Elissa
358-1958	femme	30	yes	99652014	130	183	112	31.11	72.9	99	6.2	181.8	78	8.18	9.5	19	2.57	no	0	0	False	PRE - offre 40
350-2565	femme	22	no	98505463	111	110.4	103	18.77	137.3	102	11.67	189.6	105	8.53	7.7	6	2.08	no	0	2	False	PRE - Offre WL5
343-4696	femme	34	yes	98505459	132	81.1	86	13.79	245.2	72	20.84	237	115	10.67	10.3	2	2.78	no	0	0	False	PRE - Oulidha 1000%
331-3698	femme	37	yes	96302514	174	124.3	76	21.13	277.1	112	23.55	250.7	115	11.28	15.5	5	4.19	no	0	3	False	PRE - Oulidha 2000%
357-3817	femme	37	yes	98505453	57	213	115	36.21	191.1	112	16.24	182.7	115	8.22	9.5	3	2.57	yes	39	0	False	PRE - Pack cle 4G
418-6412	homme	42	yes	98505453	54	134.3	73	22.83	155.5	100	13.22	102.1	68	4.59	14.7	4	3.97	no	0	3	False	PRE - Pass Etudiant
353-2630	homme	64	no	98505453	20	190	109	32.3	258.2	84	21.95	181.5	102	8.17	6.3	6	1.7	no	0	0	False	PRE - Taraji Mobile 150
410-7789	femme	47	yes	98505453	49	119.3	117	20.28	215.1	109	18.28	178.7	90	8.04	11.1	1	3	no	0	1	False	PRE - Tawwa
416-8428	homme	23	yes	98505453	142	84.8	95	14.42	136.7	63	11.62	250.5	148	11.27	14.2	6	3.83	no	0	2	False	PRE - TM 35mil/min

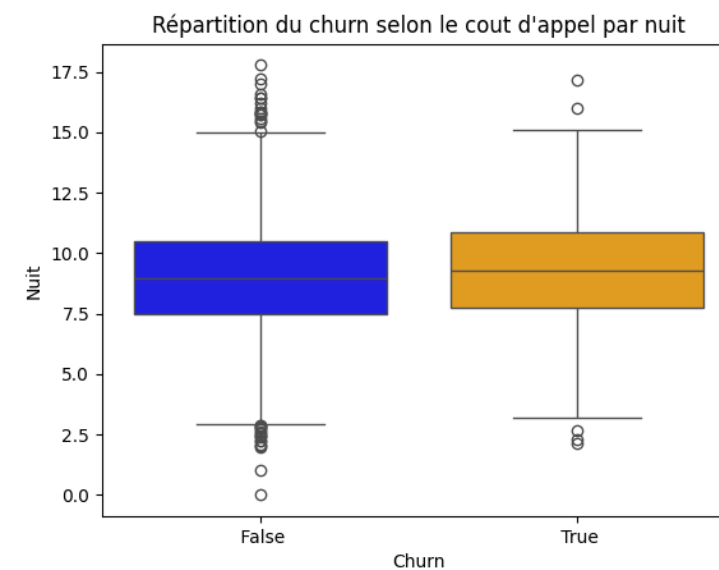
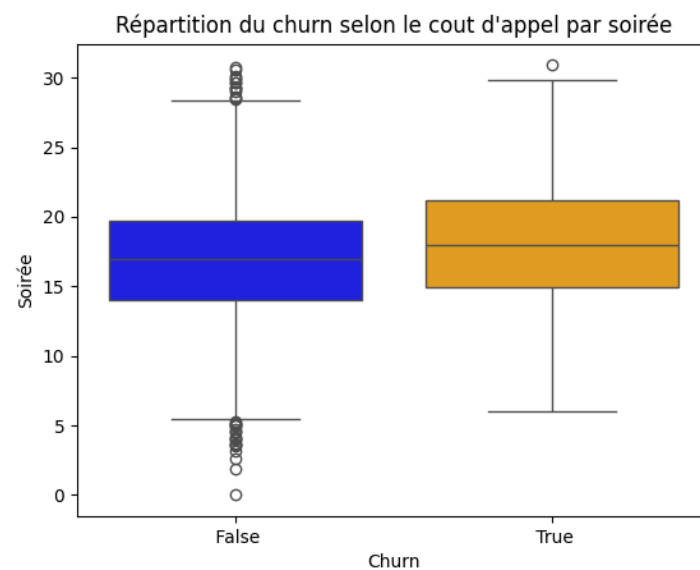
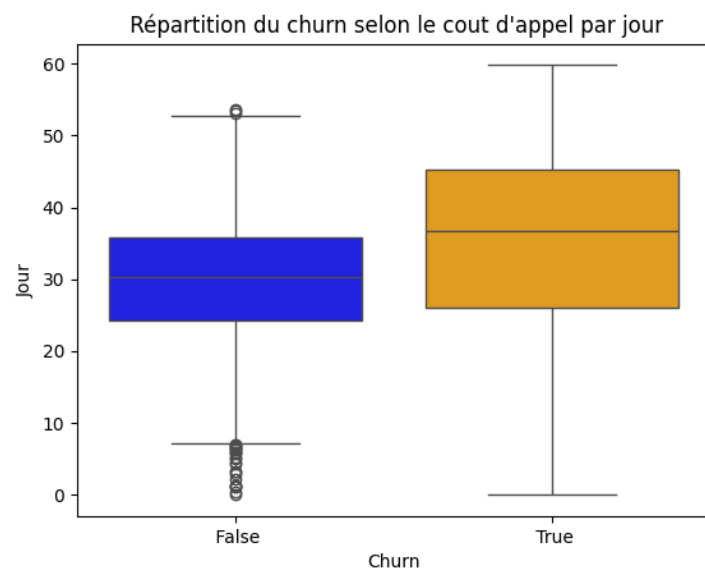
Répartition des clients en fonction de churn



Répartition du churn selon le nombre de réclamation



Répartition du churn selon le coût d'appel des différents types d'appels



Suppression des colonnes inutiles

```
id_client 5000
genre 2
age 62
marie 2
num_tel 33
nb_jours_abonne 219
duree_appel_jour 1961
nb_appel_jour 124
cout_appel_jour 1961
duree_appel_soiree 1879
nb_appel_soiree 127
cout_appel_soiree 1659
duree_appel_nuit 1853
nb_appel_nuit 132
cout_appel_nuit 1028
duree_appel_inter 170
nb_appel_inter 22
cout_appel_inter 170
active_msg_vocaux 3
nb_msg_vocaux 48
nb_reclamation 11
churn 3
offer_type 40
```

Suppression des valeurs dupliquées

```
#Vérifier s'il y a des doublons
print(df.duplicated().sum())
```

0

Traitement des valeurs manquantes

genre	0
age	0
marie	0
nb_jours_abonne	10
duree_appel_jour	0
nb_appel_jour	7
cout_appel_jour	0
duree_appel_soiree	0
nb_appel_soiree	6
cout_appel_soiree	0
duree_appel_nuit	0
nb_appel_nuit	11
cout_appel_nuit	0
duree_appel_inter	0
nb_appel_inter	12
cout_appel_inter	0
active_msg_vocaux	8
nb_msg_vocaux	0
nb_reclamation	12
churn	4
offer_type	0

Mean ()

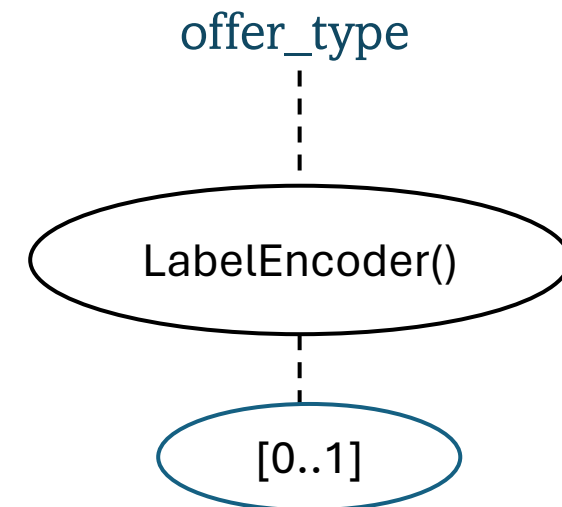
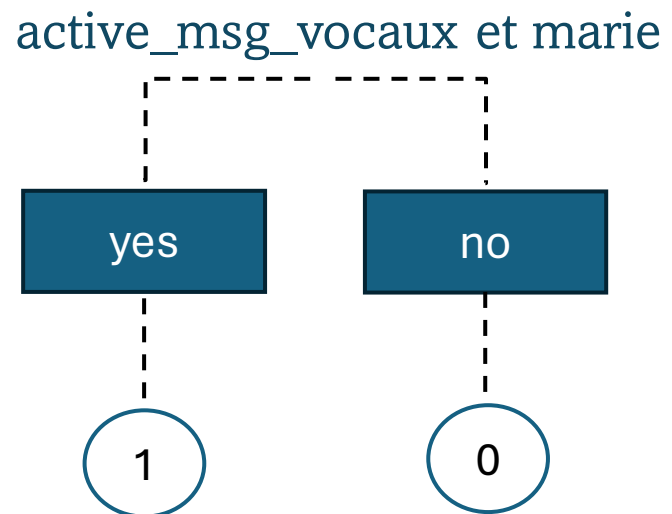
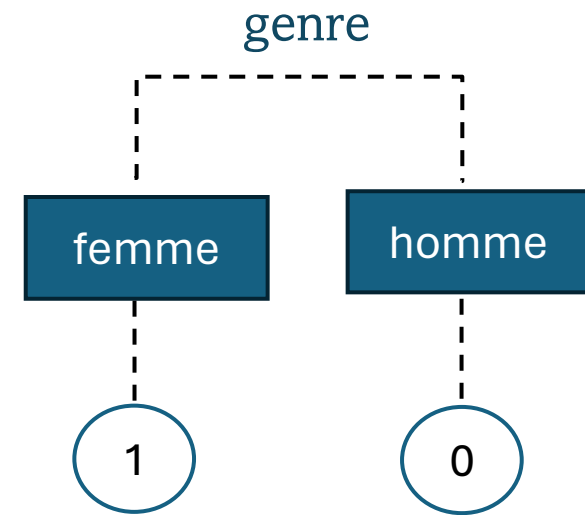
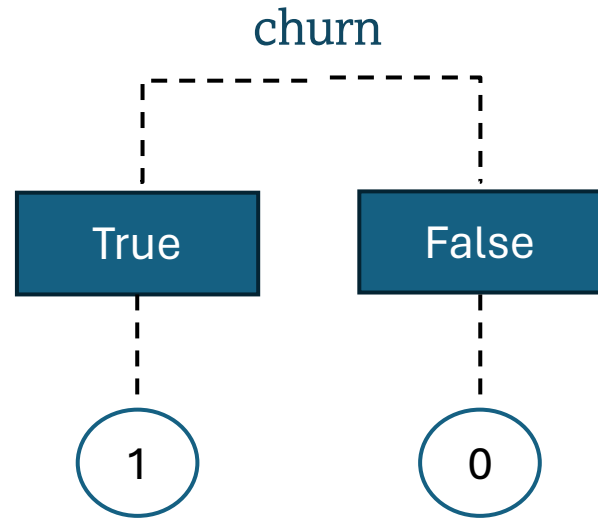
Mode ()

Ffill

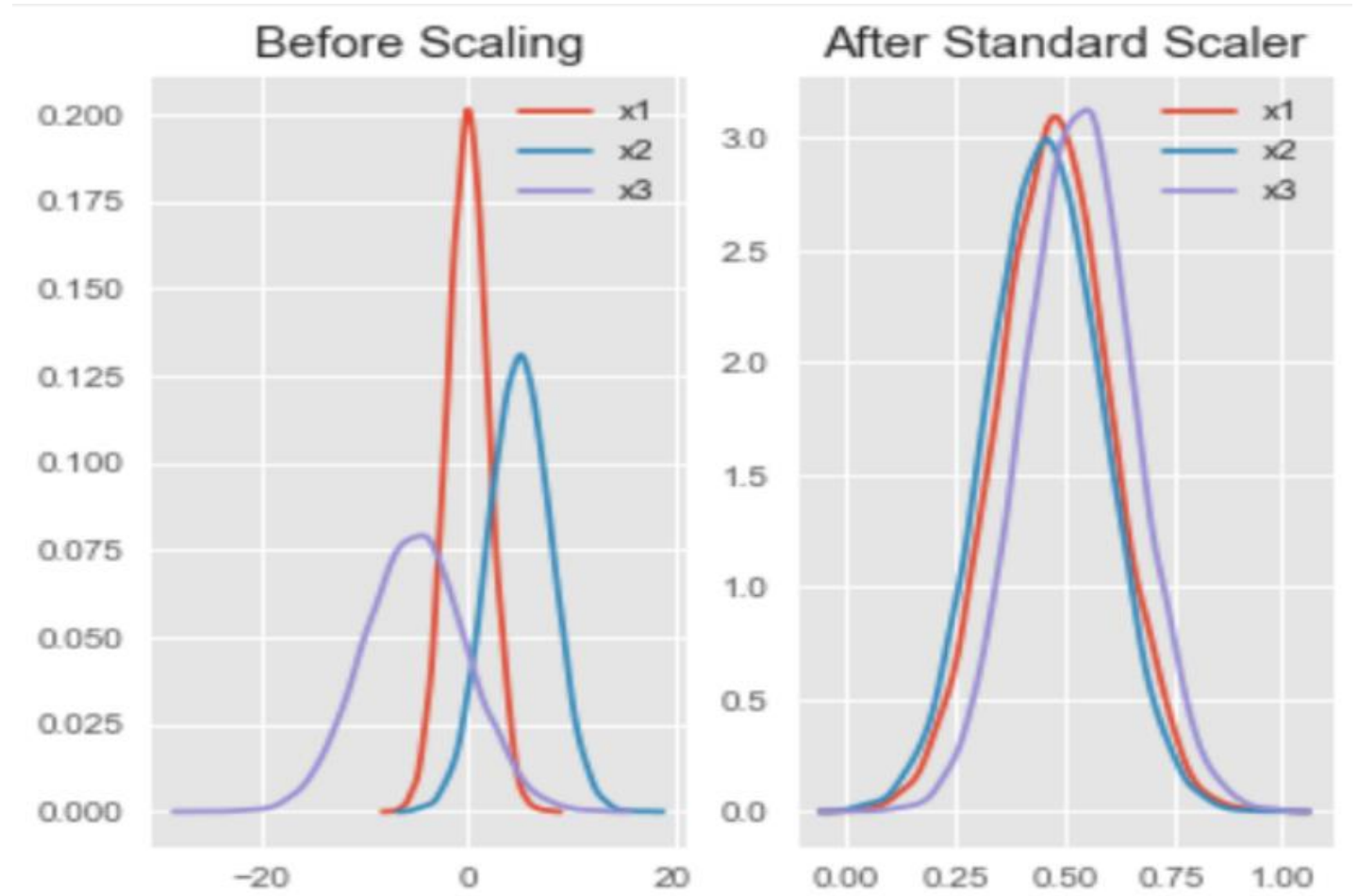
Dropna ()

Après traitement

genre	0
age	0
marie	0
nb_jours_abonne	0
duree_appel_jour	0
nb_appel_jour	0
cout_appel_jour	0
duree_appel_soiree	0
nb_appel_soiree	0
cout_appel_soiree	0
duree_appel_nuit	0
nb_appel_nuit	0
cout_appel_nuit	0
duree_appel_inter	0
nb_appel_inter	0
cout_appel_inter	0
active_msg_vocaux	0
nb_msg_vocaux	0
nb_reclamation	0
churn	0
offer_type	0



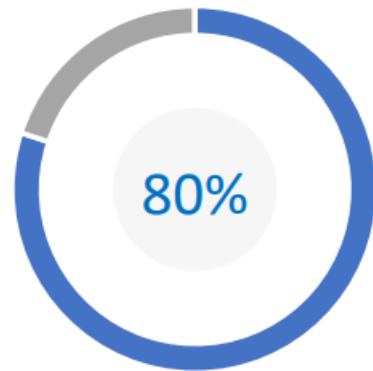
StandardScaler ()



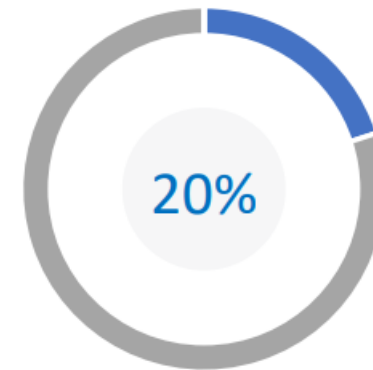
Découpage de la base de données

- **Variable prédictive:** Toutes les colonnes sauf la colonne « churn »
- **Variable prédictive:** « churn »

Données d'entraînement et de test



Données pour l'apprentissage



Données pour le test

01

Logistic
Regression

02

K-Nearest
Neighbors

03

eXtreme Gradient
Boosting
(xgboost)

04

Random
Forest

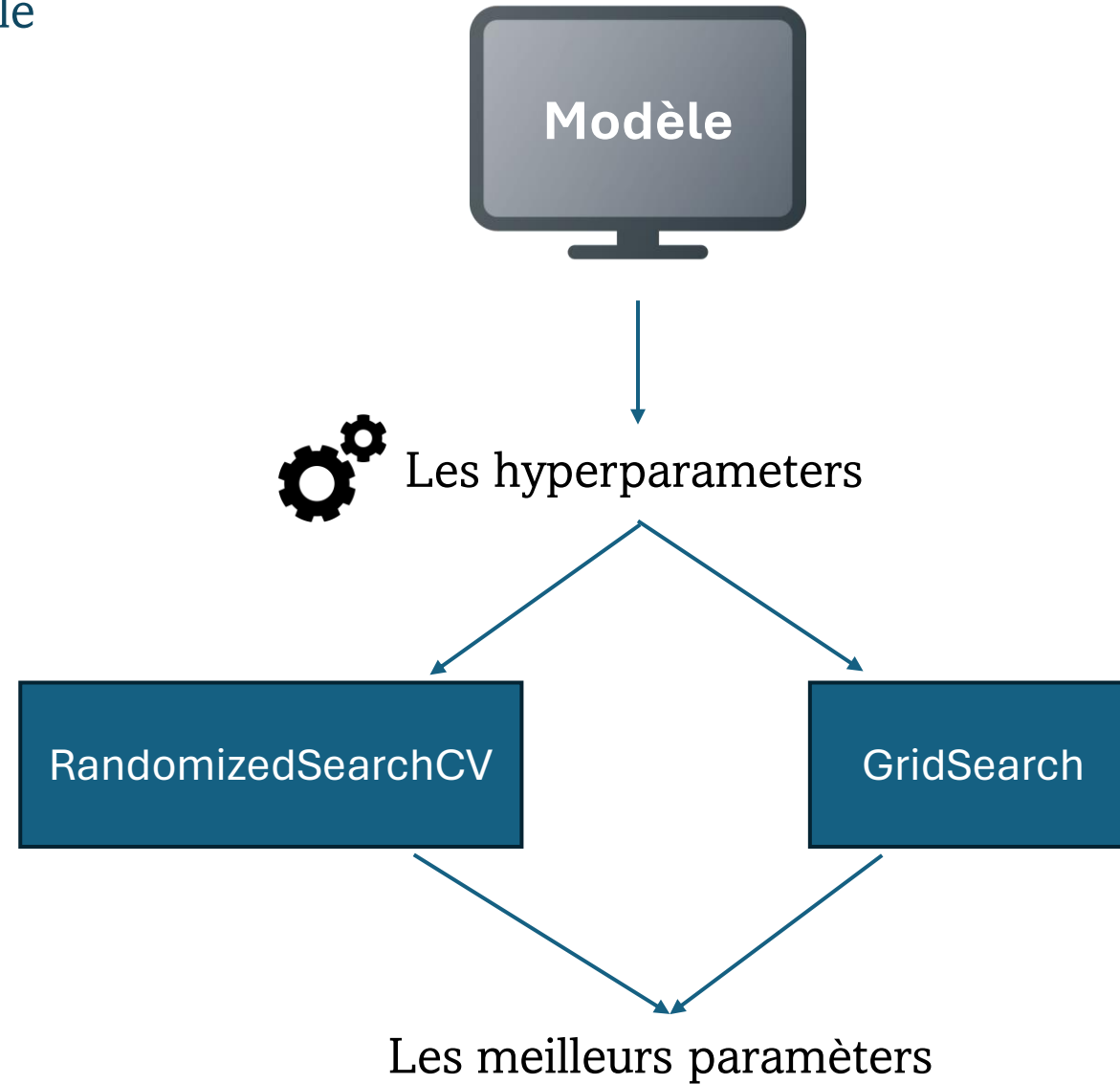
05

Decision
Tree

06





Naive
Bayes

Optimisation du modèle



Les mesures de performance

Données prédites par l'algorithme

		Non churn 	Churn 
Données réelles	Non churn 	Vrai négatif	Faux positif
	Churn 	Faux négatif	Vrai positif

Accuracy

Precision

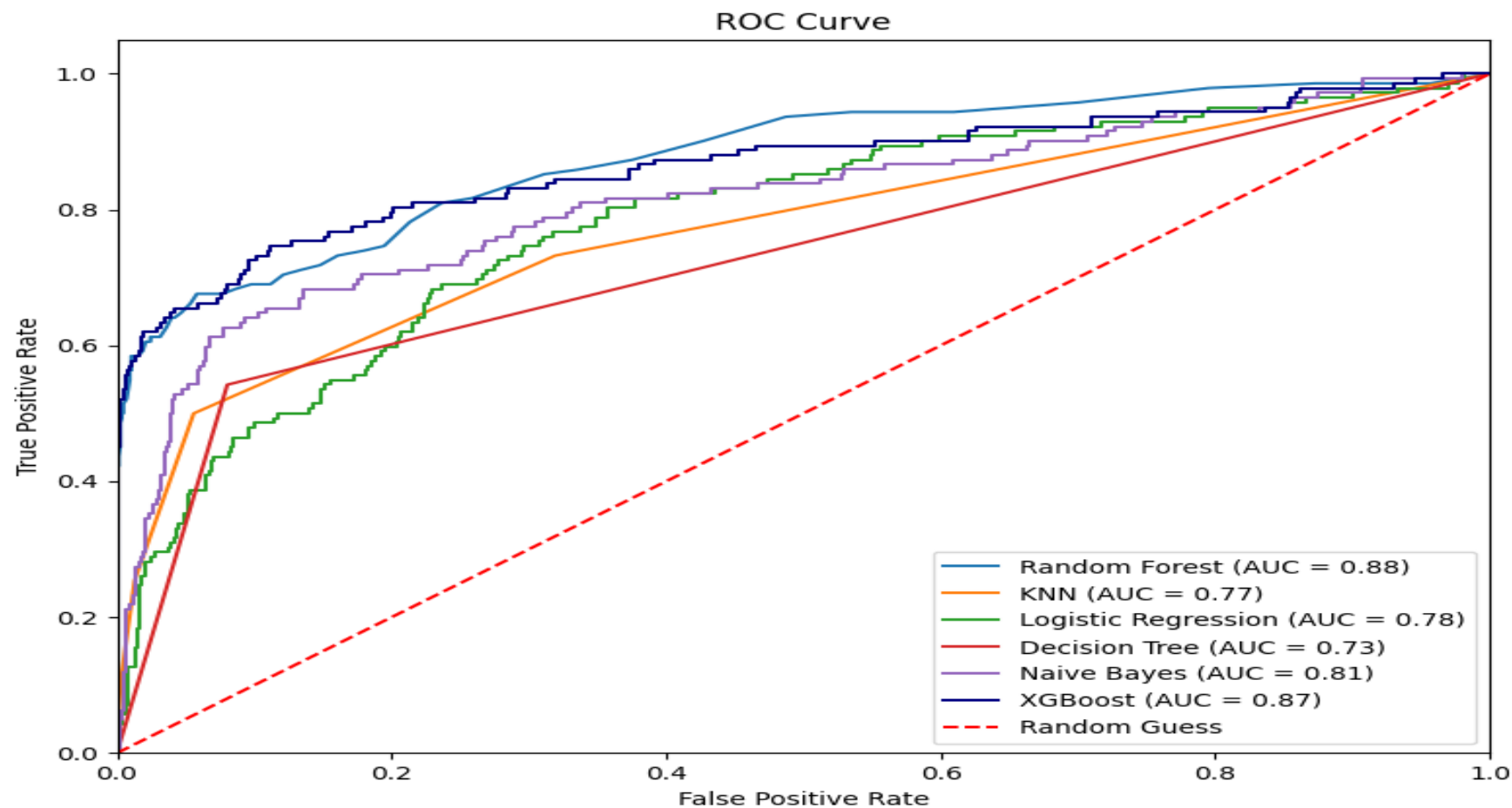
Recall

F1-Score

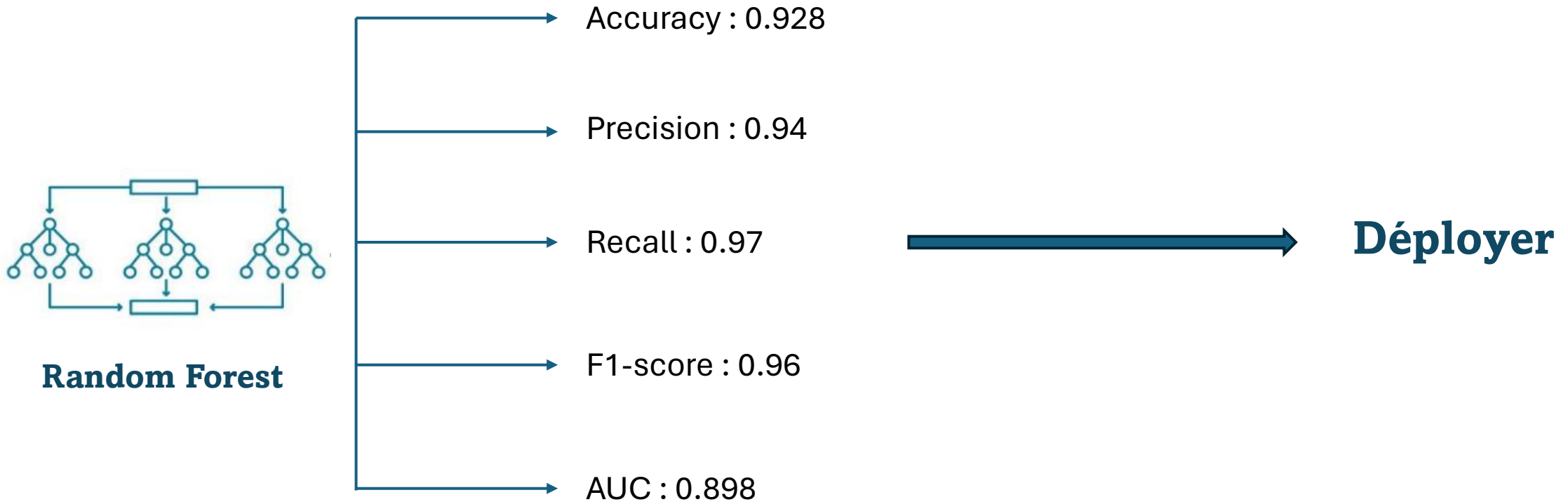
Métrique numérique

	XGBoost	Random Forest	Decision Tree	KNN	Logistic Regression	Naïve Bayes
Accuracy	0.931	0.928	0.921	0.889	0.866	0.833
Precision	0.94	0.94	0.94	0.89	0.87	0.93
Recall	0.98	0.97	0.97	0.99	0.99	0.95
F1-Score	0.95	0.96	0.94	0.94	0.93	0.94

Métrique graphique



Choix du modèle



Architecture de déploiement

Jupyter Notebook



Modèle entraîné



Sauvegarder le
modèle dans un
fichier pickle .pkl

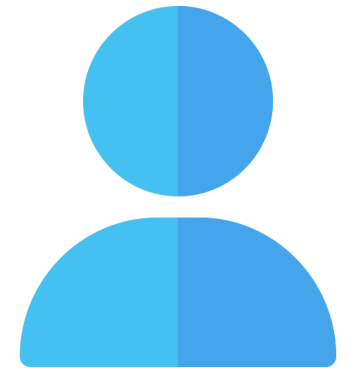
Visual Studio Code



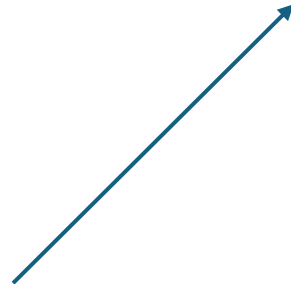
Prédiction du modèle



HTTP
REST API



Utilisateur





PRÉDICTION DU CHURN DES CLIENTS TUNISIE TELECOM

Nombre de jours d'abonnement:

118

Genre:

Femme

Nombre d'appels par jour:

82

Nombre d'appels par soirée:

70

Nombre d'appels par nuit:

73

Nombre d'appels international:

2

Activer option message vocaux:

Oui

Nombre des réclamations:

6

Âge:

26

Durée de l'appel par jour en minutes:

120,8

Durée de l'appel par soirée en minutes:

109,5

Durée de l'appel par nuit en minutes:

245,9

Durée de l'appel international en minutes:

16,9

Nombre des messages vocaux:

24

État civil:

Célibataire

Coût des appels par jour:

20,04

Coût des appels par soirée:

10,7

Coût des appels par nuit:

30,1

Coût des appels international:

8,79


Le nom d'offre tarifaire suivie:

PRE - AHL

Prédire

Résultat de prédiction: Le client résiliera son abonnement

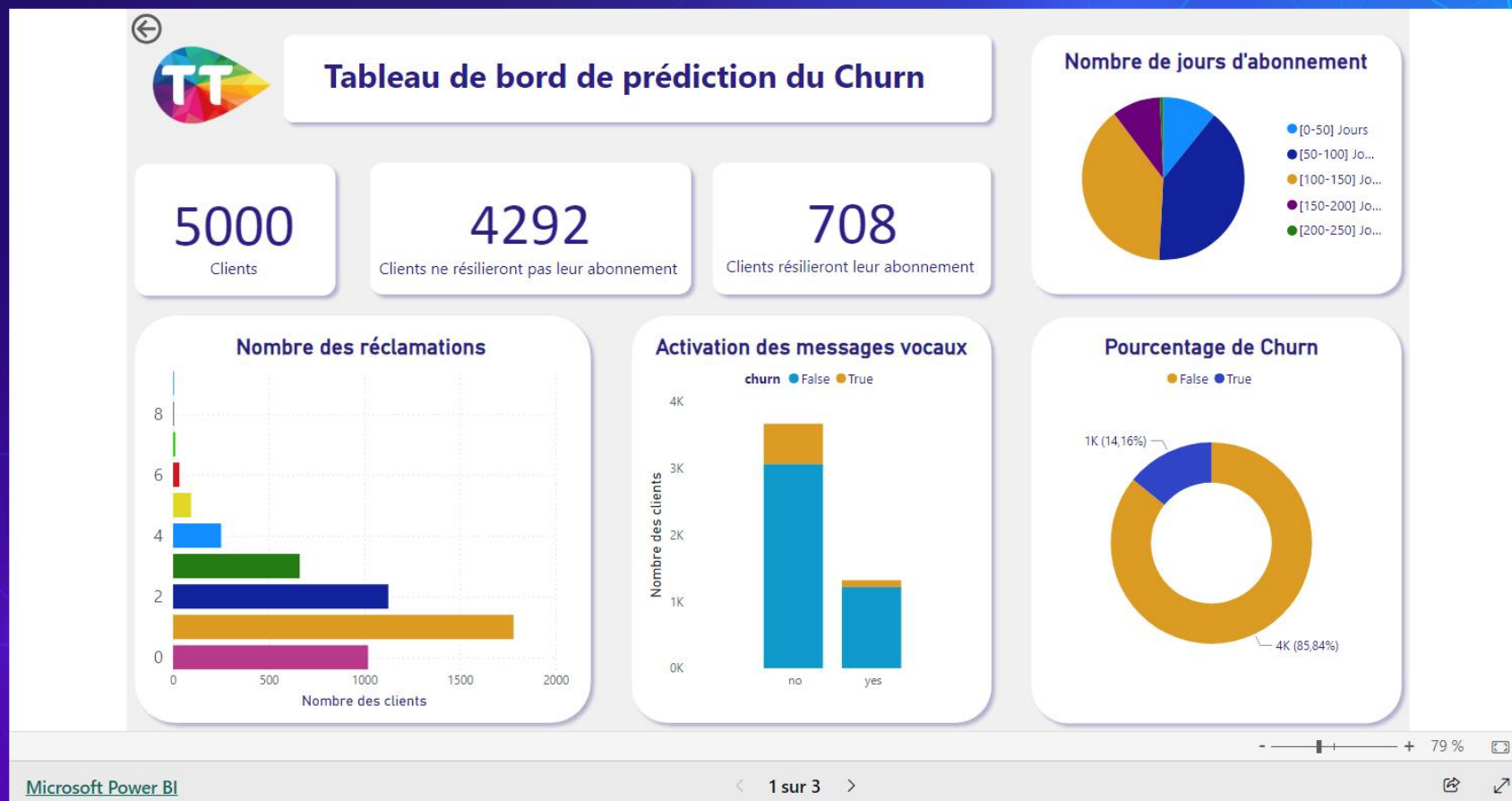
Visualisez les tableaux de bord

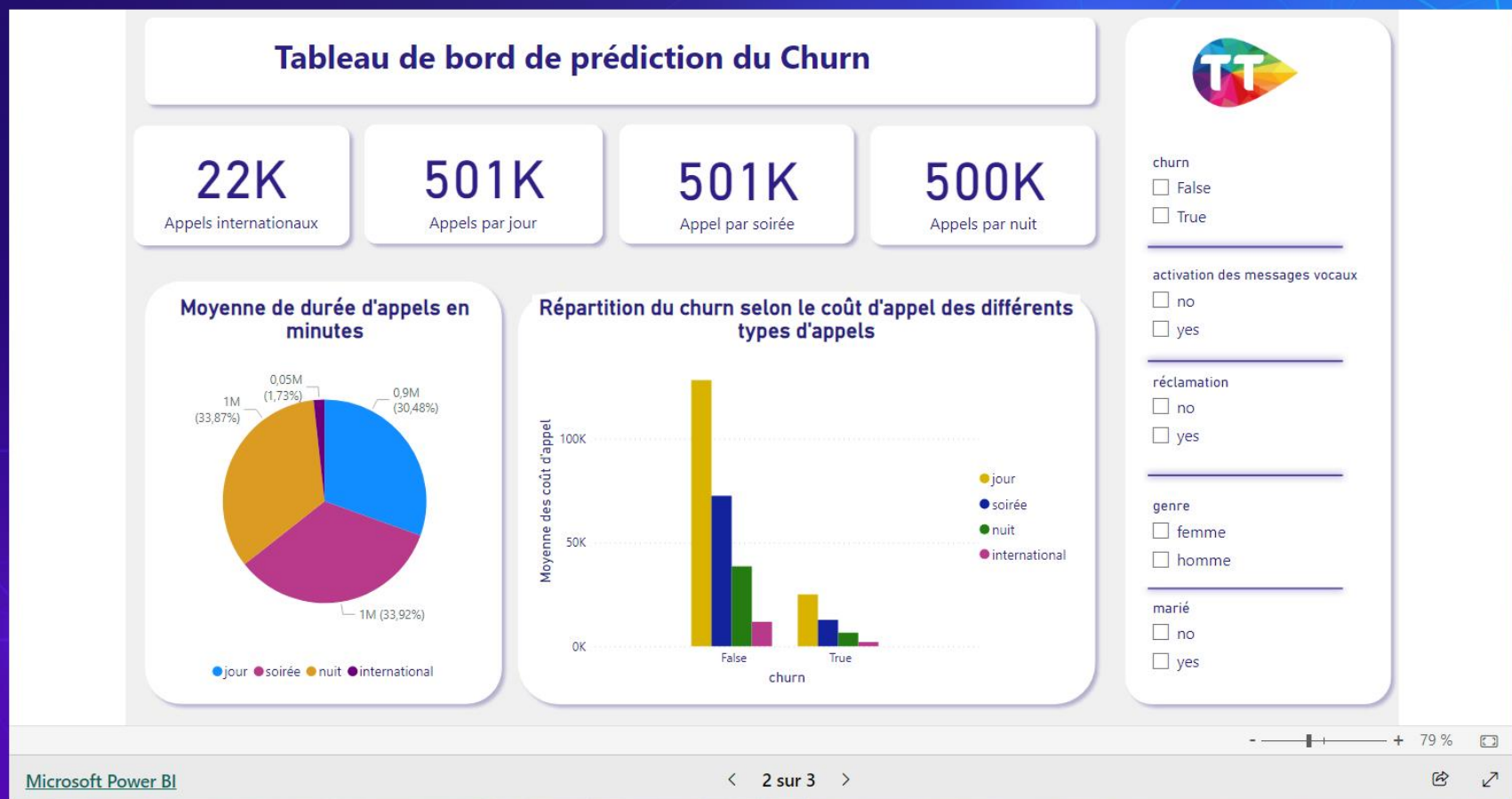


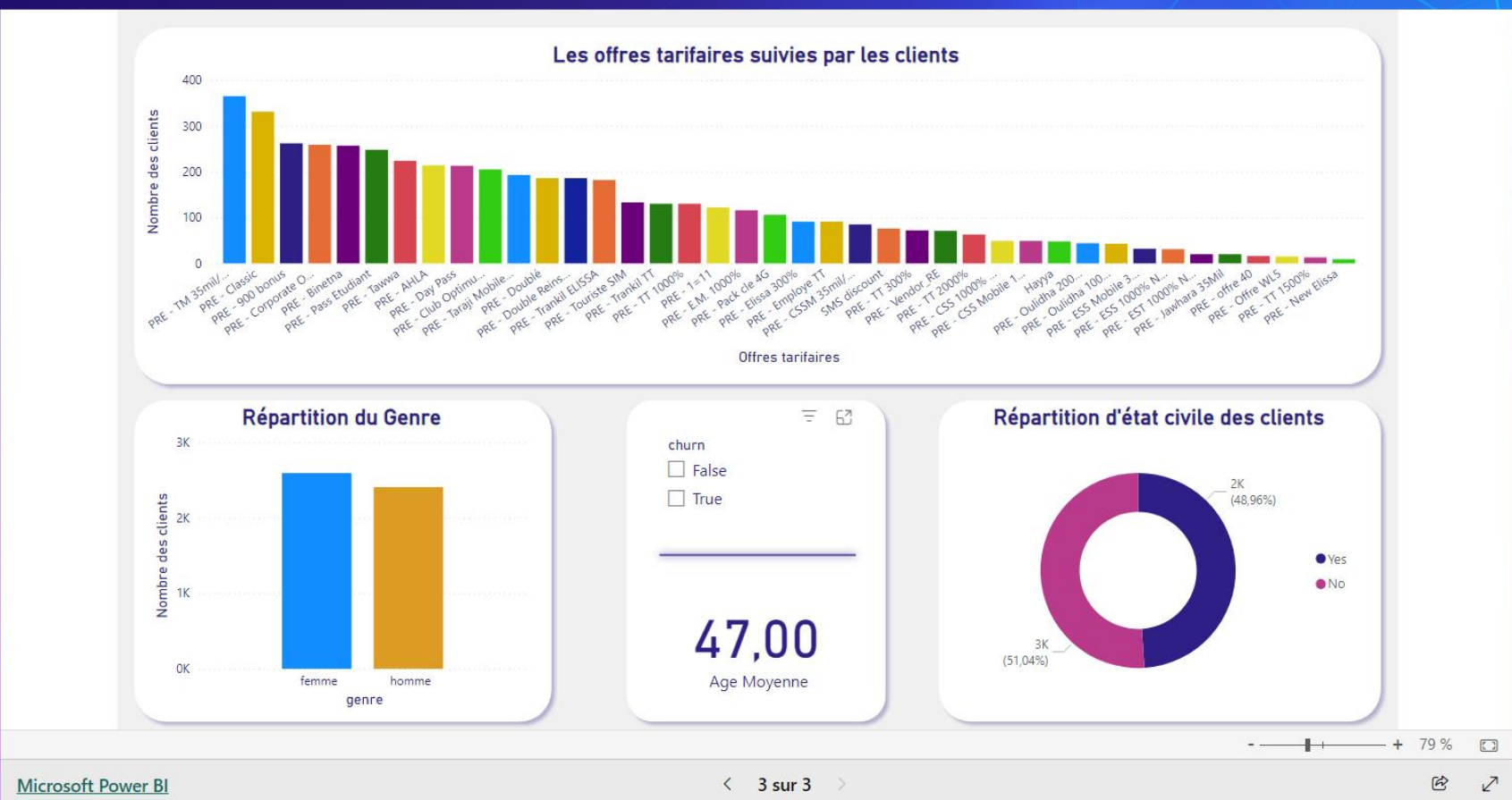
PRÉDICTION DU CHURN DES CLIENTS TUNISIE TELECOM

Nombre de jours d'abonnement:	<input type="text" value="9"/>				
Genre:	<input type="text" value="Femme"/>	Âge:	<input type="text" value="26"/>	État civil:	<input type="text" value="Célibataire"/>
Nombre d'appels par jour:	<input type="text" value="108"/>	Durée de l'appel par jour en minutes:	<input type="text" value="214,1"/>	Coût des appels par jour:	<input type="text" value="18,07"/>
Nombre d'appels par soirée:	<input type="text" value="115"/>	Durée de l'appel par soirée en minutes:	<input type="text" value="169,5"/>	Coût des appels par soirée:	<input type="text" value="14,83"/>
Nombre d'appels par nuit:	<input type="text" value="119"/>	Durée de l'appel par nuit en minutes:	<input type="text" value="189,3"/>	Coût des appels par nuit:	<input type="text" value="8,54"/>
Nombre d'appels international:	<input type="text" value="3"/>	Durée de l'appel international en minutes:	<input type="text" value="10,2"/>	Coût des appels international:	<input type="text" value="8,79"/>
Activer option message vocaux:	<input type="text" value="Oui"/>	Nombre des messages vocaux:	<input type="text" value="2"/>		
Nombre des réclamations:	<input type="text" value="1"/>			Le nom d'offre tarifaire suivie:	<input type="text" value="Hayya"/>

Résultat de prédiction: Le client ne résiliera pas son abonnement







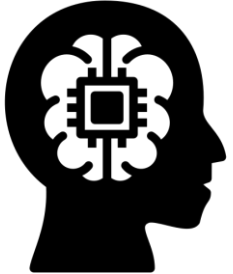


Conclusion et perspectives

Conclusion



Atteindre les objectifs fixés



Améliorer nos compétences dans le domaine de l'intelligence artificielle



Découvrir la vie professionnelle

Perspectives



Analyser les mots répétés dans les appels au service client.

Analyser les sentiments à partir des enquêtes de satisfaction, des commentaires sur les réseaux sociaux.



Merci de votre
attention