

# Projet Machine Learning

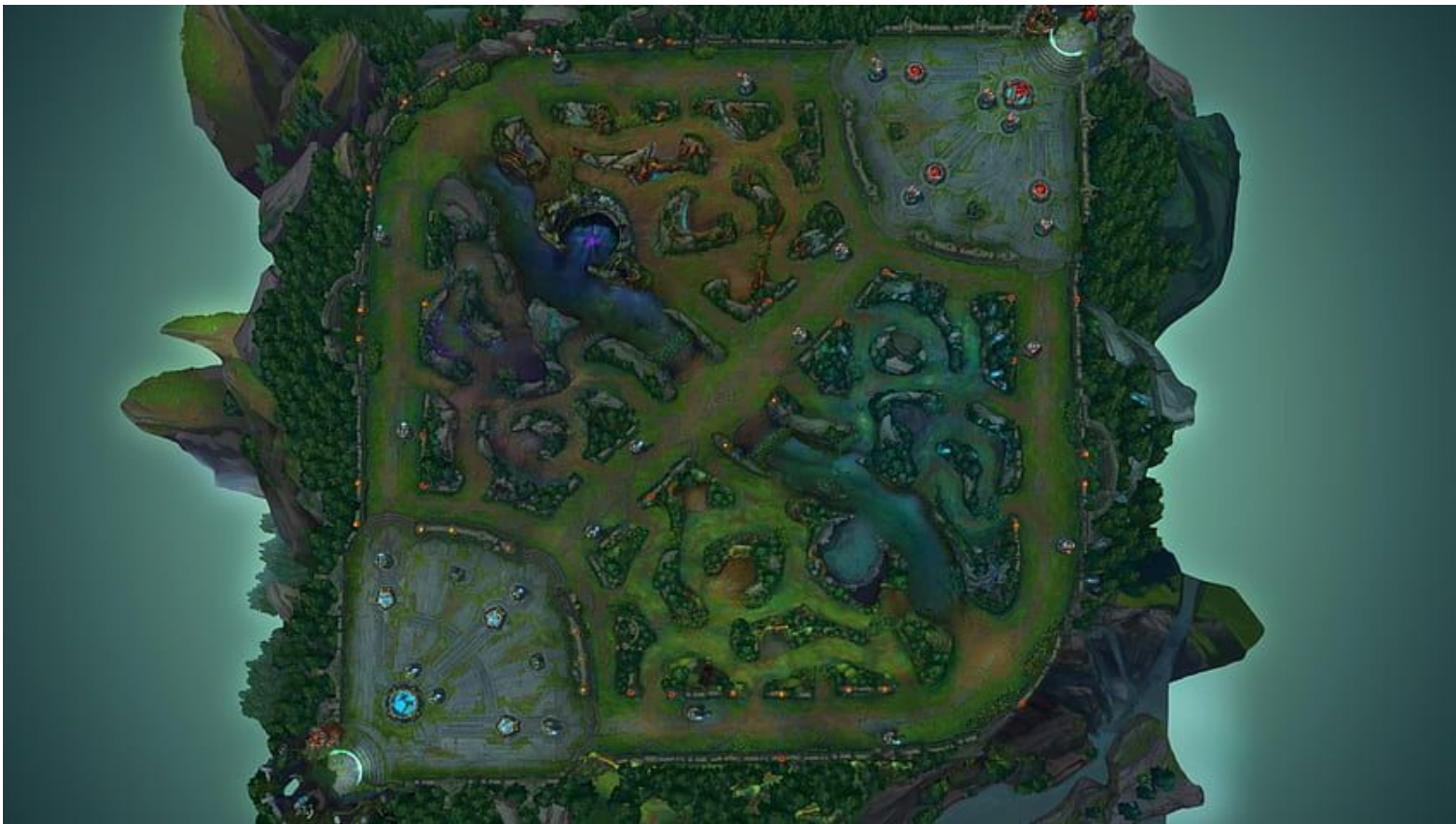


## Membre du groupe :

1. Maé Louis
2. Nadia Medjdoub
3. Maxime Eneau
4. Karim Ameur

LEAGUE OF  
LEGENDS

# Présentation du projet - Analyse de données LOL



# DataSet



**26 000 parties de  
niveau Challenger**



**65 000 parties de  
niveau GrandMaster**



**100 000 parties de  
niveau Master**

- **Caractéristiques du dataset:** 50 variables par match couvrant les performances des deux équipes
- **Données complètes** sans valeurs manquantes
- **Variables stratégiques:** ressources accumulées, objectifs sécurisés, performances individuelles  
Équilibre dans la distribution des victoires ( $\approx 50/50$ )

Ces données nous permettent d'analyser les facteurs déterminants à différents niveaux de jeu et de comparer les stratégies gagnantes selon l'expertise des joueurs.

# Préparation des données

## Liste des colonnes associées à l'équipe bleue

'blueFirstBlood', 'blueFirstTower', 'blueFirstBaron', 'blueFirstDragon',  
'blueFirstInhibitor', 'blueDragonKills', 'blueBaronKills', 'blueTowerKills',  
'blueInhibitorKills', 'blueWardPlaced', 'blueWardkills', 'blueKills',  
'blueDeath', 'blueAssist', 'blueChampionDamageDealt', 'blueTotalGold',  
'blueTotalMinionKills', 'blueTotalLevel', 'blueAvgLevel', 'blueJungleMinionKills',  
'blueKillingSpree', 'blueTotalHeal', 'blueObjectDamageDealt'

## Liste des colonnes à supprimer en raison de fortes corrélations

'blueInhibitorKills', # Fortement corrélé avec blueTowerKills (0.76)  
'blueJungleMinionKills', # Fortement corrélé avec blueTotalMinionKills (0.78)  
'blueAvgLevel', # Très corrélé avec blueTotalLevel  
'blueWardkills', # Corrélé avec blueWardPlaced  
'blueBaronKills', # Corrélé avec blueObjectDamageDealt (0.77)  
'blueFirstInhibitor' # Corrélé avec d'autres variables "First"

## # Challenger

X\_train\_challenger, X\_test\_challenger, y\_train\_challenger, y\_test\_challenger

## # GrandMaster

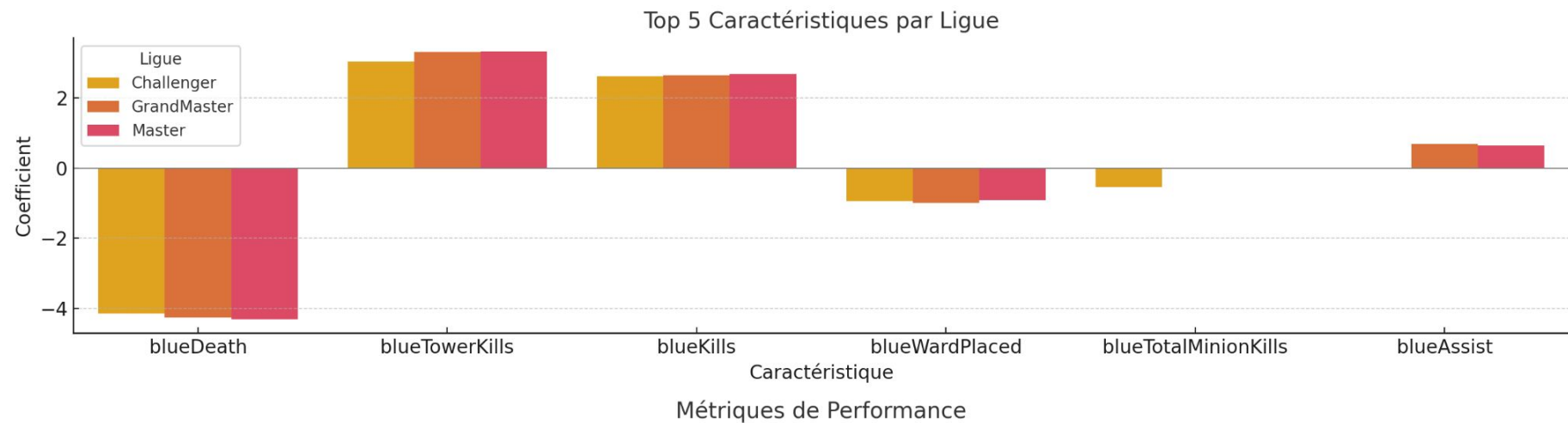
X\_train\_grandmaster, X\_test\_grandmaster, y\_train\_grandmaster, y\_test\_grandmaster

## # Master

X\_train\_master, X\_test\_master, y\_train\_master, y\_test\_master



# Régression logistique



Ligue	Accuracy	Précision	Recall	F1 Score
Challenger	0.9649	0.9661	0.9643	0.9652
GrandMaster	0.9587	0.9569	0.9602	0.9585
Master	0.9583	0.9583	0.9573	0.9578

# Pourquoi XGBoost ?

## Avantages clés :

- Gère les relations non linéaires complexes
- Robuste face aux valeurs extrêmes
- Très performant sur grands jeux de données structurés
- Prend en compte les interactions entre variables

## Notre méthode :

- Validation croisée 3 plis pour éviter le surapprentissage
- Optimisation des hyperparamètres via GridSearch
- Évaluation multi-métrique : précision, rappel, F1-score, ROC AUC
- Calibration adaptée à chaque niveau de jeu

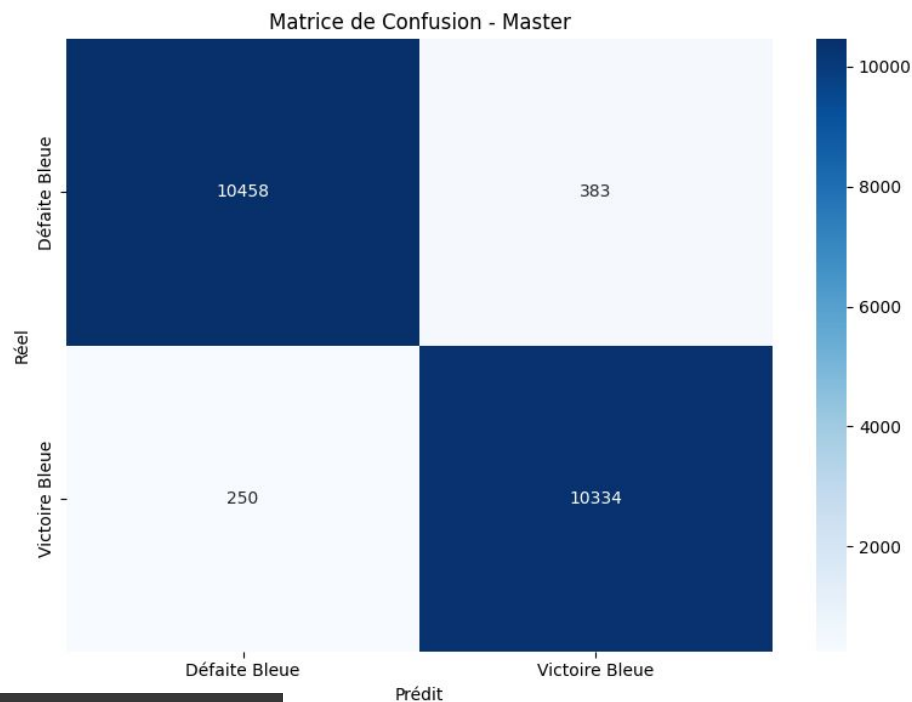
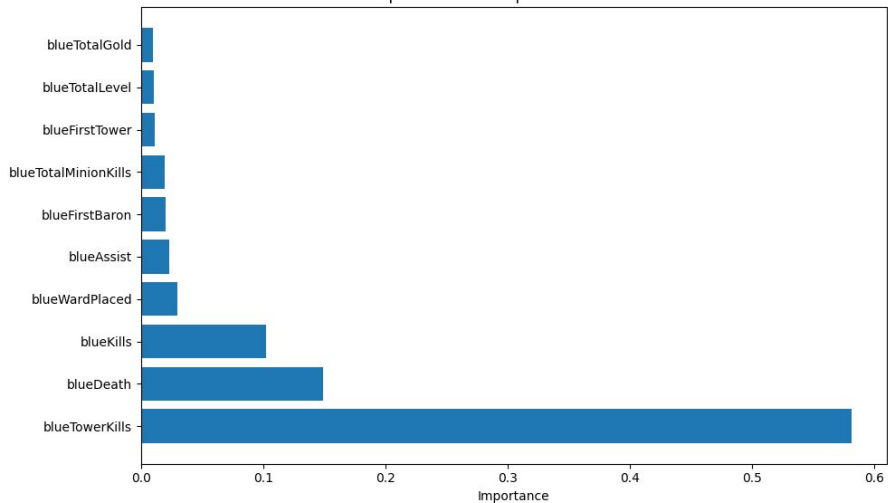
Niveau	learning_rate	max_depth	min_child_weight	n_estimators	subsample
Challenger	0.05	5	3	300	0.9
GrandMaster	0.05	5	1	300	0.8
Master	0.10	5	1	200	0.8

## Paramètres optimaux par niveau

Ces ajustements traduisent des styles de jeu différents selon le niveau, avec une plus grande prudence exigée pour modéliser les comportements des joueurs Challenger.

# XGBoost - Master

Top 10 Features Importantes - Master



==== Entraînement de XGBoost pour le niveau Master ====

Utilisation des meilleurs paramètres: {'learning\_rate': 0.1, 'max\_depth': 5, 'min\_child\_weight': 1, 'n\_estimators': 200, 'subsample': 0.8}

Accuracy avec les meilleurs paramètres: 0.9705

F1 Score avec les meilleurs paramètres: 0.9703

==== Évaluation de XGBoost pour le niveau Master ====

Accuracy: 0.9705

ROC AUC: 0.9962

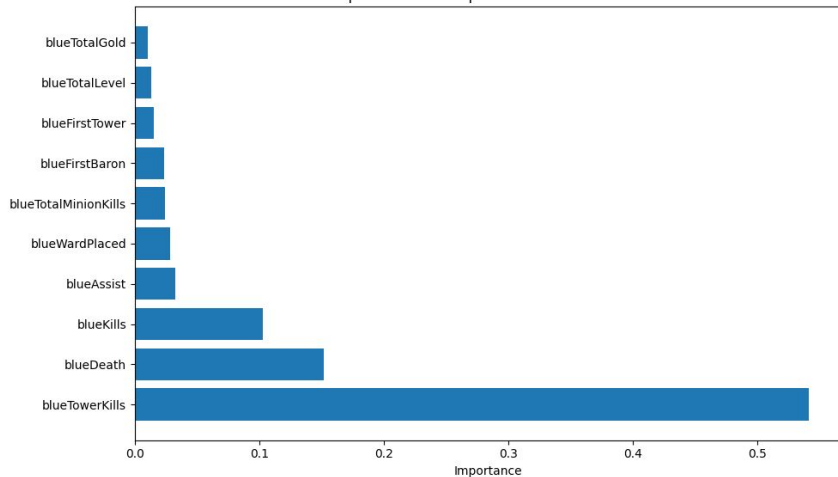
F1 Score: 0.9703

Rapport de classification:

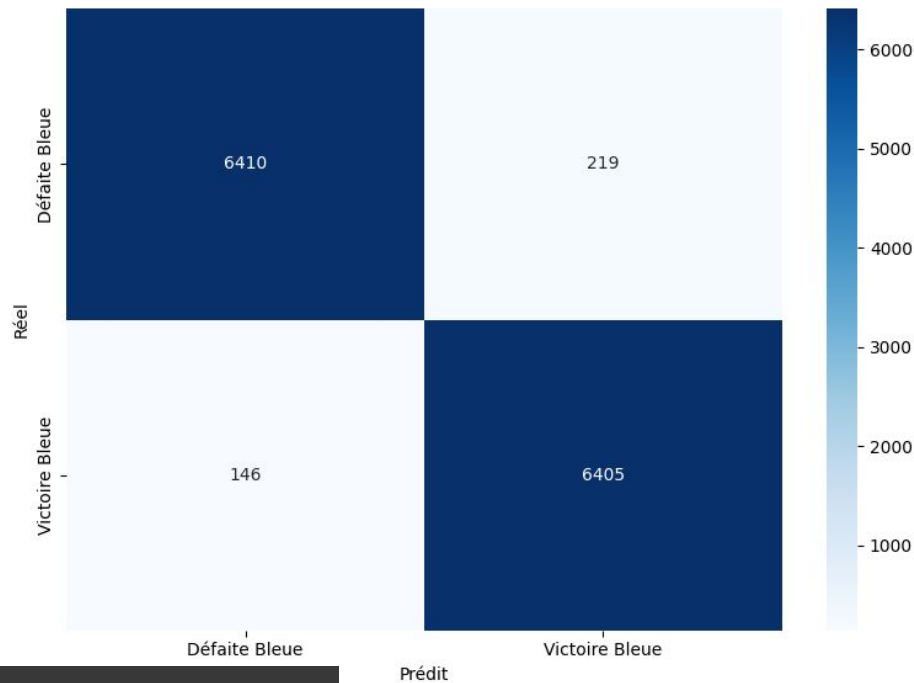
	precision	recall	f1-score	support
0	0.98	0.96	0.97	10841
1	0.96	0.98	0.97	10584
accuracy			0.97	21425
macro avg	0.97	0.97	0.97	21425
weighted avg	0.97	0.97	0.97	21425

# XGBoost - Grand Master

Top 10 Features Importantes - GrandMaster



Matrice de Confusion - GrandMaster



==== Entraînement de XGBoost pour le niveau GrandMaster ====

Utilisation des meilleurs paramètres: {'learning\_rate': 0.05, 'max\_depth': 5, 'min\_child\_weight': 1, 'n\_estimators': 300, 'subsample': 0.8}

Accuracy avec les meilleurs paramètres: 0.9723

F1 Score avec les meilleurs paramètres: 0.9723

==== Évaluation de XGBoost pour le niveau GrandMaster ====

Accuracy: 0.9723

ROC AUC: 0.9957

F1 Score: 0.9723

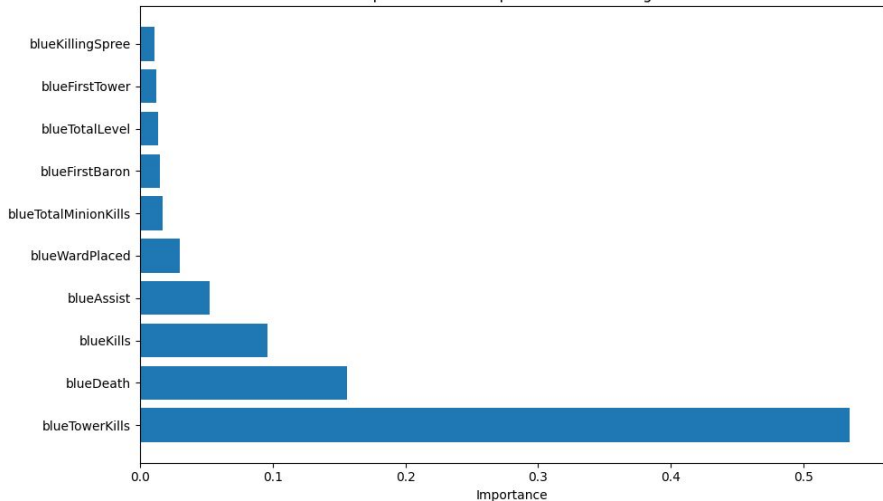
Rapport de classification:

	precision	recall	f1-score	support
0	0.98	0.97	0.97	6629
1	0.97	0.98	0.97	6551
accuracy			0.97	13180
macro avg	0.97	0.97	0.97	13180
weighted avg	0.97	0.97	0.97	13180

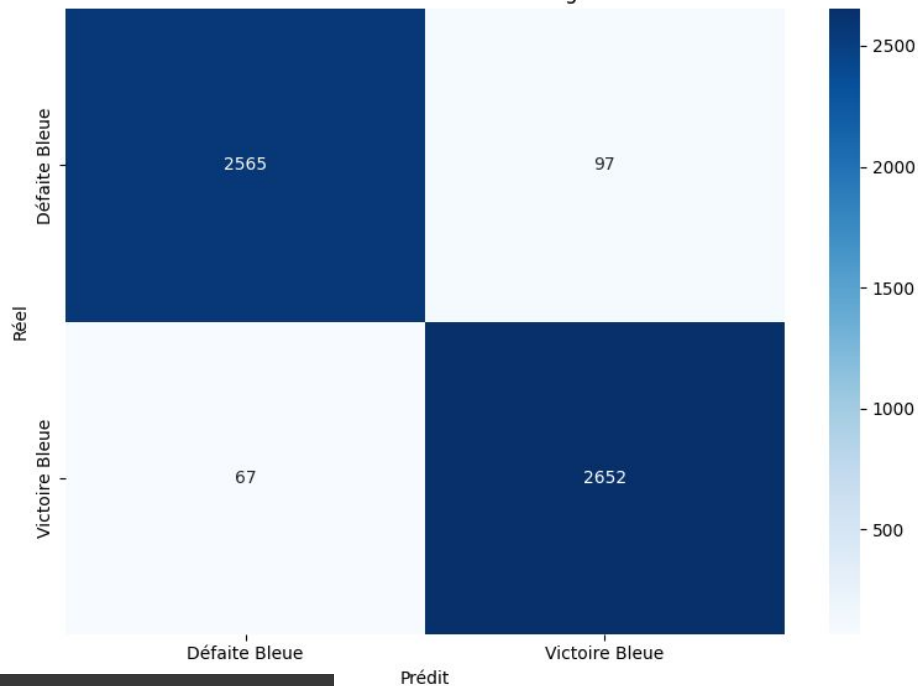


# XGBoost - Challenger

Top 10 Features Importantes - Challenger



Matrice de Confusion - Challenger



==== Entraînement de XGBoost pour le niveau Challenger ====

Utilisation des meilleurs paramètres: {'learning\_rate': 0.05, 'max\_depth': 5, 'min\_child\_weight': 3, 'n\_estimators': 300, 'subsample': 0.9}

Accuracy avec les meilleurs paramètres: 0.9695

F1 Score avec les meilleurs paramètres: 0.9700

==== Évaluation de XGBoost pour le niveau Challenger ====

Accuracy: 0.9695

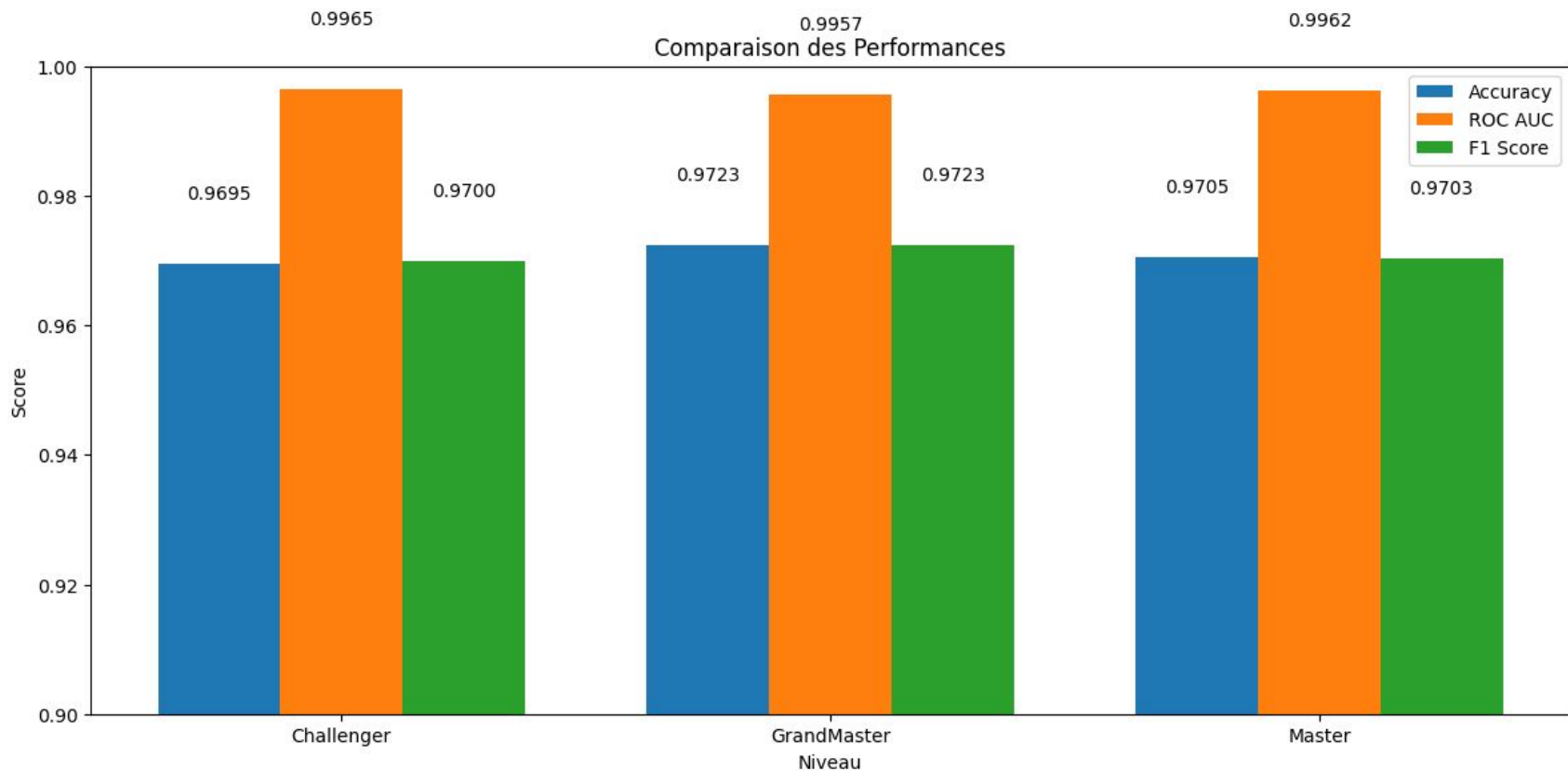
ROC AUC: 0.9965

F1 Score: 0.9700

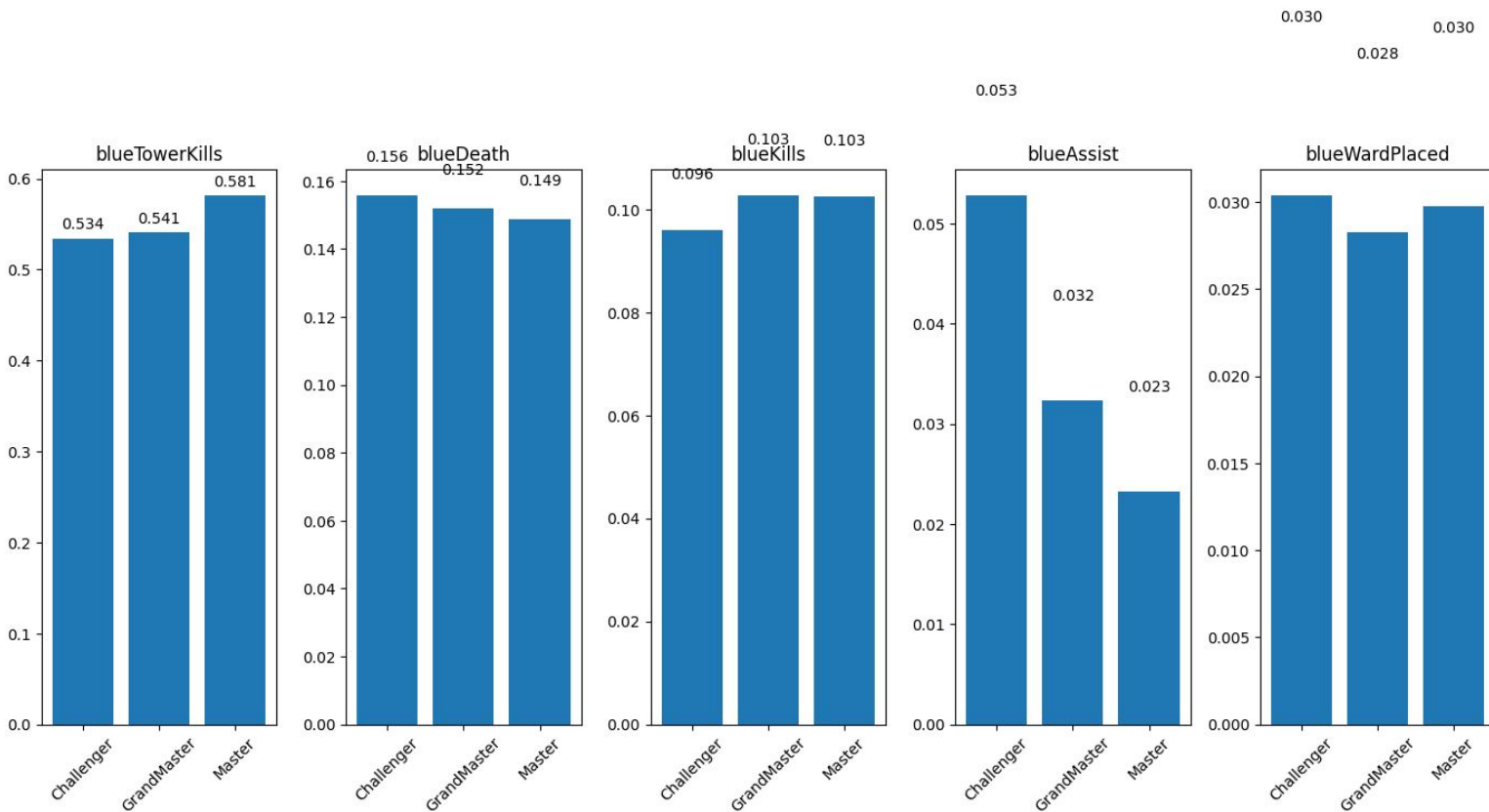
Rapport de classification:

	precision	recall	f1-score	support
0	0.97	0.96	0.97	2662
1	0.96	0.98	0.97	2719
accuracy			0.97	5381
macro avg	0.97	0.97	0.97	5381
weighted avg	0.97	0.97	0.97	5381

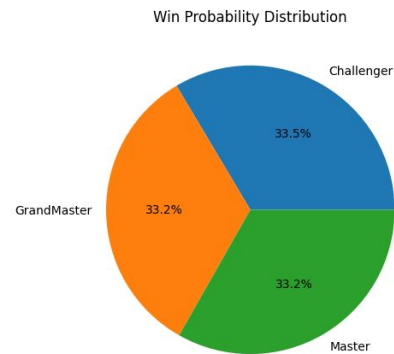
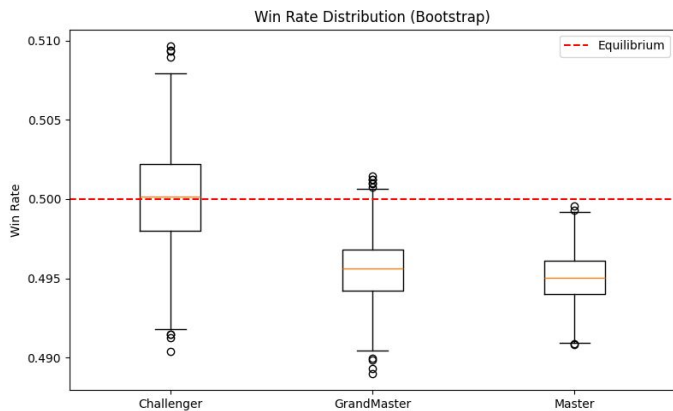
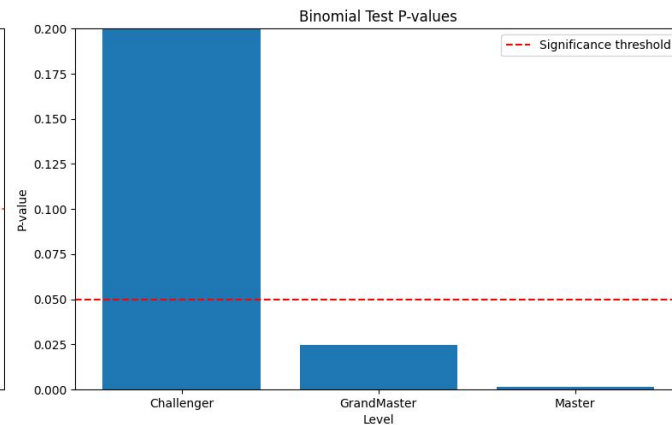
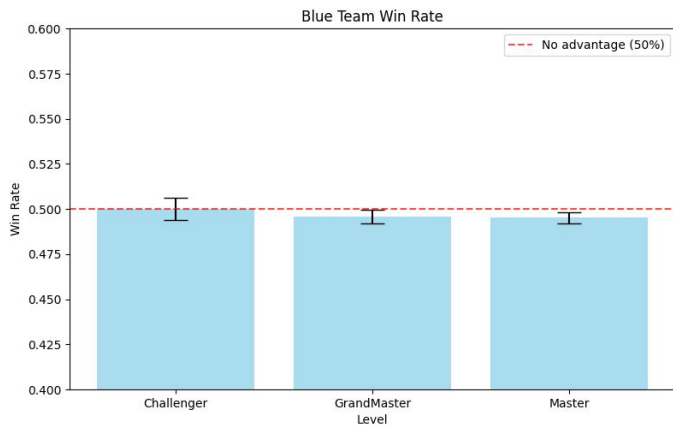
# Comparaison performances



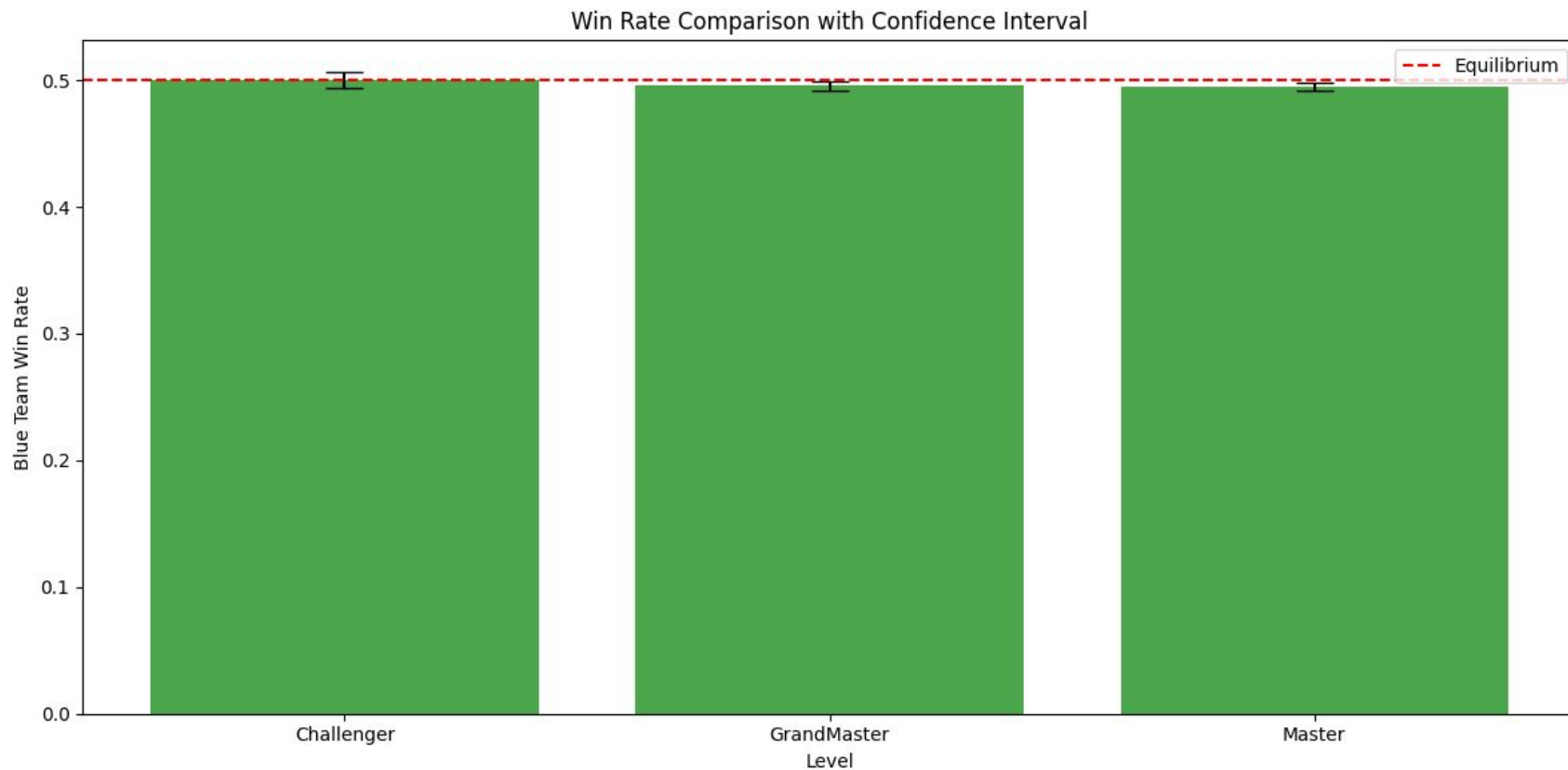
# Analyse croisée features



# Analyse croisé - Bleue - Rouge



# Analyse croisé - Bleue - Rouge



# Conclusion

- Taux de précision supérieur à 97 % sur l'ensemble des niveaux analysés
- Influence majeure des objectifs structurels (comme la destruction de tours) par rapport aux éliminations
- Classement des variables déterminantes relativement stable, quel que soit le niveau des joueurs
- Aucune tendance statistique en faveur de l'une ou l'autre des équipes, confirmant un bon équilibre du jeu

## Lectures analytiques :

- Le nombre de tours détruites (blueTowerKills) ressort comme l'indicateur le plus fiable pour prédire une victoire
- Le nombre de décès (blueDeath) a un impact négatif fort, peu importe le niveau de jeu
- Les joueurs au niveau Challenger semblent exploiter plus efficacement les avantages acquis
- L'équité entre les équipes bleue et rouge est confirmée par l'absence de biais détecté dans les résultats

## Ouvertures possibles :

- Intégrer la notion de temporalité pour identifier les moments clés d'une partie
- Étudier l'impact des compositions d'équipes et les synergies entre champions
- Réaliser une analyse en fonction des mises à jour pour observer l'impact des équilibrages sur le gameplay

Cette analyse met en lumière la pertinence des méthodes de machine learning pour décoder les dynamiques de jeu et propose des pistes concrètes pour enrichir la stratégie des joueurs comme le développement des outils d'aide à la décision.