

Rapport Séance 17/10/2018 (PE 95)

Ce qu'on a fait :

- On a mis en place une première ébauche de calendrier (https://docs.google.com/spreadsheets/d/1kXTeDiLXYAUYXsrdiHRtj3VHR-9g_AxsGaBGllvCxmQ/edit?usp=sharing). Les délais, et les différentes sous-tâches vont surement varier, c'est une première version
- On a lu le notebook plus en détail, pour relever les infos utiles pour la suite. On en retire des infos au niveau des paramètres à prendre en compte dans le modèle, un code pour scraper des données, prenant seulement l'URL en entrée
- On a amélioré le questionnaire sur les choix que font les gens lorsqu'ils parient
- On a fini le code pour notre premier modèle de prévision des résultats, et implémenté une fonction pour évaluer facilement le pourcentage de réussite des prédictions (notre modèle actuel est à environ 58% de réussite sur une saison)
Problème : On a tenté d'améliorer le code en utilisant plus de classes et de fonctions, mais les 2 versions du code ne donnent pas exactement les mêmes taux de réussite.
- On a implémenté le calcul de la distance de Finetti (tirée de cet article https://cdn.fbsbx.com/v/t59.2708-21/14690486_1087491937970617_165471620897439744_n.pdf/csam-23-297.pdf?nc_cat=111&oh=ad77a27d27163593d1c0d8f70dee4f3c&oe=5BC91ED9&dl=1)
- On a déterminé des probabilités à partir des cotes des sites de paris sportifs, et calculé la distance de Finetti entre ces probabilités et la réalité.
- Toujours à partir du document https://cdn.fbsbx.com/v/t59.2708-21/14690486_1087491937970617_165471620897439744_n.pdf/csam-23-297.pdf?nc_cat=111&oh=ad77a27d27163593d1c0d8f70dee4f3c&oe=5BC91ED9&dl=1), on a tenté d'améliorer notre modèle en prenant plus de facteurs en compte pour le calcul de lambda. On a implémenté ça en python en effectuant quelques modifications qui nous semblaient pertinentes, mais le taux de réussite obtenu est bien plus faible que celui du modèle précédent.
- On a commencé à automatiser la prévision des matchs du week-end (il reste quelques bugs à corriger, ça devrait être fonctionnel la semaine prochaine)
- On a commencé la prise en main de Git / Github (il reste quelques trucs à comprendre mais ça devrait être bon la semaine prochaine)

Objectifs pour la suite :

- On pourrait essayer de chercher le lambda qui minimise la distance de Finetti (par une méthode des moindres carrés ?).
- Continuer à travailler sur le nouveau modèle qu'on a tenté d'implémenter.
- Finir d'automatiser la prévision des matchs du week-end
- Ajuster les délais du planning, et finir de prendre en main Github
- Si possible, rendre le premier livrable (sinon, on le fera la semaine d'après les vacances)

Questions :

- Dans les objectifs du livrable 2, on avait écrit « Visualiser les rankings », qu'est-ce que cela signifie exactement ?
- Savez-vous comment trouver les dates exactes des RVP ? On a trouvé que le premier était en Décembre, et le deuxième en mars-avril, mais pas d'informations plus précises.
- Quand on regarde le planning, on voit qu'on a plusieurs livrables sur lesquels on travaille en parallèle. On aura bientôt fini les objectifs du livrable 3 par exemple. Pouvons nous changer l'ordre des livrables, et éventuellement rajouter des objectifs dans les livrables ?
- Pensez-vous que le choix du championnat que l'on étudie (Première League pour le moment) a une importance ? Serait-il plus facile de travailler sur la ligue 1 par exemple ?
- On a pensé, pour prédire des résultats sportifs, qu'on pourrait aussi utiliser des réseaux de neurones. Pensez vous que ce soit une piste intéressante à creuser ?