

**DESCRIPTION DE LA PROBLEMATIQUE DU TRAVAIL DE FIN D'ETUDES**

A déposer sur la plate-forme Moodle sous format PDF

pour le 21 mars 2025 au plus tard

**SUJET (TITRE SUGGERE) DU TFE :**

Analyse et prédiction de l'évolution des performances des nageurs en Belgique

**NOM(S), PRENOM(S) de(s) (l') étudiant(s):**

1. DEBONGNIE Nathan

2. \_\_\_\_\_

3. \_\_\_\_\_

**MOTIFS OU RAISONS POUR LESQUELS VOUS AVEZ CHOISI CE SUJET :**

Ma passion pour le sport, et plus particulièrement la natation ma amené à chercher des moyens d'améliorer les performances des nageurs et d'aider les entraîneurs à mieux connaître les nageurs. De nouveaux outils émergent afin de suivre les performances en entraînement par exemple, mais aucun outil n'est encore disponible pour analyser les résultats des compétitions. C'est à cet endroit que mes 2 passions se rencontrent. Récupérer et utiliser les résultats de compétitions de natation afin d'en retirer de l'information pertinente pour les nageurs et les clubs pourrait permettre à améliorer grandement les performances sportives.

**PRINCIPAUX OBJECTIFS POURSUIVIS ET REALISATIONS PLANIFIEES :**

**Principaux objectifs**

- Pour les clubs et entraîneurs :

- Proposer un dashboard à l'aide d'un outil BI pour présenter les performances collectives des membres du club
- Suivre l'évolution des nageurs individuellement
- Améliorer la planification des entraînements en fonction des performances

- Pour les nageurs :

- Pouvoir identifier des objectifs courts/moyen terme réalisables
- Comparer son niveau au niveau d'autres nageurs Belges au même âge

**Réalisations planifiées**

- Modélisation et création d'une base de données avec des informations globales sur les nageurs (nom, prénom, club, sexe, année de naissance, nationalité) et les résultats (nageur, distance, style, piscine (50m ou 25m), temps, points FINA, date, lieu)

- Réaliser de l'ETL pour normaliser la base de données, enlever les problèmes (données manquantes, réparation des erreurs de syntaxe, ...) et ajouter des colonnes calculées utiles (âge du nageur à une date de compétition, séparer la ville du pays ou la compétition à lieu, ...)
- Création de plusieurs rapports/applications à l'aide d'un outil de BI permettant à :
  - Un club/entraîneur de voir l'évolution des performances moyennes par compétition
  - Un club/entraîneur et un nageur de voir les informations de performances individuelles
- Création d'un modèle de prédiction de performances d'un nageur sur une course pour identifier des objectifs atteignables

#### **METHODES DE TRAVAIL UTILISEES :**

Appliquer la méthode Crisp DM semble être la meilleure approche dans le cadre de ce projet. Elle permet de structurer concrètement toutes les parties du projet étape par étape afin d'assurer que le résultat final corresponde au besoins business. Voici les 6 étapes prévues pour l'organisation du projet :

##### **1. Business Understanding**

Il s'agit de la compréhension et de la définition des objectifs du projet, en tenant compte des besoins des utilisateurs. C'est dans cette partie que les facteurs influençant la performance des nageurs doivent être identifiés pour être utilisés par la suite dans la modélisation.

##### **2. Data Understanding**

Cette étape d'analyse est nécessaire pour comprendre les données, identifier des tendances et des anomalies. Lors de cette étape, une première visualisation des résultats des nageurs par club ou par âge peut être réalisée pour comprendre plus en profondeur les données. Une première version du dashboard pourra être réalisée lors de cette étape.

##### **3. Data Preparation**

Avant de pouvoir passer à la modélisation, il faut encore s'assurer que les données soient de qualité. Il est potentiellement nécessaire de créer des nouvelles colonnes sur base des colonnes existantes, de remplacer des valeurs manquantes ou de modifier les valeurs aberrantes. Il faut ici aussi normaliser les données. Une séparation des données d'entraînement et de validation est réalisée lors de cette partie également.

##### **4. Modélisation des données**

C'est dans cette partie, une fois que les besoins sont clairs, que les données ont été identifiées et nettoyées, que l'on peut appliquer une ou plusieurs techniques statistiques afin de prédire les performances des nageurs.

##### **5. Évaluation**

L'étape suivante est l'évaluation. Il faut ici vérifier que les objectifs définis au début du projet ont été respectés. Il faut que le résultat soit exploitable pour les cibles du projet.

##### **6. Déploiement**

Dernière étape pour présenter les résultats. Il s'agit ici de mettre à jour le dashboard en intégrant les résultats du modèle de prédiction et de partager ce résultat au client. Dans le cadre de ce TFE, le rapport final fait également partie du déploiement étant donné que c'est une partie du projet final attendu.

DATE : 21 mars 2025