



La souris au bois dormant

28.09.2023

Tim Ernst, Justin Rausis, Grégory Rey-mermet & Florian Conti

Contexte

Le sommeil est un phénomène universel observé chez tous les êtres vivants, qu'il s'agisse d'hydres, de drosophiles, de souris ou d'êtres humains. De nombreuses recherches ont démontré que le sommeil ne se résume pas à une simple absence d'activité, mais qu'il s'agit d'un état comportemental activement régulé, crucial pour les performances optimales à l'état de veille, ainsi que pour l'apprentissage et la mémoire. Bien que nous puissions facilement reconnaître quand quelqu'un est endormi (ou non), il peut être difficile de déterminer si une personne fait semblant de dormir et, lorsqu'elle est effectivement endormie, dans quel état de sommeil elle se trouve : sommeil à mouvements oculaires rapides REM (Rapid Eye Movement) ou sommeil non-REM (NREM). Seule l'enregistrement de l'activité cérébrale à partir du cuir chevelu ou directement depuis la surface corticale du cerveau (c'est-à-dire l'électroencéphalogramme ou EEG) permet de déterminer de manière indiscutable si une personne ou une souris est éveillée, en sommeil NREM ou en sommeil REM.

Objectifs

Dans ce projet, le but principal sera de détecter si la souris est en train de dormir REM ou NREM ou si elle est réveillée, cela nous permettra de mieux distinguer les phases du sommeil et quels sont les éléments qui permettent de les distinguer au mieux. Nous avons pour cela à notre disposition différentes souches de souris en fonction d'un EEG et d'un EMG (Electromyographie).

La liste des objectifs fondamentaux pour notre analyse sont:

1. Développer un modèle capable de classer correctement les trois états de sommeil-réveil sur une seule souris à la fois.
2. Comprendre les éléments qui permettent de classer les phases de cette souris.
3. Généraliser le modèle pour être capable de classer correctement les phases de toutes les souris d'une même souche.
4. Assurer une généralisation satisfaisante entre les souris de la même souche.

Des objectif non fondamentaux pour notre projet peuvent être:

1. Analyser les effets de la privation de sommeil sur la classification.
2. Généraliser le modèle entre différentes souches de souris.



Périmètre

Le modèle se limite aux EEG et EMG des souris dont les souches nous ont été fournies. Le projet se limite à ses différentes souches et ne permettra donc pas de faire des prédictions sur d'autres souches de souris ou d'autres animaux.

Description de l'existant

Nous avons en notre possession un dataset fournis par Prof. Paul Franken. Ce dataset a été construit à l'aide 4 jours d'enregistrement EEG (Électroencéphalogramme) et EMG (Electromyogramme) sur 250 souris. Le signal a été enregistré à une fréquence de 200 Hz. L'état de sommeil des souris a été relevé uniquement le 3ème jour par un humain. Les jours 1, 2 et 4 ont eux été remplis par un modèle, ce n'est donc pas très pertinent de les utiliser afin de réaliser notre modèle nous allons donc nous baser principalement sur le 3ème jour.

Lors du 3ème jour, les souris ont été interdites de sommeil les 6 premières heures.

Notre dataset contient 86'400 buckets de 4 secondes. Sur chaque bucket, on a 400 bin générer grâce à une transformation de Fourier.

Pour chaque bucket nous avons en plus:

- l'état de sommeil de la souris (W => Awake, N => NREM, R => REM).
- La variance de l'EEG et EMG
- La température

Expression des besoins

I. Besoins fonctionnels

- Faire une analyse du dataset afin de:
 - Traiter les valeurs manquantes s'il y en a
 - Détecter et traiter les valeurs aberrantes
 - Détecter des différences entre les souches des souris
 - Voir si toutes les colonnes sont pertinentes
- Fournir un modèle de prédiction pour classifier l'état d'une souris en 2 catégories (endormie ou éveillée).
- Fournir un modèle de prédiction pour classifier l'état d'une souris endormie en 2 catégories (REM, NREM).
- Fournir un modèle de prédiction pour classifier l'état d'une souris endormie en 3 catégories (REM, NREM ou éveillée).
- Comparer le modèle hiérarchique du modèle qui sépare en 3 classes
- Généraliser ces modèles pour une souche de souris.
- Développer des outils d'analyses pour afficher les résultats de nos modèles
 - Utiliser le coefficient de cohen's kappa
 - Affichage du f1-score pour chaque classe
 - Affichage de graphe pour l'erreur et la précision pour test et entraînement
 - Affichage de matrices de confusion
- Comprendre le résultat du modèle qui soit bon ou mauvais
- Définir une limite de précision atteignable à l'aide de nos modèles

II. Besoins non fonctionnels

- Le modèle devrait être conçu de manière modulaire et bien documentée afin que des modifications ou des améliorations puissent être apportées facilement à l'avenir.
- Faire en sorte que les graphes utilisés pour l'analyse soit facilement compréhensible.

Déroulement du projet

Planification

Nous utiliserons une méthodologie Agile pour notre planification avec des sprint d'une durée de 2 semaines avec une réunion réalisée durant la période dédiée au projet le jeudi.

Les tâches principales du projet sont les suivantes et les dates limites sont définies dans un diagramme de Gantt en annexe.

- Planification du projet
- Preprocessing
- Présentation intermédiaire
- Développement des modèles
- Evaluation, documentation et rédaction
- Présentation finale

La planification Gantt ainsi que le tableau du résumé des heures sont disponibles via le lien ci-contre : [x JDT_GML_Mouse.xlsx](#) .

Plan d'assurance qualité

- Les réunions de sprint permettront de vérifier que les objectifs sont atteints et de définir correctement quels seront les objectifs pour le sprint suivant.
- Les pull request seront vérifiées par un autre membre du groupe afin de vérifier que le code soit fonctionnel.

Documentation

- Description détaillée des méthodes utilisées.
- Description détaillée des résultats et leurs interprétations
- Rapport final présentant les résultats et conclusions.