



Réunion Scrum Masters : Sprint S1

Tâches

- Recherche générale sur les 5 maladies : Description de la patho, facteurs de risque etc.
- Etude des variables explicatives
- Etat de l'art sur les modèles de ML déjà réalisés sur ces 5 maladies
- Synthèse dans un document

Organisation/Répartition tâches de l'équipe

- Suivi des tâches à réaliser sur Trello
- Mise en place d'un google Drive pour partage des ressources / docs en un seul lieu
- Répartition des 3 maladies avec le plus de variables (1 par personne)
- En fonction du temps restant, dispatch des maladies restantes + mise en place du doc Notion
- Réunion d'avancement quotidienne

- Milieu de semaine : alignement spécifique sur la forme à donner à la synthèse = framework

Outils utilisés

- Trello pour l'organisation
- Notion pour la synthèse de nos recherches
- Pour les recherches: Perplexity et Google Scholar nous ont permis d'accéder aux ressources scientifiques
- Discussions avec un médecin généraliste pour les questions médicales

Difficultés rencontrées

- Compréhension du jargon médical ⇒ s'améliore rapidement
- Appréhender et décrypter un article scientifique pour juger s'il est une bonne ressource pour nous ou non
- Trouver LA bonne ressource
- Synthétiser et vulgariser
- Compréhension du jargon de machine learning
- Gérer le temps (et le niveau de précision à donner au document final)

Résultats

- Doc Notion avec une page par pathologie
- Pour chaque pathologie: description / facteurs de risques / variables (incl desc / lien avec la pathologie, normes et type d'examen requis) / état de l'art du machine learning / sources

Compte-rendu

- Possibilité de mettre les ressources sur Trello.
- Alternative à Trello : Airtable

- Concernant ce qu'on trouve sur le net : considérer que c'est faux jusqu'à ce qu'on ait la preuve que c'est vrai.
- Zotero : outil de documentation scientifique pour stocker les connaissances et référencer les articles.
- Valeurs seuil : à nous de voir si on veut les utiliser ou non (et donc les trouver ou non).
- Données INRAE = données utilisées pour le test (pour maladie rénale). Pour les autres maladies, à nous d'établir la stratégie pour train et test.
- RGPD : géré pour les data sets qu'on nous donne. Réflexion à avoir de notre côté en vue de l'application web.
- Possibilité dans Perplexity de demander qu'il nous sorte juste le paragraphe qui répond à notre question dans les articles scientifiques.
- Interactions entre les maladies : voir si les données nous le permettent.
- Les maladies cardiaques : à cibler ? On a regroupé des patients qui ont des maladies cardiaques + cohorte saine. Pas de cible exacte. Risque d'avoir UNE maladie cardiaque.
- Christophe a travaillé sur des maladies cardiaques en tant que data scientist.
- garder les liens + + + +

- ☐ envoie les data sets ce soir
- ☐ invitation à partager entre nous sur le channel
- ☐ stages dans des instituts médicaux (Paris) avec au moins un partenaire présent lors des démos pour nous donner son feedback
- ☐ semaine prochaine: analyse descriptive + analyse graphique.0