

Compte rendu 2 : Planification et gestion de projet :

Introduction

Notre projet constitue pour tous ses membres un investissement. Il est donc important pour sa bonne organisation et son bon déroulement que la gestion de projet soit correctement effectuée. C'est pour cela que nous avons tout mis en œuvre pour garder une bonne cohésion de groupe et mener à bien ce projet tout au long du semestre.

Planification et gestion de projet pour le premier semestre

Nous avons élu un Team Lead pour coordonner les efforts des membres de l'équipe, nous aider à répartir les différentes tâches et activités, organiser les rôles au sein de l'équipe, fluidifier la circulation d'information entre les membres, faire un suivi de l'avancement des sous-tâches, identifier les difficultés rencontrées et motiver les troupes. Nous avons choisi de façon unanime Anamé pour ce rôle.

Au début du projet, la première chose que nous avons faite a été de réaliser un diagramme de Gantt. Pour l'élaboration de ce diagramme, nous avons utilisé Excel car nous trouvions ce logiciel plus pratique et plus esthétique. Ce diagramme nous a permis de décrire les étapes importantes et les dates limites auxquelles elles doivent être terminées, en tenant compte de la difficulté de chacune d'entre elles. Nous avons essayé de répartir les tâches sur les semaines de cours en prévoyant de les poursuivre en semaine d'alternance si nécessaire.

Chaque couleur regroupe les étapes dépendantes les unes des autres. Par exemple, le nettoyage des données se fera une fois l'étape "Parcourir les données et visualisations simple de données" réalisée.

Semaines en entreprises																																				
Septembre				Octobre								Novembre				Décembre																				
13/09 - 19/09	20/09 - 26/09	27/09 - 03/10	04/10 - 10/10	11/10 - 17/10	18/10 - 24/10	25/10 - 31/10	01/11 - 07/11	08/11 - 14/11	15/11 - 21/11	22/11 - 28/11	29/11 - 05/12	06/12 - 12/12	13/12 - 19/12																							
Semaine 2	Semaine 3	Semaine 4	Semaine 1	Semaine 2	Semaine 3	Semaine 4	Semaine 1	Semaine 2	Semaine 3	Semaine 4	Semaine 1	Semaine 2	Semaine 3																							
Création et connexion GitHub																																				
Mise en place des outils de colab (Trello, Discord, Whatsapp)																																				
	Bibliographie Résumé			Bibliographie Résumé																																
	Parcourir les données et Visualisations simples de données			Nettoyage des données			Conception et test des différents modèles de prédiction																													
							Site web simple (avec présentation projet + graphique)																													
							PPT + RENDUS					Préparation Oral		Anglais																						

Ces étapes ont été rapportées dans un tableau avec les personnes assignées à chacune des ces tâches. Cette méthode nous a permis de nous répartir équitablement le travail et de savoir clairement ce que chacun devait faire.







	Tâches	Prénoms
Web et Visualisation	Sites web	Laura
	Recherche et création de visualisation	Célia & Jean
Analyse	Nettoyage des données	Célia, Jean & Lisa
	Exploration des données avec Iramuteq	Célia
	Exploration des données avec Python	Anamé, Laura & Lisa
Accompagnement du projet	Gestion de projet	Anamé, Lisa & Matéo
Mise en place du projet	Lecture des articles	Tous
	État de l'art	Matéo & Jean
	Prise en note des réunions	Anamé & Lisa
	Récupération des données	Tous
Classification	Recherche potentielle de classifieurs	Célia, Laura, Anamé & Matéo
	Création de classification sur les données	Laura & Anamé
Finalisation du semestre	Rédaction comptes rendus	Tous
	Préparation pour l'oral	Tous

Pour les articles qui ont été mis à notre disposition, nous avons réparti les différents textes en fonction de notre niveau de compréhension de texte en anglais. Ces documents nous ont permis d'acquérir des connaissances sur le sujet. Ils ont été mises en commun à la fin de nos lectures, en écrivant chacun des résumés de textes en

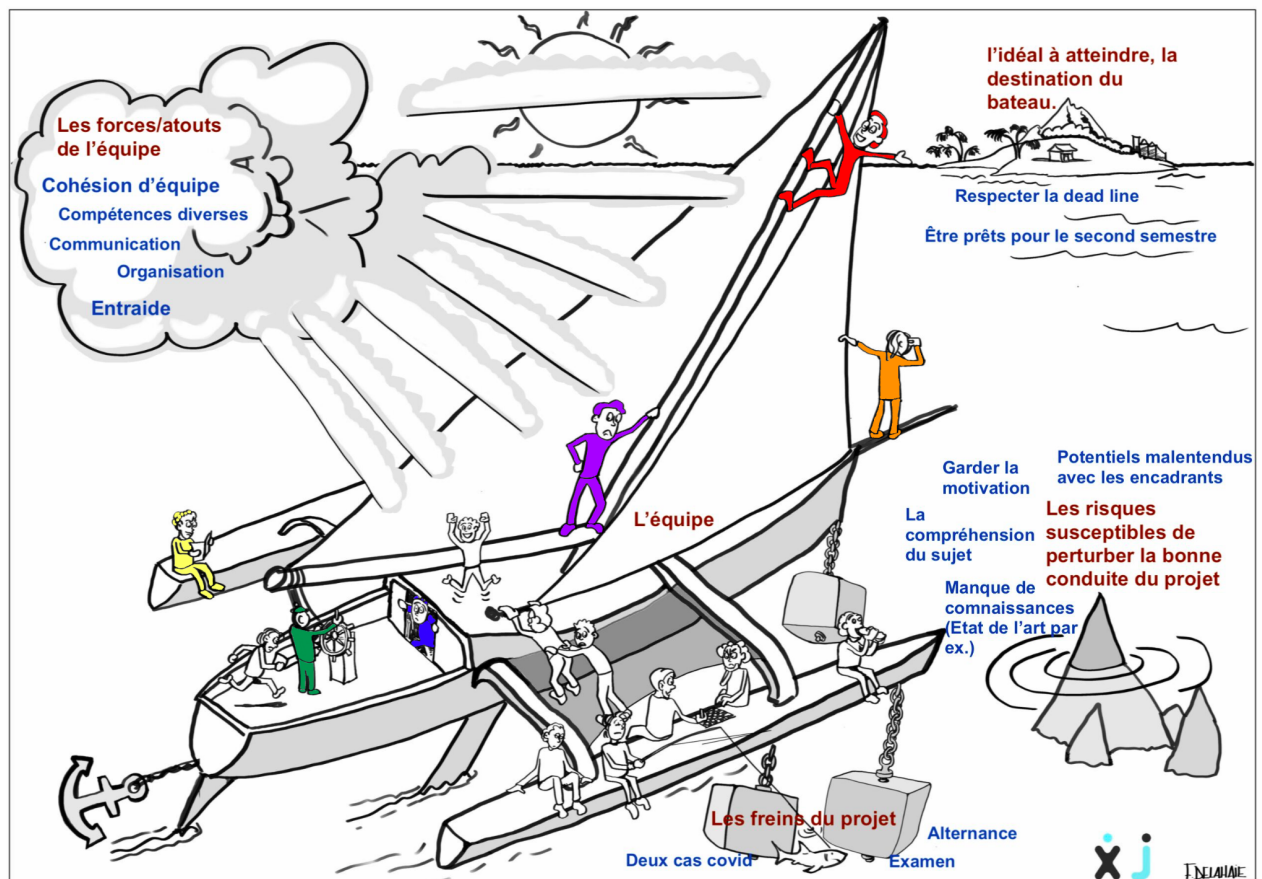
français, pour que tous les membres puissent prendre connaissance des informations contenues dans ces articles.

Articles	Prénoms
Detection of Suicide Ideation in Social Media Forums Using Deep Learning -- <i>Michael Mesfin Tadesse, Hongfei Lin, Bo Xu, Liang Yang</i>	Célia TEYSSIER & Matéo CALSACY
The integrated motivational – volitional model of suicidal behaviour -- <i>Rory C. O'Connor, Olivia J. Kirtley</i>	Jean CHABANOL & Anamé ROUMY
Unveiling Online Suicide Behavior: What Can We Learn About Mental Health from Suicide Survivors of Reddit? -- <i>Ashwin Karthik Ambalavan, Bilel Moulahi, Jérôme Azé, Sandra Bringay</i>	Laura SÉNÉCAILLE
The Interpersonal Theory of Suicide -- <i>Van Orden et al.</i>	Lisa BÉTEILLE

En début de projet, nous nous sommes concertés pour réaliser une fiche de qualification de projet qui résume le sujet. Cela nous a permis d'avoir un visuel de toutes les informations mises à notre disposition pour nous permettre de mieux nous organiser et nous assurer que tous les membres aient bien compris le sujet

<p> CONTEXTE & PRÉSENTATION DU PROJET</p> <p>Pitch du projet : Actuellement, il n'existe concrètement aucun suivi psychologique fiable des personnes ayant des pensées suicidaires. Comment savoir si une personne est au bord du suicide si elle ne vient pas d'elle-même en témoigner ? C'est pourquoi nous proposons de créer un algorithme de prédiction des TS via des posts Twitter et Reddit. Ainsi, dès que notre algorithme identifiera un cas sensible, il pourra être pris en charge et accompagné.</p> <p>Proposition de valeur : Un algorithme qui prédit si un corpus de texte est suicidaire ou non</p>	<p> L'ÉQUIPE</p> <p>Équipe Projet : SÉNÉCAILLE Laura ♥ ROUMY Anamé ♥ TEYSSIER Célia ♥ CALSACY Matéo ♥ CHABANOL Jean ♥ BÉTEILLE Lisa ♥</p> <p>Encadrants : BRINGAY Sandra LOMBARDO Florian</p>										
<p> PROBLÉMATIQUE/OBJECTIFS</p> <p>Raison du projet (Pourquoi je fais ce projet ?) quelle est la problématique : Malheureusement de nos jours, il n'y a pas assez de médiatisation sur le suicide ou de démarche publique. De plus, rien ne nous permet de prévoir à l'avance si quelqu'un va se suicider. L'intérêt de ce sujet est donc de pouvoir identifier, par une analyse textuelle, les personnes étant les plus à risque de faire une tentative de suicide pour ainsi les sensibiliser, concrètement et rapidement.</p> <p>Solution déjà existante : Chercheurs de l'Inserm et de l'Université de Bordeaux ont créé un algorithme qui identifie les 4 principaux indicateurs qui prédisent le risque suicidaire des étudiants.</p>	<p> BUSINESS & DATA</p> <p>Gain qualitatifs : Sensibiliser et pouvoir accompagner concrètement ces personnes par un corps hospitalier compétent. Recueillir plus de témoignages car souvent plus facile de poster un a "appel à l'aide" sur internet que lors d'un entretien psychologique.</p> <p>NB utilisateurs potentiels : Des centaines de personnes.</p> <p>Data ou SI source : Données sur Twitter, Reddit et Wikipédia récoltées via des API</p>										
<p> ROADMAP</p> <p>État des lieux (où en êtes-vous du projet) :</p> <table border="0"> <tr> <td>Fait</td><td>A faire</td></tr> <tr> <td>✓ En possession des données Twitter "Suicides"</td><td>✓ État de l'art</td></tr> <tr> <td>✓ En possession des données Wikipédia "Suicides"</td><td>✓ Récolte des données manquantes</td></tr> <tr> <td></td><td>✓ Exploration des données</td></tr> <tr> <td></td><td>✓ Classifications: RandomForest, SVM, K-Means</td></tr> </table>	Fait	A faire	✓ En possession des données Twitter "Suicides"	✓ État de l'art	✓ En possession des données Wikipédia "Suicides"	✓ Récolte des données manquantes		✓ Exploration des données		✓ Classifications: RandomForest, SVM, K-Means	<p> VOS ATTENTES</p> <p>Appui méthodologie / tech /suivi projet :</p> <ul style="list-style-type: none"> -Entretiens avec les encadrants pour un suivi projet -Présentation des résultats ou des problèmes rencontrés <p>→ Fluidifier la communication entre l'équipe et les commanditaires</p>
Fait	A faire										
✓ En possession des données Twitter "Suicides"	✓ État de l'art										
✓ En possession des données Wikipédia "Suicides"	✓ Récolte des données manquantes										
	✓ Exploration des données										
	✓ Classifications: RandomForest, SVM, K-Means										

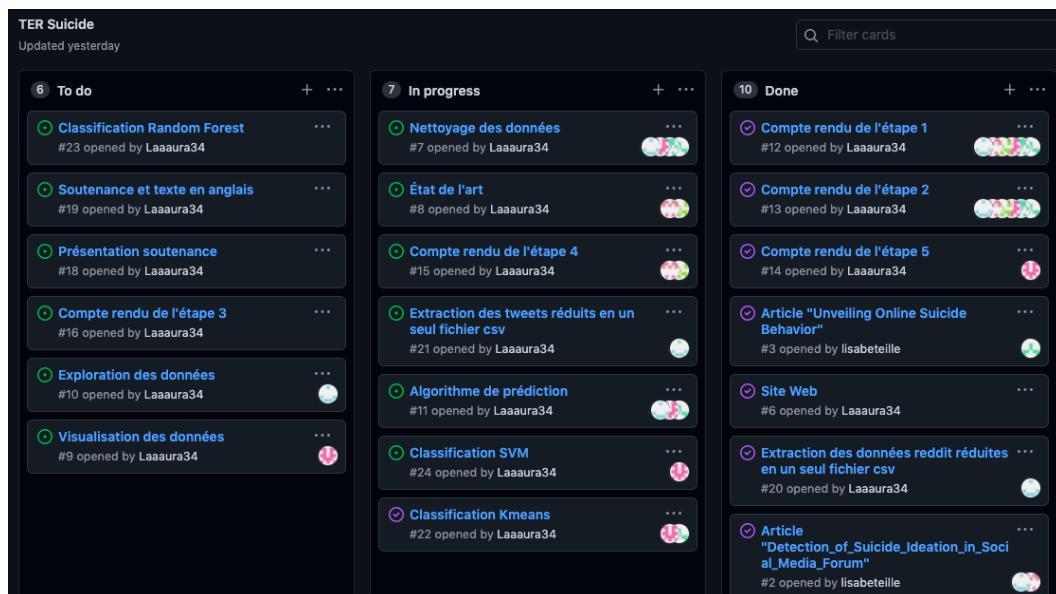
Sur le speed-boat, Mateo est le personnage en jaune car il se sentait plutôt serein, il était concentré sur sa partie qui lui a pris du temps et faisait confiance au reste de l'équipe pour diriger le bateau. Jean est le personnage bleu car il avait sous-estimé la charge de travail pour ce projet. Paniqué, il a donc dû accélérer son travail et son investissement. Laura, en orange, reste concentrée sur l'objectif pour arriver rapidement à l'idéal. Célia, en rouge, est confiante, motivée et encourageante, elle garde une vue d'ensemble sur ce que tout le monde fait. Anamé, en vert, est la conductrice du bateau car elle est la team lead, elle dirige le bateau et elle s'aide des autres membres pour se faire guider. Lisa est en violet, elle est concentrée sur ce qu'il se passe et sur les autres membres mais est sceptique sur le chemin qu'on emprunte.



Concernant l'organisation de notre projet, nous avons la chance d'avoir une bonne entente générale dans notre groupe ce qui a permis une bonne communication entre nous. Nous avons favorisé l'assignement de tâches aux personnes plus à l'aise dans ces étapes pour plus d'efficacité. Néanmoins, nous avons préféré que les personnes à l'aise prennent du temps pour expliquer ce qu'elles faisaient aux membres qui l'étaient moins. Cela leur permettait de comprendre et de progresser grâce au projet. L'assignation des tâches a été faite sur la base du volontariat.

Nous avons également élu un team lead, Anamé, qui nous a permis de faire tenir notre motivation à travailler tout le long du semestre et de vérifier si les tâches de chacun étaient respectées.

Pour l'organisation plus technique, nous avons utilisé les moyens de communication tel que Discord pour faire des réunions entre nous en appel vocal, avec un groupe dédié au projet. Whatsapp nous a également servi pour des contacts plus réguliers par messages. Google drive, Google colab et Github pour l'organisation du partage de données, de codes et de documents. Notre Github ne contient pas énormément de push, car nous avons préféré réaliser nos tâches en petit sous-groupe en présentiel l'un chez l'autre. Nous avons aussi utilisé Github pour la gestion des tâches, pour savoir celles qui nous restaient à faire, celles en cours et celles terminées. Ne pas avoir choisi Trello pour notre gestion de tâches nous a permis de nous familiariser encore plus avec Github. De plus, cela était également plus pratique pour nous, étant donné que nous allions plus souvent sur Github pour voir nos données, nous pouvions tout réunir au même endroit. L'utilisation de Github pour gérer nos tâches était facile à comprendre, ce qui nous a facilité le travail.



Il était important pour nous que tout le monde puisse avoir accès aux données et aux documents quand ils le souhaitent.

Chacun de ces appels nous permettait aussi de préparer les réunions avec Sandra et Florian, pour récolter les questions que l'on devrait leur poser et pour faire un récapitulatif de ce qu'on avait à leur présenter. Cela nous permettait de ne rien oublier et d'être plus efficace lors de ces réunions que l'on avait une fois durant nos deux semaines de cours.

Au fur et à mesure du temps, nous avons été contraint de changer notre diagramme de Gantt car certaines des tâches que nous avions prévu nous ont pris plus de temps que prévu, et d'autres moins. Certaines tâches ont été également omises lors du premier diagramme que nous avons fait. Nous avons donc refait un diagramme de Gantt qui colle parfaitement avec les avancées de ce projet.

Les problèmes les plus importants que nous avons rencontré ont été au niveau de la récupération des données Wikipédia, nous avons eu des soucis pour faire marcher l'algorithme mis à notre disposition. Mais aussi au niveau de la classification, nous avons eu des difficultés à chercher et trouver des exemples de classifications sur du texte, notamment pour afficher des k-means. Au niveau du nettoyage du texte, nous avons eu besoin d'un temps de réflexion pour régler les expressions régulières (par exemple pour transformer tous les espaces).

Pour visualiser les tâches que nous avons réellement faites, nous avons utilisé une matrice RACI. Cette matrice nous a aidé à être plus clair dans la restitution des participations de tâches.

		Anamé	Lisa	Laura	Jean	Célia	Mateo
Web et visualisation	Site web	I	I	A/R	I	I	I
Analyse	Nettoyage des données	A	I	C	I	R	I
	Exploration des données avec Iramuteq	I	I	I	C	A/R	I
	Exploration des données avec python	R	C	R	I	A	I
Accompagnement du projet	Gestion de projet	A	R	I	I	I	I
Mise en place du projet	Lecture des articles	R	R	R	R	R	R
	Etat de l'art	I	I	I	R	I	R
	Prise en note des réunions	R	R	I	I	I	I
	Récupération des données	R	R	R	R	R	R
Classification	Création de classification sur les données	R	C	R	C	C	C
	Recherche de potentiels classifieurs	R	I	R	I	I	I
Finalisation du semestre	Rédaction compte rendus	R	R	R	R	R	R
	Présentation oral / slides	R	R	R	R	R	R

A : Approbateurs
I : Informés

R : Réalisateurs
C : Consultés

Pour les améliorations à venir, nous allons essayer de respecter un peu plus notre diagramme de Gantt pour ne pas prendre de retard et nous allons essayé de faire plus attention au Git et de faire des push plus régulièrement.

Conclusion

Pour conclure, nous pouvons dire que malgré les quelques éléments à modifier pour la suite de l'année, la cohésion et la communication de notre équipe mènent au bon fonctionnement de notre gestion de travail.

Bibliographie

<https://flexjob.fr/un-atelier-un-outil-le-speed-boat/>

<https://presse.inserm.fr/predire-les-comportements-suicidaires-chez-les-etudiants-grace-a-lintelligence-artificielle/43085/>