

Réalisez un dashboard et assurez une veille technique

CONTENU

Mission - Concevez ...

Mission - Réalisez u...

Livrables et soutena...

Évaluation

SUPPORTS PÉDAGOGIQUES

Companion Nouveau

Cours

Ressources

Mission - Réalisez une veille technique



Comment allez-vous procéder ?

Cette mission continue sur le même scénario de projet professionnel que vous avez démarré précédemment.

Vous pouvez suivre les étapes pour vous aider à réaliser vos livrables.

Avant de démarrer, nous vous conseillons de :

- lire toute la mission et ses documents liés ;
- prendre des notes sur ce que vous avez compris ;
- consulter les étapes pour vous guider ;
- préparer une liste de questions pour votre première session de mentorat.

Prêt à mener la mission ?

Vous avez bien avancé dans la réalisation du dashboard, alors Michaël décide de vous confier une autre mission.

Bonjour ! Je suis Companion, votre assistant IA dédié pour toute interrogation sur vos projets.

Poser une question

De : Michaël

À : moi

Objet : Veille

Bonjour,

Merci encore pour tout ce que tu as fait jusqu'à maintenant ! J'ai vu que tu as travaillé sur de nombreux sujets chez nous en plus de celui-ci, en particulier sur des problématiques de données texte (NLP) et de données d'images.

Tu sais que nous sommes soucieux de mettre en œuvre les dernières techniques en data science sur ces deux thématiques. Pourrais-tu réaliser un état de l'art sur une technique récente de modélisation de données texte ou de données image, l'analyser, la tester et la comparer à une approche plus classique que tu as réalisée précédemment ?

Concrètement, voici ce que j'attends de ta part :

- L'état de l'art devra concerner une technique datant de moins de 5 ans, présentée dans un article.
- La technique doit être référencée sur des sites de recherche ([Arxiv](#)), des sites connus (par exemple [fastml](#), [machine learning mastery](#), [kdnuggets](#), [import AI](#), [MIT tech review](#), [MIT news ML](#)) ou des newsletters de qualité comme [data elixir](#) et [data science weekly](#).
- Tu réaliseras et nous présenteras une preuve de concept qui met en oeuvre cette nouvelle technique avec les données texte ou image que tu as déjà exploitées précédemment.
- Tu nous expliqueras rapidement les concepts et techniques dans une note méthodologique (modèle en pièce-jointe) et lors d'une présentation.

Je suis convaincu que ta curiosité nous permettra de découvrir de nouvelles techniques performantes.

Bonne journée,
Michaël

Pièce jointe :

- [Modèle de note méthodologique](#)

Bon courage pour cette mission supplémentaire !

Étapes

Étape 1 : Préparez la veille technique en NLP ou en computer vision

Avant de démarrer cette étape, je dois avoir :

- pris connaissance du modèle de note méthodologique à compléter ;
- choisis un jeu de données où j'ai déjà implémenté une démarche de computer vision ou de NLP, ou un nouveau jeu de données où je peux rapidement reproduire une démarche implémentée auparavant.

Une fois cette étape terminée, je devrais avoir :

- l'implémentation ainsi que la note méthodologique d'un Proof of Concept (POC) d'une technique de modélisation récente.

Recommandations :

- Choisissez rapidement les étapes (pre-processing, postprocessing, feature engineering etc.) et hypothèses qui resteront fixes pour le POC. L'idée n'est pas de refaire la pipeline de ML de bout en bout, mais de rapidement expérimenter.
- Une piste à explorer parmi d'autres est le Transformer et ses variantes. C'est une architecture de réseaux de neurones qui constitue le fondement de toutes les méthodes de NLP et de vision modernes comme les LLMs.
- Soyez vigilant au temps de calcul de certaines méthodes état de l'art, surtout si vous envisagez un ré-entraînement d'architectures lourdes. Vous pouvez envisager de réduire la taille du jeu de données d'entraînement par exemple.

Points de vigilance :

- Vos sources bibliographiques doivent être sérieuses (papier de recherches avec au moins quelques citations et références). Vous pouvez citer des articles de blogs (exemple Towards Data Science) s'ils sont aussi dotés d'une bibliographie avec des références solides. Attention à ne pas citer des papiers en pre-print.
- Soyez capable d'expliquer la nouvelle technique avec aisance quel que soit le niveau de détail demandé : très synthétique pour une audience métier, ou en rentrant dans les détails mathématiques si cela est demandé. D'ailleurs, cette maîtrise mathématique doit être visible dans votre synthèse de la bibliographie.

Ressources :

- [Arxiv](#)
- [ConnectedPapers](#) - outil visuel permettant de naviguer les papiers de recherche.
- [Paperswithcode](#) - répertoire de papiers qui ont été publiés avec du code ou dont les résultats ont été reproduits par la communauté Data Science.

Étape 2 : Vérifier votre travail et préparez-vous à la soutenance

- Révisez attentivement tous vos livrables pour garantir qu'ils sont complets et conformes aux instructions fournies, en mettant un accent particulier sur la clarté et l'organisation.
- Préparez une démo fluide du dashboard, en mettant en avant ses fonctionnalités clés et son déploiement Cloud.
- Assurez-vous que votre présentation résumet votre projet du manière claire et consice, limitée à 30 slides.
- Préparez-vous à répondre aux questions spécifiques sur vos choix et méthodologies pendant la phase de discussion, en réfléchissant à l'avance aux justifications de vos décisions.

Ressources:

- [Améliorez l'impact de vos présentations](#)

Précédent

Suivant

