

Proposition de stage Niveau Master / Ingénieur

Titre

Développement de modèles d'apprentissage pour la prévention d'accidents dans le BTP

Mots-clés

Science des données ; Machine et Deep learning ; LLM (Large Language Models) ; Accidentologie ; BTP

Laboratoire et entreprise impliqués

Institut Pascal - <http://www.institutpascal.uca.fr>

K-Dix - <https://www.k-dix.fr/>

Lieu de travail, période et rémunération

IUT Le Puy-en-Velay (43), pour 6 mois, démarrage en février/mars 2025, rémunération de 4,35€ / heure.

Encadrants

L'encadrement sera réalisé par

- M. Antoine Vacavant - antoine.vacavant@uca.fr
- Mme Claire Tomas - claire.tomas@k-dix.fr
- Mme Maire-Noëlle Rey - marie-noelle.rey@k-dix.fr

Contexte scientifique et technologique

La détection des accidents et la gradation de leur sévérité par des outils d'intelligence artificielle ont fait l'objet de nombreuses recherches à l'échelle internationale. Diverses techniques de machine et deep learning ont été exploitées afin de classifier des données variées en lien avec des accidents de travail. En guise d'exemple, la base de données OSHA (*Occupational Safety and Health Administration*, <https://www.osha.gov/data>) fait référence pour tester et valider les modèles d'apprentissage et de prédiction dans ce domaine. De manière similaire, à l'échelle nationale, on peut citer la base EPICEA (INRS, <https://www.inrs.fr/publications/bdd/epicea.html>) qui répertorie des comptes-rendus d'accidents. En France, les solutions logicielles portant sur la détection d'accidents sont orientées vers la vision artificielle, par exemple avec les applications SécuriSpot (<https://www.fast-point.net/>) ou Blaxtair (<https://blaxtair.com/>).

Le sujet de stage proposé en collaboration entre l'Institut Pascal et l'entreprise K-Dix est de développer une solution logicielle innovante avec un modèle de prédiction d'accidents, incorporant toutes sortes de données (statiques et dynamiques) en lien avec des salariés du BTP. Grâce à l'application K-Dix, il est possible d'obtenir de manière instantanée de très nombreuses données personnelles (anonymisées) des collaborateurs pour développer des modèles prédictifs. A plus long terme, la solution développée pourrait être déployée dans K-Dix pour aider à mieux prévenir les accidents subis dans les entreprises de BTP.

Objectifs détaillés

Le travail de ce stage se décompose de la manière suivante, et peut évoluer en fonction des compétences du/de la stagiaire et de l'avancement :

- Une première étape de bibliographie et de prise en main des travaux précédents sera nécessaire. Il s'agira de reprendre le travail d'élèves ingénieurs de l'ISIMA en 2024 (code sous forme de notebooks Python et rapport), et d'adapter ces premiers modèles aux données actualisées et anonymisées disponibles par K-Dix.

- Le stagiaire devra d'abord intégrer des données statiques, puis des données temporelles (absences, dates habilitations, etc.) sous forme de *time series* dans ces nouveaux modèles d'apprentissage et de prédiction.
- Il s'agira enfin d'intégrer une analyse des messages (non structurés) grâce à des modèles de langages. Le/La stagiaire devra comparer différentes approches et proposer une solution d'implémentation pour les modèles développés.
- Selon l'avancement du stage, différentes valorisations du travail pourront être envisagées (publication scientifique, déploiement de code pour l'entreprise, etc.).

Compétences attendues

Le/La stagiaire recruté(e) doit posséder de solides compétences en programmation, notamment avec des outils de machine learning et deep learning principalement dans le langage Python. Le sens du travail d'équipe, des capacités d'autonomie et une bonne communication sont également requis.

Pourquoi choisir ce stage ?

- Vous évoluerez dans un cadre professionnel faisant appel à vos compétences et connaissances acquises dans votre formation et à vos capacités d'innovation et d'adaptation.
- Vous serez amené(e) à manipuler plusieurs outils technologiques de pointe faisant appel au machine/deep learning et aux LLMs.
- L'environnement du stage (IUT, Institut Pascal) possède des compétences aiguisées dans l'intelligence artificielle et le développement logiciel, et vous apportera un cadre stimulant pour votre stage. L'entreprise K-Dix apportera sa solide expertise dans l'accidentologie en BTP et dans le développement de solution digitale innovante pour le bien-être au travail.
- La valorisation de ce travail a un fort potentiel pour l'entreprise.

Comment candidater ?

Envoyez dès à présent un mail à tous les encadrants du stage et avant le 30/11/2024, avec les pièces suivantes :

- Lettre de motivation
- Dernières notes obtenues
- CV à jour

Un entretien devrait être organisé à la suite de votre candidature.

Références

- [1] Luo et al. (2023) : Application of machine learning technology for occupational accident severity prediction in the case of construction collapse accidents. *Safety Science*, 2023
- [2] Sarkar et al. (2019) : Application of optimized machine learning techniques for prediction of occupational accidents. *Computers & Operations Research*, 2019
- [3] Khairuddin et al. (2023) : Harnessing the Multimodal Data Integration and Deep Learning for Occupational Injury Severity Prediction. *IEEE Access*, 2023
- [4] Khairuddin et al. (2022) : Occupational Injury Risk Mitigation: Machine Learning Approach and Feature Optimization for Smart Workplace Surveillance. *International Journal of Environmental Research and Public Health*, 2022
- [5] Khairuddin et al. (2024) : Contextualizing injury severity from occupational accident reports using an optimized deep learning prediction model. *PeerJ*, 2024
- [6] Fusun & Tufan (2021) : Comparison of machine learning methods in predicting binary and multi-class occupational accident severity. *Journal of Intelligent & Fuzzy Systems*, 2021
- [7] Sarkar et al. (2018) : Text-clustering based deep neural network for prediction of occupational accident risk: A case study. *iSAI-NLP*, 2018

- [8] Foreman et al. (2024) : Establishment-level safety analytics: a scoping review. *International Journal of Occupational Safety and Ergonomics*, 2024
- [9] Li & Wu (2023) : Deep Learning and Text Mining: Classifying and Extracting Key Information from Construction Accident Narratives. *Applied Sciences*, 2023
- [10] Montcharmont et al. (2022) : La gestion des risques professionnels par les ouvriers du BTP
- [11] Leclerc & Magnaval (2023) : Intelligence Artificielle & Génie Civil : enjeux et cas d'usages. *JIA*, 2023