**Resume de l’article**

La prédiction de la qualité des eaux, la détection des cancers font partie des principaux problèmes efficacement résolus par l’intelligence artificielle. Mais dans ce domaine, les obstacles les plus fréquents sont la disponibilité et la qualité des données. Cet article présente la mise en place d’une plateforme de collecte et d’étiquetage de données. Cette plateforme se présente comme une solution pour pallier au manque de données et vise à améliorer la qualité des données utilisées dans nos projets d’apprentissage automatique. À travers une approche méthodique, nous avons ciblé des problèmes d'intelligence artificielle dans des secteurs souffrant d'une pénurie critique de données. Une base de données adaptée a été créée, accompagnée d'une interface utilisateur intuitive pour optimiser le processus de collecte et d'étiquetage des données

**Introduction de l’article**

Le manque de données de qualité limite l'évolution constante de la recherche en intelligence artificielle. Les chercheurs et les praticiens ont des difficultés pour avoir beaucoup de données de qualité dans les entreprises pour entrainer efficacement et avoir des bons résultats dans différents thématiques en Intelligence artificielle. L’objectif est de développer une plateforme de collecte et d’étiquetage de données qui sera une alternative pour pallier le problème majeur de difficulté d’accès aux données d’entrainement de ces modèles d’IA.

En réponse à cette exigence, nous présenterons premièrement, le processus de conception, de développement et de mise en œuvre de cette plateforme novatrice de collecte et d'étiquetage de données et deuxièmement, nous présenterons les résultats obtenus de ce travail.

**Choix de l’approche**

Dans les langues locales, nous ne disposons pas de données dans chaque domaine des langues locales, nous ne disposons pas des images dans notre contexte pour mettre en place un modèle de l’intelligence artificielle, qui vérifie automatiquement, les données validées par les employés de la plateforme, c’est pourquoi nous affectons un rôle, a des utilisateurs plus expérimentés dans le domaine, pour effectuer la vérification manuelle. Nous prévoyons de développer un model de vérification automatique a base des données que nous aurons collectes avec l’approche manuelle en mettant en place un rôle expert qui sont des expert métier, qui font la vérification manuelle a place d’un modèle d’IA.

La plateforme dispose deux rôles : un rôle employé et un rôle expert, à choisir lors de l’inscription de chaque l’utilisateur. Chaque rôle conduit à des fonctionnalités définis.

Les utilisateurs qui ont le rôle employé, sont considérés comme des employés et ceux qui ont le rôle expert sont des experts du domaine. Les employés effectuent la collectée des données et l’étiquetage, les experts effectuent la vérification des données des employé avant de les enregistrer.

Nous proposons plusieurs domaines : la transcription textuelle de français en langues locales, la transcription des audios en langues locales, la classification de données en langues locales, la classification des audios en langues locales, la collecte des images. Les employés sont formés pour effectuer, la collecte et l’étiquetage.

**Objectifs de la Plateforme :**

Nous avons élaboré une plateforme polyvalente qui réponde à une diversité de besoins fondamentaux. Cette plateforme englobe la collecte et l'étiquetage de données textuelles et audio relatives aux langues locales, tout en permettant la collecte d'images. Cet outil garantit la sécurité des données, à facilite l'accès, offre la consultation des projets antérieurs, assure la formation des utilisateurs, et assume la responsabilité de la capacité d'expansion de l'outil.

Dans notre démarche, notre objectif demeure à trouver une alternative au manque de données, le développement de cet outil nous aide à surmonter ces défis en assurant une collecte et une gestion de données fluides, en accordant une attention particulière à la confidentialité, à l'accessibilité et à l'évolutivité, qui sont les éléments cruciaux de la plateforme.

Le processus en deux étapes

**La contribution de l’employé**

Après la phase d’inscription, l’employé doit choisir son domaine de compétence dans la liste de domaine proposé. Après cette première phase, il a deux possibilités, la première consiste à créer un nouveau projet ou la deuxième option qui consiste à consulter la liste de ses anciens projets.

Dans l’option création de projet, l’employé rempira un formulaire incrémental dans lequel il peut remplir autant que possible des informations en fonction du domaine choisit. Puis ensuite, il passe à la validation du formulaire si contenu respecte les règles fixées.

Les données créées par les employés sont automatiquement enregistrées dans notre base de données sécurisée. Chaque enregistrement est associé au projet correspondant, ce qui permet un suivi précis et une gestion efficace de toutes les informations collectées.

**Vérification par l'Expert**

Apres la phase d’inscription, les experts sont formes dans les mêmes domaines que l’employé. Il choisit son domaine de compétence dans la liste de domaine proposé. Après cette première phase, il a deux possibilités, la première consiste à créer un nouveau projet dans lequel il récupère les données de l'employé à partir de la base de données principale. Il sélectionne les données qui n'ont pas encore été examinées et effectue des vérifications pour s'assurer de leur qualité et de leur pertinence. Une fois vérifiées, les données sont transférées vers une nouvelle base de données dédiée à la validation et à l'approbation, la deuxième option qui consiste à consulter la liste de ses anciens projets. Les experts ont également la possibilité consulter la liste de ses projets et leur date de création, de suivre l'état d'avancement de leurs projets en cours pour être informer sur les tâches à accomplir et de s'assurer que tout progresse comme prévu.

Dans l’option création de projet, l’expert doit obtenir les données collectées par les employés pour effectuer des corrections. Mais ce sont les données des employés qui n’ont pas déjà suivi une correction que l’expert obtient. Puis ensuite il passe à la validation du formulaire.

**Sécurité des Données :**

Le service de l'authentification permet de connaitre les utilisateurs inscris dans notre plateforme, il permet de déterminer le rôle de chaque utilisateur (employé ou expert) lors de la connexion.

Chaque utilisateur choisit le rôle a effectué, cette authentification garantit que seules les personnes autorisées ont accès à ces informations et de conditionner ensuite l'accès aux différentes fonctionnalités de la plateforme en fonction du rôle.

Les employés et les experts peuvent ainsi bénéficier d'une expérience utilisateur adaptée à leurs besoins spécifiques. En effet, l'authentification joue un rôle fondamental dans le maintien de la sécurité, de l'intégrité et de la confiance de notre plateforme de collecte et d’étiquetage de données. Cela permet de garantir que seules les personnes autorisées ont accès aux informations et aux fonctionnalités.

Discussion

Sécurité des Données : La plateforme a intégré avec succès des mécanismes de sécurité avancés, garantissant la confidentialité et l'intégrité des données collectées. Les protocoles de sécurité mis en place ont non seulement répondu aux normes de l'industrie, mais ont également renforcé la confiance des utilisateurs dans la protection de leurs informations sensibles.

Accès Facile aux Données : L'interface conviviale conçue pour les chercheurs et stagiaires a considérablement simplifié le processus de soumission et de consultation des données. Les retours des utilisateurs indiquent une amélioration significative de l'efficacité, réduisant les obstacles potentiels et facilitant l'interaction intuitive avec la plateforme.

Consultation de Projets : La fonction de consultation des projets a permis aux chercheurs d'accéder rapidement à une liste détaillée de leurs projets antérieurs. Cette fonctionnalité a facilité la gestion de projets multiples et a contribué à accélérer le processus de recherche en offrant un accès immédiat aux détails pertinents.

Formation et Support : L'implémentation réussie du support et de la formation a eu un impact positif sur l'utilisation efficace de la plateforme. Les utilisateurs ont bénéficié de ressources éducatives complètes et d'un soutien dédié, améliorant ainsi leur compréhension de la plateforme et optimisant leur expérience.

Capacité d'Expansion : La conception adaptable de la plateforme a répondu aux besoins de croissance anticipée. La plateforme a démontré sa capacité à évoluer en toute fluidité, tant en termes de nombre de projets que d'augmentation d'utilisateurs et de données, garantissant ainsi sa pertinence à long terme.

Efficacité du Traitement : Les résultats de l'efficacité du traitement des données ont surpassé les attentes, assurant une expérience utilisateur optimale. Les performances rapides et efficaces ont contribué à accélérer les processus de recherche, renforçant ainsi la valeur ajoutée de la plateforme.

Contrôle d'Accès : Les mécanismes de contrôle d'accès ont été intégrés avec succès, garantissant que seuls les utilisateurs autorisés ont accès aux données sensibles. Cela a renforcé la sécurité globale de la plateforme, répondant aux normes de protection des informations sensibles.

Conlusion