



RAPPORT DE STAGE D'ÉTÉ

AskMind Biomedica Tunisie

Formation d'Ingénieur : INFO-1

Équipe de Développement :

Mokni Hamdi

Encadrant Professionnel :

M. Abderraouf Allani

 github.com/Mk-1000/AskMind

 MokniHamdi712@gmail.com

11 février 2025

Table des matières

1	Introduction	2
1.1	Contexte	2
1.2	Le Projet <i>AskMind</i>	2
2	Présentation de l'entreprise	3
2.1	Présentation générale	3
2.2	Offres et services	3
2.3	Présence de Biomedica Group en Tunisie	3
2.3.1	Implantation et effectif	3
2.3.2	Services proposés en Tunisie	3
3	Présentation du travail effectué	4
3.1	Introduction	4
3.2	Architecture du système	4
3.3	Description des tâches	5
3.3.1	Collecte et prétraitement des données	5
3.3.2	Indexation des documents	5
3.3.3	Développement du modèle de question-réponse	5
3.3.4	Évaluation du modèle	5
3.3.5	Intégration avec l'interface utilisateur	5
3.4	Défis et solutions	6
3.5	Conclusion	6
4	Interface Utilisateur	7
4.1	Introduction à l'Interface Utilisateur	7
4.2	Présentation générale de l'interface	7
4.3	Modes de Recherche et Fonctionnalités Avancées	8
4.4	Mémorisation et Amélioration Continue des Réponses	10
4.5	Gestion Manuelle et Administration	10
4.6	Conclusion de l'Interface Utilisateur	11
5	Conclusion	12

Introduction

Le stage d'été chez Biomedica a été une expérience passionnante et enrichissante, me permettant de plonger dans le domaine de l'intelligence artificielle et du traitement du langage naturel. Le présent rapport résume mon parcours et mes réalisations lors de ce stage, en mettant l'accent sur le projet captivant *AskMind*.

1.1 Contexte

Au cours de ma formation en informatique, j'ai toujours été fasciné par le potentiel de l'intelligence artificielle pour améliorer la vie des gens et résoudre des problèmes complexes. Biomedica, une entreprise réputée dans le domaine biomédical, offre des solutions innovantes pour optimiser la recherche et l'analyse de données dans le secteur de la santé. Mon stage chez Biomedica était une occasion exceptionnelle de fusionner ma passion pour l'informatique et le domaine médical.

1.2 Le Projet *AskMind*

Le projet *AskMind* est un système de question-réponse avancé qui vise à fournir des réponses pertinentes et contextuelles aux utilisateurs. En exploitant les techniques de recherche d'informations et de traitement du langage naturel, *AskMind* est conçu pour interpréter les questions des utilisateurs et extraire les informations pertinentes à partir de diverses sources de données, telles que des documents, des manuels et des spécifications.

Présentation de l'entreprise

2.1 Présentation générale

Biomedika Group est un acteur majeur spécialisé dans la fourniture de services et de produits destinés aux secteurs de la santé, de l'industrie et du nucléaire.

Fondé en 2004, le groupe a son siège social à Villeneuve-d'Ascq, en France. Avec plus de 1000 employés, il est aujourd'hui implanté dans plus de 20 pays à travers le monde.

2.2 Offres et services

Biomedika Group propose une vaste gamme de services et de produits, notamment :

- Maintenance et réparation d'équipements biomédicaux
- Gestion de la qualité
- Formation professionnelle
- Vente de dispositifs biomédicaux
- Distribution de produits nucléaires

2.3 Présence de Biomedika Group en Tunisie

2.3.1 Implantation et effectif

Implanté en Tunisie depuis 2012, Biomedika Group possède son siège social à Tunis et compte aujourd'hui plus de 50 collaborateurs.

2.3.2 Services proposés en Tunisie

En Tunisie, l'entreprise offre divers services et produits :

- Maintenance et réparation d'équipements biomédicaux
- Gestion de la qualité
- Formation spécialisée
- Vente de dispositifs biomédicaux

Présentation du travail effectué

3.1 Introduction

Lors de mon stage chez Biomedica Tunisie, j'ai eu l'opportunité de travailler sur le projet *AskMind*. Ce projet repose sur l'intelligence artificielle et le traitement du langage naturel (NLP) pour développer un système de question-réponse avancé capable de fournir des réponses précises à partir d'un ensemble de documents biomédicaux.

Ce chapitre présente une vue d'ensemble des travaux réalisés, en détaillant les différentes tâches accomplies et les technologies mises en œuvre.

3.2 Architecture du système

Le projet *AskMind* repose sur une architecture modulaire permettant de traiter efficacement les requêtes des utilisateurs. L'architecture comprend plusieurs composants principaux :

- **Frontend** : Une interface web développée avec Django permettant aux utilisateurs d'interagir avec le système.
- **Backend** : Un serveur Django gérant les requêtes des utilisateurs et orchestrant le processus de recherche.
- **Moteur de recherche** : Elasticsearch utilisé pour l'indexation et la récupération rapide des documents pertinents.
- **Module NLP** : Un modèle de traitement du langage naturel basé sur BERT et les bibliothèques NLTK et `spaCy` pour l'analyse des textes et l'extraction des réponses.

3.3 Description des tâches

Durant mon stage, j'ai participé activement à différentes phases du développement du projet *AskMind*. Voici une description détaillée des tâches réalisées :

3.3.1 Collecte et prétraitement des données

L'une des premières étapes du projet a été la collecte de documents biomédicaux à partir de différentes sources ouvertes. Une fois les données collectées, un processus de prétraitement a été appliqué :

- Nettoyage des textes (suppression des caractères spéciaux, normalisation des formats).
- Tokenisation et suppression des stopwords.
- Lemmatisation et analyse syntaxique avec `spaCy`.

3.3.2 Indexation des documents

L'indexation a été réalisée en utilisant Elasticsearch, ce qui permet d'effectuer des recherches rapides et efficaces sur les documents traités. Les documents ont été transformés en vecteurs sémantiques pour améliorer la précision des recherches.

3.3.3 Développement du modèle de question-réponse

Pour permettre une interaction naturelle avec le système, un modèle de question-réponse a été mis en place :

- Fine-tuning d'un modèle BERT pré-entraîné sur un corpus biomédical.
- Amélioration des performances grâce à l'ajout de techniques de pondération TF-IDF.
- Évaluation du modèle sur des ensembles de tests spécifiques.

3.3.4 Évaluation du modèle

L'évaluation du modèle a été réalisée en utilisant plusieurs métriques :

- Score F1 pour mesurer la précision et le rappel.
- Comparaison des réponses générées avec des réponses de référence annotées par des experts.

3.3.5 Intégration avec l'interface utilisateur

Enfin, le modèle a été intégré dans l'application web Django. Cette étape comprenait :

- Développement d'une API REST pour interagir avec le modèle NLP.
- Affichage des réponses de manière dynamique sur l'interface web.
- Tests utilisateur pour améliorer l'ergonomie et la réactivité du système.

3.4 Défis et solutions

Durant le développement, plusieurs défis ont été rencontrés :

- **Problème de performance** : L'indexation de gros volumes de données ralentissait les recherches. **Solution** : optimisation des requêtes Elasticsearch et implémentation d'une pagination.
- **Précision du modèle** : Certaines réponses étaient hors contexte. **Solution** : ajustement des hyperparamètres de BERT et ajout de filtres contextuels.
- **Interopérabilité** : Assurer une communication fluide entre le frontend et le backend. **Solution** : mise en place d'une API REST bien documentée.

3.5 Conclusion

Ce chapitre a présenté les travaux réalisés lors de mon stage chez Biomediqua Tunisie dans le cadre du projet *AskMind*. De la collecte des données à l'intégration finale du modèle NLP, chaque étape a permis d'améliorer la précision et l'efficacité du système de question-réponse.

Interface Utilisateur

4.1 Introduction à l'Interface Utilisateur

L'interface utilisateur (UI) de *AskMind* a été conçue dans une optique de convivialité et d'efficacité pour permettre aux utilisateurs d'interagir facilement avec le système. Le design a été axé sur la simplicité et l'intuitivité, afin de garantir une expérience optimale tant pour les utilisateurs non spécialisés que pour les experts biomédicaux. Ce chapitre présente les différentes interfaces, fonctionnalités, ainsi que les choix techniques associés à la création de l'interface utilisateur.

4.2 Présentation générale de l'interface

L'interface principale est composée d'une page d'accueil qui présente les différentes options disponibles pour l'utilisateur, avec une navigation fluide entre les différentes sections. Les utilisateurs peuvent interagir avec le système en utilisant différentes méthodes de recherche adaptées à leurs besoins spécifiques.

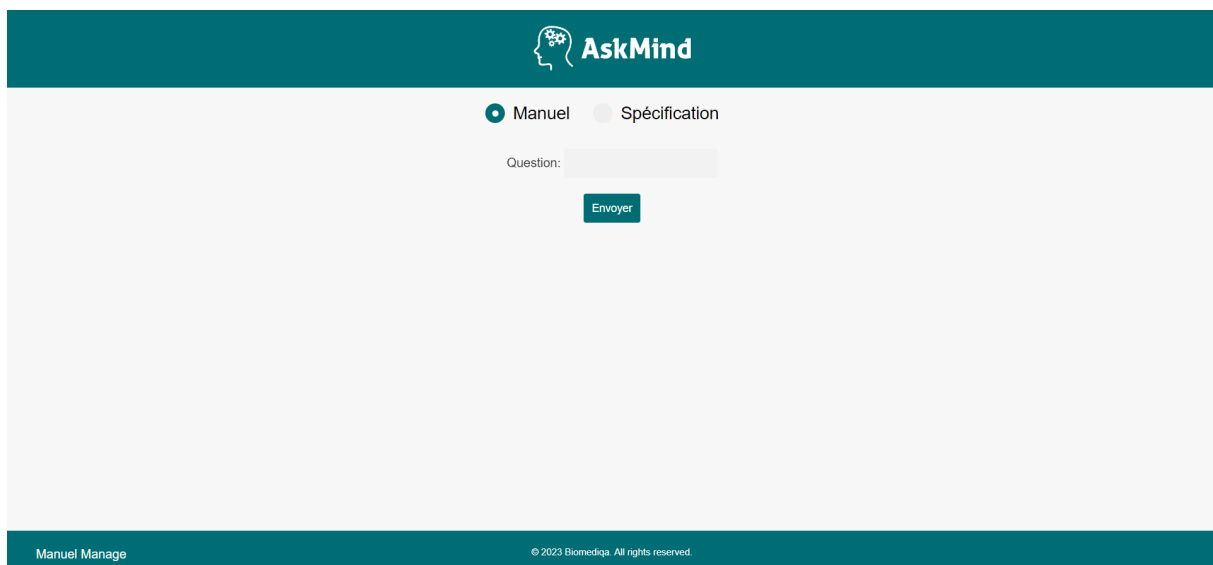


FIGURE 4.1 – Capture d'écran de l'interface utilisateur - page d'accueil

L'interface a été développée en utilisant des technologies web modernes, garantissant ainsi une compatibilité maximale avec divers navigateurs et appareils. Elle est également optimisée pour une expérience utilisateur fluide avec des temps de réponse rapides.

4.3 Modes de Recherche et Fonctionnalités Avancées

AskMind propose deux modes de recherche, chacun offrant une approche différente pour répondre aux besoins des utilisateurs :

- **Mode Manuel** : Permet aux utilisateurs de rechercher des informations dans des bases de données préalablement constituées par la gestion manuelle des jeux de données. Ce mode est particulièrement utile pour les utilisateurs ayant des besoins spécifiques et des requêtes bien définies.
- **Mode Spécification** : Ce mode repose sur des techniques avancées de recherche documentaire, utilisant des modèles d'intelligence artificielle pour analyser et extraire les informations pertinentes à partir de divers types de fichiers tels que .doc, .csv, etc.

The screenshot shows the AskMind web interface. At the top, there's a dark teal header with the AskMind logo (a head with gears) and the text 'AskMind'. Below the header, there are two radio buttons: 'Manuel' (selected) and 'Spécification'. A text input field contains the question 'Question: comment faire un suivi médical', and an 'Envoyer' button is below it. The results section, titled 'Meilleur résultat', shows a box with the following content:

Question: Un suivi médical

Réponse: Un suivi médical est :

- À enregistrer : Si le processus de création de l'employé n'est pas terminé
- À planifier : Dans le dernier trimestre d'une période de 2 ans qui commence le jour de la dernière visite à condition que le travailleur est conforme.
- Planifié : Si la dernière visite est planifiée
- À valider : Si la date de réalisation de la dernière visite médicale est passée mais la conformité n'est pas renseignée
- Conforme : Si la dernière visite est réalisée et le travailleur est apte
- Non conforme : Si l'employé n'a aucun suivi médicale ou si la dernière visite réalisé est inapte ou bien la date de la réalisation du dernière visite dépasse les 24 mois.


score de similarité: 1.0

Temps de traitement: 0.0

At the bottom of the interface, there's a dark teal footer with 'Manuel Manage' on the left and '© 2023 Biomediga. All rights reserved.' on the right.

FIGURE 4.2 – Sélection de l'option « Manuel » dans l'interface

Lorsque l'option "Spécification" est choisie, *AskMind* lance une analyse approfondie des documents, en intégrant des modèles d'intelligence artificielle pour affiner les résultats et proposer des réponses plus précises.



☐ Manuel ☒ Spécification

Question: qui est le REDACTEUR de bi

Envoyer

Indexing completed. Total time: 0.01 seconds.

Les résultats

Question: qui est le REDACTEUR de biomedica

Réponse:

Youssef LAYNAOUI

— Temps de traitement de la réponse: 16.34s

— Confiance: 0.89

— fichier: 20210223_complément_Module travailleur exposé_statut du suivi dosimétrique.pdf

— Temps de recherche des fichiers: 0.08s

Paragraphs:

cristallin mensuelle ou trimestrielle. - Un travailleur classé en catégorie B sans EPI/EPC ou avec EPI/EPC pour le s extrémités doit avoir une dosimétrie bague ou poignet trimestrielle. - Un travailleur classé en catégorie A sans EPI/EPC ou avec EPI/EPC pour le s extrémités doit avoir une dosimétrie bague ou poignet mensuelle. - Un t ravailleur concerné interv enant en zone extrémi té doit avoir un dosimètre bague ou poignet en respectant les périod ités précéden ment cités. Dosimétrie active : Le dosimètre actif est obligatoire dans deux conditions : o Un travailleur intervenant en zon es contrôlées o Un travailleur intervenant en zone d'opération MODULE : TRAVAILLEUR EXPOSE DESCRIPTIF : Complément concernant le statut du suivi dosimétrique DATE : 23/02 /2021 REDACTEUR : Youssef LAYNAOUI</p>

— Temps de recherche des paragraphes: 1.08s

Correct Incorrect

Question: qui est le REDACTEUR de biomedica

Réponse:

Youssef LAYNAOUI

— Temps de traitement de la réponse: 14.81s

— Confiance: 0.87

— fichier: 20210223_complément_Module travailleur exposé_visite médicale.pdf

— Temps de recherche des fichiers: 0.08s

Paragraphs:

Description : champ p texte de 300 c aractères • Dans cette con figuration ce tableau joue le rôle a la fois d 'un h istorique et joue le rôle de déclencheur d 'une visite médicale si une nouvelle exposition en situation d 'urgence radiologi que est entrée : o Un besoin de vis ite médicale sui te intervention en si tuation radiologique est créé . 2/ Champs électromagnétiques • Suppr imer ce champ de cette étape , nous garderons cette information uniquement dans l'étape « champ électromagnétique ». Un complément de cahier des charges du champs dans l 'étape « champ électromagnétique » sera fourni. MODULE : TRAVAILLEUR EXPOSE DESCRIPTIF : Complément concernant l'étape « Suivi médical » du processus d'enregistrement d'un travailleur exposé. DATE : 23/02 /2021 REDACTEUR : Youssef LAYNAOUI</p>

— Temps de recherche des paragraphes: 1.53s

Correct Incorrect

Manuel Manage

© 2023 Biomedica. All rights reserved.

FIGURE 4.3 – Sélection de l'option « Spécification » dans l'interface

4.4 Mémorisation et Amélioration Continue des Réponses

L'interface permet de mémoriser les réponses validées par les utilisateurs pour garantir une précision optimale lors de futures requêtes similaires. Cette fonctionnalité est essentielle pour améliorer l'efficacité du système et offrir des réponses toujours plus pertinentes au fur et à mesure de l'utilisation.

The screenshot shows the AskMind web interface. At the top is a dark teal header with the AskMind logo. Below the header, there are two tabs: 'Manuel' (inactive) and 'Spécification' (active). Under the 'Spécification' tab, there is a text input field containing the question 'qui est le REDACTEUR de bi' and a green 'Envoyer' button. Below this, a section titled 'Meilleur résultat' contains a white box with the following information: 'Question: qui est le REDACTEUR', 'Réponse: Youssef LAYNAOUI', 'score de similarité: 1.0', and 'Temps de traitement: 0.0'. At the bottom of the interface, there is a dark teal footer with the text 'Manuel Manage' on the left and '© 2023 Biomedica. All rights reserved.' on the right.

FIGURE 4.4 – Fonction de mémorisation des réponses validées

4.5 Gestion Manuelle et Administration

En parallèle de l'interface utilisateur, une interface de gestion manuelle a été mise en place pour permettre aux administrateurs de gérer les contenus du système, tels que les documents, les jeux de données et les réponses mémorisées. Cette page permet d'assurer une mise à jour constante du système pour l'adapter aux nouvelles informations biomédicales.

The screenshot shows the AskMind manual management interface. It features a dark teal header with the AskMind logo. Below the header, there is a 'Question:' label followed by a text input field. Underneath, there is an 'Answer:' label followed by a larger text area. Both the input field and the text area have a green 'Envoyer' button below them. At the bottom of the interface, there is a dark teal footer with the text '© 2023 Biomedica. All rights reserved.'

FIGURE 4.5 – Page de gestion manuelle pour les administrateurs

4.6 Conclusion de l'Interface Utilisateur

L'interface utilisateur de *AskMind* a été conçue avec un accent particulier sur l'ergonomie et l'accessibilité. Chaque fonctionnalité a été soigneusement intégrée pour faciliter l'interaction avec le système tout en garantissant la performance et la réactivité. Grâce à un design moderne et une navigation intuitive, les utilisateurs peuvent facilement accéder à l'information biomédicale, améliorant ainsi leur productivité et la qualité de leurs analyses.

Conclusion

Le projet *AskMind* a été une expérience marquante et un défi stimulant dans le domaine de l'intelligence artificielle et du traitement du langage naturel appliqué à la biomédecine. Grâce à ce stage chez Biomedica Tunisie, j'ai pu explorer des concepts avancés de modélisation de question-réponse et d'optimisation des systèmes de recherche documentaire. L'intégration des modèles d'intelligence artificielle tels que BERT et l'utilisation

d'outils puissants comme Elasticsearch ont permis de concevoir un système performant et scalable capable de répondre aux besoins des professionnels de la santé et de la recherche. L'évaluation constante des performances et les ajustements réalisés au fil du projet ont contribué à l'amélioration continue du modèle, augmentant ainsi sa précision et son efficacité. Le projet a permis de faire des avancées significatives, mais plusieurs axes

d'amélioration demeurent. Parmi eux, l'extension des fonctionnalités pour répondre à des requêtes encore plus complexes et la mise en place d'un système d'apprentissage continu pour que le modèle s'adapte aux nouvelles données biomédicales au fur et à mesure de son utilisation. D'un point de vue personnel, cette expérience m'a permis de développer

des compétences techniques solides en intelligence artificielle, tout en renforçant mes capacités de collaboration au sein d'une équipe interdisciplinaire. J'ai également acquis une meilleure compréhension des enjeux spécifiques à la biomédecine, ce qui me permettra de mieux appréhender les défis futurs dans ce domaine. Enfin, je tiens à remercier toute

l'équipe de Biomedica pour leur soutien, leurs conseils et l'opportunité qu'ils m'ont offerte de participer à ce projet innovant. Grâce à cette expérience, je me sens désormais mieux préparé à relever les défis technologiques de demain.