

Cahier des charges : Catégorisation automatique des tickets

Table des matières

| | |
|---|----|
| Table des matières | 1 |
| 1. Introduction et contexte | 2 |
| 2. Présentation du système intelligent de classification automatique..... | 2 |
| 3. Spécifications fonctionnelles..... | 3 |
| 4. Architecture technique | 5 |
| 5. Gouvernance du projet | 8 |
| 6. Indicateurs de performance (Estimatif) | 11 |
| 7. Sécurité et confidentialité | 13 |
| 8. Annexes techniques | 14 |

Rédaction :

- Ulrich Eneli Eneli
- Jean Claude Belinga
- Naomi Bibiane
- Angea Lorenza

1. Introduction et contexte

2. Présentation du système intelligent de classification automatique

Notre **système intelligent de classification automatique des tickets** représente une avancée majeure pour le Groupe Asten dans l'optimisation de la gestion des demandes d'assistance. Ce système s'appuie sur des technologies d'intelligence artificielle pour analyser, catégoriser et router automatiquement les tickets entrants vers les équipes appropriées.

Principe de fonctionnement

Le système utilise un algorithme d'intelligence artificielle spécialisé dans le traitement du langage naturel (NLP) qui analyse le contenu textuel des tickets pour déterminer leur nature. Cette analyse permet d'identifier automatiquement la catégorie appropriée et d'affecter le ticket à l'équipe concernée via notre outil de ticketing existant, créant ainsi un flux de travail fluide et efficace.

Le processus se déroule selon les étapes suivantes:

1. Réception du ticket dans le système
2. Analyse du contenu textuel par l'algorithme NLP
3. Comparaison avec une base d'exemples préétablie
4. Identification de la catégorie appropriée
5. Affectation automatique à l'équipe concernée
6. Intégration dans le flux de travail existant

Bénéfices attendus

Le déploiement de ce système apportera des avantages significatifs pour l'ensemble des parties prenantes:

| Partie prenante | Bénéfices |
|----------------------|---|
| Responsables métiers | Meilleure visibilité et catégorisation initiale plus fiable des demandes |
| Clients | Satisfaction accrue grâce à des délais de résolution plus courts et moins de tickets "perdus" |
| Équipes de support | Réduction des sollicitations manuelles permettant de se concentrer sur les cas complexes |
| Groupe Asten | Gain estimé de 20 à 40% sur les coûts opérationnels liés au service support |

| | |
|---------------------------|--|
| Ensemble des utilisateurs | Réduction de 30% des délais moyens de traitement des tickets |
|---------------------------|--|

Intégration technique

Le système reposera sur trois composants techniques principaux:

- Un **modèle NLP** pour l'analyse sémantique du contenu des tickets
- Des **algorithmes de classification** pour la catégorisation précise
- Des **modules de traitement des données** pour la préparation et la transformation des informations

Pour la phase de développement et de test, le système utilisera des simulations basées sur des fichiers CSV contenant des exemples représentatifs de tickets. Cette approche permettra de valider l'efficacité du système avant son déploiement en production.

Impact opérationnel

L'introduction de ce système intelligent transformera significativement le processus de gestion des tickets en:

- Éliminant les délais liés au tri manuel des demandes
- Réduisant les erreurs d'affectation qui génèrent des transferts multiples
- Permettant d'établir des métriques précises sur la nature des demandes reçues
- Offrant une base solide pour l'amélioration continue du service client

La technologie intelligente ne vient pas remplacer l'expertise humaine mais la compléter, en permettant aux collaborateurs de concentrer leur valeur ajoutée sur la résolution effective des problèmes plutôt que sur des tâches administratives de catégorisation.

3. Spécifications fonctionnelles

Le système intelligent de classification automatique des tickets se caractérise par des fonctionnalités distinctes adaptées aux différents utilisateurs et à leurs besoins spécifiques. Cette section détaille l'ensemble des capacités du système, des règles métier utilisées pour le tri et les interfaces utilisateur correspondantes.

Fonctionnalités par profil utilisateur

Le système propose des fonctionnalités adaptées aux différents acteurs impliqués dans le traitement des tickets:

| Profil utilisateur | Fonctionnalités principales |
|--------------------|---|
| Utilisateur final | <ul style="list-style-type: none"> • Soumission de tickets • Visualisation de la classification proposée • Contestation de la classification si nécessaire |

| | |
|------------------|--|
| Agent de support | <ul style="list-style-type: none"> • Visualisation des tickets classifiés • Consultation du score de confiance • Traitement selon la catégorie assignée |
| Superviseur | <ul style="list-style-type: none"> • Visualisation des statistiques de performance • Correction des classifications erronées • Analyse des tendances et modèles |
| Administrateur | <ul style="list-style-type: none"> • Validation et amélioration du modèle • Configuration des seuils de confiance • Gestion des règles d'affectation |

Processus de classification et d'affectation

Le système utilise un moteur d'analyse combinant traitement du langage naturel et règles métier pour traiter les tickets entrants:

1. **Importation des tickets** depuis les différents canaux d'entrée
2. **Analyse textuelle** du contenu via les algorithmes NLP
3. **Classification** parmi les 8 catégories métiers prédéfinies
4. **Calcul du score de confiance** pour chaque classification
5. **Affectation automatique** ou manuelle selon le seuil de confiance
6. **Routage vers l'équipe concernée** dans l'outil de ticketing

Règles métier et critères de décision

La logique de classification s'appuie sur des règles métier précises:

- **Classification automatique** si le score de confiance dépasse 0,80
- **Validation manuelle requise** pour les scores inférieurs à ce seuil
- **Mots-clés prioritaires** permettant d'orienter certains tickets vers des équipes spécifiques:
 - *"VPN", "accès distant"* → Équipe Infrastructure
 - *"boîte mail", "Outlook"* → Équipe Messagerie
 - *"connexion lente", "réseau"* → Équipe Réseau
- **Traitement spécifique pour les tickets multi-domaines** via une file commune pour validation humaine
- La **priorité d'un ticket** n'influence pas sa classification mais détermine son ordre d'affichage

Interfaces utilisateur et expérience de travail

Le système s'articule autour de quatre interfaces principales, chacune adaptée à un profil d'utilisateur spécifique:

1. Interface des agents de support

- Intégration native à l'outil de ticketing existant

- Affichage de la classification automatique et du score de confiance
- Indicateurs visuels distinguant les affectations automatiques et manuelles
- Option de signalement des classifications incorrectes

2. Interface des superviseurs

- Tableau de bord dédié à la supervision des classifications
- Outils de correction manuelle des tickets mal classés
- Visualisation des scores de confiance par catégorie et par période
- Rapports statistiques sur la performance du système

3. Tableau de bord supervisé

- Vue d'ensemble dynamique filtrable par:
 - Catégorie de ticket
 - Période temporelle
 - Seuil de confiance
 - Taux d'erreur de classification
- Indicateurs de performance du modèle en temps réel
- Alertes sur les anomalies de classification

4. Interface utilisateurs finaux

- Formulaire de soumission intuitif
- Affichage de la catégorie proposée
- Mécanisme simple de confirmation ou contestation
- Retour d'information sur le traitement de la demande

Cycle d'amélioration continue

Le système intègre un processus d'apprentissage permanent:

- **Collecte des corrections** effectuées par les superviseurs et utilisateurs
- **Analyse périodique des erreurs** de classification pour identifier les motifs récurrents
- **Ajustement du modèle** basé sur les retours d'expérience
- **Enrichissement continu** de la base de connaissances avec de nouveaux exemples validés

Cette approche garantit que le système s'améliore avec le temps et s'adapte aux évolutions des demandes, assurant ainsi une performance constamment optimisée pour le Groupe Asten.

4. Architecture technique

Vue d'ensemble de l'architecture

Le système intelligent de classification automatique des tickets repose sur une **architecture hybride** qui combine la flexibilité du cloud pour le traitement des modèles et la sécurité du

stockage local pour les données sensibles. Cette approche permet d'optimiser les performances tout en respectant les contraintes de confidentialité propres au Groupe Asten.

L'architecture s'articule autour de quatre composants principaux interconnectés:

| Composant | Description | Technologie |
|------------------------|---|-------------------------------------|
| Moteur d'analyse | Traitement du langage et classification | spaCy + règles métier |
| Backend applicatif | Gestion des flux et orchestration | Flask |
| Interfaces utilisateur | Visualisation et interaction | Streamlit |
| Stockage | Conservation des données et modèles | MongoDB / Système de fichiers local |

Flux de données et communications

Le système fonctionne selon un modèle de traitement par lots, avec des échanges de données structurées:

1. Les tickets sont exportés depuis l'outil de ticketing existant au format **CSV**
2. Ces fichiers sont déposés sur un **volume partagé**
3. Le moteur d'analyse traite ces fichiers et génère des résultats de classification
4. Les classifications sont réintégrées dans l'outil de ticketing

Ce choix d'architecture sans API externe à ce stade permet une intégration progressive et sécurisée, avec un minimum de modifications sur les systèmes existants.

Composants techniques détaillés

1. Moteur d'analyse et classification

Le cœur du système repose sur une combinaison de:

- **spaCy** pour le traitement du langage naturel, permettant:
 - La tokenisation du texte des tickets
 - L'extraction d'entités nommées et de concepts clés
 - L'analyse syntaxique pour la compréhension du contexte
- **Règles métier personnalisées** complétant l'analyse algorithmique:
 - Dictionnaires de termes techniques spécifiques au Groupe Asten
 - Pondérations adaptées aux priorités de l'entreprise
 - Logique de décision pour les cas particuliers

Ce moteur fonctionne en deux phases:

- Prétraitement linguistique via spaCy
- Application de la logique métier pour déterminer la classification finale

2. Backend applicatif

Développé en **Flask** (framework Python léger), le backend assure:

- La gestion du cycle de vie des tickets dans le système
- L'orchestration des traitements d'analyse
- La génération des fichiers de résultats
- Le monitoring et la journalisation des opérations

Sa légèreté permet un déploiement simplifié tout en offrant des performances adaptées au volume de tickets du Groupe Asten (estimé à 500-1000 tickets quotidiens).

3. Interfaces utilisateur

Plutôt que de créer une application séparée, le système s'intègre nativement à JIRA l'outil de ticketing déjà utilisé, gratuitement à :

- Des extensions personnalisées (Connect) pour l'affichage des classifications, pour l'affichage des catégories prédictives directement dans les tickets.
- Des composants d'interface additionnels pour les superviseurs, comme des panneaux de validation ou de reclassification.
- La réutilisation des flux de navigation familiers aux utilisateurs, workflows et transitions Jira existants pour conserver les habitudes des utilisateurs.

Cette approche minimise la courbe d'apprentissage et maximise l'adoption par les équipes.

4. Stockage des données

L'architecture de stockage hybride combine:

- **MongoDB** pour:
 - Les tickets analysés et leurs métadonnées
 - Les statistiques de performance du système
 - L'historique des classifications

Déploiement et infrastructure

L'infrastructure physique s'organise selon le modèle suivant:

- **Composants on-premise:**
 - Serveur d'application principal (Flask)
 - Base de données MongoDB
 - Stockage des données de production
 - Système de fichiers partagé

Exigences techniques

Pour garantir les performances du système, les ressources minimales requises sont:

| Composant | Spécification minimale |
|-----------------------|-----------------------------------|
| Serveur d'application | 8 CPU cores, 16 GB RAM |
| Base de données | 4 CPU cores, 8 GB RAM, 500 GB SSD |
| Stockage partagé | 1 TB avec redondance |
| Bande passante | 100 Mbps entre composants |

Sécurité de l'architecture

La conception du système intègre plusieurs niveaux de sécurité:

- **Isolation des données sensibles** sur l'infrastructure locale
- **Chiffrement des fichiers** lors des échanges entre composants
- **Journalisation exhaustive** de toutes les opérations de classification
- **Authentification renforcée** pour l'accès aux interfaces d'administration

Évolutivité et maintenance

L'architecture a été conçue pour permettre:

- **L'ajout progressif de nouvelles catégories** de classification
- **L'intégration future d'une API REST** pour des communications plus dynamiques
- Le **passage à une architecture temps réel** dans les versions ultérieures
- **L'amélioration continue des modèles** sans interruption de service

L'ensemble de ces caractéristiques techniques permet d'offrir un système robuste, évolutif et parfaitement adapté aux besoins actuels et futurs du Groupe Asten, tout en respectant les contraintes de déploiement et de sécurité de l'entreprise.

5. Gouvernance du projet

La gouvernance du projet de système intelligent de classification automatique des tickets repose sur une structure organisationnelle claire et des processus décisionnels définis, garantissant une mise en œuvre efficace et coordonnée.

Équipe de projet et responsabilités

L'équipe constituée pour ce projet réunit des compétences complémentaires permettant de couvrir l'ensemble des aspects techniques et fonctionnels:

| Membre | Fonction | Responsabilités |
|--------|----------------|---|
| Ulrich | Chef de projet | <ul style="list-style-type: none">• Cadrage global du projet• Coordination des différentes parties prenantes• Suivi des délais et des |

| | | |
|-------------|--------------------------------|---|
| | | livrables • Reporting vers la direction |
| Angea | Responsable Pôle Innovation IA | • Conception des algorithmes de NLP • Développement des modèles prédictifs • Validation technique des approches d'IA |
| Naomi | Responsable Support Client | • Définition des cas d'usage • Organisation et supervision des tests utilisateurs • Contribution à l'amélioration continue du système |
| Jean-Claude | Directeur technique | • Arbitrage sur les choix technologiques • Supervision des aspects sécurité • Validation de l'architecture technique |

Structure de pilotage

La gouvernance s'articule autour de trois niveaux de pilotage distincts :

1. **Comité opérationnel** - Réunion hebdomadaire
 - Participants : Ulrich, Angea, Naomi
 - Objectifs : Suivi d'avancement, résolution des problèmes opérationnels
 - Format: Synchronisation virtuelle via board partagé sur word cloud
2. **Comité technique** - Bimensuel
 - Participants : Ulrich, Angea, Jean-Claude
 - Objectifs: Validation des orientations techniques, arbitrage des solutions
 - Format: Réunion technique avec démonstrations et revues de code si nécessaire
3. **Comité de pilotage** - Mensuel
 - Participants: Tous les acteurs clés + représentants de la direction
 - Objectifs: Validation des jalons majeurs, allocation des ressources
 - Format: Présentation formelle avec tableaux de bord de suivi

Processus décisionnel et résolution de conflits

Le mode de décision adopté pour ce projet est **consultatif avec arbitrage final**:

- Les décisions courantes sont prises de manière collégiale lors des comités opérationnels

- Les décisions techniques font l'objet d'une consultation des experts concernés
- En cas de désaccord, le chef de projet (Ulrich) a le dernier mot après consultation des responsables techniques
- Pour les décisions stratégiques impliquant des modifications substantielles (budget, périmètre), l'arbitrage est soumis au comité de pilotage

Outils de gouvernance et suivi

Pour assurer la transparence et l'efficacité du pilotage, plusieurs outils sont déployés:

- **Board partagé sur Word Cloud:** tableau de bord commun permettant le suivi en temps réel de l'avancement des tâches
- **Référentiel documentaire** centralisé pour stocker les spécifications et comptes-rendus
- **Tableau de bord de métriques** pour suivre la performance du système de classification
- **Registre des risques** maintenu par le chef de projet et mis à jour régulièrement

Gestion des changements

Le processus de gestion des modifications est structuré comme suit:

1. Soumission d'une demande de changement documentée
2. Analyse d'impact technique et fonctionnelle
3. Évaluation des conséquences sur les délais et ressources
4. Décision d'acceptation ou de refus par le niveau approprié:
 - Impact mineur: validation par le chef de projet
 - Impact moyen: comité technique
 - Impact majeur: comité de pilotage

Communication et reporting

La communication s'organise de manière régulière et multi-canal:

- Rapports d'avancement hebdomadaires diffusés à l'ensemble des parties prenantes
- Points d'information réguliers aux utilisateurs finaux pour préparer l'adoption
- Communications spécifiques aux jalons clés du projet
- Présentation mensuelle de l'état d'avancement au comité de pilotage

Gestion de la qualité

La supervision de la qualité est intégrée à la gouvernance avec:

- Des critères d'acceptation clairement définis pour chaque livrable
- Des revues techniques systématiques pour les composants critiques
- Des tests utilisateurs supervisés par Naomie pour valider l'expérience utilisateur
- Des mesures de performance du système évaluées en continu

Cette structure de gouvernance a été conçue pour apporter agilité et rigueur au développement du système intelligent de classification, en s'assurant que les différentes expertises sont correctement représentées tout en maintenant une chaîne décisionnelle claire et efficace.

6. Indicateurs de performance (Estimatif)

Les indicateurs de performance (KPI) de notre système intelligent de classification automatique constituent le cadre d'évaluation qui permettra de mesurer objectivement l'efficacité et la valeur ajoutée apportée au Groupe Asten. Ces métriques ont été sélectionnées pour refléter à la fois la performance technique du système et son impact opérationnel sur les processus de support client.

Tableau de bord des indicateurs clés

| Indicateur | Objectif | Fréquence | Source de données | Seuil d'alerte |
|-----------------------------------|----------|--------------|---------------------------------------|----------------|
| Taux de classification correcte | > 85% | Hebdomadaire | Exports Jira avec validation humaine | < 80% |
| Réduction du temps de traitement | 30% | Mensuelle | Comparatif avant/après implémentation | < 20% |
| Taux de contestation utilisateur | < 10% | Hebdomadaire | Retours utilisateurs via interface | > 15% |
| F1-score du modèle | > 0,80 | Mensuelle | Évaluation technique du modèle | < 0,75 |
| Score de satisfaction post-ticket | > 4/5 | Continue | Questionnaires automatiques | < 3,5/5 |

Mesure de la précision technique

La performance technique du système sera évaluée principalement via deux indicateurs complémentaires:

- Le **taux de classification correcte** représente la proportion de tickets automatiquement assignés à la bonne catégorie sans intervention humaine. La mesure hebdomadaire effectuée à partir des exports Jira permet un suivi régulier et réactif de cette métrique centrale.

- La **F1-score** est une mesure technique combinant précision et rappel du modèle de classification. Calculée mensuellement, elle offre une vision plus complète de la performance algorithmique en prenant en compte:
 - Les classifications correctes (vrais positifs)
 - Les erreurs d'attribution (faux positifs)
 - Les classifications manquées (faux négatifs)

Cette double approche assure une évaluation à la fois accessible pour les décideurs et précise pour les équipes techniques.

Évaluation de l'impact opérationnel

L'efficacité opérationnelle du système est mesurée par trois indicateurs orientés résultats:

- La **réduction du temps de traitement moyen** quantifie directement le gain d'efficacité apporté par le système. L'objectif ambitieux de 30% reflète notre conviction dans la capacité transformative du système.
- Le **taux de contestation utilisateur** sert de garde-fou qualitatif, permettant de détecter rapidement les éventuels problèmes de classification non captés par les métriques purement techniques. Un taux inférieur à 10% sera considéré comme satisfaisant.
- Le **score de satisfaction post-ticket** complète l'évaluation en intégrant la perception subjective des utilisateurs finaux, élément crucial pour l'acceptation à long terme du système.

Processus de suivi et d'amélioration continue

Le suivi des indicateurs s'organise autour d'un processus structuré:

1. **Collecte automatisée** des données depuis les systèmes sources (principalement Jira)
2. **Analyse hebdomadaire** lors du comité opérationnel
3. **Revue mensuelle approfondie** en comité technique avec analyse des tendances
4. **Présentation trimestrielle** des résultats consolidés au comité de pilotage

Pour les indicateurs passant sous les seuils d'alerte définis, aucune action corrective n'est automatiquement déclenchée pour les variations mineures et ponctuelles. En revanche, une persistance sous le seuil durant deux mesures consécutives entraînera:

- Une analyse des causes racines
- Une proposition d'actions correctives
- Un suivi renforcé jusqu'au retour à la normale

Visualisation et accessibilité des indicateurs

Un tableau de bord dynamique sera mis à disposition dans l'interface administrateur permettant:

- La **visualisation graphique** de l'évolution temporelle des différents KPI
- Le **filtrage par période** pour les analyses de tendance

- La **comparaison avec les objectifs** pour une évaluation immédiate des performances
- L'**export des données brutes** pour des analyses personnalisées

Des rapports automatisés seront également générés et distribués aux parties prenantes concernées selon la fréquence définie pour chaque indicateur.

Évolution des indicateurs et méthodes de mesure

Les méthodes de mesure des indicateurs évolueront en trois phases:

- **Phase 1 (Lancement)**: Mesures manuelles avec validation humaine systématique
- **Phase 2 (Stabilisation)**: Semi-automatisation des mesures avec contrôles aléatoires
- **Phase 3 (Maturité)**: Automatisation complète avec audits périodiques

Cette approche progressive permettra d'affiner la fiabilité des mesures tout en réduisant progressivement la charge opérationnelle liée au suivi.

Benchmark et positionnement sectoriel

Pour contextualiser les performances du système, les indicateurs seront régulièrement comparés à:

- L'état de l'art académique pour les systèmes de classification
- Les benchmarks sectoriels disponibles pour des systèmes similaires
- L'historique interne de performance pour identifier les tendances d'amélioration

Cette démarche comparative permettra d'évaluer objectivement la position du Groupe Asten en matière d'automatisation de la classification des tickets et d'identifier les opportunités d'amélioration continue.

7. Sécurité et confidentialité

La sécurité et la confidentialité des données sont des piliers fondamentaux du projet, en cohérence avec les exigences internes du Groupe Asten et le cadre légal (notamment le RGPD). Le traitement de données sensibles, notamment les messages clients et les informations personnelles, nécessite une attention rigoureuse à chaque étape du cycle de vie du système.

Principes de conception sécuritaire

- **Sécurité dès la conception ("Security by design")** : L'ensemble des composants techniques a été sélectionné en intégrant les critères de sécurité dès la phase d'architecture.
- **Stockage localisé** : Les données sensibles ne transitent pas par des environnements cloud tiers. Le stockage et les traitements se font en grande partie sur l'infrastructure interne de l'entreprise.
- **Contrôle d'accès granulaire** : Une gestion fine des rôles est mise en place, avec authentification renforcée (SSO + MFA) pour les profils critiques (admin, superviseur).

- **Chiffrement des données** : Données chiffrées au repos (AES-256) et en transit (TLS 1.2 minimum). Les journaux système sont également protégés.

Conformité RGPD

- **Minimisation des données** : Seules les informations strictement nécessaires à la classification sont traitées.
- **Anonymisation** : Le contenu textuel peut être anonymisé lors des phases de test ou d'évaluation.
- **Traçabilité** : Chaque action de classification est journalisée. Les historiques sont conservés pour audit.
- **Droits des utilisateurs** : Des mécanismes sont prévus pour garantir le droit d'accès, de rectification, d'opposition et d'effacement des données traitées.

Des audits internes réguliers, couplés à des tests d'intrusion programmés, assureront la robustesse continue du système.

8. Annexes techniques

Les annexes suivantes viennent compléter le présent cahier des charges :

A. Exemples de tickets classifiés

- 10 cas représentatifs avec catégories attribuées et score de confiance
- Présentation des cas « limites » soumis à reclassification manuelle

B. Dictionnaire métier

- Liste des mots-clés associés à chaque catégorie métier
- Pondérations appliquées dans les règles de priorisation

C. Diagramme d'architecture

- Schéma des flux de données
- Positionnement des composants techniques (backend, moteur NLP, base de données)

D. Normes et protocoles

- Liste des normes de sécurité appliquées (ISO 27001, OWASP Top 10)
- Bonnes pratiques de développement sécurisé adoptées dans le code

E. Procédures de reprise

- Plan de continuité en cas de panne système
- Restauration à partir des sauvegardes (fréquence, durée estimée)

Ces annexes constituent la base de référence technique pour l'ensemble des intervenants du projet et seront régulièrement mises à jour au fil de l'avancement.