

**Modèle de dossier du candidat**

Titre RNCP “Expert en Informatique et Systèmes d’Information”

N° 36286 – Niveau 7 (Bac+5)

Ce modèle est basé sur le référentiel RNCP 36286. Il contient le sommaire minimal exigé pour cet examen. Il reprend tous les points du référentiel encadrés dans un projet unique de fn de cycle.

**1. Page de garde**





Sujet sous forme de

question

COUDRIER Georges

Référent Société:

PLESSIS Hubert

Patron

ESI-03 2023-2025

​ ​ ​ ​ ​ ​

1. **Remerciements**

Je tiens à remercier Hubert Plessis, mon tuteur en entreprise, pour son accompagnement et ses conseils tout au long de cette année. Sa disponibilité et son expérience ont été d’une grande aide pour moi dans ce projet.

Un grand merci également à mes collègues, Vivien et Sonnarin, qui ont toujours été là pour répondre à mes questions, m’encourager et partager leur savoir-faire. Travailler avec eux a été un vrai plaisir.

Enfin, je remercie tous mes professeurs pour leur enseignement et leur soutien pendant ma formation. Grâce à eux, j’ai pu développer les compétences nécessaires pour mener à bien ce projet.

Merci à vous tous !

1. **Résumé**
2. **Abstract**
3. **Sommaire**
4. **Liste des abréviations**
5. **Glossaire (défnition des termes utilisés)**
6. **Présentation personnelle**

**8.1 Introduction**

**8.1 Parcours académique et professionnel**

**8.1 Objectifs de la présentation**

**9. Présentation de la problématique**

**9.1 Contexte du projet**

**9.2 Description de la problématique**

**9.3 Objectifs du projet**

**9.4 Importance et impact de la problématique**

**9.5 Questions de recherche et hypothèses**

1. **Diagnostic général du Système d’Information et Veille technologique**

* **10.1 Diagnostic du Système d’Information**
* ​10.1.1 Observations générales
* ​10.1.2 Recommandations
  + - **10.2 Veille technologique**
      * ​10.2.1 Introduction
      * ​10.2.2 Méthodologie de la veille technologique
      * ​10.2.3 Résultats de la veille technologique
  + ​10.2.4 Recommandations et perspectives
  + ​10.2.5 Conclusion

1. **Besoins fonctionnels et Cahier des charges fonctionnel**

**11.1 Collecte et analyse des besoins**

* 11.1.1 Méthodologie de collecte des besoins
* 11.1.2 Analyse approfondie des besoins
* 11.1.3 Identifcation des contraintes et des exigences
* 11.1.4 Validation des besoins avec les parties prenantes

**11.2 Description des fonctionnalités requises**

* 11.2.1 Fonctionnalités essentielles
* 11.2.2 Fonctionnalités secondaires

**11.3 Analyse stratégique d’implémentation**

* 11.3.1 Introduction
* 11.3.3 Gestion des risques
* 11.3.4 Élaboration de la stratégie d’implémentation
* 11.3.5 Présentation des préconisations
* 11.3.6 Conclusion

**12. Implémentations techniques**

**12.1 Infrastructure et environnement de déploiement**

* 12.1.1 Infrastructure matérielle
* 12.1.2 Environnement de déploiement (cloud, on-premise, hybride)

**12.2 Technologies et outils utilisés**

* + 12.2.1 Algorithmes et frameworks
* 12.2.2 Outils de développement et de gestion de projet

**12.3 Surveillance et monitoring de la sécurité**

* 12.3.1 Outils de monitoring
* 12.3.2 Techniques de détection et de réponse aux incidents

**12.4 Plans de sauvegarde et de récupération**

* 12.4.1 Stratégies de sauvegarde des données
* 12.4.2 Plans de reprise après sinistre

**13. Gestion du projet et management**

**13.1 Élaboration du planning et gestion des ressources**

* 13.1.1 Création d'un planning détaillé
* 13.1.2 Allocation des ressources

**13.2 Suivi de l'avancement et gestion des risques**

* 13.2.1 Indicateurs de performance
* 13.2.2 Gestion des risques

**13.3 Management de l'équipe projet**

* 13.3.1 Encadrement et motivation
* 13.3.2 Développement des compétences

**13.4 Communication et coordination avec les parties prenantes**

* 13.4.1 Plan de communication
* 13.4.2 Engagement des parties prenantes

**13.5 Planifcation**

* 13.5.1 Introduction
* 13.5.2 Méthodologie de projet
* 13.5.3 Phases du projet
* 13.5.4 Planning détaillé (diagramme inclus)
* 13.5.5 Ressources et budget

**13.6 Note de cadrage du projet**

13.6.1 Contexte du projet

13.6.2 Objectifs et périmètre

13.6.3 Organisation et ressources

13.6.4 Budget et délais

13.6.5 Contraintes et risques

13.6.6 Modalités de pilotage et de communication

1. **Conception et Développement de l’Application Informatique**

**14.1 Conception**

* 14.1.1 Introduction
* 14.1.2 Décomposition de l’application
* 14.1.3 Fonctionnement de l’application
* 14.1.4 Données utilisées
* 14.1.5 Environnement informatique proposé
* 14.1.6 Spécifcations techniques

**14.2 Développement et Stratégie de Tests**

* + 14.2.1 Objectifs de la stratégie de tests
  + 14.2.2 Types de tests et outils utilisés
* 14.2.3 Plan de test détaillé
  + 14.2.4 Environnement de test

**14.3 Modalités de livraison**

14.3.1 Mise en service opérationnelle et vérifcation de service régulier (VSR)

14.3.2 Conclusion

1. **Construction et Développement d’un Modèle de « Big Data » (Option Big Data)**

**15.1 Enjeux du Big Data et l’analyse des données**

* 15.1.1 Introduction au Big Data

Le Big Data et l’IA apportent des bénéfices majeurs à NexaCore pour répondre efficacement à la problématique de churn et améliorer l’expérience utilisateur.

1. **Prise de Décision Plus Rapide** :  
   En exploitant les capacités analytiques des Big Data, NexaCore peut collecter et traiter en temps réel des données comportementales des clients, permettant des ajustements immédiats aux campagnes marketing et aux processus de fidélisation.
2. **Analyse Plus Exhaustive** :  
   Grâce au Big Data, NexaCore peut centraliser les interactions clients, les performances des campagnes et les tendances d’utilisation, offrant une vision complète et holistique pour des actions mieux ciblées.
3. **Précision des Recommandations** :  
   Les algorithmes d’IA appliqués aux données massives permettent de proposer des recommandations adaptées aux besoins spécifiques de chaque utilisateur, augmentant la pertinence des actions entreprises.
4. **Identification Proactive des Risques** :  
   Le traitement des Big Data permet de détecter précocement les signaux de désengagement client, grâce à des modèles prédictifs qui alertent sur les risques de churn.
5. **Personnalisation des Expériences** :  
   En combinant IA et Big Data, NexaCore peut segmenter les clients de manière précise et offrir des parcours utilisateur personnalisés, renforçant ainsi leur satisfaction et leur engagement.

* 15.1.2 Objectifs

L'objectif principal est de réduire le taux de churn en fidélisant les clients existants, car conserver un client coûte généralement moins cher que d'en acquérir un nouveau. Cette réduction a un impact direct sur :

* **L'augmentation des revenus récurrents** : Chaque client fidélisé contribue à la stabilité financière de l’entreprise.
* **L'amélioration de la réputation** : Un churn réduit reflète une satisfaction client accrue, attirant de nouveaux prospects grâce au bouche-à-oreille et aux témoignages positifs.
* **La diminution du coût d'acquisition client (CAC)** : Fidéliser les clients réduit la nécessité d'investir massivement dans la conquête de nouveaux clients pour compenser les pertes.

**Pourquoi repérer la probabilité de churn ?**

Identifier la probabilité de churn permet d’agir de manière préventive en :

* **Ciblant les clients à risque** : Une action spécifique peut être déclenchée (offres personnalisées, support renforcé) pour inverser leur désengagement.
* **Optimisant les ressources** : Au lieu de déployer des actions marketing générales, les efforts se concentrent sur les segments réellement à risque.
* **Améliorant l’expérience utilisateur** : Les clients identifiés comme insatisfaits peuvent bénéficier d’un accompagnement adapté, renforçant leur perception positive de NexaCore.

**Ce que cela apporte à l’entreprise**

* **Augmentation de la valeur vie client (CLV)** : Un client fidèle reste plus longtemps, générant un chiffre d’affaires accru.
* **Réduction des pertes de revenus** : Chaque client retenu contribue à stabiliser les revenus mensuels.
* **Insights stratégiques** : Les données collectées pour analyser le churn peuvent aussi être utilisées pour identifier des tendances et ajuster les offres ou services.

**Pourquoi utiliser l’IA plutôt qu’une personne ?**

1. **Scalabilité** : L’IA peut analyser des millions de données simultanément, une tâche impossible pour une personne.
2. **Précision** : Les algorithmes d’IA détectent des corrélations complexes et des patterns subtils dans les données, souvent invisibles à l'œil humain.
3. **Temps de réponse** : L’IA travaille en temps réel, permettant des actions immédiates pour contrer les risques de churn.
4. **Automatisation** : Une fois les modèles prédictifs en place, l’IA peut fonctionner de manière autonome, réduisant les coûts opérationnels liés à l’analyse manuelle.

* 15.1.3 Données disponibles

Dans le cadre du projet NexaCore, le dataset utilisé offre une vue complète des comportements d'achat et des insights sur les ventes, adapté aux besoins de prédiction de churn et d'optimisation des stratégies de fidélisation. Les données disponibles sont structurées autour des dimensions suivantes :

**1. Détails des Clients**

* **Identifiants Clients** : Chaque client possède un ID unique permettant un suivi précis de ses interactions.
* **Attributs d’Achat** : Fréquence d'achat, panier moyen et valeur client à vie (Lifetime Value).

**2. Informations Produits**

* **Identifiants Produits** : ID unique pour chaque produit.
* **Catégories Populaires** : Liste des catégories les plus achetées par les clients, permettant d'identifier les préférences générales.

**3. Métriques de Ventes**

* **Valeur Moyenne des Commandes** : Indicateur de la dépense typique des clients.
* **Valeur Client à Vie (CLV)** : Montant total attendu qu’un client dépensera sur toute la durée de sa relation avec l’entreprise.
* **Fréquence d’Achat** : Nombre de transactions réalisées par client sur une période donnée.

**4. Insights Comportementaux**

* **Temps entre Achats** : Intervalle moyen entre deux transactions d’un même client.
* **Moments d’Achat Préférés** : Analyse des heures et des jours où les achats sont les plus fréquents.
* **Catégories Fréquemment Achetées** : Produits ou services les plus populaires par client.

**5. Insights Géographiques**

* **Régions** : Les données couvrent différentes zones géographiques, notamment :
  + Amérique du Nord
  + Europe
  + Asie
  + Amérique du Sud
* Ces informations aident à adapter les stratégies marketing en fonction des spécificités locales.

**6. Saisonnalité**

* **Dates de Vente en Pic** : Identification des périodes de vente les plus intenses, souvent liées aux fêtes ou aux événements promotionnels.
* **Saisons Associées** : Correspondance des tendances de vente avec des périodes spécifiques, comme Noël ou la rentrée.

**7. Stratégies de Rétention**

* **Méthodes Appliquées** : Informations sur les campagnes de fidélisation déjà en place, telles que :
  + Remises ou promotions personnalisées.
  + Programmes de fidélité.
  + Campagnes d’e-mails ciblés pour les clients à risque.

**8. Probabilité de Churn**

* **Métriques de Probabilité** : Score calculé indiquant le risque d’attrition pour chaque client, basé sur des modèles prédictifs.

**Description Globale**

Le dataset de NexaCRM est conçu pour des analyses détaillées et des modèles prédictifs, en combinant des données démographiques, comportementales et transactionnelles. Il offre une base solide pour :

* **L’exploration des données (EDA)** : Identifier les tendances et segments clients clés.
* **Les modèles de segmentation** : Créer des groupes homogènes pour des actions marketing ciblées.
* **La prédiction du churn** : Anticiper les comportements à risque et agir en conséquence.

Ces données permettent à NexaCore de fournir une vision précise et contextualisée des clients, renforçant les capacités de personnalisation et d’engagement pour réduire le churn efficacement.

\* NexaCRM est une entreprise fictive, les données ont été récupéré sur Kaggle  
- <https://www.kaggle.com/datasets/imranalishahh/sales-and-customer-insights/data>

* 15.1.4 Analyse des données

Les données fournies permettent de réaliser une exploration approfondie afin de dégager des tendances et des corrélations entre différents comportements clients et leur risque de churn. Voici les principaux éléments d’analyse :

1. **Comportement d'Achat** :
   * Les **fréquences d’achat** et le **temps moyen entre les transactions** sont des indicateurs clés pour détecter les clients en phase de désengagement. Une augmentation des intervalles entre les achats peut signaler une baisse d'intérêt.
   * La **valeur client à vie (CLV)** peut révéler les clients stratégiques pour lesquels des actions spécifiques de rétention doivent être mises en œuvre.
2. **Données Comportementales** :
   * Les **moments d'achat préférés** et les catégories de produits régulièrement achetées permettent d’identifier des habitudes spécifiques. Un changement soudain dans ces habitudes peut indiquer une probabilité accrue de churn.
   * L’analyse des **données saisonnières** peut montrer si certains clients ont un comportement d’achat lié à des périodes spécifiques. Une absence d’activité durant ces périodes pourrait être un signal d’alerte.
3. **Interactions avec les Stratégies de Rétention** :
   * Les réponses aux campagnes de fidélisation, comme l’engagement avec les e-mails ou les remises, fournissent des insights sur la probabilité de réengagement. Les clients qui n'interagissent pas avec ces initiatives sont plus à risque.
4. **Insights Géographiques** :
   * Les comportements d'achat varient souvent selon les régions. L’analyse des performances par région peut révéler des tendances spécifiques (ex. : zones où le churn est plus élevé).
5. **Score de Probabilité de Churn** :
   * Les scores préexistants calculés à partir des données comportementales et transactionnelles permettent d’identifier directement les clients à risque pour des interventions ciblées.

**Prédiction du Churn**

Ces données sont essentielles pour entraîner des modèles d’apprentissage automatique qui détectent les clients susceptibles de churner. Voici pourquoi elles sont particulièrement adaptées :

* **Richesse des Attributs** : Les données combinent des dimensions comportementales (fréquence, préférences), transactionnelles (valeur des achats) et contextuelles (géographie, saisonnalité). Ces multiples perspectives enrichissent le modèle.
* **Corrélations Complexes** : Les modèles comme les forêts aléatoires ou les réseaux neuronaux peuvent identifier des corrélations complexes entre des signaux faibles (ex. : un léger allongement des temps entre achats) et le churn.
* **Données Historiques et Temps Réel** : La combinaison de données historiques pour l’entraînement et de données en temps réel pour la prédiction permet une réactivité accrue.

**Pourquoi ces Données Permettent de Prédire le Churn**

1. **Comportements Répétitifs** : Les clients qui montrent des tendances régulières mais les modifient subitement sont souvent à risque.
2. **Segmentation Fine** : La segmentation par préférences ou par régions permet de personnaliser les actions de rétention.
3. **Évaluation des Efforts de Rétention** : L’analyse des interactions avec les campagnes révèle l’efficacité des actions en place et leur impact sur le churn.

**15.2 Analyse de la problématique**

* 15.2.1 Identifcation des défis et des besoins spécifques
* 15.2.2 Analyse des objectifs à atteindre à travers le modèle Big Data

**15.3 Données**

* 15.3.1 Identifcation des sources de données pertinentes
* 15.3.2 Sélection des outils et des méthodes appropriés
* 15.3.3 Collecte et préparation des données
* 15.3.4 Processus d'importation et de stockage des données

**15.4 Procédures de structuration**

* 15.4.1 Organisation des données pour l'analyse
* 15.4.2 Normalisation, nettoyage et transformation des données

**15.5 Modèles**

* 15.5.1 Choix et justifcation des modèles
* 15.5.2 Mise en œuvre et évaluation des algorithmes

**15.6 Analyses des résultats et recommandations**

* 15.6.1 Analyses des résultats
* 15.6.2 Recommandations

**16. Analyses et Perspectives**

**16.1 Analyse des résultats et des performances**

* 16.1.1 Évaluation des objectifs atteints
* 16.1.2 Analyse des performances

**16.2 Retour d'expérience et leçons apprises**

* 16.2.1 Identifcation des succès et des défs
* 16.2.2 Capitalisation des connaissances

**16.3 Perspectives et recommandations**

**17. Conclusion et remerciements**

**17.1 Synthèse des principaux résultats et impacts**

* 1. **Remerciements**

**18. Bibliographie / Webographie**

1. **Annexes**