



Département : Génie Logistique et Transport

Mémoire de fin d'étude en vue de l'obtention du diplôme

## D'INGENIEUR d'État

-Filière-

**Ingénierie des Transports**

-Spécialité-

**Ingénierie des Transports**

-Thème-

# Smart Staffing Machine Learning pour l'optimisation de la planification des équipages

-Cas Air Algérie-

Réalisé par

**TALMAT Amin Mohamed**

**TAIBAOUI Mohammed**

### Les membres de Jury :

BOUZID Mouaouia Cherif	MCA	Président
REZKI Nafissa	MCB	Promoteur
ANNAD Oussama	MCB	Examinateur
ZIDANI Omar	MCA	Examinateur

Alger, le 01/07/24

## Dédicaces

*Nous souhaitons dédier ce travail à nos parents bien-aimés, qui ont toujours été présents à nos côtés, nous soutenant et nous encourageant tout au long de notre parcours académique et personnel. Leur amour inconditionnel et leurs sacrifices ont été la source de notre motivation et de notre persévérance.*

*À nos familles, pour leur soutien inébranlable et leur compréhension durant les moments de stress et de travail intense. Leur présence et leurs encouragements nous ont permis de surmonter les défis que nous avons rencontrés.*

*À nos amis proches, pour leur constant encouragement, leur amitié précieuse et leur patience. Leur soutien moral et leurs conseils ont été d'une grande aide tout au long de ce projet.*

*À tous ceux qui ont cru en nous et ont contribué, de près ou de loin, à la réalisation de ce travail, nous exprimons notre profonde gratitude.*

## **Remerciement**

Nous tenons à exprimer notre profonde gratitude envers Dieu tout-puissant pour nous avoir donné la force, le courage et la patience nécessaires à l'élaboration de ce travail.

Nos remerciements vont également à nos parents pour leurs sacrifices et leur soutien inconditionnel tout au long de notre parcours.

Nous adressons notre sincère gratitude à notre enseignante, Mme Rezki Nafissa, pour son encadrement, son aide précieuse et sa patience tout au long de l'année.

Nous remercions particulièrement nos encadrants, Mme N. Khelil, M. A. Hannou et M. N. Adjir, pour l'attention, le soutien et l'aide qu'ils nous ont apportés durant notre stage.

Nous exprimons également notre gratitude aux membres du jury pour l'honneur qu'ils nous font en participant à l'évaluation de notre mémoire de fin d'études et pour l'attention qu'ils y ont portée.

Enfin, nous adressons nos remerciements à toutes les personnes qui ont contribué, de près ou de loin, à la réalisation de ce travail avec générosité et bienveillance

## Table des matières

Les abréviations.....	8
Introduction générale.....	9
Chapitre I : Etude de l'existant.....	11
1.Introduction .....	12
2.Historique de l'entreprise : .....	12
3.Les moyens de la compagnie :.....	14
3.1. Les moyens matériels : .....	14
3.2Les moyens humains : .....	15
4.Organisation de l'entreprise : .....	15
4.1. La Direction des Programmes : .....	16
4.2. Département Frètement et Affrètement.....	17
4.3. Département Administration et Finances .....	18
4.4. Cellule AIMS et Inspection Système.....	18
4.5. Cellule Qualité.....	18
4.6. Sous-Direction Conception Programme :.....	18
4.7. Sous-Direction Contrôle et Statistiques :.....	19
4.8. Sous-Direction Régulation Programme :.....	19
5. Système d'information :.....	20
6. Destinations :.....	20
7. Objectifs de la compagnie : .....	21
8. Missions de la compagnie : .....	21
9. La planification des ressources dans les campaniens aériens.....	23
9.1. Processus .....	23
9.2. Logiciel AIMS.....	24
9.3. Méthode de Travail Crew route :.....	25
10. Système de Gestion de Base des données .....	27
11. L 'importance d'une gestion efficace de la planification pour Air Algérie .....	28
12. Conclusion :.....	28
Chapitre II : Etat de l'art.....	29
1.    Introduction .....	30
2.    Intelligence Artificielle.....	30
3.    Historique .....	31
4.    Domaines d'application de l'IA.....	32

5.	Machine Learning.....	32
6.	Deep Learning .....	33
7.	Réseau neuronal artificiel .....	33
8.	Les Types de Machine Learning .....	33
8.1.	Apprentissage supervisé .....	33
8.2.	Apprentissage non supervisé .....	35
8.3.	Apprentissage semi supervisé.....	36
8.4.	Reinforcement Learning :.....	36
9.	Processus de Machine Learning : .....	37
10.	Machine Learning et le transport aérien .....	37
10.1.	Prévision de la demande de voyageurs.....	38
10.2.	Gestion des trafics aérien.....	38
10.3.	Optimisation des revenues.....	38
10.4.	Maintenance prédictive .....	38
10.5.	Planification des opération aériennes .....	39
11.	Evolution des approches utilisées dans la planification des équipages : .....	42
12.	Conclusion :.....	43
	<b>Chapitre III : Modélisation .....</b>	<b>44</b>
1.	Introduction .....	45
2.	Ressource et Données.....	45
2.1.	Explication des codes du programme.....	46
3.	Justification de la proposition d'une solution de Machine Learning (ML) :.....	46
4.	Prétraitement des données : .....	48
5.	Présentation des outils utilisés.....	48
5.1	Python.....	48
5.3	Les bibliothèques utilisées :.....	49
5.3	Google Colab :.....	50
6.	Les différents Methods et approches utilisé : .....	51
6.1	ANN : .....	51
6.1.6.	.....	55
6.2.	Clustering : .....	56
7.	L'algorithme suggéré : .....	61
7.1.	Description de code : .....	62
7.2.	Schéma descriptif du Code.....	66
7.3.	L'affichage finale de résultat.....	67

8. Discussion des résultats.....	68
9. Discussion de notre travail : .....	70
10. Conclusion :.....	71
Conclusion Générale : .....	72
Bibliographie .....	75

## Listes des figures et des tableaux :

Figure 1:Organigramme de la Direction Générale de la Compagnie [7].....	16
Figure 2: Organigramme de la Direction des Programmes [7].....	17
Figure 3:Les étapes de planification] .....	23
Figure 4:Logo de AIMS .....	25
Figure 5:Illustration de l'interface de logiciel AIMS .....	25
Figure 6: Illustration de l'interface de logiciel AIMS .....	26
Figure 7 :Schéma de système SGBD pour une Compagnie airien [6] .....	27
Figure 8: Relation entre les défirent branches de IA [11] .....	30
Figure 9:illustration de l'apprentissage humaine et automatique [11] .....	32
Figure 10:Processus de ML [11] .....	37
Figure 11:Capture des donnée fournée.....	45
Figure 12 Notre Base de Donnée.....	48
Figure 13: Illustration d'un neurone Source : Wikipédia.....	51
Figure 14: Schéma de neurone humaine Source : Wikipédia.....	51
Figure 15: Schéma de neurone artificiel Source : Wikipédia.....	52
Figure 16:le Dendrogramme.....	60
Figure 17: Capture des données avant traitement.....	62
Figure 18: Capture des données après traitement .....	62
Figure 19: Capture des ensembles .....	63
Figure 20:Capture des résultats d'appariement initiale.....	63
Figure 21:Capture de nouvelle dataframe .....	64
Figure 22:Capture de dataframe sbnb.....	64
Figure 23; Capture de code de traitement des résultats .....	65
Figure 24: Capture des vols non couvert .....	65
Figure 25: Organigramme de l'algorithme .....	66
Figure 26:L'affichage finale de résultat .....	67
Figure 27:Illustration d'un scénario de vol entrant .....	67
Figure 28: Example de Différences dans la structure des rotations.....	69
Tableau 1:Historique de la compagnie. [2] .....	12
Tableau 2: La flotte de la compagnie [5].....	14
Tableau 3: La répartition du personnel de la compagnie [7].....	15
Tableau 4 : les évènements principaux dans le développement de l'IA .....	31
Tableau 5:Résumé des travaux similaires .....	39
Tableau 6: Exemple de l'encodage one-hot .....	53
Tableau 7:Tables des résultats.....	68

## **Les abréviations.**

DGAA	Direction Générale d'air Algérie
DP	Département programme
CA	Compagnie Aériens
ML	Machine Learning
ANN ou NN	Réseau de neurones artificiel
CNN	Réseau de neurones convolutifs
RNN	Le réseau de neurones récurrents
PN	Personnel navigant
PNT	Personnel navigant technique
PNC	Personnel navigant commercial
CPP	Crew Pairing Problem
SVM	Support Vector Machine
TPU	Tensor Processing Unit
GPU	Graphics Processing Unit
KNN	K-Nearest Neighbors

## **Introduction générale.**

Dans un contexte où l'intelligence artificielle révolutionne de nombreux domaines, et avec la croissance continue du trafic aérien, les coûts des équipages représentent la deuxième source de dépenses après les coûts du carburant [1]. L'industrie du transport aérien cherche donc à exploiter cette opportunité pour améliorer la planification des ressources, un élément clé dans la gestion des compagnies aériennes.

À mesure que l'industrie aérienne mondiale croît, la complexité des problèmes de planification, tels que la conception des horaires, l'affectation de la flotte, l'appariement des équipages et l'affectation des équipages, s'accroît également. Malgré les capacités de calcul avancées disponibles aujourd'hui, les administrateurs doivent trouver des solutions innovantes pour améliorer de manière significative l'efficacité de la planification, ce qui nécessite de nouveaux algorithmes pour résoudre ces problèmes complexes.

La revue de la littérature sur ce sujet montre que des études ont été réalisées sur l'application des algorithmes d'apprentissage automatique dans les compagnies aériennes, et que ces applications ont conduit à des améliorations significatives. Notre travail se pose donc une première question : comment intégrer efficacement le machine learning pour résoudre ce problème ? En appliquant un modèle d'apprentissage automatique, une deuxième question se pose : comment le machine learning peut-il nous rapprocher de l'optimalité en termes de temps et de coût dans la planification ?

Ce mémoire se concentre sur le défi de l'appariement des équipages, une sous-partie de la planification des équipages. Il s'agit de mettre en place un ensemble de rotations (jumelages) qui commencent et se terminent dans une même base, avec une pause de repos entre chaque vol actif, afin que chaque vol soit opéré par un équipage, tout en respectant les réglementations et les contraintes liées au personnel navigant.

L'objectif de ce travail est d'utiliser les modèles d'apprentissage automatique pour améliorer l'efficacité de l'utilisation des équipages et réduire le temps de traitement de cette opération au sein de l'entreprise.

Nous avons divisé notre travail en trois chapitres distincts :

- **Chapitre 1 :** Nous présenterons la compagnie aérienne Air Algérie et le processus de planification des équipages. Nous aborderons les défis rencontrés dans cette planification et les raisons pour lesquelles nous avons choisi d'utiliser l'apprentissage automatique.
- **Chapitre 2 :** Nous traiterons de l'intelligence artificielle, en mettant l'accent sur l'apprentissage automatique et ses différents algorithmes. Nous explorerons également les applications du machine learning dans le transport aérien et l'évolution des méthodes utilisées dans la planification.
- **Chapitre 3 :** Nous nous concentrerons sur la modélisation et la discussion des résultats obtenus.

Enfin, nous conclurons notre projet de fin d'études par une conclusion générale de nos travaux et des résultats obtenus.

# **Chapitre I : Etude de l'existant**

## **1.Introduction**

Avant de commencer notre étude, il est important de présenter l'environnement de travail, air Algérie est une entreprise publique économique organisée en société par actions (SPA) Dun capitale de 60.000.000.000 DZD, elle assure le transport pour plus 5 millions passagers (2019) vers 77 destinations (4 continents, 44 dessertes internationales et 33 domestiques) par une flotte de 56 appareils qui répondent aux normes de sécurité internationales.

## **2.Historique de l'entreprise :**

*Tableau 1:Historique de la compagnie. [2]*

1947	Création de la C.G.T. (Compagnie Générale de Transport) avec un réseau principalement orienté vers la France. [2]
1962	Recouvrement de l'indépendance de l'Algérie. [2]
1963	L'État prend 51% du capital d'Air Algérie. [2]
1970	L'État porte sa participation au capital d'Air Algérie à 83%. [2]
1973	L'État décide d'intégrer à Air Algérie la Société de Travail Aérien (S.T.A.). [2]
1974	L'État porte sa participation à 100% en rachetant les 17% encore détenus par Air France. [2]
1983	L'entreprise est scindée en deux entités distinctes, l'une pour les lignes intérieures (IAS) et l'autre pour les lignes internationales. [2]
1984	Les deux entités citées ci-dessus sont à nouveau fusionnées en une seule entreprise à laquelle revient la charge de la gestion des aérogares. [2]
1987	Air Algérie est déchargée de la gestion aérogare. [2]
1997	Air Algérie devient une société par actions avec un capital de 2,5 milliards DA. [2]
1998	Libéralisation du transport aérien. [2]
2000	Le capital d'Air Algérie est porté à 6 milliards de dinars. [2]
2001	Le capital d'Air Algérie est porté à 14 milliards de dinars. [2]
26/3/2005	Le capital d'Air Algérie est porté à 26 milliards de dinars. [2]

30/9/2005	Le capital d'Air Algérie est porté à 29 milliards de dinars. [2]
2006	Le capital d'Air Algérie est porté à 37 milliards de dinars. [2]
2007	Ouverture de la ligne directe Alger - Montréal, le capital d'Air Algérie est porté à 43 milliards de dinars. [3]
2009	Ouverture d'une ligne directe Alger - Pékin. [3]
2/4/2015	Air Algérie prend livraison de trois Airbus A330-200 commandés en 2014. [3]
1/2019	Air Algérie a trouvé un accord pour la vente de ses trois Boeing 767-300 à une entreprise privée américaine. Le dernier a quitté l'aéroport d'Alger Houari Boumediene le 27 janvier 2019 en direction des États-Unis. [2]
2/7/2019	Air Algérie lance le paiement en ligne par carte CIB et Gold d'Algérie Poste [4]. Elle lance aussi l'application Air Algérie sur IOS et Android depuis mars 2019. [3]
2020	Air Algérie a décidé de suspendre tous ses vols nationaux et internationaux en raison de la pandémie de Covid-19. [3]
1/6/2021	Air Algérie a repris partiellement ses vols vers la France (Paris et Marseille), l'Espagne (Barcelone), la Turquie (Istanbul) et la Tunisie. [3]
2/2022	Air Algérie annonce l'intention d'augmenter son offre ouvrant plus de 108 nouvelles destinations. [3]
5/2023	Air Algérie commande chez Boeing huit Boeing 737 Max 9, avec une livraison prévue en 2027. [3]
1/6/2023	Air Algérie commande chez Airbus 5 Airbus A330-900 neo et 2 Airbus A350-1000. [3]

### **3.Les moyens de la compagnie :**

#### **3.1. Les moyens matériels :**

La flotte d'Air Algérie se compose de 56 appareils modernes répondant aux normes de sécurité internationales exploité pour le transport des passagers et le cargo qui seront représenté dans le tableau suivant : [5]

*Tableau 2: La flotte de la compagnie [5]*

Illustration	Nom	Nombre	Capacité
	A330 -202	8	263
	B737-800	25	162 / 146
	B737-700	2	112
	B737-600	5	101
	ATR.72-212 A	15	66
	Hercule L 100-30	1	20 tonnes

### **3.2Les moyens humains :**

Air Algérie a fait un investissement significatif dans la formation de son personnel, ce qui a abouti à une équipe entièrement algérienne aujourd'hui. Elle est reconnue pour sa sécurité grâce à son personnel de conduite d'aéronefs qualifié. La maintenance de sa flotte est assurée par son propre personnel, garantissant ainsi une qualité constante. De plus, elle possède un centre hôtelier qui répond aux besoins en Algérie et offre une assistance aux compagnies aériennes étrangères.

Air Algérie compte un effectif de 9590 employés, les catégories de son personnel se répartissent comme suit : [6]

*Tableau 3: La répartition du personnel de la compagnie [7]*

Catégorie	Nombre
Personnel au sol	8140
Personnel navigant technique PNT	520
Personnel navigant Commercial PNC	930

### **4.Organisation de l'entreprise :**

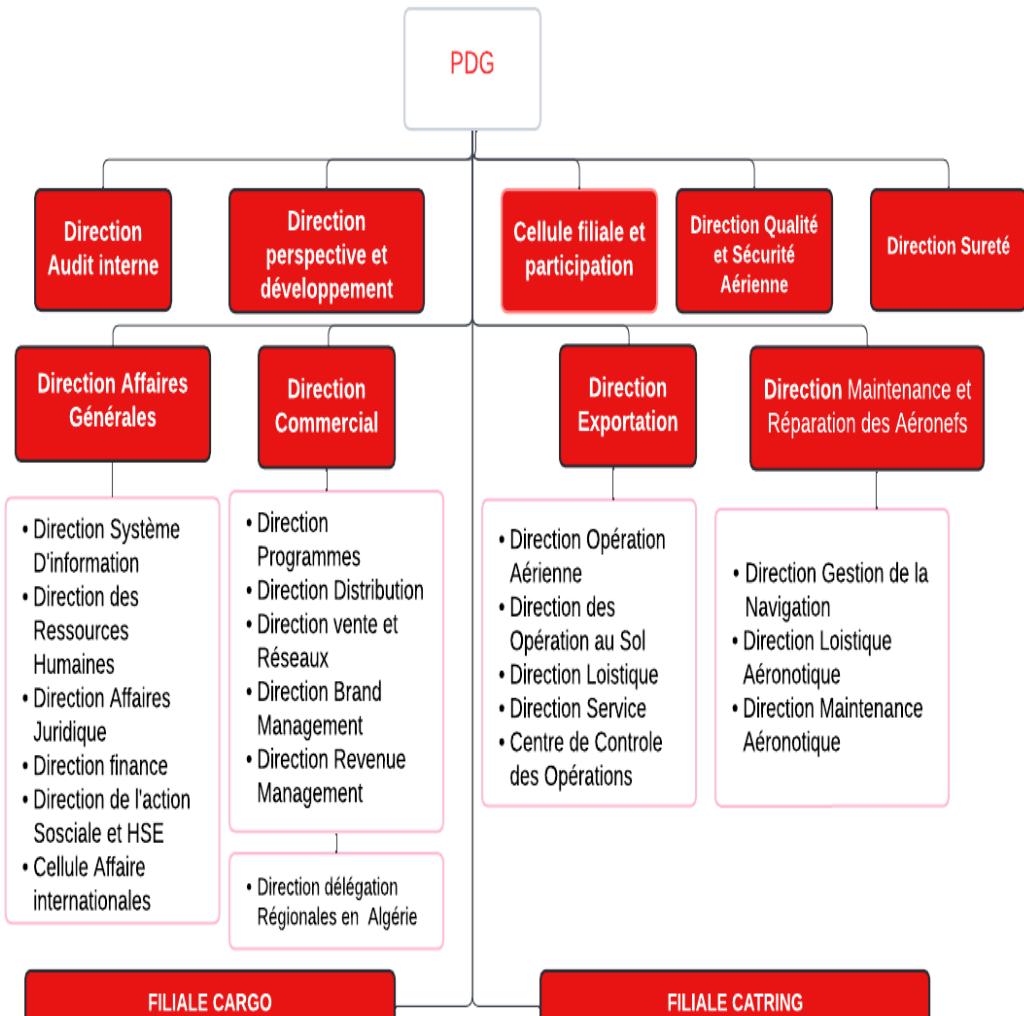
En Algérie, la compagnie aérienne nationale d'Algérie, Air Algérie, est administrée par la direction générale, également appelée D.G.A.A.

Il y a un Directeur Général et plusieurs départements au sein de la D.G.A.A. et chacun est responsable d'une fonction spécifique de l'entreprise.

La D.G.A.A. est chargée de l'application des politiques et des décisions établies par la direction de la compagnie et s'assure du respect des réglementations nationales et internationales en matière d'aviation.

La DGAA travaille en étroite relation avec les différents services d'Air Algérie, ce qui lui permet d'avoir un rôle déterminant dans la coordination de ces services et de garantir le bon déroulement des opérations de la compagnie aérienne.

Air Algérie, en sa qualité de compagnie aérienne nationale, joue un rôle clé dans l'industrie aérienne en Algérie et dans la région environnante.



Il est de la responsabilité de la Direction Générale d'Air Algérie de veiller à ce que la compagnie continue à se développer et à prospérer, en proposant des services de grande qualité à ses clients et en se conformant aux normes internationales en matière de sécurité et de réglementation.

#### 4.1. La Direction des Programmes :

Placée sous l'autorité du pôle Opérations, la Direction des Programmes a pour objectif de superviser et d'organiser la planification et la réalisation d'un programme de vols optimisé, tout en gérant efficacement le planning individuel du personnel navigant. Elle vise à optimiser l'utilisation des ressources en matière de Flotte et de Personnel Navigant.

Quatre secteurs d'opérations sont couverts, garantissant ainsi une optimisation maximale des ressources.

- La Programmation Flotte : rotation Avions.
- La Programmation PN : rotation équipage et affectation PN.
- Affrètement et frètements.
- Les Ressources Humaines.

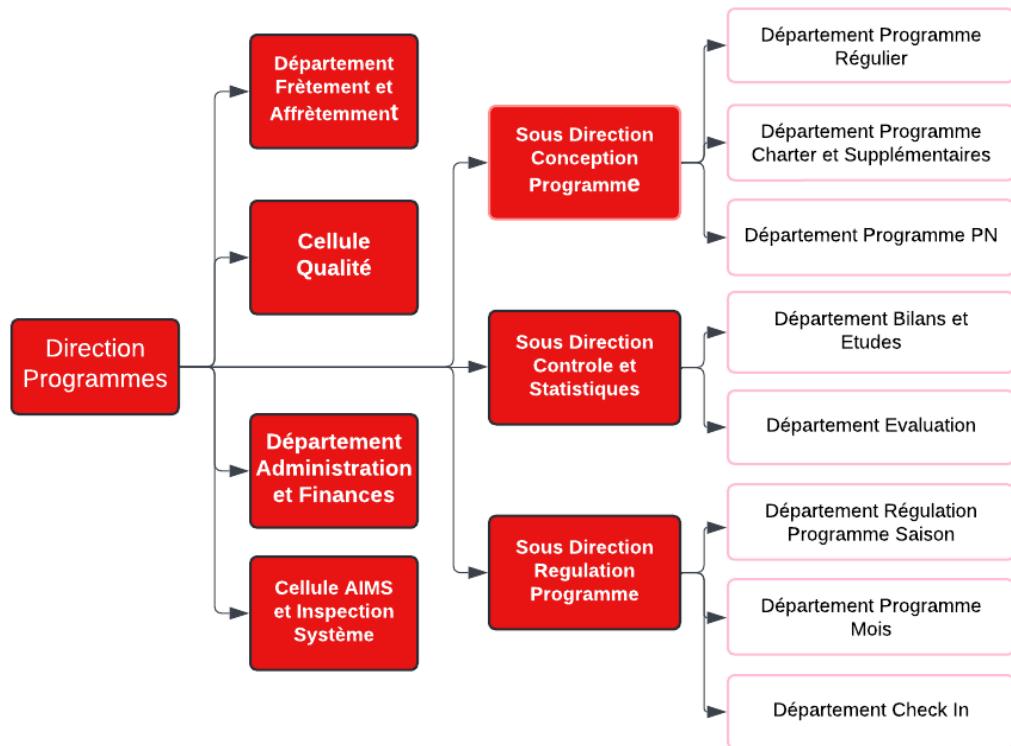


Figure 2: Organigramme de la Direction des Programmes [7]

La gestion des activités de la Direction des Programmes est assurée à travers des Départements, des Cellules et des sous-directions.

#### 4.2. Département Frètement et Affrètement

Ce département a pour mission de préparer, d'exécuter et de superviser les opérations de frètement et d'affrètement d'avions, en coordination avec les différents organes de l'entreprise, les fréteurs et les affréteurs.

Les tâches principales de ce département comprennent la gestion des appels d'offres, le suivi de toutes les demandes et l'assurance de l'obtention des autorisations nécessaires. Elle contribue également aux négociations et veille à la préparation des cahiers des charges, en plus d'analyser toutes les opportunités.

#### **4.3. Département Administration et Finances**

Ce département est chargé du suivi administratif, financier et budgétaire du personnel, en traitant les dossiers relatifs à leur formation et à leurs missions.

Il a pour mission d'établir, de surveiller et de vérifier le budget et l'ensemble des factures concernant les activités gérées par la Direction.

#### **4.4. Cellule AIMS et Inspection Système**

La cellule est chargée de la gestion des systèmes informatiques AIMS et autres villes au maintien de leur bon fonctionnement et à la mise à jour régulière de leurs bases de données. Elle assure également la maintenance de tous les systèmes de la Direction et s'occupe du contrôle de la programmation PN pour détecter toute programmation non-conforme aux normes et aux moyens, dans le but de publier le programme mensuel du PN.

#### **4.5. Cellule Qualité**

Cette cellule est chargée de l'exécution et du contrôle du système de management de la qualité et du système de management de la sécurité au niveau de la Direction des programmes. Elle a pour mission d'assister le directeur dans la mise en œuvre de la politique Qualité Sécurité, et de garantir la diffusion et la mise à disposition des documents s'y rapportant auprès de la direction. Elle collabore avec les différentes structures pour rédiger, réviser et implémenter les procédures de travail.

#### **4.6. Sous-Direction Conception Programme :**

Cette sous-direction est composée de trois départements :

- **Département Programme régulier** qui est responsable de l'élaboration des programmes saisonniers pour les périodes estivale et hivernale, qui sont représentés sous forme de grilles horaires et de fréquences hebdomadaires par type d'aéronef et par ligne, en se référant à un ou plusieurs scénarios possibles.

- **Département Programme Charter et Supplémentaires** qui élabore les programmes Hadj, Omra et autres vols charter exprimés aussi sous forme de programme horaire selon un ou plusieurs scénarios.
- **Département Programme PN** qui a comme mission l'élaboration de programme saisonnier du personnel navigant, divisé par mois, et la constitution des équipages entièrement conformes à la réglementation en vigueur, en assurant une approche transparente et équitable dans toutes les décisions.

#### **4.7. Sous-Direction Contrôle et Statistiques :**

Cette sous-direction contient deux départements dont leurs principales attributions sont :

- **Département Bilans et Études** qui sont responsable de la préparation des bilans réguliers pour la direction et collabore étroitement avec la Direction Commerciale pour mener des études de marché et anticiper les tendances futures dans le but de développer des programmes prospectifs.
- **Département Évaluation** qui s'occupe de l'inspection et de la validation régulières du programme d'exploitation et des programmes PN, ainsi que de la réalisation de contrôles ponctuels et de l'analyse des programmes de vol une fois qu'ils ont été opérationnels.

#### **4.8. Sous-Direction Régulation Programme :**

Trois départements assurent la gestion des missions dévolues à cette sous-direction :

- **Département Régulation Programme Saison** qui supervise la mise en œuvre optimale du programme d'exploitation actuel et procède aux réajustements appropriés sur le programme des aéronefs en fonction de nouveaux paramètres (situation de la flotte, demandes émergentes, nouvelles contraintes) et des aléas d'exploitation.
- **Département Régulation Programme Mois** qui est responsable de la bonne mise en œuvre du programme d'exploitation mensuel et procède aux réajustements appropriés sur les avions et les PN, jusqu'à l'envoi du programme avion au Centre de Contrôle des Opérations.

- **Département Check In** qui est responsable de la régulation du programme PN le jour J en activant les alertes appropriées et en gérant les remplacements d'équipage en collaboration avec le Centre de Contrôle des Opérations.

## **5. Système d'information :**

Un système d'information au sein d'une entreprise aéronautique se compose d'un ensemble de technologies, de processus et de données spécifiques, qui facilitent la collecte, le stockage, le traitement et la distribution d'informations au sein des divers départements et niveaux de l'organisation.

Dans le but d'optimiser la gestion de ses activités, Air Algérie a opté pour l'utilisation d'AIMS (Airline Information Management System), un système intégré qui englobe l'ensemble des aspects de l'entreprise, notamment les opérations de vol, la planification des ressources, la réservation et la billetterie, la gestion des revenus et la gestion financière. Ce système permet à la compagnie de suivre de manière continue les opérations aériennes, de gérer les équipes de vol et les appareils, d'élaborer des itinéraires, de contrôler les tarifs habituels et les tarifs promotionnels, de s'occuper des réservations et des règlements, et de créer des comptes rendus analytiques approfondis.

## **6. Destinations :**

La couverture du réseau d'Air Algérie s'étend sur 96 400 km [3], avec plus de 7,27 millions de passagers en 2023 [8] et près de 20 000 tonnes de fret transporté annuellement par la compagnie aérienne. [3]

Air Algérie met à disposition de ses clients des billets d'avion pour un large éventail de destinations sur les quatre continents, avec des vols internationaux et nationaux décollant des aéroports algériens, tels qu'Alger, Oran, Annaba, Constantine, etc. En plus des différentes escales dans plusieurs villes étrangères. [9]

## **7. Objectifs de la compagnie :**

Il convient de noter que la compagnie aérienne Air Algérie est un acteur majeur dans le développement économique du pays. Comme toute entreprise, Air Algérie a établi des buts et des objectifs à atteindre :

- Assurer de manière constante et fiable la satisfaction des besoins de la clientèle.
- Consolider la relation client et stimuler leur engagement.
- Mise en place de mesures visant à améliorer la qualité de service (sécurité, hygiène, confort et ponctualité), de la phase de réservation de siège jusqu'au moment du débarquement des clients.
- Implémenter des méthodologies techniques de production telles que la base de maintenance.
- Fournir des conditions de travail idéales pour ses employés.
- Garantir la sécurité et la facilitation des opérations d'importation et d'exportation de biens et de produits de consommation par voie aérienne.
- Améliorer la mobilité sociale grâce à la mise à disposition de la flotte à travers le territoire national et international, tout et aider à désenclaver les régions sous-développées du pays.

La politique nationale en matière de transport aérien prévoit certains objectifs que l'entreprise doit respecter, à savoir :

- Soutenir l'action de la décentralisation.
- Satisfaire la demande d'une coopération internationale multiforme.

En contrepartie, la compagnie Air Algérie cherche à augmenter de manière continue la rentabilité et assurer des profits stables sur le marché national et international. Dans le but de réaliser l'intégralité de ses objectifs, il est nécessaire pour elle de remplir ses missions tout en se conformant aux contraintes de ses ressources.

## **8. Missions de la compagnie :**

La mission principale de l'entreprise est le transport dans des conditions optimales de confort, de sécurité et de régularité des passagers, bagages, fret et courrier. De cette principale mission, découlent les missions suivantes :

- L'exploitation des lignes aériennes internationales dans le cadre des conventions et accords internationaux.
- L'exploitation des lignes aériennes intérieures et internationales en vue de garantir le transport public régulier et non-régulier des personnes, bagages et marchandises.
- La vente et l'émission des titres de transport pour son compte ou pour le compte d'autres compagnies aériennes.
- La location ou L'achat des aéronefs.
- Le transit, la prestation, les commissions, les consignations et l'assistance commerciale.
- La gestion, l'entretien régulier et le développement des installations destinées aux opérations de fret.
- Gérer et d'exploiter de manière professionnelle les installations commerciales, hôtelières et autres dans les aérogares pour promouvoir efficacement les présentations commerciales.
- L'obtention de toutes les licences, de tous les permis de survols et l'autorisation des états étrangers.

La société est également en mesure d'exécuter toute tâche de service public à la demande, dans le cadre des conventions établies avec le gouvernement.

## 9. La planification des ressources dans les campaniens aériens

L'organisation et la planification des avions et des équipages au sein des compagnies aériennes représentent une tâche extrêmement complexe et pour cela généralement sont divisées en plusieurs étapes selon un processus qui est représenté dans la figure [6] qui illustre l'ordre logique des étapes ainsi que le fait que le résultat de chaque étape sert de donnée d'entrée pour la suivante.

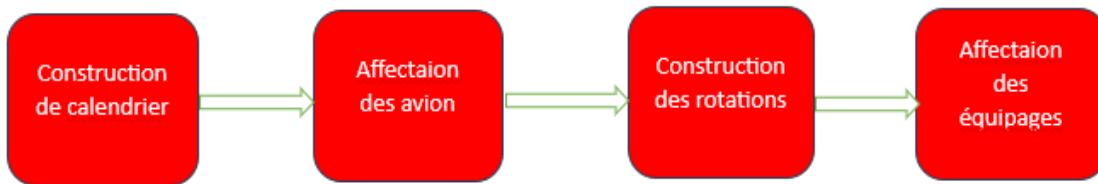


Figure 3:Les étapes de planification]

En fait, les quatre étapes sont attribuées à quatre services différents qui travaillent en parallèle sur la partie du problème de planification des ressources, ces services communiquent afin d'ajuster leurs plans les uns aux autres et de s'informer mutuellement des changements. Le première version de plans pour une période de planification données peut être mise à jour plusieurs fois jusqu'à ce que les plans soient fixés et publiés par exemple dans notre cas dans le service de création des Crew route où nous avons fait le stage après la réalisation des Crew route selon les contraintes et les réglementaires tous les modifications quel soit les vols qui peuvent être annulées ou retardées ou les équipages qui peuvent se déclarer malades, sont effectué sur la première version à cause de ça nous avons pas la vrai version des données pour faire la comparaison des résultats.

Pour résoudre ces problèmes en temps réel, les compagnies aériennes ont recours aux équipages de réserve ou aux vols charter, ou en faisant monter les équipages à bord d'un vol en tant que passagers pour couvrir un vol dans un autre aéroport, ce qu'on appelle des têtes mortes.

### 9.1. Processus

Tout d'abord, dit que le calendrier est établi, dans le but est de correspondre aux attentes du services marketing avec les flottes disponibles et avec les contraintes du réseau

par exemple les créneaux horaires disponibles pour la compagnie aérienne dans les différents aéroports.

Ensuite, dans l'étape d'affectation des avions aux segments des vols (AIRCRAFT), en prenant en compte les jours où ils sont opérationnels et ceux réservés à l'entretien, car les recettes prévues pour une étape de vol dépendent du type d'avion utilisé puisqu'il existe des aéroports où certains avions ne soient pas en mesure de décoller ou d'atterrir. Dans ce cas, la première difficulté consiste à vérifier la faisabilité du calendrier en fonction des vols disponibles. Si ce n'est pas le cas, il faudra bien entendu modifier le calendrier.

Puis, une fois les avions sont affectés, elle débute l'opération de jumelage, qui consiste en une série d'étapes de vol (séquence des vols) pour un membre d'équipage non assigné, avec un décollage et un atterrissage à la même base de l'équipage. Pendant ces séquences, le travail est généralement regroupé en plages horaires de service, telles que les journées de travail, qui sont délimitées par des pauses (arrêts de nuit). En ce qui concerne les vols de courte et moyenne durée, un jumelage peut englober environ quatre étapes de vol. Ces jumelages doivent se conformer à un grand nombre de lois gouvernementales et de conventions collectives qui diffèrent d'une compagnie aérienne à l'autre.

Tout au long des étapes de planification, la compagnie Air Algérie utilise le logiciel AIMS.

## **9.2. Logiciel AIMS**

Au sein de la compagnie Air Algérie et au niveau de département programme (DP) et comme chaque mois après la réalisation de AIRCRAFT ROUTE commence la création des CREW ROUTE de mois prochain par utilisation de AIMS (Airline Information Management System) dans le bureau planification des Crew route

AIMS INTL DWC-LLC est une entreprise qui travaille sur le développement et le soutien de solutions logicielles pour les compagnies aériennes qui évolue depuis 1983 sa base à Dubaï aux Emirats Arabes Unis et elle a des centre un peu partout dans le monde aux Grèce, et à Etats-Unis, en Athènes , son logiciel AIMS a installé dans plus de 180 compagnies aériennes et plus de 40 autorités nationales de l'aviation civile dans le monde

qui ont acceptés sur ce logiciel, en outre cette société offre des services de qualité et de sécurité selon la norme ISO [10].



Figure 4:Logo de AIMS

En fait, AIMS se concentre sur la gestion des équipages, planification et le contrôle de toutes les opérations en l'optimisation des ressources par la minimisation des coûts d'exploitation et maximisation des ressources disponibles, AIMS facilite aux décideurs de prendre les décisions plus stratégiques.

### 9.3. Méthode de Travail Crew route :

Trois ingénieurs se préparent à concevoir des itinéraires de l'équipage (appelés "Crew routes") qui doit avoir le point de départ est le même point d'arrivée. Par exemple, si un avion décolle de la base d'Alger, il devra également atterrir à la base d'Alger pour minimiser les coûts associés aux heures passées en dehors de la base.

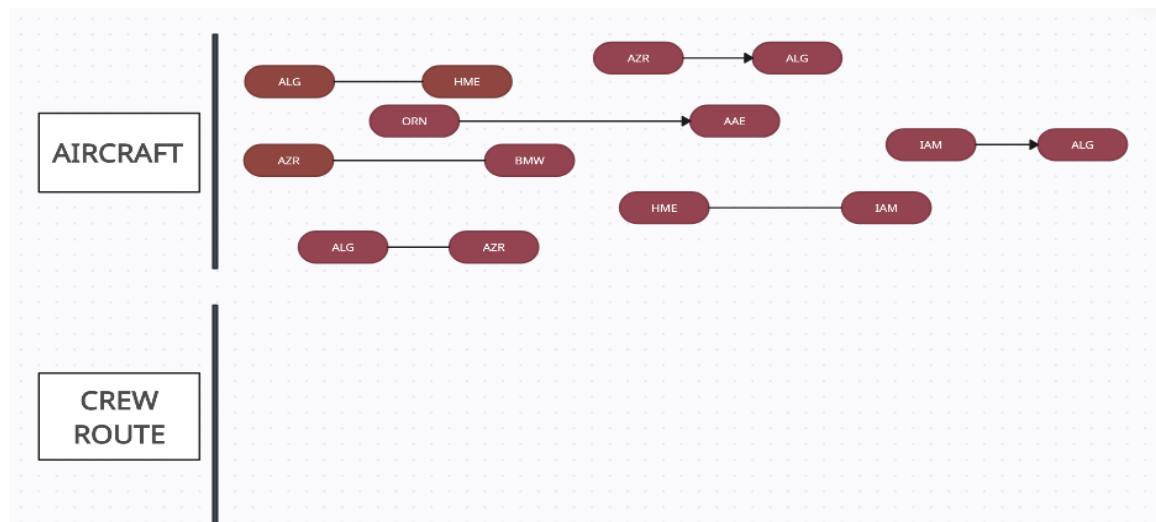


Figure 5:Illustration de l'interface de logiciel AIMS

Selon le type d'avion, le nombre d'étapes de vol quotidiennes, ainsi que les contraintes liées aux PNT (Personnel Navigant Technique) telles que la durée de travail, la durée de repos, l'heure de décollage et d'atterrissage, qui peuvent être changer, et pour

Cette opération est actuellement effectuée manuellement.

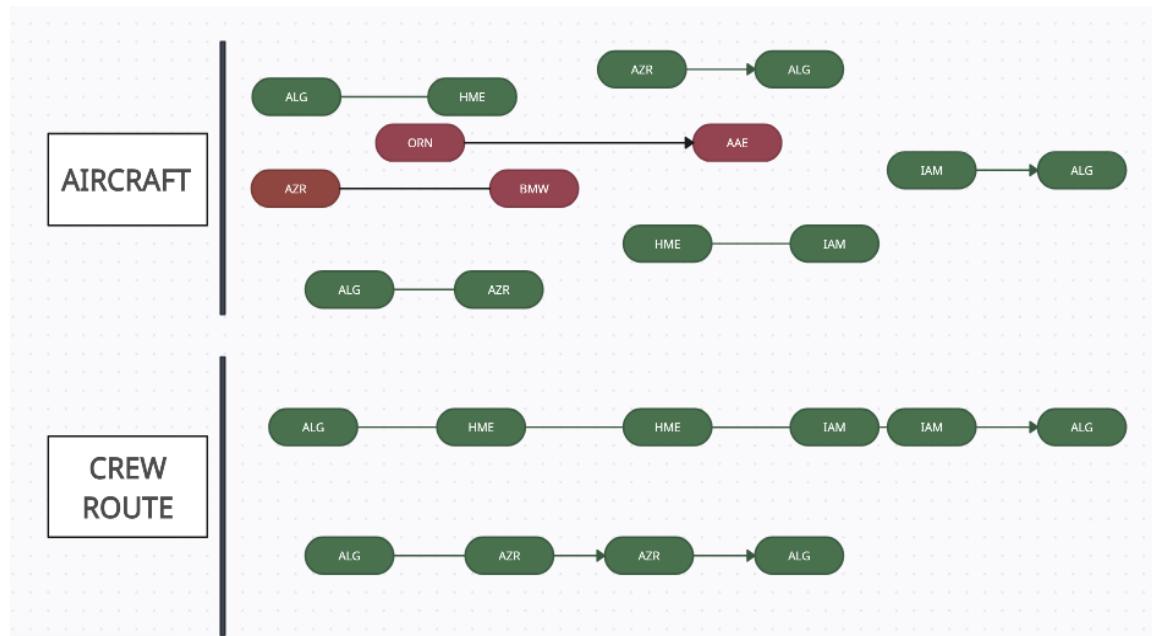


Figure 6: Illustration de l'interface de logiciel AIMS

## 10. Système de Gestion de Base des données

Logiciel AIMS utilise le Système d'information de gestion de base des données (SGBD) qui permettra aux utilisateurs dans la base des données de créer ou de supprimer, de faire des modifications, pour faciliter le stockage, l'organisation et l'accès aux données de compagnies aériennes

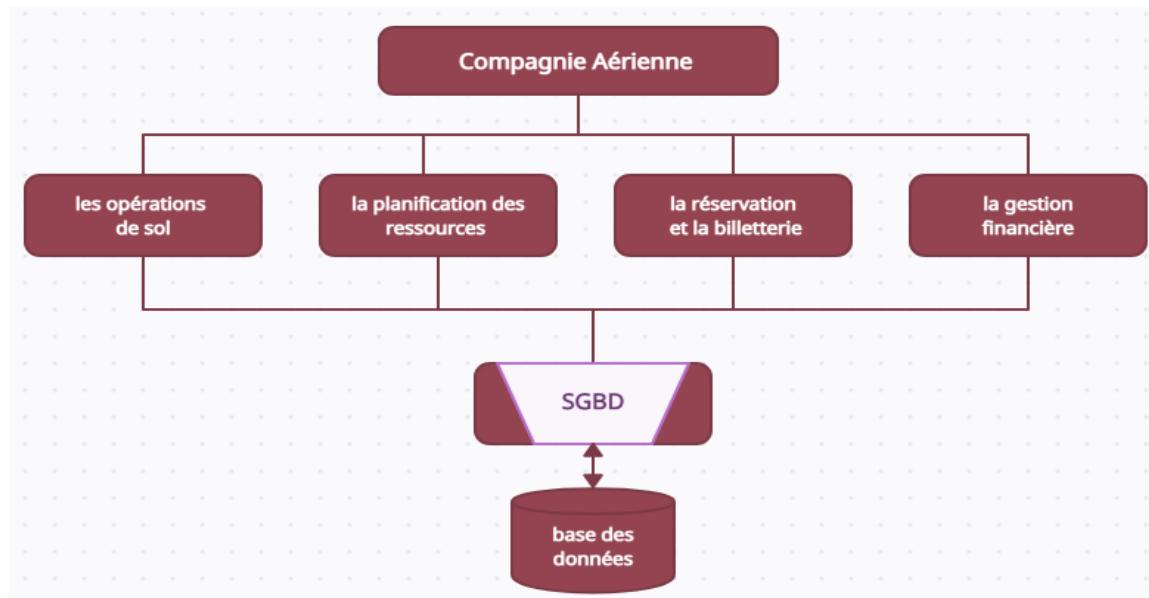


Figure 7 :Schéma de système SGBD pour une Compagnie airien [6]

Enfin, Pour terminer la planification des ressources par l'affectation des équipages, la dernière étape consiste à assigner des paires d'individus PNT et PNC, y compris les pilotes, les copilotes, les stewards et les hôtesses de l'air.) dans L'objectif est d'assurer la couverture du tous les séquences des vols (CREW ROUTE) ainsi que des exigences de formation, des congés, etc. Avec une égalité des heures de travail entre les différents équipages, tout en satisfaisant aux règles et aux réglementations de travail.

Ce problème est également connu sous le nom de "problème d'affectation des équipages" (CREW ASSIGNMENT).

## **11. L'importance d'une gestion efficace de la planification pour Air Algérie**

Dans un environnement où la présence de la concurrence est très forte, une planification optimale constitue un avantage clé pour offrir de meilleurs services à moindre coût aux clients.

Air Algérie, comme toutes les autres compagnies aériennes, travaille d'arrache-pied pour se rapprocher à une planification optimale, car une planification adéquate entraîne une réduction des coûts et une hausse de l'efficacité, ce qui équivaut à une meilleure rentabilité, l'optimisation des ressources, les avions, les équipages ( PNT,PNC), le Respect les horaires par la minimisation les retards et les annulations des vols, ce qui amélioré la ponctualité et la satisfaction des client, et d'assurer la sécurité des opérations aériennes et se conformer aux réglementations algérienne et d'éviter les problèmes d'espace aérien par le respecte des temps de vol et les périodes de repos des équipages, ainsi que l'adaptation aux changements de la demande et aux conditions météorologiques.

Une gestion de planification réussie pour Air Algérie est donc essentielle, car elle contribue à accroître l'efficacité opérationnelle, la rentabilité, la sécurité et la satisfaction des clients, tout en restant concurrentiel sur le marché.

## **12. Conclusion :**

Pour conclure ce chapitre après avoir une présentation sur l'entreprise Air Algérie, nous avons réalisé l'importance de son influence dans son secteur et sa position sur le marché.

Dans le chapitre suivant, nous allons aborder les concepts de l'apprentissage automatique, en examinant différents algorithmes et leur application dans l'industrie aéronautique, y compris l'évolution des approches dans la planification des équipages.

# **Chapitre II : Etat de l'art**

## 1. Introduction

Dans ce chapitre, nous allons aborder le sujet de l'intelligence artificielle et son développement, en nous concentrer plus spécifiquement sur la machine Learning et les différentes catégories d'algorithmes, ainsi que son application dans le transport aérien par la citation de certains travaux, et le développement des méthodes de résolution au fil des années dans la planification.

## 2. Intelligence Artificielle

L'intelligence artificielle est un concept large dans le domaine de l'informatique qui a été créé pour réaliser des services et des tâches nécessitant une intelligence humaine. Elle a pour but de développer des machines capables de s'adapter à leur environnement et de raisonner, de prendre des décisions, car il fait référence à l'apprentissage automatique, à l'apprentissage profond, aux réseaux neuronaux artificiels ou l'apprentissage par renforcement, etc. Tous ces termes sont reliés entre eux, et leur rapport est représenté graphiquement dans la figure. [11]

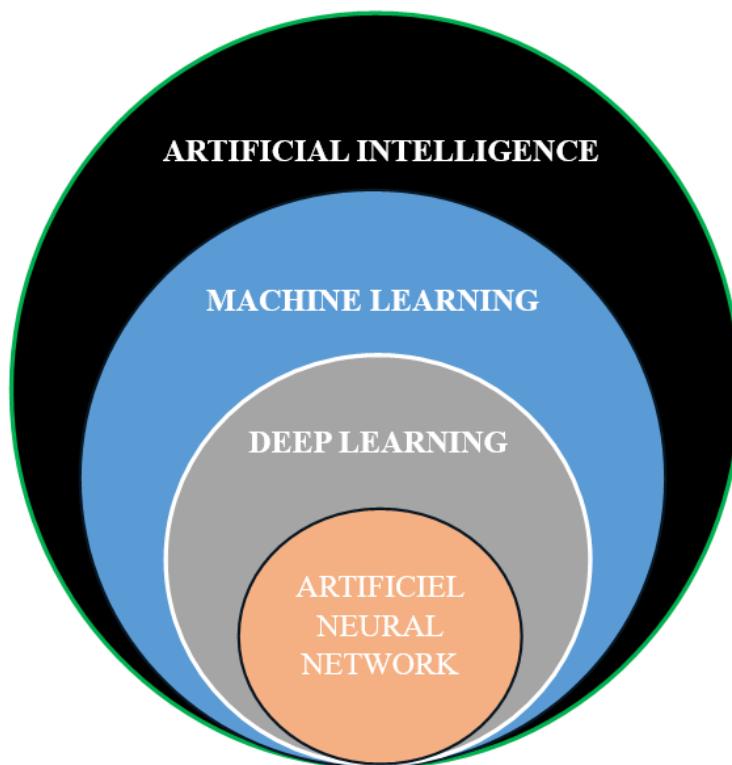


Figure 8: Relation entre les défirent branches de IA [11]

### 3. Historique

L'histoire de l'IA débute dans les années 1940, l'époque à laquelle le mathématicien et logicien Alan Turing a jeté les bases de cette science. Il est fréquemment qualifié de père de l'IA en raison de sa machine de Turing.

Pendant les années 1950 et 1960, l'intelligence artificielle a connu une croissance accélérée. Les chercheurs ont entamé la conception de programmes capables de résoudre des problèmes et de participer à des jeux, comme les échecs.

Au début du 21e siècle, l'IA a connu une évolution majeure grâce à des algorithmes plus profonds et une amélioration de ses capacités de calcul. Les scientifiques ont maintenant accès à des quantités massives de données et peuvent former des IA pour qu'elles exécutent des missions très complexes. [12]

Voici un tableau qui présente les événements ayant marqué le développement de l'intelligence artificielle : [11]

Tableau 4 : les événements principaux dans le développement de l'IA

Années	Évènement
1956	Le concept de L'intelligence Artificiel est considéré comme domaine de recherche scientifique et de technologie dans la conférence de Dartmouth
1959	Arthur Samuel introduit le terme " Machine Learning "
1966	ELIZA, le premier programme de conversation au monde
1997	L'ordinateur Deep Blue d'IBM a battu le champion du monde d'échecs Garry Kasparov
2006	Geoffrey Hinton et ses étudiants ont déclenché la vague de Deep Learning
2014	Microsoft a annoncé le lancement de son propre assistant numérique intelligent appelé Cortana
2016	Le logiciel Alpha Go de Google a remporté une victoire écrasante contre le champion du jeu de go, Lee Se-Do.

## 4. Domaines d'application de l'IA

Grâce à ses différentes méthodes, l'intelligence artificielle est largement appliquée dans de nombreuses applications telles que la vision par ordinateur, la reconnaissance de la parole, reconnaissance des visages, le traitement du langage naturel, la robotique et le contrôle des véhicules autonomes, le contrôle des systèmes logistiques les chaînes, les services aux consommateurs, l'économie, l'étude psychologique du comportement humain, les neurosciences, les systèmes de diagnostic médical pour les patients, le comportement organisationnel, l'étude des pratiques éducatives, l'étude de l'évolution, etc. [13]

## 5. Machine Learning

Machine Learning (L'apprentissage automatique) est une technologie en développement et une branche particulière de l'intelligence artificielle qui permet aux ordinateurs d'apprendre de manière autonome à partir de données historiques c'est-à-dire au lieu de prendre un ordinateur pour exécuter une tâche donnée , on lui fournit des données et des algorithmes qui lui permettent d'apprendre à partir de ces données pour exécuter la tâche elle-même et prendre de leurs expériences passées les décisions pour les nouvelles données. [11]

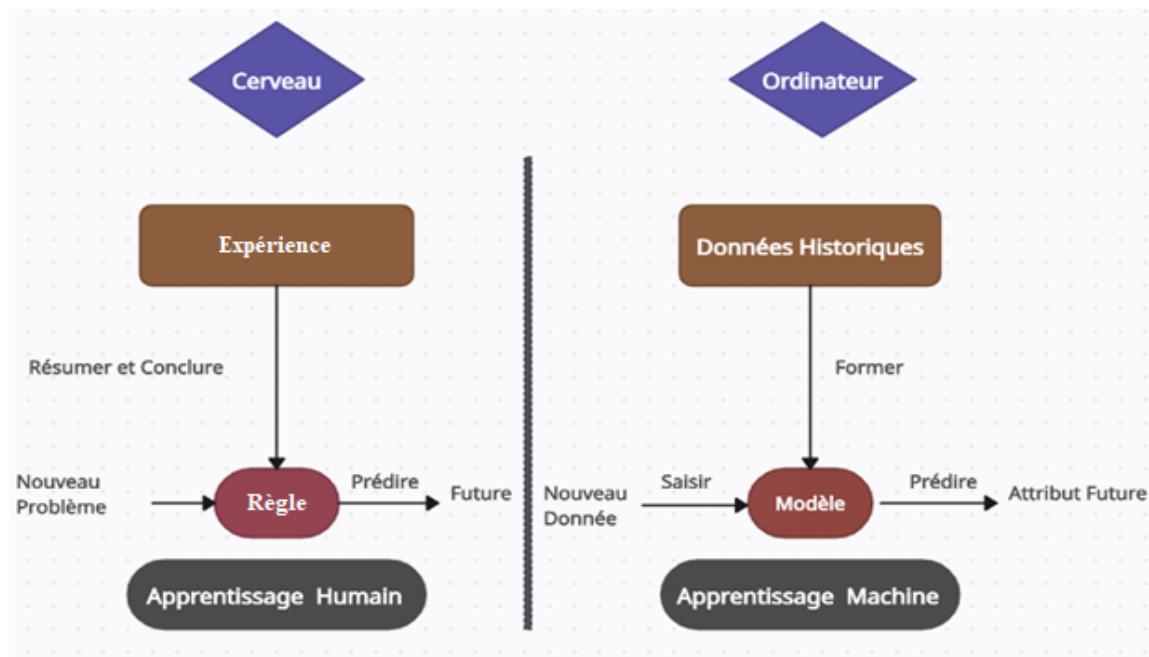


Figure 9: illustration de l'apprentissage humaine et automatique [11]

## **6. Deep Learning**

Le Deep Learning également connu sous le nom d'apprentissage profond est un type de l'intelligence artificielle issue de machine Learning qui s'appuie sur des réseaux de neurones artificiels.

Le Deep Learning utilise un nombre infini de couches de dimension limitée pour extraire de manière progressive des caractéristiques de plus en plus abstraites à partir des données brutes d'entrée.

## **7. Réseau neuronal artificiel**

Le réseau neuronal artificiel constitue la base fondamentale de l'intelligence artificielle, et son principe de fonctionnement s'appuie sur le modèle du cerveau humain. Les fondements de la conception des réseaux neuronaux ont été élaborés par des informaticiens en s'inspirant de l'agencement des neurones dans le cerveau humain. [14]

En fait Le réseau neuronal artificiel est un outil d'optimisation, puisque ses poids sont modifiés pendant le processus d'apprentissage, grâce à lui la création des conceptions de mécanismes complexes est devenue plus accessible

En particulier, dans la conception de mécanismes, les réseaux neuronaux artificiels peuvent créer un modèle, prédire des paramètres inconnus, faire une classification basée sur des données.

## **8. Les Types de Machine Learning**

L'apprentissage automatique peut être divisé en quatre catégories :

### **8.1. Apprentissage supervisé**

La méthode de travail de ces algorithmes tel que la régression linéaire et la classification est de former données d'entraînement des résultats et que les résultats sont parmi les données d'enterrement de ce modèle. [15]

**Régression :** L'algorithme de régression une approche prédictive qui consiste à étudier les succès et les échecs du passé, à en tirer des formules, et à s'en servir pour prévoir l'avenir. Dans le but d'atteindre cet objectif, il est nécessaire de déterminer le lien entre une variable qui dépend des autres et un ensemble de variables qui peuvent être modifiées,

aussi connues sous le nom de variables indépendantes. Il existe de nombreux algorithmes de régression qui peuvent être utilisés pour cette finalité. [16]

### 1) Régression linéaire :

C'est une combinaison linéaire de coefficients de régression, comprenant à la fois les constantes et les poids associés aux variables d'entrée.

La régression linéaire est une méthode déterministe et sans paramètre qui ne nécessite pas de procéder à d'autres réglages que la division des données pour l'apprentissage et le test du modèle. [17] Il existe deux types de régression linéaire :

- Régression linéaire Simple : elle est basée sur les relations entre deux variables  $X_1$  et  $X_2$ .

- Régression linéaire Multiple : elle est utilisée lorsque on a plus d'une variable d'entrée.

### 2) Régression symbolique :

Cette approche s'aligne sur des modèles comme la forme d'un arbre syntaxique constitué de symboles mathématiques aléatoires, tels que des constantes, des variables et des fonctions mathématiques, et ils sont élaborés grâce à la technique de programmation génétique stochastique et aux algorithmes évolutionnaires. [18]

## Classification :

L'algorithme de classification est conçu pour estimer la catégorie ou la classe des données fournies en entrée en fonction de leurs caractéristiques. Il vise principalement à créer un modèle à partir des données d'entraînement, dans le but de prévoir la classe ou la catégorie des nouvelles données non étiquetées. [19]

Il existe plusieurs types d'algorithmes de classification qui sont dédiés à un but spécifique et sont mis en œuvre en fonction du type de problème à résoudre.

**Les Arbres de décision :** sont des outils d'apprentissage automatique utilisés pour la régression et la classification. Elles sont formées de nœuds qui définissent un attribut spécifique des données, de branches qui symbolisent un test pour chaque valeur d'attribut

et de feuilles qui représentent la décision finale de classification. Ils aident à comprendre la manière dont les individus prennent des décisions en s'ajustant à la structure de l'arbre. [20]

**Les Machines à Vecteurs de Support (Support Vector Machines, SVM)** sont des modèles d'apprentissage supervisé très populaires qui déterminent Les fonctions de décision directement à partir des données d'apprentissage de manière à optimiser la séparation entre les frontières de décision dans un espace de caractéristiques à grande dimension. Cette stratégie de classification permet de minimiser les erreurs de classification des données d'apprentissage et d'obtenir une meilleure capacité de généralisation. [20]

**K-Nearest Neighbors (KNN)** : Il s'agit d'un algorithme qui trie les données en se basant sur les classes majoritaires parmi ses voisins dans l'espace des caractéristiques. [21]

**Classificateur Naïve Bayes** : c'est un algorithme basé sur l'application du théorème de Bayes en supposant une forte indépendance entre les différentes caractéristiques. [22]

**Forêts aléatoires** : Une méthode qui combine plusieurs arbres de décision vise à augmenter la précision de la prédiction en effectuant une moyenne des prédictions de chaque arbre. [23]

**Gradient Boosting** : c'est un algorithme qui produit un modèle de prédiction sous la forme d'un groupe de modèles de prévision approximatifs. [24]

## 8.2.Apprentissage non supervisé

Contrairement à l'apprentissage supervisé ici comme l'algorithme de Clustering, Association, on ne donne pas les résultats c'est-à-dire on lui fournit les données d'entrée et l'algorithme va regrouper selon la caractéristique et les types son comprendre ces données. [11]

**Clustering** : Il s'agit d'une méthode qui implique de regrouper des données en sous-ensembles homogènes, connus sous le nom de "clusters". [25]

Les algorithmes de clustering ont pour objectif de classifier les données en catégories de manière à ce que les éléments de la même catégorie présentent des

similitudes, tandis que les éléments de catégories différentes soient distincts. Il y a plusieurs algorithmes de clustering utilisés en machine Learning :

**K-means** : Un algorithme de partitionnement qui consiste à classer les données en groupes, où chaque point de données est assigné au groupe dont la moyenne est la plus proche du centre de ce groupe. Cette méthode de regroupement cherche à minimiser la variance à l'intérieur de chaque groupe. [26]

**Clustering hiérarchique** : Il s'agit d'un algorithme qui génère une hiérarchie de regroupements, également appelée arbre de clusters, et qui offre une décomposition hiérarchique des données sous forme de dendrogramme. Les méthodes de regroupement hiérarchique découvrent de manière récursive des clusters imbriqués, soit de manière agglomérative, soit de manière divisée. [27]

**Association** : Une méthode d'apprentissage appelée règle d'association est utilisée pour révéler les associations existantes entre les variables dans une base de données et elle identifie les éléments qui se trouvent ensemble dans la collection. [16]

### 8.3. Apprentissage semi supervisé

C'est une combinaison entre l'apprentissage supervisé et non supervisé dans ce cas cet algorithme utilise automatiquement à la fois des données qui n'ont pas des résultats avec des données qui ont des résultats pour prédire les nouvelles données. [11]

### 8.4. Reinforcement Learning :

Il s'agit d'un paradigme dans lequel une machine interagit avec son environnement, caractérisé par l'apprentissage par essais et erreurs ainsi que par les récompenses. Grâce à ce type de feedback, la machine peut apprendre à sélectionner les actions les plus optimales dans diverses situations.

## 9. Processus de Machine Learning :

Le processus se fait en trois étapes :

### **Etape 1 : Prétraitement des données**

- Importer les données
- Nettoyer les données
- Diviser en données d'entraînement et en données de test

### **Etape 2 : Modélisation**

- Construire le modèle
- Entraîner le modèle
- Faire des prédictions

### **Etape 3 : Evaluation**

- Calculer les indicateurs de performance
- Effectuer une décision

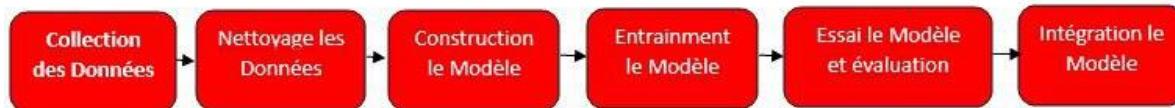


Figure 10:Processus de ML [11]

## 10. Machine Learning et le transport aérien

Il y a eu une expansion remarquable de l'association des méthodes de machine Learning et d'optimisation. Le paradigme selon lequel on prédit tout d'abord et on optimise ensuite est un modèle de modélisation traditionnel qui repose sur l'utilisation des Modèles d'apprentissage pour estimer les paramètres dans un problème d'optimisation.

L'application de ML révolutionne dans plusieurs domaines de la technologie, de la science et de l'industrie, ainsi que dans la reconnaissance d'image, la reconnaissance vocale, recommandation de produit, voiture autonomes, détection de fraude [28] ...et pour le transport aérien on trouve dans :

## **10.1. Prévision de la demande de voyageurs**

Le prévoir de la demande des passagers est une partie très importante dans la planification et la gestion à cause de son effet économiquement et financièrement c'est pour cela plusieurs compagnies aériennes utilisent la machine Learning pour aider à planifier de manière optimale et satisfaire cette demande.[35]

## **10.2. Gestion des trafics aérien**

Une CA Américaine Alaska Airlines a fait un consortium avec une société Airspace Intelligence pour utiliser une plateforme de cartographie 3D à l'aide de L'IA et ML à optimiser les itinéraires de vol [29], et aussi prévoir le retard et rendre facile à gérer, optimiser la consommation du carburant par l'estimation de carburant nécessaire pour chaque vol, développement dans la gestion des trafics aérien.

## **10.3. Optimisation des revenus**

À l'aide de l'historique des données collectées les prix des billets les jours de la semaine, les vacances, les saisons...etc., plusieurs compagnies aériennes utilisent des algorithmes tels que la régression linéaire, les réseaux de neurones selon la nature des données. [30]

## **10.4. Maintenance prédictive**

Le coût de maintenance joue un rôle essentiel dans l'économie selon les données de l'Association du transport aérien international (IATA) représente 10,3 % des coûts d'exploitation. [31] Ces coûts comprennent la maintenance, la réparation et la révision (MRO) des avions. Globalement, le marché de la MRO est prévu d'augmenter de 118 milliards de dollars à 154 milliards de dollars d'ici la fin de 2025. [32]

C'est pour cela que les Compagnies Aériens utilisent la ML pour prédire les pannes potentielles et les défaillances des avions et de faire un plan de planification de maintenance proactive. L'apprentissage automatique a la capacité de traiter de vastes ensembles de données disparates, de filtrer les points de données inutiles et de produire instantanément une représentation précise des composants individuels des avions. Grâce à ces algorithmes intelligents, plusieurs processus de surveillance basés sur l'état et de maintenance prédictive sont plus efficaces. Une étude a présenté une nouvelle technique d'apprentissage profond basée sur l'auto-encodeur et les réseaux d'unités récurrentes bidirectionnelles pour

gérer les prédictions de défaillances extrêmement rares dans la modélisation de la maintenance prédictive des avions. [33]

### 10.5. Planification des opérations aériennes

La planification des opérations aériennes est un processus complexe qui nécessite la gestion de plusieurs facteurs, à cause de ça des nombreuses compagnies utilisent l'apprentissage automatique pour améliorer la planification des opérations au sol, grâce à une gestion plus efficace des bagages, à l'optimisation des contrôles de sécurité et à la prévision précise de l'horaire des vols et de l'attribution de la flotte, ainsi qu'à l'aide à la planification des équipages. [34]

Tableau 5: Résumé des travaux similaires

Réf	Technique /Algorithme	Contribution	Résultats
[35]	-clustering - Génération de colonnes - Agrégation de contraintes - Les heuristiques basées sur ML	Cet article présente une nouvelle méthode de résolution du CPP (appariement des équipages) en associant un solveur de base amélioré avec un contrôle dynamique -un regroupement de vols basé sur l'apprentissage automatique -des techniques avancées de recherche opérationnelle pour un assemblage de solutions optimales - Il a commencé avec un solveur basique et l'a développé en incorporant des méthodes d'intelligence artificielle.	L'article indique que la méthode GENCOL-DCA surpassé le solveur de base BASELINE en termes de performance pour les CPP à grande échelle, avec une diminution moyenne des coûts de la solution comprise entre 6,8% et 8,52%, cette réduction est surtout attribuable à une baisse significative (de 69,79% à 78,11%) des coûts des contraintes globales en matière de réglementation et de limitation de l'équipage.

[36]	<ul style="list-style-type: none"> <li>- Les réseaux de neurones artificiels (ANN)</li> <li>- Les systèmes d'inférence neuro-flous adaptatifs (ANFIS)</li> <li>- les Algorithmes Génétiques (AG)</li> <li>-Les Machines à Vecteurs de Support (SVM)</li> <li>-Les Arbres de Régression</li> <li>-Les test ANOVA</li> </ul>	<ul style="list-style-type: none"> <li>- Évaluation de la performance de plusieurs algorithmes d'apprentissage automatique par rapport au modèle de régression linéaire traditionnellement dans le problème de la modélisation pour prévoir la demande de transport aérien appliquée sur la compagnie aérienne à l'Australie.</li> <li>- Identifie l'approche qui offre la plus grande précision statistique</li> <li>- Intégrer l'apprentissage automatique dans la gestion de la demande de transport aérien.</li> </ul>	<p>Les modèles ANFIS et ANN présentent un RMSE plus faible, contrairement au modèle AG qui a un RMSE multiple plus élevé</p> <p>- chaque compagnie a sa propre approche qui lui convienne.</p> <p>L'étude a également établi que les méthodes basées sur l'intelligence artificielle peuvent être mises en œuvre avec succès dans l'industrie aéronautique pour soutenir la planification et la gestion des activités.</p>
[37]	<ul style="list-style-type: none"> <li>- Arbre de Décision,</li> <li>- Réseau de Neurones</li> <li>- Machine à Vecteurs de Support (SVM),</li> <li>- Forêt Aléatoire</li> <li>- Chaîne de Markov</li> </ul>	<ul style="list-style-type: none"> <li>- Examiner les capacités de l'équipage de pilotage avant le décollage.</li> <li>- Grâce à cette approche intégrée, non seulement la classification des données de vol est améliorée, mais en plus, il offre un cadre pour améliorer les systèmes de sélection des membres d'équipage de cabine et de planification des trajets aériens.</li> <li>- Prévoir de façon proactive les performances de l'équipage de cabine afin de réduire les dangers liés à la sécurité aérienne et</li> </ul>	<p>L'utilisation de plusieurs modèles d'apprentissage automatique offre une méthode globale pour évaluer la performance de l'équipage de cockpit, ce qui permet une compréhension plus fine des facteurs de performance.</p> <p>Améliorer la Sécurité des Vols grâce à une prédition détaillée de la performance de l'équipage de pilotage.</p>

		d'accroître l'efficacité globale des opérations.	
[38]	<ul style="list-style-type: none"> <li>- la régression logistique multinomiale</li> <li>- les arbres de décision</li> <li>- Algorithme de regroupement</li> </ul>	<ul style="list-style-type: none"> <li>- prédition des choix d'itinéraires des compagnies aériennes entre deux aéroports en fonction de l'état du regroupement deux aéroports en fonction des caractéristiques des situations.</li> <li>- chaque route, telles que l'efficacité des vols, les redevances de navigation aérienne et les prévisions de trafic.</li> <li>- Explorer les différentes techniques d'apprentissage automatique pour développer de nouvelles approches de prédiction des choix d'itinéraires capables de combler le manque de données sur les itinéraires</li> </ul>	<ul style="list-style-type: none"> <li>- la performance de la régression par arbre de décision est inférieure à celle de la régression multinomiale dans certaines situations.</li> <li>- L'absence de certaines variables explicatives cruciales, telles que la disponibilité de chaque choix d'itinéraire peut influencer sur les résultats.</li> <li>- l'approche proposée offre un potentiel significatif pour améliorer les prévisions de trafic pré-tactique, afin d'optimiser les flux de trafic pour répondre à la capacité disponible</li> </ul>

## **11. Evolution des approches utilisées dans la planification des équipages :**

Depuis 1950 jusqu'à 1990 l'industrie de transport aérien utilise les techniques de recherche opérationnelle tel que la programmation en nombres entiers, la programmation linéaire par relaxation qui consiste à simplifier CPP en un programme linéaire plus facile à résoudre, Mais avec la croissance de la taille des données et les réseaux aériens la complexité de ce problème augmente ce qui rendre ces techniques inexploitables et de pousser les chercheurs à chercher des nouvelles approches.

A partir de 1990 ,l'apparition des nouveaux Techniques basées sur la génération de colonnes tel que Branchement et prix (Branch-and-Price) qui combine la génération des colonnes avec une technique de Branch-and-Bound pour la programmation en nombres entiers [39]par exemple en 1997 Barnhart et al [40] explorent une approche pour identifier des solutions quasi-optimales au problème de jumelage des équipages, en combinant une dynamique de génération de colonnes avec une procédure personnalisée de branchement et de liaison.

Au début des années 2000, l'apparition des Heuristique et des mét-heuristiques ces approches qui consiste à trouver de bonnes solutions en un temps réduit et avec l'essor des techniques d'apprentissage automatique (ML) plusieurs travaux qui ont marqués une amélioration dans la résolution de ce problème, parmi ces travaux on a Yaakoubi et al. (2020) [35] utilisent l'apprentissage automatique pour créer des solutions initiales et réorganiser les clusters du problème de planification (CP) basé sur la génération de colonnes, en parallèle Morabit et al. (2021) [41] emploient un réseau neuronal graphique pour choisir les paires à préserver pendant de la génération de dont le but est de diminuer le temps consacré au solveur, Quesnel et al. (2022) [42] prévoient les appariements seront bénéfiques en filtrant les appariements durant la génération de colonnes d'une affectation d'équipage.

Grâce à ces méthodes innovantes, les compagnies aériennes sont désormais capables de mettre en place des stratégies de jumelage d'équipages plus performantes et rentables, en prenant en compte la complexification croissante des réseaux aériens.

## **12. Conclusion :**

Ce chapitre a pour objectif de fournir un aperçu global sur l'apprentissage automatique en définissant des concepts clés et en expliquant les différents algorithmes utilisés dans ce domaine, et en mentionnant les applications de l'apprentissage automatique dans l'industrie aéronautique par la citation de certains travaux et de donner le développement des techniques utilisées pour résoudre les problèmes de la planification des équipages.

# **Chapitre III :**

# **Modélisation**

## 1.Introduction

Après avoir présenté brièvement le machine learning, nous allons, dans ce chapitre, expliquer notre base de données et la prétraiter afin d'appliquer différentes approches et méthodes d'apprentissage automatique à notre problème. Nous examinerons les avantages, les inconvénients et les limites de ces méthodes et discuterons les résultats obtenus en les comparant avec ceux de la compagnie.

## 2.Ressource et Données

Comme indiqué précédemment dans le premier chapitre, Air Algérie utilise le système AIMS tout au long du processus de planification, qui permet de relier et de coordonner les différents services et de faciliter la récupération des données.

En effet, le logiciel AIMS permet d'accéder aux différentes informations stockées telles que les programmes mensuels des vols représentés par un tableau appelé AIRCRAFT ROUTES, les programmes mensuels des rotations des vols représenté par un tableau appelé CREW ROUTES REPORT et d'afficher les informations en plusieurs versions par exemple sous la forme d'un document PDF ou d'un tableur Excel.

Aircraft Routes																		
01/06/2019 - 30/06/2019																		
Route	Dur	Operation Period	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Flight	Dep	Arr	STD	STA	Trip Day	Block Time		
1008	2	07/06/2019-07/06/2019	.	.	.	.	B738	.	.	1008 1009	ALG ORY	ORY ALG	17:35 5:40	19:55 7:50	1 2	2:20 2:10 4:30		
2008	1	16/06/2019-23/06/2019 27/06/2019-30/06/2019	.	.	.	.	B736	.	.	B736 B736	2008 2009	ALG VLC	VLC ALG	15:40 17:45	16:45 18:55	1	1:05 1:10 9:00	
1040	1	02/06/2019-02/06/2019 09/06/2019-09/06/2019 21/06/2019-28/06/2019	.	.	.	.	.	.	B737 B737	1040 1041	ALG NCE	NCE ALG	15:30 18:10	17:10 19:45	1	1:40 1:35 13:00		
1020	1	03/06/2019-10/06/2019 12/06/2019-14/06/2019 22/06/2019-27/06/2019 29/06/2019-29/06/2019	B737	.	B737	.	B737	.	.	1020 1021	ALG MRS	MRS ALG	13:20 15:50	14:50 17:30	1	1:30 1:40 31:40		
1070	1	07/06/2019-14/06/2019	.	.	.	.	B738	.	.	1070 1073	ALG MPL	MPL ORN	8:00 10:25	9:25 11:55	1	1:25 1:30 12:00		

Figure 11: Capture des données fournies

Description du vol : chaque vol selon le tableau doit comporter un code, un type d'avion spécifique, un aéroport de départ et d'arrivée, la date du vol, l'heure de décollage et d'atterrissage.

### **2.1. Explication des codes du programme**

- **Route** : c'est un code de navigation décrivant les aéroports de départ et d'arrivée pour les vols appartenant à cette rotation.

- **Operation Period** : la dates ou la période durant laquelle les vols fonctionnent sur cette route.

- **jours de la semaine (Mon, Tue, Wed, Thu, Fri, Sat, Sun)** : Ces colonnes indiquent les jours ou la semaine pendant lesquels l'avion est disponible pour les vols réguliers.

- **Flight** : Un code numérique utilisé comme identifiant de vol unique.

- **Dep** : Aéroport de départ.

- **Arr** : Aéroport d'arrivé.

- **STD** : le temps de départ standard de vol.

- **SDA** : le temps d'arrivé standard de vol.

- **TRIP Day** : le nombre de jours pour une rotation.

- **Block Time** : il représente la durée partielle du vol et la durée totale de toutes les rotations existe dans une période donnée, sans inclure le temps de service, c'est-à-dire depuis le décollage de l'avion jusqu'à son atterrissage.

- **Aéronefs (B736, B737, B738)** : ces codes représentent le type d'avion qui est disponible pour opérer le vol dans notre cas, nous avons travaillé sur le type NG.

### **3. Justification de la proposition d'une solution de Machine Learning (ML) :**

La justification de la proposition du machine learning repose sur son influence et ses avantages potentiels pour révolutionner la planification au sein de l'entreprise. En

proposant des solutions innovantes et optimisées, le machine learning constitue un investissement stratégique pour assurer une compétitivité durable à long terme.

Les algorithmes de machine learning sont capables d'analyser dénormes quantités de données, tant historiques qu'actuelles, pour créer des estimations de demande plus fiables que les approches traditionnelles. Ils optimisent l'attribution des avions aux itinéraires et aux horaires, maximisant ainsi l'utilisation des ressources tout en minimisant les temps de repos et les heures de travail. Cela permet de réduire les coûts, d'accélérer la création de nouveaux plans optimaux en temps réel, et de simuler plusieurs scénarios. Ces algorithmes sont développés sur mesure pour chaque compagnie aérienne, afin de fournir une meilleure adaptation aux horaires, une intégration aisée avec les systèmes actuels et une amélioration continue grâce à l'apprentissage automatique.

Pour Air Algérie, l'apprentissage automatique représente une opportunité d'améliorer l'optimisation de la planification des ressources et l'efficacité opérationnelle. Il permet de simplifier la gestion des défis, d'offrir une meilleure expérience à ses passagers et d'accroître sa rentabilité.

## 4. Prétraitement des données :

Dans notre étude, nous avons sélectionné les données des vols réguliers effectués par des avions de type NG entre le 1er juin 2019 et le 30 juin 2019. Ces données constituent ce que l'on appelle un ensemble d'apprentissage et comprennent un total de 2456 vols.

Après avoir importé notre base de données des vols au format CSV, nous avons effectué une suppression de certaines colonnes non pertinentes pour notre analyse, telles que Route, Durée, Jours de semaine, Jour de trajet et Temps de bloc. La figure [12] présente un échantillon typique de notre ensemble de données. Chaque ligne représente un vol individuel identifié par son code de vol (flight code), avec les détails de l'aéroport de départ, de l'aéroport d'arrivée, des dates de départ et d'arrivée, ainsi que des heures de départ et d'arrivée.

Chaque colonne de notre base de données correspond à une caractéristique ou à une étiquette spécifique, nécessaire pour notre analyse et la modélisation ultérieure.

## 5. Présentation des outils utilisés

Flight	DEP	ARR	daydep	dayarr	STD	STA
1008	ALG	ORY		1	17:35	19:55
2026	ALG	MXP		2	9:15	11:15
1023	MRS	ALG		1	10:30	12:10
1011	ORY	ALG		2	12:00	14:10
1199	NCE	CZL		1	11:10	12:35
1075	LIL	ALG		1	14:15	16:40
2014	ALG	BCN		2	15:45	17:05
1010	ALG	ORY		1	8:10	10:30
2015	BCN	ALG		2	18:05	19:25
2027	MXP	ALG		2	12:15	14:15
1022	ALG	MRS		1	8:00	9:30
1009	ORY	ALG		2	5:40	7:50

Figure 12 Notre Base de Donnée

### 5.1 Python

On a décidé de Travailler avec le python parce que :

- C'est un langage qu'on a vu dans notre formation d'ingénieur et un langage qu'on maîtrise
- Riches en bibliothèques puissantes de ML.

- Excellent pour manipuler les données qui est crucial dans le domaine de ML grâce aux bibliothèques de calcul scientifique comme : Panda et Numpy.
- La communauté active et la diversité des ressources et tutoriels.
- La flexibilité de langage et la facilité d'intégration d'autres langages de programmation.

### **5.3 Les bibliothèques utilisées :**

#### **Numpy :**

Il s'agit d'une bibliothèque Python qui fournit un objet tableau multidimensionnel, divers objets dérivés (tels que les tableaux masqués et les matrices) et un assortiment de routines pour des opérations rapides sur les tableaux, y compris des opérations mathématiques, logiques, de manipulation de forme, de tri, de sélection, d'E/S, de transformée de Fourier discrète, d'algèbre linéaire de base, d'opérations statistiques de base, de simulation aléatoire et bien d'autres choses encore. [43]

#### **Pandas :**

Il s'agit une bibliothèque Python qui fournit la manipulation rapide et efficient des données avec indexation intégrée, Outils de lecture et d'écriture de données entre des structures de données en mémoire et différents formats : fichiers CSV et texte, Microsoft Excel, bases de données SQL et autre fonctionnalités utiles. [44]

#### **Tensorflow :**

Il s'agit d'une bibliothèque Python qui offre la flexibilité, l'exécution impatiente permet une itération immédiate et un débogage intuitif, la possibilité de Construire et former des modèles de pointe sans sacrifier la vitesse ou la performance, la flexibilité et le contrôle grâce à des fonctionnalités telles que Keras Functional API et Model Subclassing API pour la création de topologies complexes. Pour faciliter le prototypage et le débogage rapide, utilisez l'exécution impatiente. [45]

#### **Matplotlib :**

Matplotlib est une bibliothèque vaste permettant de créer des visualisations statiques, animées et interactives en Python. Matplotlib rend les choses faciles faciles et les choses difficiles possibles. [46]

- Créez des graphiques de qualité professionnelle.
- Créez des figures interactives qui peuvent être zoomées et actualisées.
- Personnaliser le style visuel et la mise en page.
- Exportation vers de nombreux formats de fichiers.

### **Scikit-learn :**

Une bibliothèque de ML en Python qui fournit des outils simple et efficace pour l’analyse et la prévision des données. Elle est accessible pour tout le monde et compatible avec les bibliothèques mentionnées précédemment [47].

### **5.3 Google Colab :**

Colab est un service hébergé de bloc-notes Jupyter qui ne nécessite aucune installation et offre un accès gratuit aux ressources informatiques, y compris les GPU et les TPU [48].

On a choisi colab puisque :

- Il est particulièrement adapté à l’apprentissage automatique.
- La possibilité de diviser et d’exécuter le code en blocs qui donne une flexibilité et rend le processus de résolution des bugs plus facile.
- Toutes les bibliothèques mentionnées précédemment sont déjà installées.
- Aucune configuration requise. Il suffit de créer un compte et de commencer à coder.
- Comme son nom l’indique, plusieurs personnes peuvent créer, éditer et exécuter le même code facilement.

## 6. Les différents Methods et approches utilisé :

### 6.1 ANN :

Le réseau de neurones artificiel c'est pratiquement créé basée sur l'anatomie de cerveau humain comme il est indiqué dans les figures dessus,

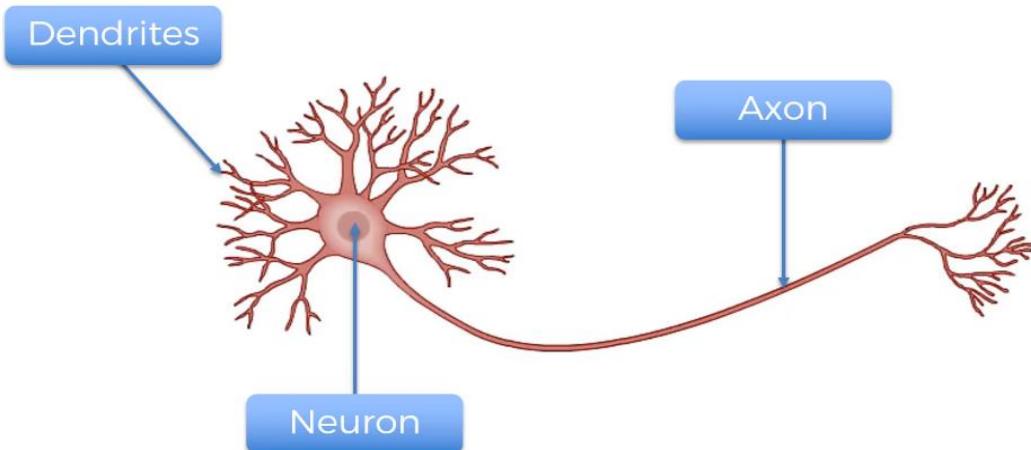


Figure 13: Illustration d'un neurone Source : Wikipédia

Vue que le cerveau humain est l'outil le plus puissant en termes de traiter les problèmes complexes, l'idée après les ANNs est de recréer ce mécanisme magnifique et essayer de l'utiliser pour résoudre des problèmes et faire des prédictions de haute qualité

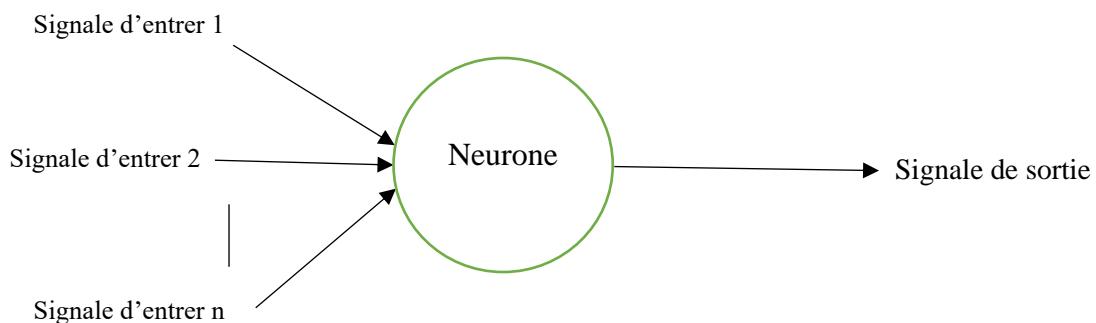


Figure 14: Schéma de neurone humaine Source : Wikipédia

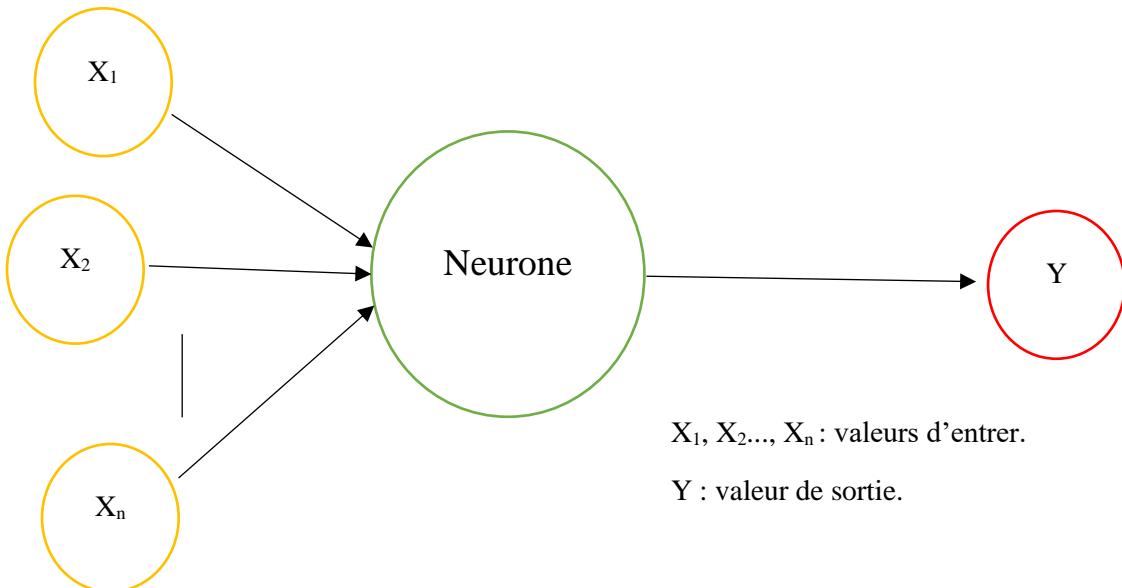


Figure 15: Schéma de neurone artificiel Source : Wikipédia

### 6.1.1. Comment on peut utiliser les ANNs dans notre problème ?

Pour adapter cette infrastructure à notre problème, nous avons considéré notre problème comme une tâche de classification afin de le rendre compatible avec l'intuition d'un réseau neuronal. Nous avons également formulé le problème de deux manières différentes.

### 6.1.2 La première approche :

Pour adapter notre problème comme une tâche de classification, où les rotations pouvant être effectuées par le même équipage sont considérées comme des classes différentes, la sortie du modèle et les valeurs d'entrée sont structurées de manière spécifique. Dans notre réseau neuronal, la structure comprend une couche d'entrée avec  $n$  neurones, où  $n$  représente le nombre de variables qui définissent un vol après le prétraitement des données. Pour notre cas, les variables incluent le départ, l'arrivée, l'heure de départ, l'heure d'arrivée, la date de départ et la date d'arrivée.

Cependant, le nombre  $n$  n'est pas nécessairement égal au nombre total d'attributs, qui est de 6 dans notre cas, car ces données ne sont pas directement compatibles avec le format requis par le réseau neuronal. Cela implique que nous devrons effectuer un encodage des données pour les rendre utilisables dans notre modèle de réseau neuronal.

### 6.1.3 Le codage des données :

Pour les attributs catégoriels tels que l'aéroport de départ et l'aéroport d'arrivée, l'encodage one-hot est recommandé. Cela transforme ces nnn catégories en nnn états ou variables, où chaque variable prend la valeur 1 si elle correspond à la catégorie respective, sinon elle prend la valeur 0. Par exemple, si nous avons les aéroports ('ALG', 'ORN', 'CZL'),

l'encodage one-hot convertira cette colonne en trois colonnes distinctes, chacune représentant un aéroport et prenant la valeur 1 lorsque l'aéroport correspondant est celui du vol, et 0 sinon.

Cela implique que les colonnes de départ et d'arrivée seront multipliées par deux fois le nombre d'aéroports présents dans la base de données.

*Tableau 6: Exemple de l'encodage one-hot*

Départ		ALG	ORN	CZL
ALG		1	0	0
ORN	→	0	1	0
CZL		0	0	1

Pour simplifier le traitement des heures et des dates de départ et d'arrivée, nous avons choisi de considérer chaque mois comme une unité de temps distincte, représentant une variation. De plus, les heures seront converties en heures de la journée, ce qui signifie que l'heure ne sera pas mesurée à partir de minuit mais plutôt sur une échelle horaire allant jusqu'à la fin du mois.

Ces choix méthodologiques visent à rendre les données compatibles avec les exigences du modèle de réseau neuronal et à faciliter l'analyse et la prédiction dans notre étude sur la planification des vols.

### **Heures de départ= (jour de départ -1) \*24+l'heure de départ h**

Pour cette approche, toutes les étapes seront implémentées à l'aide d'une boucle en Python. Après l'encodage one-hot des attributs catégoriels tels que les aéroports de départ et d'arrivée, une étape de standardisation des valeurs obtenues sera effectuée. Ainsi, le nombre total d'entrées  $n$  pour la couche d'entrée sera  $2 \times m + 2$ , où  $m$  représente le nombre d'aéroports présents dans la datasheet.

Concernant les couches cachées, il n'y a pas de règle générale fixe et cela dépend de la complexité du problème. Augmenter le nombre de couches cachées et de neurones peut améliorer la précision du modèle, mais cela rend également l'entraînement plus difficile. Dans notre cas, nous avons opté pour deux couches cachées comportant  $p$  neurones chacune.

La dernière couche du réseau neuronal sera la couche de sortie, directement liée au nombre de routes d'équipage présentes dans le dataset. Ces routes seront encodées en utilisant l'encodage one-hot, étant des données catégoriques représentées par  $Y$ .

Cette approche est conçue pour être rapide et économiser du temps, mais elle pourrait compromettre la précision des résultats et manquer de flexibilité dans la manipulation des résultats. De plus, le modèle pourrait avoir des performances médiocres avec un dataset contenant un nombre variable de routes d'équipage, ce qui pourrait limiter son efficacité.

Cette méthodologie sera mise en œuvre avec l'objectif de répondre efficacement aux besoins spécifiques de la planification des vols dans notre étude.

#### **6.1.4. Deuxième approche :**

Pour aborder ce problème comme une tâche de prédiction où le modèle combine à la fois des aspects de classification et de régression, nous devons ajuster la manière dont nous préparons les données d'entrée pour le modèle.

L'ensemble des vols restera l'entrée principale du modèle, mais cette fois-ci, la sortie pour chaque vol consistera en une prédiction du vol suivant. Cela nécessite une modification de notre approche pour nourrir le modèle en conséquence.

Cette approche hybride permettra de prédire de manière anticipée les vols suivants, ce qui est crucial pour la planification efficace des équipages dans le cadre de notre étude.

#### **6.1.5. Codage :**

Cette fois, les données d'entrée ne seront pas les attributs individuels de chaque vol, mais plutôt les relations entre les vols. Cela signifie que nous ne fournirons pas seulement l'aéroport de départ pour un vol, mais aussi les autres vols ayant le même aéroport d'arrivée et vice versa. De plus, pour l'heure et la date de départ, nous alimenterons le modèle avec la différence entre le départ de chaque vol et l'arrivée des autres vols. La valeur de sortie sera simplement le vol suivant dans la séquence.

Pour chaque attribut, nous aurons une matrice carrée de taille  $v \times v$ , où  $v$  représente le nombre de vols programmés. Cela signifie que chaque colonne d'entrée deviendra une matrice de  $v$  colonnes. Par conséquent, la couche d'entrée devra avoir  $4 \times v$  neurones, avec  $v$  étant le nombre de vols. La couche de sortie sera composée de  $v$  neurones, où chaque neurone représente un vol et sa valeur représente la probabilité que ce vol soit le vol suivant.

Cette approche devrait produire des résultats plus utiles que l'approche précédente, car elle prend en compte les relations entre les vols pour prédire le vol suivant. Cependant, sa principale limitation réside dans le fait que le nombre d'entrées est directement lié au nombre de vols, ce qui peut rendre le modèle inadapté pour des datasets comportant un nombre variable de vols.

#### **6.1.6. Suggestions :**

Lorsque les approches traditionnelles basées sur un nombre variable d'entrées et de sorties ne sont pas applicables dans la plupart des cas réels, une méthode d'exploitation efficace des réseaux neuronaux consiste à adapter les données. Cela implique d'utiliser le réseau neuronal pour réduire la complexité du problème en divisant les données en ensembles distincts.

Plus précisément, nous pouvons diviser les données en deux ensembles : l'ensemble à passer par le modèle ne contiendrait que les vols qui se répètent régulièrement tout au long

de l'année. Cela permettrait d'avoir un nombre fixe d'entrées et de sorties, ce qui résoudrait le problème de la limitation inhérente aux réseaux neuronaux.

Cette approche permet d'adapter le modèle pour répondre aux contraintes spécifiques du domaine de la planification des vols, tout en exploitant la capacité des réseaux neuronaux à traiter des données complexes et à fournir des prédictions précises dans des contextes où les structures de données peuvent varier considérablement.

## **6.2. Clustering :**

Pour générer des clusters ou ensembles de vols de manière à ce que chaque vol appartenant à un ensemble soit plus similaire aux autres vols de cet ensemble qu'aux vols des autres ensembles, nous utiliserons une approche de clustering hiérarchique. Cette méthode est adaptée car elle regroupe les objets similaires dans le même groupe en se basant sur des variables mesurées par le calcul des distances entre ces variables une à une, afin de déterminer le regroupement optimal.

Pour déterminer le nombre optimal de clusters, nous utiliserons d'abord la méthode du dendrogramme. Parmi les méthodes de liaison disponibles, la méthode de Ward est la plus appropriée pour ce problème. En effet, elle minimise l'augmentation de la variance totale au sein du cluster après la fusion de deux clusters, ce qui la rend idéale pour créer des clusters équilibrés avec une variance minimale.

L'approche de clustering hiérarchique avec la méthode de Ward nous permettra de structurer les vols en clusters significatifs, facilitant ainsi l'analyse et l'organisation des données pour la planification des équipages.

### **6.2.1 C'est quoi un dendrogramme :**

Un dendrogramme est une représentation graphique utilisée en classification hiérarchique pour illustrer l'organisation des clusters générés par les algorithmes de clustering. C'est un outil visuel crucial qui facilite l'interprétation des résultats de clustering hiérarchique. Son fonctionnement commence par considérer chaque point comme un cluster individuel, puis il fusionne progressivement les clusters les plus proches entre eux.

À chaque fusion, une ligne horizontale est tracée, dont la hauteur représente la distance entre les clusters fusionnés. Ce processus se poursuit jusqu'à ce qu'un seul cluster englobant

tous les points soit construit. Le dendrogramme permet ainsi de visualiser la hiérarchie des regroupements et de déterminer de manière intuitive le nombre optimal de clusters à partir des hauteurs de fusion.

### **6.2.2. Pourquoi le dendrogramme :**

Il offre une présentation graphique qui peut fournir des informations utiles sur la structure des données et la manière dont elles sont traitées par un algorithme de clustering. Bien qu'il ne soit pas toujours recommandé d'utiliser un dendrogramme pour déterminer le nombre de clusters dans des problèmes de clustering complexes, dans notre cas, il sera suffisant. Nos données étant relativement simples, nous utiliserons le temps comme critère principal.

L'approche consistera à observer la plus grande distance entre deux jointures sur le dendrogramme. Cette distance représentera le seuil au-delà duquel les clusters sont considérés comme distincts. C'est un point crucial où l'algorithme doit s'arrêter, ce qui devrait nous fournir des clusters de bonne qualité. Cette méthode pratique nous permettra de déterminer efficacement le nombre optimal de clusters adaptés à notre problème de planification des équipages.

### **6.2.3. Prétraitement de Data :**

Le clustering fonctionne généralement mieux avec des variables numériques, car il repose sur le calcul des distances entre les données. Ainsi, pour notre étude, les critères les plus appropriés sont l'heure et la date de départ ainsi que d'arrivée, étant donné qu'elles peuvent être converties en valeurs numériques grâce à l'utilisation de bibliothèques comme datetime en Python. Cette approche nous permettra de traiter ces données de manière similaire à ce qui a été décrit précédemment, en convertissant les données en heures de départ et d'arrivée, ce qui nous fournira deux variables numériques pour alimenter notre modèle de clustering.

#### **6.2.4. Création de modèle :**

Cette étape est vraiment simple pour la faire à l'aide de bibliothèque sklearn et plus précisément AgglomerativeClustering du class cluster, qui demande 3 paramètres principales :

**N clusters** : c'est le nombre des clusters souhaitable et ça va être déterminé à l'aide de dendrogrammes.

**Affinity** : c'est la métrique utilisée pour calculer la liaison entre les clusters, ici la métrique euclidienne sera le plus adéquat puisque nous voulons faire le clustering basée sur la distance en ligne droite.

**Linkage** : c'est la méthode de liaison et ça sera le même comme précédemment pour les mêmes raisons, 'Ward'

Et après la création de modèle on va présenter nos données pour l'entraînement de modèle et recevoir le résultat de clustering.

#### **6.2.5. Pourquoi le Clustering :**

La raison d'utiliser la méthode de clustering est de réduire la complexité du problème en divisant les données d'entrée. Cela permet de simplifier le processus de résolution en regroupant les données similaires, ce qui potentiellement réduit le temps nécessaire pour obtenir des résultats significatifs.

#### 6.2.6. L'implémentation :

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
dataset= pd.read_csv("/content/Flights.csv")
dataset['STD'] = pd.to_datetime(dataset['STD'])
dataset['STA'] = pd.to_datetime(dataset['STA'])
dataset['heure_depart_hours'] = dataset['STD'].dt.hour + dataset['STD'].dt.minute / 60
dataset['heure_arrive_hours'] = dataset['STA'].dt.hour + dataset['STA'].dt.minute / 60
dataset['STD'] = dataset['STD'].dt.time
dataset['STA'] = dataset['STA'].dt.time
dataset['depart_hours'] = (dataset['daydep']-1) * 24 + dataset['heure_depart_hours']
dataset['arrive_hours'] = (dataset['dayarr']-1) * 24 + dataset['heure_arrive_hours']
x = dataset[['depart_hours', 'arrive_hours']]
import scipy.cluster.hierarchy as sch
dendrogram = sch.dendrogram(sch.linkage(x, method= 'ward'))
plt.title('Dendrogram')
plt.xlabel('flights')
plt.ylabel('ED')
plt.savefig('my_plot.png')
from sklearn.cluster import AgglomerativeClustering
hc = AgglomerativeClustering(n_clusters=2 , affinity = 'euclidean', linkage = 'ward')
y_hc = hc.fit_predict(x)
```

Voici un schéma descriptif de Code :

- **Préparation des données** : Le code commence par importer les bibliothèques nécessaires et charger le jeu de données. Il convertit ensuite certaines colonnes en format de date et d'heure, et crée de nouvelles colonnes basées sur ces informations.
- **Création de la matrice de caractéristiques** : Une matrice de caractéristiques X est créée à partir des colonnes ‘depart\_hours’ et ‘arrive\_hours’.
- **Création du dendrogramme** : Un dendrogramme est créé en utilisant la méthode Ward pour le regroupement hiérarchique. Le dendrogramme est ensuite sauvegardé sous forme d'image.
- **Application du clustering hiérarchique** : Le clustering hiérarchique agglomératif est appliqué sur la matrice de caractéristiques X avec 2 clusters, une affinité euclidienne et une liaison de type ‘Ward’. Les prédictions du modèle sont stockées dans y\_hc.

#### 6.2.7. Les résultats :

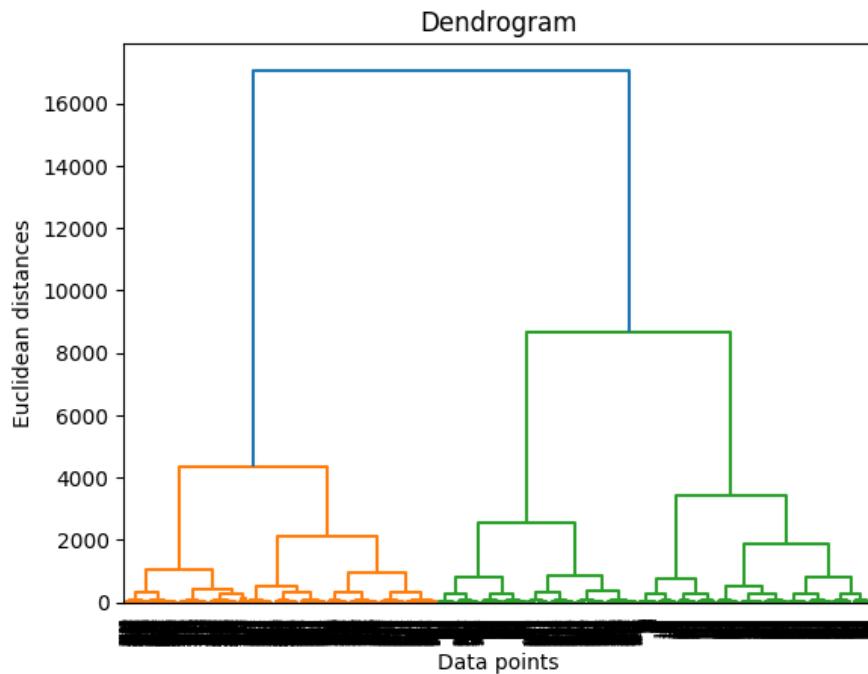


Figure 16: le Dendrogramme

#### **6.2.9. Interprétation :**

**Axe des x (Data Points) :** L'axe des x représente les différents vols qui sont regroupés.

Chaque point sur l'axe des x représente un vol.

**Axe des y (Euclidien distances) :** L'axe des y, qui signifie ‘Distance Euclidienne’, indique la distance entre les points de données dans l'espace des caractéristiques considérées par l'algorithme.

**Liens horizontaux :** Les liens horizontaux représentent les clusters formés à chaque étape de l'algorithme. La hauteur des liens indique la distance entre les clusters ; des liens plus hauts représentent des distances plus grandes.

On peut voir que le nombre de clusters recommandé ici est 2 si en coupant la plus grande ligne verticale entre deux lignes horizontales et c'est indiqué aussi par le nombre de couleurs.

Et donc notre data qui consiste de 2391 lignes (vols) va être divisé en deux ensembles de tailles 1497 et 894 vols.

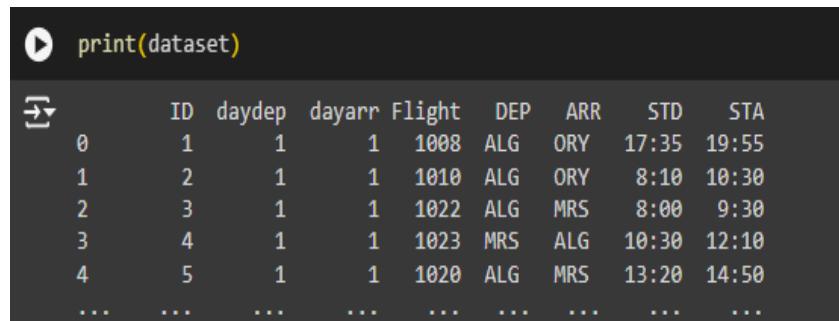
### **7. L'algorithme suggéré :**

Après avoir utilisé le clustering pour segmenter la base de données, nous avons procédé à la résolution du problème en tentant de formaliser le processus utilisé par la compagnie en un algorithme, et en développant un code Python pour générer une solution initiale qui pourrait être optimisée par la suite.

## 7.1. Description de code :

### 7.1.1. Traitement des données :

La base de données des vols est importée en format csv point-virgule et traité comme précédemment pour les heures de départ et arrivé, et un colonne ‘ID’ est ajouté qui est pratiquement une identification pour les vols pour faciliter la manipulation et l’affichage des résultats

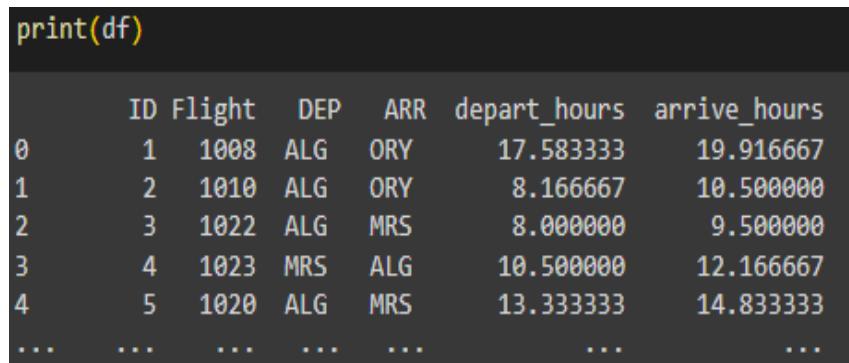


A screenshot of a Jupyter Notebook cell. The code cell contains the command `print(dataset)`. Below it is a table with columns: ID, daydep, dayarr, Flight, DEP, ARR, STD, STA. The data shows five rows of flight information.

	ID	daydep	dayarr	Flight	DEP	ARR	STD	STA
0	1	1	1	1008	ALG	ORY	17:35	19:55
1	2	1	1	1010	ALG	ORY	8:10	10:30
2	3	1	1	1022	ALG	MRS	8:00	9:30
3	4	1	1	1023	MRS	ALG	10:30	12:10
4	5	1	1	1020	ALG	MRS	13:20	14:50
...								

Figure 17: Capture des données avant traitement

Voici les données après le traitement.



A screenshot of a Jupyter Notebook cell. The code cell contains the command `print(df)`. Below it is a table with columns: ID, Flight, DEP, ARR, depart\_hours, arrive\_hours. The data shows five rows of flight information with calculated time differences.

	ID	Flight	DEP	ARR	depart_hours	arrive_hours
0	1	1008	ALG	ORY	17.583333	19.916667
1	2	1010	ALG	ORY	8.166667	10.500000
2	3	1022	ALG	MRS	8.000000	9.500000
3	4	1023	MRS	ALG	10.500000	12.166667
4	5	1020	ALG	MRS	13.333333	14.833333
...						

Figure 18: Capture des données après traitement

### 7.1.2. Logique d'appariement :

Premièrement, nous avons classé les vols en trois catégories (B2NB, NB2B, NB2NB) et les avons triés en fonction de l'heure de départ :

- B2NB : Vol qui décolle d'un aéroport de base vers un aéroport non base.
- NB2B : Vol qui décolle d'un aéroport non base vers un aéroport de base.
- NB2NB : Vol qui décolle d'un aéroport non base vers un autre aéroport non base.

ID	Flight	DEP	ARR	depart_hours	arrive_hours
48	1208	ALG	CDG	5.583333	8.000000
44	6212	ALG	QSF	6.000000	6.500000
54	6180	ALG	ORN	6.000000	7.250000
50	6190	ALG	CZL	6.333333	7.166667
18	6040	ALG	BSK	7.333333	8.166667
...					
ID	Flight	DEP	ARR	depart_hours	arrive_hours
45	1108	QSF	ORY	7.500000	9.833333
55	1076	ORN	TLS	8.500000	10.083333
15	1112	BJA	ORY	9.000000	11.333333
19	1124	BSK	ORY	9.166667	11.750000
39	1124	BSK	ORY	9.166667	11.750000
...					
ID	Flight	DEP	ARR	depart_hours	arrive_hours
51	6191	CZL	ALG	8.166667	9.000000
49	1209	CDG	ALG	9.000000	11.250000
4	1023	MRS	ALG	10.500000	12.166667
53	6173	AAE	ALG	12.166667	13.166667
34	6173X	AAE	ALG	12.166667	13.166667
...					

Figure 19: Capture des ensembles

Ensuite, le processus de séquençage démarre avec les vols B2NB, puis itère à travers les vols NB2NB et NB2B, recherchant le vol suivant qui satisfait les conditions spécifiées par des clauses "if". Si les vols respectent ces contraintes, leurs identifiants sont ajoutés à une liste et le processus se poursuit jusqu'à ce que quatre étapes soient atteintes ou que tous les vols aient été itérés. Cela génère une liste d'appariements comme décrit ci-dessus.

```
▶ print(pairings)
[36, 75, 76], [32, 72], [111, 80, 113], [116, 117], [85, 81], [97, 98, 99, 100],
```

Figure 20: Capture des résultats d'appariement initiale

Nous avons également introduit une colonne booléenne "not\_covered" initialisée à vrai pour suivre la couverture des vols, passant à faux lorsque le vol est affecté. Les séquences sont ensuite importées dans un nouveau dataframe où nous affichons l'aéroport de départ, l'aéroport d'arrivée, l'heure de départ et d'arrivée, ainsi que la période de la séquence.

index	Dep	Arr	DA	AR	Period
0	ALG	ALG	5.583333	11.250000	5.666667
1	ALG	ALG	6.000000	14.666667	8.666667
2	ALG	ORN	6.000000	12.666667	6.666667
3	ALG	ALG	6.333333	9.000000	2.666667
4	ALG	ALG	7.333333	16.333333	9.000000
...	...	...	...	...	...

Figure 21: Capture de nouvelle dataframe

Dans l'étape suivante, le code vérifie s'il existe des séquences réalisables qui commencent et finissent à la base (rotation), les marquant comme prêtes, et classe le reste en trois ensembles : séquences commençant à une base et se terminant hors base (SBNB), séquences commençant hors base et se terminant à une base (SNBB), et séquences commençant et se terminant hors base (SNBNB). Les vols non couverts de l'étape précédente sont importés dans le dataframe et considérés comme une séquence composée d'un seul vol.

Dans la suite du processus, nous introduisons le concept de "layover", où le personnel navigant passe la nuit hors de sa base mère. Cela permet de créer plus de rotations en modifiant une contrainte pour offrir davantage de liberté au code, bien que cela entraîne un coût supplémentaire en termes d'heures hors base, calculé à l'aide d'un compteur intégré

index	Dep	Arr	DA	AR	Period
7	ALG	ORY	8.166667	10.500000	2.333333
2	ALG	ORN	6.000000	12.666667	6.666667
13	ALG	CDG	12.250000	14.750000	2.500000
14	ALG	CDG	12.250000	14.750000	2.500000
15	ALG	CDG	12.250000	14.750000	2.500000
...	...	...	...	...	...

Figure 22: Capture de dataframe sbnb

dans la boucle d'itération. Et finalement, on va exporter les résultats sous Excel après les traiter et organiser dans un dataframe.

```

CrewRoutes=pd.DataFrame()
for i in range (len(output)):
    for j in range (len(output[i])):
        if isinstance(output[i][j], list):
            for k in range (len(output[i][j])):
                a=output[i][j][k]-1
                CrewRoutes=pd.concat([CrewRoutes,dataset.iloc[a:a+1]])
        else:
            a=output[i][j]-1
            CrewRoutes=pd.concat([CrewRoutes,dataset.iloc[a:a+1]])
l=["/",'/', '//', '/ ', '/ ', '/ ', '/ ', '/ ', '/ ', '/ ']
CrewRoutes=pd.concat([CrewRoutes,pd.DataFrame([1], columns=CrewRoutes.columns)],ignore_index=True)
output_file_name = "CrewRoutes.xlsx"
CrewRoutes.to_excel(output file name, index=False)

```

Figure 23; Capture de code de traitement des résultats

Enfin, nous exportons les résultats vers Excel après les avoir traités et organisés dans un dataframe. Les vols et les séquences non couverts sont également exportés vers un autre fichier Excel pour un traitement manuel ultérieur.

ID	daydep	dayarr	Flight	DEP	ARR	STD	\
298	5	5	561	ALG	CMN	2024-06-22 12:00:00	
314	5	5	6104	ALG	ORN	2024-06-22 06:30:00	
315	5	5	6168	ORN	CZL	2024-06-22 08:35:00	
316	5	5	6169	CZL	ORN	2024-06-22 11:05:00	
265	5	5	1078	ORN	LIL	2024-06-22 13:40:00	
...	...	...	...	...	...	...	...

Figure 24: Capture des vols non couvert

## 7..2. Schéma descriptif du Code

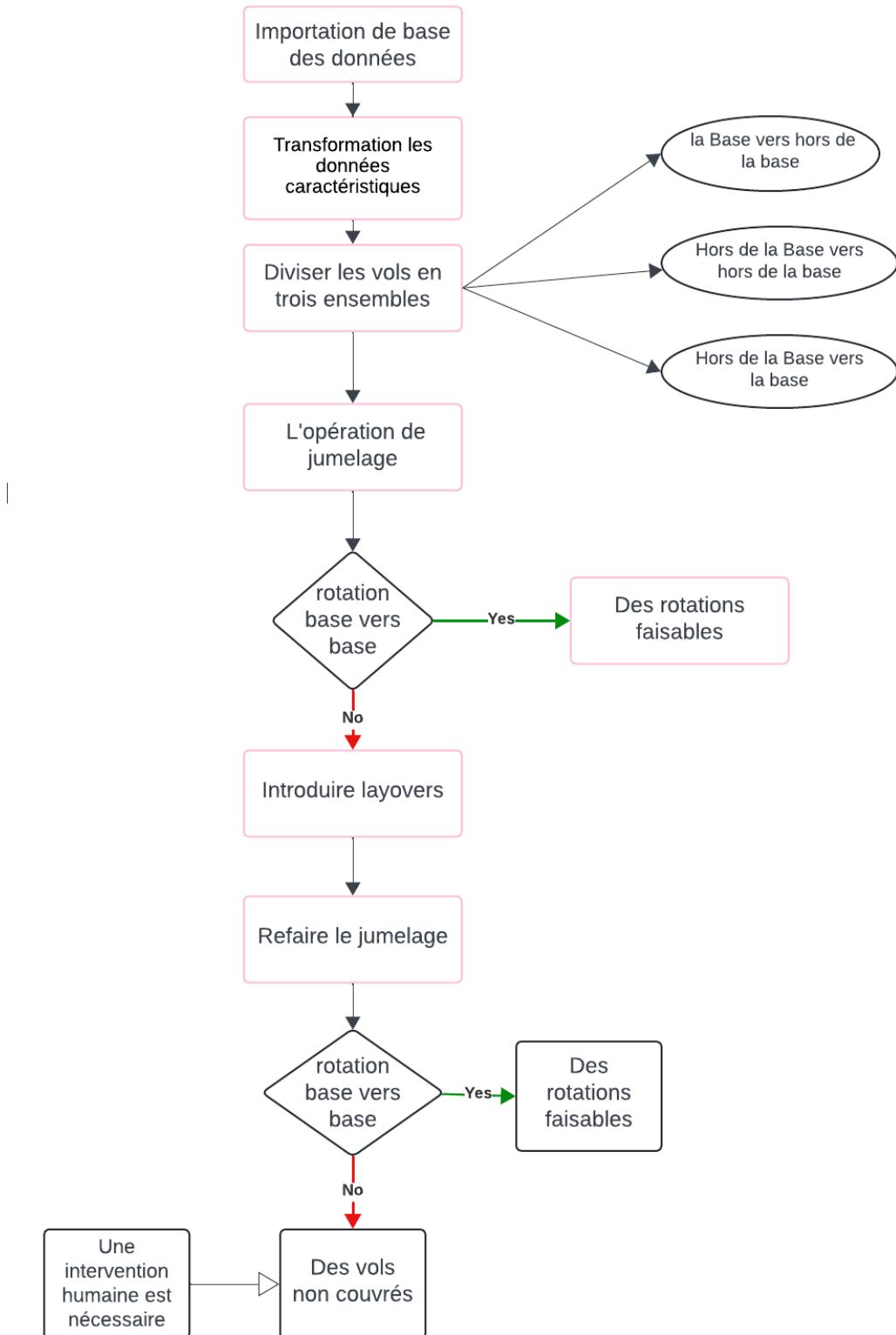


Figure 25: Organigramme de l'algorithme

### 7.3. L'affichage finale de résultat

daydep	dayarr	Flight	DEP	ARR	STD	STA
1	1 1208		ALG	CDG	5:35:00 AM	8:00:00 AM
1	1 1209		CDG	ALG	9:00:00 AM	11:15:00 AM
/	/	/	/	/	/	/
1	1 6212		ALG	QSF	6:00:00 AM	6:30:00 AM
1	1 1108		QSF	ORY	7:30:00 AM	9:50:00 AM
1	1 1109		ORY	QSF	10:50:00 AM	1:10:00 PM
1	1 6213		QSF	ALG	2:10:00 PM	2:40:00 PM
/	/	/	/	/	/	/
1	1 6190		ALG	CZL	6:20:00 AM	7:10:00 AM
1	1 6191		CZL	ALG	8:10:00 AM	9:00:00 AM
/	/	/	/	/	/	/
1	1 6040		ALG	BSK	7:20:00 AM	8:10:00 AM
1	1 1124		BSK	ORY	9:10:00 AM	11:45:00 AM
1	1 1113		ORY	BJA	12:35:00 PM	2:50:00 PM
1	1 6051		BJA	ALG	3:50:00 PM	4:20:00 PM
/	/	/	/	/	/	/
1	1 6050		ALG	BJA	7:30:00 AM	8:00:00 AM
1	1 1112		BJA	ORY	9:00:00 AM	11:20:00 AM
1	1 1125		ORY	BSK	12:45:00 PM	3:15:00 PM
1	1 6041		BSK	ALG	4:15:00 PM	5:05:00 PM

Figure 26:L'affichage finale de résultat

Avec ce fichier qui représente un tableau de données relatif aux rotations il peut aider l'entreprise dans la création des Crew route final, parmi les avantages de ce tableau :

- le tableau est claire et facile à comprendre
- les rotations sont classé selon l'ordre chronologique
- tous les rotations se terminant dans la base

Pour l'inconvénient le tableau compte des rotations de deux étapes d'où à cause de la contrainte de repos qui demande un temps de repos entre chaque vol de 30 min jusqu'à 2 heure et selon le calendrier les vols qui suivant seront après 2 heure ce qui demande intervention humaine

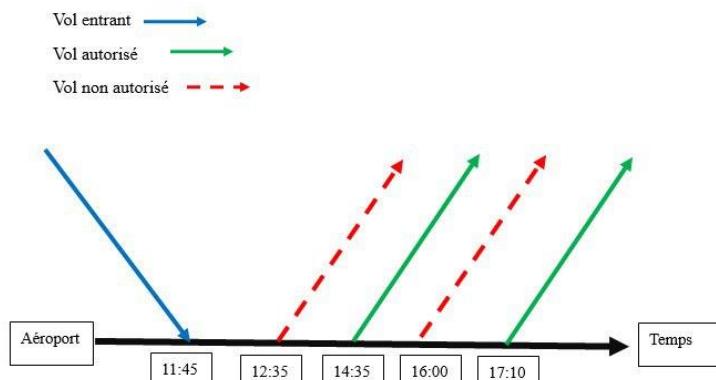


Figure 27:Illustration d'un scénario de vol entrant

## 8. Discussion des résultats

Le code a généré 750 Crew Routes (669 sans layovers) avec 511.41667h hors base et 400 vols qui n'ont pas été affectés. Après avoir terminé l'on a fait une comparaison des résultats entre notre travail et le travail réalisé par l'entreprise :

Tableau 7:Tables des résultats

	Crew route	Algorithme
Nombre des rotation	747	750
Heures hors de la base	Plus de 5000 heures	Plus de 511 heures
Nombre des vols non couvertes	0	400
Le temps d'exécution	De 5 à 10 jours	Moins de 5 minutes

On a comme des résultats :

- le nombres des rotations faisables pour équiper par les équipages
- le nombre des heures hors de la base
- le nombre des vols non couverts par les équipages
- le temps d'exécution pour l'opération de jumelage

**Réduction en les heures hors base** : ne première observation essentielle de notre algorithme d'optimisation est sa capacité à générer des rotations avec des heures hors base considérablement réduites par rapport à celles créées par l'entreprise. Cela souligne l'importance cruciale de l'utilisation d'algorithmes d'optimisation pour résoudre ce problème, non seulement pour réduire les coûts associés aux indemnités de repas et d'hôtel pour les équipages, mais aussi pour améliorer le confort des pilotes et des hôtesses de l'air.

**Nécessité d'une intervention humaine pour les vols non couverts** : Une deuxième observation concernant l'algorithme révèle que sur les 2391 vols analysés, 400 vols n'ont pas été affectés à des rotations, nécessitant une intervention humaine. Ces vols ne peuvent pas être intégrés dans des rotations en raison de contraintes spécifiques, comme par exemple le vol ALG vers ABJ qui laisse un intervalle de 5 jours avant de retourner à

l'Algérie. Dans la pratique, l'entreprise utilise des solutions telles que des vols irréguliers ou le transport des équipages en tant que passagers pour maintenir la continuité du service.

ALG	BSK	C1	ALG	BSK	C1	ALG	BSK	C1	ALG	BJA	C2
BSK	ORY		BSK	ORY		BSK	ORY		BJA	ORY	
ORY	BSK		ORY	BSK		ORY	BJA		ORY	BSK	
BSK	ALG		BSK	ALG		BJA	ALG		BSK	ALG	
Air Craft		Crew route				l'algorithme					

Figure 28: Example de Différences dans la structure des rotations

**Différences dans la structure des rotations :** Le nombre de rotations générées par l'algorithme reste presque le même malgré la présence de vols non couverts. En comparaison avec l'entreprise, où une rotation peut comporter jusqu'à 6 à 8 étapes sur plusieurs jours avec des heures hors de la base, notre approche vise à minimiser les heures hors de la base en créant des rotations de 2 à 4 étapes par jour et en assurant le retour vers la base.

**Gain de temps dans l'exécution :** Comme dernière observation, nous notons que l'algorithme d'optimisation permet de réduire significativement le temps d'exécution par rapport à la méthode manuelle utilisée par l'entreprise. Cette réduction du temps de traitement constitue un succès dans notre objectif initial. En conclusion, l'algorithme présente plusieurs avantages par rapport à la méthode manuelle de création de rotations d'équipages aériens : réduction des heures hors base et des coûts associés, amélioration du confort des équipages et gain de temps dans l'exécution de l'opération de jumelage. Cependant, il est important de noter que l'algorithme ne parvient pas à couvrir l'intégralité des vols, nécessitant une intervention humaine pour gérer les vols non affectés.

## **9. Discussion de notre travail :**

D'après les résultats obtenus et notre expérience acquise dans l'entreprise, notre travail présente plusieurs avantages et points forts significatifs. La méthode de clustering, par exemple, peut réduire efficacement le temps d'exécution du code destiné à résoudre le problème. Bien que dans notre cas spécifique l'impact de la division des données ne soit pas très prononcé, il est clair que pour des bases de données plus vastes, cet impact serait significatif. De plus, l'introduction du critère temporel dans le processus d'appariement des vols s'avère extrêmement utile pour accélérer la planification et minimiser les retards, tout en réduisant de manière significative les coûts supplémentaires associés aux heures hors base des équipages.

Un autre avantage majeur est la généralité du code proposé, ce qui le rend applicable à divers contextes opérationnels. Cependant, malgré ces points forts, notre travail présente également des limitations et des inconvénients. En tant que non-experts en Machine Learning, nous avons constaté que le code proposé est actuellement conçu pour une seule base, comme 'ALG', ce qui pourrait potentiellement affecter les résultats étant donné que la majorité des personnels navigants ont 'ALG' comme base principale, mais d'autres bases sont également pertinentes.

De plus, la nécessité de préparer les données dans un format spécifique, tel que CSV avec des points-virgules, a ajouté une couche de complexité en nous obligeant à manipuler la base de données à l'aide d'Excel pour assurer son adéquation. Cela a entraîné un travail supplémentaire, notamment la vérification minutieuse des vols, ce qui n'est pas toujours pratique.

Un autre point faible réside dans le fait que malgré l'automatisation introduite, une intervention humaine reste nécessaire pour gérer des cas particuliers qui requièrent des procédures complexes, comme l'envoi d'équipages en double ou en tant que passagers. Enfin, une optimisation plus avancée des résultats obtenus, par exemple à l'aide d'algorithmes génétiques ou de méthodes de recherche locale, n'a pas été mise en œuvre, ce qui pourrait potentiellement améliorer les performances du modèle.

En conclusion, notre travail a permis de démontrer les avantages substantiels de l'utilisation de méthodes d'optimisation basées sur le clustering dans la planification des rotations

d'équipages aériens. Toutefois, pour une adoption plus large et une efficacité maximale, il serait nécessaire de résoudre les limitations mentionnées et d'explorer des méthodes d'optimisation plus avancées dans les futures recherches.

## **10. Conclusion :**

En conclusion, après cette expérience dans le domaine du Machine Learning, il est important de souligner que le ML, tel qu'il est aujourd'hui, ne peut pas résoudre ce problème de manière isolée. Cependant, il peut apporter une valeur significative et jouer un rôle crucial dans l'optimisation et la gestion intelligente de ce type de problèmes complexes. Ce processus nécessite du temps, de l'engagement et une compréhension approfondie pour intégrer efficacement les algorithmes ML dans les opérations réelles et obtenir des résultats optimisés.

# **Conclusion Générale :**

Le travail présenté dans ce mémoire, réalisé dans le cadre du stage de fin d'études au sein de l'entreprise Air Algérie, est le fruit d'une réflexion approfondie. L'objectif principal était de répondre à la problématique posée par notre responsable de stage : optimiser les rotations des vols afin d'augmenter l'utilisation des ressources et d'améliorer la planification des équipages en termes de temps et de coût. Pour ce faire, nous avons appliqué des techniques de machine learning ainsi que d'autres algorithmes d'optimisation. C'est avec un réel enthousiasme que nous nous sommes investis dans ce projet ambitieux, d'une importance capitale pour l'entreprise.

Au cours de ce travail, nous avons exploré plusieurs approches de machine learning, notamment les réseaux de neurones artificiels (ANN). Cependant, les résultats obtenus avec cette méthode n'ont pas été satisfaisants en raison de l'incompatibilité des données avec l'intuition des ANN. Nous avons ensuite adopté l'approche de clustering pour réduire la complexité du problème et proposer une méthode générant des résultats proches de la réalité, tout en réduisant le temps de traitement et les coûts supplémentaires, tels que les heures hors base.

Notre méthode présente toutefois certaines limites. Elle ne prend pas en compte les contraintes liées aux horaires de décollage et d'atterrissage. De plus, nous avons défini l'aéroport d'Alger (ALG) comme unique base en raison du nombre important de vols qui y décollent et atterrissent. Par ailleurs, notre méthode ne traite qu'un seul type d'avion, ce qui signifie qu'elle ne peut pas gérer simultanément plusieurs types d'appareils. Les résultats obtenus ne sont pas définitifs et peuvent être améliorés.

Pour améliorer le temps d'exécution, nous recommandons de varier l'algorithme responsable de la construction des appariements des équipages ou d'apporter des modifications à l'approche de machine learning utilisée pour le regroupement des vols. Cet algorithme pourrait être combiné avec d'autres algorithmes dans des travaux futurs afin d'optimiser davantage les appariements mensuels.

Cette expérience dans le domaine de l'intelligence artificielle nous a permis d'explorer une petite partie de ce vaste domaine et de réaliser son immense potentiel. Bien que la mise en œuvre de cette technologie soit exigeante et prenne du temps, elle est extrêmement gratifiante. Par exemple, le processus de planification d'équipage, qui prend actuellement un mois, pourrait être automatisé avec un modèle intelligent. La création de ce modèle est une tâche chronophage, mais elle permettrait de réaliser des économies significatives à long terme. Avec des technologies telles que les ANN, CNN, RNN et le Q-learning, combinées à une base de données bien structurée et une compréhension approfondie des problématiques et des méthodes d'optimisation, le seul obstacle à l'existence de ce modèle est le temps.

Enfin, il est important de noter que des économies de seulement 1 % pourraient entraîner une augmentation significative du chiffre d'affaires annuel, de l'ordre de plusieurs dizaines de millions de dinars algériens.

## Bibliographie

- [1] D. Aggarwal, D. K. Saxena, M. Emmerich et S. Paulose, «On Large-Scale Airline Crew Pairing Generation,» *IEEE Symposium Series on Computational Intelligence (SSCI)*, pp. 593-600, 2018.
- [2] «Historique de Air Algérie,» [En ligne]. Available: [https://web.archive.org/web/20100107022626/https://airalgerie.dz/info/historique\\_f.htm](https://web.archive.org/web/20100107022626/https://airalgerie.dz/info/historique_f.htm). [Accès le 13 Avril 2024].
- [3] «Air Algérie,» [En ligne]. Available: [https://fr.wikipedia.org/wiki/Air\\_Alg%C3%A9rie](https://fr.wikipedia.org/wiki/Air_Alg%C3%A9rie). [Accès le 21 Avril 2024].
- [4] A. Benali, «Air Algérie lance le paiement en ligne par carte CIB,» 2 juillet 2019. [En ligne]. Available: <https://www.algerie-eco.com/2019/07/02/air-algerie-lance-le-paiement-en-ligne-par-carte-cib/>. [Accès le 21 avril 2024].
- [5] «NOTRE FLOTTE,» Air Algérie, [En ligne]. Available: <https://airalgerie.dz/experience-voyage/notre-flotte/>. [Accès le 21 Avril 2024].
- [6] B. I. M. L. Meriem, *Optimisation Opérationnelle des méthodes de contrôle de la programmation flotte - personnel navigant*, USTHB, 2023.
- [7] Air Algérie, *Document interne*, 2023.
- [8] «Air Algérie a transporté plus de 7 millions de passagers en 2023,» Algérie Presse Service, 23 Janvier 2024. [En ligne]. Available: <https://www.aps.dz/economie/165725-air-algerie-a-transporte-plus-de-7-millions-de-passagers-en-2023>. [Accès le 22 Avril 2024].
- [9] «Nos destinations,» Air Algérie, [En ligne]. Available: <https://airalgerie.dz/découvrir/nos-destinations/>. [Accès le 22 Avril 2024].
- [10] AIMS, [En ligne]. Available: <https://www.aimsairlinesoftware.com/>. [Accès le 26 Avril 2024].
- [11] Huawei Technologies Co., Ltd., Artificial Intelligence Technology, China: Springer Singapore, 2022.
- [12] «L'histoire de l'intelligence artificielle,» 22 juin 2023. [En ligne]. Available: <https://www.ia-insights.fr/histoire-de-l-intelligence-artificielle/>. [Accès le 19 avril 2024].
- [13] K. Kaygusuz, *Interdisciplinary Studies on Contemporary Research Practices in Engineering in the 21st Century II*, Karadeniz Teknik Üniversitesi, 2023.
- [14] C. B. D. H. M. J. M. Thomas Dietterich, *Introduction to machine learning*, Massachusetts Institute of Technology, 2010.

- [15] «Qu'est-ce que l'apprentissage non supervisé ?,» [En ligne]. Available: <https://www.ibm.com/fr-fr/topics/unsupervised-learning>. [Accès le 19 Avril 2024].
- [16] Shaveta, A reviewnousmachine learning, Punjab, 2023.
- [17] J. Robert, «La régression linéaire avec Python : Comment ça fonctionne ?,» 19 Janvier 2024. [En ligne]. Available: <https://datascientest.com/regression-lineaire-python>. [Accès le 19 Avril 2024].
- [18] «Régression symbolique: la méthode d'apprentissage automatique oubliée,» [En ligne]. Available: <https://ichi.pro/fr/regression-symbolique-la-methode-d-apprentissage-automatique-oubliee-189460509261205>. [Accès le 19 Avril 2024].
- [19] D. Parrochia, «Classification,» [En ligne]. Available: <https://iep.utm.edu/classification-in-science/>. [Accès le 20 Avril 2024].
- [20] T. Menzies, «Beyond Data Mining,» *Software, IEEE*, vol. 30, 2013.
- [21] «K-Nearest Neighbor(KNN) Algorithm,» 25 Janvier 2024. [En ligne]. Available: <https://www.geeksforgeeks.org/k-nearest-neighbours/>. [Accès le 20 Avril 2024].
- [22] A. A. Awan, «Naive Bayes Classification Tutorial using Scikit-learn,» Mars 2023. [En ligne]. Available: <https://www.datacamp.com/tutorial/naive-bayes-scikit-learn>. [Accès le 20 Avril 2024].
- [23] A. Crochet-Damais, «Random forest (ou forêt aléatoire) : définition et cas d'usage,» 30 May 2022. [En ligne]. Available: <https://www.journaldunet.fr/intelligence-artificielle/guide-de-l-intelligence-artificielle/1501905-random-forest-ou-foret-aleatoire/>. [Accès le 20 Avril 2024].
- [24] «Gradient Boosting in ML,» 31 Mars 2023. [En ligne]. Available: <https://www.geeksforgeeks.org/ml-gradient-boosting/>. [Accès le 21 Avril 2024].
- [25] «Clustering in Machine Learning,» 20 Mars 2024. [En ligne]. Available: <https://www.geeksforgeeks.org/clustering-in-machine-learning/>. [Accès le 21 Avril 2024].
- [26] «KMeans,» [En ligne]. Available: <https://scikit-learn.org/stable/modules/generated/sklearn.cluster.KMeans.html>. [Accès le 22 Avril 2024].
- [27] R. Kassel, «Dendrogramme : tout sur le diagramme de clustering hiérarchique,» 7 Avril 2021. [En ligne]. Available: <https://datascientest.com/dendrogramme>. [Accès le 21 Avril 2024].
- [28] "Applications of Machine learning," [Online]. Available: <https://www.javatpoint.com/applications-of-machine-learning>. [Accessed 21 Avril 2024].
- [29] Z. Tazrout, «Le défi d'Airspace Intelligence : préserver l'environnement et optimiser le trafic aérien grâce au machine learning,» 12 Juillet 2021. [En ligne]. Available:

<https://www.actuia.com/actualite/le-defi-dairspace-intelligence-preserver-lenvironnement-et-optimiser-le-trafic-aerien-grace-au-machine-learning/>. [Accès le 22 Avril 2024].

- [30] M. D.-P. J.-C. R. Giovanni Gatti Pinheiro, Optimizing Revenue Maximization and Demand Learning in Airline Revenue Management, 2022.
- [31] IATA, «Airline Maintenance Cost Executive,» 2022.
- [32] Mitsubishi Heavy Industries, «5 Strategies That Help Airlines Reduce The Cost Of Maintenance,» Forbes, 18 Septembre 2019. [En ligne]. Available: <https://www.forbes.com/sites/mitsubishiheavyindustries/2019/09/18/5-strategies-that-help-airlines-reduce-the-cost-of-maintenance/?sh=7d292e962fce>. [Accès le 23 Avril 2024].
- [33] I. K. J. ,. K. Maren David Dangut, «A rare failure detection model for aircraft predictive maintenance using a deep hybrid learning approach,» *Neural Computing and Applications*, vol. 35, p. 2991–3009, 2023.
- [34] F. Mohammad, «How machine learning is transforming airline operations,» 09 Decembre 2022. [En ligne]. Available: <https://aws.amazon.com/blogs/industries/how-machine-learning-is-transforming-airline-operations/>. [Accès le 23 Avril 2024].
- [35] F. S. L.-J. Yassine Yaakoubi, «Machine learning in airline crew pairing to construct initial clusters for dynamic constraint aggregation,» *EURO JournalnousTransportation and Logistics*, vol. 9, n° %1100020, 2020.
- [36] G. B. P. S. S. R. Graham Wild, «Machine Learning for Air Transport Planning and Management,» *AIAA*, 2022.
- [37] F. J. M. T. Naimeh Borjalilu, «COCKPIT CREW SAFETY PERFORMANCE PREDICTION BASEDnousTHE INTEGRATED MACHINE LEARNING MULTI-CLASS CLASSIFICATION MODELS AND MARKOV CHAIN,» *Aviation*, vol. 27, n° %13, p. 152–161, 2023.
- [38] O. G.-C. R. H. Rodrigo Marcos, «A Machine Learning Approach to Air Traffic Route Choice Modelling,» 2018.
- [39] Y. K. S. D. K. S. Divyam Aggarwal, «On Learning Combinatorial Patterns to Assist Large-Scale Airline Crew Pairing Optimization,» 2020.
- [40] E. L. J. G. L. N. M. W. P. S. P. H. V. Cynthia Barnhart, «Branch-and-Price: Column Generation for Solving Huge Integer Programs,» *Operations Research*, vol. 46, n° %13, pp. 293-432, 1998.
- [41] G. D. ,. A. L. Mouad Morabit, «Machine-Learning-Based Column Selection for Column Generation,» *Transportation Science*, vol. 55, n° %14, pp. 815-967, 2021.
- [42] A. W. G. D. F. S. Frédéric Quesnel, «Deep-learning-based partial pricing in a branch-and-price algorithm for personalized crew rostering,» *Computers & Operations Research*, vol. 138, 2022.

- [43] Numpy, «NumPy documentation,» Numpy, [En ligne]. Available: <https://numpy.org/doc/stable/#numpy-documentation>. [Accès le 24 Avril 2024].
- [44] Pandas, «About pandas,» Pandas, [En ligne]. Available: <https://pandas.pydata.org/about/>. [Accès le 25 Avril 2024].
- [45] TensorFlow, «Why TensorFlow,» [En ligne]. Available: <https://www.tensorflow.org/about>. [Accès le 23 Avril 2024].
- [46] Matplotlib, «Matplotlib: Visualization with Python,» [En ligne]. Available: <https://matplotlib.org/>. [Accès le 26 Avril 2024].
- [47] scikit-learn, «scikit-learn Machine Learning in Python,» scikit-learn, [En ligne]. Available: <https://scikit-learn.org/stable/index.html>. [Accès le 30 Avril 2024].
- [48] Google, «Google Colaboratory,» Google, [En ligne]. Available: <https://colab.google/>. [Accès le 2 May 2024].
- [49] «Python,» [En ligne]. Available: <https://www.python.org>. [Accès le 19 Avril 2024].
- [50] «R Project,» [En ligne]. Available: <https://www.r-project.org/>. [Accès le 19 Avril 2024].
- [51] D. Aggarwal, D. K. Saxena, M. Emmerich et S. Paulose, «On Large-Scale Airline Crew Pairing Generation,» *IEEE*, 2018.

## **الملخص**

حظيت مشكلة تخطيط الطوافم في شركات الطيران باهتمام كبير في السنوات الأخيرة. غالباً ما يتم تحليل هذه المشكلة إلى خطوتين نظراً لحجمها الكبير وتعقيدها المتزايد، أولاً إقران الطوافم ثم تعين الطوافم. يركز هذا العمل على مشكلة إقران الطوافم، والتي تُعد جزءاً أساسياً من تخطيط الطوافم وتتكون من إنشاء تسلسل (أزواج) من الرحلات الجوية، حيث تواجه الخطوط الجوية الجزائرية صعوبات حيث تستغرق وقتاً طويلاً في حل هذه المشكلة بطريقة مجدية. والهدف من ذلك هو تطبيق نماذج التعلم الآلي لحل هذه المشكلة من أجل تحسين استخدام الطوافم. استخدمنا طرق التعلم الآلي لنقليل تعقيد المشكلة من خلال تقسيم البيانات، ثم اقتربنا خوارزمية أعطت نتائج أفضل بكثير وقابلة للتطبيق.

## **الكلمات المفتاحية**

تعلم الآلة، مشكلة إقران الطوافم، خوارزمية التجميع، تحسين

## Résumé

Le problème de la planification des équipages dans les compagnies aériennes a fait l'objet d'une attention considérable au cours des dernières années. Ce problème est souvent décomposé en deux étapes en raison de sa grande taille et de sa complexité croissante, d'abord l'appariement des équipages, puis l'affectation des équipages. Ce travail se concentre sur le problème d'appariement d'équipage, également connu sous le nom de "Crew Pairing Problem", qui est une partie essentielle dans la planification des équipages et qui consiste à créer des séquences (jumelages) des vols, où chez la compagnie Air Algérie rencontre des difficultés qui nécessitent beaucoup de temps pour résoudre ce problème d'une manière réalisable. Nous avons utilisé les méthodes de Machine Learning pour réduire la complexité du problème en divisant les données puis nous avons proposé un algorithme qui donnait des résultats beaucoup mieux et réalisables.

## Les mots clés

Machine Learning, problème d'appariement d'équipage, clustering, optimisation

## **Abstract**

The problem of crew scheduling in airlines has received considerable attention in recent years. This problem is often broken down into two steps due to its large size and increasing complexity, first crew pairing, then crew assignment. This work focuses on the crew pairing problem, also known as "CPP ", which is an essential part in crew planning and consists of creating sequences (pairings) of flights, where the company Air Algeria encounters difficulties which require a lot of time to solve this problem in a feasible manner. We used the ML methods to reduce the complexity of the problem instead by splitting the data, then suggested an algorithm that gave much better and feasible results.

## **Keywords**

Machine Learning, crew pairing problem, clustering, optimization