

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

الجمهورية الجزائرية الديمقراطية الشعبية

MINISTRY OF HIGHER EDUCATION

AND SCIENTIFIC RESEARCH



المدرسة العليا في العلوم التطبيقية
École Supérieure en
Sciences Appliquées

وزارة التعليم العالي والبحث العلمي

HIGHER SCHOOL IN APPLIED SCIENCES

المدرسة العليا في العلوم التطبيقية

Mémoire de fin d'étude

Pour l'obtention du diplôme d'Ingénieur

Filière : Génie Industriel

Spécialité : Management industriel et logistique

Présenté par : DEGDEG Nesrine

Thème

Application de l'apprentissage automatique pour un problème de classification des machines de l'entreprise Mega-Papiers

Soutenu publiquement, le 30/ 07/ 2022, devant le jury composé de :

M Fouad MALIKI	MCB	ESSA- Tlemcen	Président
Mme Amina OUHOUD	MCB	ESSA-Tlemcen	Directeur de mémoire
M Hossine BEDJAOUI CHAOUCHÉ		Entreprise Mega-Papiers	Co- Directeur de mémoire
Mme Latéfa GHOMRI	Professeur	Université de Tlemcen	Examineur 1
Mme Zoulikha KOUDAD	Professeur	ESSA-Tlemcen	Examineur 2

Année universitaire : 2021/2022

Dédicace

Je dédié ce travail à moi-même.

A mon âme pure et belle

À moi qui été toujours une femme courageuse

À moi qui été toujours forte devant ce monde J'y serai jamais arrivé sans moi

Je me le dédie bien que ce ne soit pas le meilleur, mais j'ai fait ce qu'il y a de plus beau pour
moi et à ma manière.

Grand amour pour moi-même pour ma force et ma fierté, car je n'ai jamais abandonné et je
me suis toujours relevé quand je suis tombé, je ris face à la cendre, il n'y a de
dieu qu'ALLAH et de beauté que moi

Remerciement

Tout d'abord je remercie ALLAH qu'il m'a donné la patience, la volonté, l'aide et le courage pour accomplir ce travail durant toutes les années d'étude.

Un spécial remerciement pour mes parents qui tiennent une place immense et spéciale dans mon cœur, Papa tu es le meilleur père dans le monde tu es l'homme de ma vie qui m'a protégé, je n'ai pas assez de mots pour t'exprimer combien je t'aime, Maman tu es une femme fabuleuse et unique pour moi tu es la flamme de mon cœur et de ma vie ta prière été le grand secours pour mener à bien mes études je ne sais pas comment je peux te remercier et t'exprimer mes sentiments envers toi.

Merci pour vos efforts, vos conseils et vos sacrifices pour que je puisse poursuivre mes études. Vous êtes la source de mon bonheur, je ne pourrais jamais vous remercier assez parce que c'est grâce à vous que je suis ce que je suis maintenant.

Je remercie aussi mes frères Hicham, Nadjib et Mohammed Imad, vous êtes les meilleurs frères dans le monde, chacun de vous occupe une place spéciale dans mon cœur et je considère chacun de vous comme mon bras droit, je vous aime infiniment merci d'être les meilleurs soutiens pour moi.

Cette partie de remerciement est dédiée à la meilleure personne pour moi qui occupe une place spéciale dans ma vie et dans mon cœur, je ne peux pas exprimer combien je suis heureuse de t'avoir dans ma vie, je n'ai jamais imaginé que je peux trouver une personne comme toi, je ne sais pas où je serais aujourd'hui sans toi. Merci d'être à côté de moi dans mes situations les plus difficiles, et de m'avoir sentie tant spéciale, je remercie ALLAH chaque heure et chaque minute de t'avoir connue, tu mérites un énorme remerciement.

Je remercie maintenant les personnes qui j'ai l'honneur de les connaître durant ma vie :

- ❖ Chahinez NOUI tu es la plus belle et la meilleure sœur que j'ai jamais vue, j'apprécie ce que tu as fait pour moi, je veux te dire que tu es spéciale pour moi, je t'aime infiniment.
- ❖ Douaa KHELALTIA je ne sais pas comment je peux te remercier tu es la meilleure amie que je connais, tu m'as appris le sens de la vraie amitié, tu auras toujours une place spéciale dans mon cœur.
- ❖ Khadidja SEFRAOUI et Majda Mimoun merci d'être mes proches et m'encourager pour faire mon mieux et de m'avoir fait vivre les meilleurs moments de ma vie avec vous.

❖ Je remercie aussi ma sœur yasmine MAMMAR que j'ai connu cette année mais je l'aime beaucoup. Merci énormément pour votre soutien.

❖ Et aussi je remercie tous personne qui je n'ai pas mentionné.

Je remercie infiniment la meilleure encadrante Mme Amina OUHOUD, ce travail ne serait pas accomplis sans votre soutien et votre encouragement, je ne peux pas vous remercier assez, vous montrez à moi qu'il y a de bon en ce monde. Sincèrement vous êtes à moi plus qu'une encadrante vous me donnez le temps, la confiance en soi et l'encouragement nécessaire pour mener à bien ce travail.

Je remercie aussi l'entreprise Mega-Papiers qui m'a accueillir pour faire mon stage et réalisé mon projet de fin d'étude sous la supervision du Mr Houcine BEDJAOUI CHAUCHE.

Résumé :

La nature des systèmes de fabrication est de plus en plus complexe, Comportements dynamiques et parfois même chaotiques. Afin de pouvoir satisfaire efficacement la demande de produits de haute qualité, il est essentiel d'utiliser tous les moyens disponibles. L'apprentissage automatique est un domaine qui a connu des développements rapides en termes non seulement de résultats prometteurs mais aussi de convivialité. Promettant une réponse à de nombreux défis anciens et nouveaux de la fabrication, l'apprentissage automatique est largement discuté par les chercheurs et les praticiens.

Dans ce manuscrit nous avons développé un outil d'aide à la décision en se basant sur l'apprentissage automatique sur des données de l'entreprise MEGA PAPIER plus précisément filiale «Megatherm ». Les informations données par l'entreprise et de l'apprentissage automatiques peuvent aider les responsables de bien choisir entre lesdeux machines afin de réduire le temps de changement ce qui implique livrer le produit au bon moment et avec taux d'apprentissage à 89 % élevé, cela démontre que l'utilisation des données permettre une prédiction plus précise pour le choix de la machine.

Mot- Clés : Méga-papiers, machine learning, apprentissage automatique, prise de décision

Abstract :

The nature of manufacturing systems is increasingly complex, dynamic and sometimes even chaotic. In order to effectively satisfy the demand for high quality products, it is essential to use all available sources. Machine learning is a field that has seen rapid developments in terms of not only promising results but also usability. Promising an answer to many old and new manufacturing challenges, machine learning is widely discussed by researchers and practitioners.

In this manuscript we have developed a decision support tool based on machine learning and data from the company MEGA PAPER more precisely subsidiary "Megatherm". The information given by the company and the machine learning can help the managers to choose well between the two machines in order to reduce the changeover time which implies delivering the product at the right time and with 89% high learning rate.

Key words: Méga-papiers, machine learning, machine learning, decision making.

ملخص:

تزداد أنظمة التصنيع تعقيداً وديناميكية وسلوكياتها فوضوية في بعض الأحيان. من أجل التمكن من تلبية الطلب بشكل فعال على المنتجات عالية الجودة، من الضروري استخدام جميع الوسائل المتاحة. التعلم الآلي هو مجال شهد تطورات سريعة ليس فقط من حيث النتائج الواعدة ولكن أيضاً من حيث سهولة الاستخدام. يعد التعلم الآلي إجابة للعديد من تحديات التصنيع القديمة والجديدة، ويناقش التعلم الآلي على نطاق واسع من قبل MEGA PAPIER الباحثين والممارسين. في هذه المخطوطة، قمنا بتطوير أداة لدعم القرار تعتمد على التعلم الآلي وعلى البيانات من شركة يمكن أن يساعد التعلم الآلي المديرين على الاختيار بين الجهازين لتقليل وقت التغيير مما يعني تقديم "Megatherm". وبشكل أدق شركة فرعية المنتج في الوقت المناسب وبمعدل تعليمي مرتفع بنسبة 89 %، وهذا يوضح أن استخدام البيانات يسمح بالتنبؤ الأكثر دقة لاختيار الآلة.

الكلمات المفتاحية: الورق الضخم، التعلم الآلي، التعلم الآلي، صنع القرار

TABLE DES MATIERES :

RESUME	3
LISTE DES FIGURES	8
INTRODUCTION GENERALE	10
CHAPITRE I : GENERALITES SUR LA GESTION DE LA PRODUCTION	13
INTRODUCTION.....	14
1. DEFINITION DE LA GESTION DE LA PRODUCTION.....	15
2. CONCEPTS DE BASE DE LA PRODUCTION	15
2.1 <i>Produire</i>	15
2.2 <i>Temps voulu</i>	15
2.3 <i>Quantités demandées</i>	16
3. L'UTILITE DE LA GESTION DE LA PRODUCTION.....	16
3.1 <i>Planification de la production</i>	16
3.2 <i>Contrôle de la production</i>	16
3.3 <i>Amélioration du processus</i>	16
3.4 <i>Maintenance des équipements</i>	16
4. AVANTAGES DE LA GESTION DE LA PRODUCTION	16
4.1 <i>Une meilleure qualité</i>	16
4.2 <i>La réduction du taux de rebuts</i>	17
4.3 <i>La Meilleure prise de décision</i>	17
5. POURQUOI LES ENTREPRISES APPLIQUENT LA GESTION DE LA PRODUCTION.....	17
5.1 <i>Réduction des coûts de fabrication</i>	18
5.2 <i>Amélioration de la compétitivité</i>	18
5.3 <i>Atteindre les objectifs de l'entreprise</i>	18
5.4 <i>Amélioration de l'image de marque</i>	18
5.5 <i>Optimisation d'utilisation des ressources</i>	18
6. SYSTEME DE PRODUCTION	18
7. LES MODES DE PRODUCTION	19
7.1 <i>La production continue</i>	19
7.1 <i>Production discontinue</i>	19
7.2 <i>Production flow shop/job shop</i>	20
7.3 <i>Production flow shop</i>	21
7.4 <i>Production job shop</i>	21
7.5 <i>Flux poussé / flux tiré</i>	22
7.5.1 Production à flux poussé.....	22
7.5.1.1 Avantage du flux poussé.....	22
7.5.1.2 Inconvénients du flux poussé	23
7.5.2 Production à flux tiré	23
7.5.2.1 Avantage flux tiré	23
7.5.2.2 Inconvénients du flux tiré.....	23
8. LES TYPES DE PRODUCTION.....	23
8.1. <i>Production par lot</i>	24
8.2. <i>Production unitaire</i>	24
8.3. <i>Production de masse</i>	24
9. DEFINITION DE JUSTE A TEMPS	25
10. ORIGINE DU JUSTE A TEMPS.....	25
11. OBJECTIF DU SYSTEME JUSTE A TEMPS (JAT).....	25

12.	FONCTIONNEMENT DE LA METHODE JUSTE A TEMPS DANS LE STOCK ET LA LOGISTIQUE.....	25
13.	LES AVANTAGES DU JAT	25
13.1.	<i>Réduction du niveau de stock</i>	26
13.2.	<i>Détérioration des produits</i>	26
13.3.	<i>Réduction du temps</i>	26
14.	LES INCONVENIENTS DU JAT	26
14.1.	<i>Retards</i>	26
14.2.	<i>Complexité</i>	26
14.3.	<i>Adoption</i>	26
14.4.	<i>Prix variable</i>	27
15.	CAS D'APPLICATION DU JAT	27
15.1.	<i>Minimiser le flux</i>	27
15.2.	<i>Minimise les stockages</i>	27
15.3.	<i>Facilite le processus de réception et d'expédition</i>	28
15.4.	<i>Faire un contrôle</i>	28
16.	CONCLUSION.....	28
CHAPITRE II : L'INTELLIGENCE ARTIFICIELLE.....		30
INTRODUCTION.....		31
1.	ETAT DE L'ART	31
2.	HISTORIQUE.....	32
3.	DEFINITION DE L'INTELLIGENCE ARTIFICIELLE	34
4.	DOMAINES D'APPLICATION DE L'INTELLIGENCE ARTIFICIELLE.....	35
5.	LES AVANTAGES ET LES RISQUES DE L'INTELLIGENCE ARTIFICIELLE.....	35
5.1.	<i>Les avantages</i>	35
5.2.	<i>Les inconvénients</i>	35
6.	LES CATEGORIES DE L'INTELLIGENCE ARTIFICIELLE	35
6.1.	<i>Deep Learning</i>	36
6.1.1	<i>Domaine d'application</i>	36
6.1.2.	<i>Fonctionnement du Deep Learning</i>	37
6.1.3.	<i>organigramme de réseau de neurone</i>	38
6.2.	<i>Machine Learning</i>	39
6.2.1.	Apprentissage supervisé.....	40
6.2.1.1.	<i>Classification</i>	40
6.2.1.1.1.	Les algorithmes d'apprentissage de classification.....	40
6.2.1.1.1.1.	<i>K-Nearest Neighbor KNN</i>	41
a)	<i>Définition de KNN</i>	41
b)	<i>Principe de fonctionnement de KNN</i>	42
c)	<i>Organigramme du KNN</i>	43
❖	<i>Avantage du KNN</i>	43
❖	<i>Inconvénient du KNN</i>	43
6.2.1.1.1.2.	<i>Regression logistic</i>	43
❖	<i>Avantage</i>	44
❖	<i>Inconvénients</i>	45
6.2.1.1.1.3.	<i>Arbre de décision</i>	45
❖	<i>Avantage de l'arbre de décision</i>	46
❖	<i>Inconvénients de l'arbre de décision</i>	46
6.2.1.1.1.4.	<i>Forêt aléatoire</i>	46
❖	<i>Avantage de l'algorithme des arbres aléatoire</i>	47
❖	<i>Inconvénient de l'algorithme des arbres aléatoire</i>	47
6.2.1.1.1.5.	<i>SVM</i>	48

a) Organigramme SVM.....	49
❖ <i>Avantage du SVM</i>	49
❖ <i>Inconvénient du SVM</i>	50
6.2.1.1.2. Tableau récapitulatif des différents avantages et inconvénients des algorithmes de classification.....	50
6.2.1.2. <i>La régression</i>	51
6.2.2. Apprentissage non supervisé	51
6.2.2.1. <i>Le clustering</i>	51
6.2.2.1.1. <i>K-means (k-moyennes)</i>	51
6.2.3. Apprentissage par renforcement.....	51
6.3. <i>Machine Learning VS Deep Learning</i>	52
7. QUEL EST LE RAPPORT ENTRE L'INTELLIGENCE ARTIFICIELLE ET L'INDUSTRIE 4.0 ?	52
7.1. <i>Industrie 1.0 : La mécanisation</i>	53
7.2. <i>Industrie 2.0 : Production en série (massification)</i>	53
7.3. <i>Industrie 3.0 : Numérisation</i>	53
7.4. <i>Industrie 4.0 : Aujourd'hui</i>	53
7.5. <i>L'avenir de l'industrie 4.0 est l'industrie 5.0</i>	54
8. APPLICATION DE LA MACHINE LEARNING DANS LA FABRICATION.....	54
9. CONCLUSION	55
CHAPITRE III : PRESENTATION DE L'ENTREPRISE.....	56
INTRODUCTION.....	57
1. PRESENTATION DE L'ENTREPRISE	57
2. MISSION DE L'ENTREPRISE	58
3. TYPE DES PRODUITS.....	58
4. MEGA-PRINT	59
4.1. <i>Le processus</i>	59
4.2. <i>Papier offset</i>	60
4.3. <i>Papier pastoral</i>	60
4.4. <i>Papier autocopiant</i>	60
4.5. <i>Papier couché</i>	60
4.6. <i>Papier bristol</i>	61
4.7. <i>Papier adhésif</i>	61
4.8. <i>Papier sécurisé</i>	61
4.9. <i>Les cartons</i>	61
5. MEGA LABEL.....	62
5.1 <i>Domaines d'utilisation des étiquettes</i>	62
6. MEGA-PLOTTER.....	63
7. MEGATHERM.....	64
7.1. <i>Les banques</i>	64
7.2. <i>Les tickets</i>	65
7.3. <i>Le transport</i>	65
7.4. <i>Autres domaines</i>	65
7.5. <i>Le papier thermique</i>	65
7.6. <i>Fonctionnement du processus</i>	66
7.7. <i>Clients MegaPapiers</i>	67
8. CONCLUSION	67
CHAPITRE IV : APPLICATION DE L'APPRENTISSAGE AUTOMATIQUE POUR UN PROBLEME DE CLASSIFICATION DES MACHINES DE L'ENTREPRISE « ROBO 2400 ET ROBO 2401 »	69

INTRODUCTION.....	70
1. METHODOLOGIE.....	70
2. ACQUISITION DES DONNEES.....	71
3. CHOIX DE L'ENVIRONNEMENT DE TRAVAIL.....	72
3.1. <i>Principales caractéristiques du langage Python.....</i>	73
3.2. <i>Environnement de développement pour Python.....</i>	73
3.3. <i>L'environnement de développement Anaconda.....</i>	74
3.4. <i>Environnement Colab</i>	74
3.5. <i>Colab VS Anaconda (Jupyter)</i>	75
3.6. <i>Principaux modules et bibliothèques utilisés dans notre programme.....</i>	75
3.7. <i>Matplotlib VS Seaborn</i>	76
4. PREPARATION ET ANALYSE DE LA BASE DE DONNEES.....	78
4.1. <i>Visualisation globale de data</i>	78
4.2. <i>Affichage des informations statistique sur la Bdd.....</i>	79
4.3. <i>Visualisation par des histogrammes</i>	80
4.4. <i>Affichage de de la matrice de corrélation</i>	81
4.5. <i>Affichage le nuage des points</i>	83
4.6. <i>Division de data sur deux : partie d'entraînement et partie de test.....</i>	84
5. APPRENTISSAGE.....	84
5.1. <i>La matrice de confusion (confusion matrix).....</i>	85
5.2. <i>Score de précision (accuracy score).....</i>	85
5.3. <i>Les algorithmes.....</i>	86
5.3.6. <i>Comparaison entre les algorithmes</i>	89
6. DEPLOIEMENT.....	90
6.1. <i>Gradio</i>	91
6.2. <i>Streamlit.....</i>	91
6.3. <i>Gradio VS Streamlit</i>	91
6.4. <i>Déploiement de notre projet : Gradio</i>	92
6.5. <i>Déploiement de notre projet : Streamlit.....</i>	93
7. CONCLUSION.....	93
CONCLUSION GENERALE	95

Liste des figures :

Figure I. 1 : <i>objectif de la gestion de la production</i>	17
Figure I. 2 : <i>système de production</i>	18
Figure I. 3 : <i>fabrication continu</i>	19
Figure I. 4 : <i>exemple de fabrication discontinu</i>	20
Figure I. 5 : <i>job shop / flow shop</i>	20
Figure I. 6 : <i>fabrication flow shop</i>	21
Figure I. 7 : <i>fabrication job shop</i>	21
Figure I. 8 : <i>division de production en flux poussé et en flux tiré</i>	22
Figure I. 9 : <i>exemple flux tiré</i>	23
Figure I. 10 : <i>types de production</i>	24
Figure I. 11 : <i>exemple de AR SHUTTLE_G2013</i>	28
Figure II. 1 : <i>Historique de l'IA</i>	34
Figure II. 2 : <i>Relation entre IA, machine-Learning et Deep Learning</i>	36
Figure II. 3 : <i>Réseau de neurone</i>	37
Figure II. 4 : <i>organigramme de réseau de neurone</i>	38
Figure II. 5 : <i>Diagramme d'apprentissage</i>	39
Figure II. 6 : <i>Exemple de KNN</i>	41
Figure II. 7 : <i>organigramme KNN</i>	43
Figure II. 8 : <i>exemple de regression logistic</i>	44
Figure II. 9 : <i>arbre de décision</i>	45
Figure II. 10 : <i>Arbre de décision</i>	46
Figure II. 11 : <i>Forêt aléatoire</i>	47
Figure II. 12 : <i>organigramme SVM</i>	49
Figure III. 1 : <i>logo Mega Papiers</i>	57
Figure III. 2 : <i>MegaPapiers</i>	58
Figure III. 3 : <i>logo du Mega Print</i>	59
Figure III. 4 : <i>exemple d'une bobine</i>	59
Figure III. 5 : <i>le carton</i>	61
Figure III. 6 : <i>logo Mega Label</i>	62
Figure III. 7 : <i>exemple d'étiquette</i>	62
Figure III. 8 : <i>logo Mega Plotter</i>	63
Figure III. 9 : <i>rouleau plotter</i>	63
Figure III. 10 : <i>table traçante</i>	63
Figure III. 11 : <i>logo Megatherm</i>	64
Figure III. 12 : <i>exemple d'un ticket de supermarché</i>	65

Figure III. 13 : <i>exemple d'impression du fax</i>	65
Figure III. 14 : <i>atelier de MegaThetm</i>	66
Figure III. 15 : <i>clients internationaux de Megapapiers</i>	67
Figure III. 16 : <i>clients nationaux de Megapapiers</i>	67
Figure IV 1 : <i>Analyse de données et classification</i>	71
Figure IV 2 : <i>base de données</i>	71
Figure IV 3 : <i>Base de données après sélection de features</i>	72
Figure IV 4 : <i>la page de l'environnement Spyder pour Python</i>	73
Figure IV 5 : <i>la page de l'environnement Jupiter pour Python</i>	74
Figure IV 6 : <i>instruction des commande shape et info</i>	78
Figure IV 7 : <i>Résultat d'exécution de shape et info</i>	79
Figure IV 8 : <i>instruction d’Affichage des critères statistique</i>	79
Figure IV 9 : <i>les statistiques relatives à la BDD</i>	80
Figure IV 10 : <i>instruction d’affichage par des histogrammes</i>	80
Figure IV 11 : <i>résultat d’affichage par des histogrammes</i>	81
Figure IV 12 : <i>instruction d’affichage de la matrice de corrélation</i>	81
Figure IV 13 : <i>affichage de la matrice de corrélation</i>	82
Figure IV 14 : <i>instruction d’affichage par du nuage de points</i>	83
Figure IV 15 : <i>résultat du nuage de points</i>	83
Figure IV 16 : <i>les données de test et d’entrainement</i>	84
Figure IV 17 : <i>exemple matrice de confusion</i>	85
Figure IV 18 : <i>résultat K-NN</i>	87
Figure IV 19 : <i>résultat du SVM</i>	87
Figure IV 20 : <i>résulta du regression logistic</i>	88
Figure IV 21 : <i>résulta du réseau de neurone</i>	88
Figure IV 22 : <i>résultat des forêts aléatoires</i>	89
Figure IV 23 : <i>taux d’apprentissage par algorithme</i>	90
Figure IV 24 : <i>comparaison des résultats des algorithmes avec des histogrammes</i>	90
Figure IV 25 : <i>exemple d’affichage de la prise de décision</i>	92

Liste des tableaux :

Tableau II.1 : <i>récapitulatif des différents avantages et inconvénients des algorithmes de classification</i>	50
Tableau II. 2 : <i>machine Learning VS Deep Learning</i>	52
Tableau III. 1 : <i>caractéristique papier offset</i>	60
Tableau IV 1 : <i>Jupyter Notebook VS Google Colab</i>	75
Tableau IV 2 : <i>Matplotlib vs Seaborn</i>	78
Tableau IV 3 : <i>matrice de confusion</i>	86
Tableau IV 4 : <i>résultat Gradio VS streamlit</i>	91
Tableau IV 5 : <i>résultat du streamlit</i>	93



Introduction générale



Introduction générale

Le monde se développe très vite dans tous les domaines et spécialement dans le domaine de l'industrie, ce développement montre que toutes les entreprises visent à améliorer leurs systèmes de production et de fabrication.

Ce développement vient à partir du développement de la gestion de la production du processus de fabrication et de l'entreprise en général, pour plusieurs raisons citant à titre d'exemple : la minimisation des coûts, l'amélioration de l'image et la réputation de l'entreprise, l'optimisation du temps de fabrication et la raison la plus importante, c'est la bonne gestion de l'entreprise et la prise de la meilleure décision.

L'industrie représente l'un des terrains les plus fertiles pour l'intelligence artificielle. En effet, cette dernière trouve pléthore d'applications dans ce secteur d'activité. Et plusieurs facteurs expliquent cela !

C'est le buzz word de ces dernières années : l'Intelligence Artificielle (IA) ! Toutefois, plus qu'un effet de mode, l'IA est une réalité et va de pair avec la transformation numérique de l'industrie.

Production standardisée, multiplication des capteurs, tâches répétitives et chronophages : le milieu industriel gagnerait beaucoup en utilisant plus d'IA. Ces technologies intelligentes visent à améliorer la performance de l'entreprise. L'IA surpasse aujourd'hui l'humain sur certaines tâches. De plus, l'utilisation de l'intelligence artificielle permet d'effectuer d'autres tâches qu'un humain ne peut traiter seul. L'IA optimise et fiabilise la chaîne de production, et permet notamment de dégager du temps pour les salariés qui peuvent se concentrer sur les tâches à plus haute valeur ajoutée.

Dans ce projet de fin d'étude, nous avons impliqué l'IA comme outil d'aide à la décision, plus précisément l'apprentissage automatique (machine Learning) dans la chaîne de production « Megatherm » là où j'ai effectué mon stage pratique.

Notre contribution consiste à développer une application basée sur l'apprentissage automatique qui nous aide à prendre une décision sur le choix de la machine afin de réduire le temps de changement en tenant en compte la durée de fabrication ainsi la quantité à fabriquer.

Pour mener bien ce travail, nous l'avons devisé en quatre chapitres qui sont :

- ❖ Le premier chapitre : Généralités sur la gestion de la production qui comprend les différents concepts et types des systèmes de production et leurs gestions.
- ❖ Le deuxième chapitre : Intelligence artificielle : Ce chapitre explore l'utilisation de l'apprentissage automatique pour classifier les données et ses méthodologies. Nous aborderons également divers concepts importants couvrant les principes fondamentaux du l'apprentissage automatique tels que les types d'apprentissage, les différences principales entre l'apprentissage automatique et l'apprentissage profond pour la classification des données ainsi que les mesures d'évaluation, etc.
- ❖ Le troisième chapitre : Présentation de l'entreprise : Le troisième chapitre : présentation de l'entreprise : dans lequel nous avons essayé de donner un aperçu général sur l'entreprise Mega papier en présentant les différents produits qu'il fabrique ainsi que leurs process de fabrication et les leurs clients potentiels
- ❖ Le quatrième chapitre : Application de l'apprentissage automatique dans la chaîne de production « MEGATHERM ». Ce chapitre est divisé en trois parties, la première partie, c'est la partie analyse des données dans laquelle nous avons effectué une visualisation et amélioration de data. Ensuite, la partie d'apprentissage avec les différents algorithmes d'apprentissage supervisé (classification) et enfin la partie de l'essai et de déploiement qui a été réalisé avec deux interfaces gradio et streamlit.

Et enfin une conclusion qui présente le résumé de notre contribution de recherche et les suggestions de recherche possibles pour les travaux futurs.



Chapitre I : Généralités sur la gestion de la production



Introduction :

Depuis l'existence des entreprises, ils ont essayé de garantir leurs bons fonctionnements et leurs durabilités et aussi d'assurer leurs efficacités. Mais le monde industriel a vu des grands changements en terme d'évolution du marché, ce dernier a passé par 3 phases :

1. La première phase, quand l'offre est inférieure à la demande donc les entreprises produire après vendre et les principales caractéristiques de cette phase, c'est : délai fixé par le cycle de production, fabrication en série
2. La deuxième phase, c'est quand l'offre et la demande sont égaux donc le client a le choix du fournisseur donc les entreprises doit produire ce qui sera vendu et elles doivent faire des prévisions.
3. La troisième phase, c'est quand l'offre est supérieur à la demande où il y a une concurrence entre les entreprises et le client devient très exigeant, cette phase a forcé les entreprises pour : la maîtrise des coûts, des délais de livraison courts, des séries de produits personnalisés.... Etc.

Et à ce moment-là l'objectif le plus dominant, c'est « **produire ce qui est déjà vendu** ». Pour cela, les entreprises doivent être réactives ou au moins proactive.

1. **Réactive**, c'est-à-dire être efficace et ça signifie que l'entreprise doit s'adapter rapidement et d'une manière permanente à des besoins de plus en plus divers des produits.
2. **Proactive** signifié que l'entreprise est capable d'influencer le développement du marché et donc proposer des nouveaux produits avant les concurrents.

Et aussi elles doivent organiser leur production de manières à créer des produits de haute qualité, avec une grande diversité et au plus haut coût, et il est nécessaire de convertir des activités indépendantes apparemment en un processus continu en supprimant des processus précieux qui sont non utiles pour le client.

Et c'est là où la notion de la gestion de la production a été créée parce que c'est nécessaire de réduire les coûts de : fabrication, livraison et d'approvisionnement. Mais ce n'est pas assez, il est également nécessaire de réduire les temps de conception du produit par l'utilisation de l'ingénierie simultanée¹. Et de diminuer les temps de circulation et de mise à disposition de l'information, réduire les délais de prise de décision [1]

¹ (L'ingénierie simultanée est une méthode prendre en considération la durée de vie d'un produit à partir de la phase de production initiale de

Aujourd'hui dans le monde industriel et exactement dans la gestion de la production la méthode juste à temps a marqué un tournant majeur, ils visent à organiser les flux à travers la production de juste ce que le client demande donc cette méthode ne fonctionne pas avec les prévisions. [2]

Ce premier chapitre est consacré aux généralités sur la gestion de système de production, la première partie vise à définir le système de production ainsi que sa décomposition et ses caractéristiques.

Ensuite, nous exposons les activités de la gestion de la production, on étudie la complexité des problèmes de la gestion de la production.

Finalement, nous définissons le concept de juste à temps vu que l'entreprise où j'ai effectué mon stage pratique utilise ce concept dans sa gestion de la production.

1. Définition de la gestion de la production :

La production est la transformation de matériaux en quelque chose d'utile et portable. [Gershwin 94]

La notion de système de production doit être prise au sens large de système industriel manufacturier ou de processus à caractère discret, continu ou hybride et doit intégrer le point de vue social, économique, technologique et environnemental. Elle doit aussi intégrer la notion de production de services (hôpitaux, banques, grande distribution, ...) et peut inclure le secteur du bâtiment [PRO 02].

2. Concepts de base de la production :

Quand on dit production c'est-à-dire « produire, en temps voulu les quantités demandées ». Cette notion est la base de la réussite des entreprises.

2.1 Produire :

C'est-à-dire fabrication des biens ou des services pour satisfaire un besoin.

2.2 Temps voulu :

C'est-à-dire respecté les délais et les conditions de temps du client, il signifie aussi réalisé des produits en juste à temps (JAT) ça veut dire que ni trop tôt ni trop tard.

2.3 Quantités demandées :

Pour satisfaire les besoins des clients, il est nécessaire de fabriquer les quantités et respecter la demande du client ou bien faire des prévisions de production correctement. [3]

3. L'utilité de la gestion de la production :

La gestion de la production nécessite de prendre des décisions dans :

3.1 Planification de la production :

Dans cette étape, des ingénieurs prennent la décision le temps où la production doit commencer et aussi à quelle machine l'entreprise doit fabriquer pour organiser la production et mettre le processus plus efficace, et quand la production commence, il faut aussi décider le taux de production pour respecter les délais.

3.2 Contrôle de la production :

Pour assurer le bon fonctionnement du processus de fabrication les ingénieurs de qualité fait un contrôle pour vérifier que le processus respecte les normes de qualité par rapport aux temps d'exécution.

3.3 Amélioration du processus :

Dans ce cas, les ingénieurs de production ou bien on dit les responsables de la production sont obligés de suivre le processus de fabrication pour l'améliorer et le mettre mieux jour après jour, ce suivi est aussi fait pour les opérateurs pour les former par exemple.

3.4 Maintenance des équipements :

Comme l'étape précédente, les ingénieurs sont obligés de faire un suivi sur les opérateurs dans cette étape est obligés de faire un suivi sur les machines et faire des maintenances soit prédictive ou bien au moment où la machine tombe en panne, les coûts de maintenance sont généralement inclus dans le coût total des produits finis.

4. Avantages de la gestion de la production :

Sans doute, on peut dire qu'aujourd'hui la gestion de la production devient un facteur essentiel dans l'industrie, parce que les usines peuvent augmenter la rentabilité, la marge bénéficiaire et ils peuvent aussi réaliser plusieurs autres avantages comme

4.1 Une meilleure qualité :

Dans les entreprises, une bonne gestion de production donne un bon choix des équipements

et des machines avec un bon choix d'une équipe bien formé ce qui donne donc des produits de bon qualité et donc une bonne image dans le marché.

4.2 La réduction du taux de rebuts :

Dans les entreprises, le gaspillage représente une grande menace et un grand danger parce qu'il conduit l'entreprise à la faillite donc la gestion de la production lutte contre ce gaspillage et aide l'entreprise à minimiser le taux de rebut pour minimiser aussi les coûts en général.

4.3 La Meilleure prise de décision :

Avec de la gestion de la production les responsables ou bien les décideurs de prendre la meilleure décision à l'aide des meilleurs méthodes et de meilleure contrôle pour augmenter l'efficacité de l'entreprise en général. [4]

5. Pourquoi les entreprises appliquent la gestion de la production :

Aujourd'hui, les entreprises préfère utiliser la gestion de production dans ces processus pour plusieurs raisons le quel :

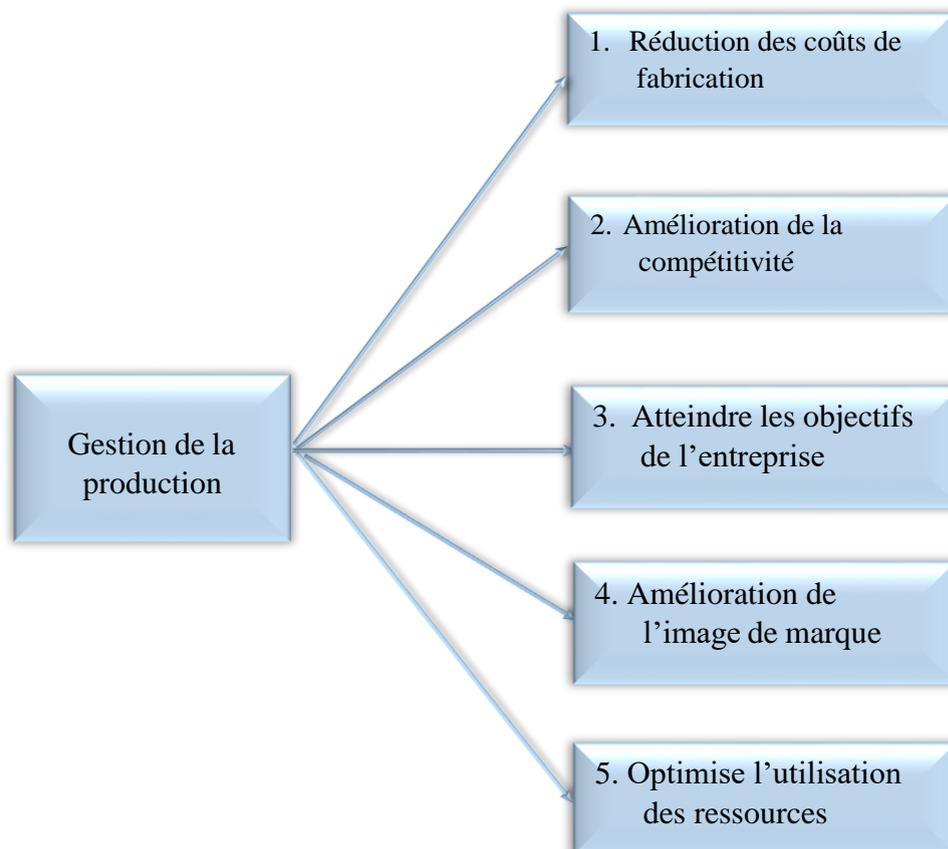


Figure I. 1 : *objectif de la gestion de la production*

5.1 Réduction des coûts de fabrication :

La gestion de la production vise à augmenter le bénéfice et à garantir une valeur ajoutée dans le marché avec un prix concurrentiel du produit à travers la réduction des coûts de fabrication.

5.2 Amélioration de la compétitivité :

Les entreprises cherchent pour rester toujours compétitive avec les autres entreprises, c'est-à-dire que les entreprises visent à produire les bons produits et à les livrer dans les bons moments pour garantir de rester dans la course et à jour avec le marché.

5.3 Atteindre les objectifs de l'entreprise :

Toutes les entreprises cherchent une satisfaction des clients, ce point est atteint par la fabrication des produits avec la bonne qualité et la livraison de ses produits en respectant les délais demandés par le client.

5.4 Amélioration de l'image de marque :

La marque et l'image de l'entreprise sont les deux clés essentiels pour garantir la continuité de l'entreprise, parce que sont les deux facteurs de confiance entre l'entreprise et le client et pour garantir aussi l'existence du nom de l'entreprise dans le marché.

5.5 Optimisation d'utilisation des ressources :

Cela signifie que la gestion de production optimise ou bien minimise l'utilisation des opérateurs et des machines avec un rendement plus. [4]

6. Système de production :

C'est un ensemble construit par des ressources humaines et physiques pour réaliser un objectif et satisfaire un besoin. [3]

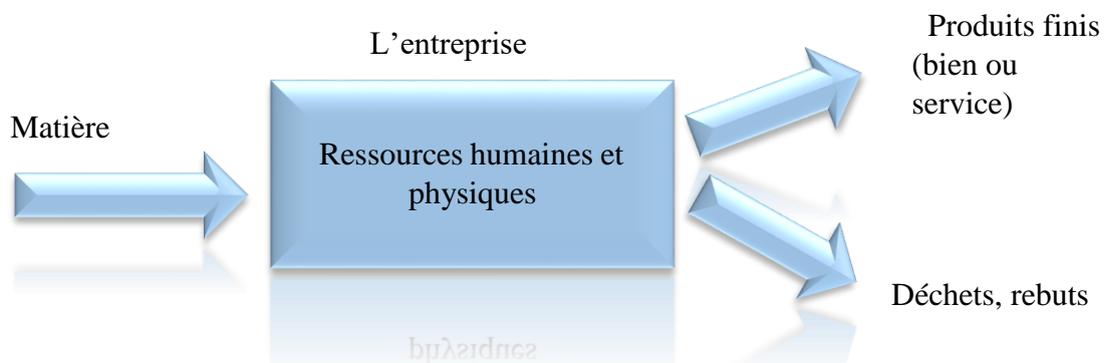


Figure I. 2 : système de production

7. Les modes de production :

On peut dire que le mode de production est la façon que les machines, les opérateurs et les ressources d'une manière générale sont traitée pour fabriquer un produit, on distingue plusieurs modes lesquels : [5]

7.1 La production continue :

C'est-à-dire production sans stockage entre deux postes de travail ce qui signifie que le processus n'est pas séparé entre deux stations de travail consécutif [3]

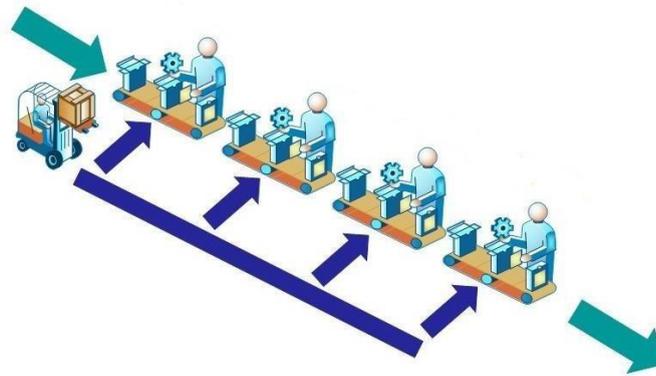


Figure I. 3 : *fabrication continu* [6]

Prenons ce simple exemple de la figure 1.3, il montre que dans la production entre les opérateurs, il n'existe pas un stock donc la production se fait d'une manière continue.

7.1 Production discontinue :

Ou elle est appelée production discrète, le produit suit un processus qui peut être divisé pour permettre de récupérer les produits semi-finis. [3]

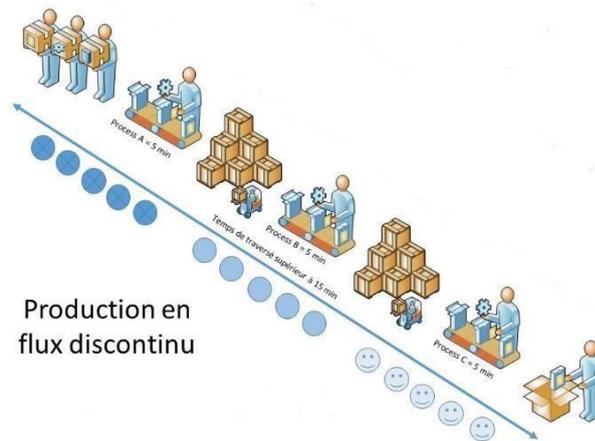


Figure I. 4 : exemple de fabrication discontinu [7]

L'exemple dans la figure 1.4 montre que dans la fabrication discontinue, on trouve des stocks tampon.

7.2 Production flow shop/job shop:

Le type de production est divisé en deux types selon l'organisation des ressources et surtout selon le type et la variété des produits.

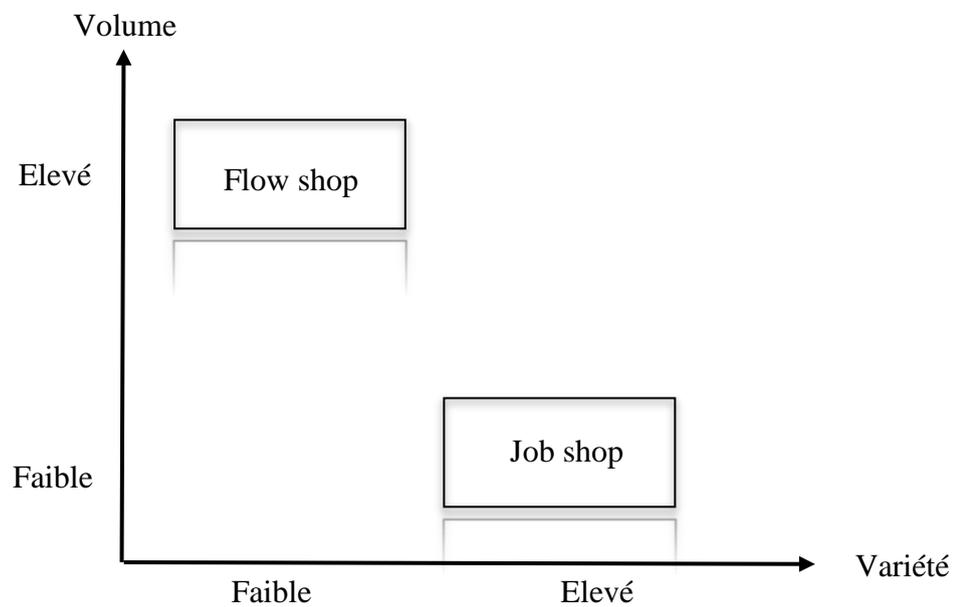


Figure I. 5 : job shop / flow shop

7.3 Production flow shop :

Les produits suivent le même enchaînement du processus avec temps de fabrication différent. [3]

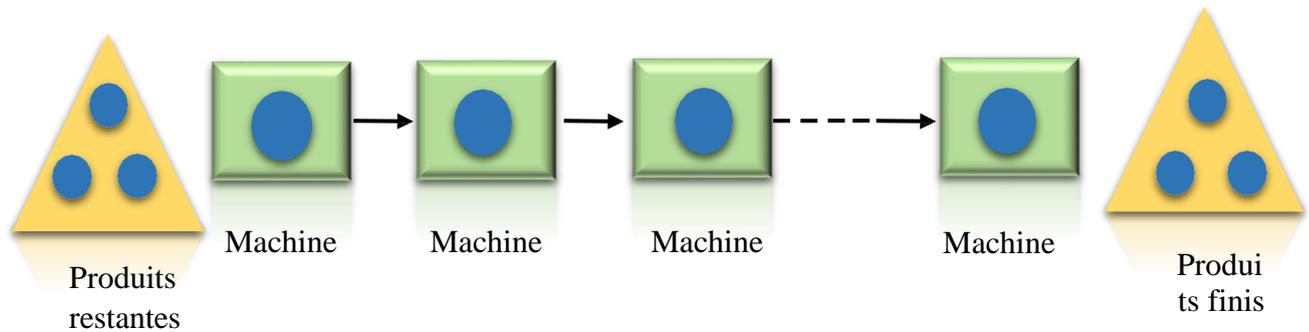


Figure I. 6 : *fabrication flow shop*

Ce type de production est caractérisé par un grand volume de fabrication, mais une variété faible.

7.4 Production job shop :

La sélection des postes de travail se fait d'une manière variable et chaque produit suit un cheminement différent par rapport à un autre type de produit. [3]

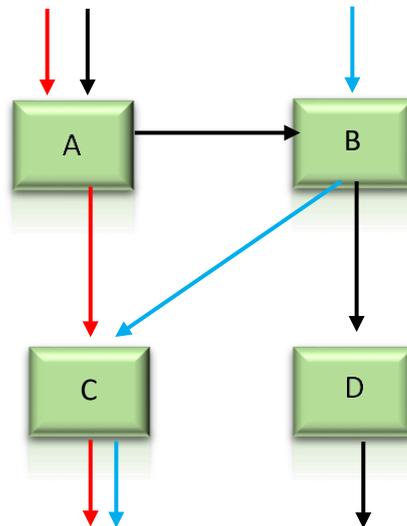


Figure I. 7 : *fabrication job shop*

Ce type de production est caractérisé par une grande variété des produits fabriqués, mais avec des petites quantités.

7.5 Flux poussé / flux tiré :

Les entreprises pour satisfaire le besoin de ces clients cherchent à améliorer leur capacité de production sachant que la performance de production d'une entreprise est en relation avec le taux de production, donc il est nécessaire de prendre une bonne décision dans le choix de type de flux. [8]

Dans la production, il existe deux types de flux : flux poussé et flux tiré.

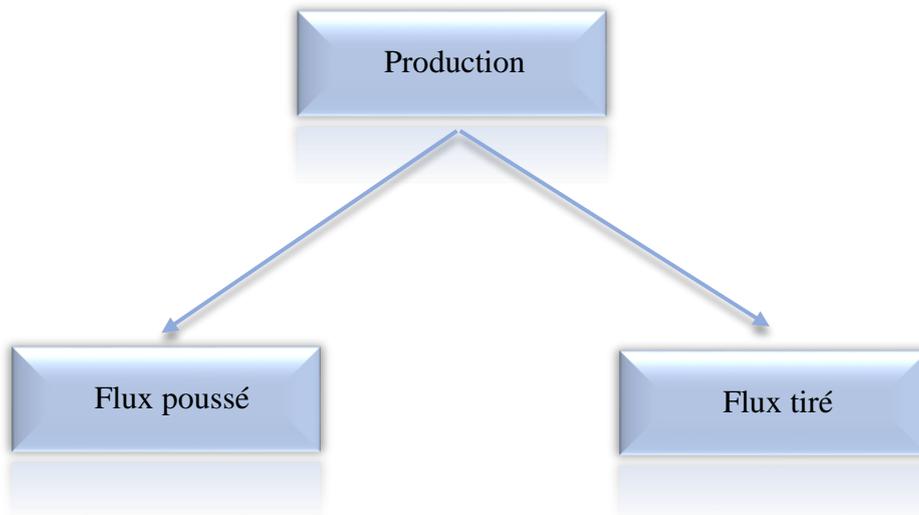


Figure I. 8 : *division de production en flux poussé et en flux tiré*

7.5.1 Production à flux poussé :

C'est-à-dire que les produits sont réalisés à travers des prévisions de la demande, elle est appliquée pour aider les entreprises d'améliorer la qualité du processus de fabrication et satisfaire plus en plus les clients, parce que cette méthode consiste à fabriquer et à satisfaire un besoin avant que le client ne demande.

Cette production est caractérisée par : chaque personne termine une tâche et pousse la tâche suivante pour débiter, tant que cette production suit le principe de production avec des prévisions donc quand le produit est fabriqué, il est stocké.

7.5.1.1 Avantage du flux poussé :

- Temps de livraison réduit.
- Taux de satisfaction des clients augmenté.
- Processus de production accéléré à cause des prévisions
- Temps de livraison réduit.

- Taux de satisfaction des clients augmenté.
- Processus de production accéléré à cause des prévisions

7.5.1.2 Inconvénients du flux poussé :

- Coût de stockage supplémentaire pour l'entreprise.
- Risque d'avoir des produits non vendus en stock.

7.5.2 Production à flux tiré :

Ce qui implique qu'on réalise les produits à partir de la demande des clients et, c'est-à-dire que c'est l'inverse de la production flux poussé, donc dans ce type le besoin et la demande du client sont les décideurs de production dans les entreprises donc il n'y a pas de stock et donc il n'y a pas des frais supplémentaires, elle est caractérisée par le JAT (juste à temps) c'est-à-dire produire au moment où la commande arrive.

7.5.2.1 Avantage flux tiré :

Optimisation de ses coûts.

- éviter la surproduction.
- Améliorer la gestion des stocks. [9]

7.5.2.2 Inconvénients du flux tiré :

- Risque de ne pas respecter les délais.
- Risque de fabriquer des produits qui ne sont pas de bonne qualité à cause du délai de fabrication limité.

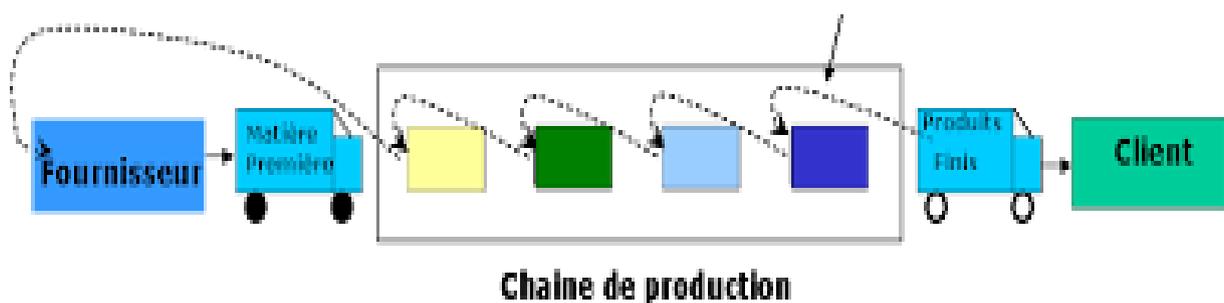


Figure I. 9 : exemple flux tiré [10]

8. Les types de production :

Le choix du type de production est une étape importante liée au type de produit fabriqué dans l'entreprise, on distingue dans ce point-là trois types : production par lot, production unitaire et

production de masse.

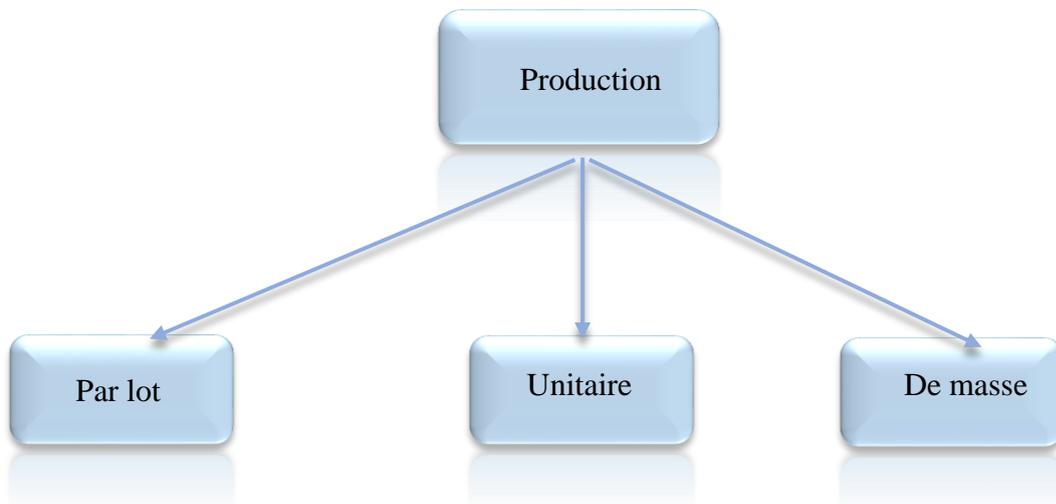


Figure I. 10 : *types de production*

8.1. Production par lot :

Fabriqués des types de produits non pas identique, mais avec même outil de fabrication, les lots de fabrication peut être petit comme il peut être grand, mais le passage d'un lot à un autre nécessite un temps de changement et de configuration ce qui implique le taux d'occupation des machines diminué et donc le taux de revient des produits fabriqués augmente. Par exemple la fabrication des différentes composantes d'un avion, elles sont fabriquées en des petits lots.

8.2. Production unitaire :

La fabrication se faite d'une manière chaque produit avec une unité ou avec petite série pour un besoin spécifique donc c'est un cas particulier de la fabrication par lot et ce type nécessite un nombre de mains d'œuvre avec une durée longue.

Par exemple la fabrication des produits de luxe comme les accessoires et les montres ou bien la fabrication des grand produit comme les avions, les bateaux...etc.

8.3. Production de masse :

Fabrication des produits standard ou bien similaire avec grande quantité. [3] Par exemple la fabrication des voitures ou bien les téléphones portables.

9. Définition de juste à temps :

La méthode JAT (juste à temps) ou bien JIT (just in time), elle est basée sur le principe de produire au bon moment la bonne quantité ou bien, on dit la quantité nécessaire et livré au bon emplacement pour organiser la production et pour réduire les coûts et le taux de gaspillage.

Aujourd'hui, la méthode, juste à temps, est appliquée aussi dans le domaine de la logistique des entrepôts pour les rendre plus efficaces.

10. Origine du juste à temps :

Cette méthode est inventée en 1950 par le japonais Taichi Ohno chez Toyota, dans le premier secteur où cette méthode est appliquée est le secteur automobile, donc Toyota a créé un nouveau système qui s'appelle « Toyota Production System ».

11. Objectif du système juste à temps (JAT) :

Le système inventé par Toyota est basé sur le principe d'éliminer ou bien supprimer toute action ou élément qui n'apporte pas de valeur dans le processus de production. Cette méthode est applicable aussi dans tous les domaines donc aussi dans le côté de matière première ou bien dans les entrepôts, pour réduire les délais, les coûts et les ressources.

Dans ce cas-là on passe du système de type Push ou bien comme on a dit avant flux poussé qui est caractérisée par le surstockage et la surproduction vers des systèmes Pull (flux tiré) pour minimiser les coûts de stockage et les délais. [11]

12. Fonctionnement de la méthode juste à temps dans le stock et la logistique :

Depuis 1950 donc depuis l'existence de cette méthode elle a un objectif fixé qui est la réduction du stockage et les délais dans les entreprises. Dans les premières utilisations de cette méthode, elle a touché directement le processus de production, le stock, la logistique ou bien le transport, le service clientèle et le service après-vente.

Pour appliquer la méthode juste à temps, il faut réduire les distances parcourues dans le processus de fabrication, il faut aussi supprimer les phases ou bien les tâches qui ne sont pas nécessaires, il est nécessaire aussi de réduire les temps de changement et les temps de transfert entre les machines et aussi fait une étude profonde sur le stockage de l'entreprise. [11]

13. Les avantages du JAT :

Après avoir parlé du concept du juste à temps et ces objectifs qui nous aident à obtenir une vision claire sur l'utilité de cette méthode on va aborder ses avantages pour confirmer cette vision :

13.1. Réduction du niveau de stock :

Ce qui implique la réduction des coûts de stockage, ce point est atteint parce que le JAT signifie que l'entreprise fabrique juste la quantité nécessaire à partir de la commande du client et non pas à partir des prévisions, donc dans ce cas-là deux choses sont optimisées : le temps de stockage et la quantité stockée.

13.2. Pas de détérioration des produits :

Par ce que la production est faite juste à partir des commandes des clients donc il n'y a pas des prévisions ou bien des quantités restantes dans le stock, car les entreprises fabriquent juste les quantités nécessaires pour ces clients.

13.3. Réduction du temps :

Le JAT vise à améliorer le processus en général ce qui implique qu'elle vise aussi à réduire les temps de chargement et déchargement des produits par ce que les tâches qui n'ont pas les nécessaires sont supprimées.

14. Les inconvénients du JAT :

Nous savons tous que rien n'est sans défaut donc le juste à temps aussi comme elle a ces avantages elle a aussi ses inconvénients :

14.1. Retards :

La méthode, juste à temps, peut causer des retards ou bien des manques dans le stock, prennent par exemple une pénurie dans le stock de matière première ce qui implique un retard dans la tâche ou bien le processus suivant et ce qui implique finalement un retard dans le délai de livraison au client.

14.2. Complexité :

La méthode, juste à temps, nécessite une forte communication et relation entre le fournisseur et l'entreprise et entre l'entreprise et le client pour appliquer cette méthode correctement.

14.3. Adoption :

Cette méthode doit être adoptée par tous les secteurs de l'entreprise, c'est ce qui signifie qu'elle est vraiment complexe.

14.4. Prix variable :

Il est clair tant que le juste à temps implique que les entreprises achètent les matières premières par la moindre quantité et selon les clients et non pas selon des approvisionnements donc il est aussi clair que le prix varie chaque période donc aussi le prix du produit fini aussi varie ce qui implique des négociations ou bien capable des problèmes entre le client et l'entreprise.

15. Cas d'application du JAT :

Dans cette partie, on donne le cas d'application de la méthode JAT, et pour appliquer cette méthode correctement il faut appliquer les trois étapes suivantes :

15.1. Minimiser le flux :

C'est une étape clé et essentiel pour la conception du layout , car il est très important de créer un bon layout par l'optimisation des distances parcourues pour optimiser les flux des marchandises.

15.2. Minimise les stockages :

Dans ce cas-là, la plupart des entreprises choisissent par exemple, le système de gestion FIFO (first in first out) c'est-à-dire le premier entre est le premier sorte pour simplifier le flux de la marchandise.

Il existe l'autre méthode pour simplifier le flux des marchandises, c'est l'automatisation des entrepôts ce qui signifiait l'élimination des opérateurs et de la main d'œuvre. [11]

Les entreprises peuvent faire un entrepôt semi automatisé comme les navettes pourpalettes – AR Shuttle qui sont des solutions pour transporter les marchandises à l'intérieurde l'entrepôt, sont guidés à distance par quelqu'un pour minimiser le taux des accidents et augmenter l'espace de stockage, aussi ses palettes augmente la capacité de logistique du côté quantitative et d'un côté qualitative. [12]



Figure I. 11 : *exemple de AR SHUTTLE_G2013* [12]

15.3. Facilite le processus de réception et d'expédition :

Dans cette dernière étape, nous soulignons sur l'importance du travail sur la phase de la réception et l'expédition des marchandises, donc avec la méthode juste à temps la réception des marchandises se fait avec des quantités inférieures ce qu'implique que le processus doit être plus rapide.

15.4. Faire un contrôle :

Quand l'entreprise fait une solution de stockage correcte et un Layout adéquat et elle fait aussi un système de chargement et déchargement rapide adapté à la méthode juste à temps, l'entreprise doit assurer la continuité du fonctionnement du système ce qui implique qu'il est nécessaire de faire un bon contrôle d'inventaire. [11]

16. Conclusion :

Chaque entreprise cherche et vise à maximiser leur profit, réduire les coûts, optimiser l'utilisation des ressources pour garder une marque et une bonne image dans le marché. Et ces objectifs sont atteints si et seulement si l'entreprise a un bon processus de fabrication, une meilleure prise de décision, des produits de bon qualité avec le minimum taux de rebuts ...etc. C'est ce qu'on appelle dans ce cas-là la bonne gestion de la production.

Aujourd'hui, la montée en puissance de l'Intelligence Artificielle et du Machine Learning permet d'anticiper les problèmes. On parle alors de gestion proactive qui repose sur une bonne utilisation des données.

Dans le chapitre suivant, nous allons voir les différents concepts liés à l'IA ainsi que les algorithmes utilisés avec leurs avantages et leurs inconvénients et aussi la relation de l'IA

avec l'industrie 4.0.



Chapitre

II

L'intelligence artificielle



Introduction :

Dans le monde aujourd'hui et exactement dans le monde industriel la concurrence entre les entreprises est devenue un facteur clé de leur développement. Chacun essaie de faire évoluer son processus de fabrication et c'est pour cela, on a vu plusieurs inventions et défis comme les usines intelligentes qui sont à cent pour cent automatisé ce qui implique la présence de l'intelligence artificielle dans ces usines et exactement la présence du machine Learning pour faire ce qu'on appelle l'apprentissage à travers une base de données.

Dans ce chapitre, nous présentons ce qu'est l'intelligence artificielle et qu'est l'apprentissage automatique, nous avons également parlé de principe de l'apprentissage et mentionner certains cas d'utilisation, nous avons mentionné les différents types d'apprentissage qui existent avec les différents algorithmes.

1. Etat de l'art :

Le domaine de l'industrie, aujourd'hui, connaît des développements impressionnants en l'absence d'informations et des données préalables (Chand & Davis, 2010) [13]. Dans ce domaine, il existe plusieurs mots-clés et noms qui sont différents d'un pays à un autre comme l'industrie 4.0 en Allemagne et smart manufacturing aux États-Unis. La disponibilité des grandes quantités des données est appelée le Big Data (Lee, Lapira, Bagheri, & Kao, 2013) [14], si on prend l'exemple des données dans le côté de qualité, ces derniers aide les entreprises pour améliorer la qualité des produits où on peut dire la qualité du processus de fabrication en général (Elangovan, Sakthivel, Saravanamurugan, Nair, & Sugumaran, 2015)[15]. On peut dire dans ce point-là la concurrence attaque directement le processus de fabrication et donc attaque les produits, la chaîne d'approvisionnement et l'entreprise en général Lu, Y. (2017) [16]. Et donc la capacité de résistance et la flexibilité sont des facteurs clé pour la compétitivité, telle qu'il est clairement exprimé que si une entreprise a la capacité de changer carrément la façon dont les produits sont fabriqués, distribués, utilisés...etc. Donc on est dans la "fabrication intelligente" (Kusiak, 2018) [17].

Pour qu'une entreprise devienne intelligente, il est nécessaire qu'elle respecte quelques conditions comme qu'elle peut gérer les contrôles informatisés, a des logiciels de gestion de production et le plus important qu'elle applique l'internet des objets (IIoT)...etc. Marquant que ces conditions seules ne sont pas suffisantes et un système de fabrication est considéré intelligent si et seulement s'il fonctionne et régulé par des technologies de contrôle intelligentes pour répondre aux événements internes et externes très rapide (Mittal et al, 2016) [18].

Cette fabrication intelligente doit embrasser le Big Data pour que le système d'information et les logiciels de gestion de la production soient enrichis de compétences analytiques (Waller et Fawcett, 2013) [19], avec une capacité d'apprentissage automatique pour garantir la compétitivité (Monostoria, 2003). [20]

Maintenant, la plupart des entreprises utilisent cet apprentissage pour leurs prises de décision, ces entreprises sont devenues plus performantes que leurs concurrents, plus productifs avec un pourcentage de 5% et plus rentable avec un pourcentage de 6%.(McAfee et al, 2012) [21]. Donc il est l'un des facteurs essentiels pour développer le système de production pour atteindre l'industrie 4.0 (Pham et Afify (2005) [22], mais malheureusement beaucoup des entreprises ne connaît pas l'importance de la collection de données pour faire un bon apprentissage ou bien quel sont les données nécessaires à mesurer à cause du manque de la connaissance (Harding et al, 2006). [23]

Marquant que cet apprentissage automatique est une branche du machine Learning et exactement de l'intelligence artificielle basant sur des données d'entrée et des algorithmes qui sont capable d'apprendre (Lu, 2017 ; Xu et al, 2018). [24] [25]

On peut dire à la fin que l'état de l'art de la machine Learning ces dernières années a marqué une grande évolution comme les voitures à conduite autonome qui sont basées sur des algorithmes d'apprentissage supervisé ou bien les jeux électroniques (Martínez- Díaz et Soriguera, 2018 ; Muller, 2002) [26] [27].

2. Historique :

L'intelligence artificielle est une discipline large et révolutionnaire donc il est vraiment difficile de donné une définition précise et claire, mais on peut la considérer comme une partie et filiale de l'informatique, elle a un but de créer des machines capables de réaliser des taches que l'être humain peut l'effectuer tels que mesurer, calculer, traiter et évaluer.

- Elle débute en 1943 avec la présentation du premier modèle mathématique d'un réseau de neurones dans l'article « A Logical Calculus of Ideas Immanent in Nervous Activity. »
- En 1950 Alan Turing a publié le test de turing qui évalue jusqu'à maintenant les IA et aussi le premier ordinateur à réseau de neurone a été produise par Marvin Minsky et Dean Edmonds (2 étudiants de Harvard).
- En 1952, la réalisation du premier programme qui est capable à jouer les échecs.
- En 1956, le mot intelligence artificiel a été prononcé pour la première fois par John McCarthy

durant la conférence « Dartmouth Summer Research Project on Artificial Intelligence » (la vraie origine de l'IA).

- En 1959, la création du terme « machine Learning » « apprentissage automatique » pendant qu'il travaillait chez IBM (international Business Machine Corporation), et aussi John McCarthy et Marvin Minsky ont créé le MIT (massachusetts Institute of Technology) Laboratoire d'informatique et d'intelligence artificielle.

- Dans les années qui suivent le doute s'installe dans le domaine de l'IA, en 1961 le rapport américain ALPAC (Automatic Language Processing Advisory Committee) souligne l'absence dans la recherche sur la traduction automatique visant à traduire la langue russe dans un cadre de guerre froide.

- En 1963, la création du « AI lab » à l'université de Stanford par John McCarthy.
- En 1973, le ministre britannique présente son rapport « Lighthill » qui souligne les déceptions de la recherche en IA, cette phase de doute dura jusqu'en 1980, et est désormais aujourd'hui décrite comme le "premier hiver de l'IA".

- Ce dernier se terminera avec la réalisation de R1 (XCON²) par Digital Equipment Corporations. Ce système professionnel est destiné pour configurer les commandes des nouveaux systèmes informatique et provoque un pic des investissements qui se prolongera pendant plus d'une décennie.

- En 1987 faces à la naissance d'alternatives moins coûteuses et moins chère le deuxième hiver de l'IA est né et les investisseurs financent la recherche sur l'IA.

- Après 10 ans, les progrès technologiques permirent un renouveau de l'intelligence artificielle, par exemple en 2008 Google a fait plusieurs avancées dans le secteur de la reconnaissance vocal et intégré cette option dans ses applications pour les smartphones.

- En 2012, Andrew Ng avec +10 millions vidéo sur YouTube a réalisé le premier réseau de neurones. Grâce à l'apprentissage profond « Deep Learning, » ce réseau est capable de reconnaissant un chat sans étudier ce que le chat est. Est donc c'est le début d'une nouvelle époque pour un apprentissage profond.

² Est un système à base de règles de production pour sélectionner automatiquement les composants du système informatique en fonction des exigences du client

- En 2016, la réussite de l'IA sur l'homme avec la réussite du système AlphaGo³
- Le Deep Learning et la machine Learning sont maintenant utilisés dans les sociétés de tous les secteurs, pour plusieurs commandes. L'IA continue de progresser, et son rêve, c'est de devenir de plus en plus proche de la réalité. [28]

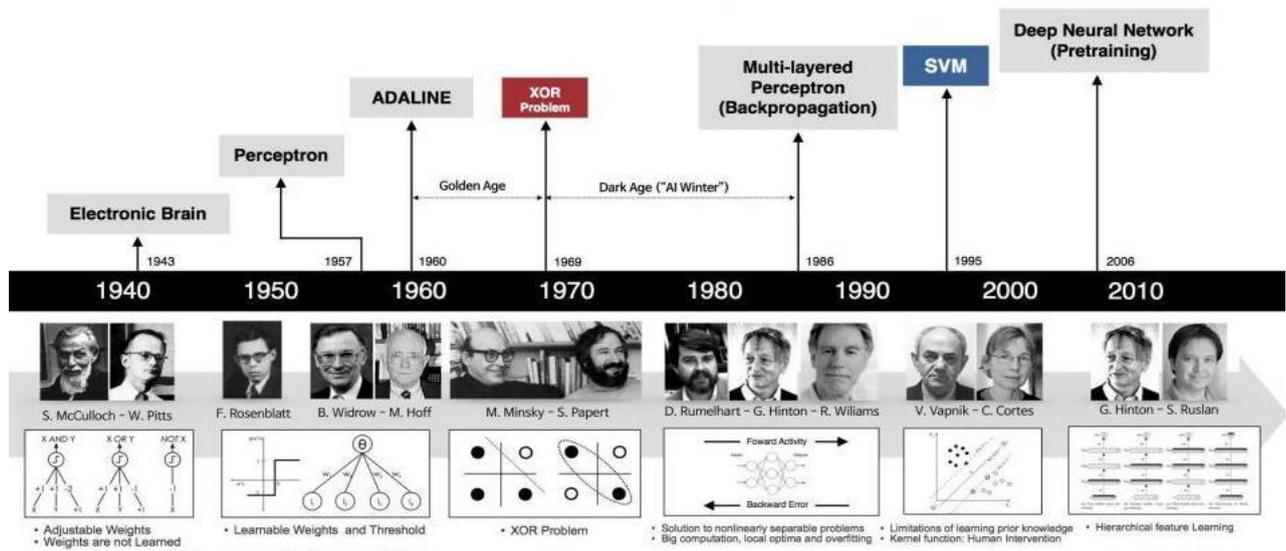


Figure II. 1 : Historique de l'IA

3. Définition de l'intelligence artificielle :

On peut dire premièrement que l'intelligence artificielle jusqu'à maintenant les chercheurs et les scientifiques ne sont pas d'accord sur une définition unique, mais il existe quelque définition dans le domaine.

- Selon ean-Claude Heudin « l'intelligence artificielle visant de donner à la machine une intelligence réelle à travers des techniques ».
- selon Eric Sibony « Machine", ne désigne pas un objet physique, mais plutôt "un système automatique capable de traiter de l'information ». [29]
- Selon E.Riche « l'IA est le domaine de l'informatique qui étudie comment faire à l'ordinateur des tâches pour lesquelles l'homme est aujourd'hui encore meilleur ».
- Selon M.Minsky « l'IA est une discipline scientifique qui consiste à faire à des machines ce que l'homme fait moyennant une certaine intelligence ». [30]

³ Est un logiciel de l'IA élaboré par la compagnie Britannique DeepMind a un but de déterminer les meilleurs joueurs de GO et c'est le premier jeu don lequel l'être humain a gagné contre la machine.

4. Domaines d'application de l'intelligence artificielle :

Personne ne peut ignorer que l'intelligence artificielle à dominer le monde, citant quelques domaines :

- Domaine de la santé (les appareils médicaux pour les diabétiques par exemple alerte seule quand il y a un risque).
- Domaine de l'industrie (l'internet des objets qui connecte les machines entre eux).
- Domaine de sécurité (l'empreinte ou bien reconnaissance faciale).
- Domaine du transport (comme on a dit dans l'état de l'art la révolution des voitures autonome).
- ...etc.

5. Les avantages et les risques de l'intelligence artificielle :

Comme on a dit dans le point précédent l'intelligence artificielle à dominer le monde et deviens un facteur clé pour tous les développements, comme tout, elle a ses avantages et ses inconvénients, commençons par :

5.1. Les avantages :

- rendre la vie plus facile.
- L'intelligence artificielle sauve les vies (dans le domaine de la santé).
- Digitalise le monde et facilitent la communication entre les machines.
- Rendre le monde très petit comme les réseaux sociaux.
- Minimise les couts des mains d'œuvre dans les entreprises.
- ---etc.

5.2. Les inconvénients :

- Risque de sécurité.
- Risque de destruction de l'humanité.
- Automatisation des armes.
- Une violation de l'intimité.
- ...etc.

6. Les catégories de l'intelligence artificielle :

Passant maintenant de l'historique de l'IA vers les domaines qui composent l'intelligence artificielle et qui sont le Deep Learning et la machine Learning.

Cette figure montre la relation entre l'intelligence artificielle, le machine Learning et le Deep Learning.

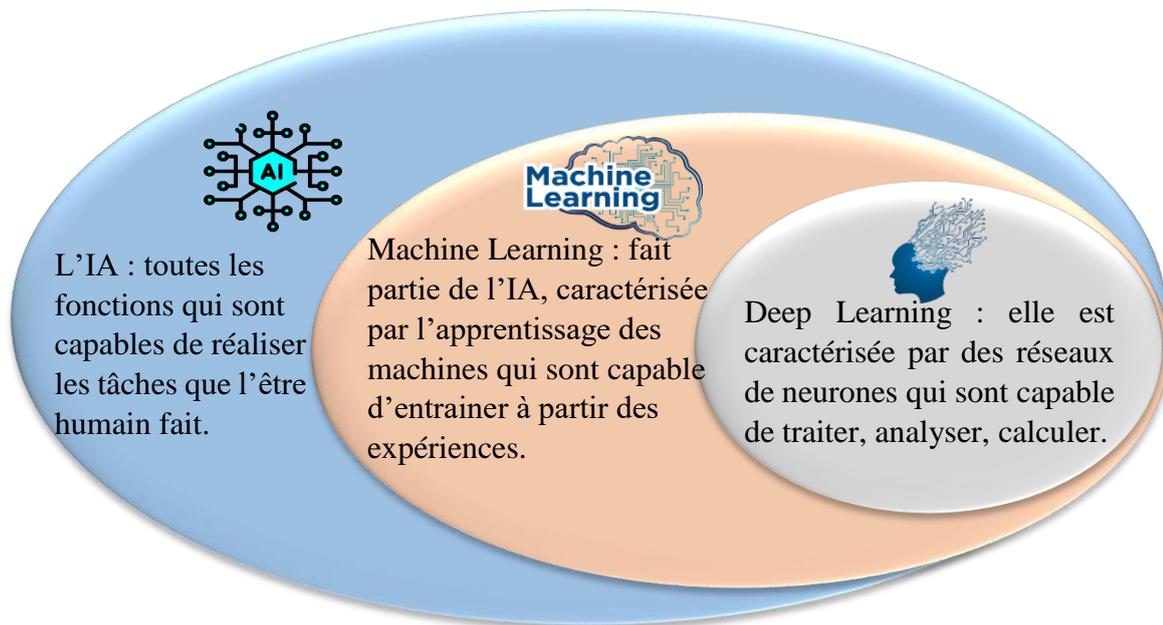


Figure II. 2 : Relation entre IA, machine-Learning et Deep Learning

Deep Learning et la machine Learning sont deux concepts utilisés dans tous les domaines et toutes les activités, dans le cas où on parle des algorithmes qui sont capables de réaliser les tâches comme le cerveau humain grâce à des réseaux de neurones pour traiter les images, le son...etc. donc on parle du Deep Learning par contre les algorithmes du machine learning sont des algorithmes qui traitent juste les données

6.1. Deep Learning :

c'est une technique de machine learning reposant sur le modèle des réseaux de neurones: des dizaines voire des centaines de couches de neurones sont empilées pour apporter une plus grande complexité à l'établissement des règles. Le Deep Learning sert que les cerveaux artificiels apprennent et s'entraînent à partir d'expériences passées comme un petit enfant.

6.1.1. Domaine d'application :

- Reconnaissance d'image (reconnaissance faciale dans Facebook ...).

- Diagnostic médical (IRM, ADN.....).
- Traduction automatique.
- Robots intelligents.

6.1.2. *Fonctionnement du Deep Learning :*

Le Deep Learning est basé sur des neurones qui sont inspirés du cerveau humain. Ces neurones sont composés de centaines de couches de neurones, chaque couche traite et analyse les données qui sont arrivées de la couche précédente et ainsi desuite et à chaque étape les mauvaises réponses sont éliminées et renvoyées vers la couche précédente, comme par exemple un système de traduction automatique, il traite les lettres en premier après il passe au traitement des mots après il fait la traduction. [31]

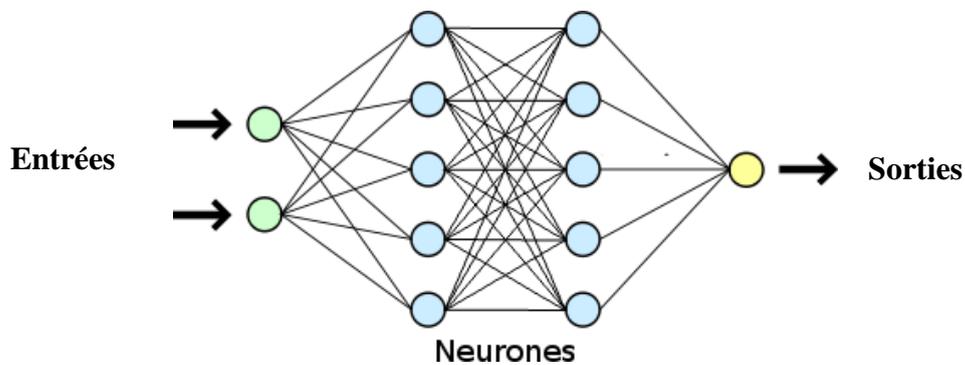


Figure II. 3 : Réseau de neurone

6.1.3. organigramme de réseau de neurone :

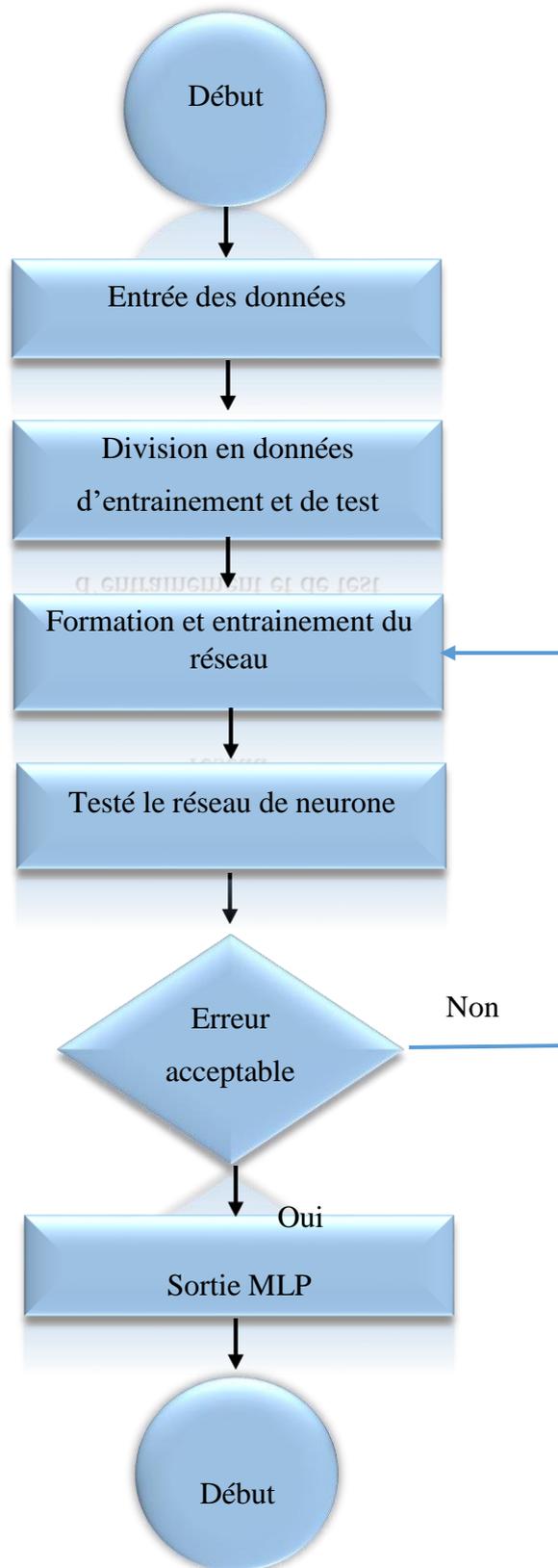


Figure II. 4 : organigramme de réseau de neurone

6.2. Machine Learning :

Le défi de l'apprentissage automatique est de permettre à un ordinateur d'apprendre à reconnaître automatiquement des modèles complexes et à prendre des décisions aussi intelligentes que possible. L'ensemble du processus d'apprentissage nécessite un ensemble de données de formation et un ensemble de données de test :

- **Ensemble de formation** : il s'agit de la base de connaissances utilisée pour former l'algorithme d'apprentissage automatique. Au cours de cette phase, les paramètres du modèle d'apprentissage automatique peuvent être ajustés en fonction des performances obtenues.
- **Ensemble de test** : utilisé uniquement pour évaluer les performances du modèle sur des données invisibles. La théorie de l'apprentissage utilise des outils mathématiques dérivés de la théorie des probabilités et de la théorie de l'information. Il y a essentiellement trois paradigmes d'apprentissage qui seront brièvement discutés dans les sections suivantes (Figure II. 1 : Diagramme d'apprentissage)

- Apprentissage supervisé
- Apprentissage non supervisé
- Apprentissage par renforcement.

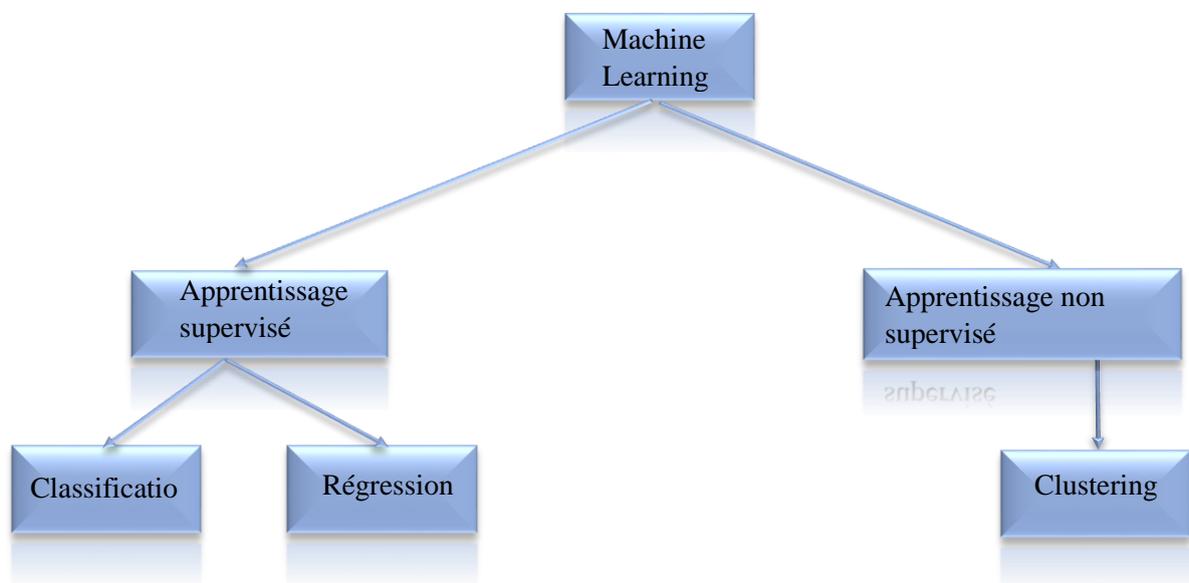


Figure II. 5 : Diagramme d'apprentissage

6.2.1. Apprentissage supervisé :

L'apprentissage supervisé est un type d'apprentissage automatique plus simple et très connu. Il est basé sur un certain nombre d'exemples pré-classifiés, dans lesquels chacun des intrants utilisés comme exemples d'apprentissage devrait appartenir à une classe (catégorie) connue a priori. Dans ce cas, la question cruciale est comment généraliser le problème de la généralisation de la prédiction pour tous les exemples. Après d'apprentissage d'un échantillon d'exemples, le système devrait produire un modèle qui devrait bien fonctionner pour toutes les entrées possibles. L'ensemble d'exemples se compose de données étiquetées, c'est-à-dire d'objets et de leurs classes associées. Cet ensemble d'exemples étiquetés constitue donc l'ensemble d'apprentissage. Le système modifie ensuite ses paramètres internes afin de minimiser cette fonction d'erreur. Par la suite, la qualité du modèle est évaluée en utilisant un deuxième ensemble d'exemples étiquetés (l'ensemble de test) afin d'évaluer le pourcentage d'exemples correctement classés et le pourcentage d'exemples mal classés.

Le contexte d'apprentissage supervisé comprend la fonction de prédiction est appelée classificateur ou *classifier* en anglais. On parle dans ce cas d'un problème de classification, qui revient à attribuer une étiquette à chaque entrée lorsque l'ensemble des valeurs de sortie est un nombre entier. Mais aussi l'apprentissage de fonctions qui prédisent des valeurs numériques lorsque la sortie que l'on cherche à estimer est une valeur continue de type réel. Cette fonction de prédiction est appelée la fonction de régression ou bien *regression function* en anglais. Il existe donc plusieurs algorithmes d'apprentissage supervisé qui ont été développés pour la classification et la régression.

6.2.1.1. Classification :

Un classificateur est un mécanisme à n entrées représentant les n entités e_1, e_2, \dots, e_n calculé à partir d'un objet à classer, et une (1) sortie. Un classificateur de classe S générera l'un des symboles w_1, w_2, \dots, w_S comme sortie, et l'utilisateur interprète cette sortie comme une décision concernant la classe de l'objet traité. Les symboles générés w_s sont les identificateurs de classes. Les algorithmes de classification tels que SVM, K-NN, Naive bayes sont des algorithmes pour classifier des données.

6.2.1.1.1. Les algorithmes d'apprentissage de classification :

En apprentissage supervisé, un algorithme reçoit un ensemble de données qui est étiqueté avec des valeurs de sorties correspondantes sur lequel il va pouvoir s'entraîner

et définir un modèle de prédiction. Cet algorithme pourra par la suite être utilisé sur de nouvelles données afin de prédire leurs valeurs de sorties correspondantes.

6.2.1.1.1. *K-Nearest Neighbor KNN:*

L'algorithme des K plus proches voisins ou K-nearest neighbors (kNN) est un algorithme de Machine Learning qui appartient à la classe des algorithmes d'apprentissage supervisé simple et facile à mettre en œuvre qui peut être utilisé pour résoudre les problèmes de classification et de régression. Dans cet article, nous allons revenir sur la définition de cet algorithme, son fonctionnement ainsi qu'une application directe en programmation.

a) Définition de k-NN :

Est un algorithme qui fonctionne à travers une base de données étiquetée avec des sorties pour qu'il puisse s'entraîner et pour qu'il puisse après utilisés des nouvelles données et donner la valeur de sortie correspondante.

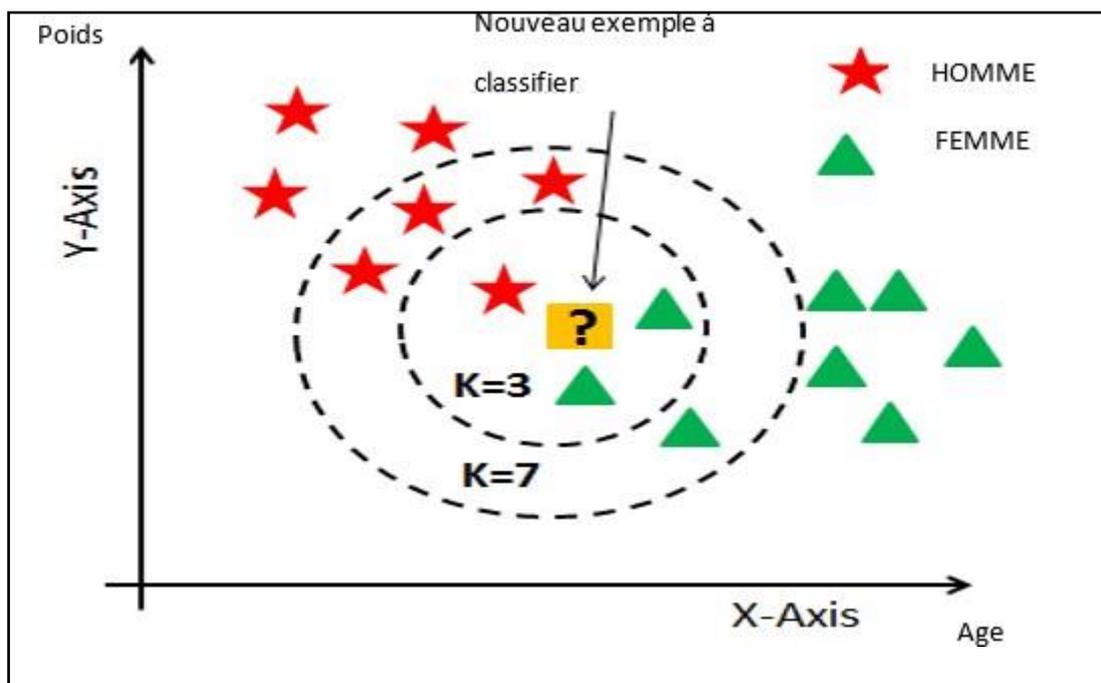


Figure II. 6 : Exemple de KNN

Prenons l'exemple de la figure II. 2, pour mieux comprendre leur fonctionnement, pour la classification entre homme et femme on a choisi deux paramètres l'âge et le poids.

Le but c'est de classifier cette entité **?**. Est-ce qu'elle est de type étoile ou bien triangle

Pour cela on a choisir par exemple $K=3$ ce qui implique qu'on a vu les trois plus proches voisins et on remarque que parmi ces trois voisins il y a deux triangles et une étoile ce qui implique que l'entité **?** est de type triangle c.-à-d. femme

b) Principe de fonctionnement de KNN :

L'intuition derrière l'**algorithme des K** plus proches voisins est l'une des plus simples de tous les algorithmes de Machine Learning supervisé :

- **Étape 1** : Sélectionnez le nombre K de voisins
- **Étape 2** : Calculez la distance

$$\sum_{i=1}^n |x_i - y_i|$$

Euclidienne

$$\sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Manhattan

Du point non classifié aux autres points.

- **Étape 3** : Prenez les K voisins les plus proches selon la distance calculée.
- **Étape 4** : Parmi ces K voisins, comptez le nombre de points appartenant à chaque catégorie.
- **Étape 5** : Attribuez le nouveau point à la catégorie la plus présente parmi ces K voisins.
- **Étape 6** : Notre modèle est prêt : [32]

c) **Organigramme du KNN :**

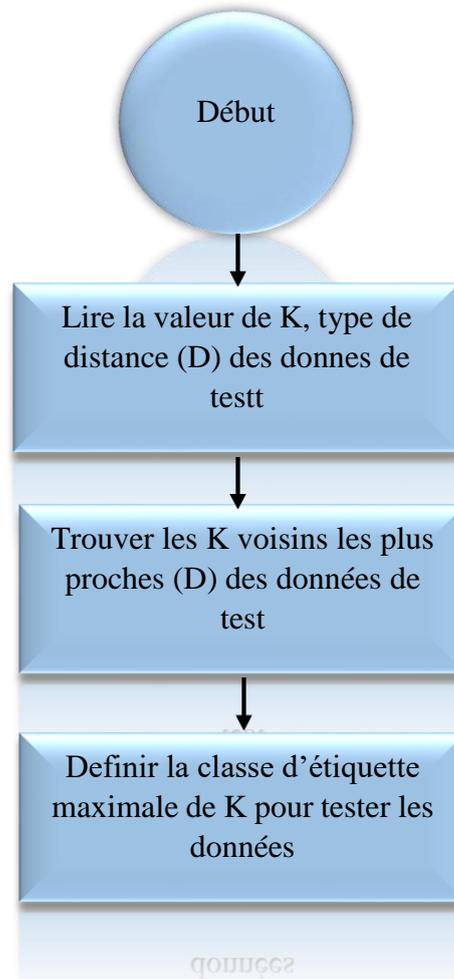


Figure II. 7 : organigramme KNN

❖ **Avantage du KNN :**

- Facile pour l'utilisation.
- Il est aussi flexible et il peut être utilisé pour la régression.
- Un seul paramètre qui est le K

❖ **Inconvénient du KNN :**

- Complexité dans le choix de la méthode de calcul de la distance.
- Complexité dans le choix de la valeur de K.

6.2.1.1.1.2. Regression logistic :

La régression logistique est un modèle statistique permettant d'étudier les relations entre un ensemble de variables qualitatives X_i et une variable qualitative Y . Il s'agit d'un modèle linéaire généralisé utilisant une fonction logistique comme fonction de lien.

Un modèle de régression logistique permet aussi de prédire la probabilité qu'un événement arrive (valeur de 1) ou non (valeur de 0) à partir de l'optimisation des coefficients de régression. Ce résultat varie toujours entre 0 et 1. Lorsque la valeur prédite est supérieure à un seuil, l'événement est susceptible de se produire, alors que lorsque cette valeur est inférieure au même seuil, il ne l'est pas.

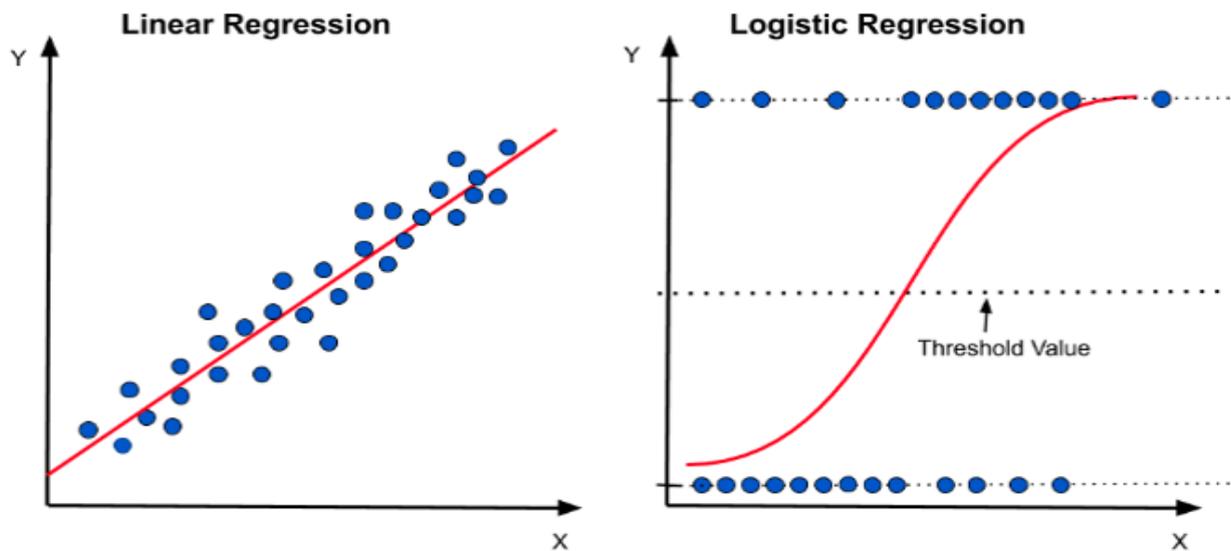


Figure II. 8: *exemple de regression logistic*

❖ *Avantage :*

- La régression logistique est plus facile à mettre en œuvre, à interpréter et très efficace à entraîner.
- Il ne fait aucune hypothèse sur les distributions des classes dans l'espace des caractéristiques.
- Il peut facilement s'étendre à plusieurs classes (régression multinomiale) et une vue probabiliste naturelle des prédictions de classe.
- Il fournit non seulement une mesure de la pertinence d'un prédicteur (taille du coefficient), mais également de son sens d'association (positif ou négatif).
- Il est très rapide pour classer les enregistrements inconnus.
- Bonne précision pour de nombreux ensembles de données simples et fonctionne bien lorsque l'ensemble de données est linéairement séparable.

❖ *Inconvénients :*

- Si le nombre d'observations est inférieur au nombre d'entités, la régression logistique ne doit pas être utilisée, sinon elle peut conduire à un sur-apprentissage.
- Il construit des frontières linéaires.
- La principale limitation de la régression logistique est l'hypothèse de linéarité entre la variable dépendante et les variables indépendantes.
- Il ne peut être utilisé que pour prédire des fonctions discrètes. Par conséquent, la variable dépendante de la régression logistique est liée à l'ensemble de nombres discrets.
- Les problèmes non linéaires ne peuvent pas être résolus avec la régression logistique car elle a une surface de décision linéaire. Les données linéairement séparables sont rarement trouvées dans les scénarios du monde réel.
- Il est difficile d'obtenir des relations complexes en utilisant la régression logistique. Des algorithmes plus puissants et compacts tels que les réseaux de neurones peuvent facilement surpasser cet algorithme.

6.2.1.1.1.3. *Arbre de décision :*

Est l'une des algorithmes du modèle d'apprentissage supervisé qui est utilisé dans plusieurs domaines tels que la planification ou bien la recherche opérationnelle, les nœuds sont les carré, notre arbre sera plus précis quand le nombre des nœuds est élevé, aussi les feuilles de l'arbre sont les derniers nœuds dans le niveau ou la décision est prise.

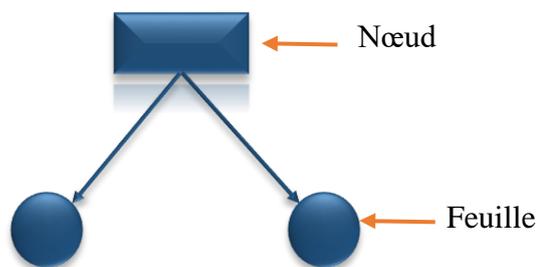


Figure II. 9 : arbre de décision

Les arbres de décision de classification ou bien classification tree qui sont utilisés pour déterminer quelle classe la variable de sortie appartient (voir figure II. 10)

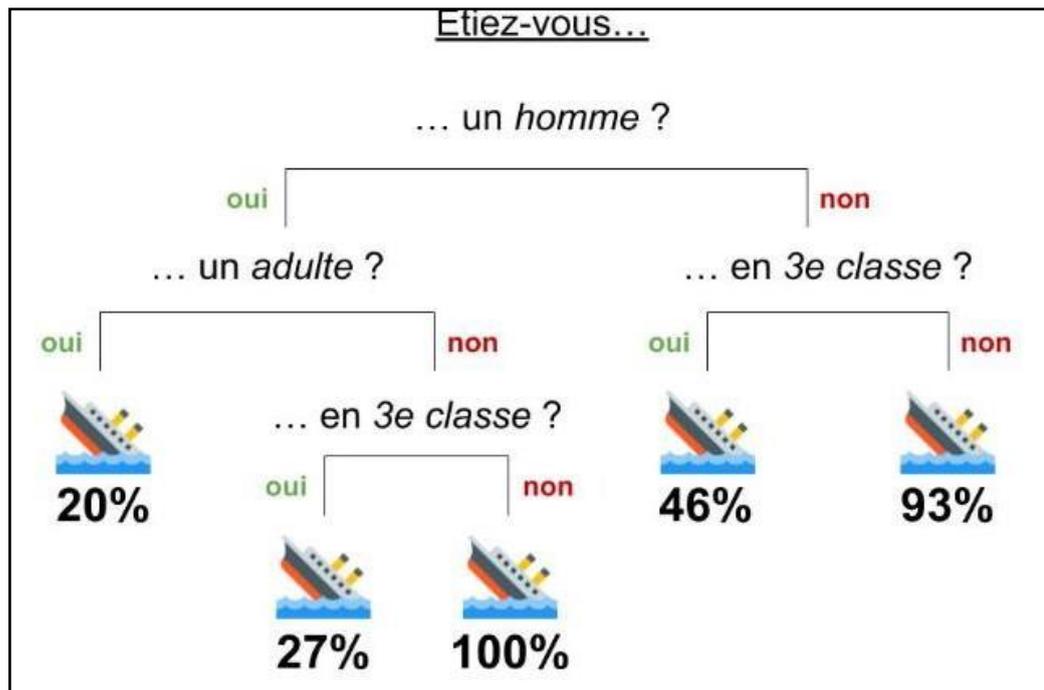


Figure II. 10 : Arbre de décision [33]

❖ *Avantage de l'arbre de décision :*

- Facile pour l'appliquer.
- Facile pour comprendre les résultats.
- Temps d'exécution petit par rapport aux autres algorithmes.

❖ *Inconvénients de l'arbre de décision :*

- Moins performant par rapport aux autres algorithmes.
- Le sur-apprentissage c'est-à-dire les données d'entraînement ne donnent pas un résultat fiable pour le test ou bien pour une nouvelle entité. [34]

6.2.1.1.1.4. *Forêt aléatoire :*

C'est un ensemble d'arbres de décision, chaque arbre du foret produit une prévision de classe et la bonne prévision est celle qui a reçu plus de vote.

Pour simplifier sa on dit que la forêt aléatoire combine plusieurs arbre de décision pour obtenir une bonne prévision.

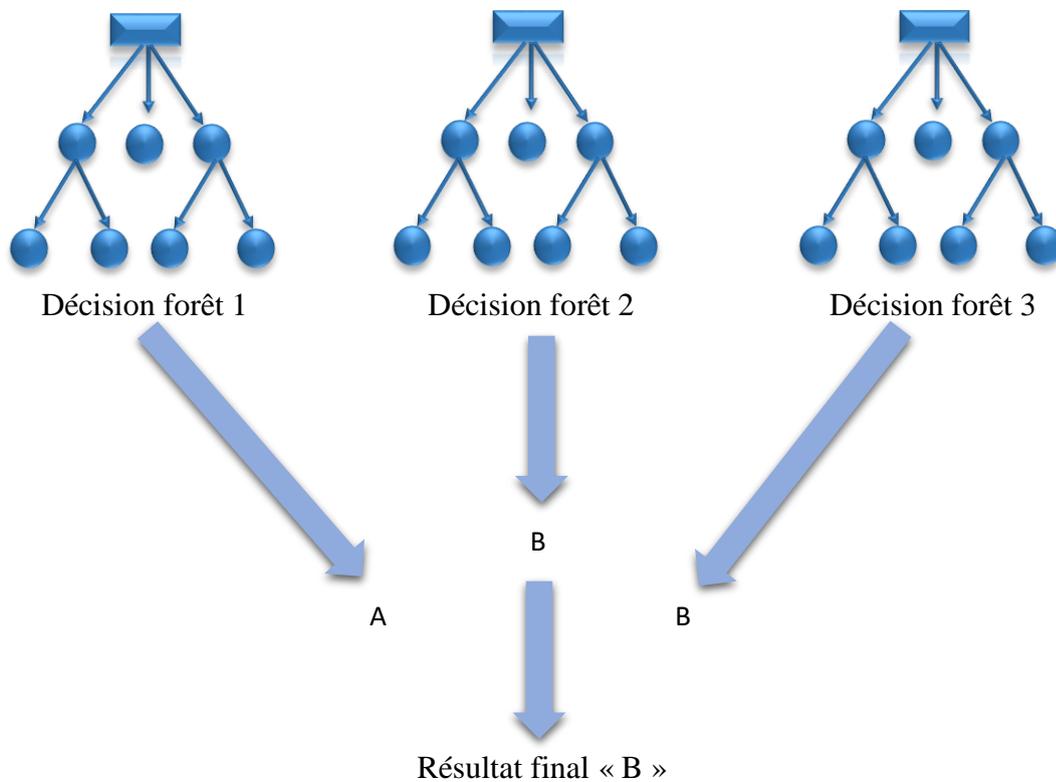


Figure II. 11: Forêt aléatoire

Si on suppose que le premier arbre a une prédiction de classe A et le deuxième classe B et le troisième aussi classe B donc dans ce cas-là la prédiction sera classe B. C'est la puissance des forêts aléatoires.

Les forêts aléatoires forment la plus grande partie des systèmes d'apprentissage automatique, qu'elle peut être utilisée dans les problèmes de classification et de régression aussi. [35]

❖ *Avantage de l'algorithme des arbres aléatoire :*

- Facile à comprendre le résultat si l'arbre est petit.
- Robuste des données non utilisables.
- Résultat un est un modèle non linéaire.

❖ *Inconvénient de l'algorithme des arbres aléatoire :*

- Le taux d'apprentissage est élevé.
- Le résultat obtenu n'est pas sûr qu'il soit la meilleure.

- Temps élevé pour l'exécution. [36]

6.2.1.1.5. SVM :

C'est l'abréviation de machine à vecteurs de support, il vise à résoudre les problèmes de classification et non pas de régression, aussi il vise à séparer les données en classe par une ligne qui s'appelle une frontière de telle sorte que la distance qui s'appelle dans ce cas une marge soit maximale.

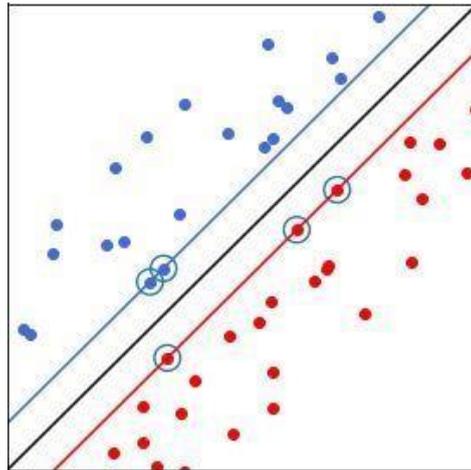


Figure II.9 : *Exemple du SVM*

Dans cet exemple on remarque que la droite noir est la frontière, les vecteurs de support sont les points bleue et rouge (les plus proches de la frontière) et la marge est la distance entre la frontière et les droites rouge et bleue, ce cas-là indique que les données sont linéairement séparable. [37]

a) Organigramme SVM :

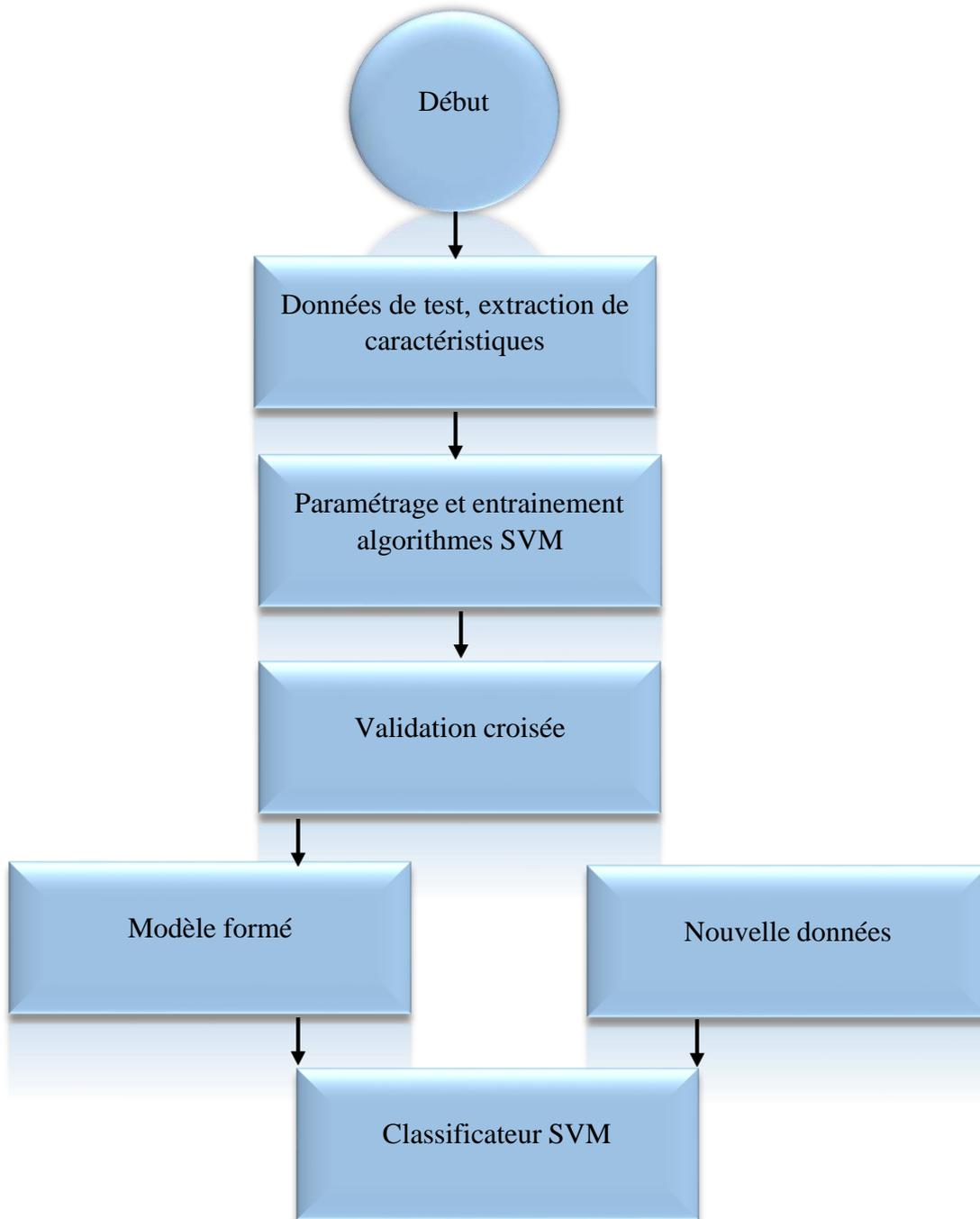


Figure II. 12 : organigramme SVM

❖ *Avantage du SVM :*

- Précision de prédiction.
- plus efficace par rapport aux autres algorithmes à cause de son sous-ensemble de points d'entraînement.

❖ *Inconvénient du SVM :*

- Il fonctionne juste avec les petites bases de données.
- Résultat moins efficace avec les valeurs extrêmes. [38]

6.2.1.1.2. Tableau récapitulatif des différents avantages et inconvénients des algorithmes de classification :

Algorithme	Avantages	Inconvénients
KNN	<ul style="list-style-type: none"> • Simple à comprendre, rapide et efficace. 	<ul style="list-style-type: none"> • Besoin de choisir manuellement le nombre de voisins «k».
Régression logistic	<ul style="list-style-type: none"> • Approche probabiliste, donne des informations sur la signification statistique des caractéristiques. 	<ul style="list-style-type: none"> • Les hypothèses de régression logistique.
Arbre de décision	<ul style="list-style-type: none"> • Fonctionne sur les problèmes linéaires / non linéaires. 	<ul style="list-style-type: none"> • Mauvais résultats sur de très petits ensembles de données, un surajustement peut facilement se produire.
Forêts aléatoire	<ul style="list-style-type: none"> • Puissant et précis, bonnes performances sur de nombreux problèmes, y compris non linéaires. 	<ul style="list-style-type: none"> • Il faut choisir le nombre d'arbres manuellement.
SVM	<ul style="list-style-type: none"> • Performant, non biaisé par les valeurs aberrantes, insensible au surajustement. 	<ul style="list-style-type: none"> • Ne convient pas aux problèmes non linéaires, pas le meilleur choix pour un grand nombre de fonctionnalités.

Tableau II.1 : *récapitulatif des différents avantages et inconvénients des algorithmes de classification [39]*

6.2.1.2. La régression :

Dans l'apprentissage automatique, les algorithmes de régression tentent d'estimer la fonction (f) des variables d'entrée (x) aux variables de sortie numériques ou continues (y). Par exemple, lorsque vous disposez d'un ensemble de données sur les maisons et que vous êtes invité à prédire leurs prix, il s'agit d'une tâche de régression car le prix sera une sortie (valeur) continue. Les exemples d'algorithmes de régression courants incluent *Linear Regression*, *Decision Tree*, *Random Forest* et *Support Vector Regression*.

6.2.2. Apprentissage non supervisé :

Parlant maintenant sur le deuxième type d'apprentissage « non supervisé » qui est une branche du machine Learning, il est destiné pour regrouper, classer et ordonner les séries de données non marquées, ce modèle n'a pas besoin de recommandation humaine.

Comme l'apprentissage supervisé ce modèle aussi est divisé en deux catégories :

6.2.2.1. Le clustering :

Où partitionnement de donnée est une méthode d'apprentissage non supervisé, vise à collecter des données, en fonction de leurs similarités où de leurs différences, il vise de regrouper les données de telle sorte que les données dans chaque groupe soient similaires le maximum et les données de chaque groupe soit différents le maximum par rapport aux données d'un autre groupe, il y a plusieurs algorithmes de clustering comme par exemple k-moyennes, classification hiérarchique et classification probabilistique...etc. [40]

6.2.2.1.1. K-means (k-moyennes) :

Il vise à résoudre le problème de clustering, il est simple et facile pour l'appliquer, il classe les données qui sont arrivées d'un certain nombre de cluster (groupe), et les données dans chaque groupe sont homogènes.

6.2.3. Apprentissage par renforcement :

C'est une méthode d'apprentissage dans le domaine de machine Learning qui permet aux machines ou bien aux ordinateurs d'apprendre par elles-mêmes à partir des erreurs donc à partir des pénalités et des récompenses.

Dans ce cas-là, l'agent de l'IA pense comment il peut accomplir une tâche donc il fait plusieurs tentatives aléatoires en parallèle pour trouver la meilleure solution.

Prenons par exemple les animaux si nous voulons leur apprendre quelque chose s'ils l'apprennent correctement, nous le leur donnons une récompense sinon nous le leur donnons une petite pénalité. [41]

6.3. Machine Learning VS Deep Learning :

	Machine Learning	Deep Learning
Domaines d'applications principaux	Classification et Régression : données tabulaires	Classification et Régression : données tabulaires, traitement de langage naturel (NLP) et Computer Vision
Volume de données d'entraînement observé dans la pratique	Des milliers d'observations	Big Data : des millions d'observations
Entraînement	"One-shot"	"Continu" par itérations (Batches & Epochs)
Matériel recommandé	CPU	GPU
Bibliothèque Python	Scikit-learn, Scipy, Panda	Keras, Tensorflow, Pytorch

Tableau II. 2 : machine Learning VS Deep Learning [42]

7. Quel est le rapport entre l'intelligence artificielle et l'industrie 4.0 ?

A partir de 1765 l'industrie a passé par des étapes de révolution qui sont expliquées par la suite :

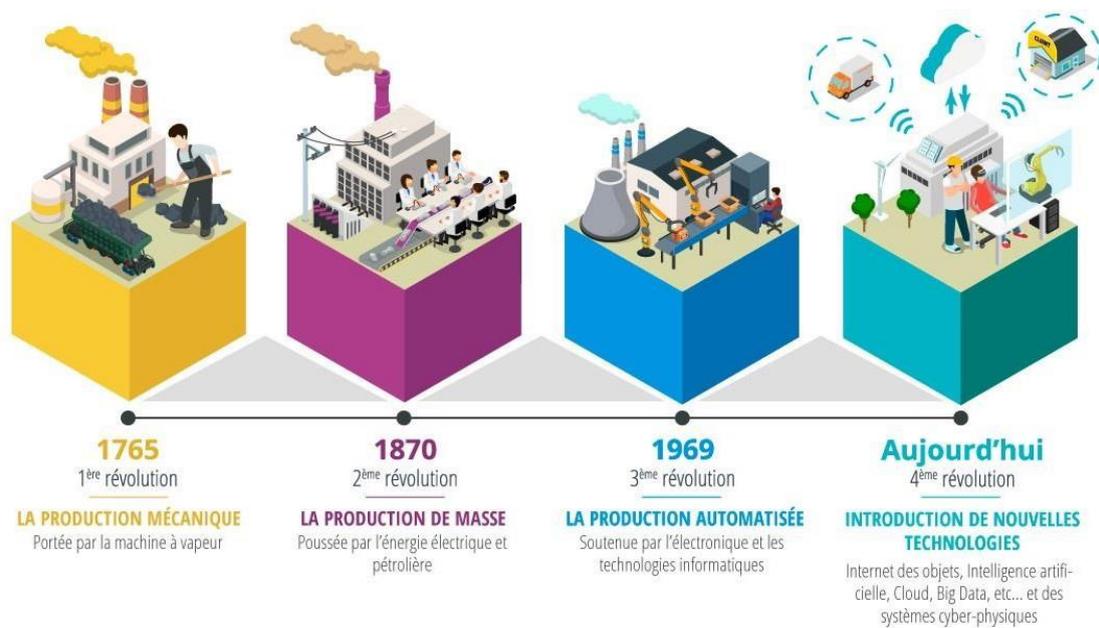


Figure II.11 : Les différentes phases de l'industrie [43]

7.1. Industrie 1.0 : La mécanisation :

Chaque révolution industrielle est une résultante des progrès scientifiques et technologiques de son temps, suite à la mise en place du premier système mécanisé « la machine à vapeur » la 1^{ère} génération de révolution industrielle est née en 1784. Dès la fin du 18^e siècle, le travail manuel a été remplacé par la production mécanisée et la capacité de production a été augmentée.

7.2. Industrie 2.0 : Production en série (massification) :

La deuxième révolution industrielle a débuté les vers 1870. Elle est basée sur la mise en œuvre de nouvelles sources d'énergie telles que le pétrole, le gaz et l'électricité. D'un autre côté, la création d'une chaîne de production et des concurrents faite avec de nouveaux moyens de transport et de communications telles que l'avion, les voitures ou le téléphone. Au cours du 19^e siècle, la production est devenue standardisée.

7.3. Industrie 3.0 : Numérisation :

La troisième révolution industrielle débute autour de 1969. C'est l'utilisation d'ordinateurs, de microprocesseurs et d'Internet, qui permettaient à plusieurs entreprises de contrôler les processus de production à l'aide des technologies tels que les API. [43]

7.4. Industrie 4.0 : Aujourd'hui :

Nous vivons aujourd'hui la 4^{ème} révolution industrielle et c'est la révolution des systèmes 3D, de l'Internet des objets (IoT) (signifier connexion entre les objets comme les capteurs), des

réseaux, du Big Data, de l'intelligence artificielle. Toutes ces technologies et d'autres encore permettant une prise de décision optimale en une fraction de seconde sous la gestion d'un ordinateur central.

Cette industrie a un but de rendre les entreprises intelligentes et à la transformer en une entité capable d'obtenir les résultats les plus optimaux, et aussi pour améliorer la productivité, la qualité, les performances, la rentabilité et l'efficacité des entreprises. [44]

7.5. *L'avenir de l'industrie 4.0 est l'industrie 5.0 :*

Avec l'utilisation des technologies comme l'intelligence artificielle et le Big Data l'automatisation des processus sera optimisée. Pour connaître les exigences des clients l'analyse de données et la solution, et aussi la fabrication à la demande est plus flexible, rapide, efficace et personnalisée.

Nous voyons que jour après jour les clients deviennent très exigeants mais les entreprises souffrent de manque de main d'œuvre.

Aujourd'hui nous vivons la quatrième révolution industrielle et les usines intelligentes ou bien « smart factories » sont devenues de plus en plus flexibles, plus productives, ils minimisent le temps d'arrêt non planifiés, accroître l'efficacité et la qualité en général des produits.

L'intelligence artificielle est la base de l'industrie 4.0 selon les statistiques de « Fédération internationale de la robotique » en 2015 le nombre de robots industriels qui fonctionnent dans les usines était 1.6 millions, en 2019, il atteint 2.6 millions et en 2021 atteint 3 millions de robots.

8. Application de la machine Learning dans la fabrication :

On va expliquer ce point par un exemple :

General Electric est une grande entreprise possède plus de 500 usines dans le monde et elle débute à les rendre intelligentes, elle fabrique les appareils électroménagers et les machines de production industrielle à travers des turbo-réacteurs.

La solution offerte par GE pour gérer et surveiller tous les aspects de fabrication c'est « Brilliant Manufacturing Suite » pour trouver tous les défauts et les défaillances possibles sur le processus de fabrication, avec cette solution leur usine Brilliant en Inde a amélioré le rendement et l'efficacité de leur installation de 18%.

Cette solution vise à rendre tous les éléments de fabrication en un seul dispositif intelligent tels que la conception, l'ingénierie ou la distribution, elle a même sa propre plate-forme Lot industrielle Predix, pour contrôler tous les éléments de fabrication cette plate-forme utilise des capteurs.

Predix possède de capacité d'apprentissage automatique pour traiter toutes les informations et fournir des domaines d'amélioration.

Par ailleurs GE a fournis plus d'un milliard de dollars pour rendre Predix sera capable de traiter plus d'un million téraoctets d'informations par jour.

9. Conclusion :

De l'intelligence artificielle à l'apprentissage automatique et puis à l'apprentissage profond, ce sont trois concepts liés à l'intelligence artificielle. Les trois concepts se combinent pour améliorer l'avenir de l'intelligence artificielle, mais ce n'est pas de l'intelligence artificielle. La nouvelle étape dans les processus de développement de systèmes d'intelligence artificielle se rapproche du défi de la réalisation de systèmes informatiques indépendants de l'intervention humaine. Autrement dit, ces systèmes sont capables d'imiter le comportement humain et la pensée.

Dans le chapitre suivant nous parlerons de l'entreprise dans laquelle nous avons accompli notre stage de fin d'étude.



Chapitre III : Présentation de l'entreprise



Introduction :

L'un des domaines de l'industrie qui se développe jour après jour est la transformation du papier, et techniquement à partir de fibres cellulosique végétale et animal le papier est fabriqué. Le papier a été découvert dans la Chine et depuis ce jour le processus de fabrication est le même, mais les types de papiers sont développés, on peut dire qu'il existe maintenant une infinité des types de papier (couché matte, couché brillant, thermique...) et aussi divers grammage et divers formats.

1. Présentation de l'entreprise :

Mega Papiers c'est une entreprise située à la zone industrielle de chetouane, Desserte 5, Tlemcen crée en 2003 qui cherche la satisfaction de ses clients (nationaux et internationaux) par la transformation de tout genre de papiers. Ce dernier est considéré comme l'un de leader du marché algérien.



Elle est composée de 70 employés et elle à peu près 350 clients.

Elle assure la bonne qualité de ses produits et la livraison dans les bons délais et aussi assure plusieurs formats, grammage et propose la possibilité de personnaliser les papiers avec impression.



Figure III. 2 : MegaPapiers

2. Mission de l'entreprise :

Cette entreprise a un seul objectif qui est la transformation de tout type de papier pour satisfaire les besoins et les exigences de ces clients avec des produits de bonne qualité.

3. Type des produits :

Megapapiers propose plusieurs gammes de produits qui diffèrent entre eux selon le type de papier, son grammage et son format .Nous en citons quelques-uns :

- Papier autocopiant
 - Papier couché
 - Cartes et cartons
 - Papier dossier & bristol
 - Papier adhésif
 - offset
 - Papier thermique
 - ...etc.
- Elle est divisée en 4 filiales : MegaPrint, MegaLabel, MegaPlotter et Megatherm.

4. Mega-Print :

« *Le papier pour imprimeurs* », Ce filial est destiné pour tout genre de papier pour les imprimeurs, Offset, pastoral, Couché, Autocopiant, Bristol, Carte et Carton

Donc la transmission de papier destiné aux imprimeries est faite dans cette filial, et selon les demandes des clients, elle offre divers types de grammages et formats de papiers. .



Figure III. 3 : *logo du Mega Print* [46]

4.1. Le processus :

- La découpe en format : c'est la transformation des bobines mère en palette cube, palette en vrac ou bien en palette papiers en rames.
- La refonte : dans cette étape, les bobineaux sont transformés ou bien sont créés à partir des bobines mère. La création d'une bobine fille, se fait par réduction du diamètre soit par réduction de la laize.
- Les standards de la découpe (refonte) de la machine sont :
 - Diamètre ≥ 1300 mm et la laize $< 1\ 650$ mm. [47]

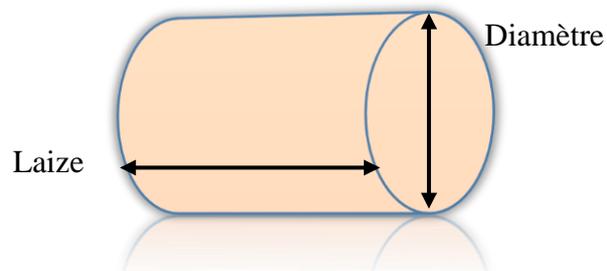


Figure III. 4 : *exemple d'une bobine*

4.2. Papier offset :

L'impression des catalogues, cahiers, livres ou bien les dictionnaires est faite par ce type de papier.

Format en (cm)	65*100	65*92	70*100				
Grammage en (gms/m ²)	60°	70°	80°	100°			
Bobines : laizes (cm)	26	40	48	50	65	93	100

Tableau III. 1 : caractéristique papier offset

4.3. Papier pastoral :

C'est un autre format de papier offset qui est caractérisé par plusieurs couleurs, son grammage est de 60° et son format est 65*100.

4.4. Papier autocopiant :

C'est un type de papier pour la transmission de l'écriture manuelle sur des multiplicatifs. Elle est caractérisée par la lisibilité des copies jusqu'à 20ans et elle existe en blanc et en couleurs. Il existe 3 types : le (CB) coated back 56°, le (CF) coated face 57° et le (CFB) coated face and back 53°.

Les deux formats standard de ce type sont 65*100 et 64*90 et le bobineau est de laize : 260, 380, 400, 480 et 500.

4.5. Papier couché :

C'est le papier dédié pour l'impression des brochures, cartes visite, dépliants...etc.

Il existe deux types de ce papier, soit le mat soit le brillant de deux faces, il est de grammages : 90, 100, 115, 130, 150, 170, 200, 250, 300 et 350, il existe un seul format c'est 70*100. Il est utilisé pour l'impression des cartes visites, flayers, brochures....etc.

4.6. Papier bristol :

Il existe deux types, soit blanc simple soit en couleur, il est de grammages : 180°, 250° et les formats sont 65*100 ou bien 32*48.

4.7. Papier adhésif :

Il est de type couché brillant avec un grammage de 80° et format de 70*100.

4.8. Papier sécurisé :

C'est le papier dédié pour les chèques, il est de grammage 95° avec un format de 45* 67.

4.9. Les cartons :

Il existe deux types :

- Soit le carton blanc gris dédié pour tous les types d'emballage, il existe en plusieurs grammages : 250°, 275°, 280°, 300°, 325°, 400° et 450° avec deux formats : 70*100 et 80*120*.
- Carton blanc bois tous bois : tous les emballages pharmaceutique ou bien cosmétique sont faits par ce type de carton. Il existe aussi comme le carton précédent en plusieurs grammages : 230°, 255°, 270°, 280°, 300°, 325° et 350° avec un seul format 70*100



Figure III. 5 : le carton

5. Mega Label :

« *L'étiquette qui colle à vos besoins* », cette filiale a pour mission d'offrir une grande variété d'étiquettes avec plusieurs formats (rouleaux). [48]



Figure III. 6: logo Mega Label [46]



Figure III. 7 : exemple d'étiquette

Il existe en plusieurs formats : 20*40, 35*45, 43*58 et 100*150.

5.1 Domaines d'utilisation des étiquettes :

- **L'agro-alimentaire** : pour l'identification, traçabilité, date de fabrication et date de préemptionetc.
- **Logistique** : décrire la nature du produit est ce qu'il est danger ou bien fragileetc.
- **Pharmacie** : les vignettes par exemple, nature du médicament...etc.
- **Chimie** : décrire aussi la nature du produit chimique est ce qu'il est dangereux, la traçabilité ...etc. [46]

6. Mega-Plotter :

« *Papier pour table traçante* Cette filiale offre une gamme avancée, divers types de papiers, formats et laizes sont disponibles pour répondre aux besoins des clients. Donc c'est la solution d'impression professionnelle pour tous les appareils de traçage et de dessin. [49]



Figure III. 8 : *logo Mega Plotter* [46]

Dans cette filiale, l'entreprise fabrique des rouleaux pour les tables traçantes et les tireuses de plan appelé communément le rouleau plotter. Le papier utilisé dans cette filiale est de type offset extra-blanc avec un grammage de 80° et 90° et un format de 914*50 ou bien 914*150. [46]



Figure III. 9 : *rouleau plotter* [46]



Figure III. 10 : *table traçante* [46]

7. Megatherm :

«*Your thermal paper* ». C'est la filiale où nous avons effectué notre stage pratique dans le cadre de projet de fin d'étude.

Cette filiale est dédiée à la production des rouleaux de papier (papier thermique), pour la satisfaction de toutes les utilités désirées comme tickets de paiement électronique, tickets de parking, tickets de transport et le fax...etc.



Figure III. 11 : *logo Megatherm* [46]

Megatherm fournit un large éventail de laize et de grammage.

Cette filiale offre aussi une impression personnalisée avec 3 couleurs. [50]

Il existe dans Megatherm plusieurs grammages tels que : 48°, 55°, 80° et 105° avec des bobineaux de 80*80*12, 80*60*12 ou bien 57*40*12

Cette filiale travaille avec trois domaines principaux :

7.1. Les banques :

Megatherm offre des rouleaux de papier pour par exemple impression dans les distributeurs des relevés bancaires (DAB).

7.2. Les tickets :

Ce domaine est le plus dominant par ce qu'il existe partout dans les supermarchés, les restaurants ou bien en dit tous impression des tickets des caisses (POS).



Figure III. 12 : exemple d'un ticket de supermarché

7.3. Le transport :

Pour toute impression, des tickets de transport soient avion, bus, train ou n'importe quelle impression des tickets de transport.

7.4. Autres domaines :

Comme l'impression des tickets de cinéma ou bien le fax (parmi la gamme la plus demandée) ou bien les stades...etc.



Figure III. 13 : exemple d'impression du fax

7.5. Le papier thermique :

Le papier thermique est un type de papier, avec l'application d'une grande chaleur ce papier réagit à cette chaleur et à cause de cela, l'écriture est imprimée sur lui et c'est pour cela, c'est un papier d'impression sans encre, il consomme dans son impression une faible énergie donc il est considéré comme la technologie d'impression la plus économique. Il est caractérisé par sa durabilité, impression de haute niveau, durée de lisibilité grande (presque 5 ans) et un temps de stockage aussi grand (presque 10 ans). Aussi, l'impression sur ce papier est plus rapide par rapport aux autres types d'impression ce qui implique un grand rendement, pour ses caractéristiques ce type de papier est utilisé dans plusieurs domaines d'impression comme l'impression des tickets des caisses, l'impression des tickets de transports, les tickets de cinéma, les vignettes sur les médicaments....etc. [51]

7.6. Fonctionnement du processus :

C'est un processus dédié pour fabriquer des produits de haute qualité qui satisfait les besoins de ces clients.



Figure III. 14 : *atelier de MegaThetm*

Megatherm a un bon matériel de découpe et d'emballage adaptés et adéquat avec les bobines et les palettes de papier thermique.

Ce matériel peut fabriquer des formats standards ou bien spécifiques suite à la demande du client.

Une bobine est constituée d'une feuille de papier de plusieurs kilomètres de long enroulée Autour d'un mandrin. Le mandrin, réalisé en carton épais, sert de support d'enroulement aux bobines filles.

Les bobines produites en sortie des machines à papier sont appelées bobines mères. Ce papier est impossible à livrer sous cette forme aux différents clients (trop volumineux et trop lourds), il est donc nécessaire de le recouper et de le reconditionner.

Ce conditionnement permet l'obtention de bobines de diamètre et de longueur plus faible qui peuvent être manutentionnées et utilisées par les clients plus facilement. Ces bobines plus petites sont appelées bobines filles.

Les bobines filles peuvent avoir des dimensions variables en fonction des besoins du client. La machine permettant ce conditionnement est appelée bobineuse. La bobineuse permet la

découpe (en largeur et en longueur de papier) des bobines mères pour s'adapter aux dimensions souhaitées par les clients.

On étudiera ici une chaîne de transformation dotée de machine bobineuse (référence) qui permet l'impression par flexographie à l'intérieur du papier équipant l'atelier de Megatherm produisant du papier thermique.

7.7. Clients MegaPapiers :

Megapapiers a deux catégories des clients : nationaux et internationaux.



Figure III. 15 : clients internationaux de Megapapiers

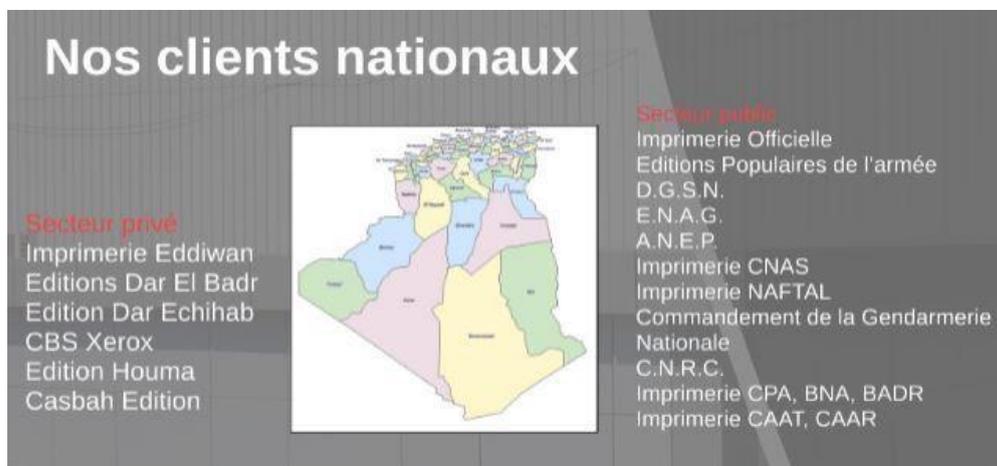


Figure III. 16 : clients nationaux de Megapapiers

8. Conclusion :

La transformation du papier est l'une des activités les plus courantes qui se développe jour après jour.

Megapapiers est une entreprise spécialisée dans la transformation de tout type de papier qui a plusieurs clients internationaux, elle est l'une des entreprises les plus connues dans le marché

national algérienne, elle a un processus de fabrication qui est adéquat par rapport aux exigences des clients.

Le responsable de la chaîne avait un problème de choix entre deux types de machine identique pour passer une nouvelle commande en diminuant le temps de changement. Pour cela, nous avons opté à développer un modèle d'apprentissage automatique basé sur des données antérieures qui sera bien détaillé dans le chapitre suivant.



**Chapitre IV : Application de
l'apprentissage automatique pour un
problème de classification des machines de
l'entreprise « ROBO 2400 et ROBO 2401 »**



Introduction :

L'évolution de l'intelligence artificielle dans le temps impose des défis aux entreprises. Dans la quatrième révolution industrielle (l'industrie 4.0), les nouvelles technologies se répondent rapidement et efficacement aux besoins des clients et des entreprises, ces derniers donc réagissent différemment avec la digitalisation et l'optimisation de production pour améliorer la productivité.

L'intelligence artificielle et l'apprentissage automatique (machine Learning) fournissent plusieurs offres et avantages aux entreprises comme des solutions optimales, moins coûteuses et efficaces pour prendre des décisions raisonnables, claires et évidentes fondées par un ensemble de données. Maintenant, l'intelligence artificielle ne traite pas juste les grandes entreprises, mais aussi les moyens et les petites entreprises. Les principales clés de cette progression sont premièrement sans doute la révolution de la technologie, le stockage des anciennes données dans les entreprises et bien sur la présence de plusieurs algorithmes, nécessaire et utile, dans le domaine d'apprentissage.

Donc comment ce phénomène se place-t-il dans les entreprises quand la concurrence exerce une forte pression sur tous les responsables économiques au même temps que les entreprises sont face à la pression du temps et aussi face aux exigences des consommateurs et au besoin de répondre rapidement aux attentes des clients ? [52]

Dans ce chapitre, nous traitons un sujet de projet de fin d'étude qui repose sur une problématique dans cette entreprise est centrée sur deux machines identiques (**ROBO 2400** et **ROBO 2401**), qui travaillent en parallèles et qui traitent les mêmes types des produits. Mais le problème qui se pose quand une commande urgente arrive à l'entreprise, cette commande est réalisable avec l'un des deux machines ou bien avec les deux à la fois ce qui implique que la décision en ce moment-là avec la contrainte du temps de changement et avec les exigences du client est vraiment difficile pour la prendre.

1. Méthodologie :

Pour mener à bien ce projet, nous avons passé par des étapes qui sont indispensable dans le domaine de machine Learning. Ces étapes sont :

- Acquisition des données.
- le choix de l'environnement de travail.
- Préparation de la base de données (data).

- Division de data en deux parties : partie d'entraînement et partie de test.
- Choix de l'algorithme d'apprentissage.
- Évaluation de ces algorithmes.
- Et enfin le test et déploiement.

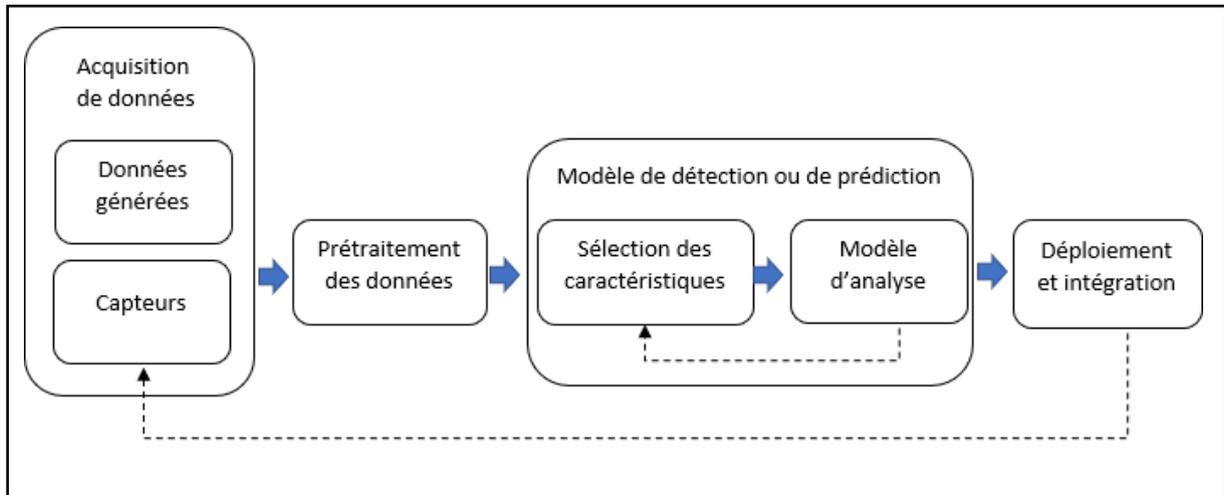


Figure IV 1 : Analyse de données et classification

2. Acquisition des données :

L'entreprise nous a donnée des données mensuelles et en détails (les quantités fabriquées par chaque machine et par jour, le nombre de cycle, les informations des bobines et de matière première... etc.) sous format Excel.

RAPPORT DE PRODUCTION													
DATE	MANDRINS			matieres premieres						produits finis			
	mandrins	quantité		marque	ref bobine	laize	grammage	metrage	poids	format	nbr cycle	quantité	poids
03/01/2021	57x13	9328		koehler		1260	55	15000	suite	57x19m	133	2926	174,29
				koehler		1260	55	15000	1039,5	57x19m	291	6402	381,34
04/01/2021	57x13	0		koehler		1260	55	15000	suite	57x19m	0	0	0,00
05/01/2021	70x13	2988		koehler		1260	55	15000	suite	210x25m	166	996	287,60
06/01/2021	70x13	3726		koehler		1260	55	15000	suite	210x25m	207	1242	358,63
07/01/2021	70x13	2682		koehler		1260	55	15000	1039,5	210x25m	149	894	258,14
10/01/2021	70x13	5652		koehler		1260	55	15000	suite	210x25m	314	1884	544,01
11/01/2021	70x13	4950		koehler		1260	55	15000	suite	210x25m	132	792	228,69
				koehler		1260	55	15000	1039,5	210x25m	143	858	247,75
13/01/2021	70x13	5310		koehler		1260	55	15000	suite	210x25m	295	1770	511,09
14/01/2021	70x13	4842		koehler		1260	55	15000	suite	210x25m	159	954	275,47
				koehler		1260	55	15000	1039,5	210x25m	110	660	190,58
17/01/2021	70x13	5598		koehler		1260	55	15000	suite	210x25m	311	1866	538,81
18/01/2021	70x13	4284		koehler		1260	55	15000	suite	210x25m	169	1014	292,79
				koehler		1260	55	15000	1039,5	210x25m	69	414	119,54
19/01/2021	70x13	6138		koehler		1260	55	15000	suite	210x25m	341	2046	590,78

Figure IV 2 : base de données

Après plusieurs réunion avec le responsable de la production nous avons pu extraire les données qui ont une relation avec le choix entre des deux machines. Ce qu'on appelle les données importantes "relevant feature" dans le vocabulaire de machine learning.

Le fichier suivant nous montre les données importantes après sélection.

format du produit	Durée de fabrication en h	quantité à fabriquer	instant t1 de la machine ROBO 2400	instant t2 de la machine ROBO 2401	machine
5719,00	24,00	20262,00	21025,00	7575,00	1,00
21025,00	188,00	34212,00	5719,00	5719,00	1,00
5719,00	71,19	77901,00	21025,00	7575,00	2,00
7974,00	40,35	29682,00	5719,00	5719,00	2,00
7919,00	21,92	24174,00	21025,00	7974,00	2,00
7915,00	31,67	40171,00	21025,00	7919,00	2,00
7919,00	25,27	26027,00	21025,00	7915,00	2,00
5710,00	96,00	94714,00	21025,00	7919,00	1,00
7969,00	34,41	23409,00	5710,00	7919,00	2,00
7960,00	38,38	28815,00	5710,00	7969,00	2,00
7940,00	8,00	5797,00	5710,00	7960,00	2,00
11024,00	8,00	2400,00	5710,00	7940,00	2,00
7529,00	16,00	7152,00	5710,00	11024,00	2,00
7978,00	24,00	10166,00	5710,00	7525,00	2,00
7974,00	11,23	6188,00	5710,00	7978,00	2,00
7940,00	36,77	29495,00	5710,00	7974,00	2,00
21025,00	624,00	132525,00	5710,00	7940,00	1,00
7978,00	16,28	8058,00	21025,00	7940,00	2,00
7974,00	15,72	10030,00	21025,00	7978,00	2,00
5730,00	10,95	5060,00	21025,00	7974,00	2,00
5710,00	69,05	104903,00	21025,00	5730,00	2,00
5741,00	17,12	10235,00	21025,00	5710,00	2,00
5719,00	2,78	2231,00	21025,00	5741,00	2,00
5730,00	12,04	10028,00	21025,00	5719,00	2,00
5710,00	45,32	64262,00	21025,00	5730,00	2,00

Figure IV 3 : Base de données après sélection de features

3. Choix de l'environnement de travail :

Nous avons utilisé python comme un langage de programmation.

Le langage de programmation Python a été créé par Guido van Rossum en 1990 et est rendu disponible sous licence libre. Son développement est aujourd'hui assuré par la « Python Software Foundation ». Python est un langage de programmation puissant et facile à apprendre, il dispose de structures de données de haut niveau et permet une approche simple mais efficace de la programmation orientée objet. Parce que sa syntaxe est élégante, que son typage est dynamique et qu'il est interprété, Python est un langage idéal pour l'écriture de scripts et le développement rapide d'applications dans de nombreux domaines et sur la plupart des plateformes.

3.1. Principales caractéristiques du langage Python :

C'est un langage open source ; Python est gratuit même pour les usages commerciaux. De nombreux outils standards sont disponibles. Python est facile à apprendre et à utiliser. C'est un langage de script qui le permet aussi créer des applications fonctionnelles rapidement et avec peu d'effort. C'est un langage interprété ; de nombreux interprètes interactifs sont disponibles. C'est orienté objet, mais la POO n'est pas obligatoire, il a aussi des outils pour le rendre possible programmation fonctionnelle ou impérative. C'est un langage très ouvert, il possède plusieurs interfaces avec C/C++/FORTRAN. Python a une excellente portabilité ; il fonctionne sur tous les systèmes d'exploitation et plateformes informatiques (les différentes variantes d'Unix, sur MacOS, BeOS, NeXTStep, MS-DOS et les différentes variantes de Windows). De nombreux modules sont disponibles dans les bibliothèques et librairies optimisées écrit en C, C++. En plus des modules intégrés dans la distribution Python standard, il existe des bibliothèques dans tous les domaines : scientifique, base de données, tests de fonctionnement et contrôle de qualité, graphiques 2D et 3D, etc ...

3.2. Environnement de développement pour Python :

- **SPYDER (Scientific Python Development Environment)** à usage scientifique de Python, doté de fonctionnalités avancées et intègre de nombreuses bibliothèques d'usage scientifique : Matplotlib, NumPy, SciPy et IPython. [53]

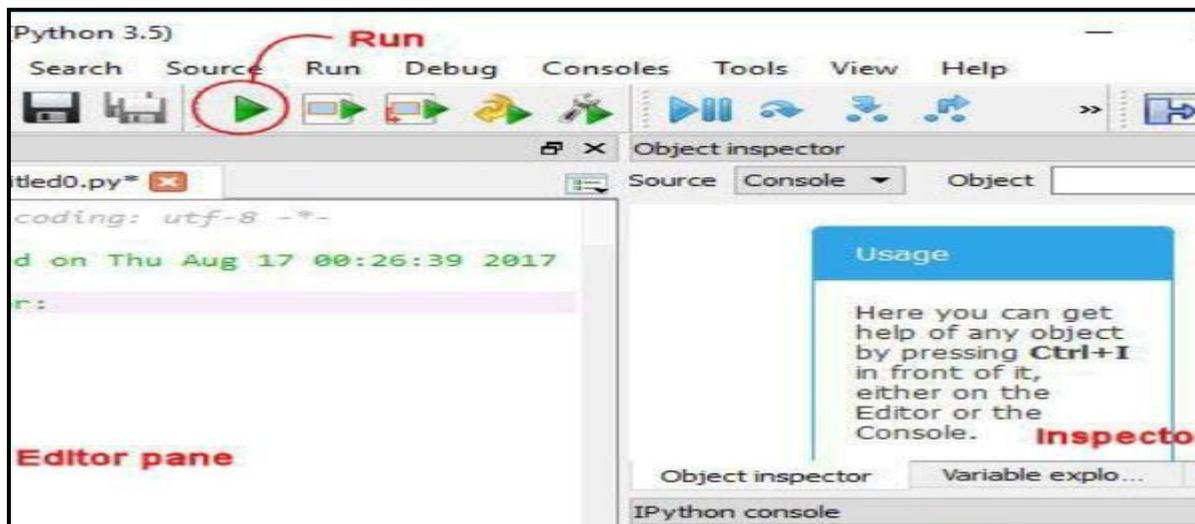


Figure IV 4 : la page de l'environnement Spyder pour Python. [53]

- **JUPITER Notebook** : L'interpréteur **IPython** (versions 0.12 à 3.x) offre la fonctionnalité *notebook* qui permet de créer des documents interactifs composés de code Python. [53]

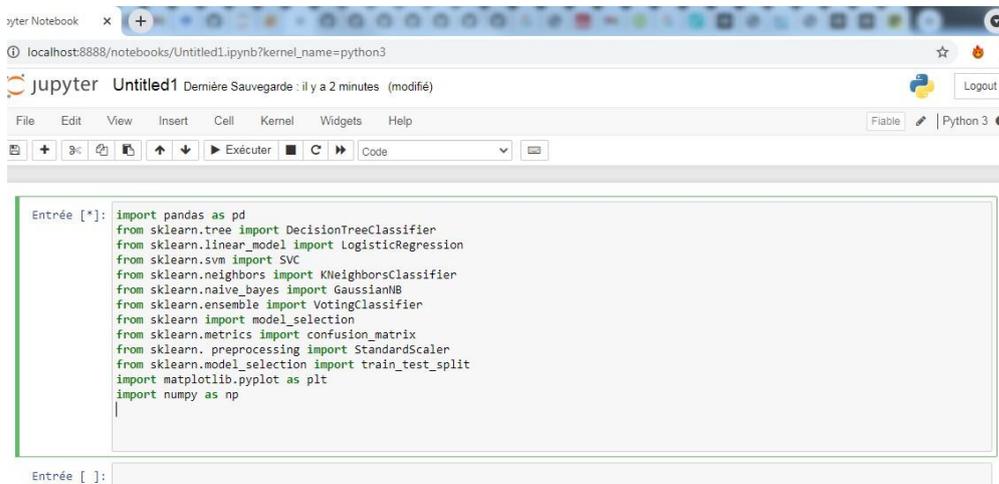


Figure IV 5 : la page de l'environnement Jupiter pour Python.

3.3. L'environnement de développement Anaconda :

L'environnement de développement **Anaconda** qui comportera la librairie standard, ainsi que les packages de base couramment utilisés pour du calcul scientifique (Numpy, Scipy, Matplotlib et bien plus encore) ainsi que IPython, Jupyter et Spyder.

Anaconda est un bundle Python3 très en vogue dans le monde scientifique. Développé par la société Anaconda Inc., libre dans sa version de base (nommée "Individual Edition"), il a l'avantage d'être multiplateforme (i.e. disponible sous Windows, macOS et GNU/Linux) et d'intégrer une grande quantité d'outils et packages Python, notamment : IPython, PIP, Jupyter, Spyder, NumPy, SciPy, Matplotlib, Pandas, Sympy, etc... [53]

3.4. Environnement Colab :

Est une plateforme utilisée par les développeurs et les informaticiens, elle offre plusieurs propriétés comme le machine Learning, le calcul, le stockage...etc. elle est aussi offert par Google ce qui implique que l'exécution est faite directement sur Google. [54]

3.5. Colab VS Anaconda (Jupyter) :

Le tableau suivant nous montre la différence entre les deux environnements de travail Colab et Jupiter

Sortie	Jupyter Notebook	Google Colab
Basé sur le cloud	Non	Oui
Synchronisation des fichiers	Non	Oui
Partage de fichier	No	Oui
L'installation de la bibliothèque	Oui	Non
Voir le fichier sans installation	Non	Oui

Tableau IV 1 : Jupyter Notebook VS Google Colab

De plus Avec Jupyter Notebook, vous devrez installer chaque bibliothèque que vous souhaitez utiliser sur votre appareil à l'aide de la commande pip .Vous serez également limité par la RAM, l'espace disque, le GPU et le CPU disponibles sur votre ordinateur. Cela étant dit, avoir les ordinateurs portables stockés sur votre matériel vs. Hébergé dans un cloud tiers est plus sécurisé, de sorte que l'installation manuelle de la bibliothèque peut en fait être un plus pour les données plus sensibles.

Comme conclusion, Google colab est un outil complet pour entraîner rapidement et tester rapidement des modèles d'apprentissage automatique sans avoir de contrainte matérielle. Sa particularité est que tout le monde peut l'utiliser (gratuit aussi !).

3.6. Principaux modules et bibliothèques utilisés dans notre programme :

- **Pandas** : est une bibliothèque qui permet de manipuler facilement des données à analyser. Il entre en jeu avant que l'ensemble de données ne soit préparé pour l'entraînement. Les pandas facilitent le travail avec des séries chronologiques et des données multidimensionnelles structurées pour les programmeurs d'apprentissage automatique.
- **NumPy** : cette bibliothèque permet d'effectuer des calculs numériques avec Python. Elle introduit une gestion facilitée des tableaux de nombres.

- **Scikit –learn** : est un outil simple utiliser pour l’analyse prédictive des données. Scikit-learn peut prendre en charge de nombreux algorithmes non supervisés et supervisés. Nous pouvons effectuer différentes tâches telles que la classification, la régression, le regroupement et d’autres tâches d’exploration de données avec cette bibliothèque Python. Cette bibliothèque comprend des algorithmes tels que les k-voisins, les forêts aléatoires, et prend en charge les machines vectorielles, etc.
- **Seaborn** : un sans pareil bibliothèque de visualisation, basée sur les fondations de Matplotlib. La narration et la visualisation de données sont importantes pour les projets d'apprentissage automatique, car elles nécessitent souvent une analyse exploratoire des ensembles de données pour décider du type d'algorithme d'apprentissage automatique à appliquer. Seaborn propose une interface basée sur un ensemble de données de haut niveau pour créer des graphiques statistiques étonnants.
- **Matplotlib** C’est une autre bibliothèque de Python dédié pour la visualisation des données, il est développé par John Hunter en 2002 qui a essayé de reproduire les fonctions de Matlab sur Python, avec une programmation de quelque ligne Matplotlib donne aux programmeurs la possibilité de tracé des diagrammes, des histogrammes et des tracés donc d’une autre façon Matplotlib permettre de visualisé les données.

3.7. Matplotlib VS Seaborn :

Features	Matplotlib	Seaborn
Fonctionnement	Il est utilisé pour réaliser des graphiques de base. Les ensembles de données sont visualisés à l'aide des barres graphiques, des histogrammes, de diagrammes circulaires, de diagrammes de dispersion, de lignes, etc.	Contient un certain nombre de modèles et de graphiques pour la visualisation des données. Il utilise des thèmes fascinants. Il aide à compiler toutes les données en un seul graphique. Il fournit également la distribution des données.
Syntaxe	Il utilise une syntaxe relativement complexe et longue. Exemple : Syntaxe pour bargraph- matplotlib.pyplot.bar(x_axis, y_axis).	Il utilise une syntaxe relativement simple qui est plus facile à apprendre et à comprendre. Exemple : Syntaxe pour bargraph-

		seaborn.barplot (x_axis, y_axis).
Transaction de plusieurs chiffres	Nous pouvons ouvrir et utiliser plusieurs figures simultanément. Cependant, elles sont fermées de manière distincte. Syntaxe pour fermer une figure à la fois : matplotlib.pyplot.close(). Syntaxe pour fermer toutes les figures : matplotlib.pyplot.close("all")	Seaborn fixe un délai pour la création de chaque figure. Cependant, cela peut entraîner des problèmes de mémoire (OOM).
Visualisation	Matplotlib est bien connecté avec Numpy et Pandas et agit comme un paquet graphique pour la visualisation de données en python. Pyplot offre des fonctionnalités et une syntaxe similaires à celles de MATLAB. Par conséquent, les utilisateurs de MATLAB peuvent facilement l'étudier.	Seaborn est plus à l'aise dans la manipulation des cadres de données Pandas. Il utilise des ensembles de méthodes de base pour fournir de beaux graphiques en python.
Souplesse	Matplotlib est un logiciel très personnalisé et robuste.	Seaborn évite la superposition des intrigues grâce à ses thèmes par défaut.
Cadres et tableaux de données	Matplotlib fonctionne efficacement avec les cadres de données et les tableaux et traite les chiffres et les axes comme des objets. Elle contient diverses API à état pour le traçage. Par conséquent, les	Seaborn est beaucoup plus fonctionnel et organisé que Matplotlib et traite l'ensemble des données comme une seule unité. Seaborn n'a pas d'état et, par conséquent, des paramètres sont nécessaires

	méthodes de type plot() peuvent fonctionner sans paramètres.	pour appeler des méthodes comme plot().
Cas d'utilisation	Matplotlib trace divers graphiques en utilisant Pandas et Numpy	Seaborn est la version étendue de Matplotlib qui utilise Matplotlib avec Numpy et Pandas pour tracer des graphiques.

Tableau IV 2 : Matplotlib vs Seaborn [55]

4. Préparation et analyse de la base de données :

L'analyse de données est un domaine issu du monde des statistiques qui vise à faire le lien entre les différentes données statistiques pour les classer, les décrire et les analyser de manière succincte.

L'objectif de l'analyse des données est d'extraire une information statistique qui permet de cerner plus précisément le profil de la donnée. Les résultats obtenus permettent ensuite d'optimiser la stratégie de la société en question en ajustant certains points.

4.1. Visualisation globale de data :

Après importation de différentes bibliothèques, avec la commande « import », nous avons utilisé la fonction **shape** pour monter la dimension de la Bdd et la fonction **info** pour nous renvoyer les informations sur cette Bdd .

```
[ ] # Import Python Libraries
import numpy as np
import scipy as sp
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

# Load dataset
df = pd.read_excel("//content/essay2.xlsx")
# Display dimensions of dataframe
print("df shape : ",df.shape)

print("-----")
# Information about the dataset
print(df.info())
```

Figure IV 6 : instruction des commande shape et info

La figure suivante montre les résultats d'exécution des deux commandes précédentes **shape** et **info**.

```
df shape : (418, 6)
-----
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 418 entries, 0 to 417
Data columns (total 6 columns):
 #   Column                                     Non-Null Count  Dtype
---  -
 0   format du produit                         418 non-null    float64
 1   Durée de fabrication en h                 418 non-null    float64
 2   quantité à fabriquer                      418 non-null    int64
 3   instant t1 de la machine ROBO 2400      418 non-null    float64
 4   instant t2 de la machine ROBO 2401      418 non-null    float64
 5   machine                                   418 non-null    int64
dtypes: float64(4), int64(2)
memory usage: 19.7 KB
None
```

Figure IV : 7 : *Résultat d'exécution de shape et info*

Les résultats montrent qu'il y a 417 lignes et 6 colonnes (format, durée en h, quantité, instant1, instant2 et machine), et le type de chaque colonne.

4.2. Affichage des informations statistique sur la Bdd :

La **fonction .describe()** fonction Pandas génère des statistiques descriptives qui résument la tendance centrale, la dispersion et la forme de la distribution d'un ensemble de données pour l'objet série donné. Tous les calculs sont effectués en excluant les valeurs NaN.

NB : valeur NaN sont des valeurs de type objet .

```
df.describe().T.style.format("{:.2f}").background_gradient(cmap='PuBu')
```

Figure IV 8 : *instruction d'Affichage des critères statistique*

	count	mean	std	min	25%	50%	75%	max
format du produit	418.00	12679.95	48081.71	56.52	5728.00	6789.50	7978.00	574413.00
Durée de fabrication en h	418.00	18.91	36.95	0.01	4.99	8.81	23.94	624.00
quantité à fabriquer	418.00	12122.17	18475.96	0.00	1418.25	5265.00	15891.75	144118.00
instant t1 de la machine ROBO 2400	418.00	10606.76	6354.57	579.00	5735.00	7969.50	21025.00	21030.00
instant t2 de la machine ROBO 2401	418.00	22388.84	82176.91	56.52	5730.00	7919.00	21025.00	574413.00
machine	418.00	1.52	0.50	1.00	1.00	2.00	2.00	2.00

Figure IV 9 : les statistiques relatives à la Bdd

4.3. Visualisation par des histogrammes :

```
# Lets know how many values for machine 1 and machine 2
print ("Total number lignes ", df.shape[0], ", ",
      df.machine.value_counts()[1], "1 and 2 are ",
      df.machine.value_counts()[2])
print("\n")
print("machine      : \n",df["machine"].value_counts())
print("\n")
print(df["machine"].value_counts(normalize=True))

# plot.bar(title='') function is used to plot bargraph
df["machine"].value_counts().plot.bar(title="machine")
```

Figure IV 10 : instruction d'affichage par des histogrammes

Cette commande est utilisée pour donner le nombre des exemples passant dans la machine 1 et le nombre des exemples passant dans la machine2.

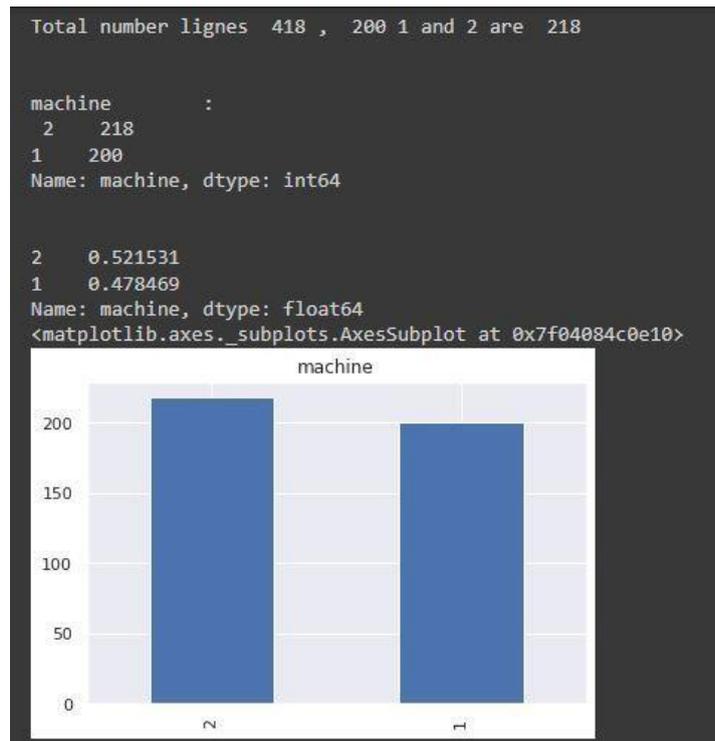


Figure IV 11 : résultat d’affichage par des histogrammes

A partir de cette résultat on a vu que le nombre de machine 1 = 200 et le nombre de machine 2 = 218, donc on remarque qu’il y a un déséquilibre dans la base de donnée.

4.4. Affichage de de la matrice de corrélation :

La corrélation explique comment une ou plusieurs variables sont liées les unes aux autres. Ces variables peuvent être des caractéristiques de données d’entrée qui ont été utilisées pour prévoir notre variable cible.

```
[20] ## Correlation matrix
      corrMatrix = df.corr()

      # plot corrMatrix
      fig, ax = plt.subplots(figsize=(10,10))
      # 'RdBu_r', 'BrBG_r', & PuOr_r are other good diverging colormaps
      # corrMatrix.style.format("{:.2f}").background_gradient(cmap="coolwarm", axis=None)

      # cmap= RdBu
      sns.heatmap(corrMatrix, fmt=".3f", annot_kws={"size":20}, annot=True,
                  linewidth=0.01, square=True, cmap="coolwarm", linecolor="black")
```

Figure IV 12 : instruction d’affichage de la matrice de corrélation

C'est l'instruction dédié pour calculer et afficher la matrice de corrélation pour voir est ce qu'il y a une forte corrélation entre les variable ou non ? C'est-à-dire est ce qu'il y a une indépendance entre les variables ou non ?

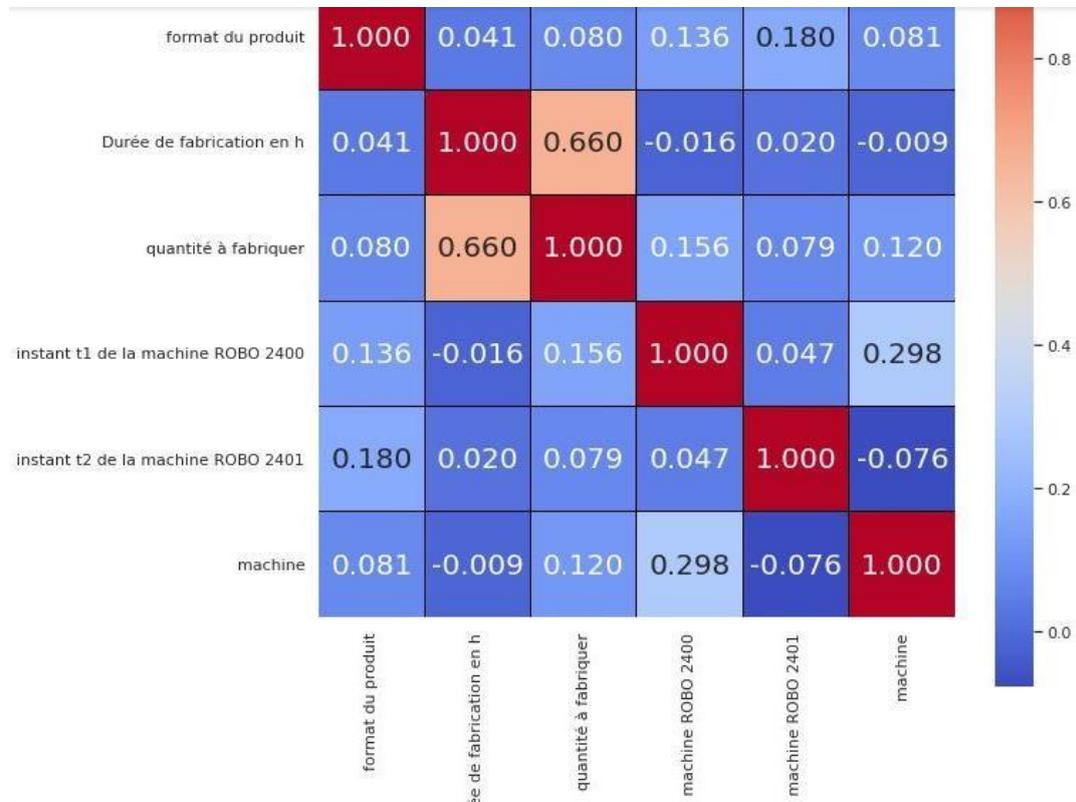


Figure IV 13 : affichage de la matrice de corrélation

- On remarque que la valeur de corrélation entre la durée de fabrication et quantité à fabriquer = 0.660 donc elle est vraiment proche de 1 donc on a une forte corrélation ou bien une forte indépendance entre ces deux variables.
- Il existe des valeurs moins corrélées par rapport à la durée de fabrication et la quantité à fabriquer par exemple instant t1 de la machine ROBO 2400 et la machine = 0.298 donc ces deux variables sont moyennement indépendant.
- On remarque que les autres valeurs sont loin de 1 ou bien -1 donc on n'a pas une forte corrélation entre les autres variables.

4.5. Affichage le nuage des points :

```
[21] # Scatterplot
plt.figure(figsize=(8,8))
sns.scatterplot(data=df, x="quantité", y="format", hue="machine")
plt.grid()
plt.show()
```

Figure IV 14 : instruction d'affichage par du nuage de points

Dans cette étape on affiche dans un graphe la relation ou bien l'indépendance entre la variable quantité à fabriquer et la variable format du produit.

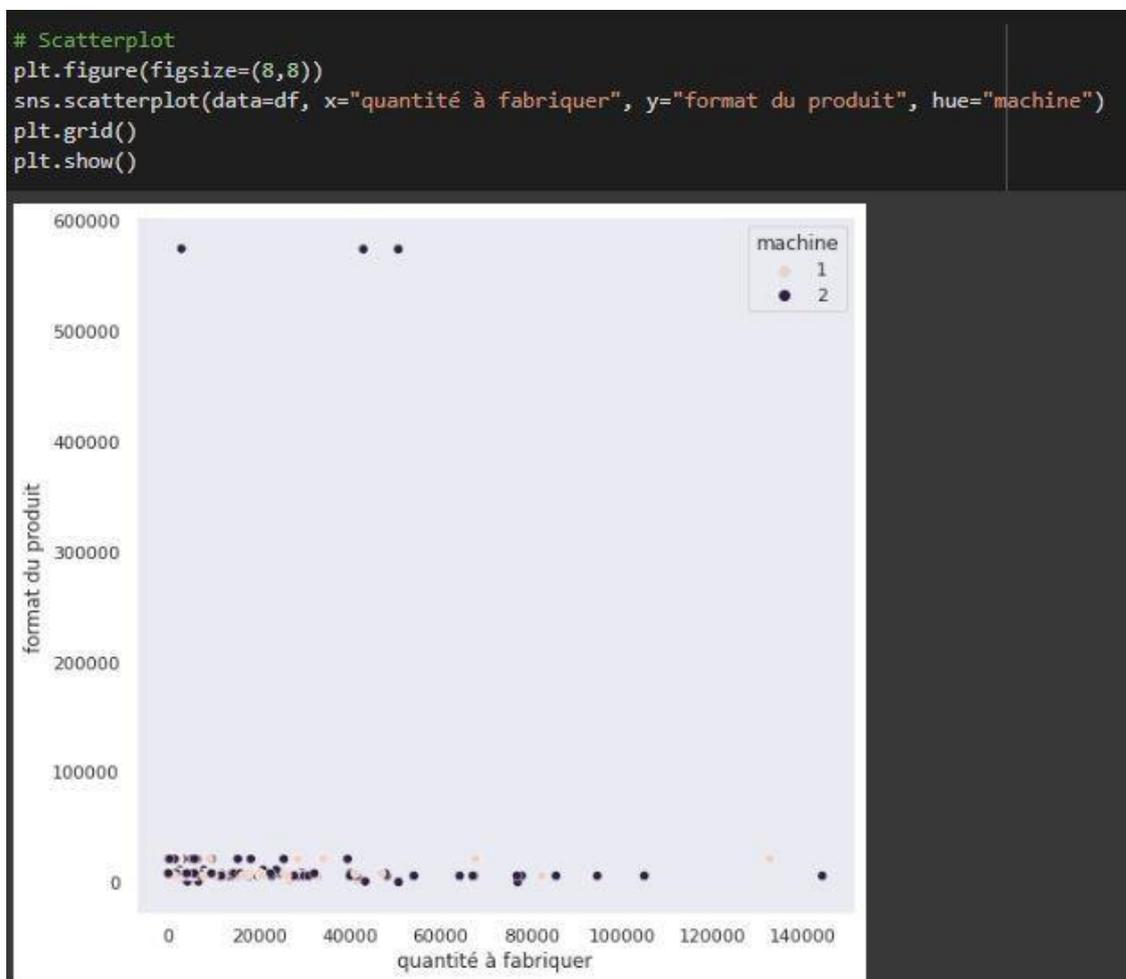


Figure IV 15 : résultat du nuage de points

À partir de cette figure on remarque que notre base de données est vraiment complexe parce que les points sont très proches et attachés l'un à l'autre donc il est vraiment difficile de les séparer.

4.6. Division de data sur deux : partie d'entraînement et partie de test :

Cette étape consiste à diviser la base de données en deux parties pour l'apprentissage, la première partie (80 %) de la base de données pour l'entraînement, la deuxième partie pour faire un test à notre programme.

```
# Split into train and test sets
from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X_scaler,
                                                  y, test_size= 0.20, random_state=42, stratify=y)

print("X_train = ",X_train.shape , " y_train = ", y_train.shape)
print("X_test = ",X_test.shape , " y_test = ", y_test.shape)

X_train = (348, 5) y_train = (348,)
X_test = (88, 5) y_test = (88,)
```

Figure IV 16 : les données de test et d'entraînement

Nous avons obtenu pour l'entraînement 348 lignes et 5 colonnes et le reste c.-à-d. 88 lignes pour le test.

5. Apprentissage :

- Nous avons utilisé plusieurs algorithmes de classification qui sont :
 - KNN
 - Logistic regression
 - Réseau de neurone
 - Forêts aléatoire
 - SVM

L'évaluation de la qualité du modèle est effectuée à l'aide de la matrice de confusion et le taux d'apprentissage ou le score de précision (accuracy score).

5.1. La matrice de confusion (confusion matrix) :

Appelé aussi tableau de contingence est un outil de machine Learning qui simplifiait les résultats permettant de mesurer des performances pour les modèles de classification et pour comprendre quelle façon le modèle est confus par rapport à la réalité, c'est-à-dire qu'il résume les résultats de prédictions et pour les comparer avec la réalité et voir comment les erreurs sont commises.

Cette matrice est calculée à travers une base de données qui contient une partie de test avec des valeurs de validation parce que cette matrice montre le nombre de classe correcte et fausse pour chaque valeur de test, elle est divisé sur 4 parties : vraie positive (true positive (TP)), vraie négative (true negative(TN)), fausse, positive (false positive(FP)) et fausse négative (false negative(FN)). Tel que :

- ❖ TP (true positive): c'est le cas où la valeur réelle est positive et la valeur de prédiction aussi positive.
- ❖ TN (true negative): c'est le cas où la valeur réelle est négative et la valeur de prédiction aussi négative.
- ❖ FP (false positive) : c'est le cas où la valeur de prédiction est positive et la valeur réelle négative.
- ❖ FN (false negative): c'est le cas où la valeur de prédiction est négative et la valeur réelle positive. [56]

		True Class	
		Positive	Negative
Predicted Class	Positive	TP	FP
	Negative	FN	TN

Figure IV 17 : exemple matrice de confusion

5.2. Score de précision (accuracy score) :

C'est une valeur pour évaluer la performance d'un modèle ou bien la fiabilité des algorithmes d'apprentissage.

$$\text{Accuracy score} = \frac{TP + TN}{TP + TN + FP + FN}$$

Exemple :

On prend l'exemple de la moyenne de 20 étudiants :

1 ➡ 12 étudiants ont la moyenne.

2 ➡ 8 étudiants n'ont pas la moyenne.

2	TP 7	FP 2
1	FN 5	TN 6
	2	1

Tableau IV 3 : matrice de confusion

TP : les étudiants qui ont la moyenne et classé juste.

TN : les étudiants qui n'ont pas la moyenne et sont classé juste.

FP : les étudiant qui n'ont pas la moyenne et qu'ils sont considéré qu'ont la moyenne.

FN : les étudiants qui ont la moyenne mais qu'ils sont considéré comme n'ont pas.

$$\text{accuracy score} = \frac{TP+TN}{TP+TN+FP+FN} = \frac{7+6}{7+6+2+5} = \frac{13}{20} = 0.65$$

5.3. Les algorithmes :

5.3.1. K-NN : voir chapitre 2 page 43

Les résultats de cet algorithme sont affichés dans la figure suivante :

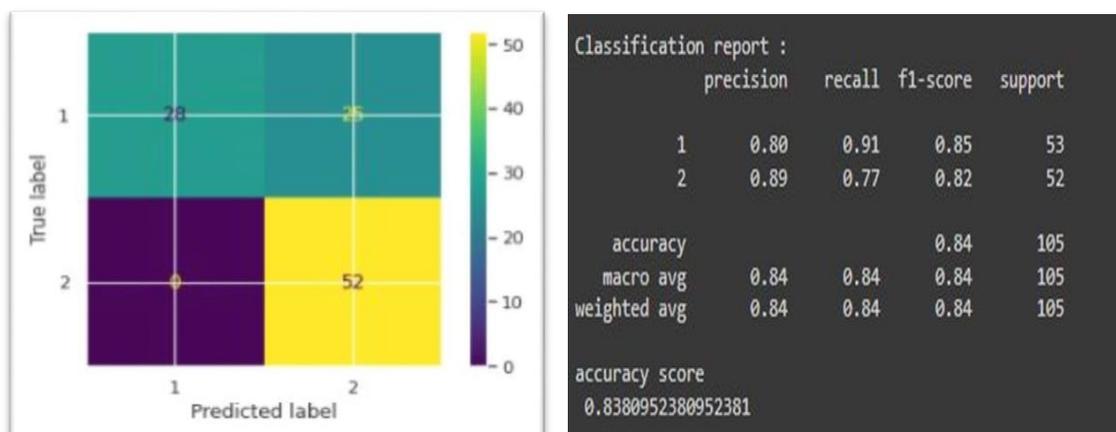


Figure IV : 18 : résultat K-NN

On remarque pour l'évaluation de la qualité de cet algorithme on a la matrix de confusion ci-dessus et l'accuracy score = 0.838 ce qu'est un taux d'apprentissage élevé et bon.

5.3.2. SVM : voir chapitre 2 page 50

Les résultats de cet algorithme sont affichés dans la figure suivante :

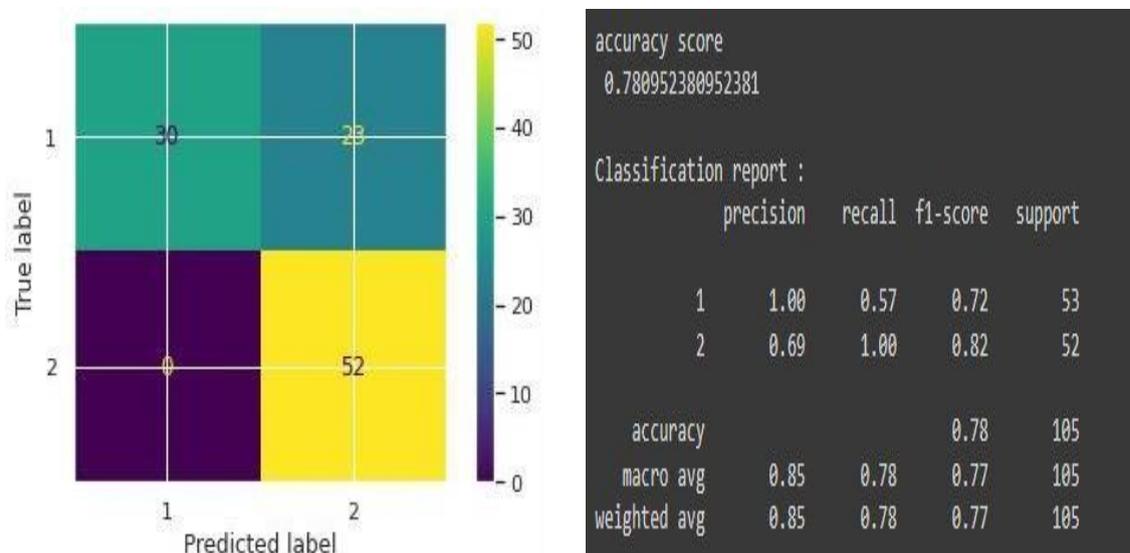


Figure IV 19 : résultat du SVM

Pour cet algorithme on a aussi la matrice de confusion avec un « accuracy score » = 0.78 qui est moins performant par rapport au résultat de l'algorithme KNN mais il reste comme même un bon résultat.

5.3.3. Logistic regression : voir chapitre 2 page 45

Les résultats de cet algorithme sont affichés dans la figure suivante :



Figure IV 20 : résultats du regression logistic

On remarque que « accuracy score » de cet algorithme = 57 qui est un mauvais résultat et qui est vraiment moins performant par rapport aux autres algorithmes.

5.3.4. Réseau de neurone : voir chapitre 2 page 39

Les résultats de cet algorithme sont affichés dans la figure suivante :

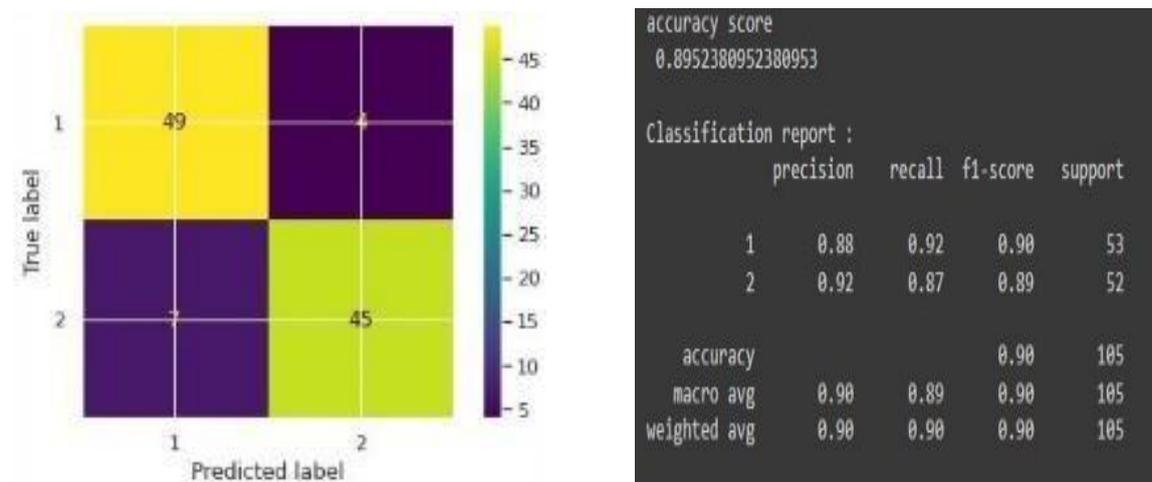


Figure IV 21 : résultats du réseau de neurone

On remarque que « accuracy score » de cet algorithme = 0.895 qui est la plus grande valeur et qui est le bon résultat par rapport aux autres algorithmes et par rapport à notre base de données.

5.3.5. Forêts aléatoire : voir chapitre 2 page 48

Les résultats de cet algorithme sont affichés dans la figure suivante :

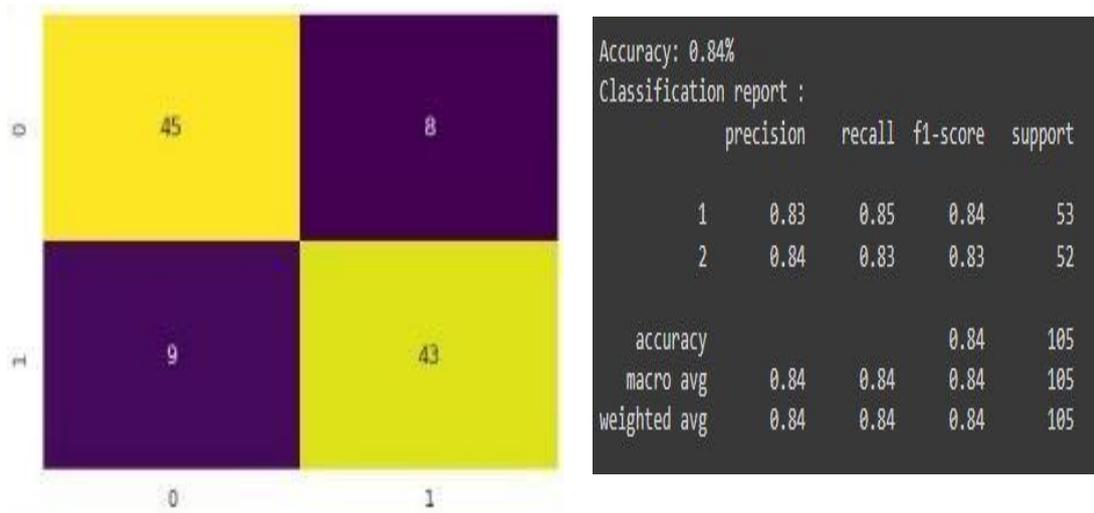


Figure IV 22 : résultat des forêts aléatoires

« accuracy score » de cet algorithme = 0.84 qui égale à « accuracy score » du KNN et plus performant par rapport aux regression logistique et SVM mais qui est moins performant par rapport au réseau de neurone, donc cet résultat reste bon.

5.3.6. Comparaison entre les algorithmes :

	Model	Accuracy
0	K-NN	83.809524
1	SVM	78.095238
2	Logistic regression	57.142857
3	Neural network	89.523810
4	Random forest	83.809524

Figure IV 23 : taux d'apprentissage par algorithme

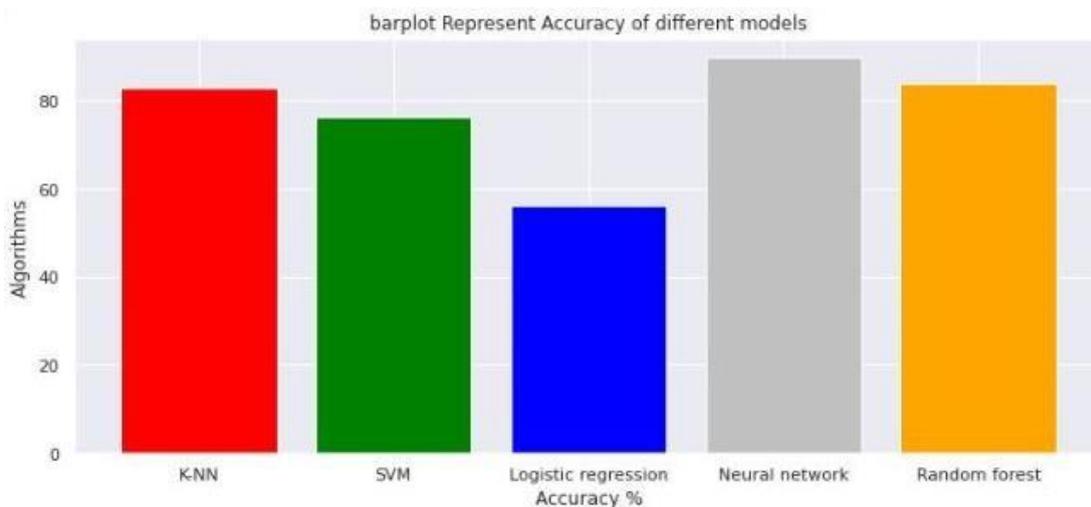


Figure IV 24 : comparaison des résultats des algorithmes avec des histogrammes

Les figures 4 et 5 ci-dessus montrent la comparaison des taux de réussites entre les différences types d'algorithme de classification que nous avons entraîné et testé. Nous avons obtenu un taux de 89,52 % pour l'algorithme neural network et un taux de 83 % pour les algorithmes KNN et random forest. Et pour l'algorithme SVM, nous avons obtenu un taux modéré égal à 78 %. Par contre, avec l'algorithme logistic regression nous avons obtenu un taux très faible = 57 %, et cela dû à la complexité de la base de données.

6. Déploiement :

Les modèles d'apprentissage automatique sont passionnants et puissants, mais ils ne sont pas très utiles en eux-mêmes. Une fois qu'un modèle est terminé, il doit probablement être déployé avant de pouvoir fournir une quelconque valeur. De plus, pouvoir déployer un modèle préliminaire ou un prototype pour obtenir des commentaires d'autres intervenants est extrêmement utile.

Récemment, il y a eu une émergence de plusieurs outils que les Data Scientists peuvent utiliser pour déployer rapidement et facilement un modèle d'apprentissage automatique. Dans ce travail, nous allons examiner 2 alternatives que nous avons utilisées pour déployer un modèle d'apprentissage automatique : Gradio, Streamlit,

6.1. Gradio :

C'est une bibliothèque du python (open source) qui permet à l'utilisateur de facilement utiliser des parties ajustables de l'UI (user interface) ou bien interface utilisateur pour les modules du ML, les APIs, ou n'importe quelle capacité subjective en deux lignes des codes. Il rendre l'utilisation des modules plus facile dans le navigateur par glisser les images, textes, ou les enregistrements de votre propre voix ou aussi des valeurs numérique comme dans notre cas.

Gradio aide aussi à construire des GUIs (Graphical User Interface) ou bien une interface graphique en ligne en quelques lignes seulement des codes ce qui est pratique pour montrer les expositions de la présentation du modèle. Il est vite, simple à réglé et prêt à utiliser et capable d'être partagé de telle façon que n'importe quelle personne peut exécuter le modèle dans la machine. [57]

6.2. Streamlit :

Elle est aussi une bibliothèque open source de python, elle donne aux utilisateurs la possibilité de créer des logiciels personnalisées et flexible par rapport aux leurs besoins à travers des lignes de programmation, généralement elle est utilisée pour les problèmes de machine Learning et pour afficher les données facilement. [58]

6.3. Gradio VS Streamlit :

	Maturité	Popularité	Simplicité	Adaptabilité	Focus
Gradio	C	B 7,413	C	B	ML model demos
Streamlit	C	A 19,514	C	B	Dashbord

Tableau IV 4 : résultat Gradio VS streamlit

Nous avons comparé ces bibliothèques sur :

- Maturité : Basée sur l'âge du projet et sa stabilité.
- Popularité : basée sur l'adoption et les étoiles GitHub
- Simplicité : Basé sur la facilité avec laquelle il est possible de commencer à utiliser la bibliothèque.
- Adaptabilité : basée sur la flexibilité et l'opinion de la bibliothèque.

- Focus : basé sur le problème que la bibliothèque résout.

6.4. Déploiement de notre projet : Gradio

Cet affichage est divisé sur deux parties les entrées et les sorties.

The screenshot shows a Gradio interface with two main sections. The left section is a form with the following fields and values:

- Format du produit: 21025
- durée de fabrication en h: 16
- quantité à fabriquer: 5000
- instant1 de la machine ROBO 2400: 21025
- instant2 de la machine ROBO 2401: 7978

At the bottom of the form are two buttons: 'Nettoyer' (grey) and 'Soumettre' (orange). The right section is a dark panel with a header 'output' and a large text area containing 'MACHINE 1'. Below this is a button labeled 'Signaler'.

Figure IV 25 : *exemple d'affichage de la prise de décision*

Cette figure montre un cas de décision pour une commande de format 210*25 avec un temps de fabrication = 16h et une quantité = 5000, dans l'instant t la machine ROBO 2400 était en train de traiter un produit de format 21025 et la machine ROBO 2401 était en train de traiter un produit de format 7978. Dans ce cas-là notre programme à affecter la commande arrivé pour que la machine 1 (ROBO 2400) peu la traiter.

6.5. Déploiement de notre projet : Streamlit

The screenshot displays a Streamlit web application interface. On the left, there is a teal sidebar with a header "Durée de fabrication" and a numeric input field containing "0,00" with minus and plus buttons. The main content area on the right has a teal background and contains several input fields, each with a header and a numeric input field with minus and plus buttons:

- format du produit: 0,00
- quantité à fabriquer: 0,00
- inst T1 de la machine ROBO 2400 =: 0,00
- inst T2 de la machine ROBO 2401: 0,00

Below these fields, the text "your data :" is followed by a table:

	format	durée	quantité	instant1	instant2
0	0.0000	0.0000	0.0000	0.0000	0.0000

Overlaid on the bottom right of the screenshot is a white box containing the following text and elements:

- accuracy_score:**
86.13861386138613%
- results**
A 2x2 confusion matrix table:

	0
0	2
- Are you satisfied with the result
- So what is your rate :

Tableau IV 5 : résultat du streamlit

7. Conclusion :

Aujourd'hui la prise de décision devient un élément essentiel pour les entreprises mais avec le temps elle devient compliqué et difficile à cause des développements technologiques, donc il est nécessaire d'utiliser les outils et les logiciels performants pour obtenir une meilleure décision.

A l'aide de notre travail on a fait un programme pour la prise de décision ou bien on peut dire pour le choix des machines (ROBO 2400 et ROBO 2401) avec plusieurs algorithmes et avec deux contraintes qui sont la durée de fabrication et la quantité à produire. Dans notre cas les réseaux de neurones donnent un taux d'apprentissage très élevé égale à 90% qui est un bon résultat par rapport à la complexité de la base de données.



Conclusion générale



Conclusion générale

Dans l'industrie, l'intelligence artificielle a d'abord pour but d'aider à la décision.

On distingue trois niveaux d'aide :

En premier lieu, l'intelligence artificielle descriptive permet, à partir d'une masse de données ou connaissances importantes, de simplifier l'information et de la présenter sous forme de tableaux de bord et de KPI (Key Performance Indicator). Un gain de temps non négligeable puisque les analyses réalisées grâce à l'IA sont plus rapides, moins coûteuses et surtout beaucoup plus fiables.

L'IA prédictive, quant à elle, fournit, sur la base de probabilités, des projections. On anticipe donc aisément les situations futures et on prévoit mieux les risques en termes de supply chain et de défauts de manufacturing par exemple.

L'intelligence artificielle ne se limite plus à donner les informations nécessaires à la prise de décision mais aiguille également les salariés en leur proposant des recommandations vis-à-vis de l'état du système. C'est ce qu'on appelle la "next best action". Dans certains cas, on l'utilise également pour décharger l'homme de certaines tâches à faibles risques ou sans enjeux stratégiques. De plus en plus d'entrepôts par exemple, confient des portions de transport logistique à des véhicules autonomes.

Quel que soit le type d'intelligence artificielle utilisé, le système va donc permettre de capter un flux de données important, de les analyser, à des fins stratégiques, tactiques ou opérationnelles.

Par conséquent, dans ce manuscrit nous avons cherché à développer un outil d'aide à la décision en se basant sur l'apprentissage automatique sur des données de l'entreprise MEGA PAPIER plus précisément filiale « *MEGATHERM* ». , à identifier les principales difficultés, et à proposer des solutions à travers des concepts et du modèle permettant de surmonter certains de ces problèmes et ainsi mieux exploiter un potentiel encore très large de cette technique. En effet, nous estimons qu'à terme, une vision intelligente et active devra permettre aux différents utilisateurs, d'obtenir une aide décisionnelle plus fiable plus représentative de la réalité.

Le présent travail a porté sur la conception d'un outil de classification des machines grâce à l'apprentissage automatique. Une étude de la littérature a été menée dans un premier temps. Et cela nous a permis d'identifier les technologies et outils existants pour la prise de décision et les algorithmes d'apprentissage machine disponible pour analyser les données.

En conclusion, les informations données par l'entreprise et de l'apprentissage automatiques peuvent aider les responsables de bien choisir entre les deux machines afin de réduire le temps de changement ce qui implique livrer le produit au bon moment. Avec un taux d'apprentissage élevé égale à 89 %, cela démontre que l'utilisation des données permettre une prédiction plus précise pour le choix de la machine.

Perspective :

Nous visons dans le long terme à améliorer ce travail, pour cela, nous proposons les perspectives suivantes :

- Tenir en compte d'autres contraintes (features) dans la base de données.
- Application de ce prototype dans les autres chaines de production de l'entreprise Megapapiers.
- Prendre en considération les qualifications des employées pour le choix de la machine.

Bibliographie :

- [1] M. P. e. C. M. - B. Alain COURTOIS, *Gestion de production*, 4^{ème} édition éd., 2003.
- [3] G. Javel, *organisation et gestion et de la production*, paris: dunod, 2010.
- [13] V. & M. Y. Cherkassky, *Another look at statistical learning theory and regularization*, 2009.
- [14] *Recent advances and trends in predictive manufacturing systems in big data environment*.
- [15] M. S. N. S. S. N. B. e. S. V. Elangovan, *Machine Learning Approach to the Prediction of Surface Roughness Using Statistical Features of Vibration Signal Acquired in Turning*, 2015.
- [16] Y. Lu, *Industry 4.0: A survey on technologies, applications and open research issues*, 2017.
- [17] Kusiak, *personnalisation : étendre les concepts de fabrication sur commande aux chaînes d'approvisionnement.*, 2017.
- [18] M. e. al, *Fabrication intelligente : Caractéristiques et technologies. Dans Conférence internationale sur la gestion du cycle de vie des produits*, 2016.
- [19] W. e. Fawcett, *Science des données, analyse prédictive et mégadonnées : A révolution qui transformera la conception et la gestion de la chaîne d'approvisionnement.*, 2013.
- [20] Monostori, *Techniques d'IA et d'apprentissage automatique pour gérer la complexité, changements et incertitudes dans la fabrication. Applications d'ingénierie de l'intelligence artificielle*, 2003.
- [21] A. B. E. e. D. T. McAfee, *Big data : la gestion révolution*, 2012.
- [22] D. & A. A. Pham, *Techniques d'apprentissage automatique et leurs applications dans fabrication*, 2005.
- [23] J. S. M. & K. A. Harding, *Exploration de données dans le secteur manufacturier : un examen.*, 2006.
- [24] Lu, *Un ensemble de neurones à convolution profonde modèle de réseau avec fusion améliorée des preuves DS pour le diagnostic des défauts de roulement.*, 2017.
- [25] L. X. E. e. L. L. Xu, *Industrie 4.0 : État de l'art et tendances futures.*, 2018.
- [26] M. & S. F. Martínez-Díaz, *Véhicules autonomes : théorie et défis pratiques. Procédure de recherche sur les transports*, 2018.
- [27] M. Müller, *L'ordinateur va. Intelligence artificielle*, 2002.
- [35] B. Ouail, *Etude et comparaison de modèles de prédiction basés sur l'apprentissage automatique*, 2021.
- [36] *Analyse de données et apprentissage statistique en R*.

- [37] «SVM - Data Analytics Post,» [En ligne]. Available: [https://dataanalyticspost.com/Lexique/svm/#:~:text=SVM%20\(Support%20Vector%20Machine%20ou,ou%20de%20d%C3%A9tection%20d%27anomalie..](https://dataanalyticspost.com/Lexique/svm/#:~:text=SVM%20(Support%20Vector%20Machine%20ou,ou%20de%20d%C3%A9tection%20d%27anomalie..) [Accès le 18 mai 2022].
- [38] «les SVM, Support Vecteur machine, Analytics & Insights,» [En ligne]. Available: <https://analyticsinsights.io/les-svm-support-vector-machine/>. [Accès le 13 juin 2022].
- [52] *Pour une recherche économique efficace*, 2020.
- [53] J.-D. B. e. S. Bancal, *Installation et utilisation de Python et outils associés*, 2020.
- [63] M. e. S. (Martínez-Díaz, *Véhicules autonomes : théorie et défis pratiques.Procédure de recherche sur les transports.*, 2018.

Webographie

- [2] «qu'est-ce que le juste à temps,» [En ligne]. Available: <https://www.piloter.org/six-sigma/juste-a-temps.htm>. [Accès le 5 juin 2022].
- [4] «Qu'est-ce que la gestion de la production ? Son utilité et ses avantages - FORCAM,» [En ligne]. Available: <https://forcam.com/fr/la-gestion-de-la-production/>. [Accès le 29 mai 2022].
- [5] «définition : mode de production,» [En ligne]. Available: https://www.toupie.org/Dictionnaire/Mode_production.htm#:~:text=D%C3%A9finition%20de%20mode%20de%20production,un%20produit%20ou%20un%20service.. [Accès le 05 juin 2022].
- [6] «Découvrez le flux continu ! - Pilotez l'amélioration continue dans l'industrie du futur,» [En ligne]. Available: <https://openclassrooms.com/fr/courses/5382991-pilotez-lamelioration-continue-dans-lindustrie-du-futur/5714251-decouvrez-le-flux-continu>. [Accès le 05 juin 2022].
- [7] «Imaginez un design innovant et au plus juste de l'usine, pour la rendre agile - Pilotez l'amélioration continue dans l'industrie du futur,» [En ligne]. Available: <https://openclassrooms.com/fr/courses/5382991-pilotez-lamelioration-continue-dans-lindustrie-du-futur/5715396-imaginez-un-design-innovant-et-au-plus-juste-de-lusine-pour-la-rendre-agile>. [Accès le 5 juin 2022].
- [8] «Production en Flux poussé ou Flux tiré : quelle méthode choisir?,» [En ligne]. Available: <https://www.picomto.com/production-en-flux-pousse-ou-flux-tire-laquelle-est-la-mieux-adaptee-a-votre-entreprise/>. [Accès le 5 juin 2022].
- [9] «Flux poussés : tout ce qu'il faut savoir pour mieux servir vos clients,» [En ligne]. Available: <https://www.supplychaininfo.eu/flux-pousse/>. [Accès le 25 mai 2022].
- [10] «lean machine square : Back to basics... Le flux tiré (# 1),» [En ligne]. Available: <http://leanmachineblog.blogspot.com/2008/10/back-to-basics-le-flux-tir-1.html>. [Accès le 25 mai 2022].

- [11] «Le système Juste à Temps (JAT) dans un entrepôt | AR Racking France,» [En ligne]. Available: <https://www.ar-racking.com/fr/actualite/blog-solutions-stockage/solutions-de-stockage/le-systeme-juste-a-temps-jat-dans-un-entrepot#>. [Accès le 28 mai 2022].
- [12] «Navettes pour Palettes - AR Shuttle | Stockage Automatique | AR Racking,» [En ligne]. Available: <https://www.ar-racking.com/fr/solutions-stockage/entrepots-automatisees/palettes/navette-pour-palettes>. [Accès le 28 mai 2022].
- [28] «intelligence artificielle : definition, histoire, utilisation, dangers,» [En ligne]. Available: <https://datascientest.com/intelligence-artificielle-definition>. [Accès le 12 avril 2022].
- [29] «Au fait, c'est quoi l'intelligence artificielle ?,» [En ligne]. Available: <https://www.europe1.fr/technologies/au-fait-cest-quoi-lintelligence-artificielle-3612572>. [Accès le 14 juin 2022].
- [30] «La naissance d'une science : les origines de l'intelligence artificielle – [1/2],» [En ligne]. Available: <https://www.journaldunet.com/solutions/dsi/1493877-la-naissance-d-une-science-les-origines-de-l-intelligence-artificielle-1-2/>. [Accès le 14 juin 2022].
- [31] «Définition | Deep Learning - Apprentissage profond | Futura Tech,» [En ligne]. Available: <https://www.futura-sciences.com/tech/definitions/intelligence-artificielle-deep-learning-17262/>. [Accès le 17 mai 2022].
- [32] «KNN : découvrez cet algorithme de Machine Learning,» [En ligne]. Available: <https://datascientest.com/knn>. [Accès le 15 juin 2022].
- [33] «Titanic _ arbre de décision (1) - Pensée Artificielle,» [En ligne]. Available: https://penseeartificielle.fr/tout-pour-bien-debuter-en-machine-learning-4/titanic_-arbre-de-decision-1/. [Accès le 18 juin 2022].
- [34] «Arbre de décision, comment ça marche? LOVELY ANALYTICS,» [En ligne]. Available: <https://www.lovelyanalytics.com/2016/08/16/decision-tree-comment-ca-marche/#:~:text=Le%20principal%20inconvenient%20des%20arbres,choix%20entre%20performance%20et%20interpr%C3%A9tabilit%C3%A9..> [Accès le 13 juin 2022].
- [36] *Analyse de données et apprentissage statistique en R.*
- [37] «SVM - Data Analytics Post,» [En ligne]. Available: [https://dataanalyticspost.com/Lexique/svm/#:~:text=SVM%20\(Support%20Vector%20Machine%20ou,ou%20de%20d%C3%A9tection%20d%27anomalie..](https://dataanalyticspost.com/Lexique/svm/#:~:text=SVM%20(Support%20Vector%20Machine%20ou,ou%20de%20d%C3%A9tection%20d%27anomalie..) [Accès le 18 mai 2022].
- [38] «les SVM, Support Vecteur machine, Analytics & Insights,» [En ligne]. Available: <https://analyticsinsights.io/les-svm-support-vector-machine/>. [Accès le 13 juin 2022].
- [39] «Avantages et inconvénients des différents modèles de classification - Acervo Lima,» [En ligne]. Available: <https://fr.acervolima.com/avantages-et-inconvenients-des-differents-modeles-de-classification/>. [Accès le 15 juin 2022].
- [40] «Qu'est-ce que l'apprentissage non supervisé | Linedata,» [En ligne]. Available: <https://fr.linedata.com/quest-ce-que-lapprentissage-non-supervise>. [Accès le 18 mai 2022].

- [41] «Reinforcement Learning : qu'est-ce que l'apprentissage par renforcement ?,» [En ligne]. Available: <https://www.lebigdata.fr/reinforcement-learning-definition>. [Accès le 20 mai 2022].
- [42] «Quelle différence entre Machine Learning et Deep Learning ?,» [En ligne]. Available: <https://datascientest.com/quelle-difference-entre-le-machine-learning-et-deep-learning>. [Accès le 15 juin 2022].
- [43] «Intelligence artificielle et machine learning dans l'industrie 4.0 : Le guide complet,» [En ligne]. Available: <https://datact.io/intelligence-artificielle-et-machine-learning-dans-lindustrie-4-0-le-guide-complet/>. [Accès le 20 mai 2022].
- [44] «Qu'est-ce que l'Industrie 4.0 - IP Systèmes,» [En ligne]. Available: <https://www.ip-systemes.com/details-qu+est+ce+que+l+industrie+4+0-798.html>. [Accès le 20 mai 2022].
- [45] «Facebook,» facebook, [En ligne]. Available: <https://web.facebook.com/megapapiers/photos/a.252989568956097/458900701698315/>. [Accès le 30 mai 2022].
- [46] «présentation Méga papiers by Meriem Djafour Baghli,» [En ligne]. Available: <https://prezi.com/p/921azcifkb7c/presentation-mega-papiers/>. [Accès le 30 mai 2022].
- [47] «MegaPrint | Megapapiers,» [En ligne]. Available: <https://www.mega-papiers.com/megaprint/>. [Accès le 30 mai 2022].
- [48] «MegaLabel | Megapapiers,» [En ligne]. Available: <https://www.mega-papiers.com/megalabel/>. [Accès le 30 mai 2022].
- [49] «MegaPlotter | MegaPapiers,» [En ligne]. Available: <https://www.mega-papiers.com/megaplotter/>. [Accès le 30 mai 2022].
- [50] «megatherm | Mega Papiers,» [En ligne]. Available: <https://www.mega-papiers.com/megatherm/>. [Accès le 30 mai 2022].
- [51] «Qu'est-ce que le papier thermique? - définition de techopedia - Matériel - 2022,» [En ligne]. Available: <https://fr.theastrologypage.com/thermal-paper>. [Accès le 30 mai 2022].
- [54] «Google Cloud Platform : Tout savoir sur la plateforme cloud de Google,» [En ligne]. Available: <https://www.lebigdata.fr/google-cloud-platform-tout-savoir>. [Accès le 11 juin 2022].
- [55] «Difference Between Matplotlib VS Seaborn - GeeksforGeeks,» [En ligne]. Available: <https://www.geeksforgeeks.org/difference-between-matplotlib-vs-seaborn/>. [Accès le 18 juin 2022].
- [56] «Confusion Matrix : l'outil de mesure de performances du Machine Learning,» [En ligne]. Available: <https://www.lebigdata.fr/confusion-matrix-definition>. [Accès le 15 juin 2022].
- [57] «Gradio Library | Create Interface For Your Machine Learning Models Using Gradio Python Library,» [En ligne]. Available: <https://www.analyticsvidhya.com/blog/2021/04/create->

interface-for-your-machine-learning-models-using-gradio-python-library/. [Accès le 15 juin 2022].

- [58] «Streamlit ou l’outil pour présenter votre travail de Machine Learning,» DataScientest, [En ligne]. Available: <https://datascientest.com/streamlit-ou-loutil-pour-presenter-votre-travail-de-machine-learning>. [Accès le 15 juin 2022].
- [59] «définition | AlphaGo | Futura Tech,» [En ligne]. Available: <https://www.futura-sciences.com/tech/definitions/intelligence-artificielle-alphago-16467/>. [Accès le 12 avril 2022].
- [60] «le machine learning, un apprentissage automatique,» [En ligne]. Available: <https://www.futura-sciences.com/tech/actualites/algorithm-machine-learning-apprentissage-automatique-91879/>. [Accès le 17 mai 2022].
- [61] «apprentissage supervisé vs apprentissage non supervisé | mobiskill,» 17 mai 2022. [En ligne]. Available: <https://mobiskill.fr/blog/conseils-emploi-tech/apprentissage-supervise-vs-apprentissage-non-supervise/>.
- [62] «Avantages et inconvénients de la régression logistique - Acerveau Lima,» [En ligne]. Available: <https://fr.acervolima.com/avantages-et-inconvenients-de-la-regression-logistique/#:~:text=La%20régression%20logistique%20est%20plus,peut%20conduire%20à%20un%20surajustement..> [Accès le 13 juin 2022].
- [64] «La vraie différence entre Machine Learning & Deep Learning | Jedha,» [En ligne]. Available: <https://www.jedha.co/formation-ia/vraie-difference-machine-learning-deep-learning#:~:text=Il%20est%20souvent%20expliqué%20que,son%2C%20le%20texte%2C%20I%27> 27. [Accès le 14 juin 2022].
- [65] «Ingénierie simultanée : Du séquentiel au simultané | Techniques de l'ingénieur,» [En ligne]. Available: <https://www.techniques-ingenieur.fr/base-documentaire/genie-industriel-th6/strategies-de-conception-pour-l-innovation-42127210/ingenierie-simultanee-a5310/du-sequentiel-au-simultane-a5310niv10005.html>. [Accès le 5 juin 2022].
- [66] «Xcon,» [En ligne]. Available: <https://stringfixer.com/fr/XCON>. [Accès le 15 juin 2022].