Table des matières

[Liste des figures](#_Toc104720528)

[Liste des Tableaux](#_Toc104720529)

[Introduction générale 1](#_Toc104720530)

[Chapitre 1. Contexte générale 2](#_Toc104720531)

[Introduction 2](#_Toc104720532)

[1.1. Problématique et Présentation du projet 2](#_Toc104720533)

[1.1.1. Contexte du Projet 2](#_Toc104720534)

[1.1.2. Problématique 2](#_Toc104720535)

[1.1.3. Planification du projet 2](#_Toc104720536)

[1.1.4. Planification du projet 2](#_Toc104720537)

[1.2. Présentation de l’entreprise d’accueil SG-ATS 2](#_Toc104720540)

[1.2.1. Group Société Générale 2](#_Toc104720541)

[1.2.2. Société générale Corporate Investment Banking 2](#_Toc104720542)

[1.2.3. Société générale Africa Technologie & Sercices 2](#_Toc104720543)

[1.2.4. L’équipe Trading Process Factory 2](#_Toc104720544)

[Conclusion 2](#_Toc104720545)

[Chapitre 2. Cadre théorique : Définitions et littérature 3](#_Toc104720546)

[Introduction 3](#_Toc104720547)

[2.1. Généralités sur les options 3](#_Toc104720548)

[2.1.1. Définition 3](#_Toc104720549)

[2.1.2. Cas d’options américaines 3](#_Toc104720550)

[2.1.2.1. Résultat à l’échéance 3](#_Toc104720551)

[2.1.2.2. Défi principale 3](#_Toc104720552)

[2.1.2.3. Contexte historique 3](#_Toc104720553)

[2.1.3. Modèles de valorisation 4](#_Toc104720554)

[2.1.3.1. Formulation du problème 4](#_Toc104720563)

[2.1.3.2. EDP 4](#_Toc104720564)

[2.1.3.3. Modèle binomiale 4](#_Toc104720565)

[2.1.3.4. Critiques et limitation  4](#_Toc104720566)

[2.1.4. Problème du fugit 4](#_Toc104720567)

[2.1.4.1. Définition 4](#_Toc104720568)

[2.1.4.2. Adaptation des approches de valorisation 4](#_Toc104720569)

[2.2. Apprentissage Automatique 4](#_Toc104720589)

[2.2.1. Concept et théorie 4](#_Toc104720590)

[2.2.1.1. Types d’apprentissage 4](#_Toc104720591)

[2.2.1.2. Technique et évolution 5](#_Toc104720592)

[2.2.2. Problème d’apprentissage 5](#_Toc104720593)

[2.2.2.1. Overfitting - Underfitting 5](#_Toc104720594)

[2.2.2.2. Solution : Régularisation 5](#_Toc104720595)

[2.3.6. Prétraitement 6](#_Toc104720596)

[2.3.6.1. Transformation 6](#_Toc104720597)

[2.3.6.2. Gestion des valeurs manquantes 6](#_Toc104720598)

[2.3.6.3. La Régularisation 6](#_Toc104720599)

[2.3.7. Modèles de régression 6](#_Toc104720600)

[2.3.7.1. Support vector machine 6](#_Toc104720601)

[2.3.7.2. Forêts aléatoires 6](#_Toc104720602)

[2.3.7.3. Gradient Boosting et XGBoost 6](#_Toc104720603)

[2.3.8. Feature Importance 7](#_Toc104720604)

[2.3.8.1. Approche globale 7](#_Toc104720605)

[2.3.8.2. Approche individuelle 7](#_Toc104720606)

[Conclusion 7](#_Toc104720607)

[Chapitre 3. État de l’art et environnement de développement 8](#_Toc104720608)

[Introuction 8](#_Toc104720609)

[3.1. Gestion des processus métiers 8](#_Toc104720610)

[3.1.1. Définition d’un processus métier 8](#_Toc104720611)

[3.1.2. Définition d’un workflow 8](#_Toc104720612)

[3.1.3. Business Model and Notation : 8](#_Toc104720613)

[3.2. Environnement de développement 8](#_Toc104720614)

[3.2.1. Outils de gestion et collaboration 8](#_Toc104720615)

[3.2.1.1. Jira 8](#_Toc104720616)

[3.2.1.2. Kibana & ElasticSearch 8](#_Toc104720617)

[3.2.1.3. Github Entreprise : 8](#_Toc104720618)

[3.2.2. Techniques et langage de programmation 8](#_Toc104720619)

[3.2.2.1. Python 8](#_Toc104720620)

[3.2.2.2. Programmation orienté objet 8](#_Toc104720621)

[3.2.2.3. Parallélisme 8](#_Toc104720622)

[3.2.2.4. Programmation asynchrone 8](#_Toc104720623)

[3.3. Connexions et Transfert de données 8](#_Toc104720624)

[3.3.1. Concept des APIs 8](#_Toc104720625)

[3.3.2. Requête et connexion aux APIs 8](#_Toc104720626)

[Conclusion 8](#_Toc104720627)

[Chapitre 4. Cas d’un processus métier réel 8](#_Toc104720628)

[Introduction 8](#_Toc104720629)

[4.1. Besoin fonctionnel 8](#_Toc104720630)

[4.1.1. Exercice anticipé à l’échelle de production 8](#_Toc104720631)

[4.1.1.1. Processus générique et réutilisable 8](#_Toc104720632)

[4.1.1.2. Scalabilité 8](#_Toc104720633)

[4.1.1.3. Efficience 8](#_Toc104720634)

[4.1.2. Intégration au système interne 8](#_Toc104720635)

[4.1.2.1. Conception du processus parent 8](#_Toc104720636)

[4.1.2.2. Exercice effective et booking de l’opération 8](#_Toc104720637)

[4.2. Modélisation et automatisation 8](#_Toc104720638)

[4.2.1. Modélisation en BPMN 8](#_Toc104720639)

[4.2.2. Preuve de concept 8](#_Toc104720640)

[4.2.2.1. Etude de faisabilité technique 8](#_Toc104720641)

[4.2.2.2. Construction d’une application interactive 8](#_Toc104720642)

[4.2.3. Automatisation du processus 9](#_Toc104720643)

[4.2.3.1. Processus parent 9](#_Toc104720644)

[4.2.3.2. Processus fils 9](#_Toc104720645)

[4.3. Collecte efficace et fiable de données 9](#_Toc104720646)

[4.3.1. Identification des sources de données 9](#_Toc104720647)

[4.3.1.1. Eliot 9](#_Toc104720648)

[4.3.1.2. MSD 9](#_Toc104720649)

[4.3.1.3. Bloomberg 9](#_Toc104720650)

[4.3.1.4. Shiva 9](#_Toc104720651)

[4.3.1.5. RDWS 9](#_Toc104720652)

[4.3.2. Connexion aux APIs et récupération de données 9](#_Toc104720653)

[4.3.2.1. PyBaseRequest 9](#_Toc104720654)

[4.3.2.2. Parallélisme et Passage à l’échelle 9](#_Toc104720655)

[Chapitre 5. Analyse et amélioration de l’efficacité et la fiabilité du processus ; 9](#_Toc104720656)

[Introduction 9](#_Toc104720657)

[5.1. Définition de besoin 10](#_Toc104720658)

[5.1.1. Données de marché 10](#_Toc104720659)

[5.1.1.1. Cours des sous-jacents 10](#_Toc104720660)

[5.1.1.2. Dividendes 10](#_Toc104720661)

[5.1.1.3. Taux de change 10](#_Toc104720662)

[5.1.2. Données internes sur les options détenues 10](#_Toc104720663)

[5.1.2.1. Contrepartie 10](#_Toc104720664)

[5.1.2.2. Contrat 10](#_Toc104720665)

[5.1.3. Détermination du point historiquement optimale 10](#_Toc104720666)

[5.1.3.1. Programmation dynamique 10](#_Toc104720667)

[5.1.3.2. Algorithme de recherche 10](#_Toc104720668)

[5.2. Etude de données 10](#_Toc104720669)

[5.2.1. Collecte et description de données 10](#_Toc104720670)

[5.2.1.1. Définition des sources et accès aux données 10](#_Toc104720671)

[5.2.1.2. Conversion vers une forme structuré 10](#_Toc104720672)

[5.2.2. Analyse exploratoire des données 10](#_Toc104720673)

[5.2.2.1. Etude statistique 10](#_Toc104720674)

[5.2.2.2. Regroupement et détection d’anomalie 10](#_Toc104720675)

[5.3. Préparation et transformation 10](#_Toc104720676)

[5.3.1. Définition du problème de prédiction 10](#_Toc104720677)

[5.3.1.1. Définition de la variable cible 10](#_Toc104720678)

[5.3.1.2. Formulation du problème d’apprentissage supervisé 10](#_Toc104720679)

[5.3.2. Transformations non linéaires 10](#_Toc104720680)

[5.3.3. Normalisation 10](#_Toc104720681)

[5.3.4. Régularisation 10](#_Toc104720682)

[5.3.5. Feature Engeneering 10](#_Toc104720683)

[5.4. Modélisation 11](#_Toc104720684)

[5.4.1. Construction du Pipeline 11](#_Toc104720685)

[5.4.2. Problème d’évaluation 11](#_Toc104720686)

[5.4.2.1. Choix de métrique d’évaluation 11](#_Toc104720687)

[5.4.3. Approche d’optimisation 11](#_Toc104720688)

[5.4.3.1. Composition optimale du pipeline 11](#_Toc104720689)

[5.4.3.2. Optimisation des hyperparamètres 11](#_Toc104720690)

[5.4.3.3. Réduction de complexité du problème d’optimisation 11](#_Toc104720691)

[5.4.4. Feature Importance 11](#_Toc104720692)

[5.4.5. Déploiement 11](#_Toc104720693)

[5.4.5.1. Intégration du modèle dans le processus 11](#_Toc104720694)

[5.4.5.2. Surveillance continue 11](#_Toc104720695)

[Conclusion 12](#_Toc104720696)

[Conclusion Générale 13](#_Toc104720697)

[Annexe A : Scripts python 15](#_Toc104720698)

[Annexe B : Modèles BPMN 15](#_Toc104720699)

[RÉFÉRENCES 15](#_Toc104720700)

[Bibliographie : 15](#_Toc104720701)

[Webographie : 15](#_Toc104720702)

# Introduction générale

# Contexte générale

## Introduction

## Problématique et Présentation du projet

### Contexte du Projet

### Problématique

### Planification du projet

### Planification du projet



## Présentation de l’entreprise d’accueil SG-ATS

### Group Société Générale

### Société générale Corporate Investment Banking

### Société générale Africa Technologie & Sercices

### L’équipe Trading Process Factory

## Conclusion

# Cadre théorique : Définitions et littérature

## Introduction

## Généralités sur les options

* + 1. Définition
    2. Cas d’options américaines

#### Résultat à l’échéance

#### Défi principale

#### Contexte historique

### Modèles de valorisation

#### Formulation du problème

#### EDP

#### Modèle binomiale

#### Critiques et limitation

### Problème du fugit

#### Définition

#### Adaptation des approches de valorisation



## Apprentissage Automatique

### Concept et théorie

#### Types d’apprentissage

#### Technique et évolution

### Problème d’apprentissage

#### Overfitting - Underfitting

#### Solution : Régularisation

### Prétraitement

#### Transformation

#### Gestion des valeurs manquantes

#### La Régularisation

### Modèles de régression

#### Support vector machine

#### Forêts aléatoires

#### Gradient Boosting et XGBoost

### Feature Importance

#### Approche globale

#### Approche individuelle

## Conclusion

# État de l’art et environnement de développement

## Introuction

## Gestion des processus métiers

### Définition d’un processus métier

### Définition d’un workflow

### Business Model and Notation :

## Environnement de développement

### Outils de gestion et collaboration

#### Jira

#### Kibana & ElasticSearch

#### Github Entreprise :

### Techniques et langage de programmation

#### Python

#### Programmation orienté objet

#### Parallélisme

#### Programmation asynchrone

## Connexions et Transfert de données

### Concept des APIs

### Requête et connexion aux APIs

## Conclusion

# Cas d’un processus métier réel

## Introduction

## Besoin fonctionnel

### Exercice anticipé à l’échelle de production

#### Processus générique et réutilisable

#### Scalabilité

#### Efficience

### Intégration au système interne

#### Conception du processus parent

#### Exercice effective et booking de l’opération

## Modélisation et automatisation

### Modélisation en BPMN

### Preuve de concept

#### Etude de faisabilité technique

#### Construction d’une application interactive

### Automatisation du processus

#### Processus parent

#### Processus fils

## Collecte efficace et fiable de données

### Identification des sources de données

#### Eliot

#### MSD

#### Bloomberg

#### Shiva

#### RDWS

### Connexion aux APIs et récupération de données

#### PyBaseRequest

#### Parallélisme et Passage à l’échelle

# Analyse et amélioration de l’efficacité et la fiabilité du processus ;

## Introduction

## Définition de besoin

### Données de marché

#### Cours des sous-jacents

#### Dividendes

#### Taux de change

### Données internes sur les options détenues

#### Contrepartie

#### Contrat

### Détermination du point historiquement optimale

#### Programmation dynamique

#### Algorithme de recherche

## Etude de données

### Collecte et description de données

#### Définition des sources et accès aux données

#### Conversion vers une forme structuré

### Analyse exploratoire des données

#### Etude statistique

#### Regroupement et détection d’anomalie

## Préparation et transformation

### Définition du problème de prédiction

#### Définition de la variable cible

#### Formulation du problème d’apprentissage supervisé

### Transformations non linéaires

### Normalisation

### Régularisation

### Feature Engeneering

## Modélisation

### Construction du Pipeline

### Problème d’évaluation

#### Choix de métrique d’évaluation

### Approche d’optimisation

#### Composition optimale du pipeline

#### Optimisation des hyperparamètres

#### Réduction de complexité du problème d’optimisation

### Feature Importance

### Déploiement

#### Intégration du modèle dans le processus

#### Surveillance continue

## Conclusion

# Conclusion Générale

# Annexe A : Scripts python

# Annexe B : Modèles BPMN

# RÉFÉRENCES

## Bibliographie :

## Webographie :