CAHIER DES CHARGES FONCTIONNEL

- PROJET DE PIPELINE DE CYBERSECURITE -

SOMMAIRE

- Identification des membres de l’équipe du projet
- Introduction et présentation du projet
- Exigences métier détaillées
- Description des données à utiliser
- Exigences métier détaillées
- Description des données à utiliser
- Liste des fonctionnalités principales
- Critères de succès spécifiques
- Contraintes temporelles et budgétaires
- Scénarios d'utilisation
- Diagramme de Gantt – Calendrier
- Impact mapping
- Terminologie
- Charte graphique

Identification des membres de l'Équipe du Projet

La Maitrise d'œuvre (MOE)

L'équipe du projet est composée des membres suivants avec leurs rôles respectifs :

Jérôme FREGEFON

Data Scientist

Responsable du développement des modèles et de l'évaluation des performances.

Serge LACOMBE

Data Protection Officier

Responsable d'assurer la conformité aux RGPD et de superviser les mesures de sécurité des données

Pauline PERCEROU

Data Scientist & Développeur Full Stack

En charge de la collecte, du nettoyage et de l'intégration des données et responsable du déploiement et de la gestion de l'infrastructure cloud.

Le Maitre d'ouvrage (MOA) (Product Owner)

Le client demandeur est la Société M2i – KPLR, sis 15 bis All. James Watt, 33700 Mérignac, représentée par son directeur Mehdi Lamrani

Introduction et Présentation du Projet

Le présent dossier fonctionnel de charges détaille les spécifications et les exigences pour le projet de mise en place d'un pipeline de cyber-sécurité pour améliorer la détection et l'identification des menaces.

L'objectif principal est d'établir un processus automatisé pour surveiller, détecter et répondre aux menaces potentielles à la sécurité des systèmes d'information de l'entreprise. Le présent projet repose essentiellement sur l'exploitation du Dataset «L'UNSW-NB15 »

Un autre objectif est de diminuer les risques d'atteinte et développer un modèle ou des modèles de prédiction des violations du système d'information, visant à optimiser la réaction aux intrusions & réduire les atteintes à l'activité des usagers du système informatique de l'entreprise.

Enfin, un autre objectif est, dans un Contexte plus large de veille et d'évaluation de la hausse des menaces en Cyber-Sécurité basé sur des statistiques globales à l'échelle mondiale.

Exigences Métier Détaillées

Les exigences métier clés pour le pipeline de cybersécurité incluent :

- Amélioration de la détection proactive des cybermenaces.
- Réduction des temps de réponse en cas d'incident.
- Renforcement de la posture globale de cyber-sécurité de l'entreprise.

Votre entreprise souhaite que le système de prédiction des menaces améliore la continuité de l'activité de l'entreprise, en réduisant les atteintes, et augmentant la satisfaction des salariés grâce à une **disponibilité accrue des ressources informatique.**

Description des Données à Utiliser

Les données pertinentes pour le pipeline de cybersécurité comprennent :

- DataSet
- Journaux d'événements du réseau.
- Alertes de sécurité provenant de systèmes de détection d'intrusion.
- Données sur les vulnérabilités connues.
- Informations sur les activités anormales des utilisateurs.

À Propos du Dataset UNSW-NB15 brut d'origine

Les paquets réseau bruts de l'ensemble de données (Dataset) UNSW-NB 15 L'UNSW-NB15 se compose de matières premières des paquets de réseau qui ont été générés par un outil appelé IXIA PerfectStorm dans le Cyber Range Lab du Centre Australien pour la Cyber-sécurité (ACCS). Il contient un hybridemoderne : d'activités normales modernes réelles,et de comportements d'attaque synthétiques contemporains.L'outil Tcpdump est utilisé pour capturer 100 Go de trafic brut (par exemple, fichiers Pcap).

Cet ensemble de données comporte 9 types d'attaques, à savoir :
Fuzzers, l'Analyse(Analysis), Backdoors(Portes dérobées), de DoS, d'Exploits, de Génériques, de Reconnaissance, de Shell code et les Vers (Worms).

Les outils Argus et Bro-ID ont été utilisés, et les 12 algorithmes ont été développés pour générer un total de 49 fonctionnalités avec une étiquette de classe.

Détail des Fichiers d'origine du Dataset

Ces fonctionnalités sont décrites dans le fichier UNSW-NB15_freatures.csv.

Le jeu de données à un total de 2,540,044 registres informatisés (enregistrements), conservés dans les 4 fichiers CSV.
4 fichiers CSV, à savoir UNSW-NB15_1.csv, UNSW-NB15_1.csv,UNSW-NB15_1.csv et UNSW-NB15_1.csv.

De plus, la table de vérité terrain est nommé UNSW-NB15_GT.csv, et la liste des fichiers d'événements est appelé UNSW-NB15_LIST_EVENTS.csv.
Une partition de cet ensemble de données est configurée comme un ensemble d'entraînement et un ensemble de tests, à savoirrespectivement : UNSW_NB15_training-set.csv et UNSW_NB15_testing-set.
Le nombre d'enregistrements dans l'ensemble de formation sont de 175 341 enregistrements et l'ensemble de test est de 82 332 enregistrements provenant de différents les types d'attaque et normal. »
Les données à utiliser comprennent les historiques de connexions, les données sur les _, les Date-Time.

Le jeu de données a été utilisée dans divers documents de recherche pour la détection d'intrusion réseau, criminalistique, la vie privée-la préservation et à la menace approches d'intelligence dans les différents systèmes, tels que les Systèmes de Réseau, de l'Internet des objets (Ido), SCADA, Industriel de l'Ido, et de l'Industrie 4.0. Les auteurs de la base de données ont accordé l'utilisation gratuite de la base de données de recherche universitaire, tandis que l'utilisation commerciale nécessite de leur approbation.

NB : Il faut prendre en compte que ces données peuvent présenter des valeurs manquantes, lacunaires, et des incohérences occasionnelles.

Liste des Fonctionnalités Principales

Le pipeline de cybersécurité doit inclure les fonctionnalités suivantes :

- Collecte de Données : Extraction et agrégation des données de sécurité à partir de diverses sources.
- Analyse et Corrélation : Identification de schémas et de comportements malveillants en utilisant des règles et des modèles d'analyse.
- Modèle d'apprentissage supervisé (machine learning et deep learning)
- Détection d'Intrusion : Identification en temps réel des activités suspectes et des comportements non autorisés.
- Réponse Automatisée : Activation de mesures de sécurité automatisées en réponse aux menaces identifiées.

- Génération de Rapports : Création de rapports détaillés sur les incidents et les tendances de sécurité.
- Maintenance et mise à jour

Critères de Succès Spécifiques

Les critères de succès du projet sont les suivants :

- Réduction du temps moyen de détection des menaces de 30%.
- Automatisation de la réponse aux incidents dans 90% des cas.
- Amélioration de l'indice de maturité en cybersécurité de l'entreprise de deux points.

Contraintes Temporelles et Budgétaires

Le projet devra être achevé en neuf mois à compter de la date de démarrage. Le budget alloué pour le projet est de 150 000 €.

Scénarios d'Utilisation

Les analystes de sécurité pourront :

- Accéder à l'interface du pipeline pour surveiller les alertes et les événements de sécurité.
- Recevoir des notifications en temps réel sur les menaces détectées.
- Activer des réponses automatisées aux incidents à haut risque.

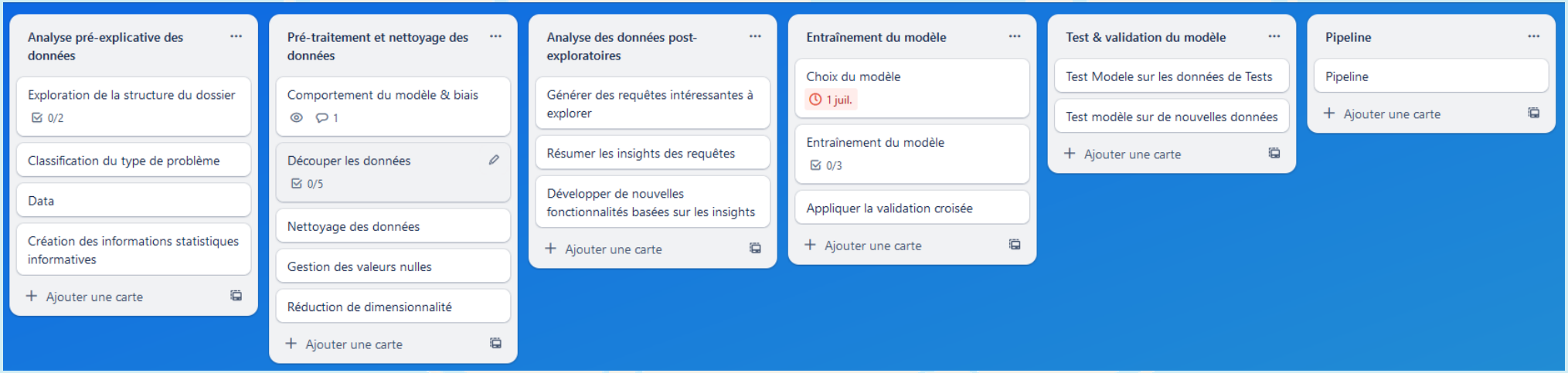


Diagramme de Gantt (Calendrier)

Impact Mapping

Objectif Principal : Mettre en place un pipeline de cybersécurité automatisé pour surveiller, détecter et répondre aux menaces potentielles à la sécurité des systèmes d'information de l'entreprise.

- Équipes de Sécurité et de développement
  - Réduction du temps de détection des menaces.
  - Amélioration de la visibilité sur les activités malveillantes.
  - Renforcement de la posture de sécurité globale de l'entreprise.
  - Réduction du risque d'attaques réussies.

    Déploiement d'un SIEM pour surveiller les activités en temps réel.
    Configuration de règles de corrélation d'événements pour identifier les modèles suspects.
    Mise en œuvre de mécanismes d'alerte pour les activités anormales.
    Intégration d'outils de threat intelligence pour une veille sur les menaces émergentes.

- Équipes de Sécurité et de développement
  - Intégration des mesures de sécurité dans le cycle de développement.
  - Collaboration plus étroite avec les équipes de sécurité pour résoudre les vulnérabilités.
  - Amélioration de la réactivité aux problèmes de sécurité identifiés.

    Intégration de tests de sécurité automatisés dans les processus de CI/CD.
    Mise en place de flux de travail pour signaler et suivre les vulnérabilités identifiées.
    Automatisation des déploiements de correctifs de sécurité.

- Utilisateurs Finaux
  - Utilisation des systèmes et des applications en toute confiance.
  - Minimisation des perturbations causées par des attaques ou des violations.
  - Architectes de données
  - Administrateurs de base de données

    Communication proactive sur les mesures de sécurité mises en place pour rassurer les utilisateurs.
    Mise en place de mécanismes de sauvegarde pour minimiser la perte de données en cas d'incident.

Formation des utilisateurs sur la détection des attaques de phishing et les meilleures pratiques en matière de sécurité.

Terminologie

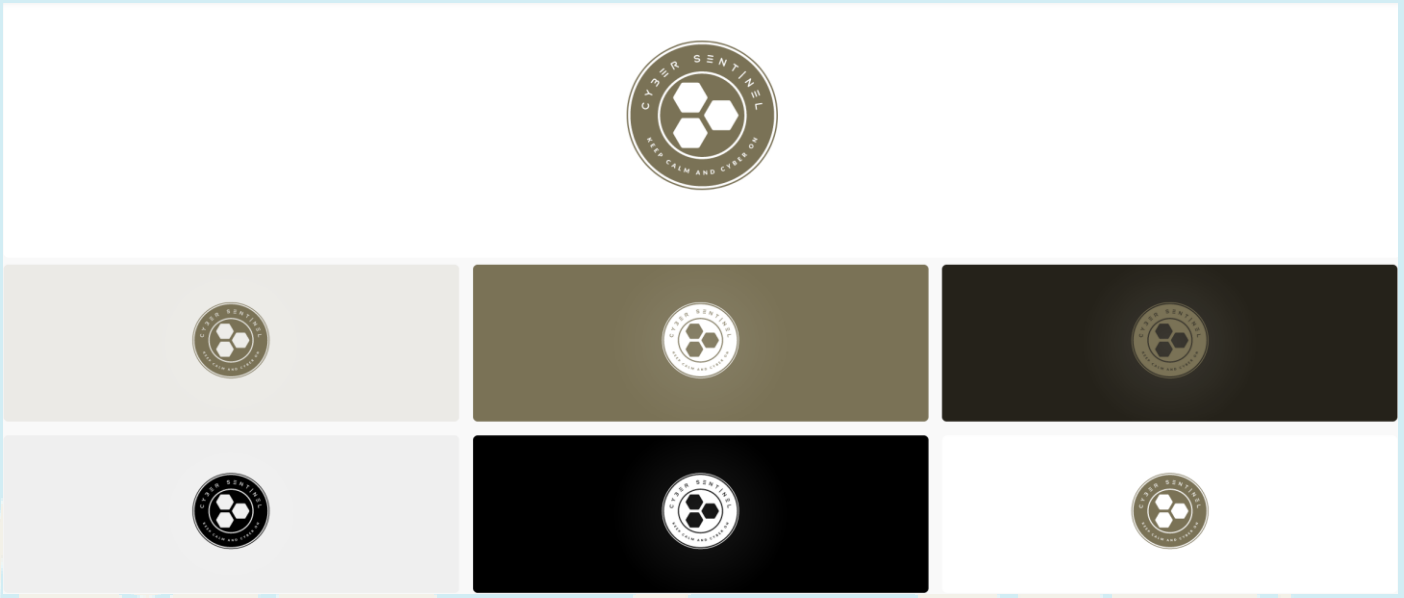
Voici quelques termes et jargon technique couramment utilisés dans le contexte d'un pipeline de cybersécurité :

- **SIEM (Security Information and Event Management) :** Plateforme de gestion des informations et des événements de sécurité qui collecte, analyse et corrèle les données de sécurité à partir de diverses sources.
- **IDS/IPS (Intrusion Detection System / Intrusion Prevention System) :** Système de détection d'intrusion et système de prévention d'intrusion qui surveillent et répondent aux activités suspectes ou non autorisées.
- **Threat Intelligence :** Informations sur les menaces provenant de sources externes pour identifier les nouvelles attaques et vulnérabilités.
- **Vulnerability Assessment :** Évaluation des vulnérabilités pour identifier les faiblesses potentielles dans les systèmes.
- **Penetration Testing (Pen Testing) :** Test de pénétration pour évaluer la sécurité en simulant des attaques réelles.
- **Zero-Day Exploit :** Une vulnérabilité de sécurité non corrigée et exploitée par des attaquants avant qu'une solution ne soit disponible.
- **SOC (Security Operations Center) :** Centre de surveillance de la sécurité qui surveille et répond aux incidents de sécurité en temps réel.
- **Incident Response :** Processus de réponse aux incidents pour gérer et atténuer les impacts d'une violation de sécurité.
- **Firewall :** Dispositif de sécurité réseau qui surveille et contrôle le trafic entrant et sortant.
- **Endpoint Protection :** Logiciels de sécurité pour protéger les appareils individuels (endpoints) contre les menaces.
- **SIEM Correlation Rules :** Règles de corrélation utilisées par un SIEM pour identifier les modèles d'activité suspecte.
- **Threat Hunting :** Recherche proactive d'activités malveillantes à l'aide de données de sécurité.
- **Phishing :** Technique d'attaque où les attaquants tentent d'obtenir des informations sensibles en se faisant passer pour une source légitime.
- **Ransomware :** Logiciel malveillant qui chiffre les données et demande une rançon pour leur restitution.
- **Patch Management :** Gestion des correctifs logiciels pour maintenir les systèmes à jour et sécurisés.
- **Multi-Factor Authentication (MFA) :** Méthode de sécurité qui exige plusieurs formes d'authentification pour accéder à un système.
- **Encryption :** Chiffrement des données pour les rendre illisibles sans la clé appropriée.
- **Security Audit :** Évaluation périodique des systèmes et des processus de sécurité pour identifier les vulnérabilités et les faiblesses.
- **Blacklist / Whitelist :** Liste d'adresses IP, de domaines ou d'applications interdits (blacklist) ou autorisés (whitelist).

- **SOC Analyst :** Analyste en sécurité opérationnelle chargé de surveiller et d'analyser les activités de sécurité.

Ces termes et expressions techniques sont couramment utilisés dans le domaine de la cybersécurité pour décrire les concepts, les technologies et les pratiques liés à la protection des systèmes d'information.

Charte Graphique



Couleurs

Blanc #FFFFFF	Doré #7a7256

Typographie

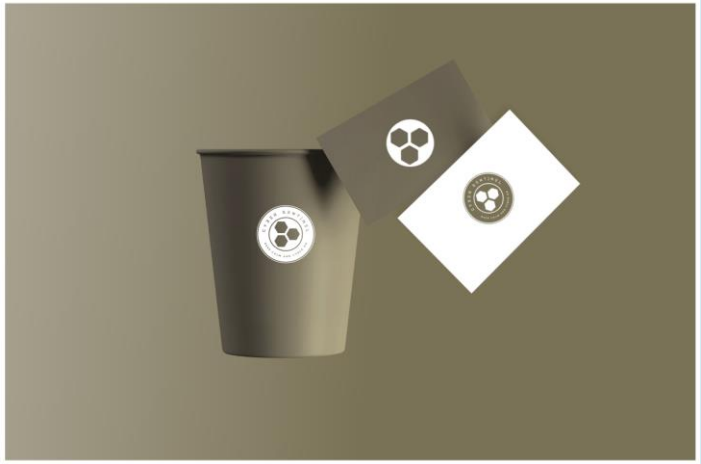
ABCDEFGH

Les polices sans serif sont à la fois minimales et attrayantes. Les marques qui les utilisent souhaitent afficher une attitude simple et sans arrière-pensée. Ce caractère ne comporte aucun élément décoratif qui distrait l'œil et il véhicule des sentiments de simplicité, de confiance et d'innovation. Son aspect propre et moderne en fait une police polyvalente.

ABCDEFGH

Syabil ExtraBold

La typographie moderne existe depuis le XVIIIe siècle. Les polices modernes sont conçues pour être simples et progressives. Alliant praticité et gaieté, elles présentent un fort contraste entre les traits épais et les traits fins. Les polices modernes sont les plus adaptées pour transmettre des sentiments d'exclusivité, de style et d'innovation et ont donc été utilisées par de nombreuses marques très populaires.



Conclusion

Ce dossier fonctionnel de charges détaille les spécifications et les exigences pour la mise en place d'un pipeline de cybersécurité visant à améliorer la détection, la réponse et la prévention des menaces pour l'entreprise.

Le respect des critères de succès, des contraintes budgétaires et temporelles ainsi que la convivialité des scénarios d'utilisation seront les pierres angulaires du succès de ce projet.