

PORTFOLIO **SOLUTIONS** & **USE-CASES**



Data has a better idea

SOLUTIONS AQUILA



ANALYSEZ ET INTERROGEZ
VOTRE BASE DOCUMENTAIRE
EN LANGAGE NATUREL



NOTRE OBJECTIF

**VOUS PERMETTRE D'ANALYSER & D'INTERROGER
VOTRE BASE DOCUMENTAIRE EN LANGAGE NATUREL
GRÂCE AU #NLP ET A LA #COMPUTER VISION**

DETECTION DE ZONE & EXTRACTION D'INFORMATION

COMMENT ?

FONCTIONNE SUR PLUSIEURS TYPES DE DOCUMENTS

- Scans (png, jpeg, etc.)
- PDF
- Word

EXTRAIT DE L'INFORMATION DE DIFFERENTES SOURCES

- Textes
- Tableaux
- image

IDENTIFIE LA STRUCTURE DU DOCUMENT

- Titres
- Sections
- Paragraphes

ASSOCIE DES TAGS A CERTAINES PARTIES DU TEXTE

- Structure de la phrase (nom, adjectif, etc.)
- Informations contextuelles (montant, date, durée, noms propres)
- Informations métiers (règles du client ou ensemble d'exemples)

NOTRE PROPOSITION



Les entreprises ont toutes des mines d'informations dormantes dans leurs GED ou bases documentaires, c'est pourquoi Aquila propose un outil complet et clé en mains qui permet d'extraire l'information et d'y avoir accès.

POUR REpondre A VOS BESOINS



ANALYSE

Analyse des documents
Identification des zones clés



EXTRACTION

Extraction du texte brut
Extraction des éléments du tableaux



TAGGING

Tagging Intelligent du texte par analyse NLP



BASE

Création d'une base documentaire
Entraînement du modèle de Q&A



Q&A

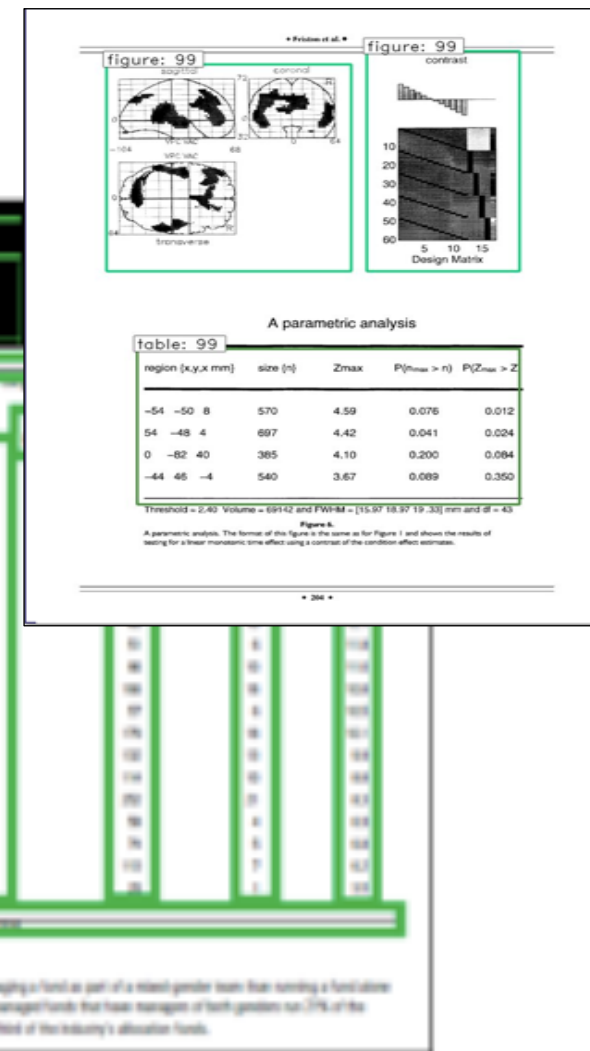
Mise à disposition de l'outil de Q&A

OBJECTIFS

- Détection des zones (textes, tableaux, images, figures) sur les documents
- Utilisation du deep learning pour le repérage des tableaux
- Analyse sémantiques des tableaux pour fusion de données
- Amélioration de la détection de colonnes dans les tableaux complexes
- Extraction d'information des graphiques

ALGORITHMES

- CNN
- NLP



EXTRACTION D'INFORMATIONS DES DOCUMENTS ET CREATION D'UN SYSTÈME DE Q&A POUR LA RECHERCHE D'INFORMATIONS

OBJECTIFS

- Numérisation et stockage intelligent d'un ensemble documentaire varié (document word, excel, scan)
- Extraction des information par zones (texte, tableaux,...)
- Création et entraînement d'un modèle spécifique de Q&A pour la recherche d'information

ALGORITHMES

- CNN
- BERT
- NLP

Question: Hello bank is only a digital bank?
Answer: digital direct bank owned by bnp paribas

Hello bank! is a **digital direct bank owned by BNP Paribas** that started operations in 2013. The bank operates in France, Belgium, Germany (using the name Consorsbank), Italy, the Czech Republic and Austria.[1] BNP Paribas has claimed that it is "the first 100% digital mobile bank in Europe". The bank is supported through different BNP Paribas retail banking subsidiaries where they exist; these are: BNP Paribas in France, BNL in Italy, BNP Paribas Fortis in Belgium. BNP Paribas, BNL and BNP Paribas Fortis branches can be used to deposit checks or to withdraw money in these countries.

Question: What is auchan?
Answer: a french multinational retail group

Auchan (French pronunciation: [oʃɑ̃]) is a **French multinational retail group** headquartered in Croix, France.[2] It was founded in 1961 by Gérard Mulliez and is owned by Mulliez Family, which has 95% stake in the company.[3] With 354 851 employees, of which 261 000 have 5% stake in the company, is the 35th

SMART VISION

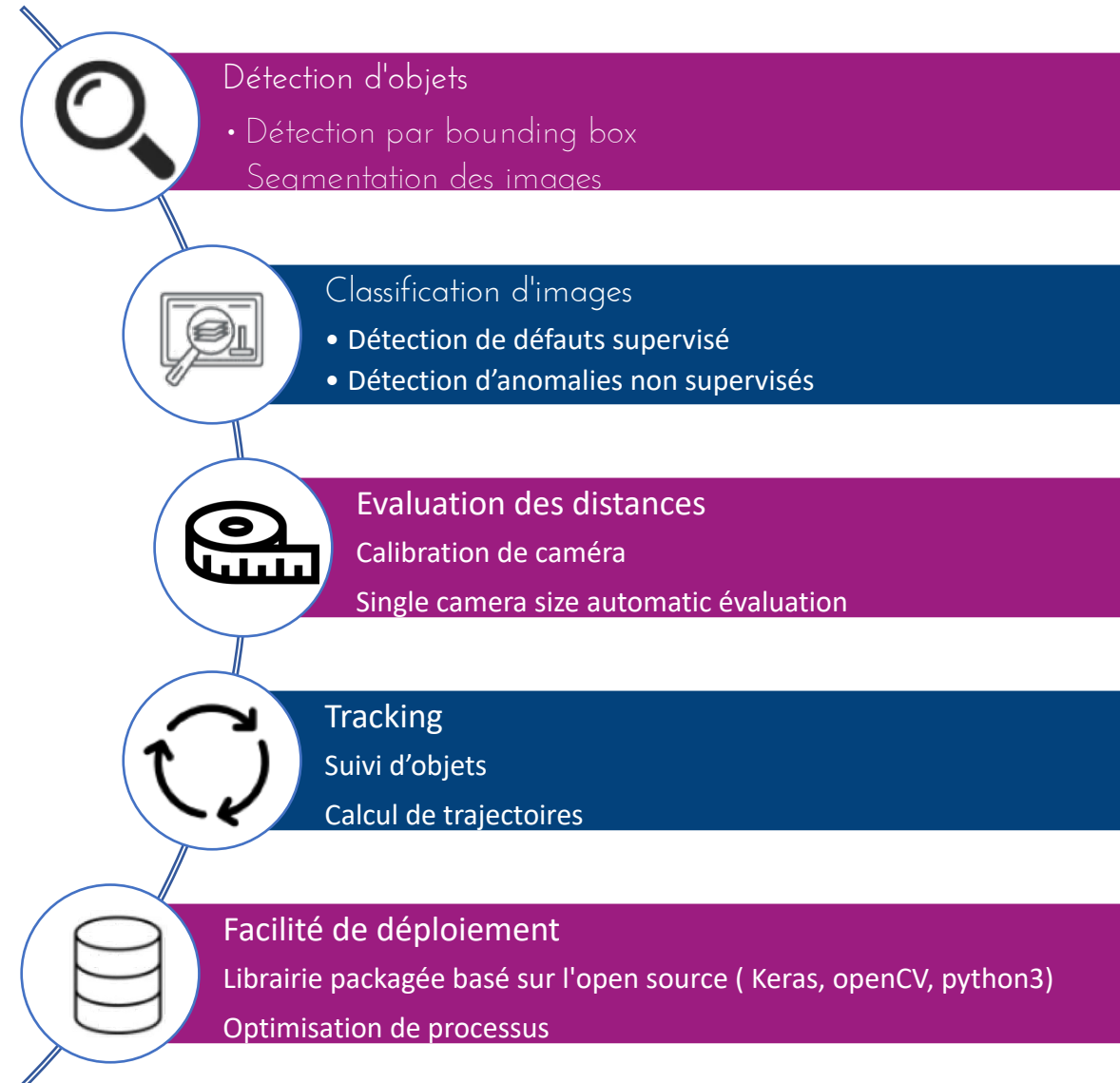
**La computer vision au service
de la qualité et sécurité**

NOTRE PROPOSITION



Les images vidéos sont une mûnes d'informations qui ne sont exploitables que depuis peu. En effet, les outils de deep learning, qui permettent de détecter, suivre et classier des objets ne sont exploiter que depuis quelques années. Smart Vision est une suite algorithmique qui s'appui sur ces algorithmes pour répondre à vos besoins de sécurité et d'audit qualité

POUR REpondre A VOS BESOINS





NOTRE OBJECTIF

VOUS PERMETTRE

D'EXPLOITER VOS VIDEOS POUR LE CONTROLE QUALITE & LA SECURITE

GRÂCE AU #DEEP LEARNING ET A LA #COMPUTER VISION

DETECTION DE ZONE & EXTRACTION D'INFORMATION

COMMENT ?

FONCTIONNE EN TEMPS REEL SUR TOUS TYPES DE FLUX VIDEO

- *Caméra de Sécurité*
- *Flux vidéo de contrôle CND*

ENTRAINEMENT SUR DE NOUVEAUX LABELS

- *Augmentation de données par GAN*
- *Optimisation des stratégies d'entraînements*

STABILISATION DES LABELS PAR TRACKING ET FILTRE DE KALMAN

- *Permet la stabilisation des modèles pour une meilleur exploitabilité des résultats*

OUTILS DE VISUALISATION DES RÉSULTATS

- *Développement d'un outil dédié aux équipes métiers*
- *Création d'outils dédiés*

DETECTION AUTOMATIQUE D'EPI POUR **CONTRÔLE D'ENTREE SUR SITE SENSIBLE**

OBJECTIFS

- Détection des équipements de sécurité et de distanciations (masques, casques ect...)
- Analyse des flux vidéos d'une caméra en temps réel
- Entraînement séparé sur la détection de chaque équipement.

ALGORITHMES

- Labellisation des données d'entraînements
- Construction d'un modèle de Deep Learning
- Utilisation de YOLO v3
- Ré- entraînement du modèle



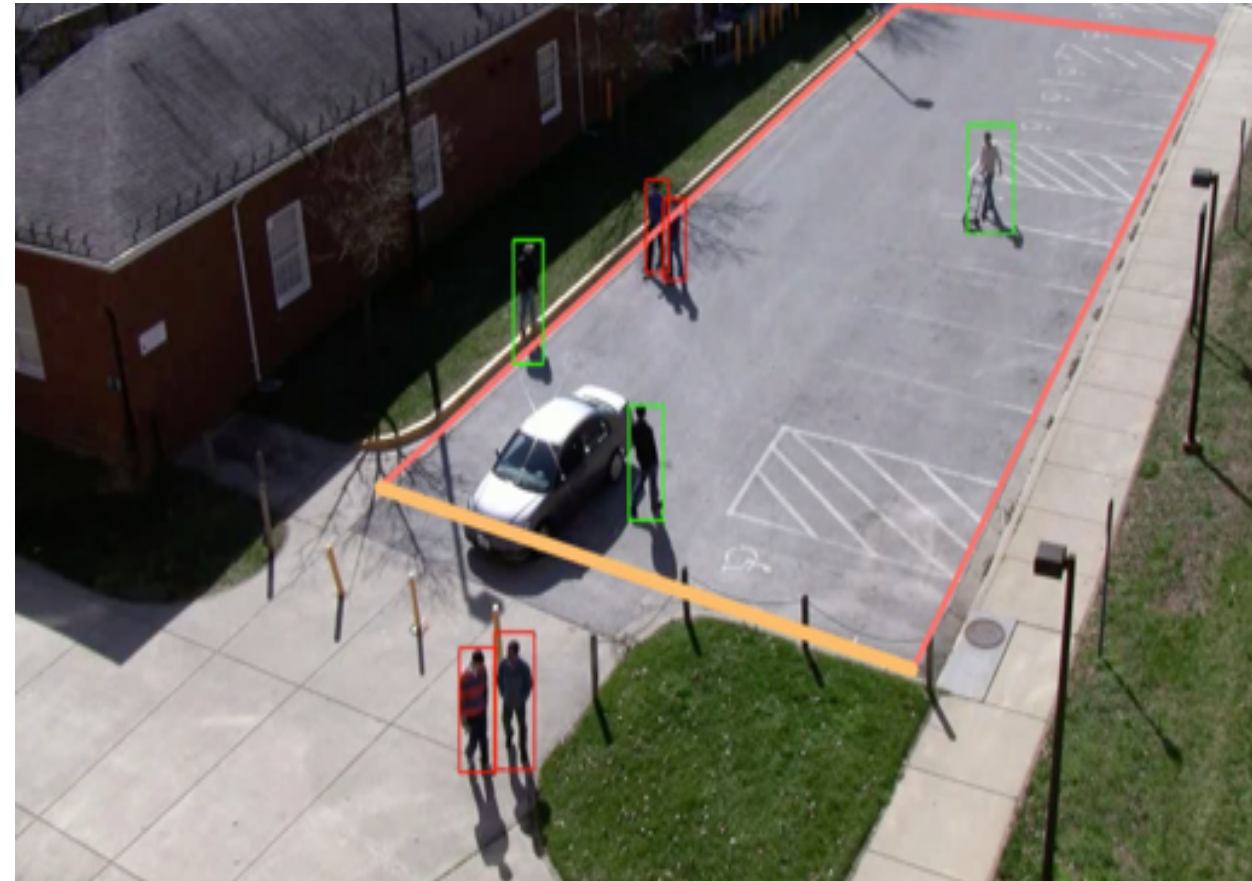
DÉTECTION DES PERSONNES ET DISTANCIATION SOCIALE

OBJECTIFS

- Détection par ordinateur des individus sur des images de caméra de surveillance fixe.
- Grâce à une calibration détermination de la matrice de projection de la caméra (outils de calcul des distances sur l'image de la caméra)
- Développement d'une application visuelle permettant aux de détecter les comportements à risques.

ALGORITHMES

- Calibration des caméras par homothéties
- Utilisation de YOLO v3
- Tracking et suivis des personnes





OUTIL DE PRÉVISION
DE SÉRIES TEMPORELLES



Prévoir la demande, de manière précise
et à des horizons temporels variés, apporte
un **avantage compétitif** aux entreprises qui leur permet :



De **réagir par avance**
aux pics de demandes



Tout **en diminuant les**
risques financiers liés aux
surplus et au stockage



NOTRE OBJECTIF

**VOUS PERMETTRE AU TRAVERS
D'UNE APPLICATION WEB SIMPLE ET FACILE
D'OPTIMISER VOS DONNÉES
& D'UTILISER VOS PRÉDICTIONS**

LES PLUS DE NOTRE OUTIL

*Automatisation
et pré-traitement
des données par
l'intelligence
artificielle*

*Utilisation de
données exogènes
pré-chargées
(Météo, données
calendaires...)*

*Spécialisation des
algorithmes
d'intelligence en
fonction de vos
données*

*Génération
de rapports et de
dashboards
à partir de
vos datas*

*Interface
d'exploration
des résultats
simple et
efficace*

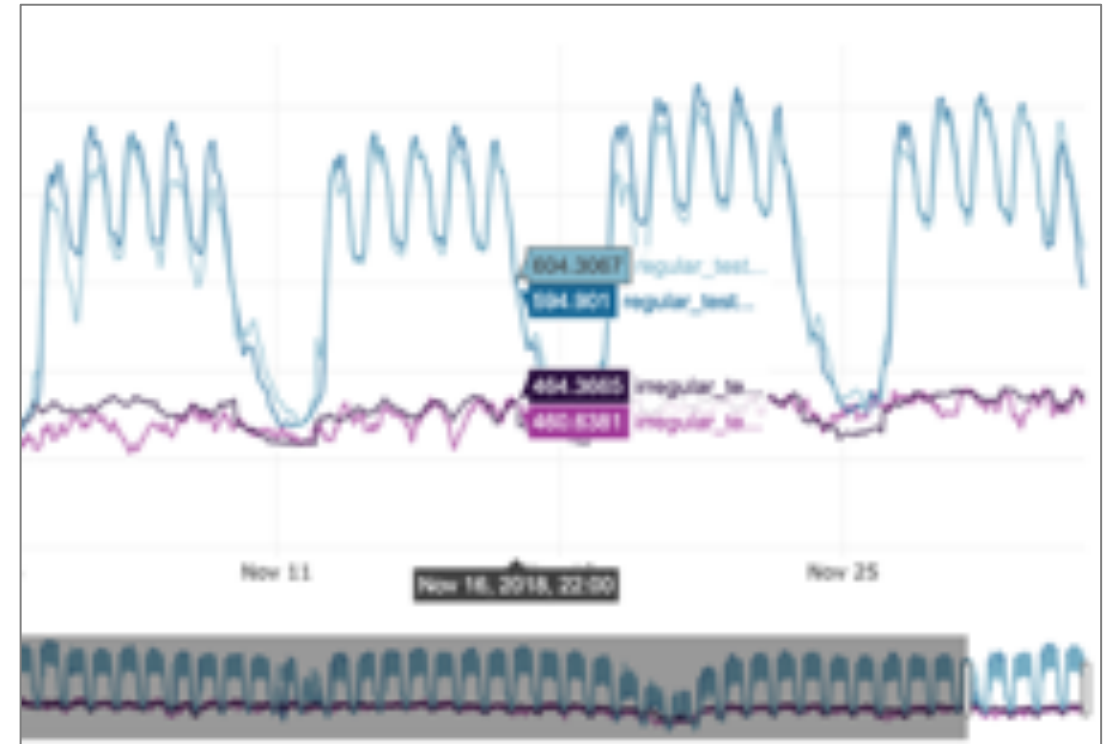
TIME FORECAST: UN OUTIL SIMPLE DE PREVISION

OBJECTIFS

- Disposer d'une application web simple et facile d'utilisation pour vos prédictions :
 - Automatisation du pré-traitement des données
 - Utilisation de données exogènes
 - Génération de rapports et visuels à partir de vos datas
 - Interface d'exploration des résultats

ALGORITHMES

- XGBoost, Random Forest
- Auto-ML



P O R T F O L I O

#1 • COMPUTER VISION

#2 • NLP (NATURAL LANGUAGE PROCESSING)

#3 • DONNÉES STRUCTURÉES

PORTFOLIO

#1 • COMPUTER VISION

#2 • NLP (NATURAL LANGUAGE PROCESSING)

#3 • DONNÉES STRUCTURÉES

DÉTECTION AUTOMATIQUE DE DÉFAUTS SUR ÉOLIENNE

OBJECTIFS :

- Détection par ordinateur des pales d'éoliennes sur des images de drones.
- Identification automatique des défauts potentiels sur les pales.
- Développement d'une application visuelle permettant aux spécialistes métiers de labéliser les défauts et de générer des rapports.
- Développement d'une méthode de labélisation automatique des défauts.

ALGORITHMES :

- Segmentation de l'image par Deep Learning
- Classification et labélisation par CNN des défauts



OBJECTIFS

- Lecture automatique des pièces jointes de mails
- Identification du type de document (RIB, CI, passeport)
- Stockage Automatique de la donnée en base de données
- Détection de documents frauduleux

- Augmentation de Dataset
- Mise en place des méthodes OCR Robustes
- Créations de métriques de filtrages



COMPTAGE DE VOITURE POUR LE SMART CITY

OBJECTIFS

- Permettre le tracking et comptage de véhicules à partir d'un système multi caméras
- Identification des zones de congestions

ALGORITHMES

- Récupération et traitement d'un large flux de données
- Création de l'identifiant unique par CNN
- Calibration des camera par pattern et keypoint matching



ÉTUDE DE RÉ-IDENTIFICATION DE PERSONNES EN BASE OUVERTE

OBJECTIFS

- Permettre l'identification et le tracking d'individus par un système multi-caméra
- Permettre la détection et la classification de nouveaux individus.

ALGORITHMES

- Définition de marqueurs et identification de personnes
- Etude des scores et labélisation



ANALYSE D'IMAGES POUR LA LOCALISATION DE DÉFAUTS SUR PANNEAUX PHOTOVOLTAÏQUES

OBJECTIFS

- Reconstruction du plan masse de l'installation
- Création d'un jeu d'images pour chaque cellule PV
- Détection et classification des défauts
- Mise en place d'un système de maintenance prédictive

ALGORITHMES

- Détection de défauts : blob lumineux
- Détection automatique des rangées et des cellules de PV
- Auto-stitch des images



DÉTECTION DE COLLAGÈNE SUR IMAGES CONFOCALES

OBJECTIFS

- Détection des collagènes sur des images confocales acquises sur le derme de plusieurs sujets.
- Réduction des paramètres avec lesquels les opérateurs interagissent.
- Calcul des métriques de fragmentation.

ALGORITHMES

- Segmentation d'image en superpixels par algorithme SLIC.
- Clustering par réseau de neurones profond (algorithme W-Net).
- Estimation du taux de fragmentation en fonction du périmètre et de la surface moyenne.



#Segmentation #CNN #AlgorithmeNonSupervisé #ImageConfocale

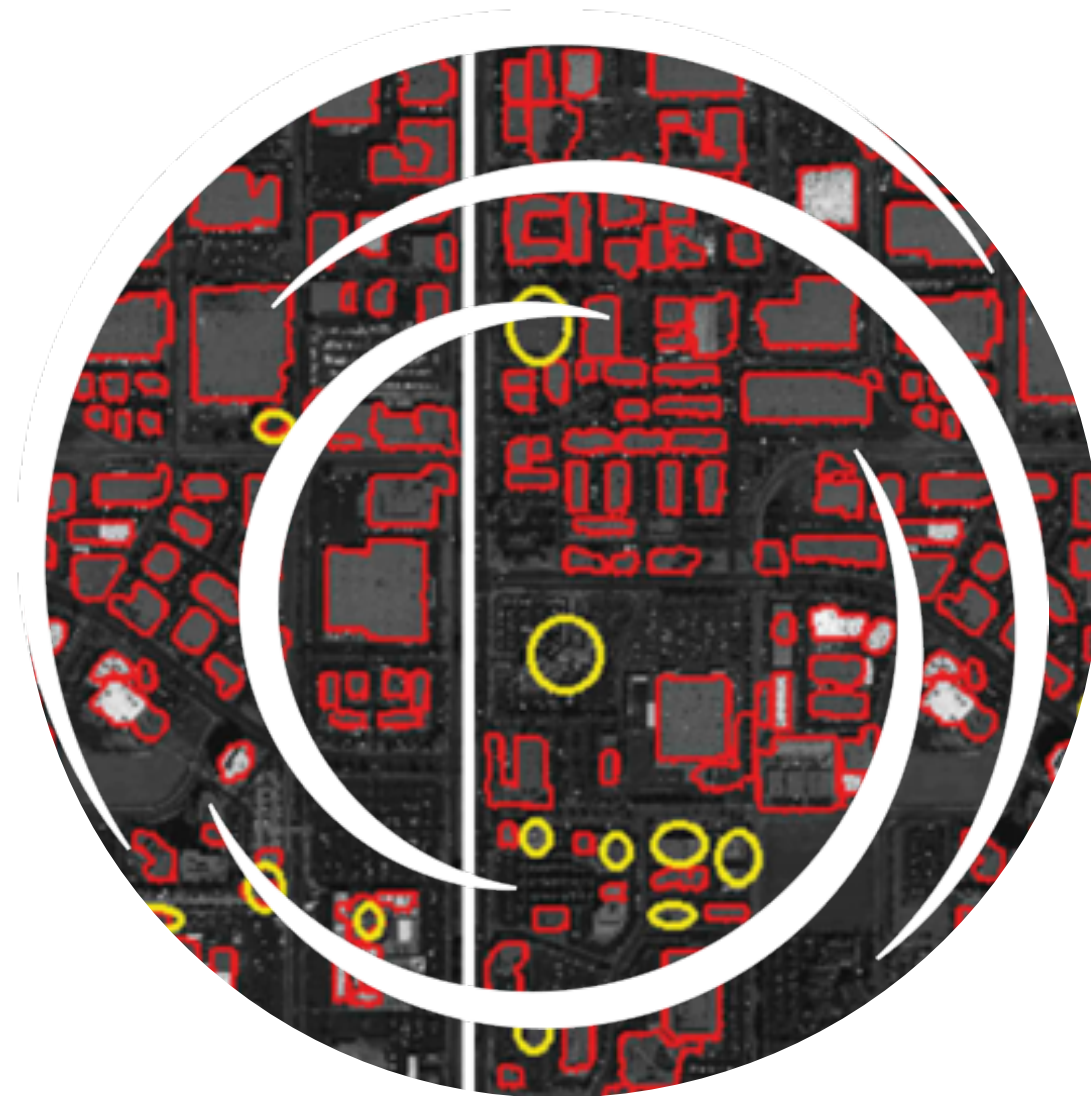
EVALUATION DE LA POPULATION PAR ANALYSE D'IMAGE SATELLITE

OBJECTIFS

- Détection de toitures par analyse d'image satellites
- Projection et forecasting de la densité de population basé sur ces observation et des données exogènes

ALGORITHMES

- Détection et identification de toitures par CNN
- Forecasting de la densité par RF
- Intégration des données Exogènes



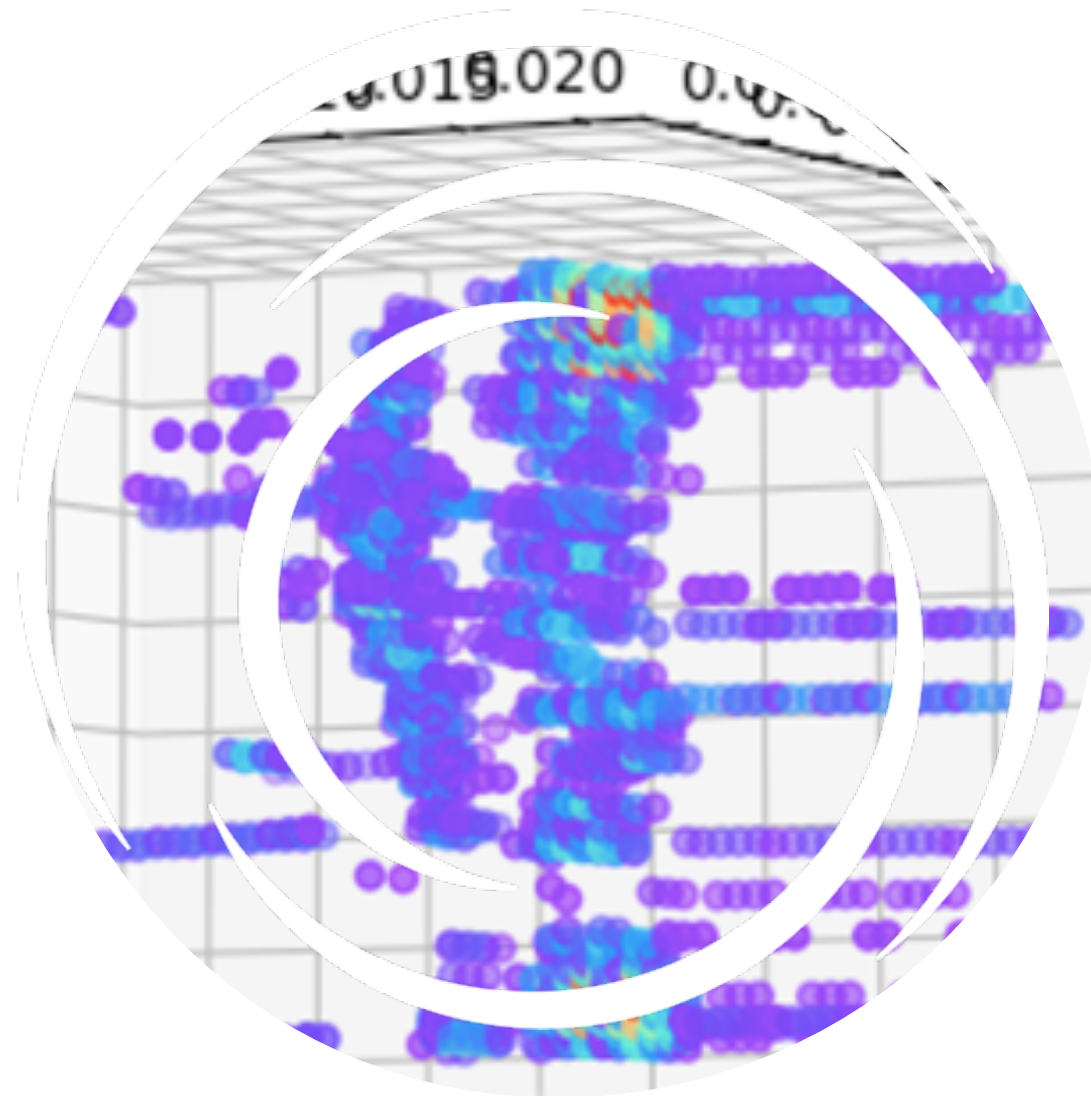
DETECTION DE DEFAUTS SUR ULTRASONS ET RECONSTITUION 3D

OBJECTIFS

- Reconstruction d'un modèle 3D à partir de coupe 2D d'ultrasons
- Détection des défauts sur automatique à partir de règles client

ALGORITHMES

- Analyse et reconstitution 3D
- Filtrage et clustering automatique des points
- Etude des défauts par apprentissage profond



PORTFOLIO

#1 • COMPUTER VISION

#2 • NLP (NATURAL LANGUAGE PROCESSING)

#3 • DONNÉES STRUCTURÉES

ANALYSE DE TEXTE ET SEGMENTATION UTILISATEUR

OBJECTIFS

- Sur un site de presse, identification des profils clients en fonction de leur historique de navigation
- Analyse sémantique du contenu des articles pour améliorer la segmentation client

ALGORITHMES

- Réseaux neuronal convolutif
- Indexation sémantique latente
- Word2Vec



EXTRACTION DE PATTERNS DANS UN CONTRAT :

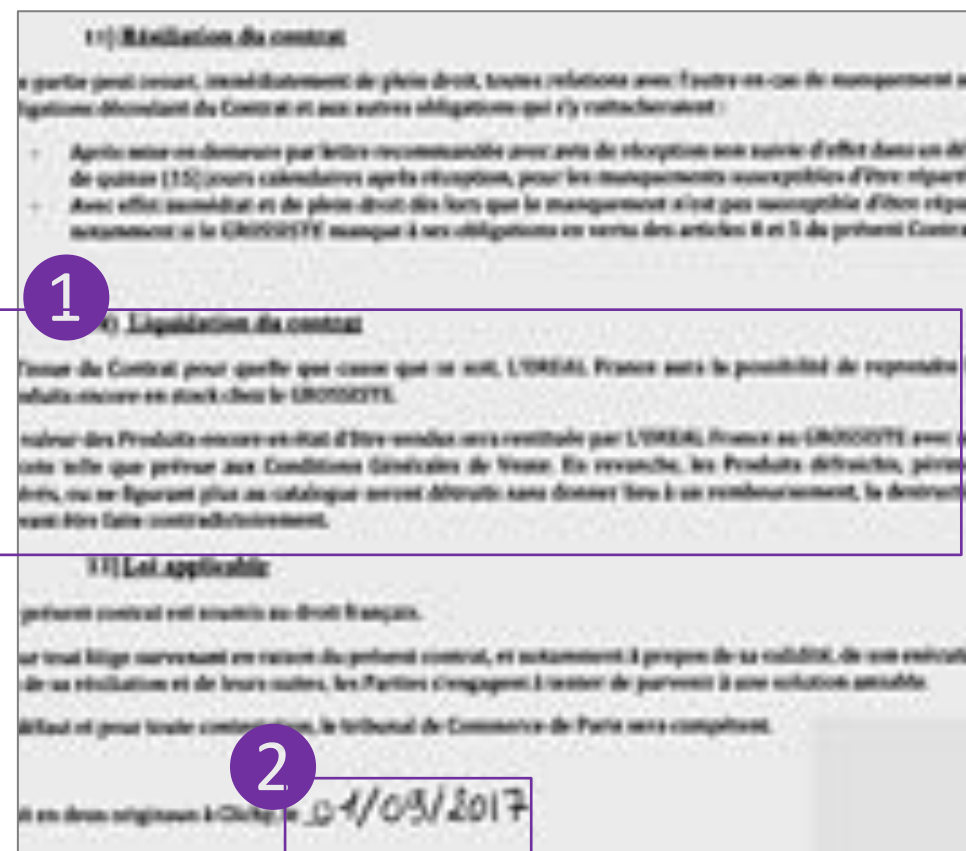
CLAUSES SPECIFIQUES, SIGNATURES ET DATE MANUSCRITE

OBJECTIFS

- Identifier et extraire les clauses particulières de nombreux contrats
- Extraire les signatures
- Vérifier l'authenticité des scripteurs par matching
- Identifier et extraire les dates manuscrites

ALGORITHMES

- Classification par CNN (LeNet-5) et SVM
- Modèle de Markov caché (lexique au format « Date »)
- Validation croisé et multi-OCR.



OCCUPATION DE LA VOIERIE À PARTIR DE L'ANALYSE D'ARRÊTÉS MUNICIPAUX

OBJECTIFS

- Identifier dans une liste d'arrêtés municipaux, les arrêtés ayant une conséquence sur l'espace public (Travaux, Manifestation, Événement Sportif...)
- Détection d'adresse et validation du périmètre impacté
- Identification des dates de début et de fin des travaux au sein de l'arrêtés
- Restitution de ces travaux au travers d'une interface visuelle avec filtres

ALGORITHMES

- Topic Modeling (LDA)
- Text Mining & NLP (TF-IDF, W2V)



PROJET CONTROVERSES

PLATEFORME GDA

OBJECTIFS

Création d'un prototype en python permettant de détecter des controverses potentielles dans des news issues des marchés financiers en langue anglaise (source Thomson Reuters) afin de faciliter le travail d'analyse des gestionnaires d'actifs (GDA)

ALGORITHMES

- Preprocessing des news (tokenisation, suppression des stopwords, lemmatisation)
- Classification supervisée via embedding word2vec pour prédiction d'un sentiment de négativité
- TF-IDF, Latent Semantic Analysis (LSA)
- Création d'indicateurs de suivi des affaires et d'alerte à la controverse (volumétrie, négativité, pertinence, verbatim controverses)

RÉSULTATS

- **Prototype codé en Python livré au client**



MAINTENANCE PREDICTIVE PAR ANALYSE DE LOGS

OBJECTIFS

- Analyser les logs machines afin de prévenir les pannes
- Avoir un score de risques continu et des point de maintenance automatique
- Extraction d'informations structurées dans des logs machine très bruités

ALGORITHMES

- Clustering des logs
- Analyse NLP pour encoder les messages
- Utilisation du Deep learning pour le calcul de score



ÉTUDE AUTOMATIQUE DES TICKETS DE MAINTENANCE

OBJECTIFS

- Classifier les tickets de maintenance
- Augmenter la vitesse et l'efficacité de traitement des tickets
- Analyse du score utilisateur et de l'impact.

ALGORITHMES

- Extraction de features des tickets par NLP
- Classification des risques d'impacts
- Création de scores automatiques



CLASSIFICATION ET LECTURE AUTOMATIQUE DE FACTURES AVEC PERTURBATIONS

OBJECTIFS

- Lecture automatique de factures réparti en + de 20 templates
- Reconnaissance et segmentations de chaque document
- Lecture automatique et robuste des informations
- Score de confiance de lecture pour toutes les informations

ALGORITHMES

- Classification par CNN et K-means avec soustraction de templates
- Lecture robuste par OCR avec score de confiance
- Validation croisée et multi-OCR



#OCR #DeepLearning

PORTFOLIO

#1 • COMPUTER VISION

#2 • NLP (NATURAL LANGUAGE PROCESSING)

#3 • DONNÉES STRUCTURÉES

PRÉVISIONS DE CONSOMMATION INSTANTANÉE DE COMPTEURS LINKY

OBJECTIFS

- Développement d'un modèle de prévision de consommation (J+2) à l'échelle d'une commune ou d'une région.
- Utilisation de données passées et de données exogènes (météo, calendrier...)
- Utilisation de la solution Big Data d'Enedis.

ALGORITHMES

- Approches classiques des séries temporelles : ARIMA, VARX
- Méthodes Deep Learning : LSTM, RNN



PRÉVISIONS DE CONSOMMATION **POUR DES CLIENTS INDUSTRIELS**

OBJECTIFS

- Développement d'un modèle de prévision de consommation (J+1 à J+14) pour des clients industriels
- Utilisation des données Météos et Exogènes
- Mise en place de clustering automatique des comptes clients

ALGORITHMES

- Séries Temporelles
- Modèles météo
- Forecasting



ANALYSE DE PERFORMANCE D'UNE UNITÉ INDUSTRIELLE PAR LA DATA

OBJECTIFS

- Mettre en place un modèle permettant de prédire la performance d'une unité industrielle grâce aux données des unités en amont. (débit, pression, température, données de prélèvement)
- Analyse des paramètres impactant la performance.
- Prévion sur des systèmes complexes à Lags variables.
- Automatisation des rapports et du système d'alerte.

ALGORITHMES

- Analyse des signaux faibles
- Modélisation de phénomène physique
- Modélisation de systèmes industrielles



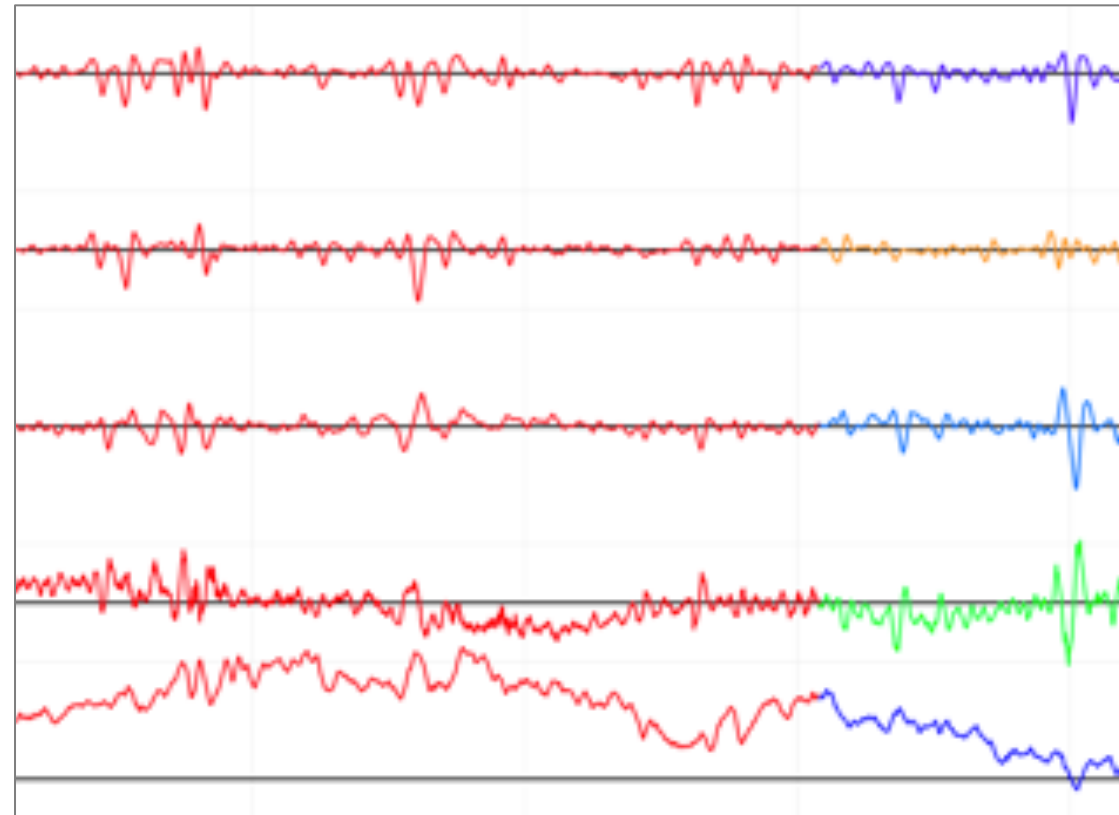
ANALYSE DE VOIES FERROVIAIRES AUTOMATIQUE ET DÉTÉCTION DE DÉFAUTS PAR CAPTEUR LASERS

OBJECTIFS

- Analyse du signal de capteurs laser pour l'analyse de la géométrie de la voie
- Détection de zone de perturbation du signal
- Analyse automatique des appareils de voies
- Gestion de labélisation active par le métier

ALGORITHMES

- Analyse du signal
- Random Forrest
- Active learning



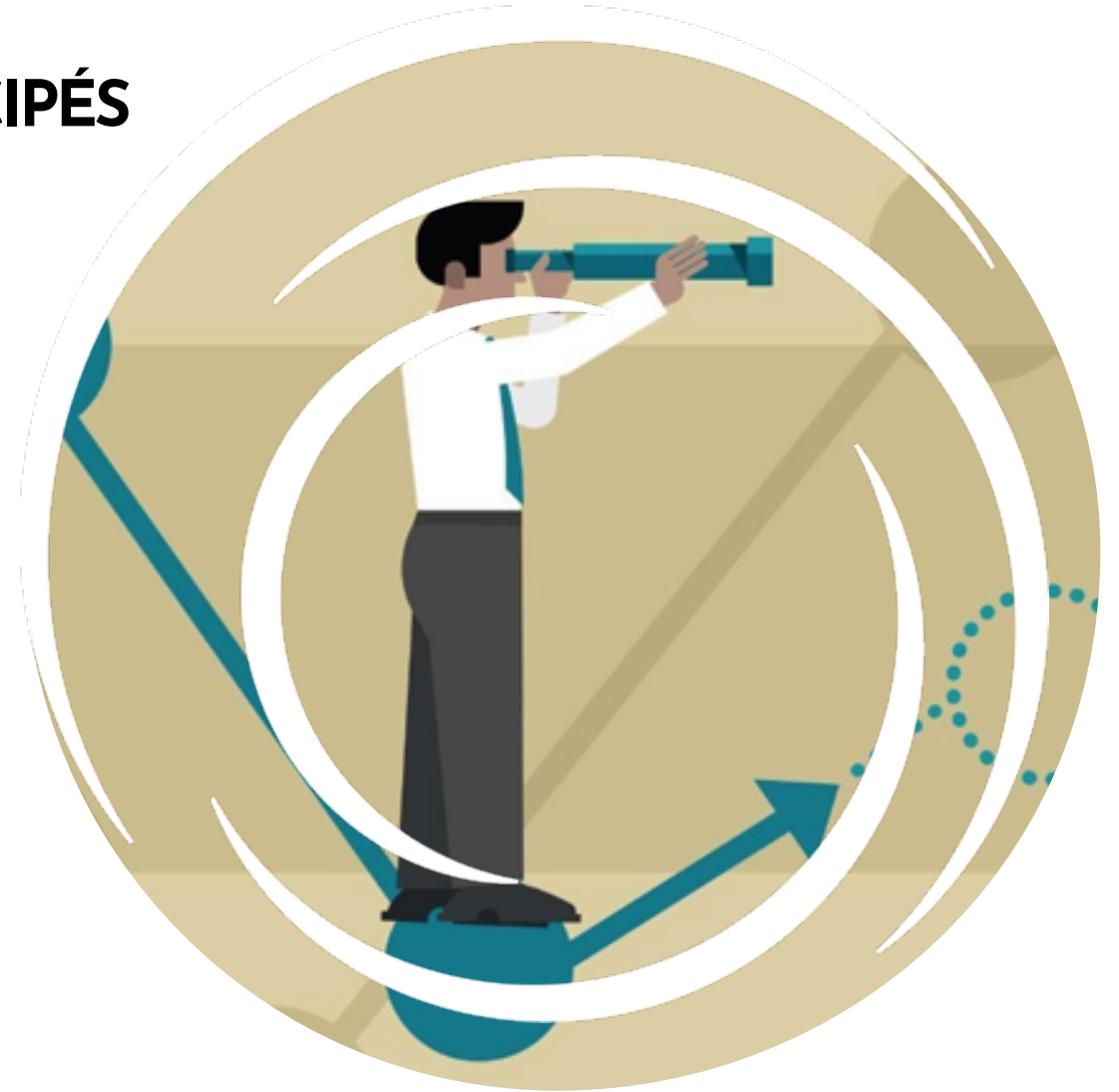
PRÉDICTION DE **REMBOURSEMENTS ANTICIPÉS**

OBJECTIFS

- Comprendre et expliquer les résultats du modèle existant
- Diagnostiquer et établir un rapport sur les méthodologies en place
- Améliorer le modèle de prédiction des remboursements anticipés

ALGORITHMES

- Explication et interprétation des modèles : SHAP, permutation importance
- Modèles : RandomForest, XGBoost ; classes déséquilibrées



CREATION D'UNE PLATEFORME DE TRAITEMENT DE DONNÉES FINANCIERE

OBJECTIFS

- Création d'une plateforme de prédiction à partir de données financières génériques.
- Gestion d'un feature engineering automatique
- Généralisation de la création de target à tout types de données
- Introduction de nouveaux indicateurs et mesure l'impact sur la performance de modèles.

FRAMEWORKS ET ALGORITHMES

- Spark/ SparkSQL, Hivemall.
- LighGBM, scikit-learn, TSFresh.



PREDICTION DU TAUX D'HYDRATATION À PARTIR DE LA FORMULATION DES PRODUITS

OBJECTIFS

- Prédire la taux d'hydratation constaté sur des mesures invivo a partir des composants et des dosages de chaque produit
- Utilisation des familles de produits pour permettre l'ajout d'informations dans le dataset
- Identification des produits impactant l'hydratation

ALGORITHMES

- Segmentation
- Clustering



ETUDE DE L'ÉLASTICITÉ DES PRIX SUR UN PRODUIT DE CONSOMMATION

OBJECTIFS

- Etablir une liste des couples (produit, type de clients) sur lesquels proposer des hausses de prix sans impacter le CA négativement.
- Analyse et recommandation de prix optimal par couple (produit, client)
- Nettoyage et remplissage des base de données

ALGORITHMES

- Réseaux de neurones
- Régression Hédonique



ANALYSE DE COMPORTEMENT UTILISATEUR SUR UN **ENSEMBLE DE SITES INTERNET**

OBJECTIFS

- Audit de la qualité des données en provenance des différents sites du groupe
- Elaboration de plans de reporting et analyses quantitatives et qualitatives des données
- Analyse du comportement de navigation des utilisateurs
- Mise en place de Dashboard de suivis
- Analyse et création des différentes KPI et proposition de stratégies d'amélioration



GÉNÉRATION AUTOMATIQUE DE PLANS D'AMEUBLEMENTS PAR DEEP LEARNING

OBJECTIFS

- Création d'un plan d'ameublement à partir du plan masse respectant des contraintes fonctionnelles et de style
- Génération automatique de la position, taille et orientation des meubles
- Créations de style à la demande

ALGORITHMES

- Neural Network
- Deep Learning
- Adversarial Neural Network



EVANGÉLISATION, WORKSHOP ET ACCOMPAGNEMENT ÉVÈNEMENT INNOVATION

OBJECTIFS

- Evangéliser des usages concrets de la datascience auprès de vos équipes métiers ou de décideurs.
- Organisation de workshop d'idéation avec vos équipes métiers pour identifier les gisements de données et les cas d'applications avec des ROI concrets pour votre entreprise.
- Création et développement de démonstrateur interactif et concret pour démontrer les usages de la datascience

EXPÉRIENCE PRÉCÉDENTES

- Journée Groupama
- Journée DIIP
- DAHER
- SALON AI
- BPI EXPERT



VOUS
SOUHAITEZ
NOUS PARLER
DE **VOTRE**
PROJET ?



Véronique SEBAN
Senior Business Manager

vseban@aquiladata.fr
+33(0)6 83 71 10 18



Antonin BRAUN
Directeur Data Science

abraun@aquiladata.fr
+33(0)6 60 61 93 50



47 rue Louis Blanc | 92 400 Courbevoie •
aquiladata.fr