

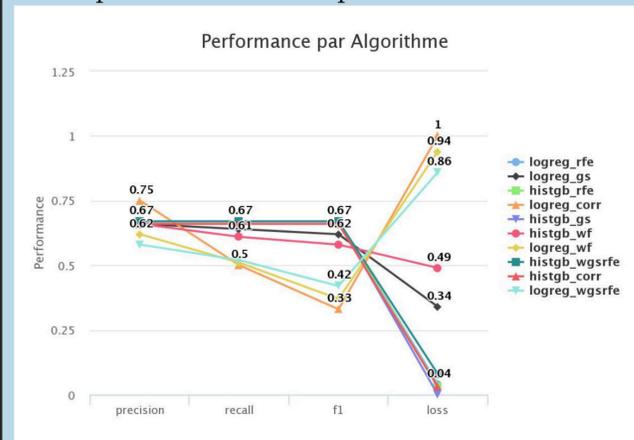
Portfolio des pratiques et des outils d'Intelligence Artificielle



Romain Boyrie
romain@boyrie.email

Réaliser un Scoring

Création d'un modèle de machine-learning pour noter les clients d'un service bancaire. **Exploratory Data Analysis, Features Engineering, Cross Validation, Grid Search**, Métriques métier et comparaison des modèles.



Code d'un Chatbot

Fabrication d'un **pipeline** pour produire et entraîner un chatbot et déploiement dans le web. Ce chatbot est capable de poser les bonnes questions aux clients pour prendre les informations de commandes de billets d'avion. Il possède une interface de monitoring et peut, de plus, avertir d'un incident, mais également afficher toute la communication ayant eu lieu avec le client. L'amélioration est ainsi conçue dans le code.

Améliorez le produit IA

Collecte de données via API REST (YELP) d'avis de consommateurs. Nettoyage du texte, puis **Topic extraction** avec LDA. Réalisation de nuages de mots. Détection des contours d'intérêts dans des images de 5 catégories avec SIFT, clustering et utilisation dans une SVM linéaire MC "cramer singer". Évaluation des modèles par matrices de confusion.

Système de recommandation

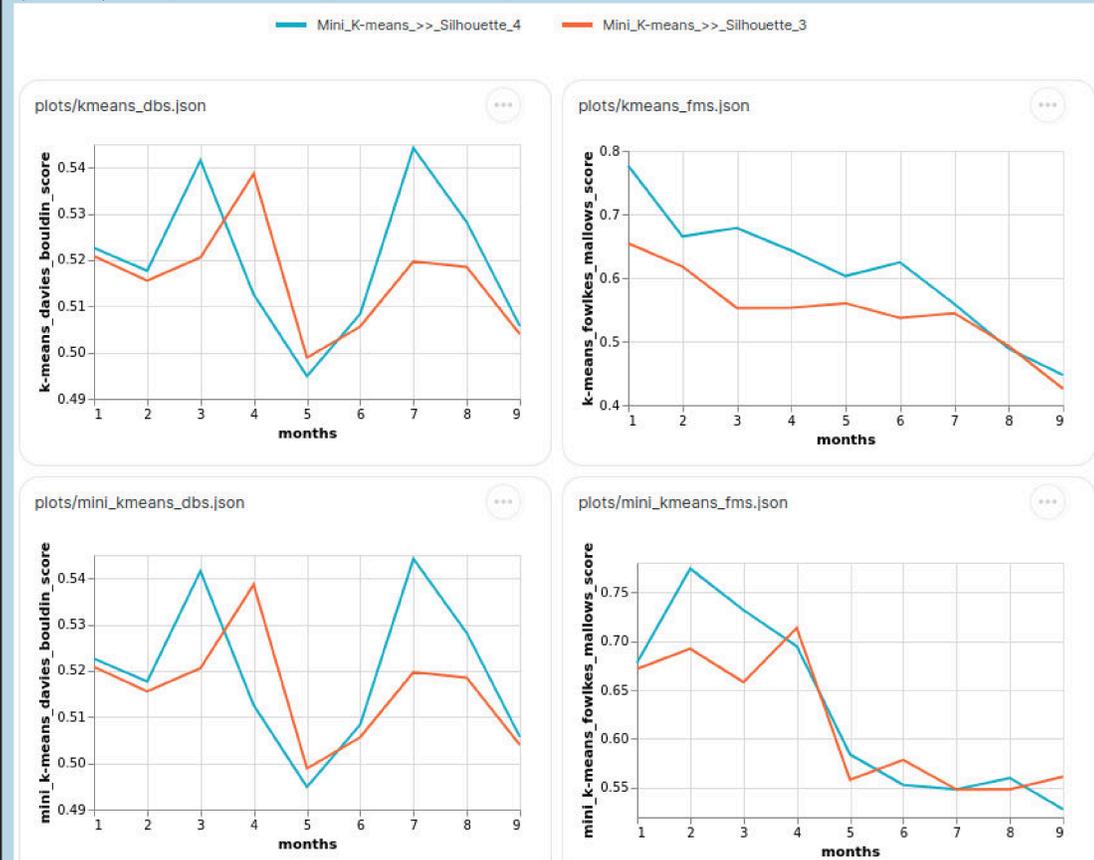
Programmation d'une IA "from stem to stern". Un pôle de data-science, un pôle mobile pour le client et un pôle de diffusion continue des recommandations par le cloud. Sélection des modèles suivant une étude comparative, en splittant les données LOO, mesure du **Hit Rate**, **RMSE**, **Diversité**, **Nouveauté**, etc. Approfondissement de **collaborative filtering** et **content-based filtering**, utilisation d'UML pour l'architecture du produit IA.

Détecter les Bad Buzz

Analyse des sentiments (+)(-)(0) sur des tweets : Pre-processing des data, **NLP**, **lemmatisation**, **stemming**, **Word2Vec**, **Bert**, **Glove** avec réseaux de neurones type **LSTM**. Programmation du calcul d'une fonction coût orientée métier pour limiter l'impact économique des bad buzz avec **TensorFlow**. Comparaison des résultats avec des solutions propriétaires (**Azure Cognitive Service**).

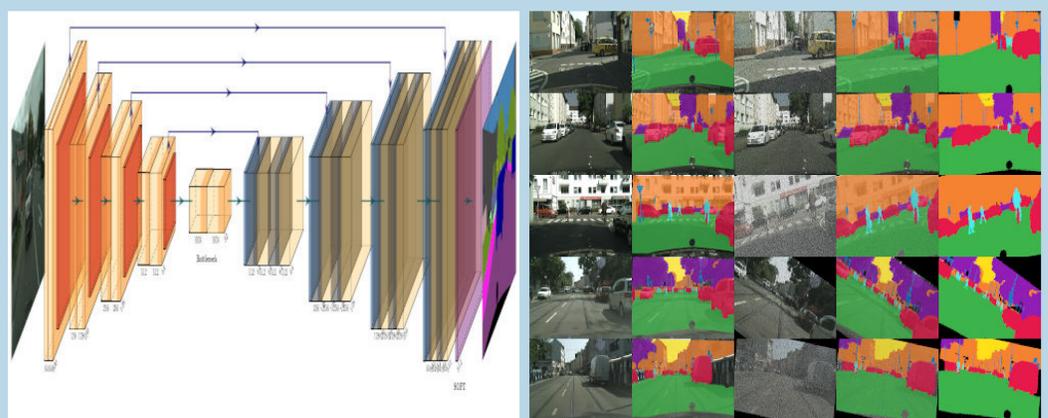
Segmentez des clients pour un e-commerce

Cleaning des data, **Exploratory Data Analysis (EDA)**, **Features Engineering**, Preprocessing des data. Représentation des clients sur une carte du monde, Diagramme Araignée, Diagramme calendrier (saisonnalité), Fréquence d'achat, Diagramme de flow (Big Data). Réduction de dimension par décomposition **PCA**, explicabilité de la variance/feature, **T-SNE**, **NMF**, **Factorisation Matricielle**. Entraînement et expérience par grid search, tirage aléatoire des hyperparamètres (gain de temps), résultat sous forme de table ou base de données **SQLite** et sur le web (**DVC**) ; aspect technique des métriques en vue d'évaluer la maintenabilité du modèle (data-flow).



Participation à la conception d'un véhicule autonome

Passage en revue des SOTA, utilisation de masques R-CNN, customisation de métriques avec **TensorFlow**, **data-augmentation**. Programme suivant une méthode **CI/CD** : publication sur **GitHub**, liaison avec **DockerHub** publication sur serveur Microsoft Azure, monitoring dans **IterativeStudio (DVC)**. Application de data-augmentation, comparée à un standard, avec 5 transformations sur des images en hautes résolutions, ainsi que sur les masques, par synchronisation.



Utilisation de réseaux de neurones pour apprentissage profond, Auto-Encoders, ResNet50, **DeepLabV3**, **Keras**. Laboratoire d'expérimentation avec comparaison des IA sous la forme de tableur et sous forme de courbes. Déploiement de l'application dans le cloud.

