

Nabil BOUCHANANE

Software Engineer Data

SYNTHESE SAVOIR-FAIRE / SAVOIR-ETRE

- Leadership / Rigueur / Sens du relationnel/ Esprit d'équipe/ Curiosité / Autonomie / Capacité d'adaptation / Flexibilité / Gestion du stress / Capacité à travailler sous pression/
- Etablir un cahier des charges / Rédaction de spécifications / Coaching / Pair Programming/ Rédaction de veilles technologiques/ Animation de Rétrospective/ Estimation des tâches/ Animation de la Backlog/ Rédaction des bonnes pratiques / Revue de code / Analyse d'impact

PRINCIPALES COMPETENCES TECHNIQUES

- **Front-end** HTML5, CSS3, JavaScript (ES6+), ESLint, TypeScript, Tailwind CSS, Bootstrap, Angular, React.js, Vue.js, Figma, Chrome DevTools
- **Back-end** Python (Django, Flask, FastAPI), Java (Spring Boot), Node.js (Express), RESTful APIs, GraphQL, Microservices, Docker, Kubernetes, Apache Kafka, Nginx, Jenkins, GitLab CI/CD
- **TDD, BDD** PyTest, JUnit, Jasmine, SonarQube, CircleCI, Postman
- **Outils / Logiciels** Git (GitHub, GitLab, Bitbucket), SVN, AWS (S3, EC2, RDS, Lambda), Google Cloud Platform (GCP), Azure ELK Stack (Elasticsearch, Logstash, Kibana) Apache Spark, Hadoop, Kafka, VS Code, PyCharm, IntelliJ
- **SGBD** PostgreSQL, MySQL, SQL Server, Oracle, MongoDB, Google BigQuery
- **Data Science** Python (Pandas, Numpy, Scikit-learn), R, Julia, Jupyter Notebook, TensorFlow, PyTorch Réseaux de neurones (ANN), LLM, NLP, Random Forest, SVM, Clustering (K-Means), PCA Power BI, Tableau, Matplotlib, Seaborn, Looker
- **Système** Microsoft Windows, GNU Linux Bash, PowerShell OAuth 2.0, JWT, SSL/TLS VMware

FORMATION

- | | |
|------|--|
| 2022 | Polytech Nantes : Diplôme d'ingénieur de l'école polytechnique universitaire de Nantes, spécialité Informatique Décisionnelle |
| 2019 | ESIR : Classe préparatoire intégrée, spécialité technologie de l'information |

LANGUES

- Anglais technique : lu, écrit, parlé.

CERTIFICATIONS

(En cours)	Professional Scrum Master I
Mars 2023	Google Cloud Big Data and Machine Learning Fundamentals
Mars 2021	TOEIC 785/990
Avril 2019	Google Analytics for digital marketing

EXPERIENCES ASSOCIATIVES

2021-2022	Ambassadeur Capgemini Student Connection – Référent Nantais
2020-2021	Vice-Président IdéeSys – Junior Entreprise Polytech Nantes

CONCOURS DE DEVELOPPEMENT

Hyblab 2021	Développement d'une application web permettant aux familles de découvrir, classer et visiter les meilleurs parcs et forêts de Nantes.
Google Hash Code 2020	Optimisation de la numérisation de livres dans plusieurs bibliothèques sous contrainte de temps et de scores de valeur.
Google Hash Code 2019	Optimisation du diaporama avec des photos et des transitions basées sur des tags.

PRINCIPALES MISSIONS

Juillet 2023 – Septembre 2024

SQUAL-IT

Software Engineer Data

Projet :

SQUALI : Création d'un outil Informatique qui intègre des modèles d'intelligence artificielle pour automatiser la correction des mauvaises pratiques de programmation et optimiser les processus de développement en qualité logiciel. Squali optimise les processus de développement, propose des corrections automatisées de code, et utilise des analyses avancées pour identifier les mauvaises pratiques ("code smells").

Réalisations :

Conception :

- Analyse de l'existant et sélection des technologies adaptées.
- Conception de l'architecture logicielle et du modèle de données.
- Étude comparative d'API (Open AI, LLama, Mistral AI) pour l'automatisation de la correction de bugs.

Développement :

- Implémentation de services RESTful pour la gestion des données et administration de bases PostgreSQL.
- Analyse de données textuelles avec Jupyter Notebook (Pandas, Numpy, Matplotlib) et mise en œuvre de modèles de Machine Learning (RandomForestClassifier, LogisticRegression).
- Création d'un agent capable d'analyser et corriger automatiquement du code informatique, en appliquant des techniques basées sur des modèles de langage pré-entraînés (LLM).
- Développement d'un frontend dynamique en AngularJS et préparation de tests automatisés avec Jasmine.

Infrastructure et déploiement :

- Déploiement de microservices avec Docker et Kubernetes sur des machines virtuelles AWS.
- Gestion des serveurs sous Unix et intégration continue des applications.

Organisation & méthodologie :

- Pilotage en mode Agile Scrum : planification, découpage des tâches, suivi quotidien (daily meeting) et rétrospectives.
- Mise en place de sprints de 3 semaines, suivi des livraisons et démonstrations via vidéoconférence.

Environnement Technique :

• *Python, JavaScript, Java, SQL, Bash, Flask, Spring Boot, AngularJS, Docker, AWS, TensorFlow, Scikit-Learn, PostgreSQL, Open AI, LLama, Mistral AI, Hugging Face.*

Novembre 2023 – Février 2024

SQUAL-IT

Software Engineer Data

Projet :

RedLetter : Développement d'une plateforme web qui facilite la veille technologique en générant des résumés personnalisés d'articles populaires sur Reddit, grâce à l'intelligence artificielle.

Réalisations :

Conception :

- Rédaction du cahier des charges et définition de l'architecture de l'application.

Développement :

- Mise en place d'une interface simple en JavaScript/HTML/CSS pour permettre aux utilisateurs de récupérer facilement les derniers articles tendances.
- Génération de fichiers de sortie (PDF, ODT, TXT) pour une consultation hors ligne des résumés.
- Extraction des articles et création de résumés personnalisés via l'API Open AI, intégrée au backend en Python (Flask).
- Automatisation du processus de récupération des articles de Reddit grâce à l'API Reddit PRAW.

Infrastructure et déploiement :

- Déploiement de l'application sur un serveur AWS EC2, avec gestion du nom de domaine via OVH.
- Mise en place d'un certificat SSL auto-signé pour sécuriser le déploiement, automatisé via un script Bash.

Environnement Technique :

• *Python, JavaScript, HTML/CSS, Flask, Bash, Open AI, Reddit PRAW, ReportLab, ODF, AWS EC2, Gunicorn, OVH.*

Janvier 2023 – Mai 2023

MANITOU GROUP

Lead Data Scientist

Projet :

Développement d'un Proof of Concept (POC) pour Manitou Group, visant à créer une extension VS Code capable de détecter et corriger automatiquement les bugs, en particulier le "code mort". Le projet a été conçu pour améliorer la qualité de leurs projets internes via l'intelligence artificielle, sans toutefois être déployé ou commercialisé.

Réalisations :

Conception :

- Recueil des besoins du client, centré sur les problèmes de "code mort" et les inefficacités dans leurs projets logiciels.
- Élaboration d'une solution complète intégrant l'automatisation de la détection de bugs et la correction de code via IA.
- Présentation de la solution, y compris une analyse des coûts de fonctionnement potentiels.

Développement :

- Développement d'une interface utilisateur dans VS Code pour sélectionner et analyser les parties du code à corriger.
- Conception et expérimentation de modèles de Machine Learning pour automatiser la détection de bugs, avec des ajustements basés sur les données des projets de Manitou.
- Mise en place d'une API pour intégrer les modèles de correction dans l'environnement de développement.

Organisation & méthodologie :

- Planification des tâches avec l'équipe et suivi du projet, en assurant un retour régulier sur l'avancement.
- Présentation de la solution en anglais devant un comité de 50 personnes à Paris, avec un focus sur l'impact potentiel de l'outil dans leurs flux de développement.

Environnement Technique :

• *Python, Flask, AngularJS, Docker, C#, TypeScript, SQL Server, SonarQube, Azure DevOps, VS Code, Git.*

Septembre 2022 – Novembre 2022

SPECTRE BIOTECH

Lead Data Scientist

Projet :

Développement d'une plateforme dédiée à la collecte et l'analyse de données EEG pour faciliter la recherche et le diagnostic en santé mentale. Ce projet inclut la création d'une interface open data pour le partage des données et l'élaboration d'un business model visant à valoriser ces informations pour les chercheurs et professionnels de santé.

Réalisations :

Conception :

- Recueil des besoins et animation de réunions avec le client pour affiner les objectifs du projet.
- Modélisation de la solution, incluant la création d'une interface de partage des données EEG et l'architecture générale de la plateforme.
- Élaboration d'un business model pour la monétisation des données EEG, conçu pour soutenir la recherche tout en créant des opportunités commerciales.

Développement :

- Mise en place d'algorithmes de machine learning pour détecter des pathologies mentales à partir de données EEG.
- Développement d'une maquette fonctionnelle pour l'interface open data, facilitant l'accès et la visualisation des données pour les chercheurs.
- Analyse des données EEG avec Python (Numpy, Pandas, Matplotlib) et expérimentation d'algorithmes avec Scikit Learn.

Infrastructure et déploiement :

- Conception de l'architecture complète du projet, planification des tâches à l'aide d'un GANTT, et intégration avec Azure DevOps pour la gestion de projet.
- Utilisation de technologies comme GraphQL, Elasticsearch et Spring Security pour garantir la sécurité et l'accessibilité des données.

Environnement Technique :

• *GraphQL, Spring Security, Elasticsearch, Kibana, Python, Jupyter Notebook, Numpy, Pandas, Matplotlib, Scikit Learn, Azure DevOps, Figma.*

Mars 2022 – Juin 2022

KUHN

Lead Data Scientist

Projet :

Développement d'un webservice intégré à l'application e-commerce de Kuhn, permettant aux utilisateurs d'identifier une pièce usée (comme une lame) à partir d'une simple photo. Ce service vise à simplifier le processus de remplacement de pièces pour les agriculteurs et à optimiser la chaîne d'approvisionnement.

Réalisations :

Conception :

- Recueil et analyse des besoins du client, en collaboration avec Kuhn, pour déterminer les fonctionnalités clés du webservice.
- Simulation des gains économiques pour Kuhn et les agriculteurs, démontrant la valeur ajoutée du service.

Développement :

- Génération d'un dataset 2D à partir de modèles 3D de lames, créés sous Blender, pour entraîner l'IA.

- Mise en place d'une IA de détection d'images pour identifier les lames sur les photos en 2D, en utilisant Pytorch et des techniques d'analyse d'image.
- Développement d'un algorithme génétique pour simuler l'usure des lames afin d'améliorer la précision des identifications.
- Conception d'une interface utilisateur en Flutter, permettant de rediriger les utilisateurs vers le site e-commerce de Kuhn après identification des pièces.

Organisation & méthodologie :

- Gestion du projet via des outils de planification comme le GANTT, en suivant régulièrement l'avancement avec le client.
- Présentation de la solution en anglais devant le client à Paris au bâtiment 147 de Capgemini.

Environnement Technique :

- *Blender, Unity, Flutter, Pytorch, Python (Numpy, Pandas, Matplotlib), Scikit Learn.*

Janvier 2022 – Mai 2023

LA BANQUE POSTALE

Rédacteur Scientifique IA/DATA

Projet :

Création de newsletters mensuelles pour La Banque Postale, mettant en avant les dernières innovations technologiques dans le domaine de l'intelligence artificielle et leur impact potentiel sur les secteurs financiers et économiques. Ce projet vise à tenir les décideurs informés des évolutions stratégiques et technologiques.

Réalisations :

Conception :

- Mise en place d'une stratégie de veille technologique pour identifier les avancées majeures en IA et nouvelles technologies.
- Collaboration avec le client pour définir le format, le calendrier éditorial, et les processus de validation pour l'année 2023.

Rédaction :

- Rédaction et publication d'articles sur des sujets innovants, notamment les impacts futurs des nouvelles technologies.
- Coordination d'une équipe de rédacteurs, avec planification des livrables pour respecter les échéances.

Organisation & méthodologie :

- Gestion du projet via des outils collaboratifs (Teams, Outlook), avec suivi des rendus et validation des articles par le client.
- Contribution à l'élaboration du devis pour le client, avec une estimation des coûts pour l'année.

Environnement Technique :

- *Microsoft Pack Office, Google Office, Teams, Outlook.*

Septembre 2021 – Août 2022

SOGETILABS

Software Engineer Data

Projet :

Cortex IA : Développement d'une interface cerveau-machine exploitant des données EEG collectées par un casque neuronal pour identifier les états mentaux et faire du contrôle d'objet. Le projet comprend le traitement des signaux cérébraux, la visualisation et l'analyse des données, ainsi que l'intégration de modèles d'apprentissage automatique pour la classification des états de pensée.

Réalisations :

Conception :

- Analyse des besoins fonctionnels et mise en place de l'architecture projet avec l'équipe technique.
- Analyse approfondie et traitement des données EEG en Python (échantillonnage, labellisation, normalisation, transformée de Fourier, transformée en ondelettes).
- Rédaction complète de la documentation technique (cahier d'expérimentation, manuel d'utilisation) et des rapports pour les recherches scientifiques (dossiers CIR).

Développement :

- Développement de l'interface utilisateur (AngularJS) pour la visualisation et l'interaction avec les données.
- Création d'un backend avec Java Spring Boot pour gérer les routes GET/POST et déployer les modèles d'apprentissage automatique.
- Mise en place d'un websocket Mosquitto en JavaScript pour faciliter la communication en temps réel.
- Migration des données EEG de MongoDB vers ELK (ElasticSearch, Logstash, Kibana) pour l'archivage et l'utilisation dans des sessions et logs IA.
- Développement et mise en production de modèles de Machine Learning (Arbre de décision, SVM, Random Forest, R-CNN) et Deep Learning pour l'analyse des états mentaux.
- Mise en place d'un pipeline CI/CD pour automatiser les entraînements et les tests des modèles IA sur des VM Azure et stocker les résultats sur ELK.

Méthodologie :

- Organisation des réunions d'équipe et des revues de sprint, utilisant des outils comme Miro pour les feedbacks et la collaboration.
- Animation de Sprint Reviews pour obtenir des retours d'équipe via des supports visuels.
- Formation de deux collaborateurs internes sur les outils DevOps et Machine Learning.

Environnement Technique :

• *Python (Poetry, Pip, Pandas, Numpy, Scikit-learn, Flask), Java, Spring, Angular, Javascript, Mosquitto, MongoDB, Docker, Kubernetes, TensorFlow, Keras, Microsoft Azure (VM, KeyVault, SQL Database), Stack ELK (ElasticSearch, Kibana, Logstash, Beats), Shell, SVN, Yaml, Qbuild, SonarQube, Doxygen.*

Septembre 2021 – Février 2022

LS2N

Data Scientist

Projet :

Développement d'un modèle d'apprentissage profond pour l'identification des scripteurs sur des papyrus grecs.

Réalisations :

Conception :

- Réalisation d'une étude bibliographique sur les méthodes existantes en vision par ordinateur pour comprendre les défis de l'analyse de manuscrits anciens.
- Formulation d'hypothèses de recherche basées sur l'état de l'art des techniques de traitement d'images.

Développement :

- Expérimentation avec des techniques de traitement d'images, telles que l'augmentation de données, la binarisation adaptative, et la réduction du bruit pour améliorer la qualité des images papyrologiques.

- Implémentation d'algorithmes de machine learning pour le clustering des sections des papyrus, permettant une première classification des fragments.
- Entraînement de modèles de deep learning, incluant des réseaux de neurones convolutifs (CNN) et des méthodes de transfer learning avec ResNet et VGG, atteignant près de 95 % de précision pour l'identification des scribes.
- Validation des résultats via des tests approfondis et optimisation des hyperparamètres pour maximiser les performances des modèles.

Méthodologie :

- Méthode Agile : Fonctionnement avec des sprints mensuels d'une durée de 1 mois, avec des réunions hebdomadaires pour suivre les avancées du projet.
- Présentation des résultats devant une équipe de papyrologues en Suisse, en mettant l'accent sur l'impact potentiel des méthodes développées pour les recherches futures.

Environnement Technique :

• *Windows, Jupyter Hub, Git, Python, Bash, Pandas, Numpy, Matplotlib, OpenCV, Seaborn, Math, Time, Pytorch, Keras, VsCode, Jupyter Notebook.*

Juin 2021 – Août 2021

CHU

Data Analyst

Projet :

Détection de pathologie cognitive en analysant des traces numériques de trajectoire issues des technologies de réalité virtuelle.

Réalisations :

Conception :

- Analyse des besoins et définition des objectifs du projet.
- Étude de complexité et optimisation des algorithmes de traitement des données.

Développement :

- Analyse et mise en forme de données sur des Dataframes à l'aide de Pandas et Jupyter Notebook.
- Visualisation des données de trajectoire en 2D/3D avec Matplotlib et Mpl_toolkits.
- Débogage d'un websocket en JavaScript pour le transfert de données en temps réel entre l'application Unity et le serveur Express JS en Backend.

Méthodologie :

- Présentation des résultats aux équipes de recherche.
- Utilisation d'une approche agile avec des sprints de 1 mois et un suivi hebdomadaire des avancées.

Environnement Technique :

• *Windows, Ubuntu, Git, GitLab, Python, Bash, JavaScript, C++, Express.js, React, Pandas, Numpy, Matplotlib, Mpl_toolkits, OpenCV, Seaborn, MySQL, VS Code, Jupyter Notebook, Unity.*

Août 2020 – Août 2021

TMG

Technicien Web

Projet :

Recherche et analyse de plateformes de streaming illégales en ligne, dans le cadre de la conformité avec la loi Hadopi.

Réalisations :

Conception :

- Analyse des exigences légales liées à la loi Hadopi et définition des objectifs de conformité.
- Élaboration d'une stratégie pour l'identification et le signalement des plateformes de streaming illégales.

Développement :

- Développement d'extensions Chrome en interne pour récupérer les liens iframe et les sources des vidéos de streaming (football, films, séries).
- Signalement et ajout des sites de streaming illégaux dans une base données MySQL.
- Utilisation de plateformes administratives sur YouTube, Facebook et d'autres réseaux sociaux pour signaler les contenus illégaux.
- Débogage d'extensions Chrome en JavaScript pour résoudre les problèmes d'interruption de fonctionnement.
- Test des liens signalés pour vérifier leur validité et leur accessibilité.

Méthodologie :

- Travail effectué principalement pendant les heures de match ou de forte affluence sur les plateformes de streaming, afin d'optimiser la détection des contenus illégaux.
- Suivi des tendances et des évolutions des plateformes de streaming illégales pour adapter les méthodes de recherche.

Environnement Technique :

• *JavaScript, Chrome DevTools, Python, Pandas, plateformes administratives (YouTube, Facebook), systèmes de gestion de tickets.*

Septembre 2020 – Mai 2021

CAPACITES

Développeur Fullstack

Projet :

Création d'un dashboard pour visualiser les données utiles au monitoring des actions effectuées par des "crowdsourcer" afin de garantir la qualité et la véracité des tests.

Réalisations :

Conception :

- Élaboration du cahier des charges avec les besoins, les plannings et les risques du projet.
- Conception des visuels de l'application avec Figma et validation avec le client.

Développement :

- Analyse et visualisation de données sur un Jupyter Notebook avec Pandas et Matplotlib.
- Recherche et définition des indicateurs pour la détection de sessions suspectes.
- Création d'un serveur backend Python Flask pour l'envoi des données.
- Développement du frontend et des dataviz avec ReactJs et D3Js.
- Déploiement de la solution web sur Heroku.

Méthodologie :

- Équipe de deux développeurs avec un chef de projet.
- Travail en mode Agile avec des cycles de trois semaines et un suivi hebdomadaire des avancées.

Environnement Technique :

Mai 2020 – Août 2020

WEENAT

Développeur Backend

Projet :

Développement d'algorithmes liés à la maintenance préventive des capteurs sur une application agrométéorologique destinée aux agriculteurs.

Réalisations :

Conception :

- Collaboration avec des ingénieurs électroniques pour définir des KPI et des seuils (threshold) sur l'état des capteurs.
- Élaboration des spécifications techniques pour l'intégration des flux de données.

Développement :

- Création de requêtes SQL et Pyspark pour l'acquisition et le traitement des flux de données.
- Développement de requêtes POST/GET dans le backend pour automatiser les traitements.
- Exportation des données sous forme de fichiers avec différentes extensions (CSV, XLSX).
- Mise en forme des données dans le frontend, incluant des fonctionnalités de filtrage.
- Prise en charge des tickets sur le backlog pour traiter et déboguer des fonctionnalités du dashboard en Django.

Méthodologie :

- Travail en collaboration avec une équipe de 6 développeurs, un tech Lead et le CEO.
- Utilisation de la méthode Agile avec des sprints de trois semaines et un suivi hebdomadaire des avancées.

Environnement Technique :

• *Git, Jira, Python, Java, SQL, Bash, JavaScript, Django, JEE, PostgreSQL, Apache Spark*
