

Bilal Bouchemella . Jarvis Consulting

Je suis étudiant à la maîtrise en Data Science and Business Analytics à HEC Montréal et diplômé en informatique de l'Université McGill, où je me suis concentré sur le génie des données, l'analyse de données et le machine learning. Je travaille présentement comme ingénieur de données chez Jarvis, où je conçois des pipelines de données, développe des composants logiciels et contribue à divers projets analytiques. J'ai auparavant travaillé comme Data Scientist et comme technicien en informatique, utilisant notamment Python et Julia pour développer des workflows analytiques et automatiser des tâches de traitement de données, tout en acquérant de l'expérience en machine learning, en traitement de données et en opérations systèmes. Ces expériences ont enrichi ma passion pour l'utilisation des données, du machine learning et de l'IA afin de créer des solutions pratiques, fiables et adaptées à des problématiques réelles.

Skills

Proficient: Python, Java, RDBMS/SQL (PostgreSQL), Julia, Pandas/Numpy, Scikit-learn, Linux/Bash, Git, Agile/Scrum

Competent: Plateformes cloud (AWS, GCP), Docker, Jira, Power BI, R

Familiar: Spark, DataBricks, Snowflake, C, C++

Jarvis Projects

Project source code: https://github.com/jarviscanada/jarvis_data_eng_BilalBouchemella

Python – Analyse clients et performance (PoC) [GitHub]: Analyse de deux années de données transactionnelles e-commerce afin d'identifier les comportements clients, les habitudes d'achat, la saisonnalité et la valeur client. Mise en place d'un pipeline analytique en Python (Pandas, NumPy, Matplotlib) sous Jupyter, connecté à PostgreSQL et Docker. Nettoyage et structuration des données, création d'indicateurs comportementaux et mensuels, calcul des métriques RFM, tendances de ventes et utilisateurs actifs pour différencier clients nouveaux et récurrents, et soutenir des stratégies de marketing ciblé et de rétention.

Java Grep Application [GitHub]: Développement d'une application Java en ligne de commande reproduisant les fonctionnalités essentielles de la commande Linux grep, en s'appuyant sur les principes de la programmation orientée objet. Parcours récursif de répertoires, lecture de fichiers ligne par ligne et filtrage via expressions régulières avec Streams et lambda expressions. Gestion du projet avec Maven et GitHub, tests manuels et unitaires avec JUnit, puis déploiement sous forme de JAR exécutable et d'image Docker via un Dockerfile personnalisé.

Agent de surveillance d'un cluster Linux [GitHub]: Conception d'un agent de surveillance pour un cluster Linux permettant à l'utilisateur de consulter et centraliser les informations matérielles de chaque machine, notamment l'utilisation du CPU, de la mémoire et du disque. Le système assure un suivi continu des performances en collectant automatiquement les données à intervalles réguliers grâce à un cron job. L'agent repose sur des scripts Bash qui récupèrent les informations système et les stockent dans une base de données PostgreSQL. L'ensemble du projet est déployé à l'aide de Docker afin de garantir une installation, une portabilité et une gestion uniformes sur différentes machines.

Highlighted Projects

Modélisation du risque de crédit par machine learning: Développement de modèles prédictifs du défaut de crédit à partir de données financières et comportementales. Création de variables pertinentes (ratios financiers, interactions, normalisation) en Python (Pandas, NumPy). Entraînement et comparaison de modèles (régression logistique, Random Forest, XGBoost) avec validation croisée, évalués via ROC-AUC, précision-rappel et statistique KS. Analyse des performances pour appuyer des décisions de gestion du risque, avec visualisations interactives sous Plotly.

Analyse de la couverture médiatique de Gladiator II: Analyse de la couverture médiatique du film Gladiator II en collectant des centaines d'articles liés au film à l'aide de Python et de NewsAPI. Appliquant un open coding et une annotation manuelle pour classer les thèmes, j'ai ensuite calculé des scores TF-IDF afin d'extraire les termes clés et visualisé les tendances de couverture avec JupyterLab et Matplotlib pour comparer la visibilité du film avec d'autres sorties. Ce projet m'a permis de renforcer mes compétences en analyse textuelle, en extraction de caractéristiques, en pipelines de collecte de données et en collaboration au sein de l'équipe grâce à une coordination régulière et au partage des responsabilités.

Système de gestion hôtelière: Développement d'un système complet de base de données hôtelière dans le cadre d'un projet d'équipe, incluant la conception du schéma ainsi que la création et la gestion des tables SQL. Développement

également d'une application Java dotée d'une interface permettant les requêtes, les mises à jour et diverses opérations. Collaboration étroite avec les membres de l'équipe, où la communication et la coordination étaient essentielles, et prise en charge de l'organisation des tâches, de la répartition du travail et de l'intégration des composants de la base de données et de l'application.

Professional Experiences

Ingénieur de données, Jarvis Consulting Group (Nov 2025 - présent): Développement de solutions en génie des données en construisant des composants logiciels, en concevant des pipelines de données et en contribuant à différents projets d'analyse et de traitement des données. Participation au codage, au débogage et à l'optimisation des traitements, tout en mobilisant de solides compétences en résolution de problèmes, en communication et en qualité du code pour assurer la fiabilité et l'efficacité des livrables.

Scientifique des données, R2 (Mai 2025 - Août 2025): Développement et optimisation de modèles de machine learning en Python et Julia pour l'analyse prédictive de séries temporelles sur de grands ensembles de données industrielles. Conception et mise en place de pipelines de nettoyage et de prétraitement afin de transformer des données brutes provenant de capteurs en caractéristiques fiables. Conception et conduite d'expérimentations permettant d'évaluer les modèles en équilibrant précision et efficacité. Mise en place de workflows reproductibles favorisant la collaboration au sein de l'équipe de science des données. Renforcement de mon expérience en développement sous Linux, en gestion de versions avec Git et en traitement de données à grande échelle dans un environnement orienté recherche.

Technicien en technologies de l'information, Tennis Canada (Mai 2024 - Avr 2025): Maintenance et mise à niveau des systèmes informatiques en collaboration avec les équipes TI, incluant le soutien technique complet sur Mac et Windows afin d'améliorer l'efficacité du personnel. Gestion de la base de données interne pour assurer la fiabilité et la disponibilité des informations opérationnelles. Utilisation de JIRA pour le suivi des demandes et la coordination des tâches. Configuration et entretien de plus de 100 ordinateurs portables pour le tournoi Omnisport Banque Nationale 2024, garantissant un fonctionnement optimal de l'équipement pour l'ensemble de l'événement.

Education

HEC Montréal (2025 - présent), Maîtrise en gestion (M. Sc.), Science des données et analytique d'affaires - Étudiant-athlète : Carabins Soccer - Cours pertinents : Machine Learning, Statistical Modelling et Apprentissage Statistique

McGill University (2021 - 2025), Baccalauréat ès sciences (BSc), Informatique - Étudiant-athlète : Capitaine Redbird Soccer - Prix Redbird Athlete Award (2021, 2022, 2023) : remis aux étudiants-athlètes s'étant distingués dans leur sport - Relevant Coursework: Data Science, Natural Language to Data Science, Applied Machine Learning, Reinforcement Learning, Database Systems, Software Design, Algorithms and Data Structures, Operating Systems, Discrete Structures, Probability, Statistics, and Statistical Computing

Miscellaneous

- Joueur de l'année de l'équipe de soccer - McGill Athletics (2022)
- Bourse de recrutement universitaire - Fondation Aléo (2021) : remise aux meilleurs étudiants-athlètes de la province pour leur excellence sportive et académique
- Joueur de soccer semi-professionnel avec plus de six ans d'expérience au Québec, évoluant dans ce sport depuis l'âge de cinq ans. Ce parcours m'a permis de développer de fortes valeurs telles que le leadership, le travail d'équipe, la discipline et la résilience.