IDRISSI YASSIR

Statisticien

yassir.idrissi08@gmail.com

Meknès,Maroc

+212663181547

FORMATION

2023-2025 **Master en Statistique appliquées et Méthodes numériques (opt: Statistique appliquées)**
Faculté des Sciences Ain chock, Casablanca

2022-2023 **Licence en Sciences Mathématiques et Applications**
Faculté des Sciences Moulay Ismail, Meknès

2021-2022 **Diplôme D'étude Universitaire Générale en Sciences Mathématiques et Applications**
Faculté des Sciences Moulay Ismail, Meknès

2014-2015 **Baccalauréat en sciences physiques**
Lycée Zitoune, Meknes

EXPERIENCE

Stage En laboratoire de la faculté de sciences Ain chock

Mars2025 - Juin2025

Dans le cadre de mon mémoire de Master, j'ai travaillé sur l'optimisation de méthodes d'apprentissage automatique appliquées à l'imagerie médicale, à travers le paradigme de l'apprentissage actif profond. J'ai conçu et implémenté un pipeline complet intégrant des réseaux neuronaux convolutifs (CNN) et diverses stratégies d'échantillonnage (incertitude, représentativité, diversité, approches hybrides), appliquées au dataset NCT-CRC-HE-100K. Les résultats ont démontré que l'apprentissage actif permet d'améliorer significativement les performances de classification (Accuracy, Recall, AUC) tout en réduisant le coût d'annotation, problématique centrale dans les données médicales. Ce projet m'a permis d'acquérir une expertise pratique dans la mise en œuvre de modèles robustes sur des données biomédicales complexes, ainsi qu'une expérience approfondie des paradigmes de l'IA (supervisé, non supervisé et par renforcement). Au-delà de l'imagerie, cette expérience a renforcé mon intérêt pour l'intégration de l'IA dans les processus décisionnels médicaux et m'a doté de compétences directement transposables aux défis du projet AI-CARE :

- Conception de pipelines IA appliqués à des données médicales hétérogènes,
- Optimisation de modèles pour un usage interprétable et économe en ressources d'annotation,
- Exploration de stratégies d'apprentissage adaptatif permettant d'améliorer les systèmes en continu à partir de données cliniques réelles.

COMPÉTENCES

Langages : R, Python, Java, SQL, Power BI, SAS

Bases de données : SQL Server**,** MongoDB, Data Warehousing,.Oracle 21c, ETL.

Statistiques descriptives : Modèles de régression, Analyse de la variance (ANOVA), Réduction de la dimensionnalité.

Statistiques inférentielles : Échantillonnage, Modélisation, Tests d'hypothèses, Estimation par intervalle de confiance.

Séries temporelles : Modèles stochastiques.

Programmation mathématique et optimisation sous contraintes : Descente de gradient, Méthode du simplexe.

Machine learning : Apprentissage supervisé, non supervisé, Méthodes d'ensemble (Voting, Stacking, Bagging, Boosting) avec **Scikit-learn**, **XGBoost**.

Deep learning : ANN, CNN, RNN, Auto-encodeurs, Transformers, Transfer learning/Fine-tuning avec **Keras** et **TensorFlow**.

Data / Text / Web Mining : Collecte, nettoyage de données, extraction de données web, analyse de contenu, extraction de règles d'association**.**

NLP & LLMs : Tokenization, Vectorisation (TF-IDF, embeddings), Transformers.

PROJETS

Statistical and machine learning models for predicting spalling in CRCP:

projet développe des modèles statistiques et d'apprentissage automatique pour prédire le spalling dans les chaussées en béton armé continu (CRCP), en utilisant des données sur les matériaux, le climat et les charges, afin d'optimiser la maintenance.

Mesure et contrôle sur le risque de marché:

Le projet se concentre sur la mesure et le contrôle du risque de marché en utilisant des outils comme la Value at Risk (VaR) pour quantifier et limiter les expositions aux risques financiers.

Anomaly detection using unsupervised machine learning algorithms(A simulation study):

Le projet compare cinq algorithmes de détection d'anomalies non supervisée sur des données simulées : One-Class SVM, One-Class SVM avec SGD, Isolation Forest, LOF et Robust Covariance.

LANGUES

Français: Courant **Anglais: Courant** **Arabe: Mternelle**