SEDJARI YASSIR

**30ans,**

**Casablanca**

**Tél :0691324249 DataEngineer**

**Email: yassir.sedjari@gmail.comga**

|  |  |
| --- | --- |
| Compétences techniques | |
| **BigData** | **HDP,Cloudera,Hadoop,HIVE,Pig,ApacheSpark,ApacheSqoop,ApacheKafka, Cassandra** |
| **Business**  **Intelligence/Reporting**  **Applications** | **Microsoft Power BI** |
| **Développement général** | **Python ,Scala, Script Shell** |
| **COMPÉTENCES TECHNIQUES** | **Statistiques/Probabilité et modélisations des données**  **Pre-processing Data/BIGDATA/Data Science/Jupyter/Python/**  **Spark/Hbase/Hive/** |
| **Analyse & Conception** | **Analyse et spécification des besoins fonctionnelles** |
| **Environnements** | **Windows,MACOS,Linux(Redhat,Centos)** |
| **Bureautique** | **Traitement de texte,Word,Excel,Powerpoint** |
| Langues | |
| **Français :C1(Autonome)**  **Anglais :C1(Autonome)** | |
| Formation académique | |
| Master BigData & Data Sciences  **Faculté des sciences et technique–Settat-**  2019 | |
| **Master Topologie Algébrique et Analyse Robotique**  **Faculté des sciences Hassan II**  **09/2016–06/2018** | |
| **Licence Mathématiques Appliquée Faculté des Sciences Hassan II**  **09/2010–07/2014** | |

|  |  |
| --- | --- |
| **Baccalauréat Sciences Technique électrique Lycée ELKHAWARIZMI**  09/2008–06/2009 | |
|  | |
| **Formation & Certifications** | |
| **Formation BIGDATA (Infrastructure, Stockage NoSQL & DataEngineer) –IBM-**  Supervised Learning with scikit\_learn\_Statement of Accomplishment  Unsupervised Learning with scikit\_learn\_Statement of Accomplishment   SCRUMFOUNDATIONPROFESSIONALCERTIFICATE  CERTIFICATIONPOWERBI  BigData avec apache Spark | |
| Expériences professionnelles | |
| 01/2021–Présent | **Data Analytiques Société : Trador**  **Système d'analyse de sentiments pour la prédiction des cours boursiers Objectif :**  **Implémentation d’un Dashboard dynamique qui indique le cout des cours boursier**  **Les techniques utilisées :**  **-langage de programmation PYTHON (librairie Flask)**  **-ML : réseau de neurones récurrents avec LSTM**  **-La méthode des moyennes mobiles** |
| 10/2020–12/2020 | ***Datascientist***  **Marketface**  **-Collecté, Analysé et interprété des données brutes des acteurs économiques marocains (entreprises, médecins, laboratoires, avocats) à partir de différents sites Web en utilisant beautifulsoup and selenuim**  **-Extraire le secteur d’activité du texte de description des entreprises à l’aide de la modélisation thématique.**  **-Création d’un système de recherche multicritères qui interagit avec cet ensemble de données à l’aide de Flask.** |

|  |  |
| --- | --- |
| 03/2020–09/2020 | **BigData & Machine Learning**  **Les Laboratoires de Sothema**  **-Ma mission était de révolutionner le processus de la chaîne d’approvisionnement en mettant l’apprentissage automatique à la dispositiondesemployésenoptimisantleurapprovisionnementpouravoir le bon article au bon endroit, au bon moment et à moindre coût :**  **\*Création d’une architecture BigData (Spark, JupyterLab) ·**  **\*Analyse statistique avancée et modèles de deep Learning (LSTM, MLP,**  **GRU) pour prédire les ventes des produits phares du laboratoire (Soclavet COQUELUSÉDAL) ·**  **\*Collaboration avec les départements "SupplyChain"et"Marketing".** |
| 09/2019–03/2020 (stage) | ***DataScientist Sentiment Analysis and creation of new Knowledge to build customer loyalty***  **Groupe Cubamare (Decora, Nolte et Miton)**  **\*Science des données et soutien à la prise de décisions : analyse des sentiments et découverte des données pour la fidélisation des clients. \*Valeur ajoutée : Le principe a l’intérêt de mon travail est d’aider DECORA à prendre des décisions valables en matière de gestion interne, de formation, d’améliorations possibles, etc., en faisant entendre l’opinion de leurs clients actuel sous potentiels.**  **Réalisations :**  **-Création et comparaison d’architectures Hadoop (Cloudera et Horton Works) et Spark.**  **-Analyse du sentiment dans 1,6 million de tweets avec Python et R –Mise en œuvre d’un modèle de régression logistique pour la segmentation géo-démographique afin de prédire le taux de départ des clients dans une entreprise. (Analyse du roulement)**  **-Avec l’accord de la société, j’ai pris l’initiative de créer deux types de Chatbots avec Rivescript et DialogFlow (Api.ai). Le premier sera intégré dans le site DECORA pour le support à distance des clients réels et le second dans la page Facebook pour marquer la formation de l’entreprise et les services destinés aux clients potentiels.** |
|  | ***PROJETPERSONNEL***  **-Développement de consommateur Kafka pour collecter les données en temps réel depuis Yahoo et tweeter**  **-Développement de job spark (payspark) streaming d’une gestion dans HDFS et base**  **-Sérialisation désérialisation en Avro des données traite**  **-Ingestion des données sur HDFS en format parquet**  **-Développement des scripts de création des tables hbase** |

|  |  |
| --- | --- |
|  | **-Développement de scripts de création des topics KAFKA**  **-Réalisation des tests unitaire**  **-Développement de job python(librairie flask) pour la visualisation des Dashboard en temps réel.** |