

PROGRAMME D'ACQUISITION DES COMPÉTENCES EN INFORMATIQUE PAR GEHMIT

PROJET : AirSense: Analyse et Prédiction de la Pollution Atmosphérique à l'Aide de la Science des Données

Mois 1: Configuration du projet et collecte de données

1. Définir les objectifs du projet:

- ☐ Décrivez clairement les buts et objectifs du projet. Cela pourrait inclure la prévision de la qualité de l'air, l'identification des sources de pollution ou la recommandation de stratégies d'atténuation.

2. Revue de littérature:

- ☐ Effectuer une revue approfondie de la littérature existante liée à la pollution de l'air, aux sources de données pertinentes et aux méthodologies utilisées dans des projets similaires.

3. Collecte et exploration des données:

- ☐ Identifier et collecter des ensembles de données pertinents. Cela pourrait inclure des données de surveillance de la qualité de l'air, des données météorologiques et des données géographiques.
- ☐ Explorez les ensembles de données pour comprendre leur structure, identifier les valeurs manquantes et obtenir des informations sur les fonctionnalités potentielles.

4. Nettoyage des données et prétraitement:

- ☐ Nettoyez et préparez les données collectées. Gérer les valeurs, les valeurs aberrantes manquantes et normaliser les données si nécessaire.

Mois 2: Développement et évaluation du modèle

1. Ingénierie des caractéristiques:

- ☐ Identifiez les caractéristiques clés qui peuvent contribuer à la prédiction de la pollution de l'air ou à l'identification de la source. Créez de nouvelles fonctionnalités si nécessaire.

2. Sélection du modèle:

- ☐ Choisissez des modèles d'apprentissage automatique appropriés pour la tâche à accomplir. Considérez des modèles comme la régression pour la prédiction ou le clustering pour l'identification de la source.

3. Formation du modèle:

- ☐ Divisez l'ensemble de données en ensembles de formation et de validation.
- ☐ Formez vos modèles sélectionnés à l'aide des données de formation.

4. Évaluation du modèle:

- ☐ Évaluez les modèles à l'aide de l'ensemble de validation. Choisissez des mesures appropriées en fonction des objectifs de votre projet (par exemple, une erreur carrée moyenne pour les tâches de régression).

Mois 3: mise en œuvre et déploiement

1. Réglage fin (Fine-Tuning):

- ☐ Affinez les modèles en fonction des résultats de validation. Ajustez les hyperparamètres pour améliorer les performances.

2. Visualisation:

- ☐ **Développer des visualisations pour présenter les résultats. Cela pourrait inclure des cartes thermiques, des niveaux de pollution, des graphiques de tendances temporelles ou des cartes géographiques.**

3. Documentation:

- ☐ Documentez l'ensemble du processus, y compris les étapes de prétraitement des données, les architectures de modèle et les résultats d'évaluation. Cette documentation sera précieuse pour la référence future et la publication potentielle.

4. Stratégie de déploiement:

- ☐ Le cas échéant, discutez et planifiez le déploiement du modèle dans un paramètre réel. Considérez comment les résultats peuvent être intégrés dans les processus décisionnels.

5. Présentation et rapport:

- ☐ Préparez une présentation complète pour les parties prenantes, résumant les objectifs du projet, la méthodologie, les résultats et l'impact potentiel.