

1 ROADMAP AI-Engineer

Bienvenue dans cette roadmap dédiée aux futurs AI Engineers et ML Engineers. Dans un monde où l'intelligence artificielle et l'apprentissage automatique jouent un rôle central dans l'innovation technologique, ces parcours professionnels sont à la fois passionnants et prometteurs. Ce guide est conçu pour vous orienter à travers les étapes clés et les compétences essentielles nécessaires pour exceller dans ces domaines.

L'IA et le ML ne sont pas simplement des tendances passagères, mais des disciplines qui transforment profondément notre façon de vivre et de travailler. Que vous soyez un débutant curieux ou un professionnel cherchant à approfondir ses connaissances, cette roadmap vous fournira les outils et les ressources nécessaires pour progresser de manière structurée et efficace.

Au fil de ce parcours, nous explorerons les concepts fondamentaux, les compétences techniques, les technologies de pointe, et les meilleures pratiques de l'industrie. Préparez-vous à plonger dans l'univers fascinant de l'intelligence artificielle et de l'apprentissage automatique, et à bâtir une carrière qui vous placera à l'avant-garde de l'innovation technologique.

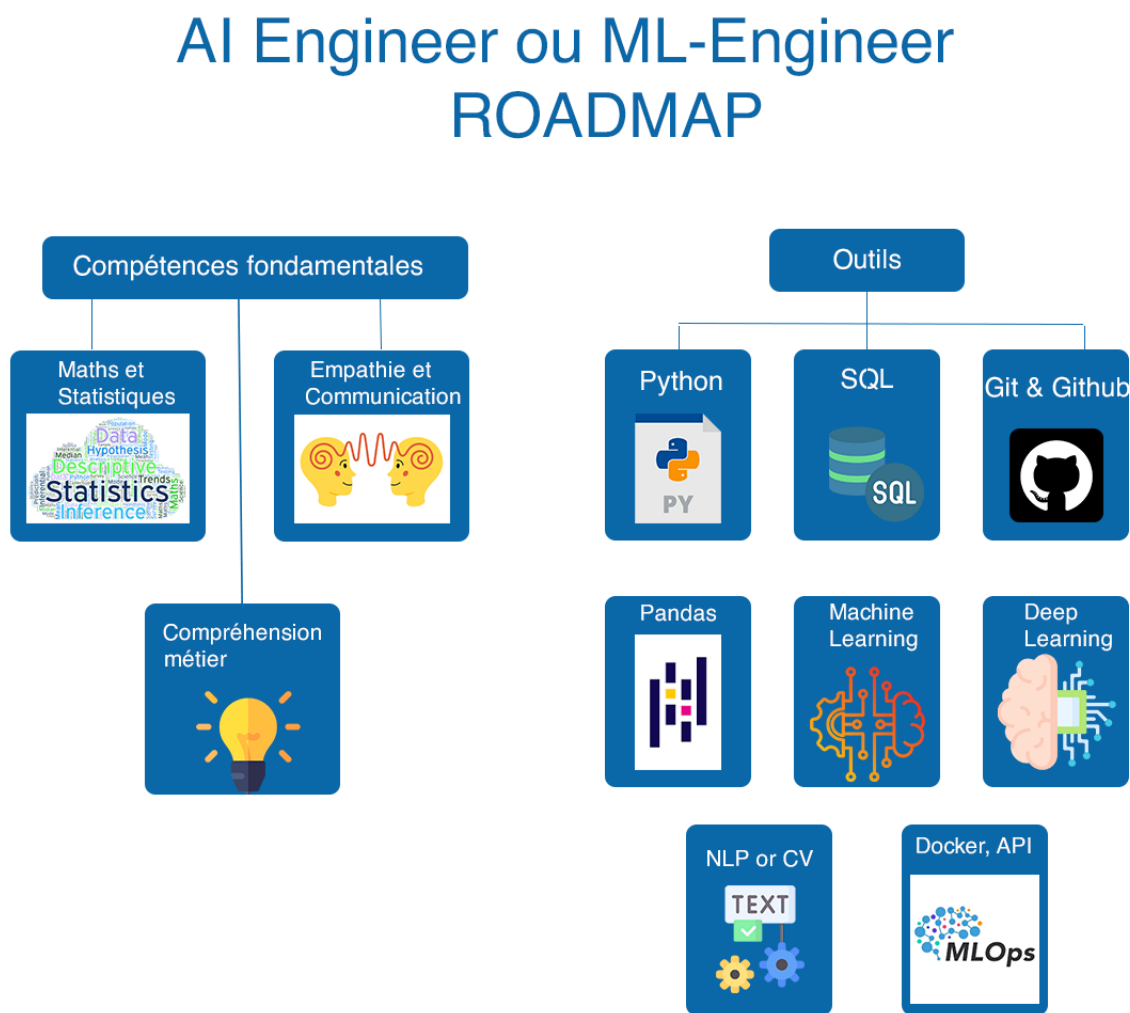


Figure 1: Caption

1.1 Semaine 1-3 : Programmation Python

- Les basiques en Python, les variables, opérateurs, conditions
- Les boucles,
- Les fonctions,
- Les dictionnaires, tuple, Dictionary, Tuple, Set,
- Gérer les exceptions,
- POO (programmation oriente objet),
- Les listes (comprehension de liste...) et les strings,
- Les modules et packages.

1.2 Structure de Données et les algorithmes en Python:

1.2.1 Data Structures

- Stack,
- Queue,
- Linked List,
- Tree,
- Graph.

1.3 Semaine 4-5 : SQL, Git and GitHub

- Principes fondamentaux de la syntaxe et de l'installation de SQL
- Création de tables, modificateurs
- Insérer et récupérer des données, SELECT INSERT UPDATE DELETE
- Agrégation de données à l'aide de fonctions, filtrage et RegEX
- Sous-requêtes, récupération de données basées sur des conditions, regroupement de données.
- Questions pratiques
- JOINS
- Concepts SQL avancés tels que les transactions, les vues, les procédures stockées et les fonctions.
- Principes de conception des bases de données, normalisation et diagrammes ER.
- Pratiquez : Entraînez vous à écrire des requêtes SQL sur des ensembles de données réels et travaillez sur des projets pour appliquer vos connaissances.

1.4 Semaine 6 : Maths et Statistiques

1. Notions de base : Statistiques descriptives et inférentielles,
2. Algèbre linéaire : Vecteurs, métriques, valeurs propres et vecteurs propres,
3. Calcul : Notions de calcul intégral et différentiel,
4. Graphiques : Histogrammes, camemberts, diagrammes en bâtons, diagrammes de dispersion,
5. Mesures de tendance centrale : moyenne, médiane, mode,
6. Mesures de dispersion : variance, écart-type,
7. Distributions : Distribution normale,
8. Corrélation et covariance,
9. Tests d'hypothèses : valeur p, intervalle de confiance, erreur de type 1 vs type 2, Test Z.

1.5 Semaine 7 : EDA

L'analyse exploratoire des données (Exploratory Data Analysis - EDA) est une étape cruciale dans le processus de développement de modèles en intelligence artificielle et en apprentissage automatique. Elle permet de comprendre la structure, les distributions et les relations présentes dans les données. Nous allons utiliser des bibliothèques populaires comme pandas, numpy, matplotlib, et seaborn.

1.6 Semaine 8 : Preprocessing

- Handling NA values, outlier treatment, data normalization,
- One hot encoding, label encoding,
- Feature engineering,
- Train test split,
- Cross validation.

1.7 Semaine 9 : Algorithme de Machine Learning

1. Regression Linéaire,
2. Regression Logistique,
3. Decision Tree,
4. Gradient Descendant,
5. Random Forest,
6. Regression Ridge et Lasso,
7. Naive Bayes,
8. Support Vector Machine,
9. KMeans Clustering.

1.8 Semaine 10-11: MLOPS

- FastAPI,
- Docker.

1.9 Semaine 12-14 : Natural Language Processing

- Extraction de données (scrapping, extraction de données des pdf, des json..),
- Exploration et pre-traitements des données (convertir les données en caractere miniscule, suppression de la ponctiation, suppression des stop words, tokenization, stemmatisation, lemmatization,..),
- Conversion de texte en features (count vectorizer, One hot encoder,..),
- Trouver la similarité entre des textes,
- Speech-to-text,
- text-to-speech,
- Génération de texte,
- Auto-completion,
- Résumé automatique de texte,
- Clustering de documents.

References