

Skill Radar : Analyse et prévision de l'évolution des compétences Data Science & IA dans les offres d'emploi

Préparé par :

Boulaalam Yassine
Bouskine Othmane
Bouhafa Taha

Encadré par:

Prof. Imad Sassi

Le : 27/05/2025

Sommaire

1

Introduction

4

Implementation des modèles

2

Contexte et enjeux

5

Visualisation

3

Données et méthodes

6

Automatisation

7

Conclusion

Introduction

L'IA et la Data Science redéfinissent le marché de l'emploi technologique

Notre projet répond à ce défi par :

- Une analyse en temps réel des besoins du marché
- Des modèles prédictifs sur les compétences d'avenir
- Un outil opérationnel pour les organismes de formation"

Big Data And Artificial Intelligence Market

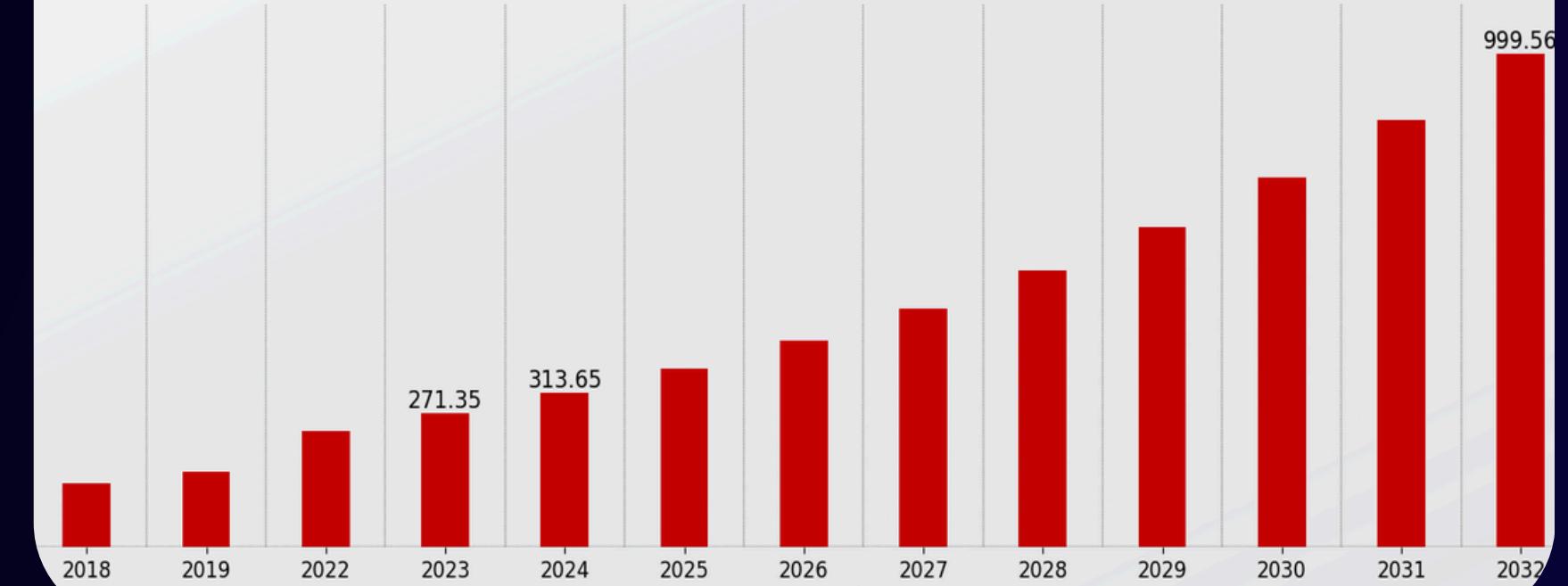


Figure : Croissance fulgurante du marché Big Data & IA : une transformation inévitable
(market research future)

Ce graphique illustre l'explosion du marché IA/Data justifiant notre étude :

- Un secteur multiplié par 4 en 10 ans (2018-2032)
- Une croissance qui valide l'urgence à maîtriser ces technologies
- Contexte idéal pour notre analyse des compétences stratégiques.

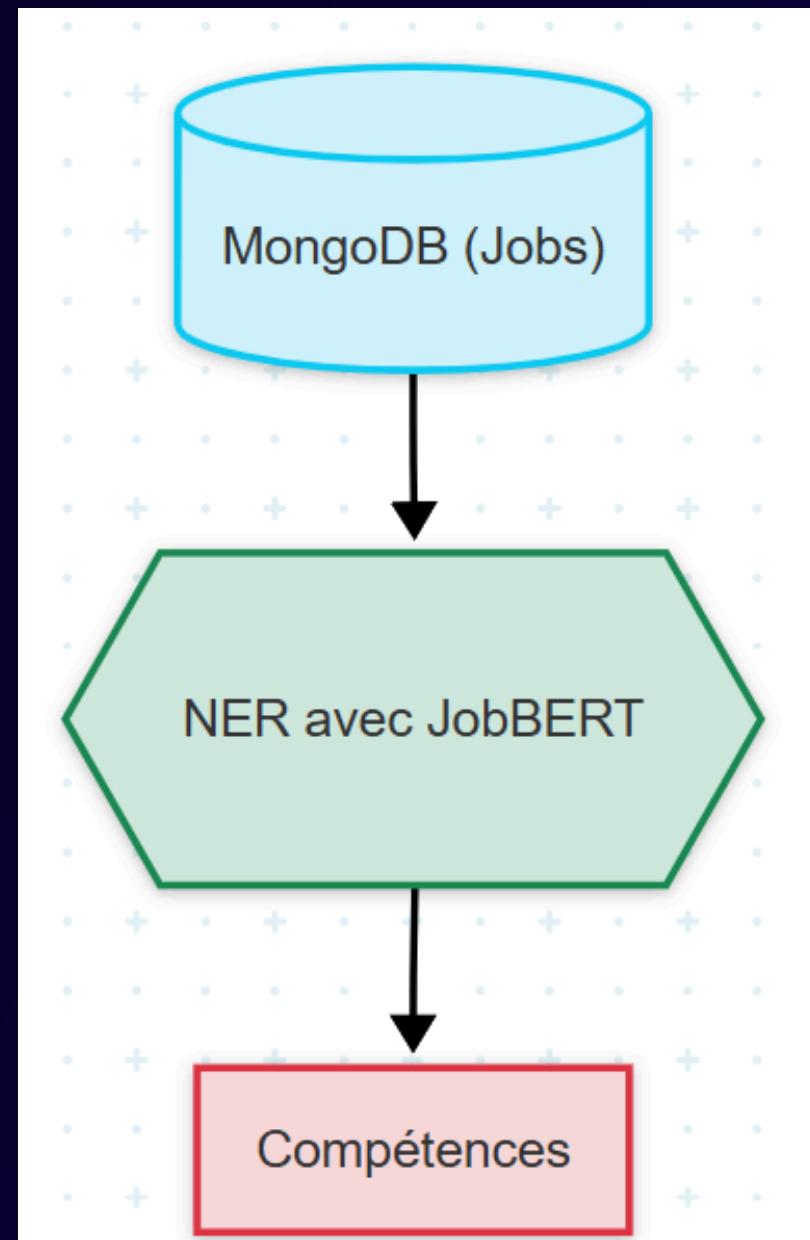
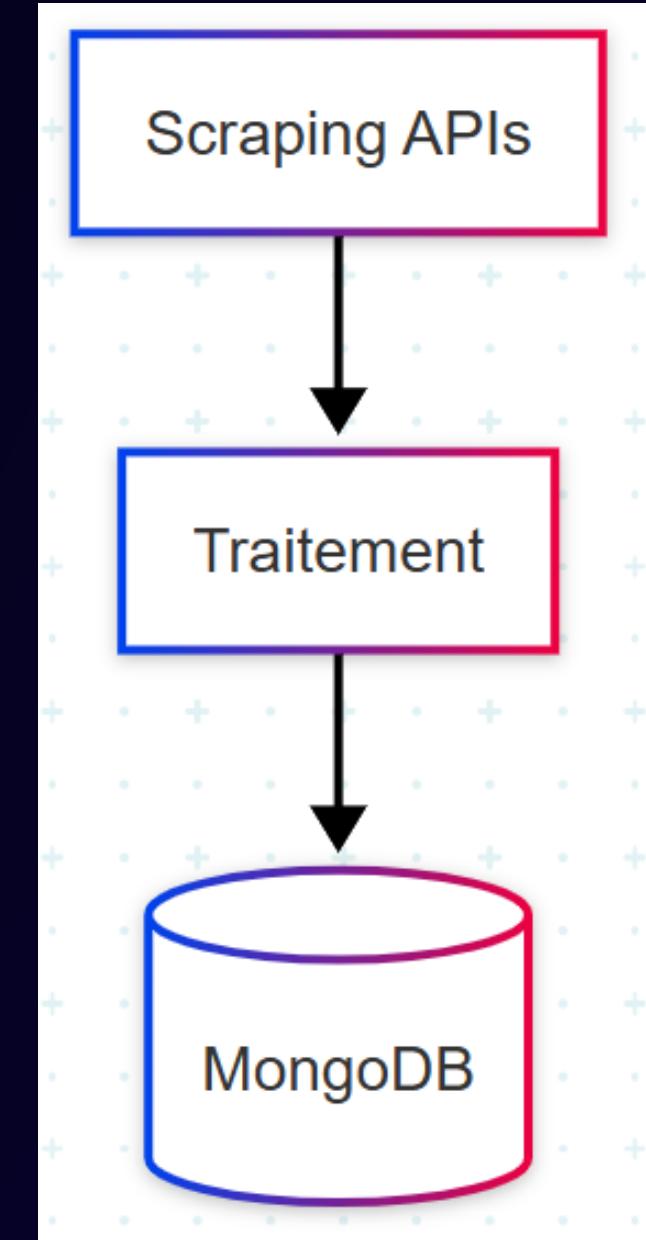
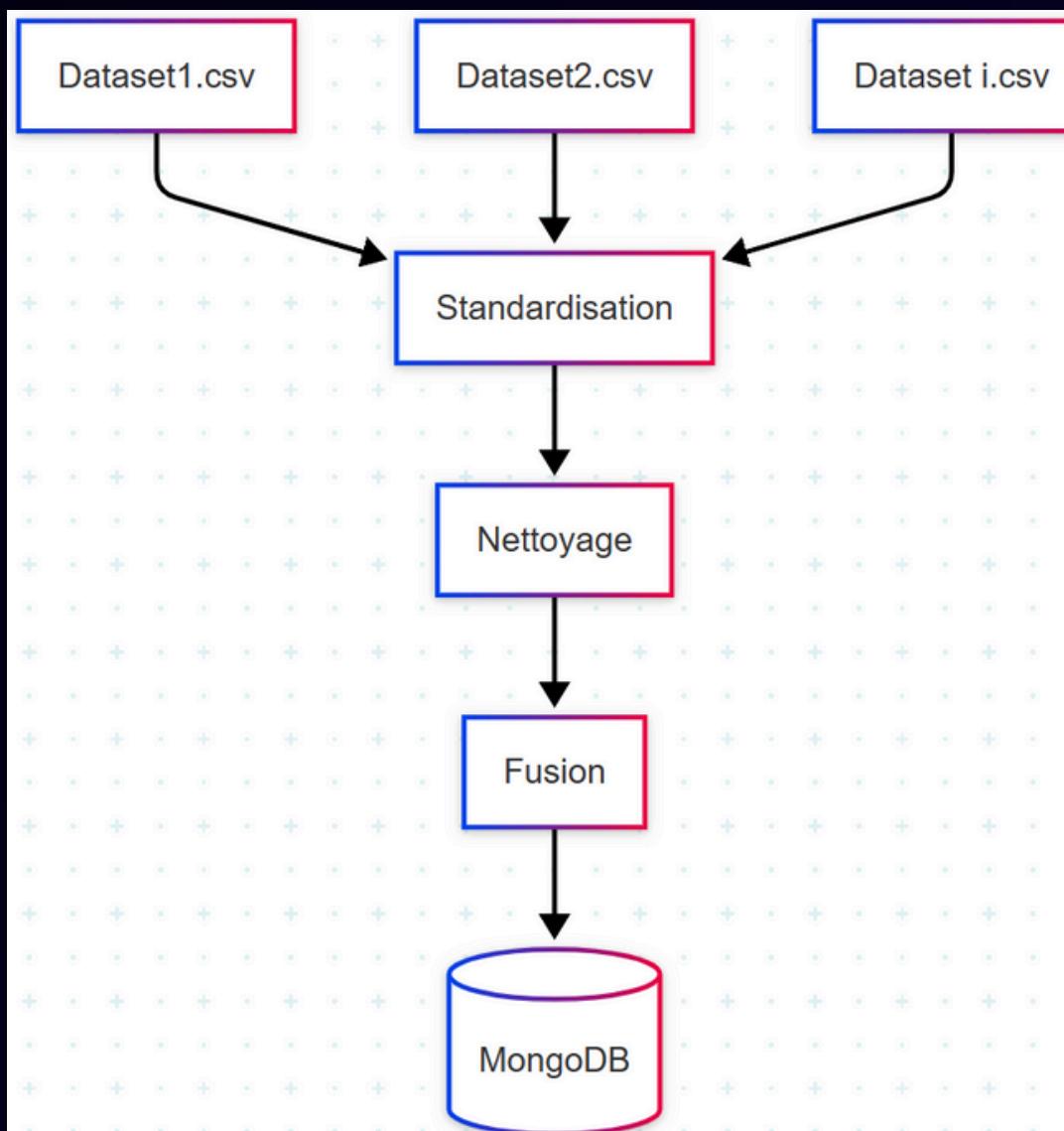
Contexte & Enjeux

Points clés :

- L'évolution accélérée des technologies IA et Data Science
- Le décalage croissant entre formations et besoins réels des entreprises
- La difficulté à anticiper les compétences stratégiques pour demain

Données et méthodes

1. Sources de données
2. Scraping APIs
3. Pré-traitement
4. Base de données
5. Analyse NLP



Implementation des modeles

Estimation des salaires

Prédire les salaires selon poste/compétences pour aider les grilles salariales

Modèle 1: XGBoost

Configuration :

- Paramètres optimaux trouvés : best_params = { 'Learning_rate': 0.1 , 'max_depth': 9, 'n_estimators': 100 , 'subsample': 0.7 , 'random_state': 42 }

Modèle 2: Deep Learning

Architecture: model = keras.Sequential([layers.Dense(256, activation='relu', input_shape=(X_train_scaled.shape[1],)), layers.Dropout(0.3), layers.Dense(128, activation='relu'), layers.Dropout(0.3), layers.Dense(64, activation='relu'), layers.Dense(1)])

Performances :

Critère	XGBoost	Deep Learning
Précision (R^2)	0.29	0.98

Implementation des modèles

Recommandation de compétences

Recommander des compétences complémentaires basées sur les compétences déjà acquises

Modèle 1 : Régression Logistique

Approche : Régression logistique multi-label, utilisant la stratégie One-vs-All

Modèle 2: k-NN

Approche : Recommandation basée sur la cosine similarity.

Modèle 3:Réseau de Neurones (Deep Learning)

- Architecture : Couches denses ($512 \rightarrow 256 \rightarrow 252$ neurones) avec Dropout pour éviter le surapprentissage. Fonction d'activation ReLU (couches cachées) et sigmoïde (couche de sortie).

Performances :

Critère	Régression Logistique	K-NN	Deep Learning
Précision	0,97	0,86	0,98
Rappel	0,84	0,79	0,84
F1-score	0,90	0,82	0,90

Implementation des modèles

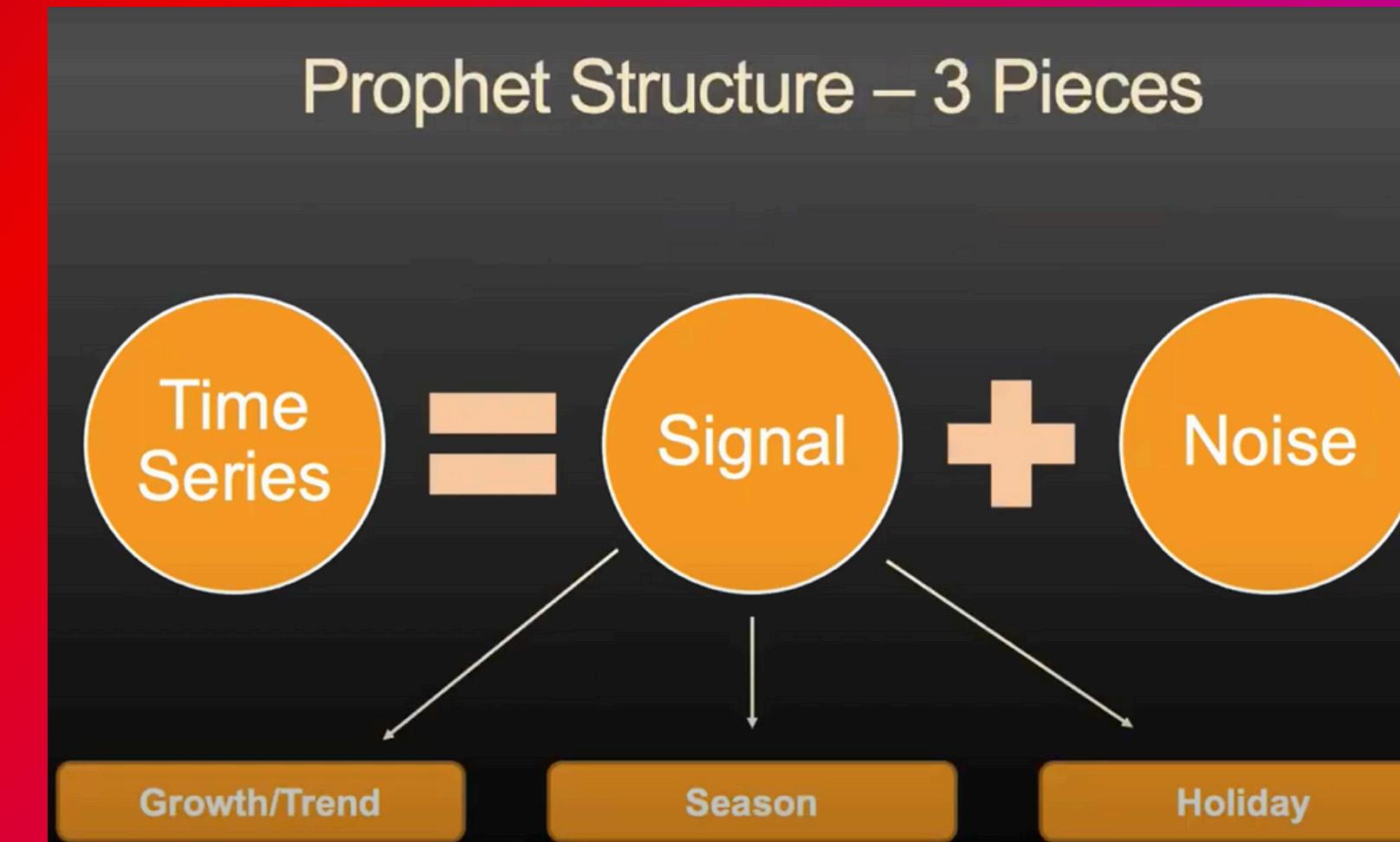
Prévision de demande

Prédire la demande future pour les compétences niche

Préparation des données:
Filtrage intelligent :
250+ compétences retenues via 3 critères :

- Volume
- Potentiel
- Nettoyage

- Meilleur RMSE moyen (42 pour ARIMA vs 94 pour PROPHET).
- Gère bien les spécificités des compétences niche.



Visualisation

Ce tableau de bord interactif a pour objectifs de :

- **Explorer les tendances de l'emploi dans les domaines de la data science et de l'IA selon le titre de poste, l'entreprise, le pays, et la période.**
- **Prévoir la demande de compétences techniques dans le temps.**
- **Recommander des compétences pertinentes à un utilisateur selon son profil.**
- **Estimer les salaires en fonction du poste et des compétences détenues.**

dashboard

Predictions

i About

This dashboard helps you:

 Explore job trends by title, company, country, and more

 Forecast demand for tech skills over time

 Recommend relevant skills based on your current profile

 Estimate salaries by job title and skillset

Ideal for job seekers, data professionals, and recruiters tracking the evolution of AI and data roles.

Skill Radar: Trends in Data Science & AI Jobs

• Top Job Titles

Top Job Titles by Number of Offers



Job Offers by Country

Number of Job Offers per Country



[dashboard](#)[Predictions](#)

About

This dashboard helps you:

 Explore job trends by title, company, country, and more

 Forecast demand for tech skills over time

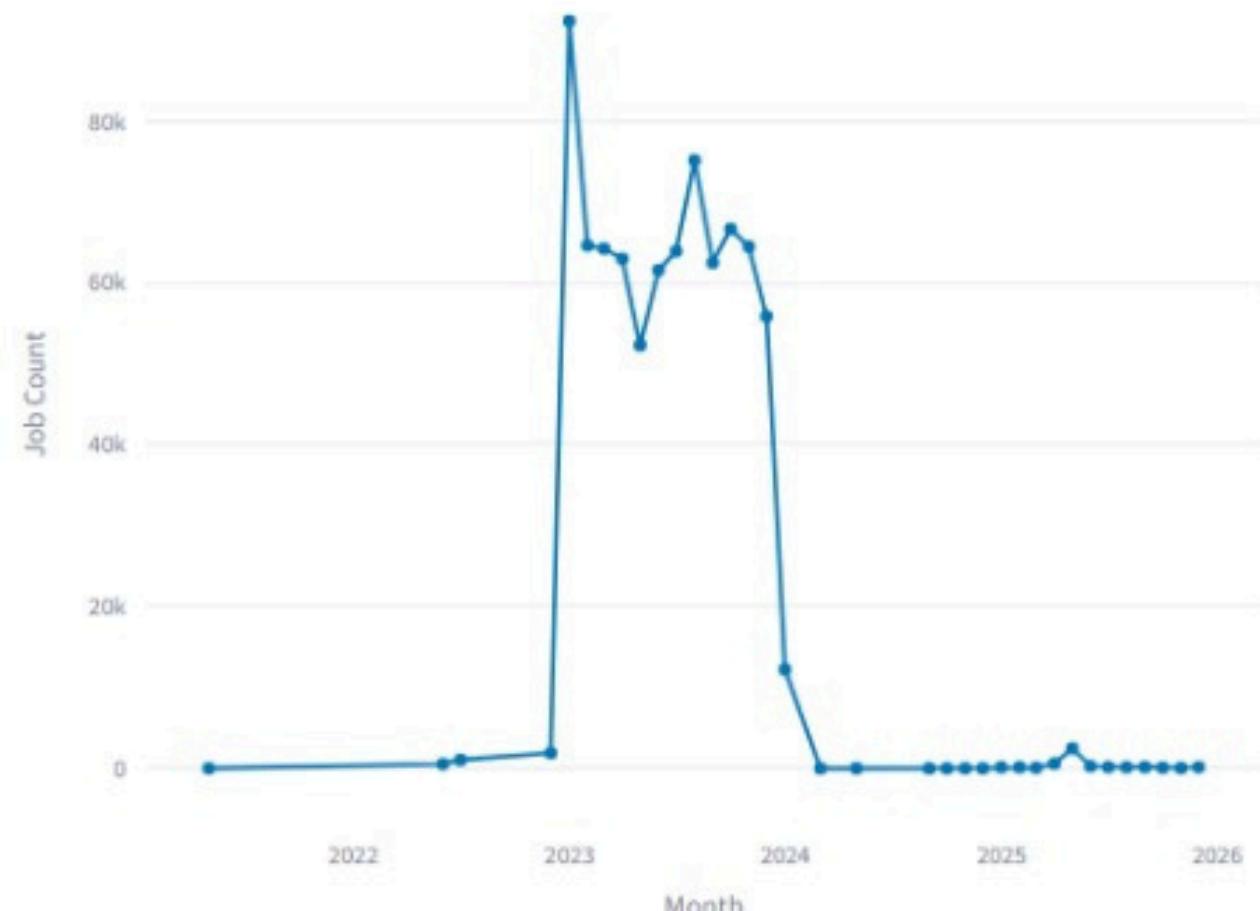
 Recommend relevant skills based on your current profile

 Estimate salaries by job title and skillset

Ideal for job seekers, data professionals, and recruiters tracking the evolution of AI and data roles.

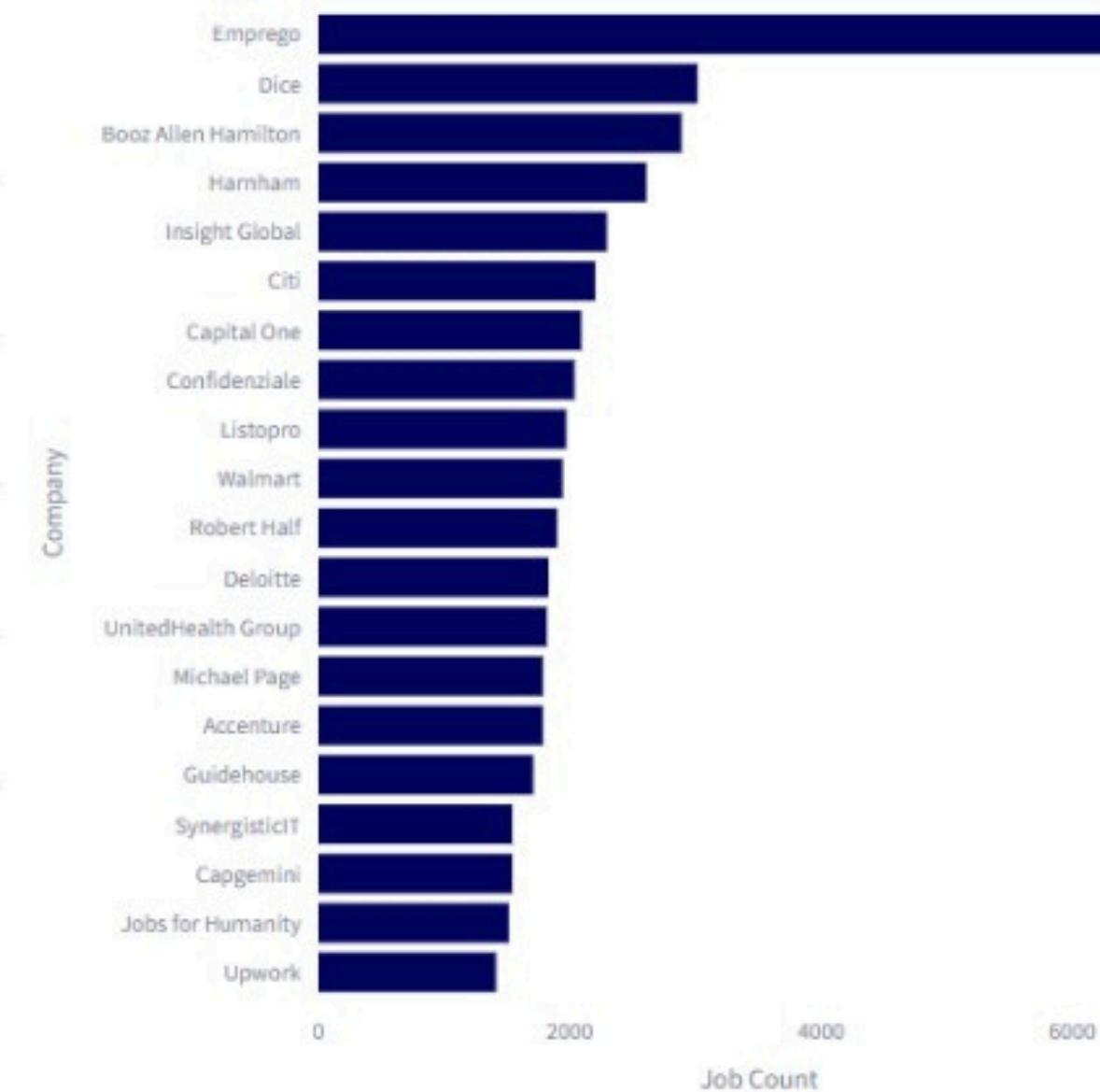
Job Offers Over Time

Number of Job Offers per Month



Top Companies by Job Offers

Top Companies by Number of Job Offers



[dashboard](#)
[Predictions](#)

About

This dashboard helps you:

 Explore job trends by title, company, country, and more

 Forecast demand for tech skills over time

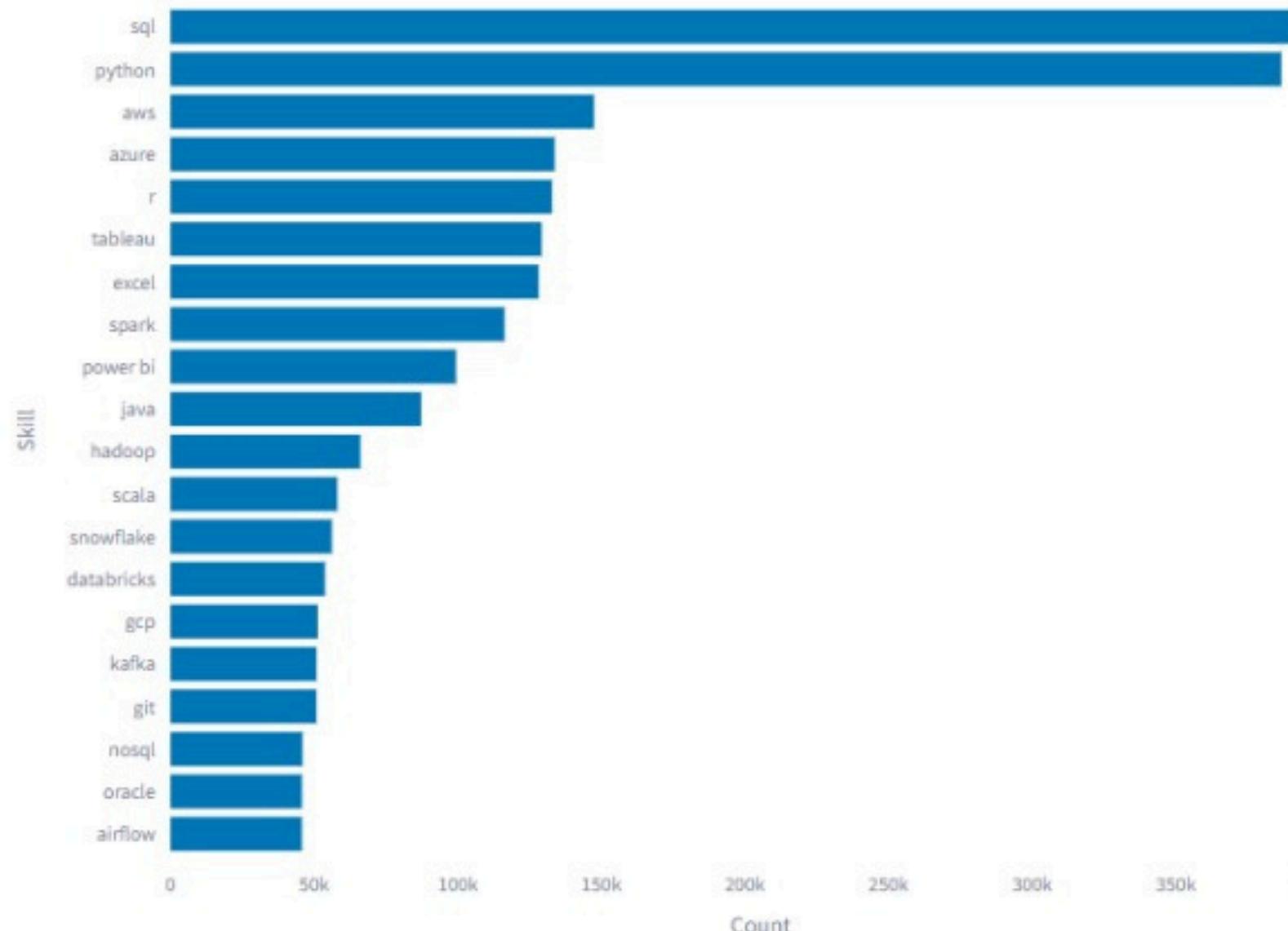
 Recommend relevant skills based on your current profile

 Estimate salaries by job title and skillset

Ideal for job seekers, data professionals, and recruiters tracking the evolution of AI and data roles.

Top In-Demand Skills

Top Requested Skills

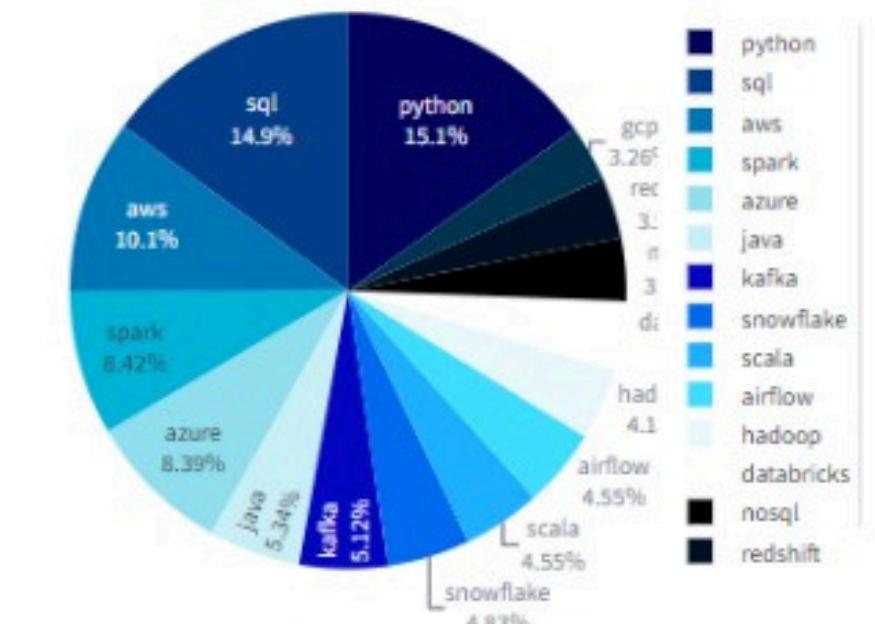


• Skill Distribution per Job Title

Select a Job Title

Senior Data Engineer

Skill Distribution for Senior Data Engineer





dashboard

Predictions

Select One Prediction Tool

Choose the type of prediction you want to perform:

Skill Forecast

Skill Recommendation

Salary Estimation

Deploy 

Skill Recommendation

Select your known skills

python x pytorch x angular x 

Number of skills to recommend
5 
3 20

 Recommend

Recommended skills:

- tensorflow
- sql
- keras
- nltk
- java

dashboard

Predictions

 Select One Prediction Tool

Choose the type of prediction you want to perform:

Skill Forecast

 Skill Recommendation

 Salary Estimation

Deploy 

Prediction Center

Salary Estimation

 Select job title

Machine Learning Engineer 

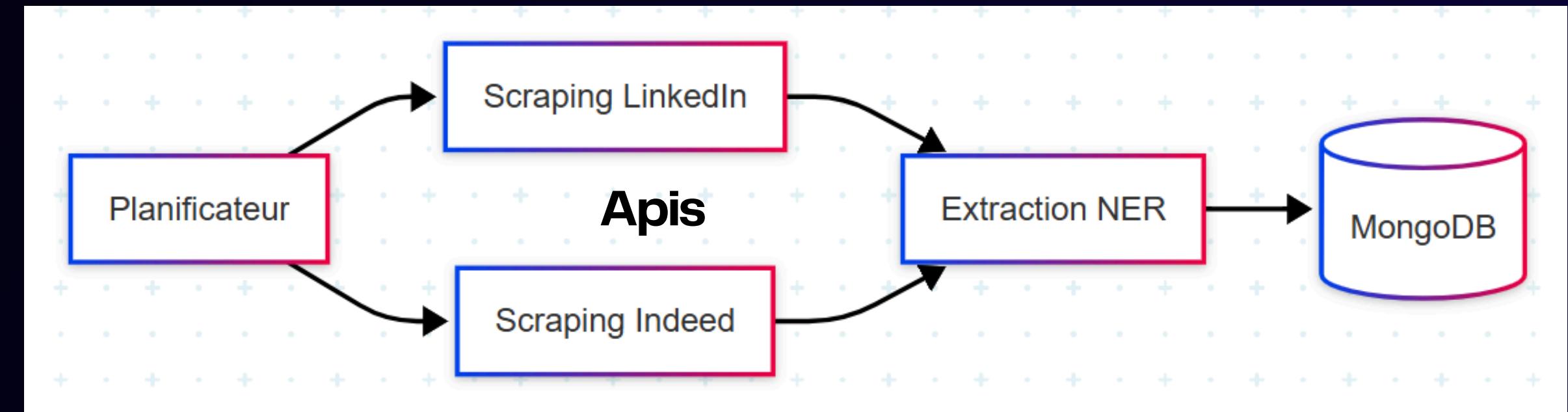
 Select skills

python  scikit-learn  ansible  

 Estimate Salary

 Estimated Salary: 52,598.48 €

Automatisation du pipeline



Nous avons automatisé notre pipeline de collecte et de traitement des données avec une orchestration quotidienne à 18h45 via la bibliothèque schedule. Ce système allie une extensibilité (grâce à une architecture modulaire permettant d'intégrer facilement de nouveaux scripts de scraping ou d'analyse). Cette solution garantit une mise à jour régulière de nos données tout en minimisant l'intervention manuelle.

Conclusion

- +700 000 offres analysées (LinkedIn, Indeed, Kaggle)
- Extraction automatique des compétences via JobBERT
- 3 modèles développés : estimation salariale, recommandation, prévision des compétences
- Intégration dans un tableau de bord interactif (Streamlit)
- Pour les professionnels : visualiser les compétences stratégiques & estimer leur valeur
- Pour les entreprises : anticiper les besoins en recrutement de manière proactive
- Réduction du travail manuel grâce à l'automatisation complète du pipeline



"Skill Radar transforme la veille marché en insights actionnables, pour une adaptation continue aux mutations de l'IA."

**Thank You
For Your Attention**