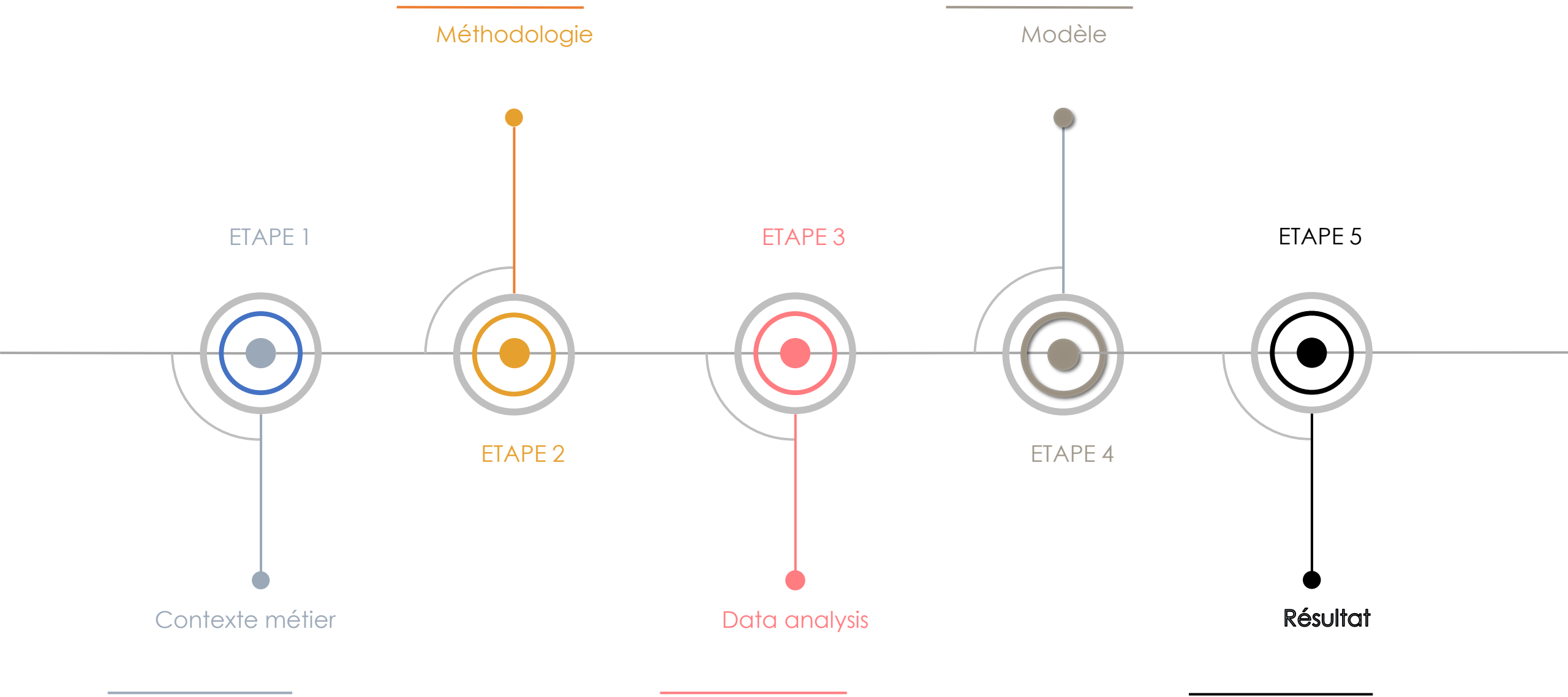




Mise en place d'un modèle de prevision du prix de maison a seattle

PÂQUARSE DELVICH VAN MAHOVI



Problématique (1 / 2)

- Les prix de l'immobilier ont progressé en moyenne de 4,5 % par an depuis 20 ans, soit un niveau bien supérieur à l'inflation et aux revenus des ménages. (Arnaud Sylvain, 2021).
- Monde en 2023 → Sortir d'une crise sanitaire et une guerre en Europe
- Incertitude sur les marchés financiers dont l'immobilier
- Outil pour estimer le prix d'une maison serait un vrai **gagne-temps**

Problématique (2/2)

1

Les acheteurs et les vendeurs peuvent utiliser les prévisions de prix pour prendre des décisions informées sur l'achat ou la vente d'une propriété

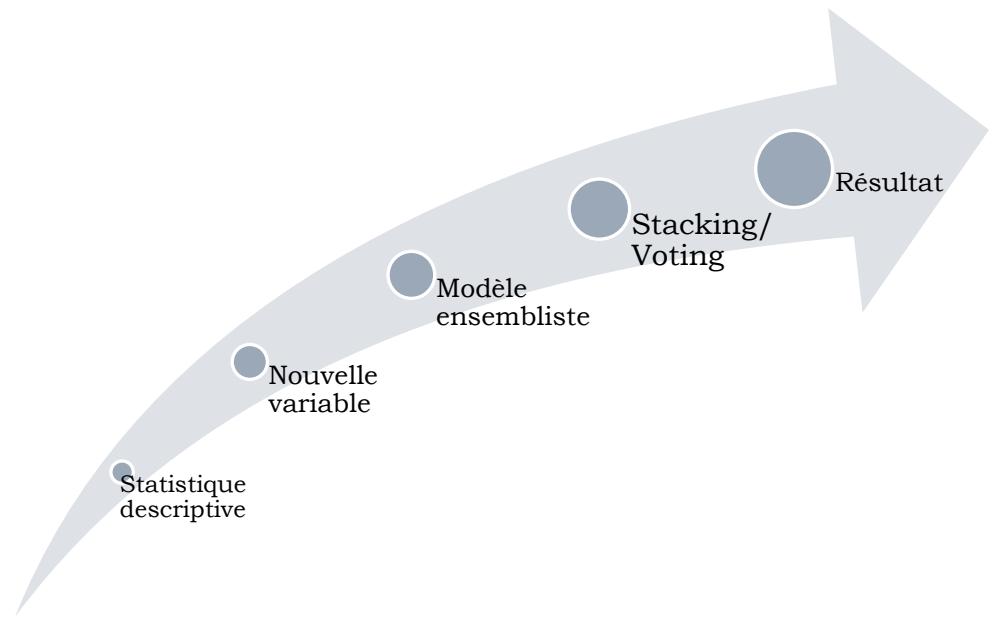
Pour les institutions financières : Les prévisions de prix de l'immobilier sont également importantes pour les institutions financières qui prêtent de l'argent pour l'achat de propriétés

2

3

Les investisseurs immobiliers utilisent les prévisions de prix pour évaluer les risques et les rendements potentiels des investissements immobiliers.

Méthodologie



- ❖ Statistique descriptive : Histogramme...
- ❖ Nouvelle variable : variable recodée, variable crée...
- ❖ Modèle ensembliste : xgboost...
- ❖ Stacking : Technique de stacking...

Data Analysis

QUELQUES NOUVELLES VARIABLES CRÉES ...

- ☐ Part de la maison sur le terrain ;
- ☐ Part du logement au dessus du sol ;
- ☐ Distance par rapport au centre ville

QUELQUES VARIABLES RECODÉES

- ☐ La date de la dernière rénovation ;
- ☐ La date à laquelle la maison à été vendue

Modèles

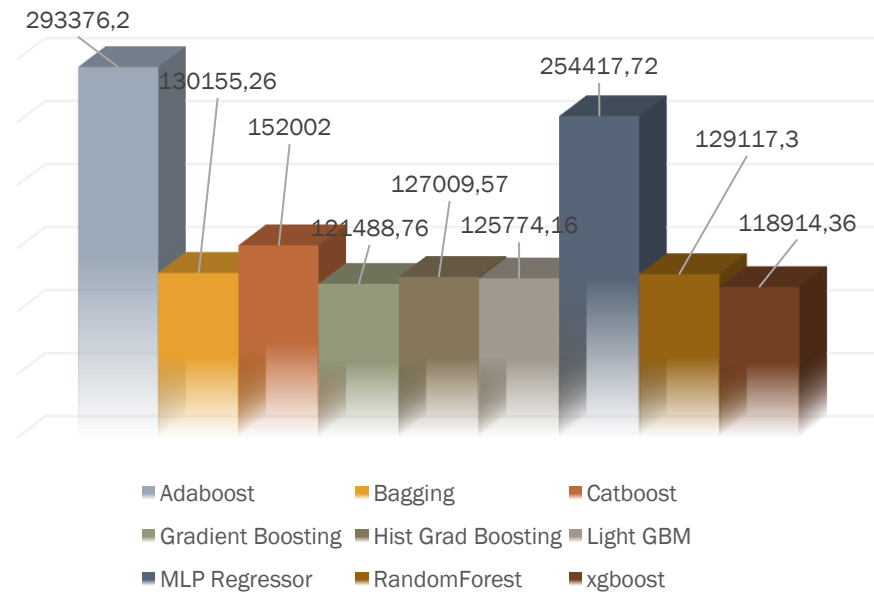
- ❖ Xgboost
- ❖ RandomForest ;
- ❖ Adaboost ;
- ❖ Bagging
- ❖ Gradient Boosting
- ❖ Histogram Gradient Boosting
- ❖ Light GBM
- ❖ Catboost
- ❖ MLP Regresspr

Modèles de stacking

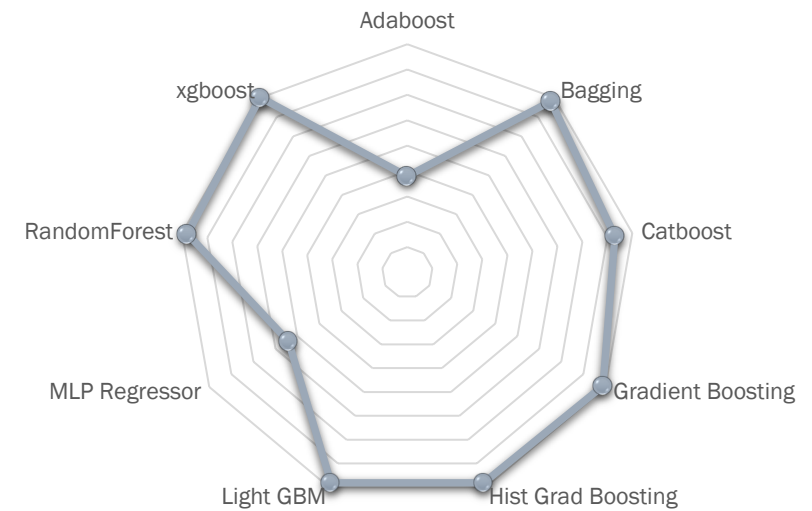
- ✓ Stacking Regressor
- ✓ Voting Regressor

Résultat

RMSE

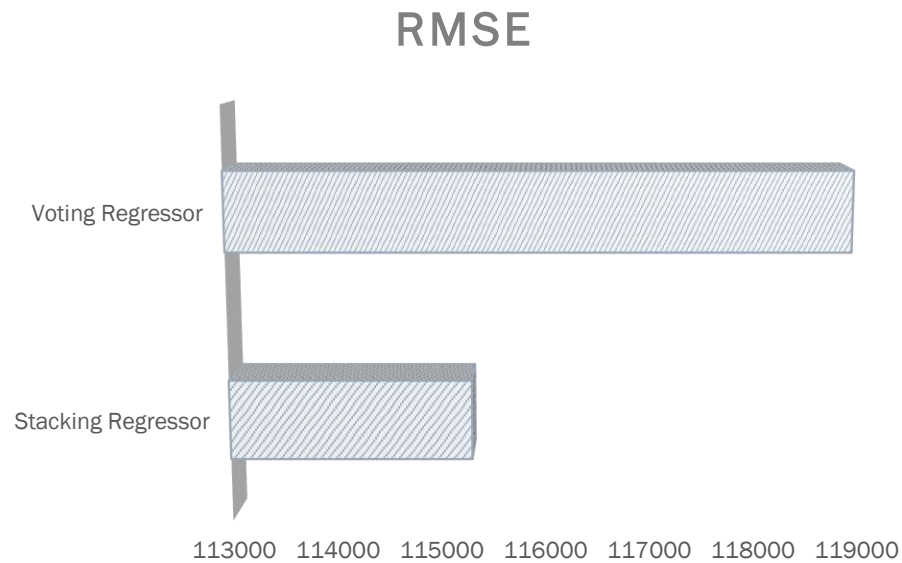


R2 SCORE



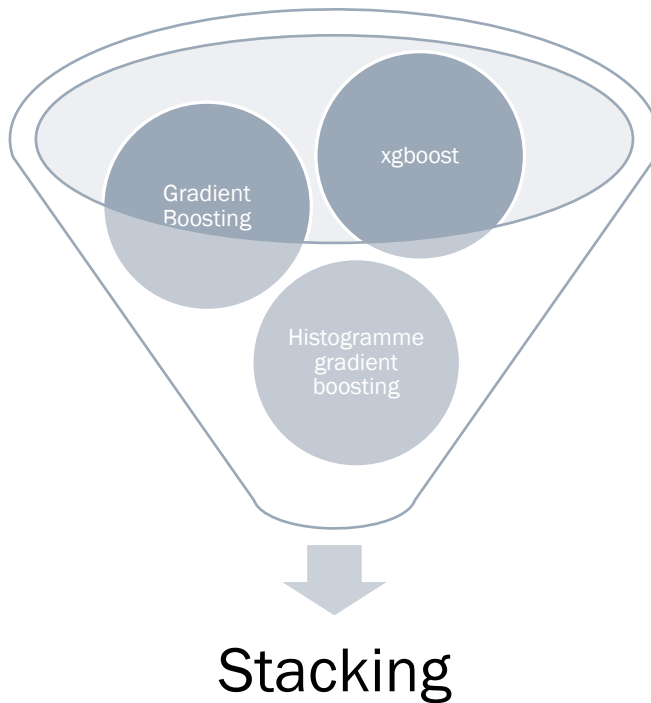
Résultat : Stacking VS Voting

RMSE



➤ Stabilité du stacking Regressor

Modèles : Modèle final



POURQUOI LE STACKING ?

- ☐ Améliorer les performances de prédiction par rapport à un modèle unique
- ☐ C'est une technique puissante et performante pour de nombreux problèmes de prédiction ;
- ☐ Aide à améliorer la stabilité des résultats ;
- ☐ Réduction du surapprentissage ;
- ☐ Plus efficace en MLOps (apparemment...)

Résultat : Modèle de Stacking

±
306

Base	RMSE
Echantillon test	115.306
Echantillon validation (prof.)	115000

Stabilité des résultats pour les professionnels de l'immobilier

Pour aller plus loin ...

Pour aller plus loin...

FORCES

- ❖ Le tour des modèles ensemblistes optimisés ;
- ❖ L'initiation aux modèles de stacking ;
- ❖ Stabilité des réponses du modèles ;
- ❖ Réduction de l'aléatoire ;
- ❖ Application des réseaux de neurones sur des thématiques de régression/classification

FAIBLESSES

- Intégré plus d'indicateur (Variables tel que la proximité des transports en commun, les commerces...)
- Faire une revue plus complète des facteurs influençant l'achat de maison par les particuliers
- Une meilleur connaissance des indicateurs des agences immobilières
- Une meilleure compréhension de la mentalité américaine

Difficulté principale : sortir des zones plateaux pour la RMSE

Merci...

DES QUESTIONS ?