

**RAPPORT DE PROJET :**

# « Prédiction des prix de l'immobilier »

*Projet réalisé par :*

Oumaima BAIROUK  
Khadija BENDIB

*Projet encadré par :*

Chuan Xu

## REMERCIEMENTS

Tout d'abord, nous tenons à remercier tout particulièrement et à témoigner toute notre reconnaissance notre professeur, pour leur dévouement et leur soutien dans la concrétisation de ce projet et pour nous avoir accordé toute la confiance nécessaire pour élaborer ce projet librement, et avoir mis à notre disposition tous les moyens disponibles.

## Table des matières :

<a href="#">1.Introduction de Prédiction des prix de l'immobilier :</a>	4
<a href="#">2.Introduction de jeu des données sur Kaggle :</a>	5
<a href="#">✓ Les variables</a>	5
<a href="#">3.Méthode appliquée</a>	7
<a href="#">✓Prétraitement de données</a>	7
<a href="#">✓Structure de modèle :</a>	8
<a href="#">✓Fonction de perte :</a>	8
<a href="#">✓Méthode d'optimisation</a>	8
<a href="#">✓Evaluation de modèle IA</a>	8
<a href="#">4.Conclusion :</a>	9

## Les figure :

<a href="#"><u>Figure 1 : les données de fichier train</u></a> .....	7
<a href="#"><u>Figure 2:les données nettoyées</u></a> .....	8
<a href="#"><u>Figure 3:La fonction de perte</u></a> .....	8
<a href="#"><u>Figure 4:Le résultat sur Kaggle(Score)</u></a> .....	9

## **1. Introduction de Prédiction des prix de l'immobilier :**

Après 1,1 million de ventes estimées pour 2022, le volume devrait tomber à 950 000 unités sur les douze prochains mois, avec une progression moyenne des prix de 3%. Et derrière une moyenne nationale se cachent irrémédiablement des disparités régionales.

Dans ce concours, il est demandé de prédire des prix de l'immobilier, à partir de variables strictement.

Si comme on demande à un acheteur de maison de décrire la maison de ses rêves, et il ne commencera probablement pas par la hauteur du plafond du sous-sol ou la proximité d'une voie ferrée est-ouest. Mais l'ensemble de données de ce concours de terrain de jeu prouve que beaucoup plus influence les négociations de prix que le nombre de chambres ou une clôture blanche.

## 2. Introduction de jeu des données sur Kaggle :

Avec 79 variables explicatives décrivant (presque) tous les aspects des maisons résidentielles à Ames, Iowa, ce concours vous met au défi de prédire le prix final de chaque maison :

### 1. Les variables

- SalePrice : le prix de vente de la propriété en dollars il s'agit de la variable que nous essayer de prédire.
- MSSubClass : La classe bâtiment
- MSZoning : Le classement général de zonage
- LotFrontage : Pieds linéaires de rue reliés à la propriété
- LotArea : Taille du terrain en pieds carrés
- Street : Type d'accès routier
- Alley : Type d'accès ruelle
- LotShape : Forme générale de la propriété
- LandContour : Planéité du bien
- Utilities : Type d'utilitaires disponibles
- LotConfig : Configuration des lots
- LandSlope : Pente de la propriété
- Neighborhood : Emplacements physiques dans les limites de la ville d'Ames
- Condition1 : Proximité d'une route principale ou d'une voie ferrée
- Condition2 : Proximité d'une route principale ou d'une voie ferrée (si une seconde est présente)
- BldgType : Type de logement
- HouseStyle : Style d'habitation
- OverallQual : Qualité générale des matériaux et de la finition
- OverallCond : Évaluation de l'état général
- YearBuilt : Date de construction originale
- YearRemodAdd : Date de rénovation
- RoofStyle : Type de toit
- RoofMatl : Matériau du toit
- Exterior1st : Revêtement extérieur sur maison
- Exterior2nd : Revêtement extérieur de la maison (si plus d'un matériau)
- MasVnrType : Type de placage de maçonnerie
- MasVnArea : Surface de placage de maçonnerie en pieds carrés
- ExterQual : Qualité des matériaux extérieurs
- ExterCond : État actuel du matériau à l'extérieur
- Foudation : Type de fondation
- BsmtQual : Hauteur du sous-sol
- BsmtCond : État général du sous-sol
- BsmtExposure : Murs de sous-sol au rez-de-chaussée ou au rez-de-jardin

- BsmtFindType1 : Qualité de la superficie finie du sous-sol
- BsmtFinSF1 : Type 1 pieds carrés finis
- BsmtFinType2 : Qualité de la deuxième surface finie (si présente)
- BsmtUnfSF2 : Type 2 pieds carrés finis
- BsmtUnfSF : Pieds carrés non aménagés du sous-sol
- TotalBsmtSF : Total de pieds carrés de sous-sol
- Heating : Type de chauffage
- HeatingQC : Qualité et état du chauffage
- CentralAir : Climatisation centrale
- Electrical : Système électrique
- 1stFlrSF : Pieds carrés du premier étage
- 2ndFlrSF : Pieds carrés au deuxième étage
- LowQualFinSF : Pieds carrés finis de mauvaise qualité (tous les étages)
- GrLivArea : Pieds carrés de surface habitable au-dessus du niveau du sol (au sol)
- BsmtFullBath : Salles de bain complètes au sous-sol
- BsmtHalfBath : Demi-salles de bain au sous-sol
- FullBath : Salles de bain complètes au-dessus du sol
- HalBath : Demi-bains au-dessus du sol
- Bedroom : Nombre de chambres au-dessus du sous-sol
- Kitchen : Nombre de cuisines
- KitchenQual : Qualité de la cuisine
- TotRmsAbvGrd : Nombre total de chambres au-dessus du niveau du sol (n'inclut pas les salles de bain)
- Functional : Évaluation de la fonctionnalité de la maison
- Fireplaces : Nombre de cheminées
- FireplaceQu : Qualité cheminée
- GarageType : Emplacement du garage
- GarageYrBlt : Année de construction du garage
- GarageFinish : Finition intérieure du garage
- GarageCars : Taille du garage en capacité de voiture
- GarageArea : Taille du garage en pieds carrés
- GarageQual : Qualité garage
- GarageCond : État des garages
- PaveDrive : Allée pavée
- WoodDeckSF : Superficie de la terrasse en bois en pieds carrés
- OpenPorchSF : Porche ouvert en pieds carrés
- EnclosedPorch : Porche fermé en pieds carrés
- 3SsnPorch : Superficie du porche trois saisons en pieds carrés
- ScreenPorch : Superficie du porche moustiquaire en pieds carrés
- PoolArea : Superficie de la piscine en pieds carrés
- PoolQC : Qualité de la piscine
- Fence : Qualité de la clôture

- MiscFeature : Fonctionnalité diverse non couverte dans les autres catégories
- MiscVal : \$Valeur de la fonctionnalité diverse
- MoSold : Mois vendu
- YrSold : Ans
- SaleType : Type de vente
- SaleCondition : Condition de vente

### 3. Méthode appliquée

#### 1. Prétraitement de données

Voici notre jeu de données avant le traitement :

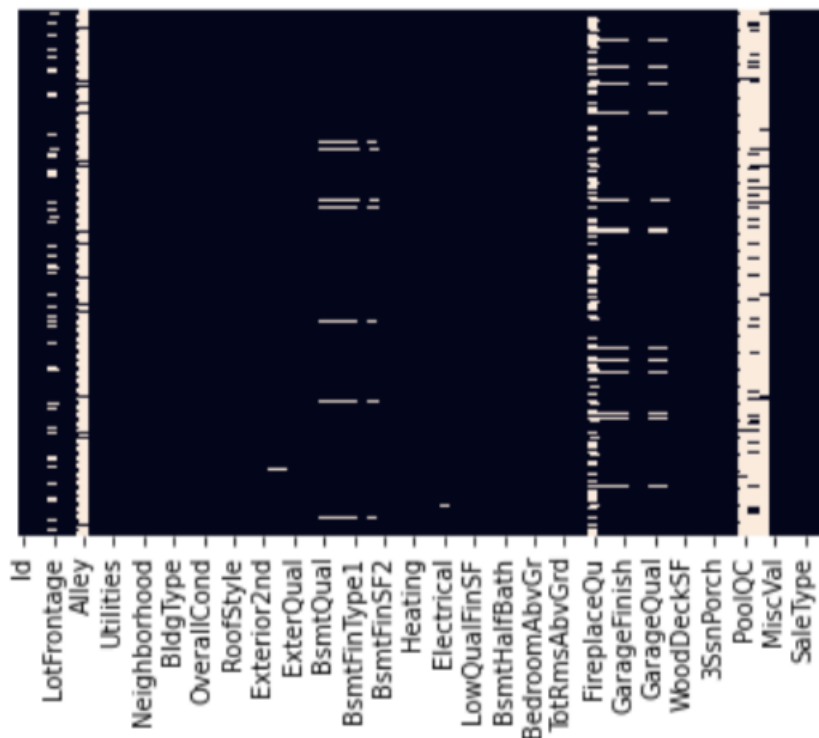


Figure 1 : les données de fichier train

- ❖ Les valeurs manquantes des variables quantitatives sont remplacées par la moyenne de la colonne.
- ❖ Nous avons supprimé les colonnes qui contiennent seulement des valeurs manquantes.
- ❖ Concernant les variables qualitatives, dans un premier temps on les a remplacées par la valeur dominante mais après nous avons constaté qu'on n'a pas besoin de ces valeurs, on va travailler justement sur les valeurs quantitatives par la suite nous avons supprimé les valeurs qualitatives.



Nous avons utilisé l'algorithme d'optimisation Adam car il est le meilleur optimizer en IA et il regroupe deux méthodes Momentum + RMSprop.

## 5. Evaluation de modèle IA

Dans cette figure, on constate qu'il y'a une diminution de la fonction de perte, alors on déduit que notre modèle apprend bien sur le testSet et trainSet .

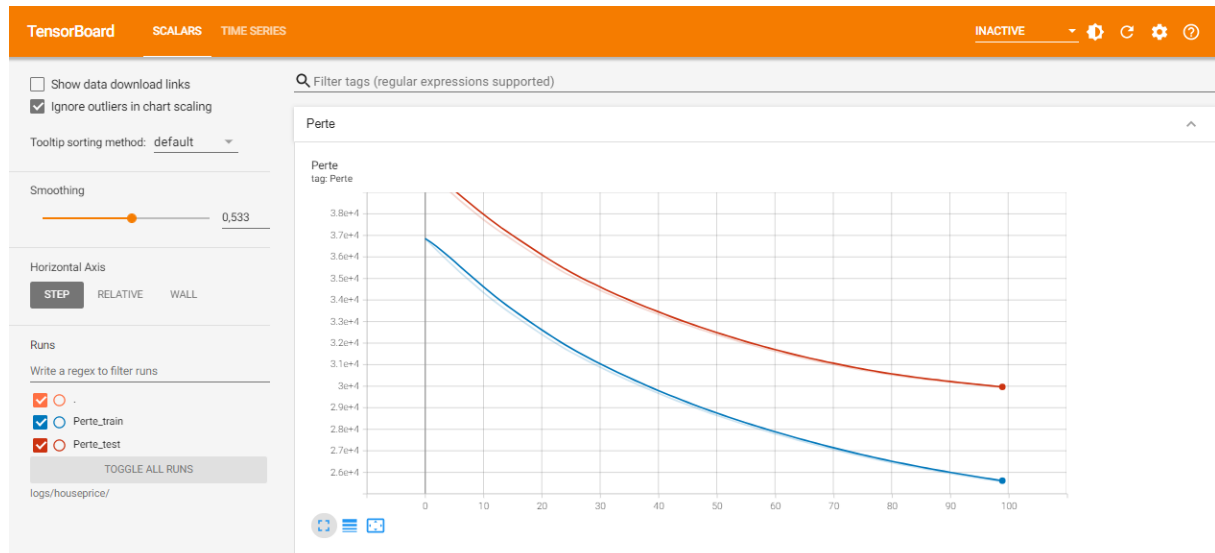


Figure 3: La fonction de perte

## 4. Conclusion :

Après le nettoyage de notre jeu de données et les traitements qu'on a effectué on a réussi de trouver un bon modèle avec des bonnes prédictions.

2280	Oumaima Bairouk		0.14708	1	1s
	Your First Entry! Welcome to the leaderboard!				

**Figure 4: Le résultat sur Kaggle(Score)**