



Prédiction de la maladie de Parkinson

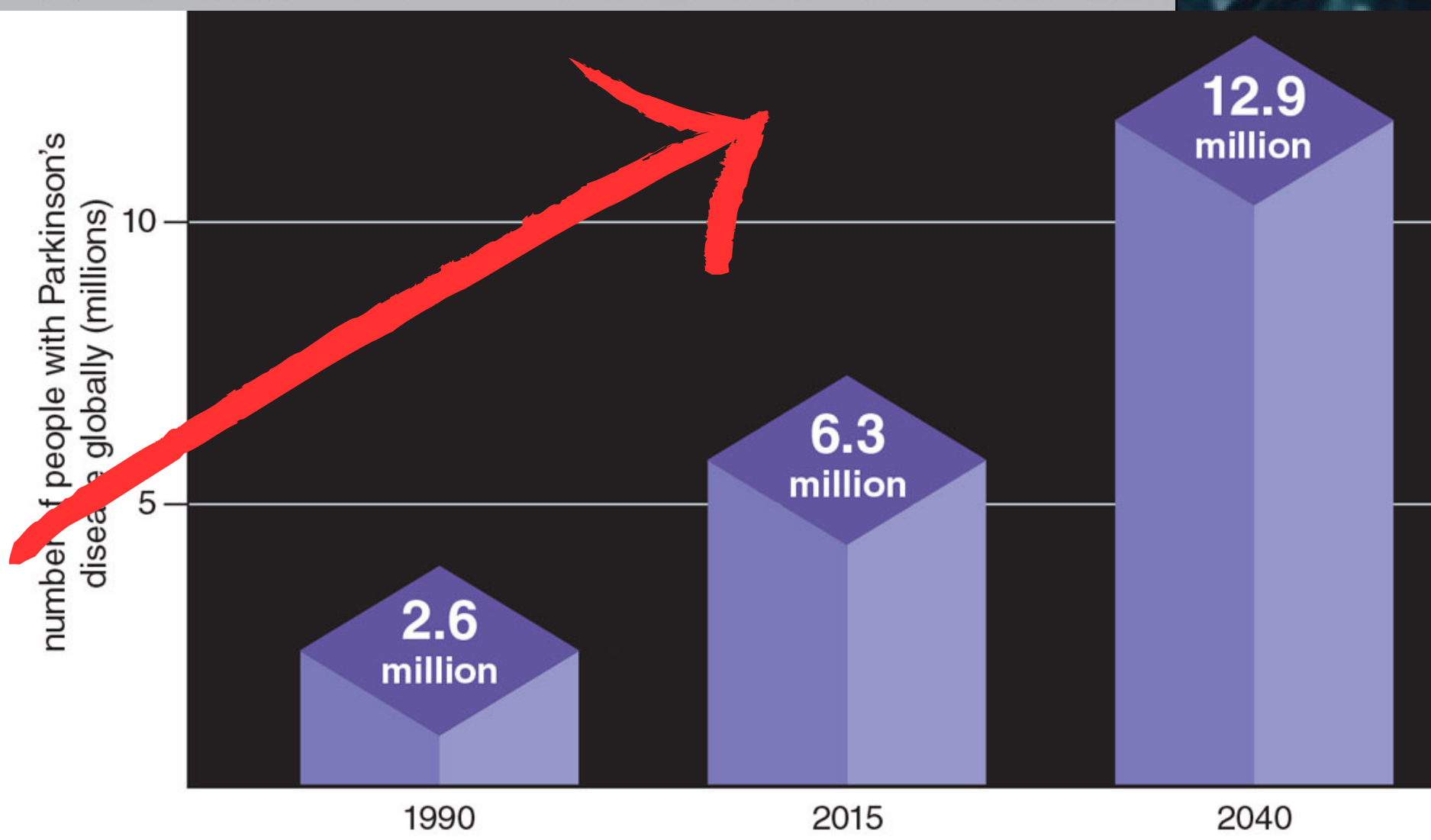
Sommaire

- PARTIE 1 : Introduction
- PARTIE 2 : Caractéristiques des Données
- PARTIE 3 : Algorithmes Utilisés
- PARTIE 4 : Comparaison de modèles
- PARTIE 5 : Intégration

Introduction



PARKINSON'S DISEASE DETECTION USING MACHINE LEARNING



Caractéristiques des Données



Parkinsons Telemonitoring

Donated on 10/28/2009

Oxford Parkinson's Disease Telemonitoring Dataset

Tableau des données

Variable Name	Role	Type	Demographic	Description
subject#	ID	Integer		Integer that uniquely identifies each subject
age	Feature	Integer	Age	Subject age
test_time	Feature	Continuous		Time since recruitment into the trial. The integer part is the number of days since recruitment.
Jitter(%)	Feature	Continuous		Several measures of variation in fundamental frequency
Jitter(Abs)	Feature	Continuous		Several measures of variation in fundamental frequency
Jitter:RAP	Feature	Continuous		Several measures of variation in fundamental frequency
Jitter:PPQ5	Feature	Continuous		Several measures of variation in fundamental frequency
Jitter:DDP	Feature	Continuous		Several measures of variation in fundamental frequency
Shimmer	Feature	Continuous		Several measures of variation in amplitude
Shimmer(dB)	Feature	Continuous		Several measures of variation in amplitude

Shimmer:APQ3	Feature	Continuous		Several measures of variation in amplitude
Shimmer:APQ5	Feature	Continuous		Several measures of variation in amplitude
Shimmer:APQ11	Feature	Continuous		Several measures of variation in amplitude
Shimmer:DDA	Feature	Continuous		Several measures of variation in amplitude
NHR	Feature	Continuous		Two measures of ratio of noise to tonal components in the voice
HNR	Feature	Continuous		Two measures of ratio of noise to tonal components in the voice
RPDE	Feature	Continuous		A nonlinear dynamical complexity measure
DFA	Feature	Continuous		Signal fractal scaling exponent
PPE	Feature	Continuous		A nonlinear measure of fundamental frequency variation
motor_UPDRS	Target	Continuous		Clinician's motor UPDRS score, linearly interpolated
total_UPDRS	Target	Continuous		Clinician's total UPDRS score, linearly interpolated
sex	Feature	Binary	Sex	Subject sex '0' - male, '1' - female

LE NOMBRE D'INSTANCES : 5875
NOMBRE DE CARACTÉRISTIQUES : 19

Objectif des données

PRÉDIRE LES SCORES

Motor UPDRS
Total_UPDRS

À PARTIR DE 16 MESURES VOCALES

Seuils cliniques :

FAIBLE SÉVÉRITÉ (LÉGÈRE):
MOTOR UPDRS SCORE: 0 TO 20

SÉVÉRITÉ MODÉRÉE :
MOTOR UPDRS SCORE: 20 TO 40

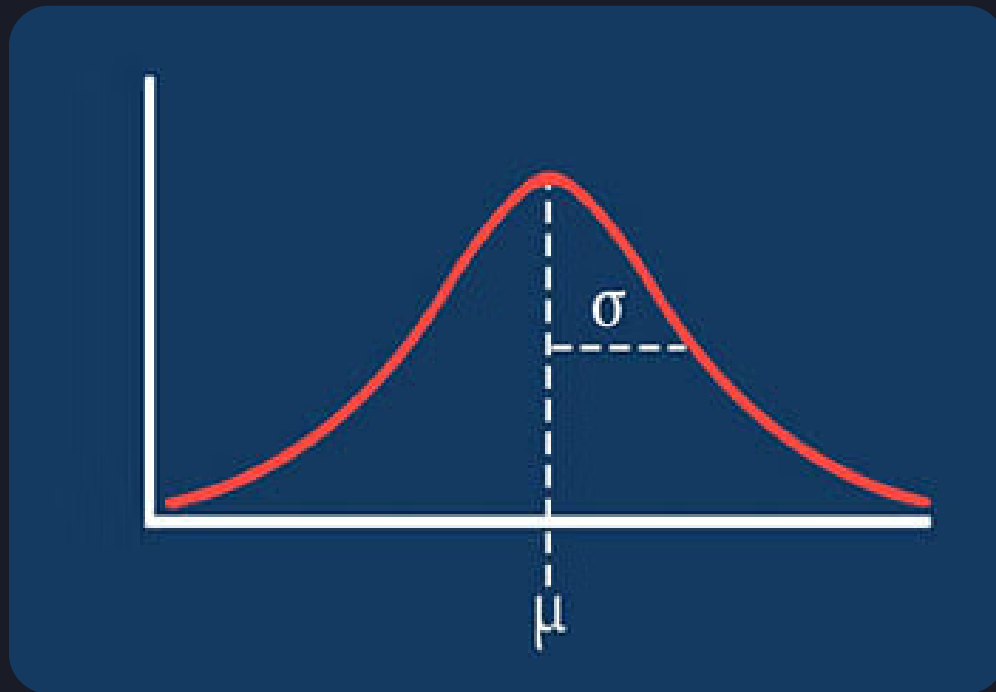
SÉVÉRITÉ ÉLEVÉE (SÉVÈRE) :
MOTOR UPDRS SCORE: 40 AND ABOVE



NB : LE MAXIMUM DES SCORES UPDRS DE LA DATASET SONT :
MOTOR_UPDRS : 39.511
TOTAL_UPDRS : 54.992

Prétraitement des Données

- LA NORMALISATION DES DONNÉES
- LE TRAITEMENT DES VALEURS MANQUANTES
- ÉLIMINATION DE LA COLONNE « SUBJECT »

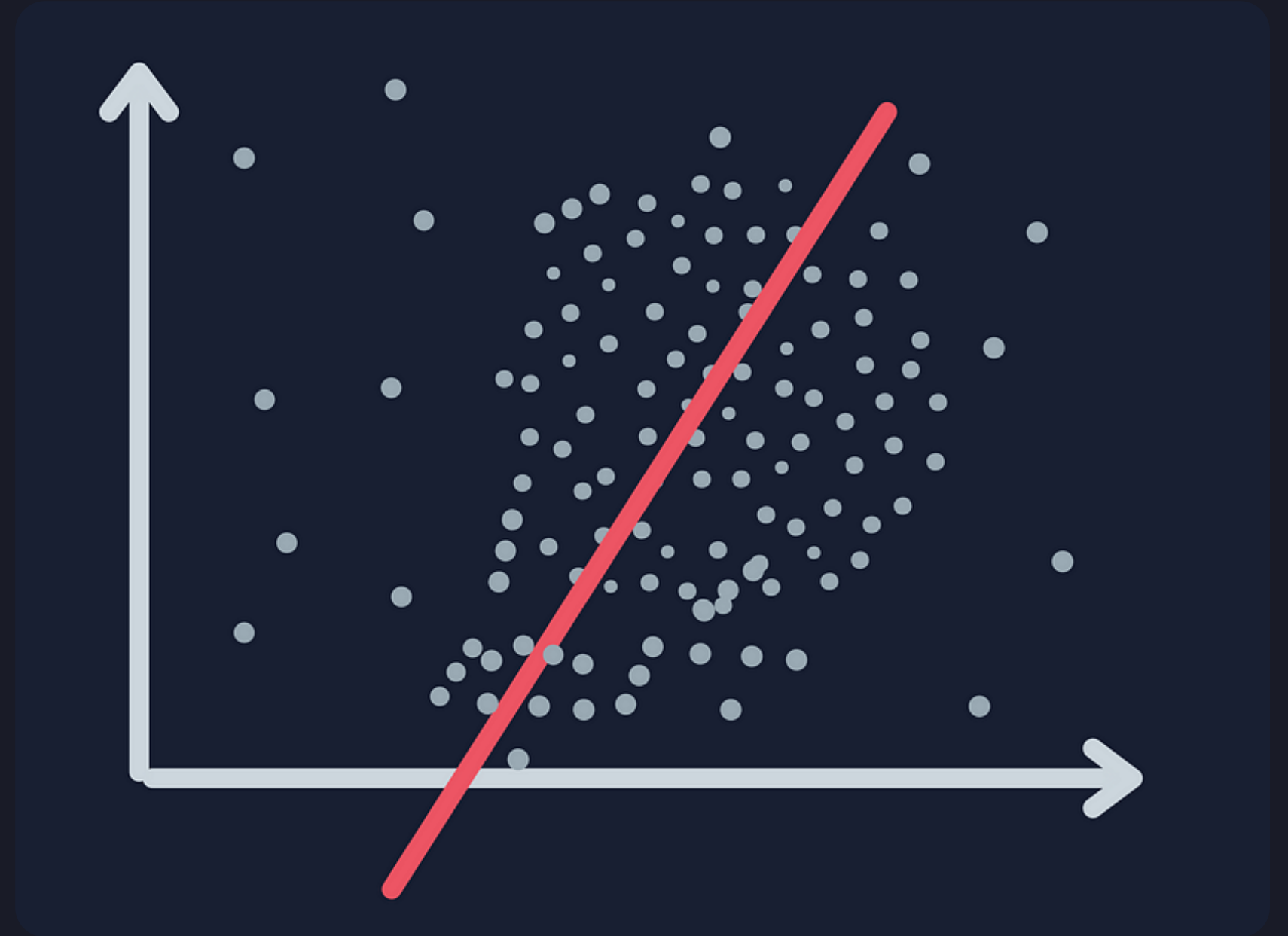


Variable Name	Role	Type
subject#	ID	Integer



Algorithmes Utilisés

Régression linéaire



MÉTRIQUES DE PERFORMANCE DU MODÈLE

★ ERREUR QUADRATIQUE MOYENNE
(MSE - MEAN SQUARED ERROR)

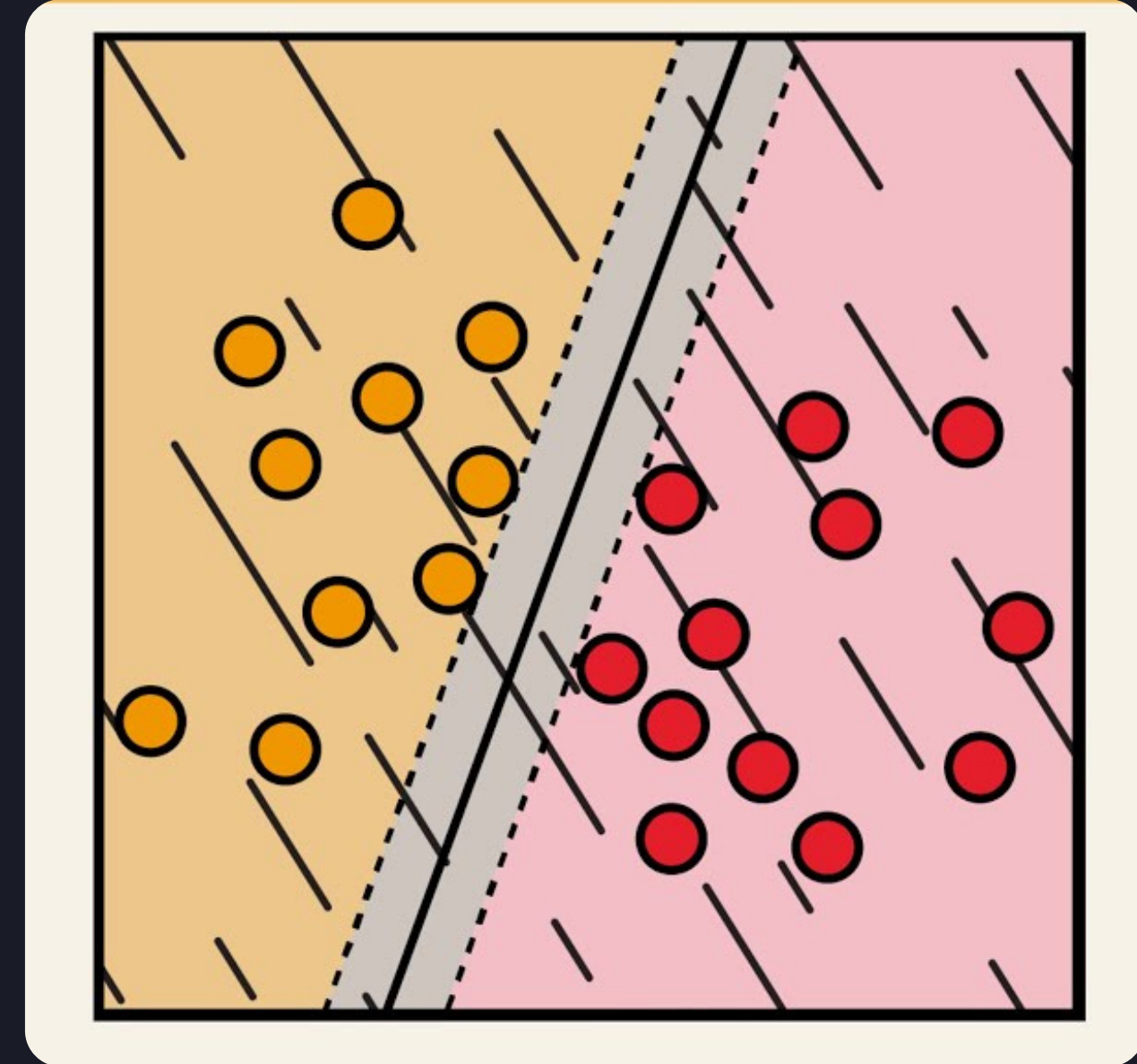
★ COEFFICIENT DE DÉTERMINATION
(R² - R-SQUARED)

```
score 0.6679576465646273  
Multivariate UPDRS Regression:  
Mean Squared Error: 0.04456191257720407  
R-squared: 0.13841810630890372
```

MODÈLE NON PERFORMANT



Svm

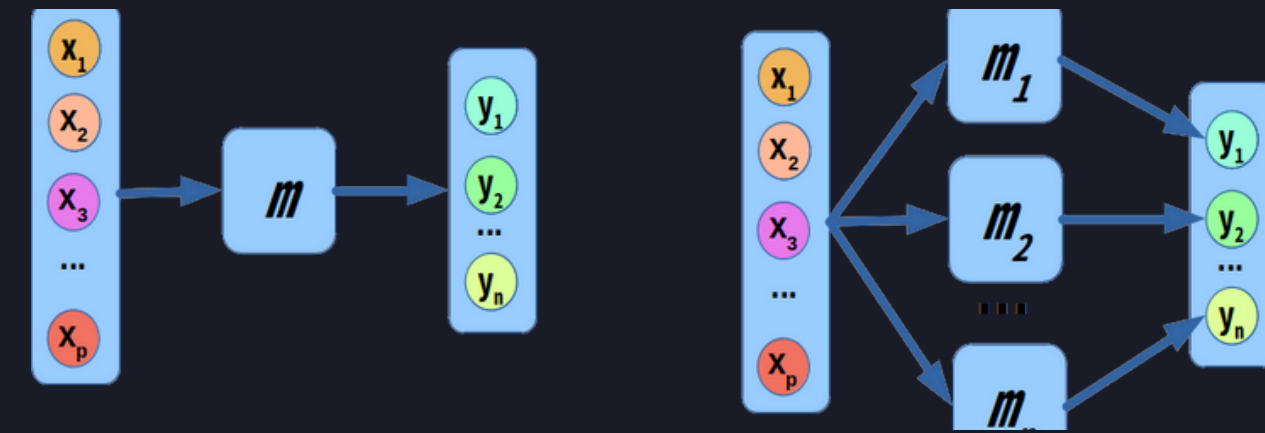


ARCHITECTURE DU MODÈLE

- ★ NOYAU RBF CE NOYAU TRANSFORME NOS DONNÉES DANS UN ESPACE DE DIMENSION SUPÉRIEURE



★ UTILISATION DE MULTIOUTPUTREGRESSOR



- ★ PARAMÈTRE DE RÉGULARISATION C **FIXÉ À 100** A ÉTÉ UTILISÉ POUR CONTRÔLER LA TRADE-OFF ENTRE LA COMPLEXITÉ DU MODÈLE ET SON APTITUDE À BIEN S'ADAPTER AUX DONNÉES.

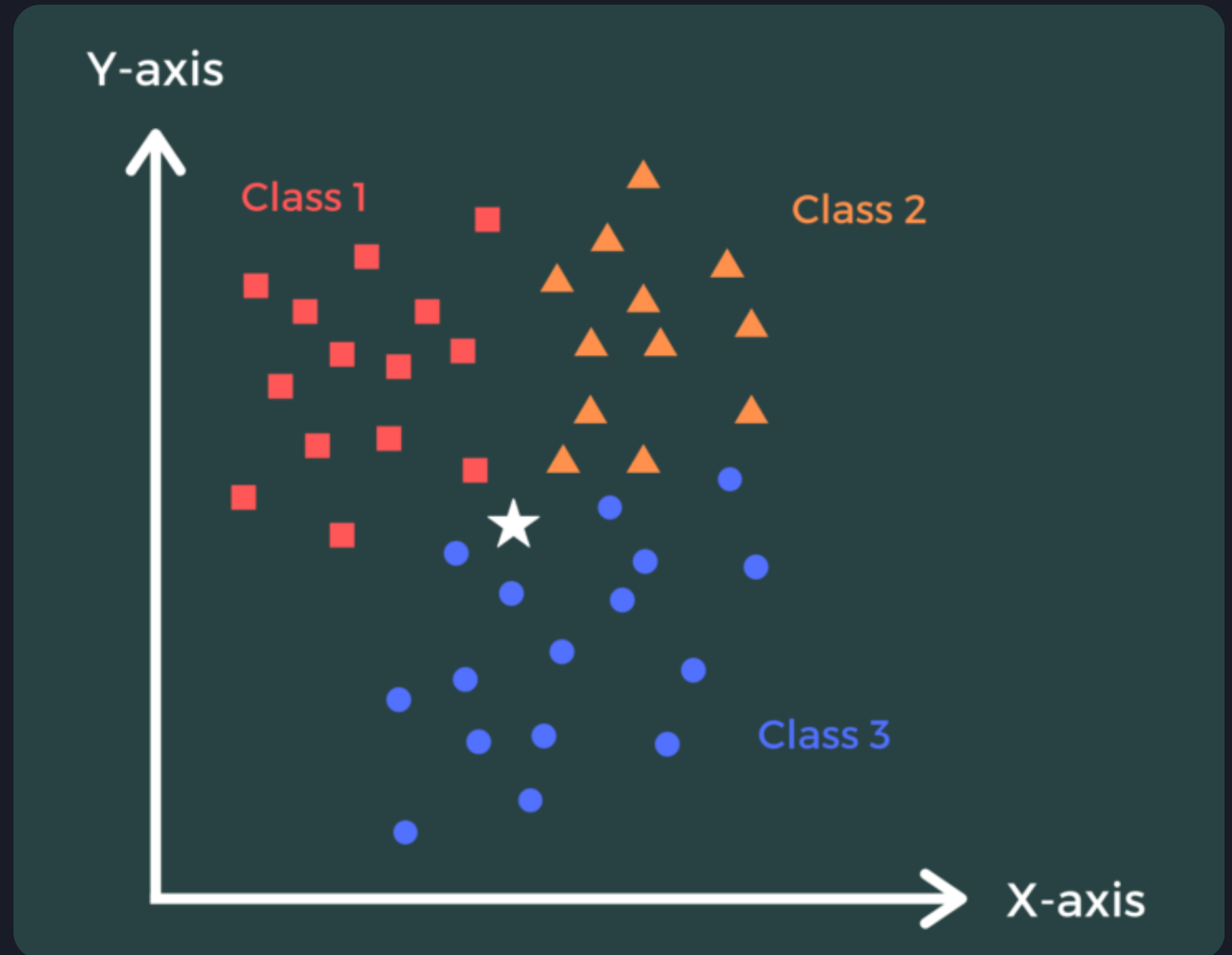
PERFORMANCES ET RÉSULTATS

★	MEAN SQUARED ERROR (MSE)	0.0218
★	COEFFICIENT DE DÉTERMINATION (R^2)	0.5788
★	RACINE CARRÉE DE L'ERREUR QUADRATIQUE MOYENNE (RMSE)	0.1476
★	ERREUR ABSOLUE MOYENNE (MAE)	0.113
★	CORRÉLATION	-1

MODÈLE PERFORMANT MODEREMMENT



Knn



ARCHITECTURE DU MODÈLE

- ✦ CRÉATION ET ENTRAÎNEMENT DU MODÈLE KNEIGHBORS REGRESSOR
- ✦ DÉFINIR LA GRILLE DES PARAMÈTRES À TESTER
- ✦ ENTRAÎNEMENT DU MODÈLE AVEC RECHERCHE DES MEILLEURS PARAMÈTRES
- ✦ LES MEILLEURS PARAMÈTRES TROUVÉES:: {'METRIC': 'MANHATTAN',
'N_NEIGHBORS': 6, 'WEIGHTS': 'DISTANCE'}

PERFORMANCES ET RÉSULTATS

★	★	COEFFICIENT DE DÉTERMINATION (R^2)	0.7644
	★	RACINE CARRÉE DE L'ERREUR QUADRATIQUE MOYENNE (RMSE)	0.1102
	★	ERREUR ABSOLUE MOYENNE (MAE)	0.065
	★	CORRÉLATION	1

Decision tree



ARCHITECTURE DU MODÈLE

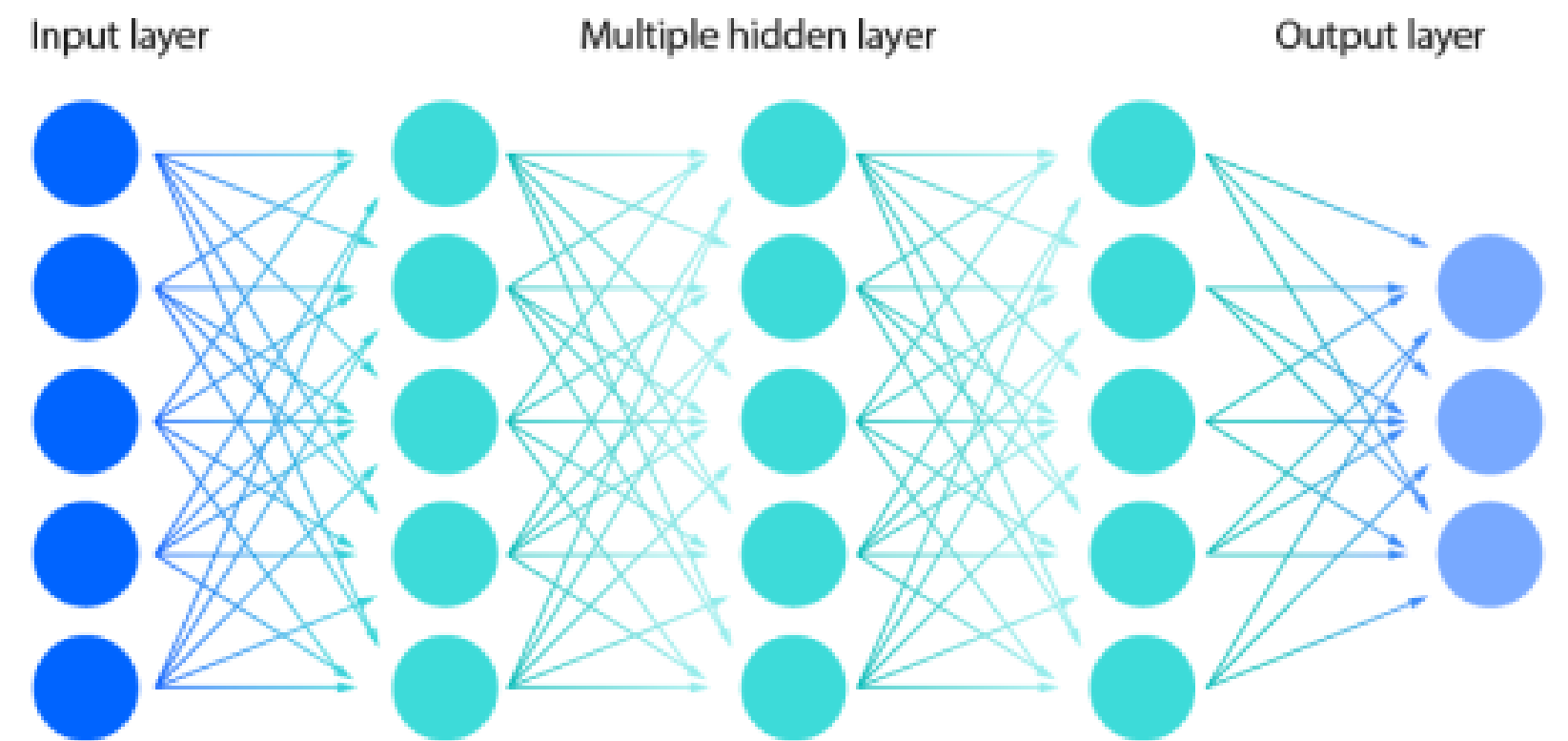
- ✦ CRÉATION ET ENTRAÎNEMENT DU MODÈLE ARBRE DE DÉCISION
- ✦ AJUSTEMENT DE MAX_DEPTH(LA PROFONDEUR MAXIMAL DE L'ARBRE)
- ✦ FAIRE DES PRÉDICTIONS SUR LA PARTIE TEST
- ✦ ÉVALUATION DU MODÈLE

PERFORMANCES ET RÉSULTATS

	★	MEAN SQUARED ERROR (MSE)	0.036
	★	COEFFICIENT DE DÉTERMINATION (R²)	0.928
★	★	RACINE CARRÉE DE L'ERREUR QUADRATIQUE MOYENNE (RMSE)	0.06
	★	ERREUR ABSOLUE MOYENNE (MAE)	0.024
	★	CORRÉLATION	1

Réseau de neurones

Deep neural network



ARCHITECTURE DU MODÈLE

★ NOTRE MODÈLE EST UN RÉSEAU NEURONAL SÉQUENTIEL

★ UNE COUCHE D'ENTRÉE (19,)

★ COUCHES CACHÉES

★ COUCHES SORTIE

★ 30 ÉPOQUES AVEC UN BATCH SIZE DE 140:



★ PERFORMANCES ET RÉSULTATS

```
37/37 [=====] - 0s 2ms/step - loss: 0.0098 - mse: 0.0098  
Perte (MSE) sur l'ensemble de test : [0.009790387935936451, 0.009790387935936451]
```

★ CALCUL DU COEFFICIENT DE DÉTERMINATION R^2 ENTRE LES VALEURS RÉELLES ET PRÉDITES



```
37/37 [=====] - 0s 2ms/step  
 $R^2$  Score: 0.810784473250429
```

★ CES RÉSULTATS DÉMONTRENT L'EFFICACITÉ DE NOTRE MODÈLE SÉQUENTIEL

MODÈLE PERFORMANT



Comparaison de modèles



Les valeurs R2 squared des modèles

Modèle	R2 squared
Support Vector Machine (SVM)	0.57
k-Nearest Neighbors (KNN)	0.76
Decision tree	0.92
Régression linéaire	0.13
Réseaux de neurones	0.81



les valeurs de Mean Squared Error (MSE), Root Mean Squared Error (RMSE) et Mean Absolute Error (MAE)

Modèle	Mean Squared Error (MSE)	Root Mean Squared Error (RMSE)	Mean Absolute Error (MAE)
Decision Tree	0.0036869396963509172	0.06072017536495524	0.024113214557162484
SVM	0.021774184742448445	0.147560783213049	0.11300990486148446
KNN	0.12149384393918164	0.11024377712372554	0.06535725350839787
Réseaux de neurones	0.009790387935936451	0.08361810104576031	0.05020479635927863
Régression linéaire	0.04	0.20099474755413203	0.10049737377706602

Intégration





SCAN ME

Merci !

Notre équipe



HAMDAOUI KHALIL



MEHJOUBI YOSRI



NAAMANE NAJWA



EL OUATIKI ABDESSAMAD