Prétraitements

-> Remplissage des valeurs non renseignées par des 0

-> Premier encodage (en lettres) des catégories

-> Concaténation des catégories en une seule

-> Suppression des doublons par combinaison

-> Suppression de la colonne « User\_ID »

Encodage

-> Utilisation d’un encodage séquentiel pour toutes les colonnes sauf pour « Purchase » puisqu’elle contient des valeurs continues.

-> Pour la colonne « Purchase » l’encodage utilisé est une normalisation par minimum (=0) et maximum (=1) a été utilisée.

Entrainement

-> Utilisation des models disponible sur Sk-Learn ; LinearRegression, DecisionTreeRegressor, RandomForestRegressor, SVC (SVM), DecitionTreeClassifier, RandomForestClassifier. En donnant des paramètres différents après analyse des métriques pour obtenir le model qui donne les résultats les plus exploitables.