

Université du Québec à Rimouski

Campus de Lévis

TP2 – RAPPORT DE CONCEPTION

Par

Bastien Goulet

Travail présenté à M. Yacine Yaddaden

Dans le cadre du cours Programmation orientée objet II

INF11207-MS

9 mars 2021

Le programme est automatisé. Je n’ai pas fait d’interface interactive en raison du manque de temps. En effet, le programme crée les deux arbres de décision et leur fournit deux échantillons chacun afin de tester le fonctionnement. Faisons un tour du programme.

## Lecture et fusion des fichiers

Le programme lit le contenu des fichiers .csv utiles grâce à la méthode LireCsv()de la classe Csv.

Text

Description automatically generated with medium confidence

Le contenu des fichiers est inséré dans un DataTable. Ce type de données est utilisé un peu partout dans ce travail. Il est la manière privilégiée de transporter les ensembles de données.

Pour créer le fichier *animaux.csv*, j’utilise la méthode FusionnerTables() de la classe Csv, comme ci-dessous :

*Text

Description automatically generated with low confidence*

## Construction de l’arbre décisionnel

Par la suite, le programme procède à la construction des arbres de décision en utilisant les DataTables générés. La méthode de la classe ArbreDecision qui effectue ce traitement nécessite également deux autres paramètres : l’attribut correspondant à la classe du tableau (le pivot) et la liste des attributs excluant la classe qui fournit à l’arbre les détails des colonnes et des valeurs du tableau. La méthode NouveauNoeud() est décrite ci-dessous :

Text

Description automatically generated with medium confidence

Pour créer l’arbre, cette méthode récursive doit déterminer les attributs les plus significatifs. Elle y parvient grâce aux méthodes statiques créées à cet effet dans les classes Gains, Entropie, Probabilites et NombreLignes. Ces différentes classes travaillent ensemble afin de calculer le gain d’informations de chaque attribut et prendre la décision quant à l’attribut le plus significatif. J’invite le lecteur à consulter le code source pour observer le fonctionnement en détails de ce système. Je vais me contenter de présenter dans ce document le calcul du gain d’informations.

Graphical user interface, text, application

Description automatically generated

## Test d’un échantillon

La classe ArbreDecision met à la disposition du programme la méthode TesterUnEchantillon() qui lui permet de confier un échantillon à l’arbre qui prendra une décision par la suite en fonction des valeurs de l’échantillon. Cette méthode utilise en réalité la méthode ci-dessous qui contient la logique de décision.

Text

Description automatically generated with medium confidence