Sharon Kasturiarachi

MIS 64060 Fundamentals of Machine Learning

Dr. Rouzbeh Razavi

Kasturiarachi-Assignment 5

Column variables

Text

Description automatically generated

A close-up of a document

Description automatically generated with medium confidence

Finding the number of missing values



# after converting the variables

Text

Description automatically generated

A picture containing text

Description automatically generated

Normalizing the data for each variable to be treated equally by the distance measure

Text, table

Description automatically generated

Diagram

Description automatically generated

Chart

Description automatically generated with medium confidence

Diagram

Description automatically generated with medium confidence

Chart, histogram

Description automatically generated

Complete – considers the maximum distance between clusters

Single – considers the smallest distance between clusters

Average- considers the average distance between clusters

Ward - the pair of clusters with the lowest distance is merged

The agglomerative coefficient is highest in the ward method.

The ward method will be used in the hierarchical clustering analysis.

Text, application

Description automatically generated

Determining the optimal number of clusters using the elbow method

Chart, line chart

Description automatically generated

Chart, line chart

Description automatically generated

I chose 6 clusters

Diagram, schematic

Description automatically generated

A picture containing text

Description automatically generated

#store the clusters in a data frame along with the cereals data

cereals\_clusts\_hc <-cbind(clusters, cereals\_data.clean)

Table

Description automatically generated

Table

Description automatically generated

Text

Description automatically generated



To find a cluster of healthy cereals to support a healthy diet. They should be high in fiber and low in sugar content

Lowest sugar is 0.00 and maximum fiber is 14

Let’s examine the clusters

Graphical user interface

Description automatically generated with low confidence

**Cluster 1 has 3 cereals** - 100%-Bran, All-Bran, All-Bran-with-Extra-Fiber- fall under healthy breakfast cereals

**Cluster 2 has 10 cereals** - 100%-Natural-Bran, Clusters, Cracklin'-Oat-Bran, Crispy-Wheat-&-Raisins, Life, Muesli-Raisins Dates &Almonds, Muesli-Raisins Peaches & Pecans, Quaker Oat Squares, Raisin Nut Bran, Great Grains Pecan-

**Cluster 3 has 21 cereals** – Apple Cinnamon Cheerios, Cap’n’Crunch, Apple Jacks, Cinnamon Toast Crunch, Cocoa Puffs, Corn Pops, Count Chocula, Froot Loops, Frosted Flakes, Fruity Pebbles, Golden Crisp, Golden Grahams, Honey Graham Ohs, Honey Nut Cheerios, Honeycomb, Lucky Charms, Multi-Grain Cheerios, Nut &Honey Crunch, Smacks, Trix, Wheaties Honey Gold.

**Cluster 4 has 10 kinds of cereals** – Basic 4, Fruit & Fibre Dates, Walnuts, and Oats, Fruitful Bran, Just Right Fruit & Nut, Mueslix Crispy Blend, Nutri-Grain Almond Raisin, Oatmeal Raisin Crisp, Post Nat. Raisin Bran, Raisin Bran, Total Raisin Bran

**Cluster 5 has 21 kinds of cereals**- Bran Chex, Bran Flakes, Cheerios, Corn Chex, Cornflakes, Crispix, Double Chex, Grape Nuts Flakes, Grape-Nuts, Just Right Crunchy Nuggets, Kix, Nutri-grain Wheat, Product 19, Rice Chex, Rice Krispies, Special K, Total Corn Flakes, Total Whole Grain, Triples, Wheat Chex, Wheaties,

**Cluster 6 has 9 kinds of cereal** – Strawberry Fruit Wheats, Shredded Wheat spoon size, Shredded Wheat 'n' Bran, Shredded Wheat, Raisin Squares, Puffed Wheat, Puffed Rice, Maypo, Frosted Mini-Wheats

Graphical user interface, application, table

Description automatically generated