Cereal Analysis Based On Ratings By Using Machine Learning Techniques With IBM Watson

1. INTRODUCTION

1.1 Overview

For kids who grew up in the 1980s and 1990s, sugary cereals flooded the television landscape and many of these commercials shared a common theme of advocating cereal as the "perfect and complete breakfast". Overtime, as more people became health-conscious, fortified cereal turned out to be the new solution to address any health concerns. We knew cereal was the real deal when supermarkets started to dedicate an aisle for it. Walking along these aisles, we observed labels often screaming "25% more iron", "10% more calcium!", "High in Vitamin A, C and E!" and the list goes on.. However, the question to ask is: on what basis does our consumers choose their cereals? Is it based on taste, nutritional values or price?

In this analysis, we will disclose the main attributes that influence cereal ratings. We will also be predicting ratings using a simple linear regression model.

1.2 Purpose

The purpose is to find the high dietary food that is predicted on the basis of rating of the food. A customer can get high dietary food by the rating of the food given to it from the cereals and ingredients present. The model can predict the rating of the food more accurately by giving the inputs which are the cereals and ingredients present in the food. it is important to know the relationships between things like calories, sugar, vitamins, and how these factors related to the rating of a cereal. It would be very useful for manufacturers to be able to receive higher ratings on their cereals while at the same time offering healthier options for consumers.

2. LITERATURE SURVEY

2.1 Existing Problem

A customer wants to buy some food items with high dietary benefits so that he wants to know which food item has high dietary benefits. It is so difficult to choose an item. Usually a customer expects to consume dietary cereals with high proteins, fiber and low sugars, fats. So it is difficult to find the beneficiary cereals by manual examination. Therefore, Predicting a brand with high dietary cereals became a big issue.

2.2 Proposed System

We use machine learning algorithms to predict the food with a high beneficiary diet. The model can predict the rating of the food more accurately by giving the inputs which are the cereals and ingredients present in the food. The most notable of findings is that healthier cereals are generally disposed to receiving better ratings. It would be very useful for manufacturers to be able to receive higher ratings on their cereals while at the same time offering healthier options for consumers. The rating is predicted using the neural networks model.

3. THEORETICAL ANALYSIS

3.1 Block Diagram

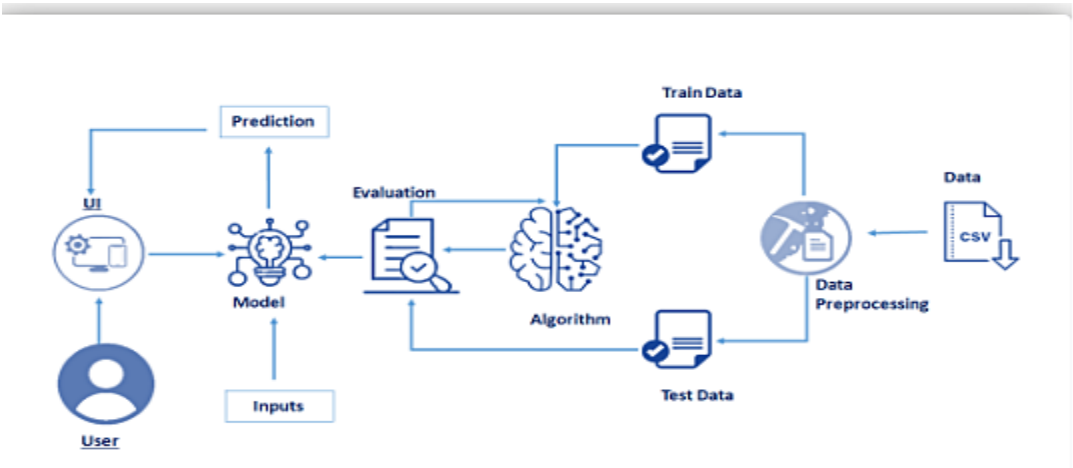

Fig 1: Block Diagram

3. Hardware/Software Designing

Hardware Requirements:

RAM: 8 GB

Storage: 250 GB

Processor: Intel i5 Core Processor

Software Requirements:

Operating System: Windows 10 Home

Anaconda: Anaconda must be installed as it provides jupyter notebook and spyder.

IBM Academic initiative accounts is required to access IBM services

- IBM Watson Studio: IBM Watson studio helps data scientist and analyst prepare data and build models at scale across any cloud.
- IBM Watson Machine Learning: IBM watson machine learning helps data scientists and developers accelerate AI and machine-learning deployment.
- IBM Cloud Object Storage:

IBM cloud object storage makes it possible to store practically limitless amounts of data,simply and cost effectively.

4. EXPERIMENTAL INVESTIGATIONS

Our aim is to build a machine learning model that estimates the ratings of food. For this purpose, we collect the dataset for training purposes. From the dataset we have chosen the necessary attributes that contribute to our prediction. The dataset consist of 16 attributes and has no missing values. There are 15 input variables and 1 output variable. Here we are going to build a machine learning model that find the high dietary food that is predicted on the basis of rating of the food. After some preprocessing on the dataset, we are building our model.

In this analysis, we will disclose the main attributes that influence cereal ratings. We will predicting ratings using neural networks model.

5. FLOWCHART

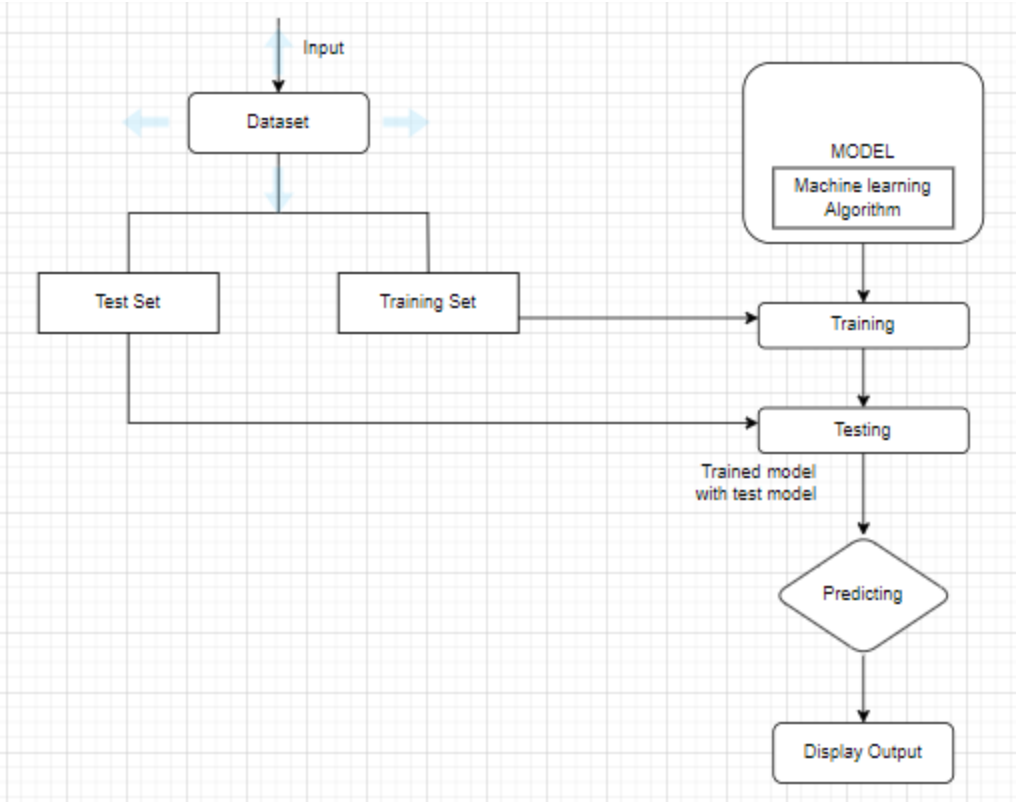

Fig 2: Flowchart

6. RESULT

Fig 3: Home Page

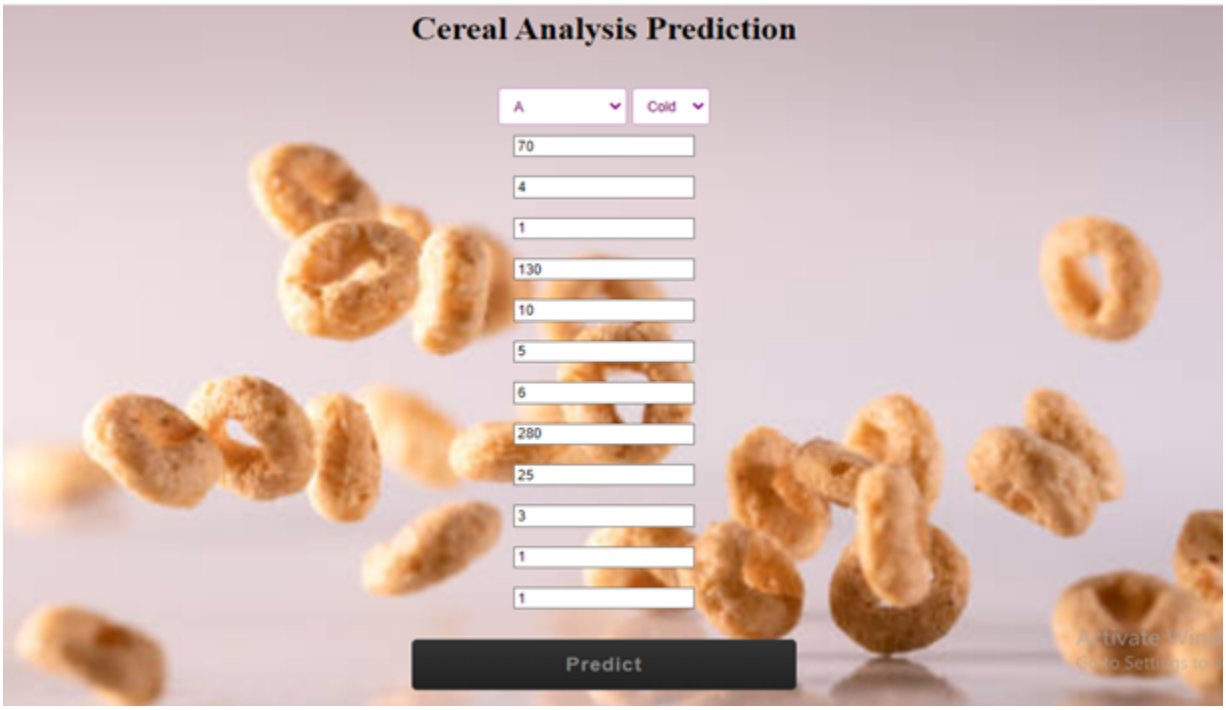

Fig 4: Predict Page

Fig 5: Result Page

7. ADVANTAGES & DISADVANTAGES

7.1 Advantages

- Help to choose high beneficiary cereal
- Critical role in human diets
- Encouraging progress toward better nutritional quality
- Toward a sustainable dietary revolution

7.2 Disadvantages

- Attributes and dataset collected should be correct. Otherwise it will affect the accuracy of our model.
- Require more dataset to increase the accuracy of model.

8. APPLICATIONS

An individual can use this application as a cereal rating calculation to choose high dietary food and improve their health. It can also be used in the medical industry to provide high beneficiary diet food to the patients for improve their health. It will reduce pressure on medical staff and save time, providing great convenience in medical treatments. Helps to keep track on contents of food intake on a daily basis which helps in health maintenance.It would be very useful for manufacturers to be able to receive higher ratings on their cereals while at the same time offering healthier options for consumers.

9. CONCLUSION

The more calories that a cereal has, the less likely it is to receive a high rating. Manufacturers that want to bring in high ratings should create cereals that are high in fiber, protein, and potassium and avoid creating cereals with high calorie counts or lots of sugar or fat. Cereals with high ratings are more likely to be placed on the first or third shelf, because that is generally where the consumers' eyes gravitate.

10. FUTURE SCOPE

In future, we can include the aspects to accurately predict the cereal ratings more effectively.An improved form of this application developed and used in the health sector for better suggestions of food and thereby we can maintain health.It would be very useful for manufacturers to be able to receive higher ratings on their cereals while at the same time offering healthier options for consumers.

11. BIBLIOGRAPHY

https://www.kaggle.com/code/kianwee/analysis-on-cereal-prediction-on-ratings/notebook

www.wikipedia.org

12. APPENDIX

Source code

```
from flask import Flask, render_template, request

import requests

import json

    # NOTE: you must manually set API_KEY below using information retrieved from your IBM Cloud account.

API_KEY = "bJSrKnWH-dYvye3Ig4-MxTGXynisAZKTwdMhOaIXTXsk"

token_response = requests.post('https://iam.cloud.ibm.com/identity/token', data={"apikey": API_KEY, "grant_type": 'urn:ibm:params:oauth:grant-type:apikey'})

mltoken = token_response.json()["access_token"]


header = {'Content-Type': 'application/json', 'Authorization': 'Bearer ' + mltoken}

app = Flask (__name__)

import pickle

model = pickle.load(open('cerealanalysis.pkl','rb'))


@app.route('/')

def helloworld():

    return render_template('base.html')


@app.route('/assesment')

def prediction ():

    return render_template('index.html')


@app.route('/predict', methods = ['POST'])
```

```
def admin():
    a=request.form["mfr"]
    if (a == 'a'):
      a1, a2, a3, a4, a5, a6, a7=1,0,0,0,0,0,0
    if (a == 'g'):
      a1, a2, a3, a4, a5,a6,a7 = 0,1,0,0,0,0,0
    if (a == 'k'):
      a1, a2, a3, a4, a5, a6, a7=0,0,1,0,0,0,0
    if (a == 'n'):
      a1, a2, a3, a4, a5, a6, a7=0,0,0,1,0,0,0
    if (a == 'p'):
      a1, a2, a3, a4, a5, a6, a7=0,0,0,0,1,0,0
    if (a == 'q'):
      a1, a2, a3, a4, a5, a6, a7=0,0,0,0,0,1,0
    if (a == 'r'):
      a1, a2, a3, a4, a5, a6, a7=0,0,0,0,0,0,1
    b= request.form["type"]
    if (b=='c'):
        b=0
    if (b== 'h'):
        b=1
    c= request.form["Calories"]
    d= request.form["Protien"]
    e= request.form[ "Fat"]
    f= request.form["Sodium"]
    g= request.form[ "Fiber"]
    h= request.form["Carbo"]
```

```
    i= request.form["Sugars"]
    j= request.form["Potass"]
    k= request.form[ "Vitamins"]
    l= request.form[ "Shelf"]
    m= request.form["weight"]
    n= request.form["Cups"]


    t=[[int (a1), int(a2), int(a3), int(a4), int(a5), int(a6), int (a7), int (b), int(c), int(d), int(e), int(f)
,int(g), int(h),int(i),int(j),int(k),int(l),int(m),float(n)]]
    y = model.predict(t)
    return render_template("prediction.html", z = y[0][0])


 # NOTE: manually define and pass the array(s) of values to be scored in the next line
    payload_scoring = {"input_data": [{"field":
[["mfr","type","G1","G2","G3","G4","G5","G6","calories","protein","fat","sodium","fiber","carb
o","sugars","potass","vitamins","shelf","weight","cups"]], "values": t}]}


    response_scoring = requests.post('https://us-
south.ml.cloud.ibm.com/ml/v4/deployments/d18d52cb-55ec-40df-9e62-
b8de982c3585/predictions?version=2021-10-28', json=payload_scoring,
headers={'Authorization': 'Bearer ' + mltoken})
    print("Scoring response")
    predictions=response_scoring.json()
    print("Final Prediction:")
    #print(predictions['predictions'][0]['values'][0][0])
    print(predictions)
if __name__ == "__main__":
   app.run(debug=False)
```