SESHADRI RAO GUDLAVALLERU ENGINEERING COLLEGE

PROJECT REPORT

MILK GRADING SYSTEM USING IBM WATSON

Prepared by

CH.Kousik Sai

19481A0316

Gudlavalleru Engineering College

<u>CONTENT</u>

- INTRODUCTION
- PURPOSE
- THEORITICAL ANALYSIS
  - Block diagram
  - Hardware/software designing
- EXPERIMENTAL INVESTIGATION
- FLOW CHART
- PROJECT FLOW
  - Data set collection
  - Data Preprocessing/Analysis
    - Importing the libraries
    - Reading the data set
    - Data visualization
  - Model Building
    - Training and Testing the model
    - Evaluation of Model
  - Application Building
    - We create an HTML pages
    - Build a python code
- RESULT
- ADVANTAGES
- DISADVANTAGES
- APPLICATIONS
- CONCLUSION
- FUTURE SCOPE
- REFERENCES

INTRODUCTION :

Grading refers to the classification of milk into different grades.

THE PURPOSE of grading milk is to separate the available supply of potable milk into classes differing in superiority. Nearly all food products are graded in some way. so that the consumer may select milk for particular purposes according to his desires and pocketbook.

Certified milk is practically the only stable grade, rules for production being laid down by the American Association of Medical Milk Commissions.

There is a serious question as to whether or not it is possible in the present state of the industry to enforce uniform grades universally.

The main problem here is not just the feature sets and target sets but also the approach that is taken in solving these types of problem.

The quality of any dairy product is determined by sensory tests, smell, taste, touch and sound. Of these, taste and smell are the most important in judging and grading.
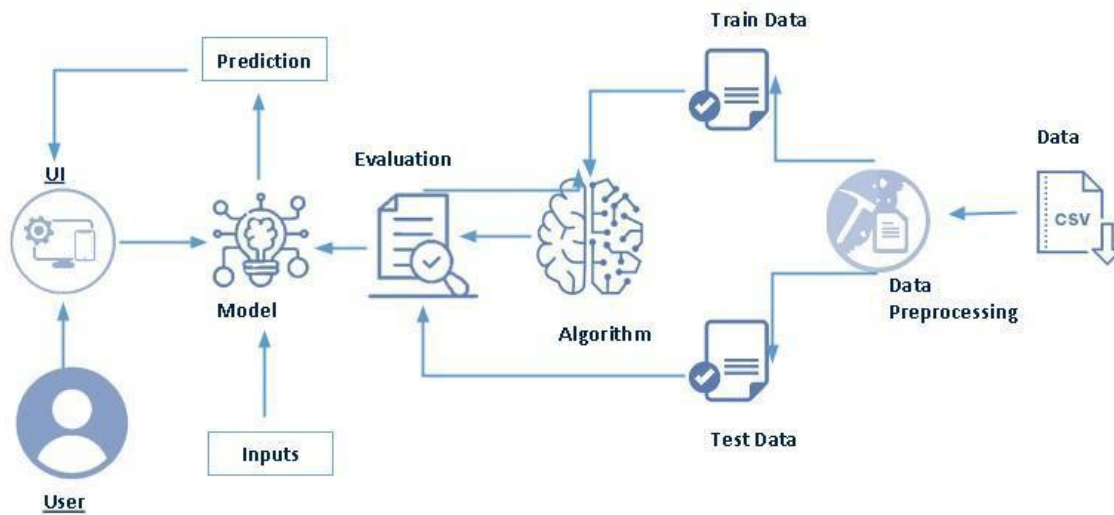
PURPOSE :

The main aims at creating an application from where the user can know the quality of milk is GOOD/BAD/MODERATE by giving some data values to predict.

To know the quality of milk we will do some tests like ph test, temperature test, turbidity test etc. we will also find the taste, odor, fat, colour.

THEORTICAL ANALYSIS :
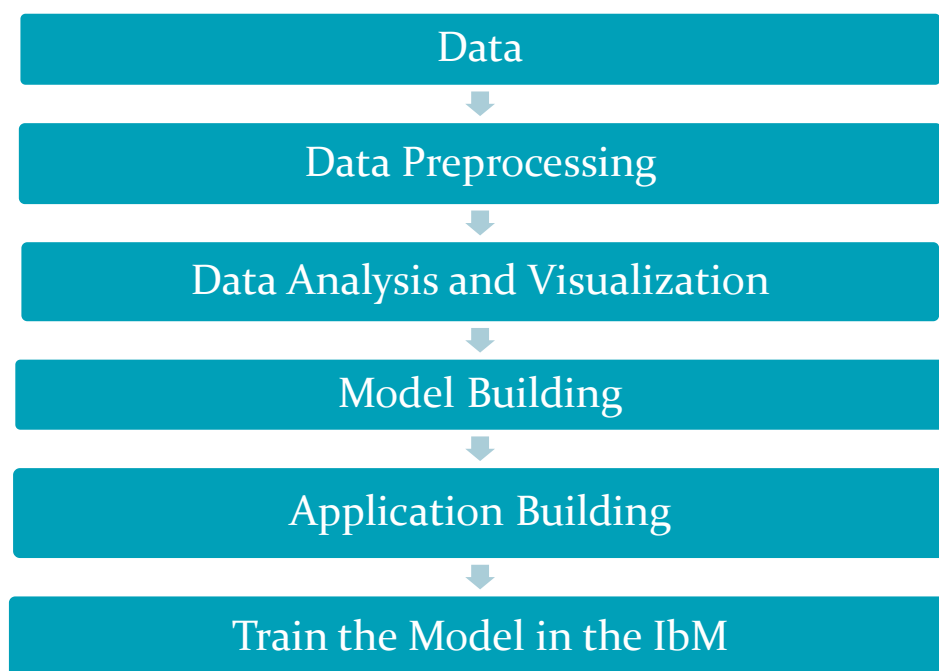
BLOCK DIAGRAM

Hardware/Software designing

- JUPYTER NOTEBOOK
- HTML,CSS,JS
- FLASK
- NOTEPAD
- ALL NECESSARY PYTHON PACKAGES

EXPERIMENTAL INVESTIGATION :

Modeling and prediction of Milk quality are very important for the health. Developing a model by using advanced artificial intelligence algorithms can be used to measure the future milk quality. In this proposed methodology, the advanced artificial intelligence algorithms, namely, NARNET and LSTM models were used to predict the Milk quality. Moreover, machine learning algorithms such as SVM, KNN, and Naive Bayes were used to classify the data.

The proposed models were evaluated and examined by some statistical parameters. For the milk quality prediction, the result has revealed that the performance of the NARNET model is slightly better than the LSTM model based on the obtained value. However, the SVM algorithm has achieved the highest accuracy of the prediction of the Milk quality concepts as compared with KNN and Naive Bayes algorithms. After examining the robustness and efficiency of the proposed model for predicting the milk quality, in future work, the developed models will be implemented to predict the Milk quality.

FLOW CHART

Data

↓

Data Preprocessing

↓

Data Analysis and Visualization

↓

Model Building

↓

Application Building

↓

Train the Model in the IbM

PROJECT FLOW :

- Data Collection.
    - Collect the dataset or Create the dataset
- Data Preprocessing.
    - Import the Libraries.
    - Importing the dataset.
    - Checking for Null Values.
    - Data Visualization.
    - Taking care of Missing Data.
    - Label encoding.
    - One Hot Encoding.
    - Feature Scaling.
    - Splitting Data into Train and Test.
- Model Building
    - Training and testing the model
    - Evaluation of Model
- Application Building
    - Create an HTML file
    - Build a Python Code

Dataset

The dataset used in this study is collected from certain locations in India. It contained different samples from different Indian states during the period. The dataset has 8 significant parameters, pH, Temperature, Taste, Odor, Fat, Turbidity, colour, Grade.

This dataset was obtained from
https://www.kaggle.com/datasets/prudhvignv/milk-grading

DATA PREPROCESSING/ANALYSIS

The processing phase is very important in data analysis to improve the data quality. In this phase, the milk quality has been calculated from the most significant parameters of the dataset. Then, Milk samples have been

classified on the basis of the milk quality index values. For obtaining superior accuracy, the -score method has been used as a data normalization technique.

Importing The Libraries

It is important to import all the necessary libraries such as pandas, numpy, matplotlib.

- Numpy- It is an open-source numerical Python library. It contains a multi-dimensional array and matrix data structures. It can be used to perform mathematical operations on arrays such as trigonometric, statistical, and algebraic routines.
- Pandas- It is a fast, powerful, flexible, and easy to use open-source data analysis and manipulation tool, built on top of the Python programming language.
- Seaborn- Seaborn is a Python data visualization library based on matplotlib. It provides a high-level interface for drawing attractive and informative statistical graphics.
- Matplotlib- Visualization with python. It is a comprehensive library for creating static, animated, and interactive visualizations in Python

Reading the dataset

You might have your data in .csv files, .excel files

Let's load a .csv data file into pandas using read csv() function. We will need to locate the directory of the CSV file at first (it's more efficient to keep the dataset in the same directory as your program).

If your dataset is in some other location, Then see below command

Data=pd.read_csv(r"File_location/filename.csv")

```
data=pd.read_csv('Milk Grading (1).csv')
data.head()
```

	pH	Temprature	Taste	Odor	Fat	Turbidity	Colour	Grade
0	6.6	35	1	0	1	0	254	1.0
1	6.6	36	0	1	0	1	253	1.0
2	8.5	70	1	1	1	1	246	0.0
3	9.5	34	1	1	0	1	255	0.0
4	6.6	37	0	0	0	0	255	0.5

- After reading the data we have to check the file size and data analysis and cleaning.
    - To check the file size use this command
        - data.shape
    - For data analysis use this command
        - data.info()

Data Visualization

Data visualization is where a given dataset is presented in a graphical format. It helps the detection of patterns, trends and correlations that might go undetected in text-based data. Understanding your data and the relationship present within it is just as important as any algorithm used to train your machine learning model. Machine learning models will perform poorly on data that wasn't visualized and understood properly.

To visualize the dataset we need libraries called Matplotlib and Seaborn. The Matplotlib library is a Python 2D plotting library that allows you to generate plots, scatter plots, histograms, bar charts etc.

Model Building

- After visualization and train test model we evaluate the model.

```
from sklearn.metrics import accuracy_score
svc1=SVC(kernel='rbf',gamma='auto',break_ties='bool')
svc1.fit(x_train,y_train)
y_train_pred1=svc1.predict(x_train)
y_test_pred1=svc1.predict(x_test)
print("train accuracy",accuracy_score(y_train_pred1,y_train))
print("test accuracy",accuracy_score(y_test_pred1,y_test))

train accuracy 0.961038961038961
test accuracy 0.9433962264150944
```

- We solve the model on 'milk.pkl'.

```
import pickle
pickle.dump(svc1,open('milk.pkl','wb'))
```

Application Building

- Creating 3 html pages to this project.
    - Home.html
    - Predict.html

> o Submit.html
- Building python code.


```python
In [1]: from flask import Flask, render_template, request
        import numpy as np
        import pickle
        import requests
        import pandas as pd

In [2]: model = pickle.load(open('milk.pkl','rb'))
        app = Flask(__name__)

In [3]: @app.route('/')
        def home():
            return render_template('home.html')
        @app.route("/pred")
        def pred():
            return render_template('predict.html')

In [4]: @app.route("/predict",methods=['POST'])
        def predict():
            x=[]
            x.append(request.form['ph'])
            x.append(request.form['temperature'])
            x.append(request.form['taste'])
            x.append(request.form['odor'])
            x.append(request.form['fat'])
            x.append(request.form['turbidity'])
            x.append(request.form['colour'])
            x = np.array([x])
            pred = model.predict(x)
            if pred=='Good':
                res='Good'
            elif pred=='Moderate':
                res='Moderate'
            else:
                res='Bad'
            return render_template('submit.html',prediction=res)

In [5]: if __name__=="__main__":
            app.run(debug=False)

        * Serving Flask app "__main__" (lazy loading)
        * Environment: production
          WARNING: This is a development server. Do not use it in a production deployment.
          Use a production WSGI server instead.
        * Debug mode: off

        * Running on http://127.0.0.1:5000/ (Press CTRL+C to quit)
        127.0.0.1 - - [16/Sep/2022 20:37:55] "GET / HTTP/1.1" 200 -
        127.0.0.1 - - [16/Sep/2022 20:37:56] "GET /favicon.ico HTTP/1.1" 404 -
        127.0.0.1 - - [16/Sep/2022 20:37:58] "GET /pred HTTP/1.1" 200 -
```
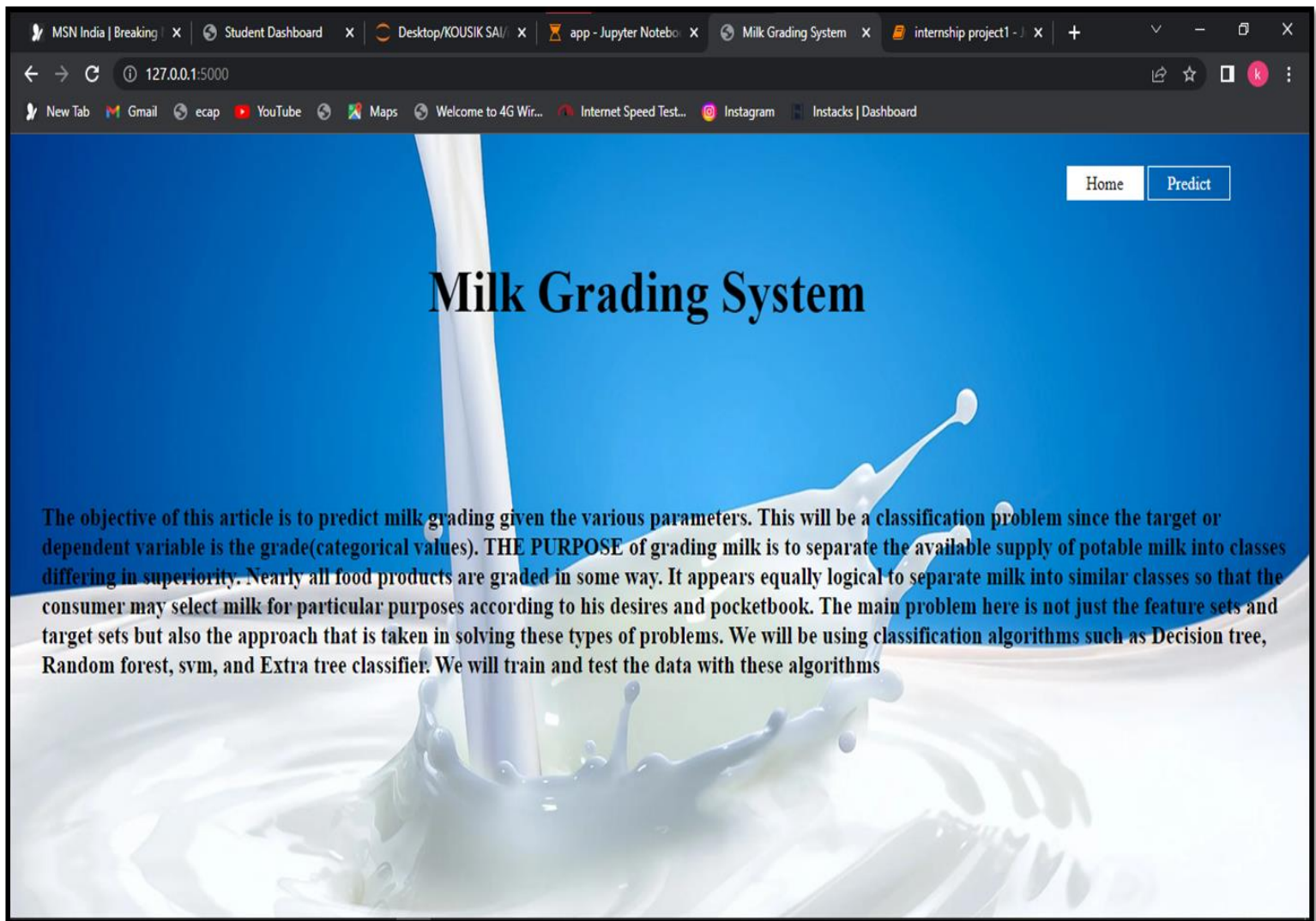
Result :

For validating the developed model, the dataset has been divided into 70% training and 30% testing subsets. While the ANN and LSTM models were used to predict the milk quality the SVM and KNN, and Naïve Bayes were utilized for the Milk quality prediction.

- While running the python code we will get one link address by clicking on it. It will redirect to this page.

- Interference will be shown as above image.
- By clicking on predict button which is shown on right corner. It will redirect to another page as shown in image.
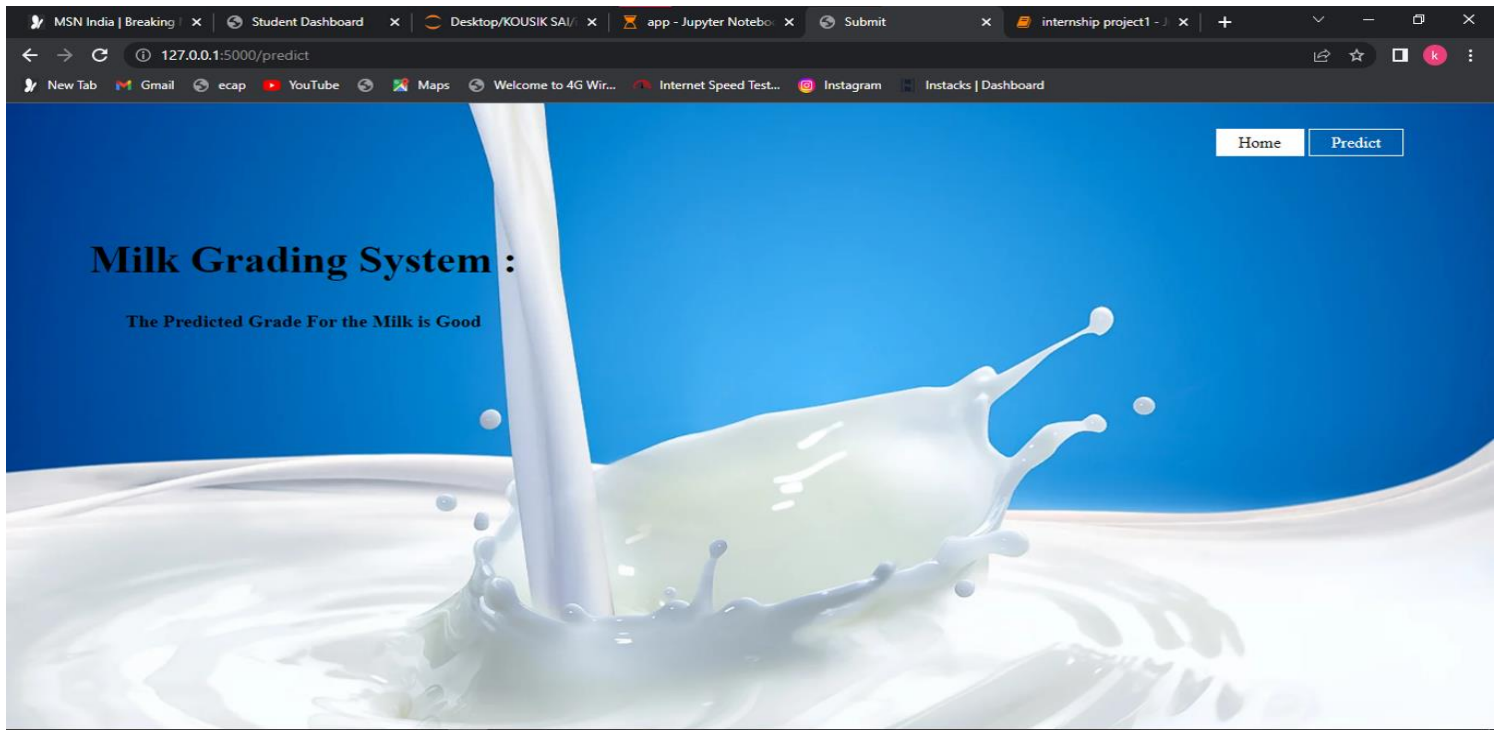
- To Know the quality of milk we have to enter the tested values of milk like pH , Temperature, Taste, Odor, Fat, Turbidity, Colour.
- After entering the required values click on the submit button. It will redirect to another page as shown in image below.
- By comparing the entered data values to prepared model it will predict the milk is Good/Bad/Moderate.

- Another output screenshots are:

Advantages :

- The predicted values are accurate.
- It will analyze the given milk sample.
- It is the advanced technology used for prediction.
- We can also get the graphical representation by using the data visualization.

Disadvantages :

- Accuracy should be maintained to get exact values.

APPLICATIONS:

- Milk prediction for drinking, household livelihoods, food security and nutrition etc.
- To produce manufactured Dairy products.

CONCLUSION :

We have created an application of **MILK GRADING SYSTEM** (predict the quality of milk), which as inputs pH , Temperature, Taste, Odor, Fat, Turbidity, Colour. Which shows output as **GOOD / MODERATE / BAD**. So, we can easily predict the quality of milk by using this application.

FUTURE SCOPE :

The paper presents the economical solution to avoid contamination of milk in rural and urban areas. The quality of milk is monitored using reduction test and the future prediction of milk contamination is achieved using machine learning algorithms. This system helps to save the health of the humans who are drinking the milk regular. The future scope of this project is to predict the pasteurized milk quality and to avoid contaminants.

REFERENCES :

https://repository.ipb.ac.id/handle/123456789/28783

http://ecoursesonline.iasri.res.in/mod/page/view.php?id=6059

https://biodatamining.biomedcentral.com/articles/10.1186/s13040-019-0200-5