

COMP9321 – Assignment3

Team FreshCoke

- Aim of the service:

This dataset is related to breast cancer with digit values, which is measured by cancer images. We provide a clear and friendly interface for users to review the collected cancer data. At the same time, users can predict diagnosis (either malignant or benign) by input feature values related to the cancer cell. Thus, users can use this system to study the features of cancer and make predictions.

- The Datasets:

Breast Cancer Wisconsin (Diagnostic) Data Set from Kaggle.

- Communication channel:

Daily communicating through the “Google Hangouts”.

- Code repository:

Code repository on Bitbucket: https://bitbucket.org/comp9321_freshcoke/comp9321_ass3/src

- Each member's role in the project:

* Machine Learning:

z5140081-Tianwei zhu; z5084093-Haoxiang Zhao

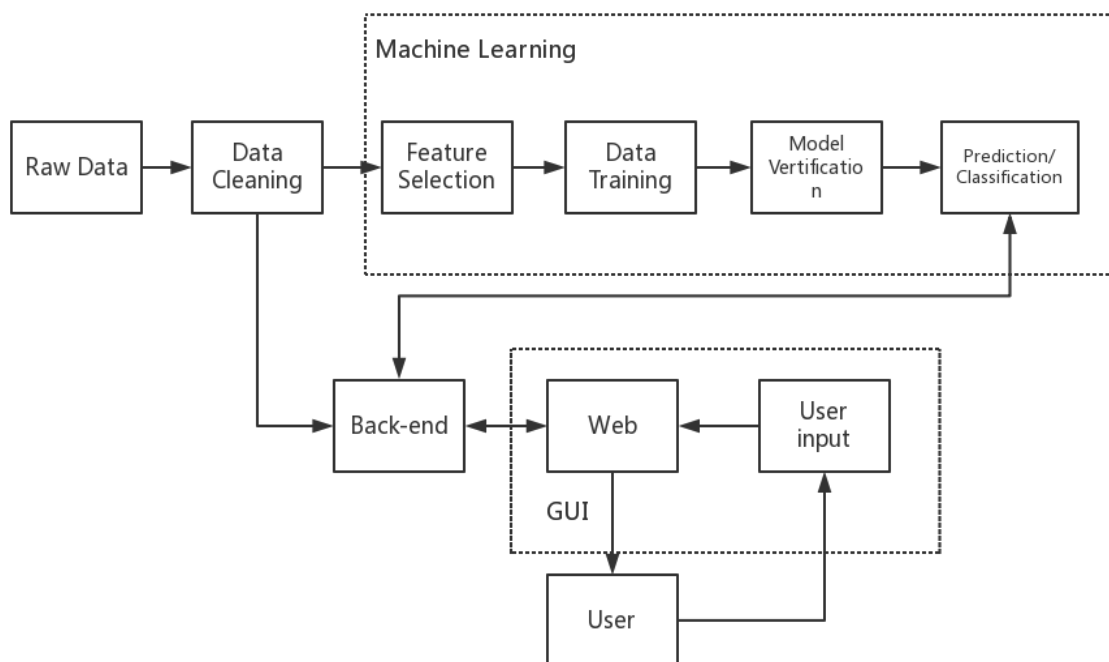
* REST APIs:

z5149974-Yuchen Xiao; z5147201-Zhenyang Lu; z5180103 - Yubo Sun

* Front-end / UI:

z5149974 - Yuchen Xiao; z5147201-Zhenyang Lu; z5180103 - Yubo Sun

- Project documentation:



Data cleaning: Some datasets contain useless or uncomplete data, we do this step before make use of them.

Feature selection: To provide a more accurate machine learning result (prediction or classification), pre-processing features of data is important.

Data training and model verification: We use Cross-validation to test our model, then modify parameters in ML to reach a reasonable result.

Back-end: Flask is used to support HTML. This part is charge of transmitting data from web page and machine learning.

Web page and user input: There will be a HTML based website, which allow users to review the dataset and input values to make prediction/classification.

- Web design:

