K. N. Toosi University of Technology

Faculty of Physics
Educational Group of
Atomic-Molecular and Astronomy

Special Topics I Final Projects
(Project 5 - Blood Donation)

Dr. Mohammad Hossein Zhoolideh

Academic Year 1403-1404

Overview

The Red Cross reported in January 2019 that "Nationwide, the Red Cross saw 27000 fewer blood donations over the holidays than they see at other times of the year." Blood collection managers face a major and ongoing challenge when predicting the blood supply. By using machine learning to identify patterns in the data, more lives can be saved by helping to forecast future blood donations.

Please download the data from this link. You will be working with information gathered from the Blood Transfusion Service Center donor database in a city of Taiwan, for this project. The dataset includes a random sample of 748 donors.

Note: The given data is raw. To answer this question, you must first preprocess the data using the Pandas package.

Important Points

Be sure to

- Leave appropriate comments for different parts of your code.

- Completely explain about the algorithm(s) you use to answer this question.

- Use **model selection**, **feature engineering** and **feature scaling** in your code.

- Measure your model performance using model evaluation metrics and interpret the obtained result(s).

- If you used a specific book or article in your project, mention it in your notebook.

A part of your score will be allocated to these items.

* You should write all the steps of your project in the **Jupyter notebook** and upload it as a file with the **.ipynb** extension on the vc site.