Project design phase

Part -1

Proposed solution:

- Water is one of the largest resources on earth. People need water to sustain life, including drinking water. It is important to know whether drinking water – human life resource – is enough for everyone now and in the future.

- Also, the authors build a model that allows people to predict the potability of a water resource by the data of each factor of that resource. According to the research, the features of water are not related to each other. All the features should meet a specific standard in order to get potable water.

Methodology:

- Machine learning models
- Data
- Training process

The establishment of simulation model:

- The overall data clearly shows that water resources are 61% non-potable, while only 39% are potable. This contrast is a strong indication that the amount of water currently potable is under serious threat. More and more pollution which breeds bacteria occupies large space. Such bacteria pollute the groundwater and air, and then threaten safety of drinking water. At the same time, because of over-exploitation of groundwater, some areas have produced phenomena like shrinkage of lakes and disappearance of the mud flats. These all would cause the reduction of water storage capacity and water self-purification. This will deteriorate the status of water portability.
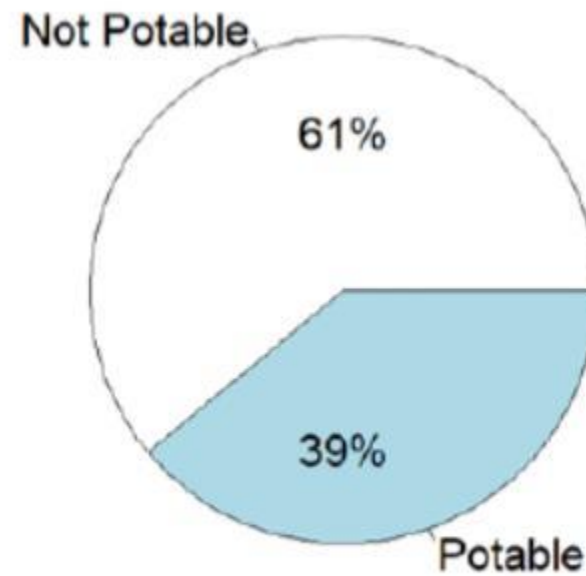
Pie chart:

Figure 1. Pie chart of the water potability

Solution architecture:

- Drinking Water Distribution Mains carry portable water from many water resources like reservoirs, river, and water tanks for industrial, commercial and residential consumers purposes through complex buried pipe networks.

- The proposed work represents the development of novel low cost fuzzy based water quality monitoring system using wireless sensor networks which is capable of analyzing the physiochemical parameters of water quality such as pH, temperature, conductivity, oxidation reduction potential and turbidity.

Solution architecture diagram:
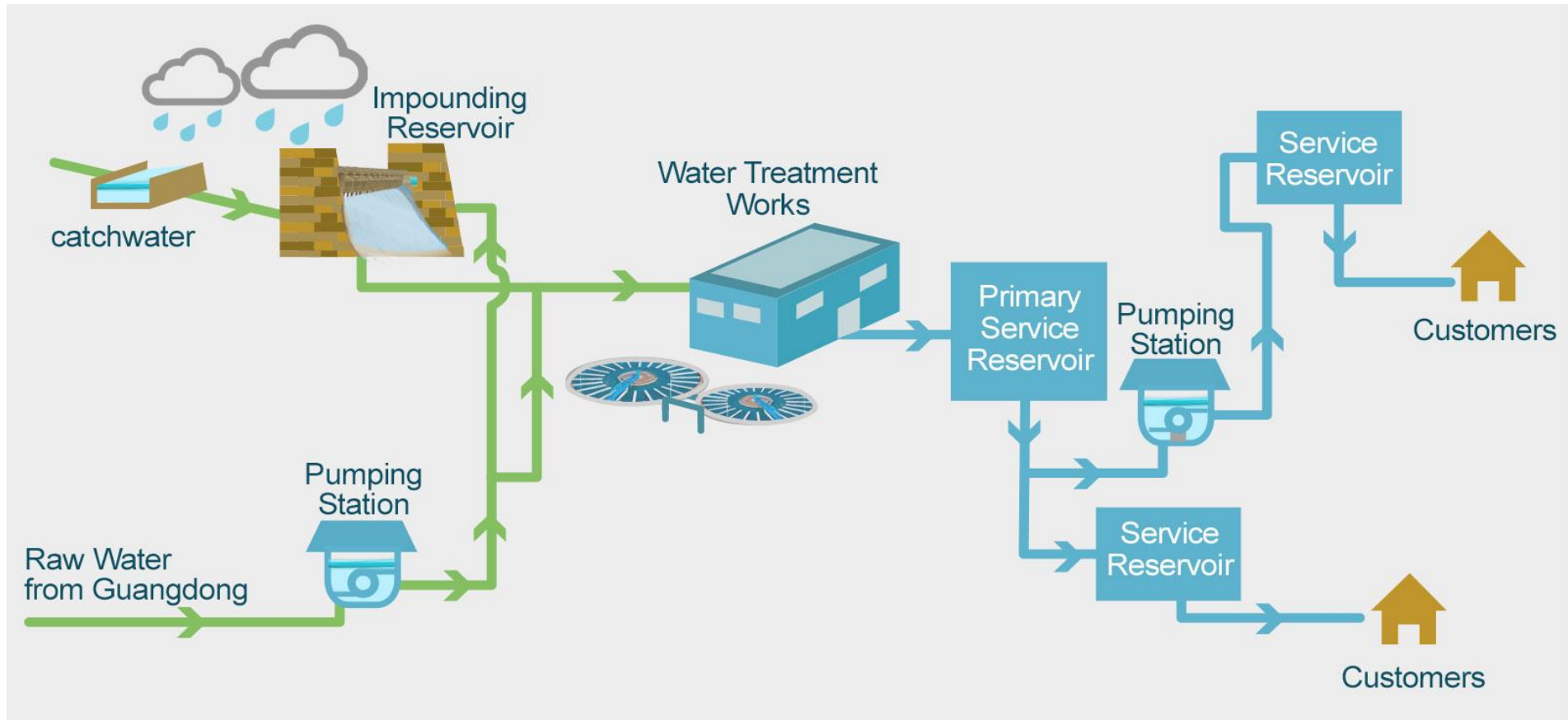
Conclusion:

- Water quality relates to everyone's life. The adequacy of water resources not only affects people's life safety, but also deeply affects the development and stability of society.

- By using statistical models, including binomial regression and the K-nearest neighbor algorithm (K-NN), the authors got a model in the end which can help to predict whether a water resource is potable. Although HLGOF of the logistic regression model is quite high, which indicates that there are not much differences between real values and predicting ones, the significance levels of most features treated by logistic regression are also very high. As a result, the false positive rate is high. The model still needs improvement