Assignment 1 for COMP1004
Name: Wang Ruijie
Student NetID: 22103808d

Q1
a)
i) First, I renamed the two default classes, and opened the webcam to separately capture the images of drinking water and not drinking water. Around 70 images each class are captured in each class to ensure a comprehensive training for the model to improve its accuracy. Then click to make the model trained by my computer.

Not drinking water ✏

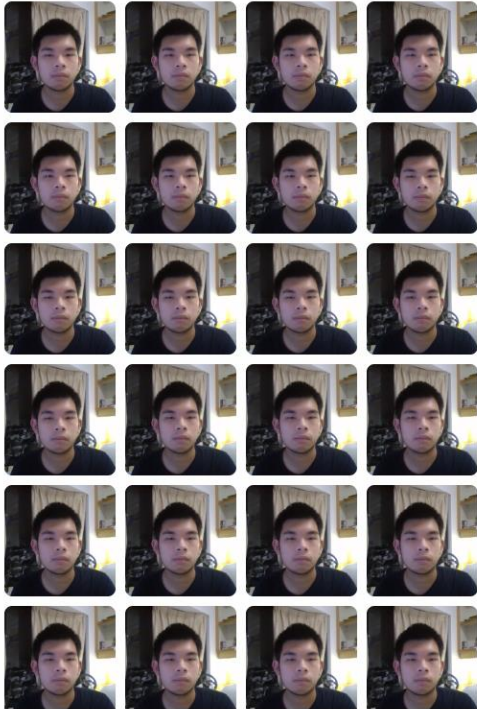I tested the model by using the images of drinking water and not drinking water, and it turned out to be fine.
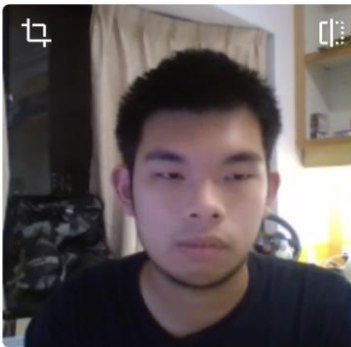
ii) The ways to confuse the model are to hold the bottles and to pretend to grasp something. They turned out to be false positive. (Let "Drinking water" be positive, and let "Not drinking water" be negative")
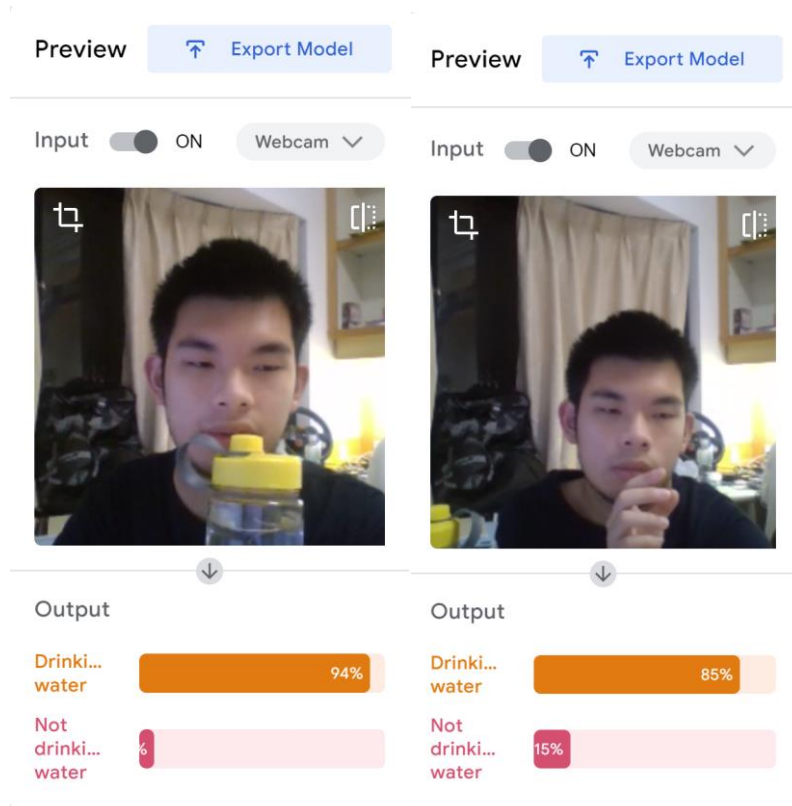
iii) Another method of confusing the model is to the way I drink water. It turned out to be false negative.
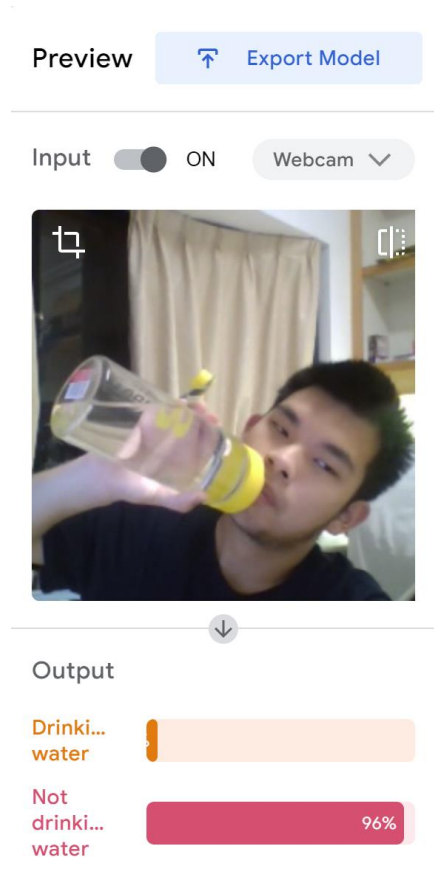
iv) To improve the model, we are supposed to add more images that are taken from different angles or contain different ways of drinking water. Also, we need different bottles.
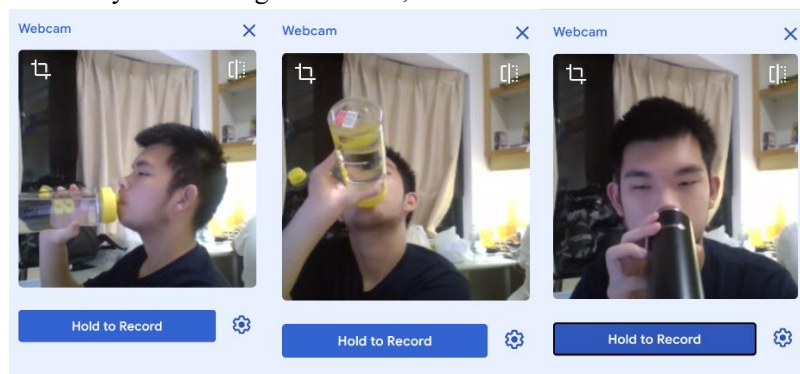
The model is also expected to identify the confusing images by adding them to the class of "Not drinking water".

Eventually, we could see that the model works better than before. All these errors mentioned above are corrected now.

b)

i) The target variable is the classification of the homes according to the city in which it is located.

ii) It depends. When the prediction accuracy is not enough satisfactory because of lack of recursion, it could be effective to add more layers to improve the accurac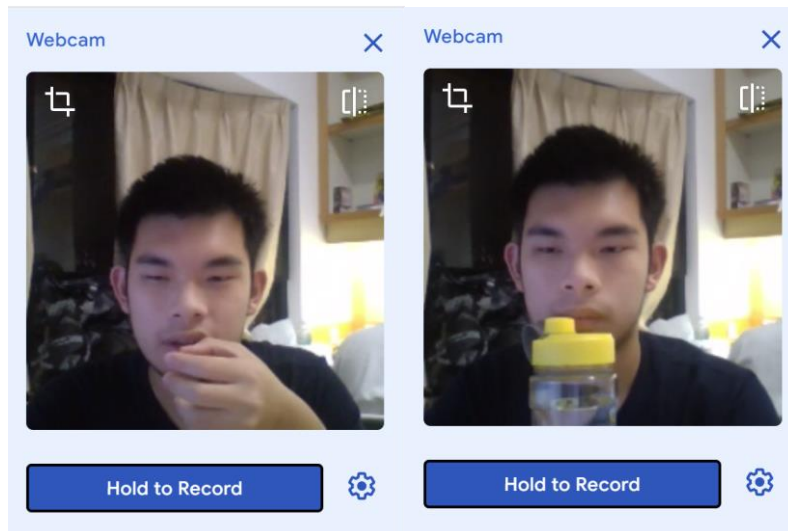y. However, more layers might result in overfitting if the details in the layers added are irrelevant, and therefore the prediction accuracy might decrease.

Q2

a)

i) There exists intended or unintended bias on the ground of race and sex in facial recognition systems, including lower recognition accuracy when it comes to people of color or females.

ii) The cause should be the gender and race skews in the datasets used for machine learning, which includes a preference for people in power or creators of the programs who are overwhelmingly white and principally men.

iii) First, the accuracy and the efficiency of the recognition should be guaranteed to make the application of these systems could work smoothly. Second, the facial recognition systems are supposed to avoid using skewed datasets and displaying skewed performance on the grounds of gender and race to ensure equality. Third, abusing of the technology of facial recognition are expected to be stopped, for these systems are responsible to protect user's privacy from being harmed. This requires the systems strictly obey the law and enhance information security.

b)

Value of b	-31 or -32
Classification accuracy	80%
Number of false positive(s)	1

Number of false negative(s)	1

Q3
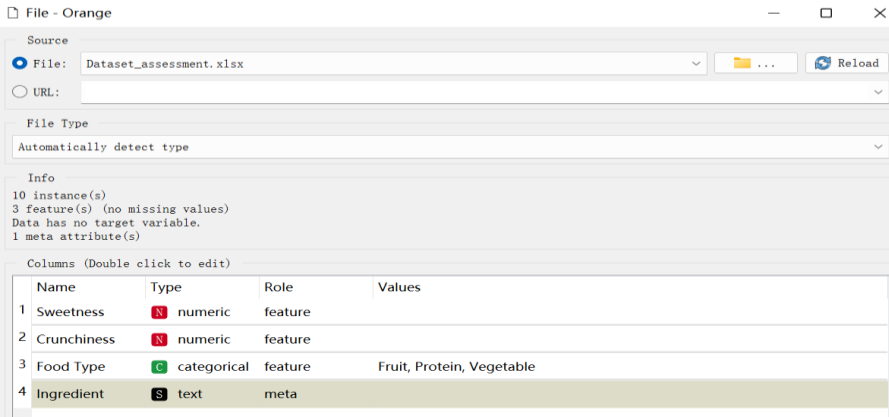
a)

1) First, create a new Excel document and upload all the data.
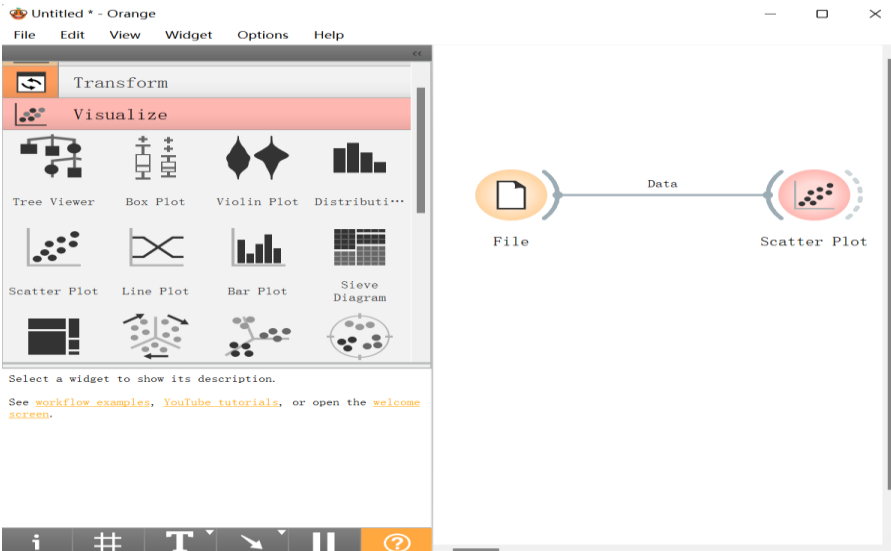
	A	B	C	D
1	Ingredient	Sweetness	Crunchiness	Food Type
2	Apple	10	9	Fruit
3	Bacon	1	4	Protein
4	Banana	10	1	Fruit
5	Carot	7	10	Vegetable
6	Celery	3	10	Vegetable
7	Chesese	1	1	Protein
8	Grape	8	5	Fruit
9	Green Bean	3	7	Vegetable
10	Nuts	3	6	Protein
11	Orange	7	3	Fruit

2) Second, create a "File" widget and upload the dataset to the file in Orange.

3) Third, choose the "Scatter plot" widget and connect it with the "File" widget.

4) And then, modify the options, and eventually a completed scatter plot is displayed on my screen.
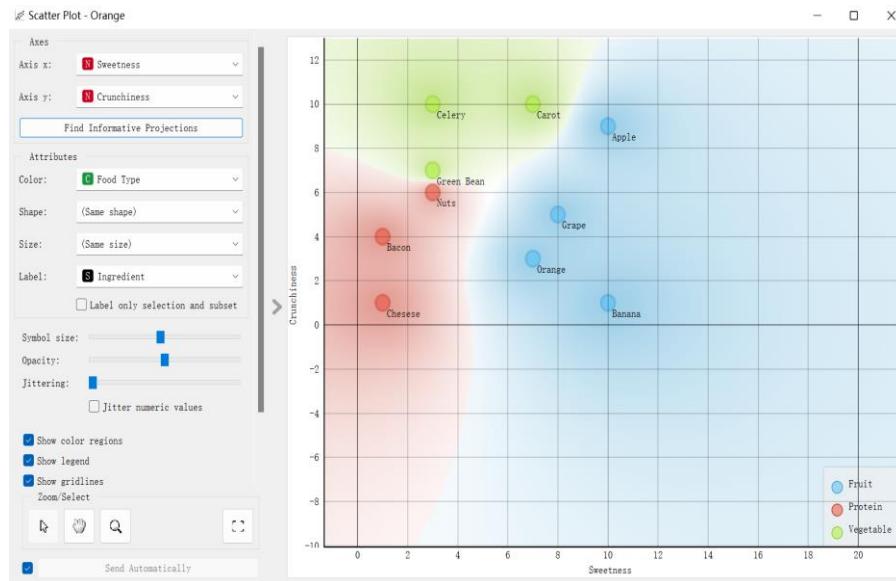
b)

**1)** First, create a new Excel document which contains the test data.

	A	B	C	D
1	Ingredient	Sweetness	Crunchiness	Food Type
2	X	8	9	

**2)** Similarly, create another "File" widget and upload the data for test.

File (1)

3) What to do next is to choose a "Select Columns" widget and choose "Food Type" as the target variable, and "Crunchiness" and "Sweetness" as the features, for the roles of the attributes of the dataset.
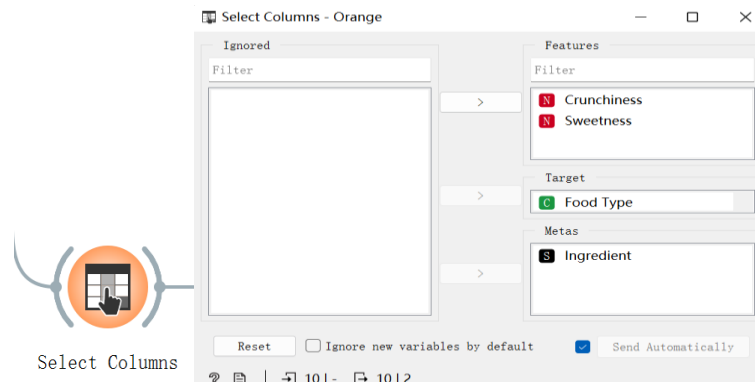

4) Then, create a "kNN" widget and connect it with "Select Columns", and remember to connect

the previous "File" bidget with the "Select Columns" in order to complete this model that could be learned by our computers. As required, we are supposed to modify the options in "kNN" as followed.
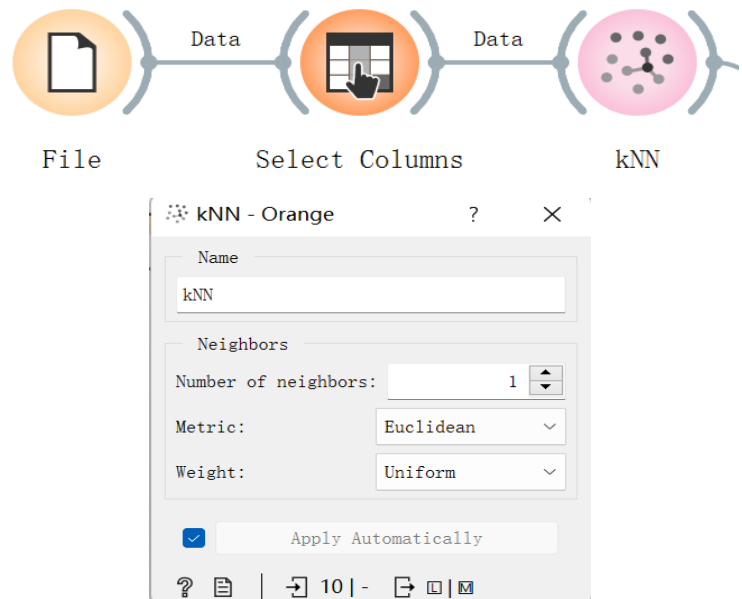
5) To make a prediction, we should choose "Prediction" widget, and connect it with the model and the test data separately.
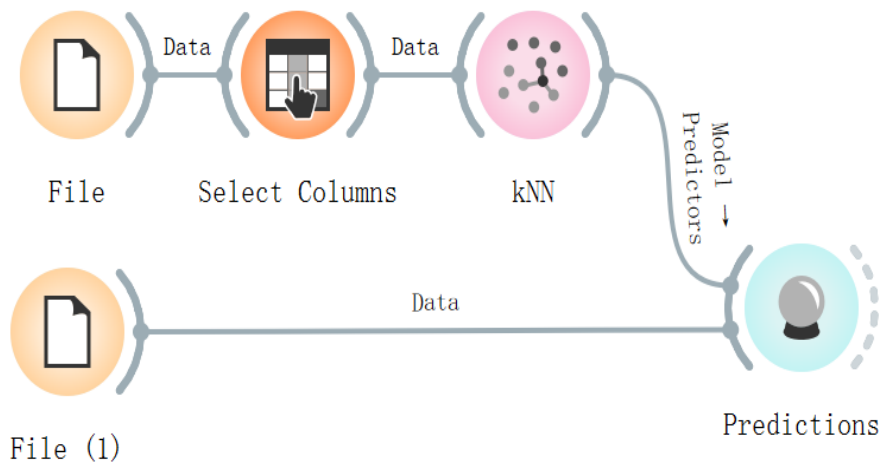
6) We could finally have the prediction outcome. The model tells us that it should be a kind of vegetable.

	kNN	Ingredient	Sweetness	Crunchiness
1	Vegeta...	X	8	9