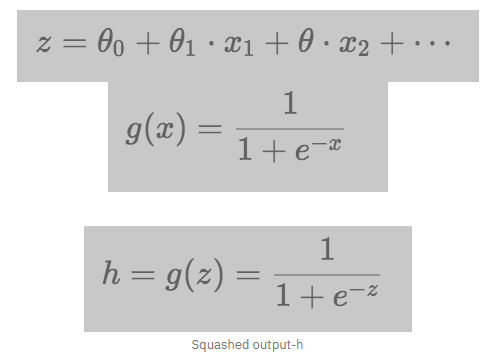
**Introduction:**

Logistic regression is the most famous machine learning algorithm after linear regression. In a lot of ways, linear regression and logistic regression are similar. But the biggest difference lies in what they are used for. Linear regression algorithms are used to predict/forecast values, but logistic regression is used for classification tasks. There are many classification tasks done routinely by people. For example, classifying whether an email is a spam or not, classifying whether a tumor is malignant or benign, classifying whether a website is fraudulent or not, etc. These are typical examples where machine learning algorithms can make our lives a lot easier. A simple, rudimental and useful algorithm for classification is the logistic regression algorithm.



**Dataset:**

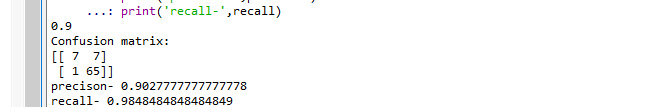
The dataset contains several parameters which are considered important during the application for master’s Programs. The parameters included are: 1. GRE Scores (out of 340) 2. TOEFL Scores (out of 120) 3. University Rating (out of 5) 4. Statement of Purpose and Letter of Recommendation Strength (out of 5) 5. Undergraduate GPA (out of 10) 6. Research Experience (either 0 or 1) 7. Chance of Admit (ranging from 0 to 1)

**Method:**

We need do some pre-processing before the data can be fed for logistic regression. In the dataset, the values are in terms of probabilities between 0 and 1. For logistic regression, we need to convert these values to binary i.e. either 0 or 1. To do this, we take a threshold of 0.6. Any value greater than the threshold is rounded off to 1 and values less than threshold and rounded to 0. Now, the dataset is ready for processing. We consider columns 1 to 8 as factors for predicting the admission. The data is split into training (80%) and test (20%) and then perform logistic regression. The accuracy of the model is checked by means of a confusion matrix, prediction score, precision score and recall score.

**Result & Interpretation (in terms of Confusion matrix and accuracy metrics):**

The accuracy we got after using logistic regression on the dataset was 90%. The confusion matrix and accuracy scores screenshot are attached below.



**Conclusion & Scope:**

Hence, using Logistic Regression we can predict if something like whether customer will stay or not, whether someone will get admission or not. The drawback is that all the data that is used for prediction must be encoded in terms of numbers.