# **Logistic Regression**

Krishna Kumar Veeraputhiran

Grand Canyon University

DSC 540 – O500 – Machine Learning for Data Science

Dr. Aiman Darwiche

August 25,2021

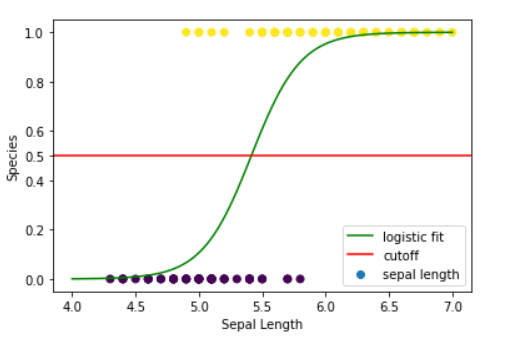
# **Logistic Regression**

Logistic regression is a model which is used to model the probability of a discrete outcome based on the input variable. Logistic regression is a transformation of a linear regression using the sigmoid function. In simple words the predictions are transformed to binary outcomes using the logistic function.

The logistic regression prediction probabilities can be given as below,

**p= eβ0+β1x / (1+ eβ0+β1x)**

In the above equation p is the predicted probability output and β0 is the intercept and β1 is the coefficient for the input variable (x). As we can see above the equation **β0+β1x** is of the form **mx + c** which represents the linear equation. The Logistic regression probability outcome when plotted against the input variable provides a non linear relationship (sigmoid curve). The below plot is an outcome for Iris dataset to predict between the species Setosa and Versicolor while using the sepal length as the input variable.



As we can see above the Predicted outcome has a non-linear relationship with the input variable and hence we can say that the logistic regression is a nonlinear regression problem.

In a logistic regression the Odds can be calculated from the base probability function. The odds can be given in relation to the probability as below,

**Odds = p/1-p**

The above equation can be further simplified as below,

**Odds = p/1-p = eβ0+β1x**

Odds is the ratio of probability of a particular outcome occurring to the probability of it not occurring. With respect to our example, the odds of being Setosa is the ratio of probability of being a Setosa to the probability of not being setosa. The log of odds yields a function that goes from – infinity to infinity and this is called logit. According to Gopal (2019), the transformation to log(odds), with logit as dependent variable, converts the problem to a linear regression task. So this can be formulated as,

**Log(odds) = log(p/1-p) = β0+β1x**

Hence we can say that the logistic regression can be treated as an equivalent to linear regression while using the logit function.

# **References**

Gopal, M. (2019). *Applied machine learning*. McGraw-Hill Education.