• Maximum Likelihood Estimation (MLE)

Machine Learning

Dr. Jagendra Singh

MAXIMUM LIKELIHOOD ESTIMATION

- Maximum Likelihood Estimation (MLE) is a probabilistic based approach to determine values for the parameters of the model.

- MLE is a widely used technique in machine learning, time series, panel data and discrete data.
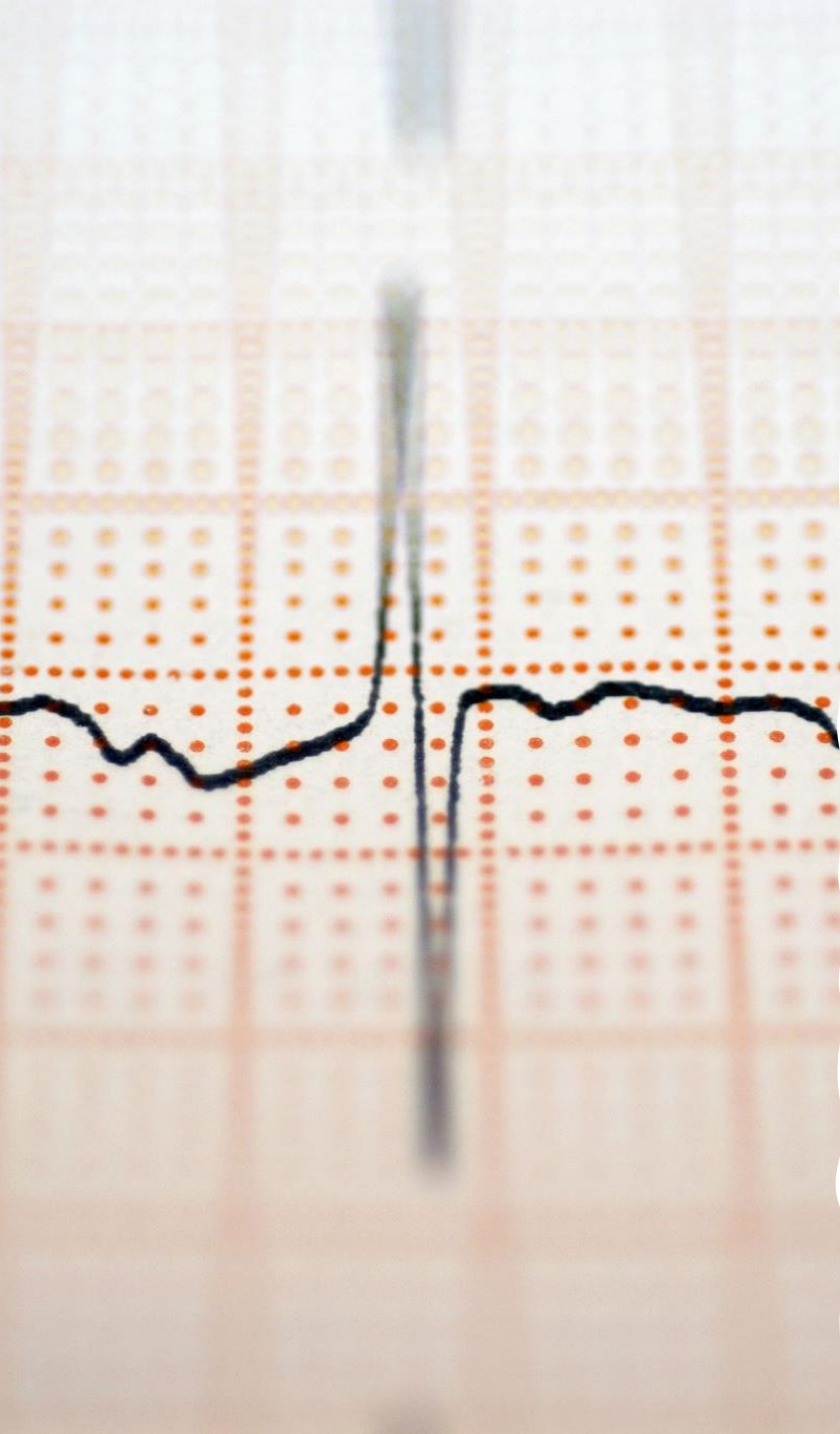
MAXIMUM LIKELIHOOD

- The motive of MLE is to maximize the likelihood of values for the parameter to get the desired outcomes.

- Following are the topics to be covered.

- What is the likelihood?

- Working of Maximum Likelihood Estimation

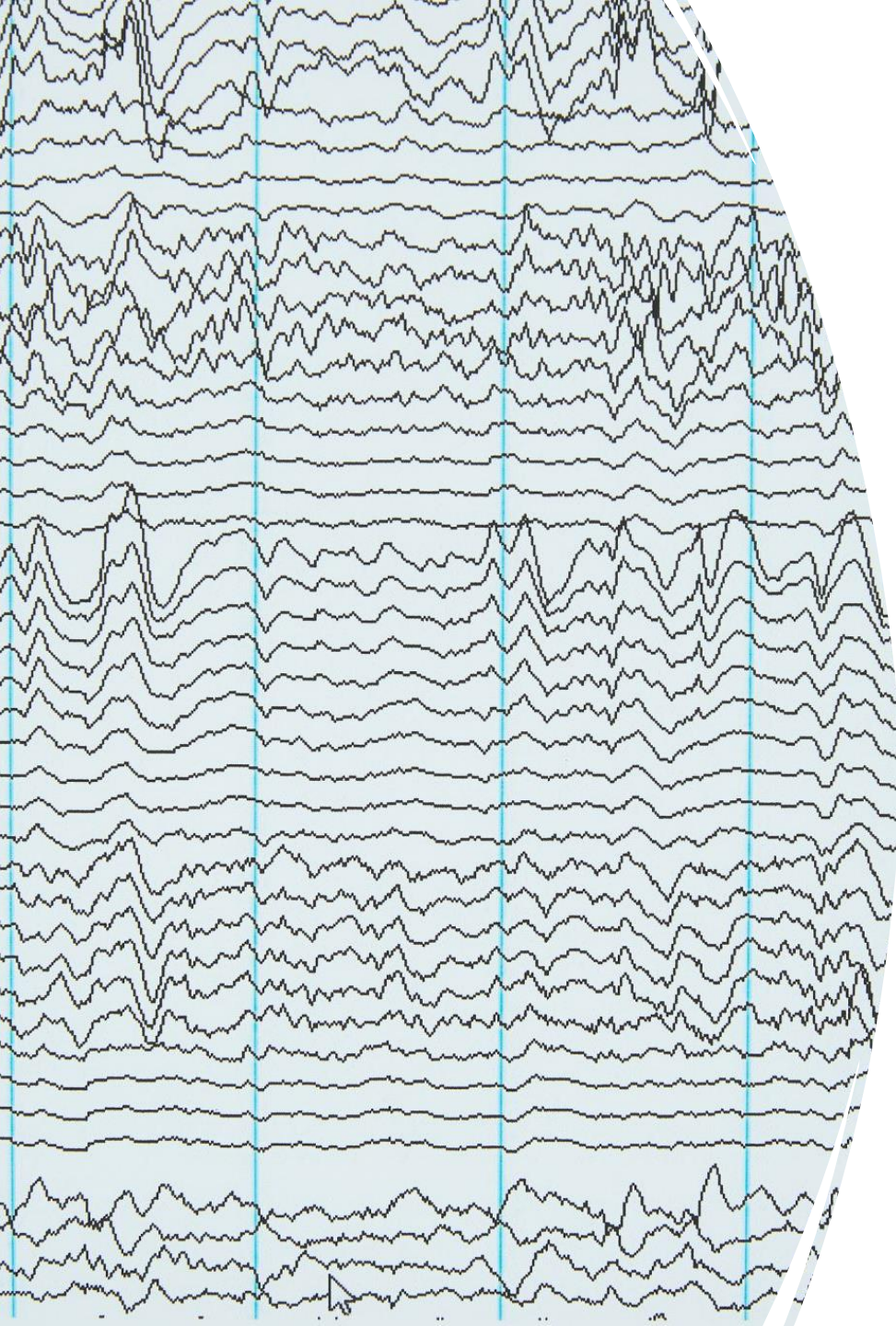- Maximum likelihood estimation in machine learning

MAXIMUM LIKELIHOOD

- For understanding the concept of Maximum Likelihood Estimation (MLE) we need to understand the concept of Likelihood first and how it is related to probability.

WHAT IS THE LIKELIHOOD?

- The likelihood function measures the extent to which the data provide support for different values of the parameter.

- It indicates how likely it is that a particular population will produce a sample.

MAXIMUM LIKELIHOOD

- The likelihood function is different from the probability density function.

- Likelihood describes how to find the best distribution of the data for some feature or some situation.

MAXIMUM LIKELIHOOD

- Probability describes how to find the chance of something given a sample distribution of data.

- Let's understand the difference between the likelihood and probability density function with the help of an example.

- Consider a dataset containing the weight of the customers. Let's say the mean of the data is 70 & the standard deviation is 2.5.

MAXIMUM LIKELIHOOD

- When [Probability](#) has to be calculated for any situation using this dataset, then the mean and standard deviation of the dataset will be constant.

- But in the case of Likelihood, mean and standard deviation of the dataset will be varied to get the maximum likelihood for weight > 70 kg.
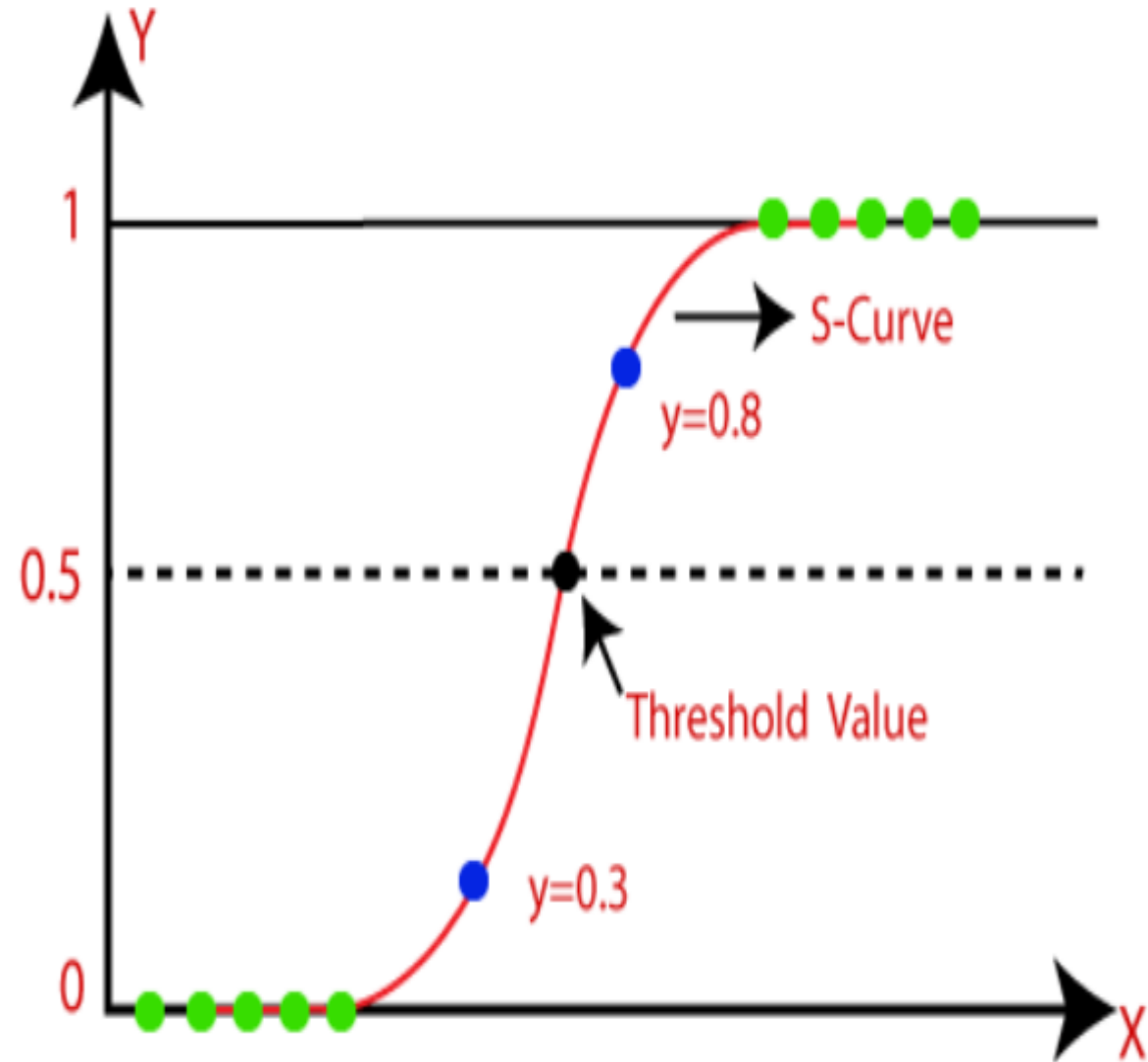
WORKING OF MAXIMUM LIKELIHOOD ESTIMATION

- The maximization of the likelihood estimation is the main objective of the MLE.

- Let's understand this with an example.

- Consider there is a binary classification problem in which we need to classify the data into two categories either 0 or 1 based on a feature called "salary".

WORKING OF MAXIMUM LIKELIHOOD ESTIMATION

- So MLE will calculate the possibility for each data point in salary and then by using that possibility, it will calculate the likelihood of those data points to classify them as either 0 or 1.

- It will repeat this process of likelihood until the learner line is best fitted. This process is known as the maximization of likelihood.

WORKING OF MAXIMUM LIKELIHOOD ESTIMATION

- The above explains the scenario, as we can see there is a threshold of 0.5 so if the possibility comes out to be greater than that it is labelled as 1 otherwise 0.

- Let's see how MLE could be used for classification.

MAXIMUM LIKELIHOOD ESTIMATION IN MACHINE LEARNING

- MLE is the base of a lot of supervised learning models, one of which is Logistic regression.

- Logistic regression maximum likelihood technique to classify the data.

- Let's see how Logistic regression uses MLE.

- MLE procedures have the advantage that they can exploit the properties of the estimation problem to deliver better efficiency and numerical stability.

MAXIMUM LIKELIHOOD ESTIMATION IN MACHINE LEARNING

- These methods can often calculate explicit confidence intervals.

- The parameter "solver" of the logistic regression is used for selecting different solving strategies for classification for better MLE formulation.

THANK YOU