Tutorial for using this in real project:

<https://towardsdatascience.com/kernel-density-estimation-explained-step-by-step-7cc5b5bc4517>

Citation:

<https://www.itm-conferences.org/articles/itmconf/pdf/2018/08/itmconf_sam2018_00037.pdf>

KDE allows flexible modeling of complex, non-linear distributions of features like crime locations and times.

Also choose naïve bayes model -> infer on past data rather than making summary

## Naïve Bayes Model

Naïve Bayes Model is a cheap learning algorithm which works by applying Bayes’ Rule, and it assumes that every feature is conditionally independent (Citation here). In the context of crime prediction, a set of crimes with types to be predicted is represented by a one-hot vector . Given features as evidence, the probability is the product of likelihood and prior distribution i.e.

It works well with high-dimensional data and is commonly used in criminal analysis (Citation here), because it summarizes the past data and thus make inference for the future, which align with our research goal. Compared with DT, it is relatively robust to large dataset, but it is limited by its independence assumption.

## Kernel Density Estimation

Kernel Density Estimation (KDE) is a non-parametric model that used to fit probability density function (PDF) for a random variable. It does not take any assumptions but uses two hyperparameters, kernel function and bandwidth to fit a curve for given data points.

KDE applied on Time:

KDE applied on Location (Coordinate):

KDE applied on both: