	Decision tree	Naïve Bayesian Classifier	Logistic Regression	KNN classifier
Input Parameter	The attributes of a tuple	The posterior probabilities of Hypothesis H based on additional information	The measure of the total contribution of all independent variables used in the model.	A new unknown tuple for which a class has to be assigned.
Principle	The attributes of a tuple are tested against the decision tree and a path is traced from the root to a leaf node which holds the prediction for that tuple	Given a tuple X, the classifier will predict that X belongs to the class having the highest posterior probability conditioned on X.	Given x representing the exposure to some set of risk factors, LR predicts the probability of occurrence of an event by fitting data to a logistic curve, f(x), which represents the probability of a particular outcome	Nearest-neighbor classifiers compare a given test tuple with training tuples that are similar and described by n attributes and are stored in n-dimensional space ➔ Find the k-nearest tuples from the training set to the unknown tuple
ALG. FORMULA	At start, all the training tuples are at the root. Then, tuples are partitioned recursively based on selected attributes. The test attributes are selected on the basis of a heuristic or statistical measure (e.g., information gain). We stop when all samples for a given node belong to the same class or there are no remaining attributes for further partitioning ➔ majority voting.	At start, compute P(C) The prior probability ofC_i. Each class can be computed based on the training tuples. Then compute each independent probability for attribute xi in reference with Class C and multiply P $(X_i\|$C). Example: The probability of class C to be yes is P(C) = 9/14. The amount of attribute Xi being test with respect to C is P$(x_i\|C_i)$ = $\frac{xi=test}{Ci=yes} = \frac{2}{9}$ Finally, compute P(X$\|C_i$)P(C_i) for each class ➔ The naïve Bayesian Classifier predicts C=yes for tuple X	Since x is defined as x =$\beta_0 + \beta_1 X_1 + \cdots + \beta_k X_k$ the LR is $\frac{e^{\beta_0+\beta_1 X_1+\cdots+\beta_k X_k}}{1+e^{\beta_0+\beta_1 X_1+\cdots+\beta_k X_k}}$ Estimate the parameters using the Maximum Likelihood Function and then by computing the partial derivatives of the log likelihood, equate each partial derivative to zero, and solve the resulting nonlinear equations	The ALG assigns for the unknown tuple the most common class among its k-nearest neighbor. When k=1 the unknown tuple is assigned the class of the training tuple that is closest to it. To measure the distance we can use the Euclidean distance $\sqrt{\sum_{i=1}^{n}(x_{1i}-x_{2i})^2}$ **Choose K:** If k=1 the classification will be 1:1 sensitive to other data. If k=n we'll suffer high noise. Go through various K's and choose one giving lowest misclassification error!