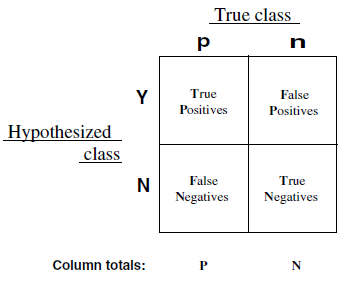
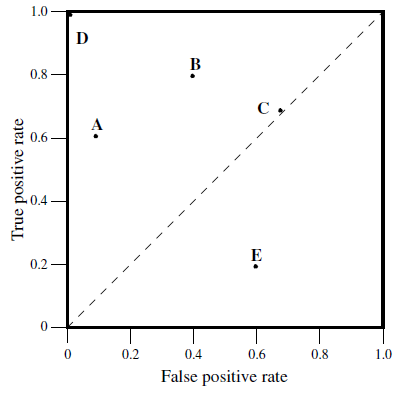
## Zusammenfassung – An introduction to ROC analysis

Tom Fawcett führt den ROC-Chart (Receiver Operating Characteristic oder dt. Grenzwertoptimierungskurve) zum Jahre 1975 zurück. Nach der geschichtlichen Exkursion wird festgestellt, dass der ROC-Chart in den vergangenen Jahren immer präsenter wird.

Um den ROC-Chart besser zu verstehen, wird zunächst die Vier-Felder-Matrix (Siehe Abbildung 1) eingeführt. Diese 2x2 Matrix beinhaltet zwei Spalten, die die wahre Klassifizierung, und zwei Zeilen, die die Hypothesenklassen beinhalten. Dabei beinhaltet die wahre Klassifizierung die echte Zuordnung zur Klasse und die Hypothesenklasse die Zuordnung durch ein Modell. Dabei gilt zu beachten, dass die folgenden zwei Formeln aus der Vier-Felder-Matrix für den ROC-Chart relevant sind.

Abbildung

Der ROC-Chart ist ein zweidimensionaler Graph, der auf der y-Achse die Werte der *tp rates* und auf der x-Achse *fp rate* visualisiert. Alle Werte eines Klassifikators über der gestrichelten Diagonalen konnten durch das Model positiv klassifiziert werden. Befindet sich ein Klassifikator auf der Diagonalen, könnte die Zuweisung der Werte auch zufällig entstanden sein. D.h. wenn ein Klassifikator auf der Diagonalen liegt, enthält er mit großer Wahrscheinlichkeit keine nützlichen Informationen. Der untere Teil bedeutet, dass der Klassifikator schlechter als zufällig abschneidet. Eine Möglichkeit Informationen aus den unteren Teil zu gewinnen, besteht darin die Klassifikatoren zu negieren. Das übergeordnete Ziel des Graphens ist das Erkennen, welche Klassifikatoren nützlich sind. Besonders gewinnbringend sind die Klassifikatoren in der linken oberen Ecke (0,1).