

SUPPORT VECTOR MACHINES

A SUPPORT VECTOR MACHINE IS USED TO BUILD BINARY CLASSIFIERS

POINTS: EMAILS
CATEGORIES: SPAM OR HAM

(THIS MEANS THAT GIVEN A SET OF POINTS, A SUPPORT VECTOR MACHINE WILL CLASSIFY THOSE POINTS INTO 2 CATEGORIES)

IN ADDITION, SUPPORT VECTOR MACHINES MAKE THEIR CLASSIFICATION DECISION ON THE BASIS OF A "LINEAR FUNCTION" OF THE POINT'S COORDINATES

IF A POINT IS $X = (x_1, x_2, x_3, \dots, x_n)$

A LINEAR FUNCTION IS SOMETHING LIKE

$$f(X) = ax_1 + bx_2 + cx_3 + \dots + zx_n$$

LASTLY, SVMs INVOLVE AN EXPLICIT TRAINING STAGE WHEN THE MODEL "LEARNS" FROM A SET OF TRAINING DATA

THE SUPPORT VECTOR MACHINE WILL RUN A TEST LIKE: IF $f(X) > 0$, EMAIL IS SPAM, ELSE EMAIL IS HAM

ALSO, SUPPORT VECTOR MACHINES DO NOT INVOLVE EXPLICIT ASSUMPTIONS ABOUT THE PROBABILITY DISTRIBUTIONS OF THE POINTS

(NAIVE BAYES CLASSIFIERS, FOR INSTANCE, ASSUME THAT THE DISTRIBUTIONS OF DIFFERENT FEATURES ARE INDEPENDENT)

"A SUPPORT VECTOR MACHINE IS A SUPERVISED MACHINE-LEARNING APPROACH USED TO BUILD LINEAR, NON-PROBABILISTIC BINARY CLASSIFIERS"

