Here the SVM classifier is expected to find a hyper-plane to separate testing examples as posi- tive and negative. Wu et al. (2004) extend the basic SVM to a probabilistic version. Its goal is to estimate

Name	Description
Position	1/sentence no.
Doc_First	Whether it is the first sentence of a document
Para_First	Whether it is the first sentence of a paragraph
Length	The number of words in a sentence
Quote	The number of quoted words in a sentence

where A and B are estimated by minimizing the

negative log-likelihood function using training

data and their decision values f. Then p is ob- $_i$

$$y_i(w^T\varphi(x_i)+b)\geq 1-\xi_i$$