



Istituto di Analisi dei Sistemi ed Informatica "Antonio Ruberti"
Consiglio Nazionale delle Ricerche

colloquia@iasi



*"La mente è come un paracadute.
Funziona solo se si apre"
A. Einstein*

Formulazioni continue e discrete del problema di selezione degli attributi (Feature Selection) nel contesto SVM

Prof. Manlio Gaudioso

Dipartimento di Ingegneria Informatica, Modellistica, Elettronica e Sistemistica, Università della Calabria

Abstract

La scelta degli attributi più significativi nella classificazione binaria in contesto Support Vector Machine (SVM) viene usualmente modellata come problema di ottimizzazione sparsa. Quest'ultima consiste nella ricerca di un punto di minimo di una funzione di molte variabili caratterizzato da un numero piccolo di componenti diverse da zero. In questo contesto, quindi, la pseudo-norma L_0 gioca un ruolo importante.

Nel seminario viene analizzata la possibilità di esprimere proprietà di tale pseudo-norma L_0 ricorrendo a quelle di una norma poliedrale nota come k -norma e definita come somma di componenti massimali (in modulo) di un vettore. In generale, condizioni sui valori della pseudo-norma L_0 sono esprimibili come vincoli su differenze di k -norme, e, quindi, in forma di differenza di funzioni convesse (vincoli DC).

Sulla base di tali proprietà vengono definiti modelli di ottimizzazione sia di tipo continuo che discreto e vengono presentati i risultati della loro implementazione nell'ambito SVM per dataset comunemente adottati come benchmark.