

A gépi tanulás egyes konvex optimalizációs feladatainak közelítő megoldása bázisfüggvények ritka kombinációjával

Kovács Kornél¹ és Kocsor András²

¹MTA-SZTE Mesterséges Intelligencia Kutatócsoport, Szeged Aradi vrt., H-6720,
e-mail: kkornel@inf.u-szeged.hu

²MTA-SZTE Mesterséges Intelligencia Kutatócsoport, Szeged Aradi vrt., H-6720,
e-mail: kocsor@inf.u-szeged.hu

A gépi tanulási feladatok sok esetben olyan klasszifikációs, illetve regressziós problémára vezetnek, amelyeket érdemes konvex optimalizáció formájában megoldani. Tekintsük a klasszifikációs feladat következő általános, konvex optimalizációs formáját:

$$\inf_{f(\underline{x}) \in \text{Span}(S)} \sum_i L(f(\underline{x}_i) y_i),$$

ahol S folytonos függvények egy véges halmaza, és $L: R \rightarrow R$ módszerenként változó konvex veszteség-függvény. Belátható, hogy a fenti forma a következő alakra hozható:

$$\inf_{\underline{\alpha}} g(\underline{\alpha}), \quad g(\underline{\alpha}) = \sum_i L \left(\sum_j \alpha_j f_j(\underline{x}_i) y_i \right).$$

Megmutatjuk, hogy ez a formalizmus magában foglal több, széles körben használt gépi tanuló algoritmust, mint például bizonyos „boosting” (AdaBoost, logisztikus regresszió), és Support Vector Machine (Least-Squared, Smooth) változatokat. A célfüggvény optimális értékének meghatározását végezhetjük nemlineáris Gauss-Seidel (GS) iterációval, amely a paramétervektornak egyetlen komponensét változtatja meg minden lépésben. Ha $\nabla g(\underline{\alpha})$ Lipschitz-folytonos, akkor az iteráció konvergenciája bizonyítható. A GS iteráció alkalmazását indokolja az alacsony memóriagigénye, mégis a gyakorlatban felmerülő igen nagy problémákra a megoldás kivihetlenné válik a rendkívül sok iterációs lépés miatt.

Mindezen érvek miatt természetesen adódik az igény közelítő megoldást szolgáltató (heurisztikus) eljárások alkalmazására. Egy ilyen eljárás családot fogunk definiálni, amelyek gyorsan és hatékonyan határoznak meg – a klasszifikációs szempontokat tekintve – olyan szuboptimális megoldásokat, amelyek annak ellenére, hogy nem optimálisak jól viselkednek egy adott klasszifikációs feladat esetén. Az eljárásainkat a gépi tanulás egy részterületét képező ún. tulajdonság-szelekciós módszerek felhasználásával fogjuk definiálni. A felhasznált eljárások a következők lesznek: szekvenciális előretartó kiválasztás, (l,r)-keresés, szekvenciális lebegő előretartó kiválasztás és mindezek általánosítása.

A javasolt módszer család olyan közelítő megoldásokat keres, amelyek esetén a paramétervektornak maximálisan csak előre rögzített számú nemzérus eleme van. A módszereink létjogosultságát kísérleti úton igazoljuk a gépi tanulási algoritmusok tesztelésére széles körben használt UCI Repository adatbázis-gyűjtemény egyes problémáinak hatékony megoldásával.