

Statisztikus tanuláselmélet: klasszifikáció és regresszió sztochasztikus garanciákkal

Csáji Balázs Csanád

MTA SZTAKI: Magyar Tudományos Akadémia
Számítástechnikai és Automatizálási Kutatóintézet
csaji.balazs@sztaki.mta.hu

A statisztikus tanuláselmélet¹ egyik alapkérdése, hogy megfigyelések egy mintájából hogyan következtethetünk induktívan az adatokat generáló rendszerre úgy, hogy minél jobb legyen a modellünk általánosító képessége. A tanuláselméletben – az aszimptotikus eredményekkel szemben – előtérbe kerülnek a kismintás statisztikai tulajdonságok valamint fontos cél az a priori információk minimalizálása; így a fókusz az eloszlás-független és nem-aszimptotikus garanciákkal rendelkező konstrukciókon van.

A nemparametrikus tanulási módszerek egy jelentős része kernel alapú vagy „kernelizálható” – például a szupport vektor gépek – és elméleti megalapozásukban kulcsszerepet játszanak a reprodukáló magú Hilbert terek.²

A statisztikus tanulási módszerek szoros kapcsolatban állnak a konvex optimalizálással, mivel gyakran konvex optimalizálási feladatokra vezethetőek vissza, sőt a meglévő algoritmusok „kernelizálása” is tipikusan a hozzájuk tartozó (primál) konvex feladatok Wolfe duálisain keresztül történik.

A megtanult modellek a véletlentől is függenek, mivel zajos megfigyelési adatokra épülnek; így egy természetes kérdés, hogy mennyire bízhatunk a kapott megoldásokban? Az előadásban áttekintjük a statisztikus tanuláselmélet néhány fontos eredményét, különös tekintettel a (tipikusan regularizált) klasszifikációs és regressziós módszerekhez adható sztochasztikus garanciákra.³ A klasszikus VC (Vapnik–Chervonenkis) dimenzió alapuló eredmények és a kereszt validációs bizonytalansági korlátok mellett néhány újramintavételezésen valamint eloszlások (reprodukáló magú) Hilbert terekbe ágyazásán⁴ alapuló sztochasztikus garanciát is bemutatunk.

Köszönetnyilvánítás: Az előadás az NKFIH 111 797 és a 125 698 számú projektjeinek és az MTA Bolyai ösztöndíjának támogatásával készült.

¹Vapnik, V. N. (1998). *Statistical Learning Theory*. Wiley-Interscience.

²Hofmann, T., Schölkopf, B., & Smola, A. J. (2008). Kernel Methods in Machine Learning. *The Annals of Statistics*, 36:1171–1220.

³Cucker, F., & Smale, S. (2002). On the Mathematical Foundations of Learning. *Bulletin of the American Mathematical Society*, 39:1–49.

⁴Muandet, K., Fukumizu, K., Sriperumbudur, B. & Schölkopf, B. (2017). Kernel Mean Embedding of Distributions: A Review and Beyond. *Found. and Trends in Mach. Learning*.