MŰEGYETEM 1782

Budapesti Műszaki és Gazdaságtudományi Egyetem

Méréstechnika és Információs Rendszerek Tanszék

Valyon József

Kiterjesztett LS-SVM és alkalmazása rendszermodellezési feladatokban

(Extended LS-SVM for System Modeling)

PhD. értekezés tézisei

Témavezető:
Dr. Horváth Gábor
Budapesti Műszaki és Gazdaságtudományi Egyetem
Méréstechnika és Információs Rendszerek Tanszék

Budapest, 2007. február

Budapesti Műszaki és Gazdaságtudományi Egyetem

Méréstechnika és Információs Rendszerek Tanszék

1117 Budapest, XI. Magyar Tudósok körútja 2.

I. épület E329. szoba

Tel: (1) 463 3587, Fax: (1) 463 4112,

Email: valyon@mit.bme.hu

A disszertációban bemutatott munkát részben az Országos Tudományos Kutatási Alapprogram (OTKA) támogatta a T 046771 számú szerződés keretében.

TARTALOMJEGYZÉK

1. Bevezetés, előzmények

A környező világ megértésének és megismerésének egyik legfontosabb eszköze a rendszermodellezés. A rendelkezésre álló adatoktól, információktól, valamint az alkalmazott módszerektől függően számos lehetőség van egy vizsgált rendszer modelljének elkészítésére.

A leggyakoribb esetben amikor a vizsgált rendszer belső felépítése nem ismert és csak a rendszer be- illetve kimenetén mért adatok állnak rendelkezésre, fekete doboz modellezéssel egy viselkedési modell készíthető. A modellezési feladat egy úgynevezett „ill-posed" probléma, ahol az adatok gyakran zajosak, hiányosak, nem elegendőek a rendszer leírására.

Tipikus példa erre a Linz-Donawitz konverteres (LD konverter) acélgyártási probléma amit a disszertációban szereplő megoldások valós, ipari tesztelésére használunk [1].

A mérési adatokból történő tanulás, illetve a fekte doboz modellezési feladatoknál a kernel módszerek, ezeken belül pedig –az elsősorban Vladimir Vapnik nevéhez köthető– Szupport Vektor Gépek (Support Vector Machine – SVM) [2] egyre szélesebb körben kerülnek alkalmazásra. Ennek köszönhetően ez a témakör napjainkban igen intenzíven kutatott terület, mely fokozatosan egy önálló kutatási területté nőtte ki magát a hagyományos intelligens rendszerek témaköreiben. Kutatásaim során valós mérésekkel gyűjtött –rendszerint zajos- adatok alapján történő rendszermodellezéssel, illetve osztályozási és függvény approximációs modellezési feladatokkal foglalkoztam.

A kernel módszerek, illetve kernel gépek működésének lényege, hogy az eredeti megfogalmazásában még komplex nemlineáris megoldást igénylő feladatot, illetve a feladatot reprezentáló mintákat, nemlineáris transzformációk segítségével egy a bemeneti mintatér dimenziójánál több dimenziós térbe transzformálják, ahol az már lineárisan megoldható. Az egyik legismertebb és legelterjedtebben használt kernel alapú megoldás a szupport vektor gép, amely számos előnyös tulajdonsága miatt hamar a kutatások középpontjába került.

A SVM modell egyik legnagyobb előnye, hogy egy garantált felső korlátot ad az approximáció általánosítási hibájára. Egy másik fontos jellemzője, hogy a tanulási algoritmus törekszik a modell méretének, komplexitásának minimalizálására (ritka, azaz „sparse" modellt alkot). Ez a hiba rovására történik, de mértéke egy paraméterrel szabályozható. A szupport vektor gép, ahogy nevéből is látszik úgynevezett szupport vektorokat[1] választ ki a tanító minták közül, melyek a kialakított modell alapját adják. A modell komplexitása, az SVM neurális értelmezése szerint a neuronok száma, megegyezik a szupport vektorok számával.

A hagyományos SVM alkalmazásának legnagyobb akadálya a módszer nagy algoritmikus komplexitása és a nagy memóriaigény, ami nehezíti a valós problémákban történő felhasználást. Ezen problémára számos megoldás született. Ezek az algoritmusok többnyire iteratív megoldások, melyek a nagy optimalizálási feladatot kisebb feladatok sorozatára bontják. Az egyes szeletelési, „chunking" algoritmusok főként a feladat dekomponálás módjában -a részfeladatok meghatározásában- különböznek [3]. Egy másik megközelítés az eredeti optimalizálási feladat átfogalmazása, úgy, hogy annak megoldása hatékonyabb legyen. Egy ilyen megközelítés a Simított SVM (SSVM - Smooth Support Vector Machine) és az ebből származó Redukált SVM (RSVM – Reduced Support Vector Machine) [4], valamint a kutatásaim kiindulásául szolgáló Least-Squares SVM [5].

Az LS-SVM bevezetése Johan Suykens nevéhez fűződik. A módszer lényege, hogy az eredeti SVM kialakításában végrehajtott módosítások révén (pl. négyzetes hibakritérium alkalmazása) a korábban kvadratikus programozást igénylő optimalizálási feladat egy egyszerű lineáris egyenletrendszer megoldására módosul. Az eljárás alkalmazásának egyik legnagyobb hátulütője, hogy nem válogat szupport vektorokat a tanítóminták közül, azaz minden tanítópont része a megoldásnak. A megoldásként adódó modell ezért nem ritka (sparse), így különösen nagy mennyiségű tanítópont felhasználása esetén, a modell túlságosan nagy. Erre a problémára Suykens egy iteratív metszési (pruning) eljárást vezet be, ami azonban jelentősen növeli a módszer számításigényét, valamint a méret csökkentésével arányosan jelentős romlást eredményez a modell pontosságában. Egy másik hasonló megoldás a rögzített méretű LS-SVM (Fixed Size LS-SVM), ahol előre megadható a felhasználni kívánt szupport vektorok száma. Az LS-SVM

[1] A szupport vektorok magyar fordítása a „tartó" vagy „támasztó" vektor lehet, hiszen ezeken a vektorokon alapul megoldás (ezek „tartják" a megoldást). A tartó vektor elnevezés azonban nem terjedt el a magyar szakirodalomban, ezért a szupport vektor elnevezést használom.

alkalmazhatóságát nagymértékben segíti a súlyozott verzió (a Weighted LS-SVM) kidolgozása, ami a zajos adatok, különösen a kiugróan hibás adatok, az „outlierek” hatásának csökkentését tűzi ki célul.

2. Kitűzött kutatási feladatok

A különböző rendszermodellezési feladatokra számos módszer, eljárás áll rendelkezésre, melyek tulajdonságaikat tekintve is eltérőek. A modellezési folyamat illetve a megoldás tekintetében a legfontosabb jellemzők az alábbiak:

- Az elkészített modell komplexitása.
- A modell pontossága.
- A modellezési eljárás algoritmikus komplexitása.
- A modellezési eljárás során eldöntendő kérdések száma, nehézsége.

A kutatás eredeti célja az acélgyártási probléma szupport vektor alapú megoldása volt. Az itt felhasználható alapmódszerek (az SVM és az LS-SVM) egyike sem teljesíti teljes körűen a fenti kritériumokat, ezért egy új, a két módszer előnyeit egyesítő megoldást kellett kidolgoznom.

A valós feladatokban általában nagy mennyiségű adattal kell dolgozni, ezért különösen fontos, hogy – ellentétben a hagyományos LS-SVM-el – a kiszámított modell komplexitása független legyen a tanítópontok számától, ugyanakkor a modell pontossága érdekében minden rendelkezésre álló adatot felhasználjunk.

A modell pontosságának biztosítására elengedhetetlen, hogy a megoldás képes legyen a zajos adatok kezelésére. Itt elsősorban a normál eloszlású zaj, az ismert mértékű (eloszlású) zaj, illetve az „outlierek” hatásának csökkentésére kell eszközöket biztosítani.

A modellezési folyamat megkönnyítése érdekében olyan megoldásokat, módszereket kell felhasználni, amelyek algoritmikus komplexitása lehetővé teszi nagy bonyolultságú, illetve nagy mennyiségű tanítóponttal leírt rendszerek modellezését is.

A modellkészítési feladat része a hiperparaméterek jó (lehetőleg optimális) megválasztása is, így a szabad paraméterek számának csökkentése, vagy a helyes megválasztás megkönnyítése jelentősen egyszerűsítheti a feladatot.

A kutatás célja tehát egy olyan szupport vektor alapú eljárás kidolgozása volt, mely lehetővé teszi egy ritka, de ugyanakkor pontos, könnyen kialakítható és egyben hatékonyan kiszámítható modell készítését.

3. A kutatás módszertana

A dolgozatban általánosan használt vizsgálati módszerek a rendszermodellezés, illetve a függvényapproximáció módszereiből, a lineáris algebrából, a matematikai statisztikából, a neurális hálózatok, valamint az optimalizáció elméletéből ismert megközelítések. Számos irodalmi eredményre támaszkodom, különösen a lineáris egyenletrendszer megoldásai, illetve a statisztikai módszerek kapcsán.

A kutatás során alapvetően az LS-SVM hátrányainak kiküszöbölésére törekedtem, így az egyes téziseket elsősorban az LS-SVM felírásából származtatom. Ez valójában annyit jelent, hogy az eredeti probléma kernel térbeli alakját, azaz az LS-SVM végső lineáris egyenletrendszerét az eredeti feladattal ekvivalens felírásként kezelem és a redukciós valamint optimalizálási lépéseket ezen reprezentációból kiindulva keresem. Az irodalomban a hasonló célokat legtöbbször az eredeti (primal) problémából, illetve a feladatot már linearizáló jellemző (feature) térbeli alakból származtatják. Ebből a megközelítésből mindig olyan eredmény születik, ami a kernel térbeli megadáshoz rögtön egy megoldási módot is rendel, ellentétben az általam alkalmazott megközelítéssel, ami innen kiindulva számos további módosítást és megoldási módot tesz lehetővé. A kernel térbeli felírás használatának egy másik fő előnye, hogy itt –ellentétben a valójában nem ismert, nem kiszámított, jellemző térbeli feladattal- egy egzakt, véges és a gyakorlatban is kiszámított, valamint már lineáris reprezentáció felhasználásával dolgozom.

A fenti megközelítésből adódóan, az eredmények először az LS-SVM egyenletrendszerének módosítását, azaz kernel térben történő műveleteket jelentenek, melyekhez később társulhatnak a jellemző térbeli, illetve az eredeti problématérbeli értelmezések.

Az eredmények, módszerek értékelése során a legtöbb esetben több jellemző együttes figyelembevételére van szükség. A legfontosabb jellemzők:

- A kialakított modell pontossága.
- A kialakított modell bonyolultsága (mérete).
- Az eljárás algoritmikus komplexitása.
- Az eljárás implementációjának nehézsége.

A disszertációban javasolt módszerek ezen jellemzőinek értékelésére is kitérek.

A javasolt algoritmusok matematikai háttere, motivációja a legtöbb esetben megadható, ugyanakkor a probléma (ill-posed) jellege miatt az eredmények értékelésére nincsenek egzakt matematikai eszközök, erre általában „benchmark” problémákat és kereszt kiértékelésen („cross validation”) alapuló szimulációt alkalmaznak. A disszertációban bemutatott eljárások értékeléséhez az irodalomban a módszerek működésének bemutatására és kiértékelésére legelterjedtebben használt problémákat használtam fel. Az eredetileg kitűzött valós ipari rendszermodellezési probléma, az LD konverteres acélgyártás modellezése egy regressziós feladat, ezért a kutatások során elsősorban a regresszió megvalósítására koncentráltam. A tématerületen alkalmazott benchmarkok azonban főként osztályozási feladatok, a regressziós feladatok inkább csak szemléltetésre alkalmas problémák. Mivel a javasolt megoldások mind regressziós mind pedig osztályozós esetre működnek, az alapmódszert (LS^2-SVM) az osztályozós esetre is implementáltam és teszteltem. A regressziós megoldásokat az egyszerű mintapéldákon túl egy bonyolultabb idősor előrejelzési feladatban is teszteltem, majd az acélgyártási problémára alkalmaztam.

A kutatás során a célkitűzések között nem szereplő, de a modellépítésben szerepet játszó részfeladatok (pl. hyperparaméterek meghatározása, kernelek kiválasztása) tekintetében a leggyakrabban használt értékeket illetve módszereket használtam.

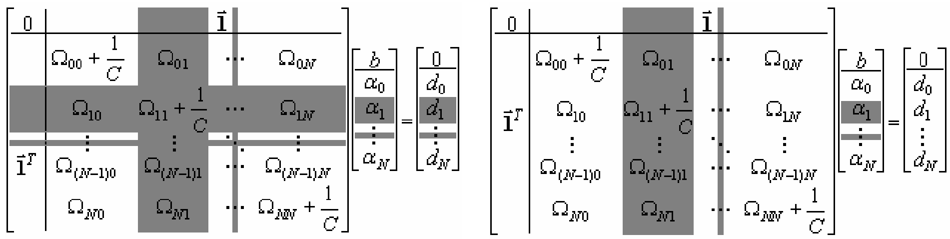
Az ismertetett eredmények ellenőrzését, a javasolt módszerek tesztelését egy saját készítésű MATLAB „toolbox” segítségével végeztem. A „toolbox” kialakítása jól illeszkedik a kiterjesztett LS-SVM keretrendszer lehetőségeihez. Az első tézisnek megfelelő redukált egyenletrendszer implementációja mellett, „plug-in”-szerűen támogatja többek közt a második téziscsoport szerinti redukciós eljárások, valamint a harmadik téziscsoportnak megfelelő különböző megoldások beillesztését.

4. Új tudományos eredmények

Az első téziscsoport a kernel tér redukciójának, azaz egy kevesebb dimenziós kernel tér alkalmazásának lehetőségét mutatja be. Ez nemcsak egy ritka (sparse), azaz kisebb komplexitású, de ugyanakkor egy pontos, minden bemeneti mintát figyelembevevő megoldáshoz vezet. A második téziscsoportban a megfelelő („optimális”) redukció meghatározására adok módszereket. Az alkalmazott megoldás –a redukált kernel térben- túlhatározottá teszi a problémát, ami további elemzésre, optimalizálásra ad lehetőséget. A harmadik téziscsoportban megmutatom, hogy a redukált kernel térben túlhatározott feladat statisztikai elemzésével zaj hatása, típusától függetlenül, jelentősen csökkenthető.

1. **Tézis: Kidolgoztam egy új részlegesen redukált LS-SVM megvalósítást, ami lecsökkenti a szupport vektorok számát, ugyanakkor az összes mintapontot felhasználja a modell kiszámításában. Ezt kihasználva a modell komplexitása úgy csökkenthető, hogy a ritka LS-SVM pontossága nem, vagy nem jelentősen romlik a teljes LS-SVM eredményeihez képest.**

   A ritka LS-SVM kialakítására irányuló kutatásaim során egy olyan részleges redukciós eljárást dolgoztam ki, amely lehetővé teszi, hogy az összes tanítóminta felhasználása mellett egy redukált, kevesebb tanítómintán –szupport vektoron- alapuló modellt hozzunk létre. Az LS^2-SVM fő célja, hogy csökkentse a modell komplexitását, azaz a rejtett rétegbeli nemlineáris neuronok számát. A javasolt módszer egy járulékos előnye, hogy a modell számításának algoritmikus komplexitása is csökken. A ritka (sparse) LS-SVM kialakításához használt metszési (pruning) eljárás bizonyos –kevésbé fontosnak tartott- mintapontokat elhagy, és egy új kvadratikus egyenletrendszert old meg. A részleges redukció ezzel szemben minden tanítópontot megőriz, de nem minden pontot használ fel kernel középpontnak. Az LS-SVM számítását leíró egyenletrendszerben (lásd 1. ábra) a kernel középpontoknak az oszlopok, míg a tanítópontok által reprezentált feltételeknek a sorok felelnek meg, így a részleges redukció egy túlhatározott egyenletrendszerre vezet.



a.) b.)

1. ábra. Az LS-SVM egyenletrendszer hagyományos "teljes" (a.) és javasolt "részleges" (b.) redukciója. Az egyenletekből a szürke elemek kerülnek törlésre.

Az oszlopok törlése, de a sorok megtartása biztosítja, hogy a neuronok (kernelek) száma csökkenjen, míg az összes ismert megkötést (tanítópontot) figyelembe veszünk. Ez a kulcsa, hogy a megoldás komplexitása a pontosság megtartása mellett csökkenthető legyen. A módosított, részlegesen redukált egyenletrendszert alapesetben (normál eloszlású zaj esetén) „Least Squares" értelemben oldjuk meg, ezért nevezzük ezt a módszert Least Squares LS-SVM-nek vagy röviden LS^2–SVM-nek.

A redukált LS-SVM négyzetes hibaminimumra optimalizáló megoldása (LS^2–SVM) megfeleltethető a hagyományos LS-SVM megoldásnak, valamint látható, hogy redukció nélkül a két megoldás azonos.

1.1. Kidolgoztam egy új részleges redukciós eljárást, mellyel úgy alakítható ki ritka (sparse) megoldás, hogy közben az összes tanítópontból rendelkezésre álló információ felhasználható a megoldás kialakításához.

1.2. Igazoltam, hogy a részleges redukció eredményeképpen –mivel az összes mintapont felhasználásra kerül- egy ritka megoldás a hagyományos –mintapontokat elhagyó– megoldásnál kisebb hibával érhető el, másrészről pedig egy adott pontosságú megoldás lényegesen kisebb modellel is elérhető.

1.3. Megmutattam és szimulációkkal igazoltam, hogy a részleges redukció felhasználásával a regularizáció szerepe megváltozik, mert a túlhatározott egyenletrendszernek köszönhetően már nem ez a felelőse a túlilleszkedés (overfitting) elkerülésének.

1.4. Az eredeti LS-SVM felírásból részleges redukcióval kialakított modellt megfeleltettem egy (a jellemző tér helyett) a kernel térben regularizált modellnek.

2. Tézis: Új több szempontból is előnyös eljárásokat adtam a részlegesen redukált modell szupport vektorainak meghatározására, valamint megmutattam, hogy erre a már ismert válogatási eljárások is alkalmazhatók.

A részleges redukción alapuló ritka megoldáshoz először meg kell határozni a redukált modell alapját jelentő szupport vektorokat. A szupport vektorok kiválasztására, illetve meghatározására számos módszer alkalmazható, melyek nem csak az eredmény –a kiválasztott vektorokon alapuló modell- pontosságában, hanem például algoritmikus komplexitásban is eltérnek. Fontos továbbá, hogy a módszer választ adjon arra a kérdésre is, hogy hány szupport vektorra van szükség. Munkámban több –irodalmi, illetve saját javaslaton alapuló- kiválasztási eljárást tárgyalok és hasonlítok össze a fenti szempontok szerint, valamint javaslatot teszek egy hatékony, a kernel matrix bázisának meghatározásán alapuló eljárásra. A szupport vektorok meghatározására kidolgozott eljárásom csökkenti az LS-SVM és az SVM közötti különbségeket, mert az általa meghatározott szupport vektor szám – hasonlóan az eredeti SVM-hez – egy paraméterrel állítható, mely a modell komplexitása és pontossága közötti súlyozást adja meg. Megmutattam, hogy a hagyományos LS-SVM-hez használt metszési (pruning) eljárás, valamint a redukált bázisú ridge regresszióban alkalmazott jellemzővektor kiválasztási (FVS - Feature Vector Selection) eljárás a redukált modell kialakításához is felhasználható.

2.1. Új algoritmust adtam a szupport vektorok hatékony kiválasztására. A kidolgozott algoritmusban a hagyományos SVM-hez hasonlóan egy paraméterrel kontrolálható a modell méret csökkentése (a szupport vektorok száma) és pontosság fontossági viszonya. Az új módszer algoritmikusan is hatékony az ismert eljárásokhoz képest. Szimulációval igazoltam, hogy ez az eljárás a kialakított modell pontossága tekintetében is jó eredményekre vezet.

2.2. Megmutattam, hogy az irodalomból ismert szupport vektor kiválasztási eljárások (pl. pruning, FVS) alkalmazhatók a kiterjesztett LS-SVM elkészítésében is.

2.3. Megmutattam, hogy a redukált modell – előnyei, például a pontosság, megtartása mellett – alkalmas a rögzített méretű LS-SVM-nek (Fixed Size LS-SVM) megfelelő előre megadott méretű modell megvalósítására is.

2.4. Megmutattam, hogy a hagyományos metszési (pruning) eljárás kiválasztási kritériumának az ellentétét felhasználva egy új módszer nyerhető (inverse pruning). Szimulációkkal alátámasztottam, hogy bizonyos hiperparaméter beállítások esetén ez a megközelítés jobb eredményre vezet.

3. Tézis: Megmutattam, hogy a részlegesen redukált LS-SVM túlhatározottsága révén – a kernel térben – további optimalizálásra (főként a zaj hatásának csökkentésére) nyújt lehetőséget.

Az eredeti LS-SVM esetében a nemlineáris transzformációval létrehozott térben (kernel tér) csak egyetlen lineáris megoldást kereshetünk, mert a tér dimenzióinak száma megegyezik az ide leképezett pontok számával. A részleges redukció esetén azonban a kernel térben sokkal több pont áll rendelkezésünkre, mint ahány dimenziós megoldást keresünk. Ennek eredményeképp egy olyan túlhatározott rendszer áll elő, amelynek már nem csak egy lehetséges megoldása van, hanem különböző szempontokból optimális megoldások kereshetők.

A részleges redukcióval létrehozott kernel térben lehetőség van az eredeti LS-SVM-nek megfeleltethető négyzetes hiba alkalmazására, de a tanítóminták statisztikai jellemzőinek figyelembevételével egyéb, más szempontokból optimális megoldások is kereshetők. Ezek a zajos adatok kezelésében kivételes lehetőségeket nyújtanak. Megoldást mutatok a zaj hatásának csökkentésére, valamint „outlierek” detektálására, hatásának csökkentésére.

A zajos adatok kezelése kapcsán valójában a kernel térben keresett **lineáris** megoldás optimalizálására vonatkozó lehetőségeket használtuk ki. Megmutatom, hogy a részleges redukció lehetővé teszi a kernel térbeli megoldás még általánosabb felfogását is.

3.1. Megmutattam, hogy az egyes tanítómintákhoz tartozó hiba (zaj) mértékek ismeretében a súlyozott négyzetes hiba minimalizáló eljárás alkalmazásával – négyzetes hibakritérium értelmében – optimális megoldás adható.

3.2. Megmutattam, hogy míg a hagyományos LS-SVM metszési (pruning) és súlyozási (weighting) eljárások –bár a céljuk nem ellentétes- nem alkalmazhatók egyszerre, ugyanakkor a kiterjesztett LS-SVM egyszerre tud ritka és súlyozott megoldást adni.

3.3. Megmutattam, hogy a kernel térbeli lineáris megoldásban a robusztus statisztika (approximáció) eszközeinek alkalmazásával a zaj hatása, különös tekintettel a kiugró hibák (outlierek) torzító hatására, csökkenthető, illetve kiszűrhető.

3.4. Megmutattam, hogy a feladat a kernel térben tovább általánosítható. A lineáris megoldások mellett létrehozhatók lokálisan lineáris, sőt nemlineáris modellek is. A lokálisan lineáris modellek az inkrementális tanulást teszik lehetővé, míg a nemlineáris modellek SVM vagy LS-SVM megvalósításai több rétegű szupport vektor gépeket eredményeznek.

5. Az új eredmények gyakorlati hasznosítása

Napjainkban számos valós ipari rendszermodellezési probléma vár megoldásra. Ezen feladatok egy jelentős részében, csakúgy, mint a bemutatott acélgyártási probléma esetén, nagy mennyiségű és többnyire zajos adat alapján kell hatékony fekete doboz modellt alkotni. A disszertációban tárgyalt módszerek lehetővé teszik, hogy az utóbbi évtizedben nagy sikerrel alkalmazott kernel gépek ilyen problémák megoldására is hatékonyan felhasználhatók legyenek. A modellezés célja egy hatékonyan számítható, de legfőbbképpen ritka és robusztus kernel alapú megoldás készítése.

Az LS-SVM hátrányainak kiküszöbölésére ad megoldást az első és a második téziscsoport. Az első téziscsoportban szereplő részleges redukciós módszer a pontosság megőrzése mellett biztosítja a ritka megoldást. A második téziscsoportban bemutatott kiválasztási algoritmus megfelelően támogatja a szupport vektorok meghatározását, valamint egy paraméteren keresztül eszközt biztosít a ritkaság és a pontosság közötti fontosság szerinti súlyozására. A harmadik téziscsoport a részleges redukciós megoldásból adódó túlhatározott egyenletrendszer megfelelő megoldásaival a zajos adatok kezelésére ad hatásos eszközöket. A téziscsoport a valóságban leggyakrabban előforduló esetekre (normál eloszlású zaj, ismert zaj paraméterek, „outlierek") egyedi megoldásokat ad, valamint bemutatja, hogy a statisztikai módszerek hogyan alkalmazhatók a modellkészítés során.

A javasolt megoldások hasznosságát és alkalmazhatóságát a tématerületen elterjedt „benchmark" problémákon, valamint egy valós, ipari feladaton is bemutattam. Az acélgyártás rendkívül összetett fizikai-kémiai folyamat, ami a gyártási technológiából adódóan egy dinamikus, nagy zajjal terhelt rendszerben megy végbe. A magas hőfokon, zárt rendszerben lezajló összetett folyamatokról jelenleg csak pontatlan, az ipari alkalmazás igényeit messze nem kielégítő közelítő modellek állnak rendelkezésre. Ezt az ipari rendszert –mivel a működő LD konverterben mérések nem hajthatók végre - legpontosabban a bemenetek, valamint a kimenet írja le. A javasolt kiterjesztett LS-SVM jól alkalmazható volt erre a problémára:

- A részleges redukciónak köszönhetően a nagy mennyiségű, ugyanakkor a folyamat egészét még így sem elégségesen leíró adat felhasználása mellett, egy jelentősen redukált ritka modellt alakítottam ki, melynek pontossága megfelel a korábban neurális hálózatokkal elért eredményeknek.
- Megmutattam, hogy a robusztus megoldásokkal az acélgyártási problémában fennálló zaj hatása csökkenthető, az esetleges „outlierek" hatása jól kiküszöbölhető.

Az eredmények alapján a részleges redukciós eljárással nagyfokú ritkaság érhető el, a modell pontosságának jelentős romlása nélkül. Kijelenthető, hogy a részleges redukció segítségével olyan mértékben ritka, egyszerű modellek is létrehozhatók, amelyek a hagyományos LS-SVM metszés (pruning) alkalmazásakor nagyon –elfogadhatatlanul- nagy hibával járnának együtt. A vizsgálatok alapján második tézisben javasolt szupport vektor kiválasztási eljárás (RREF alapú módszer) általában egy jó megoldást megalapozó kiválasztást eredményez, számos más előnye, mint például a szabályozható szupport vektor szám, algoritmikus hatékonyság stb. mellett. A zaj hatásának csökkentésére, szűrésére javasolt módszer ugyancsak igen hatásosnak bizonyult, különösen a nagy kiugró hibájú pontok, az úgynevezett „outlierek" kiszűrésében.

Az LS-SVM módszer kiterjesztéseit a szakterület közössége mind a megjelent folyóirat, mind pedig konferencia cikkekből megismerhette. Fontos megemlíteni, hogy az LS-SVM Least-Squares kiterjesztésének, azaz az LS^2-SVM-nek az alapötlete a BME intelligens rendszerek szakirányain az utóbbi években bemutatásra került, valamint szerepel a 2007-ben megjelenő „Neurális hálózatok" című könyvben, mely egyben a szakirány tankönyve is.

Az irodalomban elterjedt megoldások leggyakrabban az eredeti, illetve a már lineáris jellemző (feature) térben keresik a kívánt (redukált, robusztus stb.) megoldást. A tézisekben bemutatott a kernel térbe leképezett feladatot redukáló, illetve a kernel térben optimalizáló eljárások számos további kutatási lehetőséget rejtenek.

6. A disszertációhoz tartozó publikációk

6.1. Könyvek

Altrichter Márta, Horváth Gábor, Pataki Béla, Strausz György, Takács Gábor, **Valyon József**, „Neurális hálózatok", Panem, 2007.

6.2. Folyóiratcikkek

J. Valyon and G. Horváth, „A Robust LS–SVM Regression", *International Journal of Computational Intelligence, Vol. 3 No. 3.,* pp. 148-153, 2006.

J. Valyon and G. Horváth, „Extended Least Squares LS–SVM", *International Journal of Computational Intelligence, Vol. 3 No. 3.,* pp. 234-242, 2006.

J. Valyon and G. Horváth, „A Weighted Generalised LS–SVM", *Periodica Polytechnica Electrical Engineering*, 47/3-4, pp. 229-252., 2003.

J. Valyon and G. Horváth, „A Sparse Least Squares Support Vector Machine Classifier", *Periodica Polytechnica Electrical Engineering*, pp. 17-23., 2004.

Valyon József és Horváth Gábor, „Least squares szupport vektor gépek adatbányászati alkalmazása", *Híradástechnika*, pp 33-39, 2005.

6.3. Nemzetközi konferenciák

P. Berényi, **J. Valyon** and G. Horváth: „Neural Modeling of an Industrial Process with Noisy Data", *IEA/AIE*-2001 *The Fourteenth International Conference on Industrial & Engineering Applications of Artificial Intelligence & Expert Systems*, June 4-7, 2001, Budapest in Monostori, L, Váncza, J., Ali Moonis (eds.) *Lecture Notes in Artificial Intelligence*, Springer, pp. 269-280, 2001.

J. Valyon and G. Horváth, „Reducing the complexity and network size of LS-SVM solutions", POSTER: *NATO-ASI Conference*, Leuven, 2002.

J. Valyon and G. Horváth, „A generalized LS–SVM", *SYSID'2003,* Rotterdam, pp. 827-832, 2003.

J. Valyon and G. Horváth, „A Sparse Least Squares Support Vector Machine Classifier", *Proceedings of the International Joint Conference on Neural Networks IJCNN 2004*, Budapest, pp. 543-548, 2004.

G. Horváth**, J. Valyon**, Gy. Strausz, B. Pataki, L. Sragner, L. Lasztovicza, N. Székely, „Intelligent Advisory System for Screening Mammography", *Proceedings of IMTC/04, 21th IEEE Instrumentation and Measurement Technology Conference*, Como, Italy, pp. 2071-2076, 2004.

J. Valyon and G. Horváth, „A Robust LS-SVM regression", *Enformatica conference,* Prague, pp. 148-153, 2005.

J. Valyon and G. Horváth, „A Sparse Robust Model for a Linz-Donawitz Steel Converter", *IEEE Instrumentation and Measurement Technology Conference*, Poland, May 1-3, 2007.

6.4. Hazai konferenciák

J. Valyon and G. Horváth, "Employing Support Vector Machines for Nonlinear System Modelling", *Proceedings of the Mini-Symposium 2001*, pp. 34-35, 2001.

J. Valyon and G. Horváth, "A comparison of the SVM and LS–SVM regression, from the viewpoint of parameter selection", *Proceedings of the 2002 Mini-Symposium*, pp. 18-19, 2002.

J. Valyon and G. Horváth, „A Weighted Generalised Least-Squares Support Vector Machine", *Proceedings of the 10th Phd Mini-Symposium*, pp. 30-31., 2003.

J. Valyon and G. Horváth, „A Sparse Least Squares Support Vector Machine Classifier", *Proceedings of the John von Neuman Phd Conference*, pp. 23-26, 2003.

J. Valyon and G. Horváth, „Controlling the complexity and network size of LS–SVM solutions", *$(CS)^2$ - The Third Conference of PhD Students in Computer Science*, Szeged, Hungary, pp. 105, 2002 (best paper award).

7. Irodalomjegyzék

[1] G. Horváth, B. Pataki, and Gy. Strausz: "Black box modeling of a complex industrial process", Proc. of the 1999 IEEE Conference and Workshop on Engineering of Computer Based Systems, Nashville, TN, USA, pp. 60-66, 1999.

[2] V. Vapnik, "The Nature of Statistical Learning Theory", New–York: Springer–Verlag., 1995.

[3] Yuh – Jye Lee and O. L. Mangasarian, "RSVM: Reduced Support Vector Machines", Proceedings of the First SIAM International Conference on Data Mining, Chicago, 2001.

[4] J. C. Platt. "Sequential Minimal Optimization: Fast Algorithm for Training Support Vector Machines", Microsoft Research Technical Report MSR–TR–98–14, 1998.

[5] J. A. K. Suykens, V. T. Gestel, J. De Brabanter, B. De Moor, J. Vandewalle, "Least Squares Squares Support Vector Machines", World Scientific, 2002.