

MODERNÉ METÓDY ANALÝZY DÁT A MODELOVANIA

URBANÍKOVÁ Marta, SR

Resumé

S rozvojom informačných technológií sú úzko späté moderné metódy analýzy dát a modelovania systémov založené na teórii neurónových sietí, genetických algoritmoch a fuzzy logike. Cieľom príspevku je popísať uvedené teórie a možnosti ich využitia v oblastiach podnikových procesov, financií a makroekonomických analýz.

Kľúčová slova: informačné a komunikačné technológie, neurónové siete, genetické algoritmy, fuzzy logika

MODERN METHODS OF DATA ANALYSIS AND MODELING

Abstract

With the developing information technologies are closely linked modern methods of data analysis and systems modeling based on the theory of neural networks, genetic algorithms and fuzzy logic. The aim of this paper is to characterize this theories and their applicability in the areas of business processes, finance and macroeconomic analysis.

Key words: information and communication technology, neural network, genetic algorithm, fuzzy logic

Úvod

Informačné a komunikačné technológie zaujímajú rozhodujúce miesto v živote spoločnosti. Podnikové, finančné, organizačné a ekonomické systémy predstavujú zložité dynamické systémy. Modely sú zjednodušené matematické zápisy týchto systémov. Využívajú sa predovšetkým na skúmanie systémov, ich optimalizáciu a predikciu správania sa systémov. Modely môžeme zostaviť z hľadiska potrieb na makroekonomickej aj mikroekonomickej úrovni.

Neurónové siete patria k analytickým nástrojom, ktoré možno zahrnúť pod pojem umelá inteligencia. Systémy umelej inteligencie sú úzko späté s rozvojom informačných technológií. Inšpirácia k vzniku neurónových sietí prišla z biologických systémov. Ide vlastne o snahu simulovať funkciu mozgu. Jednou z najvýznamnejších vlastností neurónových sietí je to, že sú vlastne univerzálnym aproximátorom funkcií. Často sa stáva, že máme veľmi zložitý systém, ktorého popis je prakticky nemožný, alebo by vyžadoval absurdne veľa strojového času. Máme však dáta, ktoré do systému vstupujú, a k nim zodpovedajúce výstupy. V tejto situácii môžeme použiť vhodnú neurónovú sieť a pokúsiť sa ju naučiť chovať sa ako sledovaný systém. Toto je dôležitý moment, ktorý determinuje aplikačné uplatnenie neurónových sietí v praxi.

Neurónové siete sú často označované aj termínom black box, pretože nemôžeme poznať detailne vnútornú štruktúru systému. Na vnútornú štruktúru systému, ktorý je modelovaný čiernou skrinkou, kladieme iba niekoľko predpokladov, ktoré umožnia popísať chovanie systému funkciou, ktorá robí transformáciu vstup – výstup. Neurónové siete je vhodné použiť v prípade, keď významnú úlohu v modelovanom procese hrá náhoda a kde deterministické závislosti sú natoľko zložité a poprepájané, že ich nedokážeme separovať a analyticky identifikovať. Sú teda vhodné pre modelovanie zložitých, často nezvratných, strategických rozhodnutí.

Systémy umelej inteligencie, genetické algoritmy, teóriu fuzzy logiky možno použiť v mnohých oblastiach analýz dát a modelovania. V mikroekonómii možno využiť neurónové siete na modelovanie činnosti firiem, modelovanie produkčných funkcií, pri marketingových štúdiách, pri riadení procesu výroby, atď.

1. Moderné metódy analýzy dát a modelovania

Medzi moderné metódy analýzy dát a modelovania systémov môžeme predovšetkým zahrnúť neurónové siete, genetické algoritmy a teóriu fuzzy logiky. Neurónová sieť je masívne paralelný výpočtový systém, ktorý má schopnosť uchovávať informácií a umožňuje ich ďalšie spracovanie, pričom napodobňuje činnosť ľudského mozgu v zbieraní poznatkov v procese učenia a uchovávaní týchto poznatkov s využitím medzineurónových spojení - synaptických váh.

Ak sa na uchovávanie údajov v sieti používajú aj iné parametre ako medzineurónové spojenia, napr. prahové koeficienty alebo parametre prechodových funkcií neurónov, hovoríme o adaptívnych sieťach. Základným prvkom neurónovej siete je neurón. Vo všeobecnosti má niekoľko vstupov od iných neurónov alebo z okolitého prostredia a jeden výstup. Operácia, ktorou neurón transformuje svoje vstupy na výstup je spravidla veľmi jednoduchá. Zložitá činnosť neurónovej siete spočíva v spojení mnohých takýchto jednoduchých elementov do celku.

Podľa smeru, ktorým sa šíria signály v neurónovej sieti ich rozdelíme na dopredné a rekurentné. Dopredné neurónové siete sú také, v ktorých sa signál šíri iba od vstupných neurónov (neuróny, ktorých vstupmi sú signály z prostredia) cez skryté neuróny k výstupným neurónom (neuróny, ktorých výstup vedie do prostredia). Rekurentné neurónové siete sú také, v ktorých sa signál môže pohybovať aj smerom od výstupov ku skrytým častiam alebo dokonca až k vstupom. Je zrejmé, že rekurentné neurónové siete sú zložitejšie ako dopredné.

Jednou zo základných vlastností neurónových sietí je schopnosť učiť sa. Podľa toho, na akom princípe je táto schopnosť v neurónových sieťach realizovaná, rozoznávame dva typy učenia sa: kontrolované a nekontrolované. Kontrolované učenie je založené na tom, že sieť má pri svojom učení k dispozícii množinu vstupov a k nim požadovaných výstupov. V procese učenia sa upravujú parametre siete, t.j. váhy spojení tak, aby sa minimalizoval rozdiel medzi obdržanou odozvou siete na dané vstupy a výstupmi požadovanými pre tieto vstupy. Nekontrolované učenie je také, pri ktorom má sieť k dispozícii iba vstupy. K nim prislúchajúce výstupy sieť generuje na základe vlastností samotných vstupných údajov. Tieto výstupy nie sú vopred známe.

Dopredné neurónové siete s kontrolovaným učením sú veľmi často používané pre ich ľahkú implementovateľnosť a všestrannosť ich použitia. Používajú sa pri klasifikácii údajov, aproximácii funkcií, predikcii alebo identifikácii a modelovaní neznámych systémov. Najbežnejšie používaným typom doprednej neurónovej siete s kontrolovaným učením je viacvrstvový perceptrón. Neuróny sú usporiadané v jednej vstupnej, jednej výstupnej a jednej alebo viacerých skrytých vrstvách.

Z rekurentných neurónových sietí s kontrolovaným učením je potrebné spomenúť Hopfieldove neurónové siete. Spôsob spojenia neurónov, tzv. topológia siete v tomto type neurónovej siete je veľmi zvláštna: každý z neurónov je vstupný a zároveň výstupný a je spojený so všetkými ostatnými neurónmi. Siete tohto typu sa používajú na realizáciu pamäte. Za zmienku možno stojí aj fakt, že Hopfieldovou neurónovou sieťou je možné riešiť aj tzv. problém obchodného cestujúceho.

Medzi dopredné neurónové siete s nekontrolovaným učením patria najmä Kohonenove samoorganizujúce sa mapy. Topologicky sú veľmi jednoduché: obsahujú iba jednu vstupnú a

jednu výstupnú vrstvu neurónov, skryté neuróny neobsahujú. Samoorganizujúce sa mapy je však možné hierarchicky usporiadať, a vtedy sa stávajú veľmi silným nástrojom divízneho zhľukovania údajov. Ich využiteľnosť môže ďalej vzrásť, ak sú spojené s fuzzy množinami

Posledným typom neurónových sietí sú rekurentné siete s nekontrolovaným učením. Topológia spojenia neurónov a spôsob učenia týchto sietí je spravidla založený na Adaptive Resonance Theory. Tento spôsob učenia umožňuje naučiť sieť správne reagovať na nové vzory bez toho, aby „zabudla“ vedomosti naučené v minulosti.

Genetické algoritmy umožňujú riešiť predovšetkým optimalizačné úlohy ekonomickej praxe. Proces reprodukcie zahŕňa selkciu, kríženie a mutácie. Pri výpočtoch je nutné zostaviť tabuľku, určiť účelovú funkciu, zadať hodnoty ktoré budú optimalizované, obmedzujúce podmienky, veľkosť populácie, dĺžku chromozómu, parameter kríženia a prechodu do ďalšej generácie.

2. Možnosti využitia genetických algoritmov, neurónových sietí a fuzzy logiky

Genetické algoritmy, neurónové siete a fuzzy logika zohrávajú dôležitú úlohu pri predikcii časových radov. Na rozdiel od štatistických metód nepotrebujú žiadne predpoklady o modeloch, stačia im iba vstupné dáta. Takisto si dokážu poradiť aj s nízkou kvalitou dát. Ďalšia z výhod je schopnosť zovšeobecňovania, keďže sú schopné aproximovať ľubovoľnú funkciu do ľubovoľného stupňa presnosti. Pre zlepšenie výsledkov sa využívajú kombinácie týchto metód. Napríklad štruktúra neurónovej siete je menená pomocou genetických algoritmov. Uplatnenie predikcie časových radov možno nájsť vo finančníctve, napr. predpoveď menových kurzov, cien akcií, komodít, indexov finančných nástrojov, v makroekonómii a mikroekonómii pri predikcii hospodárskych ukazovateľov.

Prostriedky fuzzy logiky, genetické algoritmy a neurónové siete môžeme využiť pri optimalizácii výrobného procesu. Výrobné podniky sú pod vplyvom silnej domácej i zahraničnej konkurencie, ktorá neustále narastá s globalizáciou trhového prostredia. Uvedená situácia núti výrobcov prispôbiť sa novým okolnostiam a pružne reagovať na požiadavky svojho okolia. Pružnosť vzhľadom na trh je podmienená pružnosťou výroby a skracovaním výrobných cyklov výrobkov. Splnenie požiadavky na skrátenie celkového výrobného cyklu výrobku, od jeho návrhu až po jeho dodávku k zákazníkovi s čo najmenšími nákladmi vo výrobe, je najdôležitejší predpoklad úspechu podniku. V tejto zložitej situácii zohráva dôležitú úlohu použitý systém pre plánovanie a riadenie výroby. Veľké množstvo informácií, ktoré je žiaduce spracovať pri úlohách plánovania a riadenia výroby vyžaduje nevyhnutne podporu výpočtovej techniky v tejto oblasti. PPS (ProductionPlanningSystem) majú za úlohu plánovanie a riadenie výroby tak, aby bola optimálna z kapacitného, ekonomického a časového hľadiska. Ide o úlohy ekonomických kalkulácií, vytvárania optimálnych ročných, mesačných, denných výrobných plánov s ohľadom na výrobné kapacity výrobných zariadení a pod. Pri optimalizácii priebehu výrobného procesu treba podčiarknuť význam genetických algoritmov. Problém optimalizácie úzko súvisí s minimalizáciou nákladov, ktoré majú úzku súvislosť so ziskom. Čím kratšia je doba výroby výrobkov, čím nižšie sú nároky na počet strojov a počty pracovníkov, tým nižšie sú náklady.

Pri tvorbe neurónových sietí, využití genetických algoritmov a fuzzy logiky máme viaceré možnosti. Buď priamo naprogramovať, väčšinou v jazyku C++ alebo JAVA, alebo využiť komplexný software, ktorý umožňuje vytvoriť neurónovú sieť priamo vo svojom prostredí, napr.: MathWorks MATLAB – Neural Network Toolbox, Genetic Algorithm, Fuzzy Logic Toolbox, Wolfram MATHEMATICA, STATISTICA, Neuro Solutions

Záver

Neurónové siete, genetické algoritmy a teória fuzzy logiky vo významnej miere prispievajú k rozvoju analýzy dát a modelovania. Predstavujú významné nástroje pre podporu rozhodovania. Samotná tvorba neurónovej siete, počnúc prípravou dát, voľbou počtu vrstiev, voľbou počtu neurónov v jednotlivých vrstvách, voľbou tréningových metód, aktivačných funkcií, atď. je stále vecou skúmania a experimentovania. Možno práve z tohto dôvodu si neurónové siete získali mnoho pozornosti aj medzi odbornou verejnosťou.

Pomocou neurónových sietí je možné popísať aj zložité vysokodimenzionálne dáta, ktoré nie je možné popísať jednoduchými pravidlami alebo zákonitosti platiace na množine týchto dát nie je jednoduché objaviť napr. štatistickými metódami.

Modelovanie systémov pomocou neurónových sietí je vysoko abstraktné: tá istá sieť môže byť bez zmeny štruktúry modelom pre rôzne systémy, ak sa na učenie tejto siete použijú rôzne údaje. Táto vysoká miera abstrakcie je na druhej strane aj nevýhodou neurónových sietí, pretože znalosti obsiahnuté v naučenej neurónovej sieti sú rozptýlené v jej parametroch, váhach medzineurónových spojení, a nie je možné ich zo siete získať v explicitnom tvare. Tento nedostatok neurónových sietí sa však darí eliminovať pomocou kombinácie neurónových sietí a iných systémov, napr. fuzzy regulátorov.

Literatura

1. DOSTÁL, P. *Pokročilé metody analýz a modelování v podnikatelství a veřejné správě*. Brno: Akademické nakladatelství CERM 2008. ISBN 978-80-7204-605-8.
2. FANTA, J. *Neurónové sítě ve společenských vědách*. Praha: Univerzita Karlova v Praze, Nakladatelství Karolinum. 2000. ISBN 80-246-0175-3.
3. KAMRUZZAMAN, J., BEGG, R., SARKER, R. *Artificial Neural Networks in Finance and Manufacturing*. Idea Group Publishing .2006 ISBN 1-59140-672-2.
4. MILLER, R.M. *Computer Aided Financial Analysis*. Addison-Wesley 1990, ISBN 0-201-12337-1.
5. SKŘIVÁNKOVÁ, V., SKŘIVÁNEK, J.. *Kvantitativne metody finančných operácií*. Bratislava, IURA Edition. 2006. ISBN 80-8078-074-9
6. ŠPIRKOVÁ, J. *Actuarial Functions in Second Pillar Pension Applications of Mathematics and Statistics in Economy Proceedings August 26–29, 2010, Demänovská Dolina, Slovakia*, 360-372, ISBN 978-80-89438-02-0

Článok bol napísaný s podporou projektu KEGA 035UKF-4/2011 Meranie úrovne regionálneho rozvoja s využitím matematických metód (Študijná e-pomôcka pre študijný program Geografia v regionálnom rozvoji)

Lektoroval: **doc. Ing. Viera Papcunová, Ph.D.**

Kontaktní adresa:

Marta Urbaníková, Doc. RNDr.CSc.

Ústav manažmentu a informačných technológií,

Fakulta prírodných vied UKF Nitra, Trieda Andreja Hlinku 1, 94901 Nitra, SR, tel. 00421 905364794,

e-mail: murbanikova@ukf.sk