

NEURONOVÉ SÍTĚ PRO CHEMII II. APLIKACE

VLADIMÍR KVASNIČKA, ŠTĚPÁN SKLENÁK a JIŘÍ
POSPÍCHAL

Katedra matematiky, CHTF STU, 812 37 Bratislava

*Věnováno prof. RNDr. Ing. Alexandrovi Tkáčovi, DrSc. při
příležitosti jeho 70. narozenin*

Došlo dne 24.II.1992

Obsah

1. Úvod
2. Aplikace neuronových sítí v chemii
 - 2.1 Spektroskopické aplikace
 - 2.2 Řízení procesů
 - 2.3 Chemická reaktivita
 - 2.4 Bioaplikace
 - 2.5 Ostatní aplikace
3. Závěr

1. Úvod

V předchozím článku¹ jsme popsali teorii a implementaci neuronových sítí se strategií zpětného šíření². Tento článek bude popisovat aplikace těchto sítí v chemii.

2. Aplikace neuronových sítí v chemii

Aplikace neuronových sítí³ jsou velmi rozmanité, od rozpoznávání textů až po řízení procesů. V této práci se zaměříme pouze na aplikace v chemii a ve velmi malé míře i ve fyzice. Téměř výhradně se v chemických aplikacích používá vrstevnatá neuronová síť s adaptací strategie zpětného šíření². Nebude-li napsáno jinak, bude se ve zbyvajícím textu článku rozumět tento typ sítě.

Aplikace v chemii lze rozdělit na několik oblastí:

- 1) Spektroskopie (často uplatněná v analytické chemii) nejrozličnějšího typu MS, NMR, UV/VIS, IR a další.
- 2) Řízení procesů v chemickém inženýrství
- 3) Chemická reaktivita
- 4) Bioaplikace - struktura proteinů, rozpoznávání promotorních míst v DNA a další
- 5) Ostatní aplikace

2.1. Spektroskopické aplikace

Jedná se o nejčastější a nejúspěšnější aplikace neuronových sítí v chemii. Spektra jsou nejrozličnějšího druhu: MS, NMR, UV/VIS, IR, gama, rentgenová fluorescenční. Dále můžeme zmínit chromatografická data.

V práci⁴ se používá neuronová síť pro klasifikaci nízko rozlišených hmotnostních spekter neznámých sloučenin. Zjišťuje se přítomnost či absence 100 chemických skupin. Autoři v práci diskutují představy o designu neuronových sítí, o jejich statistických vlastnostech a efektech různých metod tréninku. Autoři zjistili, že neuronová síť nabízí slibnou alternativu k existujícím metodám spektroskopické klasifikace. Neuronová síť je vysoce flexibilní, rychlá a tolerantní k vstupním chybám, klasifikuje MS spektra lépe než další metody publikované v literatuře a má další žádoucí vlastnosti. Autoři práce⁴ očekávají, že může být dosaženo velkého zlepšení interpretace MS spekter kombinací s dalšími zdroji dat, např. s IR, ¹H-NMR a ¹³C-NMR spektry. Výpočty byly prováděny na počítačích HP 9000/370, Sun-4, Sparcstation I.

Další oblastí aplikace neuronových sítí je IR spektroskopie. V práci⁵ autoři použili perceptron (t.j. neuronovou síť bez skrytých neuronů) jako model pro automatické interpretace IR spekter. Model byl trénován užitím databáze IR spekter organických sloučenin známých struktur. Užití perceptronu se ukazuje v této práci jako nadějná praktická metoda pro automatickou interpretaci IR dat organických sloučenin.

Titíž autoři v práci⁶ používají už mnohvrstevnou neuronovou síť (t.j. mající skryté neurony) pro identifikaci funkčních skupin v organických sloučeninách, která je založená na IR spektroskopii. Získané výsledky jsou pod-

statně lepší než u jednoduchého lineárního modelu publikovaného dříve⁵.

Autoři⁵ zprogramovali neuronové sítě v jazycích Pascal a C. Předběžné výpočty prováděli na počítači VAX 3500, další na superpočítači CRAY.

Práce⁷ se zabývá rozpoznáním píků IR spekter. Autoři v práci⁸ měřili blízká IR spektra proteinů pšenice a UV/VIS spektra farmaceutických produktů. Pro tyto druhy spekter (simulovaných i reálných) vyhodnocovali výsledky neuronovými sítěmi a porovnávali je s principal component regression, PCR. Pro reálná IR spektra byly výsledky PCR mírně lepší než u neuronových sítí, v případě UV/VIS dává neuronová síť lepší výsledky než PCR.

V práci⁸ bylo použito počítače Zenit 80386 33 MHz, v budoucnu se autoři chystají použít speciální paralelní hardware.

Titíž autoři v práci⁹ popisují metody pro detekci a modelování nelineárních oblastí spektrálních odpovědí. Simulovaná i experimentální UV/VIS data byla použita pro studium schopnosti multivariačních lineárních modelů pro aproximaci nelineárních odpovědí. Zdroje reálných a zdánlivých nelinearit autoři simulovali včleněním nelineárních funkcí odpovědi (např. rozptýlené světlo, koncentračně závislý posun vlnových délek a koncentračně závislé absorpční šířky).

V článkách^{10,11} se popisuje použití neuronových sítí typu asociativní paměti pro řešení automatické radioizotopové identifikace užitím energetického spektra.

V práci¹⁰ byly použity dva typy detektoru Ge(Li) a Na(Tl). Trénink neuronové sítě musí být vždy proveden pro týž spektrometr a totéž geometrické uspořádání. Izotopy ²²Na, ⁶⁰Co, ¹³³Ba, ¹³⁷Cs, ⁵⁷Co byly použity v tréninkové množině. Výsledky výpočtů pro naměřenou testovací množinu udávají, kolikrát je spektrum daného izotopu obsaženo v testovacím spektru směsi.

Výsledky vycházejí velmi dobře, s přesností lepší než 5 %. Statistické fluktuace, ztěžující identifikaci spekter, dělají neuronové sítě ještě atraktivnějšími než ostatní metody rozpoznávání. Problém nastane, když spektrum obsahuje izotop, pro který není neuronová síť natrénována. Výsledek je pak chybný, neboť neuronová síť špatně interpretuje píky.

V práci¹¹ stejní autoři provádějí pomocí neuronové sítě typu asociativní paměti automatickou analýzu radiačního spektra. Používají tyto izotopy: ²²Na, ⁶⁰Co, ⁵⁴Mn, ¹³⁷Cs, ¹⁰⁹Cd, ¹³³Ba, ⁵⁷Co a NaI detektor. Spektra byla mnohem složitější než v předešlé práci¹⁰.

Ukázalo se, že neuronová síť správně zpracovává vzor-

ky, které mají řádově stejnou či nižší koncentraci, jako ty vzorky, které byly použity v tréninkové množině. Výsledky obdržené pomocí neuronové sítě vycházejí ve všech případech lepší než při použití standardní metody.

Neuronové sítě jsou zvláště vhodné pro spektra s nízkým rozlišením a jsou velmi nadějnou i jednoduchou metodou pro řešení rozpoznávání radiačního spektra. Výpočty¹⁰ byly realizovány na PC AT.

V článku¹² se používá neuronové sítě jako matematického modelu pro kalibraci dat rentgenové fluorescenční spektroskopie. Autoři použili dvě tréninkové metody: strategie zpětného šíření a genetický algoritmus. V případě malé tréninkové množiny (13 členů), vytvořené pomocí dat z literatury pro Fe/Ni/Cr, strategie zpětného šíření dává srovnatelně dobré výsledky jako jiné metody z literatury.

Poslední spektroskopickou oblastí, kde se vyskytují aplikace neuronových sítí, jsou ¹H-NMR a ¹³C-NMR spektroskopie. Aplikace lze rozdělit na dvě části, první část obsahuje rozpoznávání, identifikaci a interpretaci naměřených NMR spekter, druhá část se týká předpovědi chemických posunů pro látky určité chemické skupiny.

V práci¹³ autoři použili neuronovou síť pro rozpoznávání ¹H-NMR spekter 6-ti druhů (t.j. je použito 6 výstupních neuronů) cukerných alditolů. Výpočty byly prováděny na počítači IBM PC AT 80286.

Článek¹⁴ se zabývá analýzou pulsů NMR spekter. Použitím neuronové sítě se 100 vstupními, 200 skrytými a 16 výstupními neurony se autoři snaží získat 16 Fourierových koeficientů ze spektra.

Ve velmi zajímavé práci¹⁵ je neuronová síť použita k identifikaci ¹H-NMR spekter komplexních velikých oligosacharidů. Pro rozpoznávání jsou využívány oblasti silného překryvu pásů (spektrum vyššího řádu) i vyřešené oblasti spektra (spektrum 1. řádu). Pouze malé procento signálů ze spektra je u těchto látek separovatelné v individuálně rozložené NMR multiplety. Nejvíce signálů je spojeno v oblasti silného překryvu. Tento typ spektra je typický pro spektra mnoha dalších biologicky důležitých molekul, včetně DNA, RNA a proteinů. Autoři použili jednodimensionální 500 MHz ¹H-NMR spektra. Autoři testovali, jestli se na rozpoznávání spektra účastní překryté i nepřekryté (vyřešené oblasti). ¹H-NMR spektra 5 oligosacharidů byla rozdělena na dvě množiny, jedna obsahuje pouze nepřekryté (vyřešené oblasti), druhá množina obsahuje pouze překryté (spektrum vyššího řádu) oblasti. Neuronové sítě pro oba druhy oblastí zkonvergovaly a byly schopny rozpoznat každé ze spekter. Rozpoznání překrytých oblastí je méně intuitivní, a proto autoři interpretovali

tento výsledek myšlenkou, že neuronová síť může rozlišit mezi spektry, u kterých rozdíly nejsou objeveny lidským pozorováním. I překrytá oblast obsahuje informaci nezbytnou k rozeznání mezi oligosacharidy.

Neuronové sítě se jeví jako nadějný nástroj v rozpoznávání NMR spekter, není potřeba žádných definovaných pravidel a výpočty jsou rychlejší než klasické knihovní hledání. Autoři¹⁵ používali komerčně přístupný software neuronových sítí na Silicon Graphics 4D/220 GTX počítači.

V práci¹⁶ je využíváno neuronových sítí pro automatické přiřazení proteinů.

Jak již bylo zmíněno výše, druhá část prací, používající neuronové sítě v NMR spektroskopii, se týká předpovědi ¹³C-NMR chemických posunů. V článkách^{17,18} odvozuje řešení paradigmatu strategie zpětného šíření v takové formě, aby bylo použitelné pro gradientovou optimalizaci účelové funkce metodou proměnné metriky.

V práci¹⁷ popisujeme odhad chemických posunů v ¹³C-NMR spektroskopii pro acyklické alkany, které byly popsány pomocí grafového formalismu.

V článku¹⁸ neuronová síť počítá odhad ¹³C-NMR chemických posunů sekundárních uhlíkových atomů v acyklických alkanech. Opět se trénovací i testovací objekty popisují pomocí grafového formalismu.

Výsledky ukazují, že neuronová síť je schopna dobře předpovídat chemické posuny ¹³C-NMR spektroskopie.

V práci¹⁹ jsme počítali ¹³C-NMR posuny *ipso*, *ortho*, *meta* a *para* uhlíků benzenového kruhu monosubstituovaných benzenů. Používali jsme dva způsoby zadávání struktury substituentu.

Jednodušší způsob vycházel z práce²⁰. Složitější popis byla tabulka spojení reaktantů, která byla převzata z práce²¹.

Je zajímavé¹⁹, že jednodušší deskriptor poskytoval lepší výsledky než složitější. Je to dáno zřejmě tím, že neuronová síť nebyla schopna v případě složitějšího popisu správně "rozpoznat" zadání, neboť asi bylo pro ni příliš "složitě".

Práce²² popisuje neuronové sítě jako nástroj rozpoznávání vzorů chromatografických dat sedmi různých tříd proudových paliv. Data byla popsána pomocí binárních vzorů, parametry a architektura neuronové sítě byly optimalizovány tak, aby dávaly co nejlepší výsledky. Výpočty se ukázaly jako životaschopné a užitečné pro klasifikaci chromatografických dat proudových paliv.

2.2. Řízení procesů

Jednou ze zajímavých aplikací použití neuronových

sítí v chemii je řízení chemických procesů.

Jako ilustrační příklad můžeme uvést práci²³. Autoři použili dvoustupňovou neuronovou síť pro diagnostiku procesu reakce $C_7H_{10} \rightarrow C_7H_8 + 4 H_2$

První stupeň neuronové sítě rozlišuje mezi příčinami chyb a určuje druh chyby, druhý stupeň odhaduje velikost chyby. Jako vstup do 1. stupně slouží koncentrace C_7H_{10} , teplota ohřivače a výstupní signál řídicí jednotky. Neuronové sítě mají kapacitu pro učení a uložení informace o chybách procesů a jsou použitelné pro klasifikaci chybových stavů.

2.3. Chemická reaktivita

Velice lákavé a nadějné jsou aplikace neuronových sítí v oblasti chemické reaktivity. Tento přístup sice nemůže poskytnout více informací než seriózní rozbor založený na metodách kvantové chemie, ale přesto je využitelný pro řešení této problematiky. Kromě exaktních přístupů k problematice chemické reaktivity založených na principech kvantové mechaniky, existuje i mnoho dalších přístupů, které poskytují lepší či horší výsledky.

Jedním z výše uvedených přístupů jsou expertní systémy vytvořené z pravidel reaktivity daných explicitně. Neuronová síť nepotřebuje žádná pravidla, učí se z příkladů. Nedovede ale vysvětlit důvody svého rozhodování.

V práci²¹ se autoři zabývali elektrofilní aromatickou substitucí monosubstituovaných benzenů. Pro reprezentaci dat byly zvoleny dva postupy. Více úspěšný - tabulka spojení reaktantů, méně úspěšný - reprezentace nábojového vektoru.

Tabulka spojení reaktantů je vhodná metoda pro reprezentaci informace, odpovídající chemické struktuře, pro počítání.

Reprezentace nábojového vektoru - pomocí programu MOPAC autoři spočítali náboje na uhlíkových atomech aromatického kruhu pro dané sloučeniny.

Dva výstupní neurony (po vynásobení 100) udávají procenta frakce *ortho/para* a *meta* produktu.

Výsledky pro obě neuronové sítě byly srovnány s předpověďmi expertního systému CAMEO a s předpověďmi tří experimentálních chemiků.

Neuronové sítě předpovídaly výtěžky elektrofilní aromatické substituce benzenů korektně, výpočty s tabulkou spojení reaktantů dávaly lepší výsledky než druhý přístup. Není ale možno předpokládat, že tabulka spojení reaktantů je obecně dobrý prostředek popisu chemických entit, viz¹⁹.

Zajímavé může být spojení přístupu neuronových sítí s expertními systémy.

V článku²⁴ se popisují předpovědi produktů pro tři typy reakcí:

- 1) Pro Markovnikovu adici HX na alkeny se předpovídá typ produktu.
- 2) Pro Dielsovu-Alderovu a retro Dielsovu-Alderovu cykloadici nesymetrických dienů a dienofilů s nesymetrickými dienofily se předpovídá typ produktu a též se testuje možnost rozpojení sloučeniny v příslušné retro Diels-Alderově reakci.
- 3) Pro Zajcevovu E1 eliminaci alkenů se předpovídá typ produktů.

Reprezentace reakce se prováděla pomocí modifikované Dugundjiho-Ugiho matice. BE matice je modifikovaná v tom, že prvky diagonály obsahují protonová čísla příslušných atomů.

Pro každou reakci byla zvlášť navržena neuronová síť. Protože BE matice je symetrická, stačilo použít její dolní trojúhelník. Počet prvků dolního trojúhelníku (i s diagonálou) určoval počet vstupních neuronů. Modifikovaná BE matice se ukázala jako vhodný deskriptor popisu sloučenin (reakcí) pro neuronové sítě.

V práci²⁰ jsme použili neuronových sítí pro předpovězení výtěžků *meta* produktu nitrace monosubstituovaných benzenů.

Pro vstup již navázané funkční skupiny jsme se pokusili použít takové parametry vyjadřující strukturu, kterými by bylo možné zpracovat většinu organických sloučenin. Popis struktur byl založen na popisu struktur použitým již Zouem a spol.²⁵ Používali jsme tyto vstupy: přítomnost kladného nebo záporného náboje na atomu, volný elektronový pár na atomu, hlavní kvantové číslo mínus jedna, počet vodíkových atomů a konečně násobnost vazby spojující daný atom směrem k reakčnímu centru. Tyto údaje byly použity pro dvě úrovně vzdáleností od benzenového jádra.

Použili jsme dvě různé topologie neuronových sítí spolu s odpovídající adaptací vstupu, jednu klasickou a druhou se stavbou neuronové sítě odrážející vnitřní logickou strukturu vstupních informací. Obě topologie dávaly kvalitativně dobré výsledky předpovědi výtěžků *meta* produktu nitrace monosubstituovaných benzenů. Výpočty byly provedeny na počítači IBM PC 80386.

V práci²⁶ autoři řešili rozdělení kondenzace na aldovou, Claisenovu, Michaelovu a enaminovou, kde za ústřední můžeme považovat skupinu C=O. Skupiny, které byly pak navázány na jedné i druhé straně uhlíkového

řetězce, jehož je C=O středem, byly pak ohodnoceny také svou "vzdáleností" od skupiny C=O, vyjádřenou celkem osmi vstupními neurony pro čtyři úrovně okolí C=O na pravé i levé straně.

Síť byla trénována na 34 příkladech a testována na 60 dalších. Asi v 80 % případů byly získány rozumné předpovědi.

Poslední popsaná práce²⁷ se liší od všech předešlých. Autoři prováděli lipasou katalyzovanou reakci mastných kyselin s glukosidy a dostávali 6-O-monoestery či 2,6-O-diestery po 22 hodinách vzájemného reagování. První výstupní neuron udával výtěžek 6-O-monoesteru a druhý 2,6-O-diesteru. Jako vstupní deskriptor byly brány tyto údaje: 1) poměr mastná kyselina/glukosid [mol/mol], 2) reakční teploty v stupních Celsia, 3) tlak v mbar, 4) množství enzymu v %, 5) rychlost otáček míchadla.

Výsledky testovací množiny byly velice povzbudivé, chyby oproti experimentálním hodnotám činily pouze několik procent. Dále autoři zjišťovali, pro které reakční podmínky dostanou více než 88 %-ní výtěžek 6-O-monoesteru. Vygenerovali soubor s 9900 reakčními podmínkami, kde parametry deskriptoru pokrývaly všechny reálné hodnoty, a testovali tuto velkou množinu. Pro tři data dostali výtěžek vyšší než 88 %, a proto jedny z těchto tří reakčních podmínek vyzkoušeli a dostali výsledky velmi blízké vypočteným. Ukázalo se tedy, že neuronová síť může poskytovat velmi dobré výsledky pro plánování reakčních podmínek enzymatické reakce. Doufejme, že tyto výsledky se budou moci rozšířit a zobecnit i na jiné typy reakcí.

Výpočty byly prováděny na počítači IBM-PS/2 model 50.

2. 4. B i o a p l i k a c e

Další oblastí, kde se s úspěchem používá neuronových sítí, jsou bioaplikace, t.j. struktura proteinů, rozpoznání promotorních míst v DNA a jiné.

V práci²⁸ je tento druh aplikací souhrnně popsán. V článcích^{28,29,30} autoři zjišťovali pomocí neuronových sítí sekundární strukturu proteinů z primární.

Vstup pro neuronové sítě³⁰ tvořilo "okno" 13 aminokyselin. Každá aminokyselina měla přiřazen jeden z 21 vstupních neuronů, celkově tedy bylo použito 273 vstupních neuronů. Toto "okno" se posunovalo po proteinu z jednoho konce na druhý. Pro každou pozici "okna" na proteinu se pro prostřední aminokyselinu "okna" předpovídala sekundární struktura proteinu. Úspěšnost předpovědi byla přibližně 63-65 %, což je více než u metod dříve používaných v literatuře.

Sekundární struktura proteinů není pouze lokální vlastnost posloupnosti aminokyselin, a proto je odhadován teoretický limit rozpoznání sekundární struktury pro lokální metody 70 %.

Autoři prací^{28,29,30} optimalizovali struktury svých neuronových sítí, zkoušeli vliv přítomnosti homologických proteinů v tréninkových množinách. Celkem se dají hodnotit aplikace neuronových sítí v této oblasti jako úspěšné.

V další práci se autoři snažili o předpověď terciární struktury proteinů³¹. V článku³² neuronové sítě rozpoznávaly promotorní oblasti v DNA s úspěšností 94-99 %. Pro kódování DNA bylo použito "okno" o "šířce" 6 bází, každá báze byla kódována dvěma neurony (t.j. celkově bylo použito 12 vstupních neuronů).

2.5. Ostatní aplikace

Neuronové sítě byly použity v nejrůznějších oblastech chemie někdy s většími, jindy s menšími úspěchy. Jsou použitelné jako nelineární model pro mnoho závislostí a vztahů, které buď nejsou vůbec přesně známy, či jsou, ale pro svou složitost jsou nepoužitelné pro praktické výpočty. Ale i neuronové sítě mají svá omezení. Dopředu se nedá říci, jestli se podaří daný problém vyřešit pomocí neuronové sítě či nikoli. Neuronové sítě nemohou přinést nic zcela nového, pouze zobecňují už poznané. Velice důležitý je správný výběr deskriptoru a designu neuronové sítě. Jinak lze neuronové sítě doporučit především pro řešení takových úloh, jako je zobecnění závislostí atd.

Autoři v práci³³ popisují teoretické základy aplikace neuronových sítí pro "pole" elektrod. V článku³⁴ stejní autoři použili "pole" iontově selektivních elektrod pro současné stanovení H^+ , K^+ , Ca^{2+} , NO_3^- , Cl^- iontů. Natrénovali neuronovou síť pro různé koncentrace všech použitých iontů v roztoku a pak natrénovanou neuronovou síť použili pro současné stanovení neznámých koncentrací těchto iontů. Výsledky měly chybu 8 %.

V článku³⁵ autoři používali neuronovou síť pro výpočty vlastních energií energetických hyperploch harmonických oscilátorů.

Stejní autoři v práci³⁶ počítali neuronovou síť tepelné kapacity c_p pro polymery při nízkých teplotách. Předpokládali c_p v teplotním rozsahu 10-100 K ze série c_p pro teploty 110-360 K. Měření c_p při nízkých teplotách je mnohem obtížnější. Jako vstup pro 25 neuronů byly c_p pro teploty 110, 120, ..., 350, 360 K. 10 výstupních neuronů udávalo c_p pro teploty 10, 20, ..., 80, 90 K s průměrnou chybou 0,34 %. Změřené c_p pro teploty 100-360 K dovolí dobře

extrapolovat c_p pro nižší teploty, pro které nelze c_p dobře měřit. Znalost c_p pro nízké teploty je potřebná pro vyčíslení celkového tepelného obsahu v polymerech. Výpočty byly prováděny na pracovní stanici VAX 2000.

Článek³⁷ se zabývá expertním systémem pro vyhodnocování mechanismu elektrodoých reakcí.

Práce³⁸ používá křemenný krystal jako chemický senzor pro stanovení látek ve voňavkách. Vyhodnocování rezonančních frekvenčních posunů bylo prováděno neuronovou sítí.

Článek³⁹ popisuje inteligentní stroje pro analýzu a sběr experimentálních dat. Zabývá se virtuálními přístroji, expertními systémy a neuronovými sítěmi.

V práci⁴⁰ je popisováno zjišťování endo či exo konformace norborenů pomocí neuronových sítí. Jako vstup sloužily chemické ^{13}C -NMR posuny.

Velmi nadějnou oblastí použití neuronových sítí jsou QSAR studie⁴¹⁻⁴⁹. Autoři v práci⁴² zkoumali vztah struktura - aktivita pro deriváty karbochinolu. Jako vstup sloužilo 6 přepočítaných QSAR parametrů. Neuronová síť poskytla lepší výsledky než standardní metody.

Článek⁴¹ popisuje: 1) výpočty antikarceroenních aktivit 16 karbochinonových derivátů pomocí neuronových sítí, 2) vztah struktura - vůně, 3) vztah struktura - chuť. Neuronové sítě dosáhly lepších výsledků než standardní metody popsáné v literatuře.

3. Závěr

Příklady uvedené v předcházející kapitole dostatečně ilustrují možnosti použití neuronových sítí v chemii, jmenovitě mnohovrstevných neuronových sítí s aplikací strategie zpětného šíření. Třemi nejzávažnějšími úkoly při návrhu použití neuronových sítí pro řešení konkrétního chemického problému se jeví: konstrukce vhodných deskriptorů popisujících strukturu chemického objektu, návrh topologie neuronové sítě a konečně výběr vhodné tréninkové a testovací množiny. Současné aplikace neuronových sítí v chemii řeší tyto problémy ad-hoc způsobem bez hlubšího předzpracování (např. statistického) vstupních údajů. Toto způsobuje obtížné vyhodnocení a vzájemné porovnání výsledků výpočtů různých autorů. Proto se domníváme, že aplikace neuronových sítí v chemii stojí v současnosti před vážným teoretickým problémem seriózního statistického vyhodnocení jak jejich vstupů, tak i jejich výstupů. Teprve na základě této analýzy bude možno kvalifikovaně posuzovat schopnost predikce a klasifikace neuronové sítě a vzájemně porovnávat práce různých autorů.

LITERATURA

1. Kvasnička V., Sklenák Š., Pospíchal J.: *Chem. Listy* 87, 79 (1993).
2. Rumelhart D. E., McClelland J. L.: *Parallel Distributed Processes*, Vols. 1-2. MIT Press, Cambridge 1986.
3. Zupan J., Gasteiger J.: *Anal. Chim. Acta* 248, 1 (1991).
4. Curry Bo., Rumelhart D. E.: *Tetrahedron Comput. Methodol.* 3, 213 (1990).
5. Robb E. W., Munk M. E.: *Mikrochim. Acta (Wien)* 1, 131 (1990).
6. Munk M. E., Madison M. S., Robb E. W.: *Mikrochim. Acta (Wien)* 2, 505 (1991).
7. Wythoff B. J., Levine S. P., Tomellini S. A.: *Anal. Chem.* 62, 2702 (1990).
8. Long J. R., Gregoriou V. G., Gemperline P. J.: *Anal. Chem.* 62, 1791 (1990).
9. Gemperline P. J., Long J. R., Gregoriou V. G.: *Anal. Chem.* 63, 2313 (1991).
10. Olmos P., Diaz J. C., Perez J. M., Gomez P., Rodellar V., Aguayo P., Bru A., Garcia-Belmonte G., De Pablos J. L.: *IEEE Trans. Nucl. Sci.* 38, 971 (1991).
11. Olmos P., Diaz J. C., Perez J. M., Garcia-Belmonte G., Gomez P., Rodellar V.: *Nucl. Instrum. Methods Phys. Res., Sect. A*, A312, 167 (1992).
12. Bos M., Weber H. T.: *Anal. Chim. Acta* 247, 97 (1991).
13. Thomsen J. U., Meyer B.: *J. Magn. Reson.* 84, 212 (1989).
14. Gezelter J. D., Freeman R.: *J. Magn. Reson.* 90, 397 (1990).
15. Meyer B., Hansen T., Nute D., Albersheim P., Darvill A., York W., Sellers J.: *Science* 251, 542 (1991).
16. Kjaer M., Poulsen F. M.: *J. Magn. Reson.* 94, 659 (1991).
17. Kvasnička V.: *Chem. Papers* 44, 775 (1990).
18. Kvasnička V.: *J. Math. Chem.* 6, 63 (1991).
19. Kvasnička V., Sklenák Š.: *Collect. Czech. Chem. Comm., (zasláno)*.
20. Kvasnička V., Pospíchal J.: *J. Mol. Struct. (THEOCHEM)* 235, 227 (1991).
21. Elrod D. W., Maggiora G. M., Trenary R. G.: *J. Chem. Inf. Comput. Sci.* 30, 477 (1990).
22. Long J. R., Mayfield H. T., Henley M. V., Kromann P. R.: *Anal. Chem.* 63, 1256 (1991).
23. Watanabe K., Matsuura I., Abe M., Kubota M., Himmelblau D. M.: *AICHE J.* 35, 1803 (1989).
24. Elrod D. W., Maggiora G. M., Trenary R. G.: *Tetrahedron Comput. Methodol.* 3, 163 (1990).
25. Zou Y., Johnson M. A., Tsai C.-C.: *J. Chem. Inf. Comput. Sci.* 30, 442 (1990).
26. Luce H. H., Rakesh G.: *Tetrahedron Comput. Methodol.* 3, 143 (1990).
27. Kirk O., Barfoed M., Bjoerklig F.: *Tetrahedron Comput. Methodol.* 3, 239 (1990).
28. Petersen S. B., Bohr H., Bohr J., Brunak S., Cotterill R. M. J., Fredholm H., Lautrup B.: *Trends Biotechnol.* 8, 304 (1990).
29. Holley L. H., Karplus M.: *Proc. Natl. Acad. Sci. USA* 86, 152 (1989).
30. Qian N., Sejnowsky T. J.: *J. Mol. Biol.* 202, 865 (1988).
31. Friedrichs M. S., Wolynes P. G.: *Tetrahedron Comput. Methodol.* 3, 175 (1990).
32. Lukashin A. V., Anshelevich V. V., Amirikyan B. R., Gragerov A. I., Frank-Kamenetskii M. D.: *J. Biomol. Struct. Dyn.* 6, 1123 (1989).
33. van der Linden W. E., Bos M., Bos A.: *Anal. Proc. (London)* 26, 329 (1989).
34. van der Linden W. E., Bos M., Bos A.: *Anal. Chim. Acta* 233, 31 (1990).
35. Darsey J. A., Noid D. W., Upadhyaya B. R.: *Chem. Phys. Letters* 177, 189 (1991).
36. Darsey J. A., Noid D. W., Wunderlich B., Tsoukalas L.: *Makromol. Chem. Rapid Commun.* 12, 325 (1991).
37. Palys M., Bos M., van der Linden W. E.: *Anal. Chim. Acta* 248, 429 (1991).
38. Chang S. M., Awasaki Y., Suzuki M., Tamiya E., Karube I.: *Anal. Chim. Acta* 249, 323 (1991).
39. Barker P. G.: *Anal. Proc. (London)* 28, 110 (1991).
40. Aoyama T., Suzuki Y., Ichikawa H.: *Chem. Pharm. Bull.* 37, 2558 (1989).
41. De Saint Laumer J. Y., Chastrette M., Devillers J.: *Výzkumná zpráva, Laboratoire de Chimie Organique Physique, Lyon*.
42. Aoyama T., Suzuki Y., Ichikawa H.: *J. Med. Chem.* 33, 905 (1990).
43. Aoyama T., Ichikawa H.: *Chem. Pharm. Bull.* 39, 372 (1991).
44. Livingstone D. J., Hesketh G., Clayworth D.: *J. Mol. Graphics* 9, 115 (1991).
45. Andrea T. A., Kalayeh H.: *J. Med. Chem.* 34, 2824 (1991).
46. Aoyama T., Suzuki Y., Ichikawa H.: *J. Med. Chem.* 33, 2583 (1990).

47. Hussain A. S., Yu X., Johnson R. D.: *Pharm. Res.* 8, 1248 (1991).
48. Rose V. S., Croall I. F., MacFie H. J. H.: *Quant. Struct.-Act. Relat.* 10, 6 (1991).
49. Andrea T. A., Kalayeh H.: *Pharmacochem. Libr.* 16, 209 (1991).

V. Kvasnička, Š. Sklenák and J. Pospíchal
(Department of Mathematics, Slovak Technical University,

Bratislava): Neural Networks in Chemistry II. Applications

Different applications of neural networks in chemistry are described. In particular, spectroscopic applications are discussed, where this approach offers powerful algorithmic tools for the classification and interpretation of spectra of different nature. Neural networks are also very effective for the classification and prediction of chemical reactivity of organic molecules.