

České vysoké učení technické v Praze

Fakulta stavební

Katedra betonových a zděných konstrukcí



BAKALÁRSKA PRÁCA

Využitie neurónových sietí pre tvorbu materiálových modelov

Utilization of Neural Networks in Material Modeling

Vypracovala: Patrícia Benková

Vedúci práce: Ing. Martin Petřík, Ph.D.

Rok: 2018



ZADÁNÍ BAKALÁŘSKÉ PRÁCE

I. OSOBNÍ A STUDIJNÍ ÚDAJE

Příjmení: Benková	Jméno: Patrícia	Osobní číslo: 438935
Zadávací katedra: Katedra betonových a zděných konstrukcí		
Studijní program: Stavební inženýrství		
Studijní obor: Konstrukce pozemních staveb		

II. ÚDAJE K BAKALÁŘSKÉ PRÁCI

Název bakalářské práce: Využití neuronových sítí pro tvorbu materialových modelů	
Název bakalářské práce anglicky: Utilization of Neural Networks in Material Modeling	
Pokyny pro vypracování:	
1) V práci bude teoreticky popsána problematika neuronových sítí a jejich aplikace	
2) Vytvoření programu v jazyce Python pro tvorbu materiálových modelů betonu založených na neuronových sítích	
Seznam doporučené literatury:	
Tariq Rashid - Make your own Neural Network	
Yaser S. Abu-Mpstafa - Learning from Data	
Joel Grus - Data Science from Scratch	
Mario Collepardi - Moderní beton	
Jméno vedoucího bakalářské práce: Ing. Martin Petřík, Ph.D.	
Datum zadání bakalářské práce: 21. 02. 2018	Termín odevzdání bakalářské práce: 27. 5. 2018
Údaj uveďte v souladu s datem v časovém plánu příslušného ak. roku	
Podpis vedoucího práce	Podpis vedoucího katedry

III. PŘEVZETÍ ZADÁNÍ

Beru na vědomí, že jsem povinen vypracovat bakalářskou práci samostatně, bez cizí pomoci, s výjimkou poskytnutých konzultací. Seznam použité literatury, jiných pramenů a jmen konzultantů je nutné uvést v bakalářské práci a při citování postupovat v souladu s metodickou příručkou ČVUT „Jak psát vysokoškolské závěrečné práce“ a metodickým pokynem ČVUT „O dodržování etických principů při přípravě vysokoškolských závěrečných prací“.

Datum převzetí zadání	Podpis studenta(ky)
-----------------------	---------------------

Čestné prehlásenie

Prehlasujem, že som túto bakalársku prácu vypracovala samostatne a iba za odborného vedenia vedúceho bakalárskej práce Ing. Martina Petříka, Ph.D.

Ďalej prehlasujem, že všetky podklady, z ktorých som čerpala, sú uvedené v priloženom zozname použitej literatúry.

V Prahe dňa 27. 5. 2018

.....

Patrícia Benková

Podakovanie

Na tomto mieste by som sa rada poďakovala všetkým, ktorí mi boli pri spracovaní tejto práce nápomocní či už radou, podkladmi, alebo podporou. Moje najväčšie ďakujem patrí vedúcemu práce Ing. Martinovi Petříkovi, Ph.D. za odborné vedenie, cenné rady, motiváciu a trpezlivosť počas riešenia bakalárskej práce. Tiež by som rada poďakovala všetkým účastníkom seminára za inšpiratívnu a priateľskú atmosféru na hodinách. Zvlášťne poďakovanie chcem vyjadriť rodine za motiváciu a podporu počas štúdia.

Abstrakt

V tejto práci je predstavené využitie umelých neurónových sietí pri predikcii správania materiálov. Hlavnou motiváciou je tvorba pracovného diagramu betónu raného veku s cieľom vymodelovať vhodnú neurónovú sieť a porovnať jej výstupy s reálnym chovaním betónu počas zaťažovania. Ako základ pre modelovanie slúžili výsledky zo zaťažovacích skúšok betónu veľmi raného veku, ktoré sa v minulosti uskutočnili na Fakulte stavebnej ČVUT. Na začiatku práce sú najprv objasnené princípy procesov prebiehajúcich v umelých neurónových sieťach a následne je vysvetlený postup pri tvorbe vlastného modelu umožňujúceho predpovedať analyzované vlastnosti. Samotná tvorba pracovného diagramu betónu sa nachádza v záverečnej časti práce.

Kľúčové slová

Neurónové siete, python, materiálový model, betón raného veku, predikcia chovania.

Abstract

This thesis deals with a prediction of material behavior with a usage of artificial neural networks. The main motivation is to create a stress-strain diagram of early-age concrete with the aim to model a suitable neural network and compare its outputs with the real behavior of concrete specimens. As an input for the model, test results of concrete specimens from experiments at CTU's Faculty of Civil engineering were used. The thesis starts with the summarised principles of processes in artificial neural networks and is followed by the steps required to construct a model able to predict the analyzed properties. The construction of the model itself is at the end of this work.

Keywords

Neural Networks, python, material model, early age concrete, prediction of behaviour.

Obsah

Abstrakt	iii
Zoznam obrázkov	vi
Zoznam tabuliek	viii
Zoznam symbolov	ix
1 Úvod	1
2 Neurónové siete	3
2.1 Úvod do neurónových sietí	3
2.2 História a použitie	5
2.3 Mechanizmus neurónových sietí	7
2.3.1 Priebeh signálov sieťou	7
2.3.2 Aktivačná funkcia	9
2.3.3 Vyhodnotenie výstupov zo siete	13
3 Tvorba vlastnej neurónovej siete	16
3.1 Python	16
3.2 Maximálny moment na prostom nosníku	17
4 Zloženie betónovej zmesi	31
4.1 Chemické prísady	31
4.2 Superplastifikátory	32
4.3 Spracovateľnosť betónovej zmesi	33
4.4 Aproximácia nespojitej funkcie	38

5	Pevnosť betónu raného veku	41
5.1	Pracovný diagram betónu	41
5.2	Problematika betónu raného veku	43
5.3	Materiálové modely pre betón raného veku	45
5.4	Experimentálne skúmanie vlastností betónu raného veku	46
5.5	Tvorba modelu umelej neurónovej siete	49
5.6	Analýza výsledku	54
6	Záver	56
A	Tabuľka aktivačných funkcií	60
B	Vzor forward propagation	61
C	Náhodnosť počiatočných hodnôt váh	64

Zoznam obrázkov

2.1	Porovnanie biologického neurónu s umelým	4
2.2	Príklad architektúry siete	5
2.3	John Hopfield	6
2.4	Sieť 3x3	7
2.5	Násobenie vstupných dát váhami	8
2.6	Aplikácia aktivačnej funkcie	10
2.7	Skoková funkcia	11
2.8	Lineárna funkcia	11
2.9	Funkcia sigmoidu	12
2.10	Graf hyperbolického tangensu	13
3.1	Grafické porovnanie hodnôt M_{max}	19
3.2	Priebeh tréovania siete	20
3.3	Grafické porovnanie hodnôt M_{max} po tréovaní	20
3.4	Príklad overfittingu	21
3.5	Príklad overfittingu na riešenom príklade 2D	22
3.6	Príklad overfittingu na riešenom príklade 3D	22
3.7	Priebeh tréovania siete po rozdelení dát	24
3.8	Výstupy zo siete po rozdelení dát 2D	25
3.9	Výstupy zo siete po rozdelení dát 3D	25
3.10	Vplyv regularizácie na overfitting	27
3.11	Vplyv regularizácie pre odlišné λ	28
3.12	Priebeh tréovania siete po zavedení λ	29
3.13	Zmena váh počas tréovania s použitím λ	29
3.14	Zmena váh počas tréovania bez použitia λ	30

3.15	Konečné výstupné dáta 3D	30
4.1	Schéma vývoja straty spracovateľnosti betónu so superplastifikátorom SLCA	33
4.2	Skúška sadnutia kužela	34
4.3	Závislosť spracovateľnosti zmesi na množstve plastifikátora SLCA	35
4.4	Výstupy zo siete pred trénovaním pre rôzne seedy	36
4.5	Výstupy zo siete po trénovaní pre rôzne seedy	37
4.6	Odlišné miery regularizácie pre np.random.seed(1)	37
4.7	Závislosť spracovateľnosti zmesi na množstve plastifikátora SLCA - upravené	38
4.8	Vplyv rozloženia vstupných dát na výstupy zo siete	40
5.1	Charakteristický tvar pracovného diagramu betónu v tlaku	42
5.2	Zjednodušené pracovné diagramy	43
5.3	Produkcia tepla v rôznych fázach hydratácie	45
5.4	Deštruktívny stav skúšaného betónu	47
5.5	Spriemerované pracovné diagramy z experimentu	48
5.6	Porovnanie výstupov z nenatréovaných sietí s rôznymi počtami ne- urónov s experimentom	50
5.7	Porovnanie výstupov z natréovanej siete 2-3-1 s experimentom	52
5.8	Odlišné miery regularizácie na výstupy z natréovanej siete	52
5.9	Výsledné optimálne výstupy zo siete pre všetky časy	53
5.10	Výsledné optimálne výstupy zo siete pre medzilahlé časy	53
5.11	Výsledná plocha pre všetky časy	54
5.12	Výsledná plocha fuzzy riadenia	55
B.1	Vzorový príklad	61
B.2	Grafické porovnanie	63
C.1	Vplyv náhodnosti počiatočných váh na výsledný výstup	65

Zoznam tabuliek

3.1	Tabuľka vstupných dát	18
3.2	Porovnanie hodnôt M_{max}	19
3.3	Porovnanie hodnôt M_{max} po natrénovaní	20
3.4	Tabuľka trérovacích a testovacích dát	24
3.5	Výstupy zo siete pre trérovacie a testovacie dáta	26
4.1	Tabuľka poskytnutých dát k trérovaniu a testovaniu siete	35
5.1	Vlastnosti skúšaného betónu (Kalafutová, 2007)	47
5.2	Namerané výsledky pre čas $t = 8 \text{ hod}$ (Kalafutová, 2007)	49
5.3	Porovnanie výstupov z natrénovaných sietí s experimentom pre čas $t = 8 \text{ hod}$ (Kalafutová, 2007)	51
A.1	Tabuľka aktivačných funkcií	60
B.1	Porovnanie vstupných a výstupných dát	63

Zoznam symbolov

Tvorba neurónových sietí

X_i ... vstupné dáta

w_{ij} ... hodnota váhy na synapse medzi i-tým neurónom a j-tým neurónom nasledujúcej vrstvy

$S^{(i)}$... suma súčinov signálov a váh v i-tej vrstve

$f(x)$... aktivačná funkcia

$a^{(i)}$... výsledok aplikácie aktivačnej funkcie v i-tej vrstve

y_i ... očakávaný výstup zo siete

\hat{y}_i ... skutočný výstup zo siete

J ... penalizačná funkcia

λ ... regularizačný parameter

Materiálové vlastnosti a zvyšné veličiny

$l [m]$... rozpon nosníka

$q [kN/m]$... spojité zaťaženie

$M_{max} [kNm]$... maximálny ohybový moment na nosníku

$f_{cm} [MPa]$... medzná hodnota pevnosti betónu v tlaku

$E_{cm} [MPa]$... sečný modul pružnosti betónu

$E_{c0} [MPa]$... dotyčnicový modul pružnosti betónu

$\sigma [MPa]$... nãpatie

$\varepsilon_{c1} [-]$... pretvorenie na medzi pevnosti

$\varepsilon_{c0} [-]$... medzné pretvorenie

$f_{ck} [MPa]$... charakteristická pevnosť betónu v tlaku

$f_{cd} [MPa]$... návrhová pevnosť betónu v tlaku

Kapitola 1

Úvod

Vývoj a spracovanie materiálových vlastností je veľmi komplexný proces. Modely neurónových sietí sú v týchto prípadoch extrémne nápomocné a to hlavne vďaka ich schopnosti odpozorovať a naučiť sa určité súvislosti tak, ako sa to deje v mozgu pri procese učenia sa. Ďalšou ich výhodou je rozpoznávanie a dispozícia umožňujúca zatriedenie.

Materiálové modely predstavujú základný nástroj pre popis odozvy skutočných materiálov na rôzne podmienky zaťaženia. Slúžia tiež ako idealizácia skutočnosti a neoddeliteľne patria k inžinierskym výpočtom pre navrhovanie konštrukcií v praxi. Tiež sú dôležitou súčasťou výskumu pre pochopenie a popísanie procesov prebiehajúcich v materiáloch, ktoré sú vystavené mechanickému namáhaniu. Na presnosti a relevantnosti materiálového modelu je postavený úspech výpočtov, v ktorých sú ako základ použité práve hodnoty z daných modelov.

Náplňou tejto práce je vyšetriť vhodnosť použitia umelých neurónových sietí pre vyjadrenie materiálových modelov. Ako hlavná úloha bola zvolená tvorba pracovného diagramu v krátkodobom jednoosom namáhaní betónu veľmi raného veku v závislosti na čase. V práci sú popísané postupy pri tvorbe optimálnej neurónovej siete, ktorá spĺňa dané nároky. Cieľ práce avšak nespočíva v nájdení náhrady dnes používaných analytických metód, ale v preskúmaní možností použitia umelých neurónových sietí pre modelovanie správania betónu ako možnú alternatívu v prípadoch, kedy môže byť toto jednoduchšie alebo relevantnejšie.

Na začiatku práce je uvedený prehľad a postupné vysvetlenie najzákladnejších

vlastností umelých neurónových sietí, a tiež je stručne popísaná inšpirácia v reálnych biologických sieťach. Sú zoskupené dôležité dátumy a poznatky v oblasti histórie vývoju neurónových sietí. Z dôvodu dôležitosti správneho pochopenia všetkých procesov, ktoré v umelých neurónových sieťach prebiehajú, sú okrem celej druhej kapitoly k dispozícii ešte dve prílohy s informáciami, ktoré uvedenú problematiku rozširujú a formou ručného vzorového príkladu znázorňujú.

V tretej kapitole je zvolený príklad, na základe ktorého je možné vytvoriť vlastný model umelej neurónovej siete. Príklad sa musí vyznačovať jednoduchosťou a možnosťou overenia správnosti získaných výstupov. Preto je zvolené stanovenie veľkosti maximálneho momentu na prostom, rovnomerne zaťaženom nosníku. Dostupný vzorec na výpočet M_{max} slúži len na porovnanie výstupov a pre model umelej neurónovej siete nie je známy. Na demonštrovanom príklade sú tiež vysvetlené pojmy overfitting a regularizácia.

Štvrtá kapitola slúži na overenie, že predchádzajúci model siete bol naprogramovaný správne a že svojimi výstupmi dokáže reagovať aj na iné situácie. V danej kapitole je tiež vyskúšaná odozva neurónovej siete na aproximáciu nespojitých funkcií, s ktorými sa je tiež možné pri vedeckej analýze stretnúť.

Jadrom bakalárskej práce je piata kapitola, ktorá obsahuje samotné riešenie stanoveného cieľa práce. Základ pre jej tvorbu a následné vyhodnotenie výsledkov tvoria dáta z experimentov uskutočnených na katedre betónových a zděných konštrukcií Fakulty stavebnej ČVUT (Kalafutová, 2007). Skúšky sa venovali betónom veľmi raného veku v časoch od troch do ôsmich hodín.

Tvorba modelu umelej neurónovej siete a všetky s tým súvisiace výpočty boli uskutočnené pomocou programovacieho jazyku Python. Na tvorbu textu a väčšiny grafov bol použitý textový editor L^AT_EX.

Kapitola 2

Neurónové siete

V nasledujúcej kapitole bude stručne načrtnutý úvod do umelých neurónových sietí a budú spomenuté významné medzníky v histórii danej problematiky. Bude vysvetlený základný proces prebiehajúci pri získavaní výstupov z modelu umelej neurónovej siete. Jedná sa o priebeh signálov sieťou a názov je inšpirovaný angličtinou (tzv. forward propagation). V rámci neho bude vysvetlená aplikácia aktivačnej funkcie a spätné vyhodnotenie získaných výstupov (tzv. backpropagation).

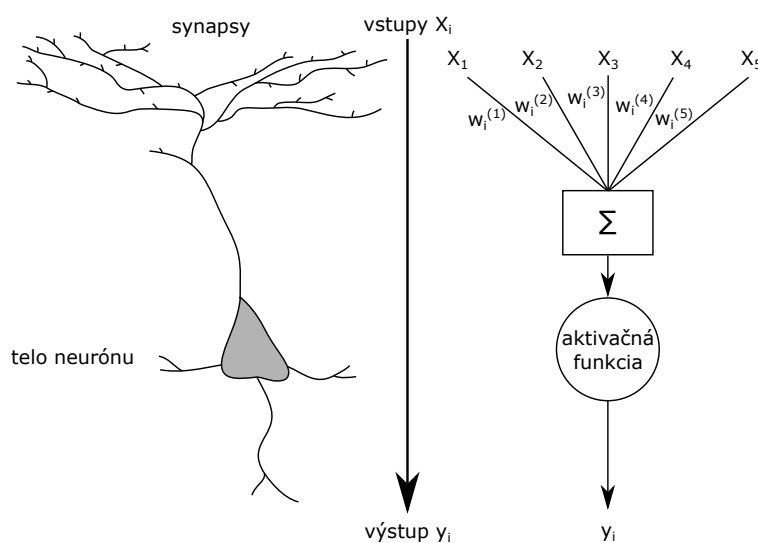
Na rozšírenie kapitoly slúžia prílohy A a B nachádzajúce sa na konci práce po zozname použitej literatúry. Príloha A tabulkovou formou poskytuje prehľad dnes najčastejšie používaných aplikačných funkcií a v prílohe B sa nachádza vzorový ručne vypočítaný príklad forward propagácie.

2.1 Úvod do neurónových sietí

Umelé neurónové siete sú počítačové modely inšpirované biologickými neurónovými sieťami, ktoré sa snažia zachytiť a vyjadriť ich správanie a adaptívne vlastnosti. Biologický nervový systém je tvorený obrovským množstvom nervových buniek (neurónov), v ktorých neustále prebiehajú zložité biologické deje. Vďaka nim sú jednotlivé nervové bunky schopné prijať signály z iných buniek, a tiež k nim vyslať svoje vlastné signály. Komunikáciu medzi bunkami zabezpečujú a sprostredkujú špecializované spoje, tzv. synapsy. Synapsy pôsobia na signály pomocou váh a môžu ich buď zoslabiť, alebo zosilniť (Floreano and Mattiussi, 2008).

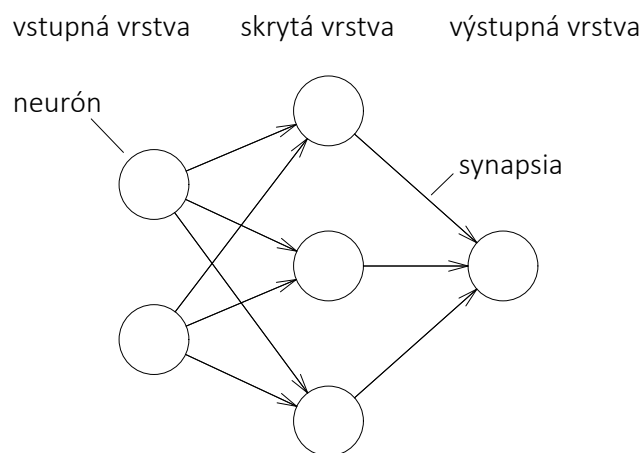
V prípade umelej neurónovej siete sa za základnú stavebnú jednotku považuje „umelý neurón“, ktorý je v porovnaní s biologickým neurónom menej komplexný. Umelé neuróny sú vzájomne zoskupené a prepojené vo vrstvách prostredníctvom synáps, a tým sa stotožňujú s biologickou nervovou sústavou. Zoskupenie vrstiev sa vyznačuje nasledujúcou štruktúrou – prvá vrstva sa nazýva vstupná vrstva, posledná sa označuje termínom výstupná vrstva a medzi nimi sa nachádza určitý počet skrytých vrstiev. Synapsy majú v umelých neurónových sieťach jednoduchú prácu. Použijú hodnotu, ktorá do nich vstupuje, prenášobia ju určitou váhou (váhy sú označované symbolom w) a vygenerujú výsledok. Neuróny zabezpečujú priebeh viacerých operácií. Ich úlohou na začiatku procesu je zhromaždiť všetky signály z neurónov v predchádzajúcich vrstvách, sčítať ich a následne použiť aktivačnú funkciu.

Na obrázku č. 2.1 je schematicky porovnaný biologický neurón (ľavá časť obrázka) s umelým neurónom (pravá časť).



Obr. 2.1: Porovnanie biologického neurónu s umelým
(Floreano and Mattiussi, 2008)

Kvôli názornosti zobrazenia umelých neurónových sietí sú neuróny reprezentované kružnicami a čiary medzi nimi predstavujú synapsy. Na obrázku č. 2.2 je znázornená umelá neurónová sieť s dvoma neurónmi vo vstupnej vrstve, tromi neurónmi v jednej skrytej vrstve a jedným neurónom vo výstupnej vrstve.



Obr. 2.2: Príklad architektúry siete

2.2 História a použitie

Dôležitú úlohu v histórii umelých neurónových sietí zohráva modelovanie samotnej neurónovej bunky. Aby bolo možné nasimulovať biologický nervový systém, je v prvom rade potrebné poznať skutočné správanie neurónov v mozgu.

Prvý model umelej neurónovej siete predstavila v roku 1943 dvojica amerických vedcov W.McCulloch a W.Pitts (McCulloch and Pitts, 1943). Výsledkom ich práce bola prostá neurónová sieť prezentovaná pomocou elektrických obvodov s jednoduchými logickými prepínačmi, ktorá sa však z technologického hľadiska vyznačovala nedostatkami. Keďže hodnoty synaptických váh boli presne dané, sieť sa nebola schopná kontinuálne učiť. Model tiež zlyhával pri riešení nebinárnych problémov. Aj napriek tomu sa vedcom podarilo ukázať, že ich modelom siete sa dajú realizovať všetky procesy, ktoré na svoj opis potrebujú konečný počet symbolických výrazov.

Postupný vývoj umelých neurónových sietí narastal so zlepšovaním výkonnosti počítačov. O hypotetické simulovanie sietí sa medzi prvými neúspešne pokúsil americký vedec N.Rochester. Ďalšie pokusy iných vedcov boli úspešnejšie a znamenali vo vývoji neurónových sietí posun vpred. Rok 1959 možno označiť za prelomový. Americkým vedcom B.Widrovi a M.Hoffovi sa podarilo vytvoriť model prvej neuró-

novej siete, ktorá riešila problém z reálneho sveta. Pomenovali ju ADALINE (ADaptive LInear Elements). Zoskupenie niekoľkých typov uvedenej siete bolo základom vzniku rozmernejšej siete MADALINE (Multiple ADaptive LInear Elements). Sieť ADALINE mala na starosti rozpoznávanie binárnych vzorov. Úlohou siete MADALINE bolo odstraňovanie ozviev pri telefonovaní pomocou adaptívneho filtra. Oba modely siete sa síce klasifikujú ako staršie systémy, ale zdroj (Roberts, 2009) uvádza možnosť ich zaradenia do oblasti komerčnej prevádzky.

Nasledujúce roky so sebou priniesli mnoho zaujímavých pokrokov. Za zmienku spomedzi veľkého množstva stojí napr. práca amerického vedca J.Hopfielda z roku 1982. Ten ako prvý ukázal obojsmerný typ siete. Do tohto roku boli známe a prezentované len jednosmerné spoje medzi neurónmi. V uvedenom roku bol takisto predstavený ešte jeden nový typ siete. Jednalo sa o hybridnú viacvrstvovú sieť, ktorá na riešenie problému používala v každej vrstve odlišnú stratégiu.



Obr. 2.3: John Hopfield (Hopfield, 2007)

Medzi jednu z najväčších výhod neurónových sietí sa radí možnosť učenia. V sieťach sa neučia neuróny, ale práve ich spomenuté spojenia. Pri učení sa kladie dôraz na nastavenie takých hodnôt váh, aby bolo možné napríklad správne aproximovať danú funkciu, alebo predikovať správanie materiálu.

Neurónové siete tvoria v súčasnosti významnú časť počítačovo-orientovanej umelej inteligencie. Predstavujú akýsi univerzálny prístup k štúdiu a napodobňovaniu procesov prebiehajúcich v ľudskom mozgu. Vyznačujú sa širokospektrálnym uplatnením. Medzi odbory ich aplikácie možno zaradiť lingvistiku, riadenie procesov alebo zložitých priemyselných systémov (Kvasnička, 2014).

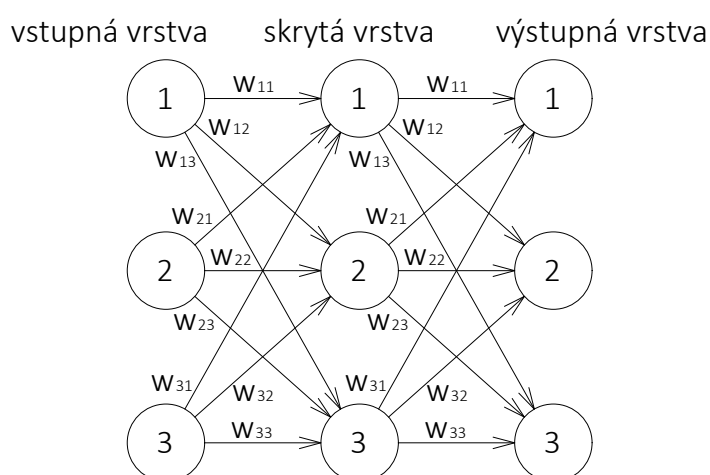
Za veľmi zaujímavú ukážku pracovania neurónových sietí je možné považovať výsledky práce S.Welcha, ktorý vo svojich názorných príručkách vysvetlil základné vlastnosti a pravidlá neurónových sietí. Menované parametre názorne demonštroval na konkrétnom príklade – štúdia bola zameraná na vygenerovanie výsledného počtu bodov v teste. Základnými faktormi, ktoré mali výsledok ovplyvniť bol čas vynaložený na prípravu a zároveň počet hodín spánku pred testom (Welch, 2014).

2.3 Mechanizmus neurónových sietí

V nasledujúcej časti práce budú vysvetlené základné matematické operácie, ktoré sa uskutočňujú v umelých neurónových sieťach. Medzi ne patrí priebeh signálov zo vstupných dát cez skryté vrstvy a aplikácia aktivačnej funkcie pre vygenerovanie výstupu. Ten sa následne vyhodnotí pomocou chýb (rozdiel požadovaného výsledku a výstupu) a budú hľadané vhodné postupy na minimalizáciu chýb (Rashid, 2016).

2.3.1 Priebeh signálov sieťou

Proces bude znázornený na neurónovej sieti s troma vrstvami. Vo všetkých troch z nich sú k dispozícii tri neuróny a každý neurón bude spojený s každým neurónom v nasledujúcej vrstve tak, ako to zobrazuje obrázok č. 2.4.

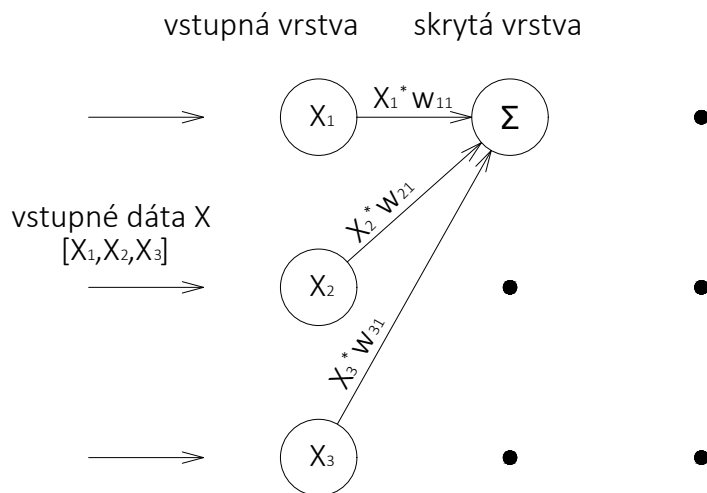


Obr. 2.4: Sieť 3x3

Sila spojenia medzi jednotlivými neurónmi je vyjadrená pomocou váh označovaných ako w_{ij} . Napríklad váha w_{23} informuje, že sa jedná o spojenie medzi druhým neurónom v ľubovoľnej vrstve a tretím neurónom v nasledujúcej vrstve. Analogicky váha w_{12} vyjadruje spojenie medzi prvým neurónom vo vrstve a druhým neurónom v nasledujúcej vrstve. Hodnoty váh sú na začiatku procesu vždy nastavené ako náhodné čísla.

V prvej vrstve neurónov, nazývanej vstupná vrstva, sú zhromažďované vstupné dáta, ktoré na začiatku procesu do siete vstupujú a v prvej vrstve siete budú ukladané. Vstupy budú prezentované maticovým zápisom pomocou matice \mathbf{X} . V tejto vrstve neprebiehajú žiadne výpočty a tiež nebude aplikovaná aktivačná funkcia. V nasledujúcej skrytej vrstve bude na začiatku procesu hodnota jednotlivých vstupných dát X_i vynásobená hodnotou váhy w_{ij} uloženou na konkrétnej synapse a dodatočne budú všetky prenášobené hodnoty zosumarizované v individuálnych neurónoch. Na základe daného pravidla sa v prvom neuróne skrytej vrstve nachádza hodnota prezentovaná vzťahom:

$$X_1 * w_{11} + X_2 * w_{21} + X_3 * w_{31}. \quad (2.1)$$



Obr. 2.5: Násobenie vstupných dát váhami

Podobne sa v druhom a treťom neuróne vyskytujú hodnoty dané výrazmi:

$$\begin{aligned} X_1 * w_{12} + X_2 * w_{22} + X_3 * w_{32}, \\ X_1 * w_{13} + X_2 * w_{23} + X_3 * w_{33}. \end{aligned} \tag{2.2}$$

Zoskupením váh a súm výsledkov v jednotlivých neurónoch do maticovej podoby, možno všetky výpočty potrebné na realizáciu predchádzajúcich dejov vyjadriť pomocou maticového násobenia. Maticové násobenie je dané rovnicou:

$$(\mathbf{A} * \mathbf{B})_{ij} = \sum_{p=1}^n a_{ip} * b_{pj}, \tag{2.3}$$

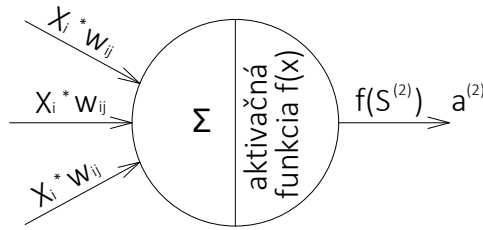
kde \mathbf{A} a \mathbf{B} sú jednotlivé matice, n je počet stĺpcov matice \mathbf{A} a indexy i, j vyjadrujú počty riadkov resp. stĺpcov. Hodnoty váh budú ukladané v matici $\mathbf{W}^{(1)}$. Index 1 definuje, že sa jedná o váhy používané medzi prvou a druhou vrstvou. Sumy pre násobených vstupných dát budú prezentované maticou s označením $\mathbf{S}^{(2)}$, keďže ide o súčty vstupujúce do druhej vrstvy siete. Výsledkom použitia rovnice 2.3 je matematicky vyjadrený priebeh signálov medzi vstupnou a skrytou vrstvou (Welch, 2014):

$$\mathbf{S}^{(2)} = \mathbf{XW}^{(1)} = \begin{bmatrix} X_1 & X_2 & X_3 \end{bmatrix} * \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{bmatrix}. \tag{2.4}$$

2.3.2 Aktivačná funkcia

Na základe pozorovaní v biologickej neurónovej sieti bolo zistené, že neuróny nie vždy reagujú na dané signály. K docieleniu ich reakcie a vyprodukovaní určitého výstupu z neurónu, je nevyhnutné, aby hodnota signálu dosiahla určitú hranicu, odborné nazývanú prahová hodnota. Až po jej prekročení bude signálu umožnené šírenie do ďalších vrstiev (Rashid, 2016).

V umelých neurónových sieťach sa k napodobneniu opísaného správania používa aktivačná funkcia $f(x)$. Pomocou nej je opísané výstupné správanie neurónu. Po vypočítaní váženého súčtu vstupných dát (podrobnejšie rozpracované v časti 2.3.1) sa jednotlivé hodnoty uložené v neurónoch pretransformujú použitím aktivačnej funkcie a do nasledujúcich vrstiev bude šírená zmenená hodnota signálu, viď obrázok č. 2.6.



Obr. 2.6: Aplikácia aktivačnej funkcie

Daný proces je vyjadrený rovnicou:

$$\mathbf{a}^{(2)} = f(\mathbf{S}^{(2)}). \quad (2.5)$$

Ďalšou z možností využitia aktivačnej funkcie je zoskupenie hodnôt signálov do určitého intervalu v závislosti na druhu funkcie a riešeného problému (Floreato and Mattiussi, 2008). V prvej formulácii neurónovej siete mali neuróny len binárny výstup (0 alebo 1) (McCulloch and Pitts, 1943).

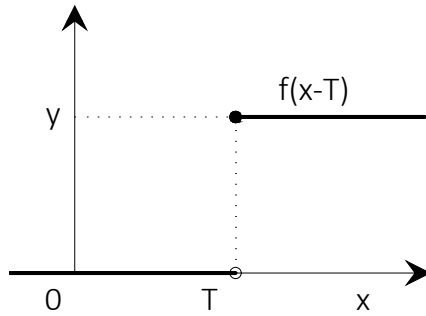
Častým príkladom na vysvetlenie aktivačných funkcií je skoková funkcia, ktorá popisuje správanie reálneho biologického neurónu. Ak je hodnota celkového signálu vstupujúceho do neurónu menšia ako hodnota T (z angl. prahová hodnota = threshold value), potom výstup z neurónu je nulový. Naopak po prekročení hodnoty T neurón vyprodukuje určitý výstup. Skoková funkcia je daná predpisom (Floreato and Mattiussi, 2008):

$$f(x - T) = \begin{cases} 0 & x < T \\ y & x \geq T \end{cases} \quad (2.6)$$

Grafické vyjadrenie:

Z dôvodu potreby tréovania neurónových sietí (vysvetlené v časti 2.3.3) musia byť aktivačné funkcie spojité, aby bolo možné vypočítať ich deriváciu. Skoková funkcia danú podmienku nespĺňa, a preto sa v umelých neurónových sieťach používajú funkcie, ktoré sú diferenciovateľné (Grus, 2015).

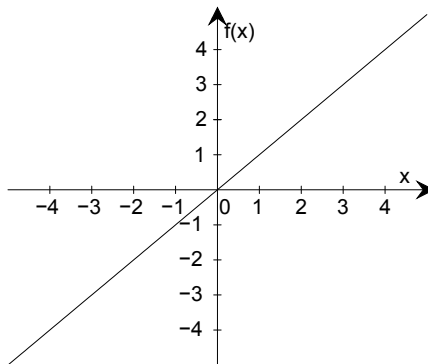
Aktivačné funkcie sa rozdeľujú na dva základné typy, a síce lineárne a nelineárne.



Obr. 2.7: Skoková funkcia

- Lineárna funkcia

Lineárna funkcia, označovaná ako identická aktivačná funkcia, je prezentovaná priamkou a výstup z nej nie je žiadnym spôsobom obmedzený kvôli jej definičnému oboru $(-\infty, \infty)$. Daná funkcia z uvedeného dôvodu neumožňuje riešiť problém v odlišnosti jednotiek vstupných dát. Na obrázku č. 2.8 je znázornený graf funkcie $f(x) = x$.



Obr. 2.8: Lineárna funkcia $f(x)=x$

- Nelineárne funkcie

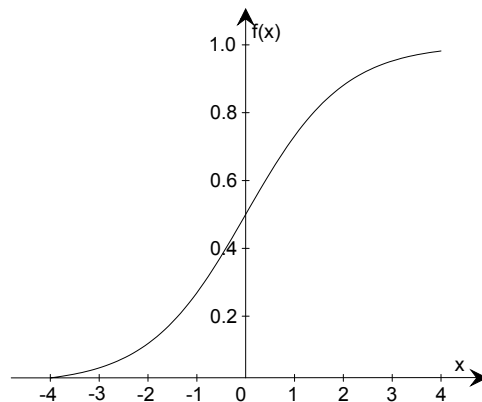
Nelineárne aktivačné funkcie sú používané častejšie v porovnaní s lineárnou funkciou. Umožňujú lepšie prispôsobenie modelu siete odlišnostiam v jednotkách vstupných dát. Základným kritériom pre rozdelenie nelineárnych aktivačných funkcií je in-

terval, v ktorom sa nachádzajú ich funkčné hodnoty. V umelých neurónových sieťach sú známe a využívané viaceré druhy nelineárnych aktivačných funkcií. Medzi často používané sa zaraďuje funkcia sigmoidu alebo funkcia hyperbolického tangensu.

Funkcia sigmoidu sa tiež označuje pojmom logistická funkcia. V počiatkovej fáze rastie približne exponenciálne a jej rast sa následne asymptoticky zastaví. Skupina logistických funkcií sa často využíva v empirických vedách. Sigmoid je daný rovnicou:

$$f(x) = \frac{1}{1 + e^{-x}}. \quad (2.7)$$

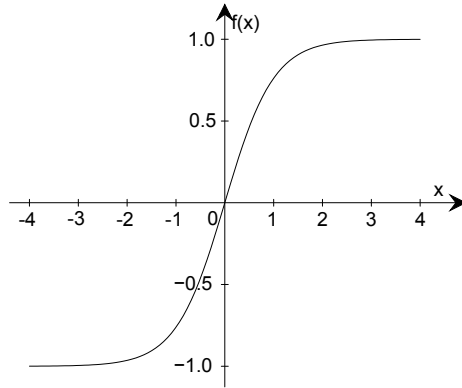
V kladnej časti definičného oboru sa hodnoty funkcie limitne blížia k hodnote 1 a v zápornej časti k hodnote 0. Práve vďaka oboru hodnôt na intervale (0,1) je sigmoid vhodný na tvorbu modelov, kde je cieľom stanoviť pravdepodobnosť ako výstup zo siete.



Obr. 2.9: Funkcia sigmoidu

Funkcia hyperbolického tangensu sa tiež zaraďuje medzi logistické funkcie, ale jej funkčné hodnoty sú oproti sigmoidu v intervale $(-1, 1)$. Graf hyperbolického tangensu je znázornený na obrázku č. 2.10 na nasledujúcej strane.

V prílohe A sú v tabuľke zhromaždené ďalšie druhy používaných aktivačných funkcií. Po aplikácii aktivačnej funkcie v skrytej vrstve neurónovej siete budú zopakované predchádzajúce kroky. Hodnoty prezentované jednotlivými neurónmi budú znovu prenasobené hodnotami váh uložených na spojeniach medzi neurónmi, ďalej budú prenasobené hodnoty sčítané a bude aplikovaná aktivačná funkcia. Tým sú



Obr. 2.10: Graf hyperbolického tangensu $f(x)=\tanh(x)$

vyjadrené výstupy zo siete označované ako \hat{y} . Celý proces je zapísaný rovnicami

$$\mathbf{S}^{(3)} = \mathbf{a}^{(2)}\mathbf{W}^{(2)}, \quad (2.8)$$

$$\hat{y} = f(\mathbf{S}^{(3)}), \quad (2.9)$$

kde $\mathbf{S}^{(3)}$ predstavuje maticu s prenásobenými súčtami vstupujúcimi do tretej vrstvy a $\mathbf{W}^{(2)}$ je matica s uloženými hodnotami váh medzi druhou (skrytou) a tretou (výstupnou) vrstvou. Konkrétny vzorový číselný príklad predošlých popísaných procesov je zobrazený v prílohe B. Priebeh signálov sieťou a aplikácia aktivačnej funkcie sú v anglickej terminológii ako celok označované názvom *forward propagation*.

2.3.3 Vyhodnotenie výstupov zo siete

Po prvom vygenerovaní výstupu zo siete sa hodnoty \hat{y} líšia od požadovaného výsledku y . Odlišnosť je spôsobená faktom, že sieť ešte nebola natrénovaná. Trénovanie siete v sebe zahŕňa proces hľadania takej kombinácie hodnôt váh, aby vzniknutá chyba bola čo najmenšia. Na vyjadrenie celkovej chyby je v teórii umelých neurónových sietí zavedená tzv. penalizačná funkcia J , ktorá je vyjadrená rovnicou:

$$J = \sum \frac{1}{2}(y_i - \hat{y}_i)^2. \quad (2.10)$$

Funkcia J závisí od vstupných dát a počiatočných, náhodne vygenerovaných, hodnôt váh. Jedná sa teda o funkciu dvoch premenných. Minimálnu hodnotu funkcie

viacerých premenných, a tým pádom ideálnu kombináciu hodnôt váh, je možné určiť pomocou parciálnych derivácií. K dispozícii sú dve sady váh $\mathbf{W}^{(1)}$ a $\mathbf{W}^{(2)}$. Preto aj jednotlivé derivácie budú počítané podľa dvoch premenných (konečný tvar úpravy oboch derivácií vyjadrujú rovnice 2.18 a 2.20):

$$\frac{\partial J}{\partial \mathbf{W}^{(1)}} = \frac{\partial \sum \frac{1}{2}(y - \hat{y})^2}{\partial \mathbf{W}^{(1)}}, \quad (2.11)$$

$$\frac{\partial J}{\partial \mathbf{W}^{(2)}} = \frac{\partial \sum \frac{1}{2}(y - \hat{y})^2}{\partial \mathbf{W}^{(2)}}. \quad (2.12)$$

Najprv bude použitím pravidiel derivácie upravená rovnica 2.12:

$$\frac{\partial J}{\partial \mathbf{W}^{(2)}} = \frac{\partial \sum \frac{1}{2}(y - \hat{y})^2}{\partial \mathbf{W}^{(2)}} = \sum \frac{\partial \frac{1}{2}(y - \hat{y})^2}{\partial \mathbf{W}^{(2)}}. \quad (2.13)$$

Pre uľahčenie výpočtu nebude braná do úvahy operácia sumácie. Po celkovom vyjadrení parciálnej derivácie pre jeden prípad sa na konci výpočtu určí suma všetkých derivácií.

$$\frac{\partial J}{\partial \mathbf{W}^{(2)}} = -(y - \hat{y}) \frac{\partial \hat{y}}{\partial \mathbf{W}^{(2)}} \quad (2.14)$$

Dosadenie rovnice 2.9:

$$\frac{\partial J}{\partial \mathbf{W}^{(2)}} = -(y - \hat{y}) \frac{\partial \hat{y}}{\partial \mathbf{S}^{(3)}} \frac{\partial \mathbf{S}^{(3)}}{\partial \mathbf{W}^{(2)}}. \quad (2.15)$$

Pre pokračovanie vo výpočte je potrebné určiť deriváciu aktivačnej funkcie. Bude použitá funkcia sigmoidu vyjadrená rovnicou 2.7 a zobrazená na obrázku č. 2.9.

$$f'(x) = \frac{e^{-x}}{(1 + e^{-x})^2} \quad (2.16)$$

$$\frac{\partial J}{\partial \mathbf{W}^{(2)}} = -(y - \hat{y}) f'(\mathbf{S}^{(3)}) \frac{\partial \mathbf{S}^{(3)}}{\partial \mathbf{W}^{(2)}} \quad (2.17)$$

Dosadenie rovnice 2.8 a konečná úprava:

$$\boxed{\frac{\partial J}{\partial \mathbf{W}^{(2)}} = (\mathbf{a}^{(2)})^T \delta^{(3)}}, \quad (2.18)$$

kde

$$\delta^{(3)} = -(y - \hat{y}) f'(\mathbf{S}^{(3)}). \quad (2.19)$$

Podobným postupom bola upravená tiež rovnica 2.11 na konečný tvar:

$$\boxed{\frac{\partial J}{\partial \mathbf{W}^{(1)}} = \mathbf{X}^T \delta^{(2)}}, \quad (2.20)$$

kde

$$\delta^{(2)} = \delta^{(3)} (\mathbf{W}^{(2)})^T f'(\mathbf{S}^{(2)}). \quad (2.21)$$

Celý algoritmus tréovania umelej neurónovej siete je označovaný ako metóda spätnej propagácie (z angl. *backpropagation*) (Grus, 2015).

Kapitola 3

Tvorba vlastnej neurónovej siete

V nadväzujúcej časti práce budú na názornom príklade ukázané prebiehajúce procesy vysvetlené v kapitole 2. Tiež budú popísané zatiaľ nezmiernené vlastnosti umelých neurónových sietí. K danému účelu bol vybraný jednoduchý, ľahko overiteľný príklad, kvôli potrebe naučiť sa naprogramovať vlastnú umelú neurónovú sieť. Pre vytvorenie kódu bol zvolený programovací jazyk Python.

3.1 Python

Python je označovaný ako programovací jazyk, ktorý je pravdepodobne najľahšie osvojiteľný pre začiatočníkov. Jeho kód je zrozumiteľný pre čítanie aj zápis a tiež sa vyznačuje stručnosťou. V porovnaní s ostatnými jazykmi ako je napr. C++ alebo Java stačí obvykle napísať menej riadkov pre vytvorenie ekvivalentného programu. Jednou z veľkých výhod jazyka Python je, že okrem štandardnej knihovne je k dispozícii veľké množstvo knihovní tretích strán, ktoré poskytujú výkonné možnosti práce s jazykom. Jedná sa napr. o numerickú knihovňu Numpy, alebo grafickú knihovňu Matplotlib, ktorá umožňuje vizualizovať dáta a vytvárať grafy (Summerfield, 2010). Obe spomenuté knihovne boli použité pri tvorbe kódu. Pri písaní kódu bol tiež použitý zdroj (Idris, 2015).

3.2 Maximálny moment na prostom nosníku

Prvým príkladom k overeniu vlastností umelých neurónových sietí bude stanovenie veľkosti maximálneho momentu M_{max} na prostom nosníku zaťaženého spojitým zaťažením q . Rozpon nosíka je označený písmenom l . Podľa vzorca sa veľkosť maximálneho momentu stanoví ako

$$M_{max} = \frac{1}{8}ql^2. \quad (3.1)$$

Pre model umelej neurónovej siete je vzorec pre výpočet neznámy a modelu budú poskytnuté tri dvojice vstupných dát s hodnotami zaťaženia $q[kN/m]$ a rozponu $l[m]$. Po vytvorení umelej neurónovej siete budú porovnávané výsledky pre maximálny moment získané dosadením do vzorca s výstupmi zo siete. Postup pri programovaní siete a pri práci s výsledkami je zhrnutý v nasledujúcom algoritme:

Algorithm 1 Postup pri tvorbe modelu siete

- 1: Zoskupenie a normalizácia vstupných dát
 - 2: Priebeh signálov sieťou
 - 3: Vyhodnotenie výstupu
 - 4: Trénovanie neurónovej siete a vyhodnotenie nového výstupu
 - 5: Overenie prítomnosti overfittingu
 - a) rozdelenie dát na trénovacie a testovacie
 - b) regularizácia
 - 6: Trénovanie upraveného modelu siete
 - 7: Vyhodnotenie výstupu
-

V tabuľke 3.1 sú zhromaždené hodnoty vstupných dát. Pred ich použitím je potrebné dáta upraviť tak, aby s nimi neurónová sieť bola schopná správne pracovať aj napriek odlišnostiam v ich jednotkách. V tomto príklade sa každá hodnota vydolí najväčšou poskytnutou hodnotou a výsledkom bude zoskupenie veľkostí q , l a M_{max} do intervalu $(0, 1)$. Týmto spôsobom nazývaným normalizácia dát bude dosiahnutá správna manipulácia siete s rôznymi jednotkami v dátach.

Tabuľka 3.1: Tabuľka vstupných dát

Vstupné dáta X_i				Normalizácia dát		
	q	l	M	q	l	M
1	3	7	18,375	0,3	0,7	0,147
2	8	5	25	0,8	0,5	0,2
3	10	10	125	1,0	1,0	1,0

Po normalizácii vstupných dát bol napísaný základný zdrojový kód umožňujúci vytvorenie architektúry siete. Pre väčšiu prehľadnosť kódu bolo zvolené objektovo orientované programovanie (OOP), pri ktorom sú časti programu zoskupované do celkov nazývaných objekty. Objekty sú vytvárané pomocou preddefinovaných kritérií (tried) a každý objekt má svoje priradené funkcie, tiež nazývané metódy (Summerfield, 2010).

Triede Neurónová sieť bola na začiatku priradená metóda, pomocou ktorej je možné zadávať počty neurónov v jednotlivých vrstvách. Postupne boli pridané metódy vpred a sigmoid zabezpečujúce priebeh signálov sieťou a aplikáciu aktivačnej funkcie (popísané v častiach 2.3.1 a 2.3.2).

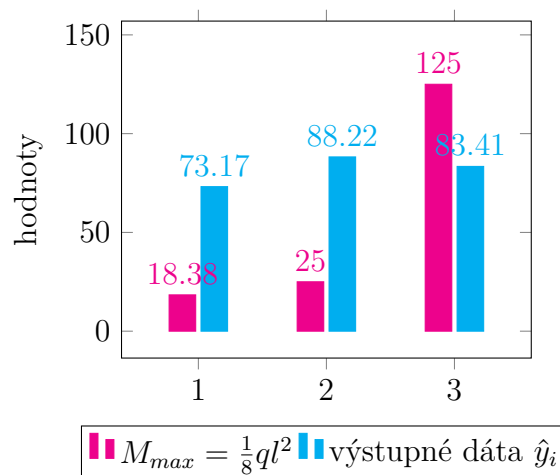
```
# definovanie triedy Neurónová sieť
class Neuronova_Siet(self, vstupna_v, skryta_v, vystupna_v):
    def __init__(self):
        # vzhľad siete
        self.vstupna_v = vstupna_vrstva
        self.skryta_v = skryta_vrstva
        self.vystupna_v = vystupna_vrstva
        # nastavenie váh - náhodné počiatkové hodnoty
        self.W1 = np.random.randn(self.vstupna_v, self.skryta_v)
        self.W2 = np.random.randn(self.skryta_v, self.vystupna_v)
        # priebeh signálov sieťou - maticové násobenie
    def vpred(self, X):
        self.S2 = np.dot(X, self.W1)
        self.a2 = self.sigmoid(self.S2)
        self.S3 = np.dot(self.a2, self.W2)
        vystup_NS = self.sigmoid(self.z3)
        return vystup_NS
    def sigmoid(self, x):
        # aktivačná funkcia sigmoid
        return 1/(1+np.exp(-x))
# vytvorenie modelu siete - zadanie počtu neurónov vo vrstvách
NS = Neuronova_Siet(2,3,1)
# vygenerovanie výstupu zo siete
vystup_NS = NS.vpred(X)
```

Po zadání vstupných dát do siete boli porovnané hodnoty M_{max} získané pomocou vzorca 3.1 s výstupmi zo siete.

Tabuľka 3.2: Porovnanie hodnôt M_{max}

	$M_{max} = \frac{1}{8}ql^2$	Výstupné dáta \hat{y}_i
1	18,375	73,17
2	25	88,22
3	125	83,41

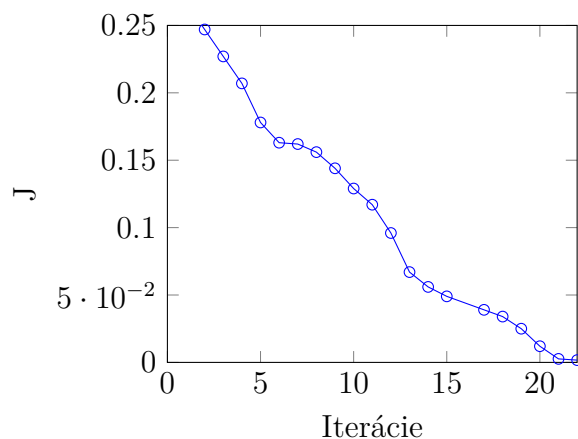
Grafické porovnanie:



Obr. 3.1: Grafické porovnanie hodnôt M_{max}

Za účelom získania presnejších výstupov bola pomocou zmien v kóde umelá neurónová sieť natrénovaná. Triede Neurónová sieť boli pridané ďalšie metódy umožňujúce výpočet penalizačnej funkcie J a derivácie aktivačnej funkcie (vysvetlené v časti 2.3.3). Tiež bola naprogramovaná nová trieda s názvom Tréning s metódami pre hľadanie ideálnej kombinácie hodnôt váh.

K dosiahnutiu takej kombinácie váh, pri ktorých bude penalizačná funkcia J čo najmenšia, bolo po spustení programu potrebných 22 iterácií. Na obrázku č. 3.2 sú graficky znázornené hodnoty J v jednotlivých opakovaníach.

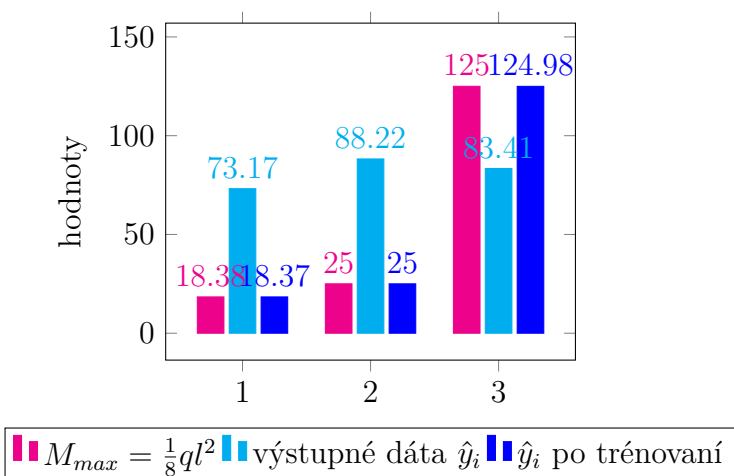


Obr. 3.2: Pribeh tréovania siete

Výsledky boli znovu porovnané pomocou tabuľky aj graficky. Po natrénovaní siete sú výstupy presnejšie.

Tabuľka 3.3: Porovnanie hodnôt M_{max} po natrénovaní

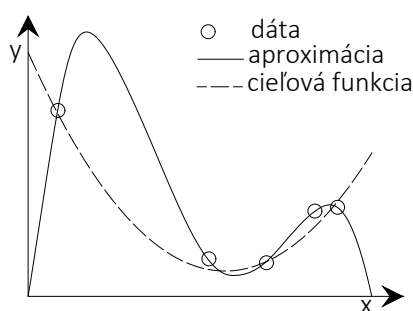
	$M_{max} = \frac{1}{8}ql^2$	Výstupné dáta \hat{y}_i	\hat{y}_i po tréovaní
1	18,375	73,17	18,374
2	25	88,22	25,001
3	125	83,41	124,98



Obr. 3.3: Grafické porovnanie hodnôt M_{max} po tréovaní

Ďalším krokom pri práci na modeli siete bolo overenie prítomnosti overfittingu. Overfitting doslova preložené znamená aproximovať dáta viac než je vhodné. Je to jav, keď schopnosť neurónových sietí extrahovať a reprezentovať závislosti v pozorovaných dátach už nevedie k zlepšovaniu výstupov zo siete, ale môže mať opačné účinky. Nastáva napr. v prípade, keď model učenia siete je zložitejší ako je potrebné k prezentácii cieľovej funkcie, alebo naopak, keď súbor pozorovaných dát obsahuje len funkcie jednoduchšie ako cieľová funkcia (Abu-Mostafa et al., 2012). Hlavnou príčinou prítomnosti overfittingu je snaha o úplne presnú aproximáciu vstupných dát.

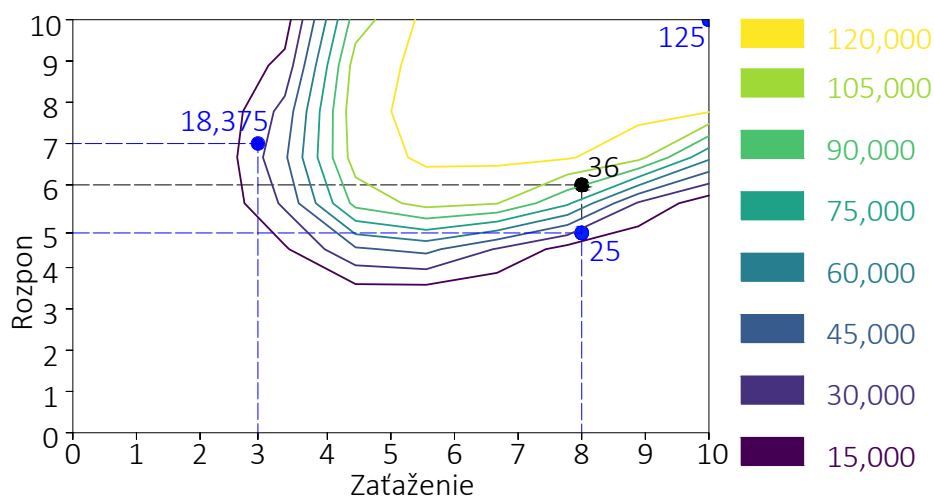
V spomenutom zdroji je tiež uvedený názorný príklad overfittingu. Je uvažovaný jednoduchý jednorozmerný problém regresie (aproximácia funkcie) a k dispozícii sa nachádza 5 bodov (vstupné dáta). Cieľová funkcia pre umelú neurónovú sieť nie je známa, ale obecný model, ktorý najlepšie reprezentuje päť bodov v rovine je polynóm 4.stupňa. V skutočnosti je cieľovou funkciou polynóm 2.stupňa s malým množstvom pridaného šumu v dátových bodoch. Použitím polynómu 4.stupňa je síce dosiahnutá úplne presná aproximácia vstupných dát, pretože graf funkcie prechádza cez všetkých päť bodov, ale výsledok vôbec neodpovedá cieľovej funkcii. Celý proces je znázornený na obrázku č. 3.4.



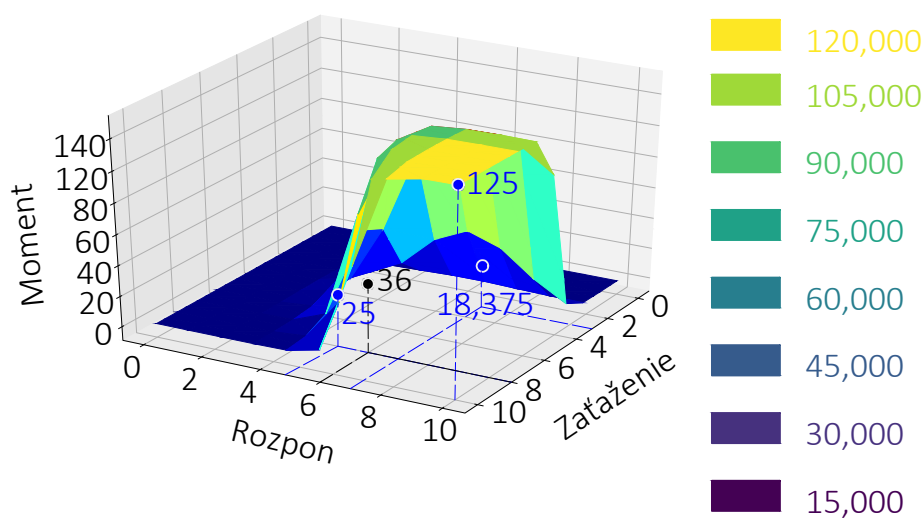
Obr. 3.4: Príklad overfittingu (Abu-Mostafa et al., 2012)

Podobným spôsobom bola overená prítomnosť overfittingu aj vo vzorovom príklade. Po natrénovaní siete výstupy súhlasia s hodnotami maximálneho momentu M_{max} , získaného dosadením do rovnice 3.1, ale správne vytvorený model siete by mal odpovedať aj iným hodnotám pre zaťaženie q a rozpon nosníku l . Preto boli

výstupy vykreslené pre rôzne kombinácie q a l . Zo získaného grafu na obrázku č. 3.5 vyplýva, že sa aj v danom prípade vyskytuje overfitting. Zatiaľ, čo použité vstupné dáta ležia na správnych vrstevniciach (body zobrazené modrou farbou), tak pri ostatných bodoch je už výstup zo siete nepresný. Ako názorný príklad bol čiernou farbou zobrazený bod so súradnicami $q = 8[kN/m]$ a $l = 6[m]$, pre ktorý správna hodnota vychádza $M_{max} = \frac{1}{8}ql^2 = \frac{1}{8} * 8 * 6^2 = 36[kNm]$, ale výstup z neurónovej siete udáva hodnotu približne 91,542[kNm]. Graf bol tiež kvôli lepšej názornosti vykreslený v 3D.



Obr. 3.5: Príklad overfittingu na riešenom prípade 2D



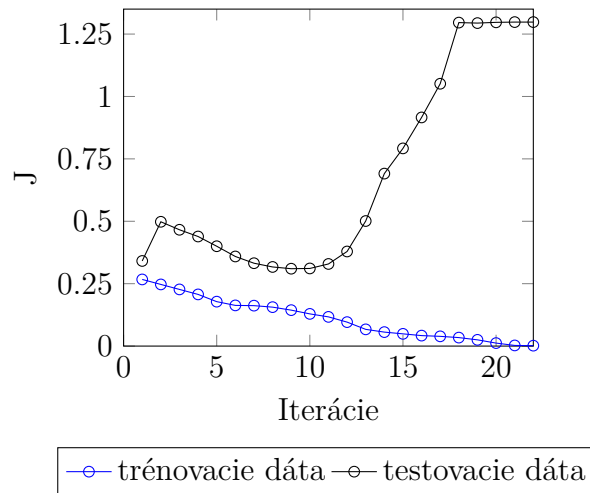
Obr. 3.6: Príklad overfittingu na riešenom prípade 3D

Účinným spôsobom, ako zabrániť overfittingu, je rozdeliť množinu vstupných dát na tréningové a testovacie. Tréningové dáta sú také, v ktorých bude algoritmus pre učenie neurónových sietí nachádzať určité súvislosti a bude sa z nich učiť tak, ako to bolo v prípade pôvodných vstupných dát. Počas učenia modelu siete budú používané len tréningové dáta. Pre overenie kvality naučeného systému a nových výstupných dát sa použijú testovacie dáta, ktoré sa líšia od tréningových. Za dobre naučený model umelej neurónovej siete sa pokladá taký, ktorý bude správne vyhodnocovať tréningové a rovnako aj testovacie dáta (Abu-Mostafa et al., 2012).

Aby sa vstupné dáta mohli rozdeliť na tréningové a testovacie, bolo k nim pridaných ďalších sedem dát so zadanými hodnotami zaťaženia q [kN/m] a rozponu nosníka l [m]. Šesť dát bolo použitých na tréningovanie siete a zvyšné štyri dvojice slúžili na testovanie siete. Všetky hodnoty pridaných nových vstupných dát sú uvedené v tabuľke 3.4. Tiež bol upravený kód metód prislúchajúcich k triede Tréning, ktorým bol okrem predošlého argumentu zadaný takisto argument testovacie dáta. Vďaka zmenám v kóde sa dosiahol správny spôsob tréningovania siete iba na tréningových dátach. Následne bol opäť vykreslený priebeh tréningovania modelu siete. Z uvedeného grafu na obrázku č. 3.7 možno vyčítať, že neurónová sieť stále správne vyhodnocuje iba tréningové dáta. S narastajúcim počtom iterácií potrebných k určaniu optimálnej kombinácie hodnôt váh, sa hodnota penalizačnej funkcie J u tréningových dát znižuje, ale u testovacích začne po 10. iterácii znovu narastať. Tým sa tiež zvyšujú nepresnosti vyhodnotenia testovacích dát.

Tabuľka 3.4: Tabuľka tréningových a testovacích dát

Vstupné dáta X_i				Normalizácia dát			
		q	l	M	q	l	M
Tréningové dáta	1	3	7	18,375	0,3	0,7	0,147
	2	8	5	25	0,8	0,5	0,2
	3	10	10	125	1,0	1,0	1,0
	4	2	2	1	0,2	0,2	0,008
	5	10	2	5	1,0	0,2	0,04
	6	9	4	18	0,9	0,4	0,144
Testovacie dáta	7	5,5	7	33,6875	0,7	0,55	0,27
	8	8	6	36	0,6	0,8	0,288
	9	2	8	16	0,2	0,8	0,128
	10	6	9	60,75	0,6	0,3	0,486

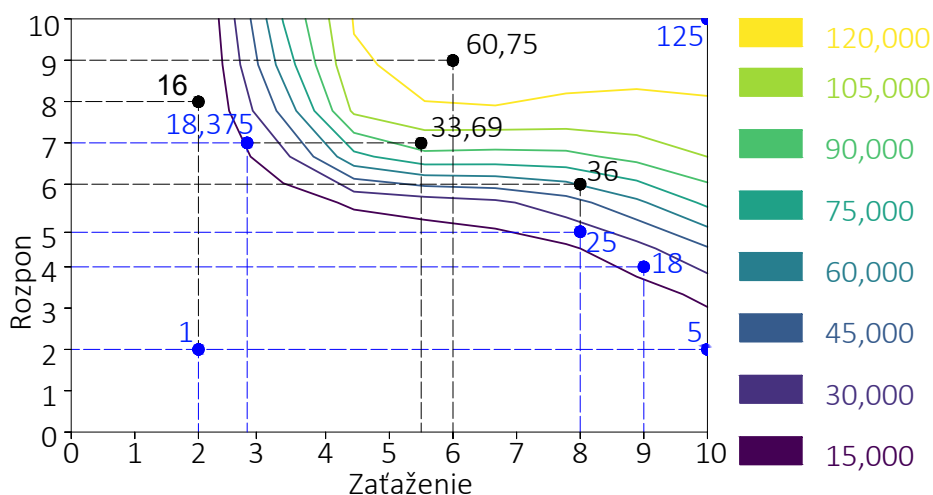


Obr. 3.7: Priebeh tréningovania siete po rozdelení dát

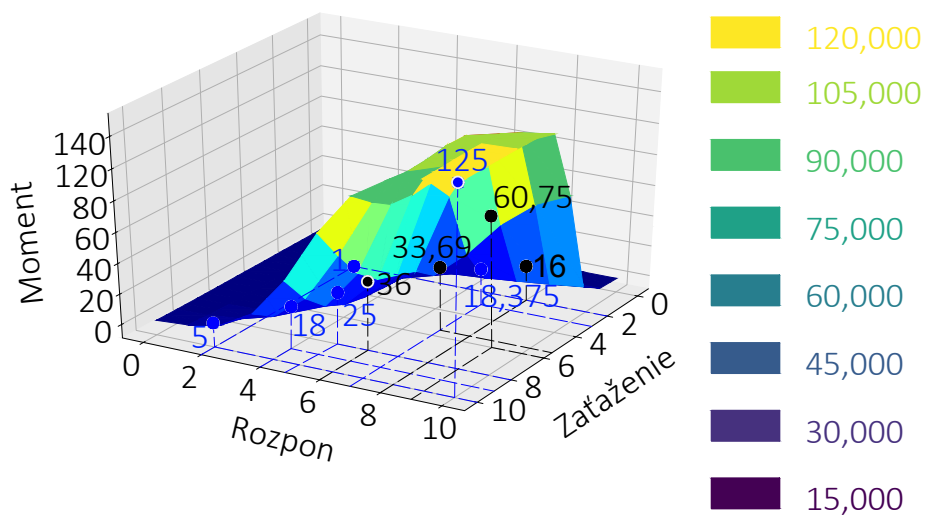
Vykreslenie výstupov zo siete po pridaní a rozdelení vstupných dát je zobrazené v 2D aj 3D na obrázku č. 3.8 a 3.9. Z uvedených grafov tiež vyplýva nesprávne zaobchádzanie modelu so sadou testovacích dát.

Výstupy zo siete boli tiež zhrnuté do tabuľky 3.5. Nesprávne zaobchádzanie modelu siete s testovacími dátami je spôsobené množstvom vstupných dát. Zdroj

(Abu-Mostafa et al., 2012) uvádza, že k dosiahnutiu správnych výstupov z umelej neurónovej siete je potrebné mať k dispozícii minimálne 10-násobok počtu spojení medzi jednotlivými neurónmi v sieti. V riešenom príklade sa pracuje so sieťou, ktorá má 2 vstupné neuróny, 3 neuróny v skrytej vrstve a jeden výstupný neurón (viď zdrojový kód a obrázok č. 2.2). V sieti sa potom nachádza celkovo 9 synapsíí.



Obr. 3.8: Výstupy zo siete po rozdelení dát 2D



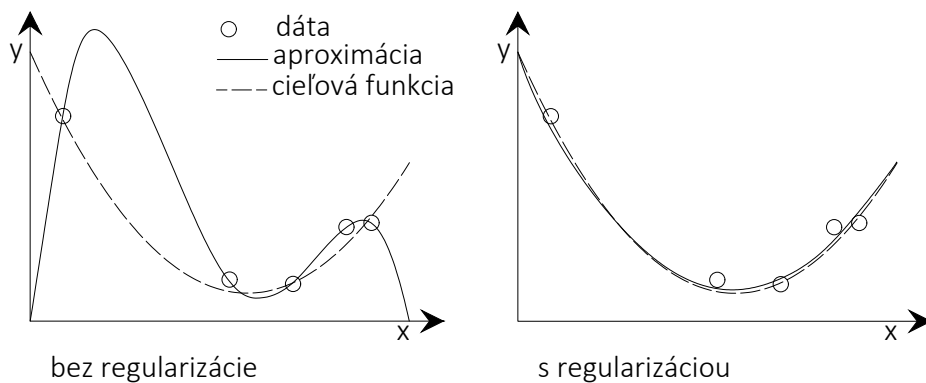
Obr. 3.9: Výstupy zo siete po rozdelení dát 3D

Tabuľka 3.5: Výstupy zo siete pre trérovacie a testovacie dáta

Vstupné dáta X_i				Výstupné dáta \hat{y}_i	
		q	l	$M_{max} = \frac{1}{8}ql^2$	\hat{y}_i
Trérovacie dáta	1	3	7	18,375	18,374
	2	8	5	25	25,001
	3	10	10	125	124,94
	4	2	2	1	0,98
	5	10	2	5	4,97
	6	9	4	18	17,96
Testovacie dáta	7	5,5	7	33,6875	91,28
	8	8	6	36	62,38
	9	2	8	16	9,76
	10	6	9	60,75	123,19

To znamená, že najmenešie potrebné množstvo vstupných dát pre uvedenú neurónovú sieť, je 90 dvojíc dát so zadanými hodnotami zaťaženia $q[kN/m]$ a rozponu nosníka $l[m]$. V stavbárskom odvetví je veľmi často poskytnutých menej výstupov z meraní ako je ideálny počet k natrérovaniu umelej neurónovej siete. Preto bol zavedený ďalší spôsob používaný k odstráneniu overfittingu nazývaný regularizácia (z angl. regularization). Obmedzuje algoritmus učenia siete tak, aby sa chyby zmenšovali a to tiež v prítomnosti šumu. Už aj malá, vhodne zvolená miera regularizácie môže doceliť lepšiu aproximáciu. Na obrázku č. 3.10 je ukázaný vplyv regularizácie na prítomnosť overfittingu popísaného na obrázku č. 3.4.

V rámci regularizácie existuje viacero možných metód vhodných na použitie. Medzi známe patria metóda *Weight Decay*, často vyskytujúca sa metóda *Soft Order Constraint* založená na princípe použitia Legendreových polynómov alebo metóda vychádzajúca z definovania tzv. rozšírenej chyby (z angl. augmented error). Všetky spomenuté metódy sú založené na matematických odvodeniach pre nájdenie optimálnej miery a spôsobu regularizácie.

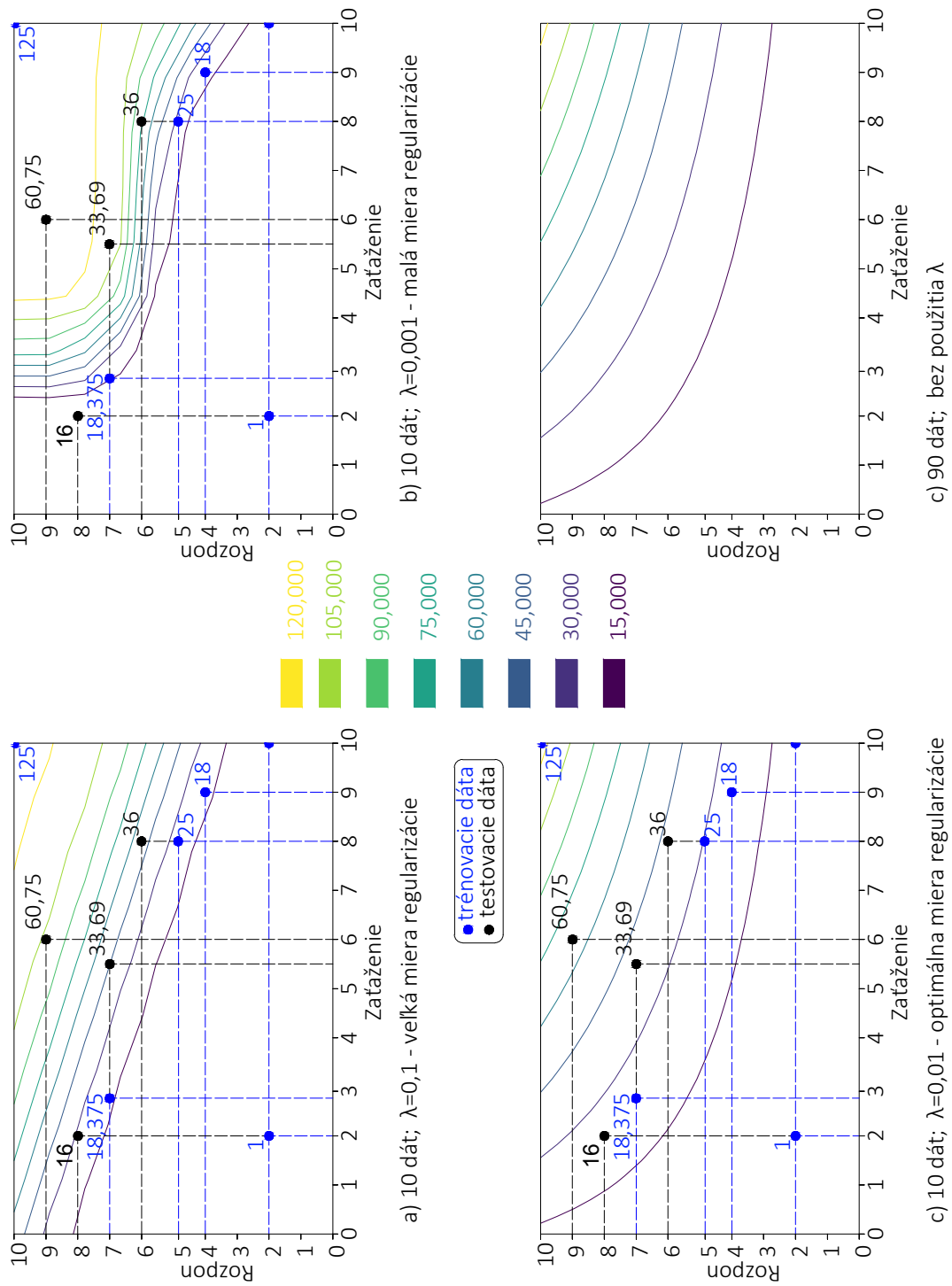


Obr. 3.10: Vplyv regularizácie na overfitting (Abu-Mostafa et al., 2012)

Pri značnom množstve v praxi úspešne používaných techník sa zavedie tzv. regularizačný parameter λ , ktorý penalizuje príliš komplexné modely siete (Abu-Mostafa et al., 2012). Jedná sa o heuristické postupy, ktoré sú využívané pri hľadaní takej hodnoty parametru λ , pre ktorú budú výstupy z modelu neurónovej siete odpovedať reálne očakávaným hodnotám. Mnohokrát sa hodnota stanoví metódou pokus-omyl.

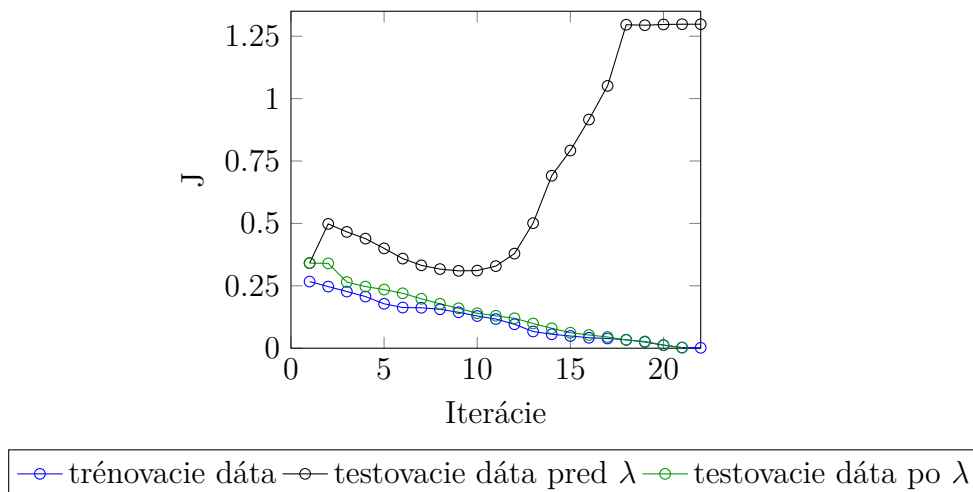
V riešenom príklade boli následne vykreslené 2D grafy reprezentujúce výstupy z modelu siete pre odlišné hodnoty zavedeného parametra λ . Aby sa dosiahlo porovnanie rôznych výstupov na rovnakom modeli siete (s výnimkou hodnoty λ), bol použitý príkaz `np.random.seed(x)`, ktorý zabezpečuje rovnaké vygenerovanie počiatkových náhodných váh pri každom novom spustení kódu. V danom príklade bol použitý `np.random.seed(1)`. Výsledky porovnania sú zobrazené na obrázku č. 3.11 a odkazujú na to, že pri použití optimálnej miery regularizácie bol dosiahnutý rovnaký výstup zo siete ako pri použití 90 vstupných dát. Výhodou použitia regularizačného parametra je teda dosiahnutie obmedzenia počtu potrebných vstupných dát k vytvoreniu správneho modelu umelej neurónovej siete a získaniu odpovedajúcich výstupov.

V prílohe C sú graficky porovnané výstupy zo siete pre rôzne hodnoty `np.random.seed(x)`.



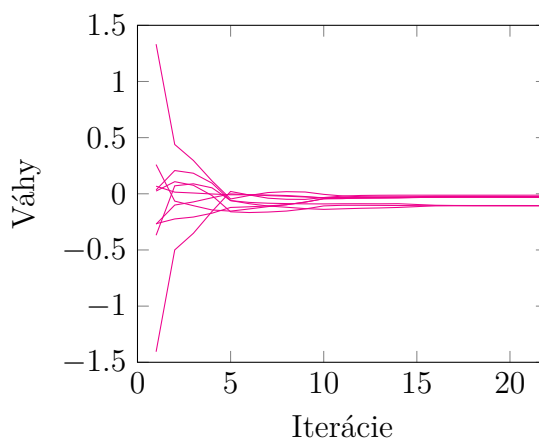
Obr. 3.11: Vplyv regularizácie pre odlišné λ

Použitím vhodnej miery regularizácie bol tiež vyriešený problém nesprávneho zaobchádzania siete s testovacími dátami zobrazený na obrázku č. 3.7. Po opätovnom vykreslení priebehu tréovania sa hodnota penalizačnej funkcie J zmenšovala u oboch skupín vstupných dát.



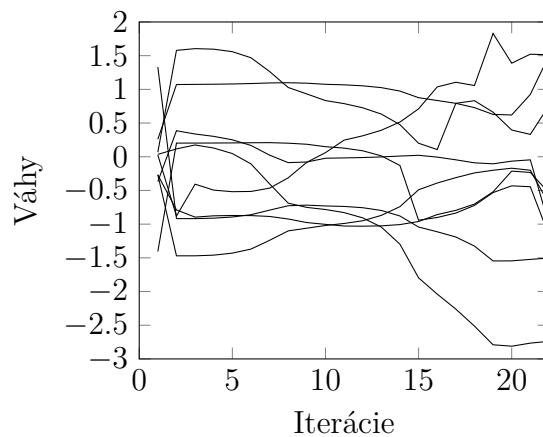
Obr. 3.12: Priebeh tréovania siete po zavedení λ

Vplyv λ na priebeh tréovania možno vidieť aj po vykreslení hodnôt váh počas tréovania. V prvej iterácii sú hodnoty individuálnych váh nastavené ako náhodné čísla, ale použitím príkazu `np.random.seed(x)` sú vygenerované hodnoty zhodné pre oba modely siete. V prípade použitia λ sa počiatočné hodnoty váh postupne zmenšujú a na konci procesu sú buď rovné, alebo blízke nule.



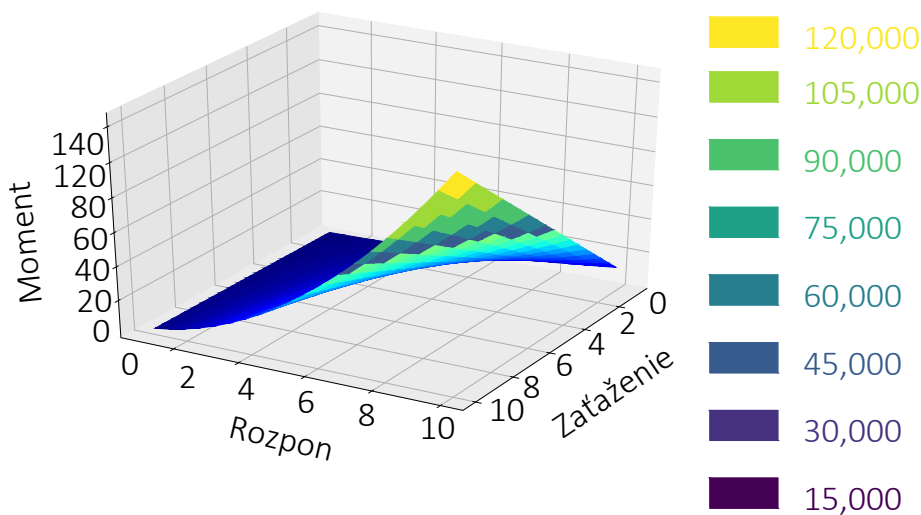
Obr. 3.13: Zmena váh počas tréovania s použitím λ

Práve vďaka postupne znižujúcim sa veľkostiam váh sa zredukoval počet vstupných dát nutných k vytvoreniu správneho modelu siete. Na nasledujúcom obrázku sú vykreslené zmeny hodnôt váh počas tréningu siete bez použitia regularizačného parametru λ . V danom prípade sa hodnoty jednotlivých váh zväčšovali a preto zostalo množstvo potrebných vstupných dát nezmenené.



Obr. 3.14: Zmena váh počas tréningu bez použitia λ

Následne bol graf dosiahnutý použitím optimálnej miery regularizácie vykreslený aj v 3D.



Obr. 3.15: Konečné výstupné dáta 3D

Kapitola 4

Zloženie betónovej zmesi

V nasledujúcej kapitole bol na overenie výstupov z vytvoreného modelu siete vybraný príklad, pri ktorom už nie je k dispozícii vzorec na overenie správnosti výsledkov. V celej časti práce boli informácie čerpané prevažne zo zdroja (Collepari, 2006). V úvodnej časti knihy sú vysvetlené základné definície a termíny, ktoré sa týkajú betónu a železobetónu. Jedná sa napríklad o suroviny potrebné k výrobe betónu, alebo popis postupu výroby. V ďalších častiach sa píše o vlastnostiach betónu ako sú spracovateľnosť, pórovitosť, či trvanlivosť. Záverečné kapitoly sa venujú návrhu betónovej zmesi. Analyzovaný pojem znamená vypracovanie zloženia betónovej zmesi na základe požadovaných konečných vlastností betónu. Návrh je založený na určitých experimentálnych vzťahoch medzi zložením betónu na jednej strane a vlastnosťami vstupujúcich surovín na strane druhej.

4.1 Chemické prísady

Základné zložky betónu tvoria cement, hrubé a drobné kamenivo, voda, prímеси a prísady. Chemické prísady sa v relatívne malých množstvách pridávajú tesne pred, alebo v priebehu miešania betónu. Obvykle sú to kvapalné látky (na rozdiel od minerálnych prímеси v práškovej forme) a umožňujú zlepšiť vlastnosti čerstvého aj stvrdnutého betónu. Množstvo pridanej chemickej látky sa spravidla udáva ako percento z hmotnosti cementu. Podľa špecifickej funkcie, ktorú plnia chemické prísady po pridaní do betónovej zmesi, môžu byť rozdelené do viacerých kategórií, napr.

urýchľujúce, spomaľujúce, prevzdušňujúce, hydrofobizačné prísady, prísady znižujúce obsah vody, alebo polymérne látky superplastifikátory.

4.2 Superplastifikátory

Superplastifikátory sú označované ako najdôležitejšie chemické prísady, ktoré silne znižujú spotrebu zámesovej vody potrebnej pri zachovaní rovnakej spracovateľnosti, ako u kontrolného betónu. Za jeden z najvýznamnejších pokrokov v rámci technológii betónu za posledných 30 rokov sa považuje práve použitie superplastifikátorov.

Aby sa použité prísady mohli označovať uvedeným pojmom, uvádza európska norma *EN 943-2* kritériá, ktoré prísady musia splniť:

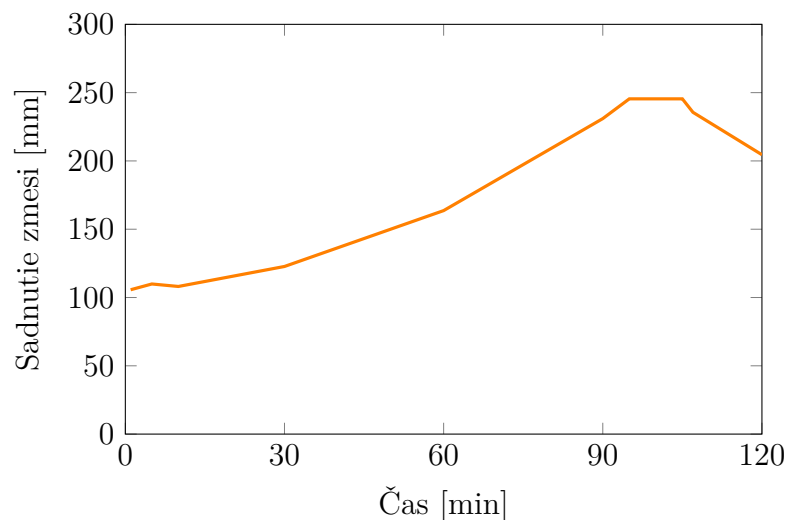
- sadnutie minimálne 120 mm oproti kontrolnému betónu so sadnutím 30 mm,
- sadnutie superplastifikovanej zmesi po 30 minútach (minimálne 30 mm) musí byť rovnaké ako sadnutie kontrolnej zmesi hneď po rozmiešaní,
- obsah vzduchu v superplastifikovanej zmesi má obsah vzduchu v kontrolnom betóne prevyšovať maximálne o 2%.

Vymenované požiadavky bezpečne spĺňajú všetky komerčne vyrábané superplastifikátory.

Mechanizmus pôsobenia superplastifikátorov je založený na dispergácii. V cementovej paste betónovej zmesi sa vyskytujú väčšie aglomeráty cementových častíc. Sú to častice vo forme hrudiek dosahujúce veľkosť až niekoľko milimetrov. Pôsobením superplastifikátorov sa rozložia a tým sa dosiahne zvýšená tekutosť cementovej pasty a teda aj betónovej zmesi. Popísaný proces sa označuje pojmom *dispergačný efekt*.

Pre inšpiráciu k ďalšiemu pracovaniu s modelom umelej neurónovej siete bol z použitého zdroja (Collepari, 2006) vybraný graf znázorňujúci schému vývoja straty spracovateľnosti s rôznymi superplastifikátormi na báze polyakrylátov. Ako konkrétny plastifikátor bola zvolená prísada označovaná ako SLCA (slump-loss controlling agent; modifikovaný superplastifikátor). V grafe na obrázku č. 4.1 je schematicky

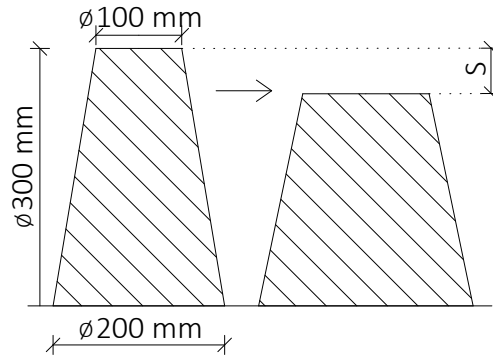
znázornené správanie čerstvého betónu v prítomnosti superplastifikátoru na bázi SLCA.



Obr. 4.1: Schéma vývoja straty spracovateľnosti betónu so superplastifikátorom SLCA (Collepari, 2006)

4.3 Spracovateľnosť betónovej zmesi

Jednou zo základných vlastností betónovej zmesi je spracovateľnosť a vyjadruje jej schopnosť pohybovať sa a byť zhutňovaná. Spracovateľnosť súvisí s konzistenciou betónu a jednou zo skúšok určených pre meranie konzistencie je skúška sadnutím kužela. Skúška popisuje, ako veľmi čerstvý betón tečie a meria sa pri nej pokles výšky betónového vzorku naplneného do formy tvaru kužela nazývaného Abramsov kužel. Spodná podstava kužela má priemer 200 mm, horná podstava 100 mm a celková výška je 300 mm. Na základe odmeranej veľkosti sadnutia s presnosťou na 10 mm sa určí trieda konzistencie betónovej zmesi, ktorá nadobúda zatriedenie medzi S1-tuhá až S5-tekutá (ČSN, 2009). Na nasledujúcom obrázku je schematicky znázornený priebeh skúšky sadnutia kužela.



Obr. 4.2: Skúška sadnutia kužela

Pre účely pracovania s modelom umelej neurónovej siete bol graf na obrázku č. 4.1 upravený nasledovne:

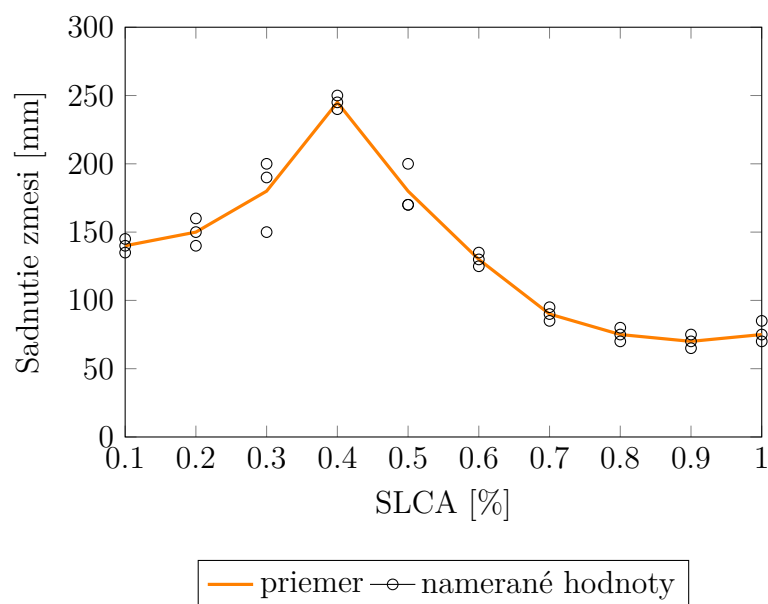
- schéma vývoja spracovateľnosti bude vykreslená pre jeden konkrétny čas,
- na vodorovnej osi sa miesto jednotlivých časov bude nachádzať množstvo použitého plastifikátoru SLCA vyjadreného ako % z hmotnosti cementu (rozmedzie od 0,1% do 1% po kroku 0,1%),
- pre každé množstvo SLCA sú k dispozícii 3 namerané hodnoty veľkosti sadnutia. Kvôli odchýlkam v meraní sa poskytnuté hodnoty od seba líšia.
- Pre trénovanie a testovanie siete budú používané všetky namerané dáta náhodne rozdelené na trénovacie a testovacie v pomere 6:4. Úlohou neurónovej siete tak bude na základne zadaných vstupných dát určiť závislosť veľkosti sadnutia betónovej zmesi na množstve použitého plastifikátoru SLCA. Modifikovaný graf bude pre neurónovú sieť neznámy a bude slúžiť na overenie správnosti výstupov zo siete.
- Jednotky na zvislej osi zostanú nezmenené,
- tvar závislosti bude upravený na zložitejšiu funkciu.

Po uskutočnení vymenovaných úprav už graf **nezodpovedá realite**, ale slúži len k testovaniu vlastností vytvoreného modelu umelej neurónovej siete.

V tabulke 4.1 sú zoskupené dáta, ktoré budú ďalej použité k trénovaniu a testovaniu siete, tiež bol vykreslený upravený graf.

Tabuľka 4.1: Tabuľka poskytnutých dát k trénovaniu a testovaniu siete

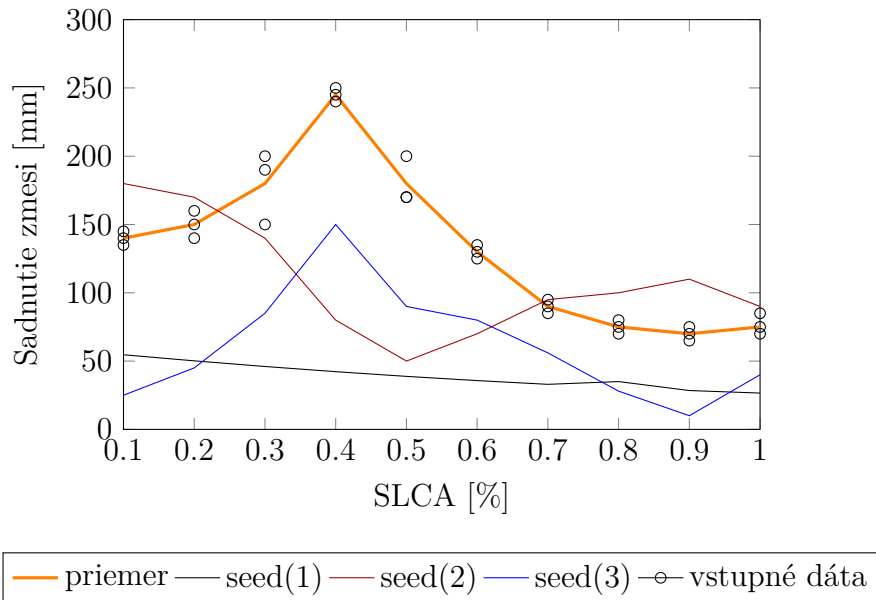
%SLCA	Veľkosť sadnutia zmesi [mm]			
	č.1	č.2	č.3	priemer
0,1	135	145	140	140
0,2	160	150	140	150
0,3	190	200	150	180
0,4	250	245	240	245
0,5	170	170	200	180
0,6	135	130	125	130
0,7	90	95	85	90
0,8	75	70	80	75
0,9	75	65	70	70
1,0	70	85	75	75



Obr. 4.3: Závislosť spracovateľnosti zmesi na množstve plastifikátora SLCA

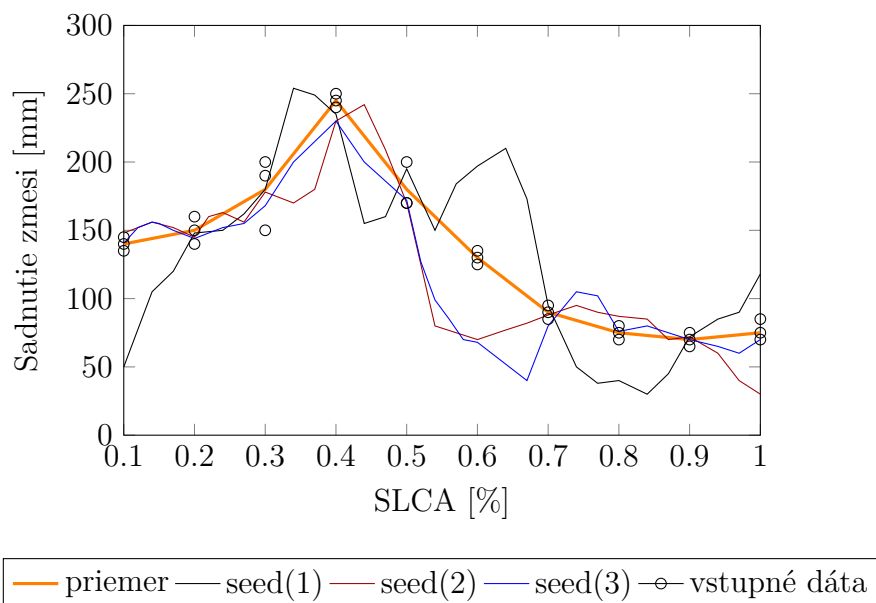
Po upravení modelu sa s umelou neurónovou sieťou pracovalo podobne ako v kapitole 3. Najprv bola vytvorená architektúra siete. Keďže ako vstupné dáta sú k dispozícii len množstvá použitého plastifikátora vzhľadom k hmotnosti cementu, postačí vo vstupnej vrstve siete jeden neurón. Takisto sa jeden umelý neurón nachádza vo výstupnej vrstve a konečne v skrytej vrstve boli navrhnuté tri neuróny. Celkovo sa teda v umelej sieti nachádza $2 * 3 = 6$ synapsí a k ideálnemu natrénovaniu siete by bolo potrebných $10 * 6 = 60$ vstupných dát (Abu-Mostafa et al., 2012).

Po vytvorení modelu boli vykreslené výstupné dáta z nenatrénovanej siete pre rôzne hodnoty príkazu *np.random.seed(x)*. Za argument x boli postupne dosadené hodnoty 1, 2 a 3. Význam príkazu bol vysvetlený v kapitole 3 na strane 27 a jeho použitím sa dosiahne rovnaké počiatočné nastavenie váh pri opakovanom spustení programu.



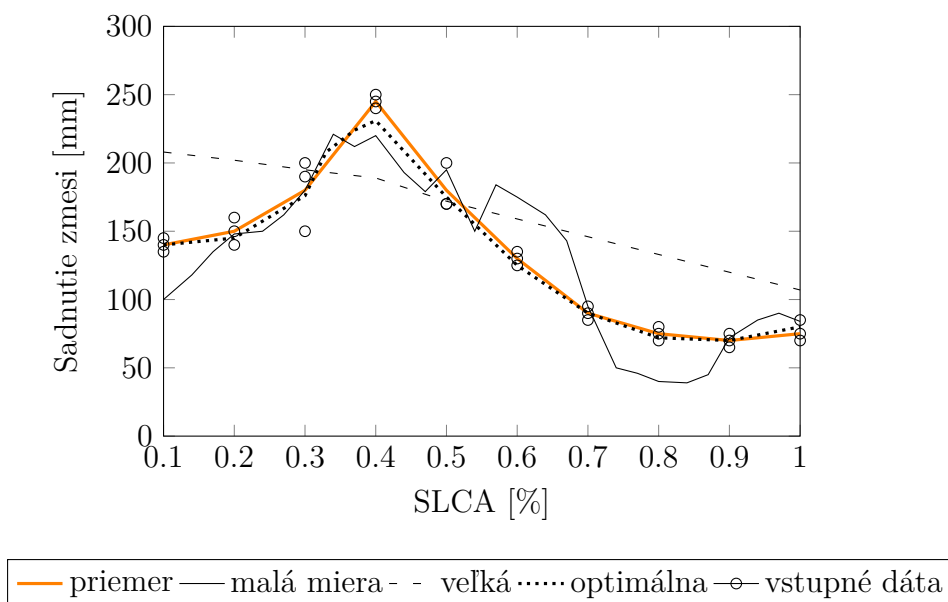
Obr. 4.4: Výstupy zo siete pred trénovaním pre rôzne seedy

Ďalej bola sieť natrénovaná. Množina vstupných dát bola opäť rozdelená na podmnožinu trénovacích a testovacích dát. Po vykreslení výstupov sa sice odhady zo siete zlepšili ale aj v danom príklade je zrejmy overfitting vyplývajúci z nedostatočného množstva vstupných dát. Sieť stále správne vyhodnocovala iba poskytnuté dáta ale na zvyšné dáta reagovala nesprávne. Problém je možné riešiť zavedením regularizačného parametra λ .



Obr. 4.5: Výstupy zo siete po trénovaní pre rôzne seedy

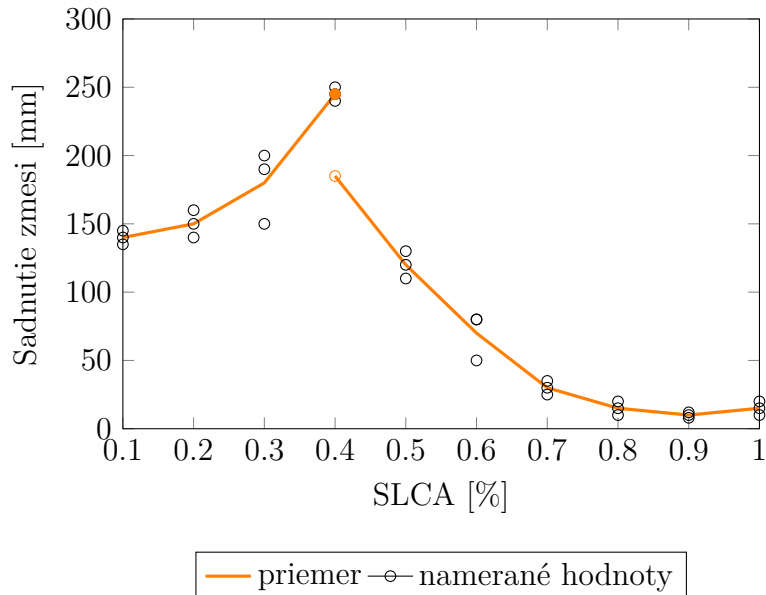
Následne boli vykreslené výstupy zo siete po zavedení regularizačného parametra λ . Bola porovnaná malá, veľká a optimálna miera regularizácie. Porovnanie sa pre-
 viedlo na prípade, keď bol použitý príkaz `np.random.seed(x)` s argumentom 1 (pred-
 chádzajúci obrázok - čierna farba).



Obr. 4.6: Odlišné miery regularizácie pre `np.random.seed(1)`

4.4 Aproximácia nespojitej funkcie

Po úspešnom natrénovaní modelu siete na upravenom grafe zobrazenom na obrázku č. 4.3, bol graf upravený ešte raz. Posunutím druhej časti závislosti pozdĺž zvislej osi vznikla nespojitá funkcia.



Obr. 4.7: Závislosť spracovateľnosti zmesi na množstve plastifikátora SLCA - upravené

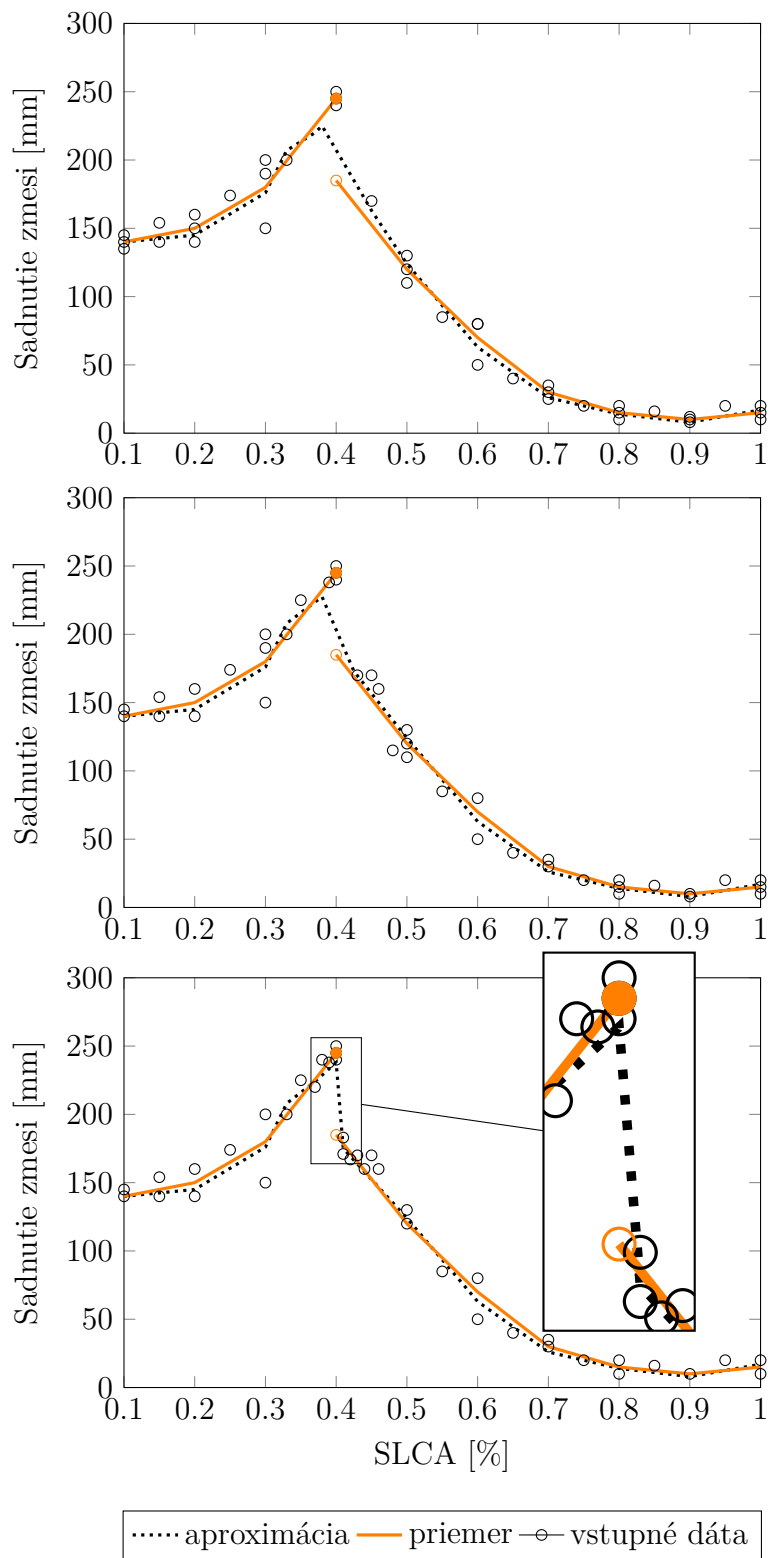
Podobne ako v predchádzajúcej časti práce, aj v danom prípade spočívala úloha neurónovej siete v priblížení sa k cieľovej funkcii. Problém aproximácie diskontinuálnych funkcií neurónovou sieťou je často potrebný tiež v praktických aplikáciách ako sú napr. spracovanie biologických informácií, alebo tvorba systémov riadenia pohybu (Llanas et al., 2008).

Pri pokusoch o aproximáciu nespojitých funkcií vyžaduje model umelej neurónovej siete vyšší počet neurónov v skrytých vrstvách a tiež väčší počet výcvikových iterácií. Aj napriek tomu je možné dosiahnuť len určitú mieru presnosti - výstupná funkcia bude spojitá (Selmic and Lewis, 2002).

Preto boli modelu umelej neurónovej siete pridané ďalšie neuróny a tiež jedna skrytá vrstva. Celkovo sa tak v sieti nachádzal jeden neurón vo vstupnej vrstve, štyri neuróny v oboch skrytých vrstvách a jeden neurón vo výstupnej vrstve. Každý

neurón bol prepojený s každým neurónom v nasledujúcej vrstve siete model siete preto obsahuje spolu 35 synapsií. Pri nespojitých funkciách je všeobecne dôležité rozloženie tréningových a testovacích dát, ktoré musia byť na jednej strane rozložené približne rovnomerne, aby bola neurónová sieť schopná aproximovať každú časť funkcie. Daná požiadavka sa zhoduje s predpokladmi pri predchádzajúcich úlohách. Na strane druhej bude aproximácia tým presnejšia, čím viac vstupných dát sa bude nachádzať práve v okolí skoku.

Na nasledujúcom obrázku je znázornený vplyv rozloženia tréningových a testovacích dát na výsledný aproximovaný tvar funkcie. Pre každú znázornenú variantu bolo použitých spolu 40 vstupných dát rozdelených na tréningové a testovacie dáta v pomere 6:4. V každom prípade je vykreslené rozdielne rozloženie vstupných dát a finálna podoba výstupu z umelej neurónovej siete, ktorá už bola natrénovaná a bol pre ňu tiež nájdený optimálny regularizačný parameter umožňujúci redukciu počtu vstupných dát a zamedzenie overfittingu. Výstupy zo siete pred tréningom a po tréningu neboli znázornené, pretože priebeh tréningu siete bol podobný s predchádzajúcimi úlohami.



Obr. 4.8: Vplyv rozloženia vstupných dát na výstupy zo siete

Kapitola 5

Pevnosť betónu raného veku

Nasledujúca kapitola bude zameraná na vytvorenie konkrétneho materiálového modelu využitím neurónových sietí. Ako hlavná problematika bola zvolená pevnosť betónu raného veku. Cieľom je nájdenie vhodného modelu umelej neurónovej siete a jeho využitie k vývoju pracovného diagramu betónu počas tuhnutia na základe reálnych nameraných dát.

Za účelom uvedenia do danej problematiky bude na začiatku kapitoly najprv stručne objasnený pojem pracovný diagram a budú vysvetlené deje prebiehajúce pri tuhnutí betónu. Následne bude popísaný priebeh skúšok, pri ktorých boli získané vstupné dáta a ďalej tvorba modelu umelej neurónovej siete. Na konci kapitoly budú zhrnuté získané výsledky.

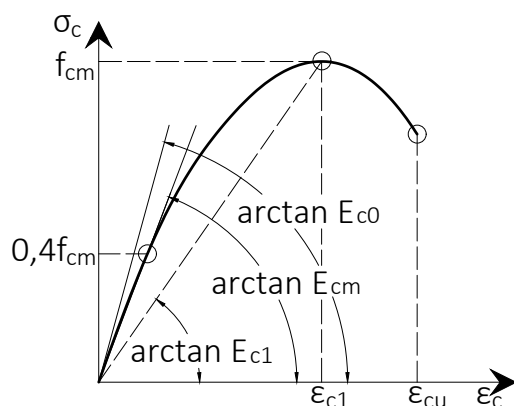
5.1 Pracovný diagram betónu

Pracovný diagram graficky znázorňuje závislosť napätia σ na pomernom pretvorení ε . Je to základný nástroj, pomocou ktorého je možné popísať mechanické správanie materiálu a zároveň ním idealizovať skutočnosť. Model pracovného diagramu je možné použiť nie len na inžinierske výpočty pre navrhovanie konštrukcií v praxi, ale tiež na výskumné účely. Preto hrá veľmi dôležitú, často rozhodujúcu, úlohu práve správna voľba modelu pracovného diagramu.

Charakteristickým znakom pracovného diagramu betónu je jeho výrazná nelinearita. K jeho popisu boli v minulosti vyvinuté rady rôznych analytických modelov,

umožňujúcich jednoduchšiu prácu s častokrát zložitými fyzikálnymi dejmi. Použitelnosť odvodených modelov závisí na ich aplikácii. Medzi nenáročné aplikácie sa zaraďuje napr. výpočet medznej únosnosti železobetónového prierezu a vystačia mu jednoduchšie modely. So zložitou aplikácií stúpajú aj nároky na použitý druh materiálového modelu. Je pre nich potrebné mať k dispozícii relevantný model, ktorý dokáže vystihnúť všetky dôležité charakteristiky správania materiálu.

Na obrázku č. 5.1 je zobrazený charakteristický tvar pracovného diagramu betónu v tlaku pre analýzu konštrukcie.

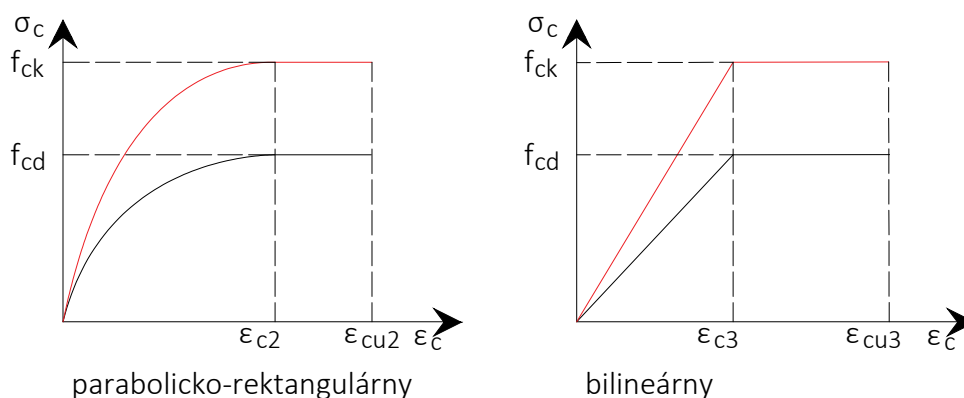


Obr. 5.1: Charakteristický tvar pracovného diagramu betónu v tlaku (Hana Hanzlová, 2015)

Význam jednotlivých vyznačených veličín:

- f_{cm} - medzná hodnota pevnosti betónu v tlaku v MPa,
- $0,4f_{cm}$ - do napätí, ktoré odpovedá približne 40% pevnosti v tlaku, nevznikajú v betóne žiadne trhliny. Preto možno v uvedenom rozmedzí označiť správanie sa betónu ako ideálne elastické a možno ho nahradiť pomocou sečného modulu pružnosti E_{cm} .
- E_{c0} - dotyčnicový modul pružnosti. Charakterizuje krivku len v bode, pre ktorý je hodnota napätia σ nulová.
- ϵ_{c1} - pretvorenie na medzi pevnosti ,
- ϵ_{cu} - medzné pretvorenie .

Medzi spomenuté modely, ktoré aproximujú skutočné správanie betónu pri jednoosom krátkodobom namáhaní len schematicky patria napr. parabolicko-rektangulárny alebo bilineárny návrhový diagram. Používajú sa pri posúdení medznej únosnosti železobetónového prierezu a sú zobrazené na nasledujúcom obrázku. Červené krivky označujú charakteristickú pevnosť betónu v tlaku f_{ck} a čierne návrhovú f_{cd} .



Obr. 5.2: Zjednodušené pracovné diagramy (Hana Hanzlová, 2015)

5.2 Problematika betónu raného veku

Betón sa kvôli svojim vlastnostiam zaraďuje medzi najpoužívanejšie stavebné materiály. Veľkú výhodu má vďaka univerzálnosti jeho použitia, dobrej únosnosti v tlaku, nehorľavosti alebo tiež cene. Na strane druhej má nevýhodu v tom, že pri monolitických stavbách je ukladaný ako tekutá zmes a až postupom času sa z nej stáva tuhý materiál s požadovanou pevnosťou. Preto je nutné po uložení zmesi pri výstavbe zahrnúť technologickú pauzu, aby sa štruktúra betónu mohla vyvinúť do fázy, v ktorej už je možné betón zaťažiť.

V súčasnej dobe je ale rýchlosť výstavby veľmi dôležitým faktorom a to nie len z finančného hľadiska. Jednou z možností, ako vyhovieť požiadavke skrátenia času výstavby je práve skrátenie technologických prestávok, ktoré vznikajú medzi jednotlivými betonážami, odbedňovaním a následným zaťažovaním. Z uvedeného dôvodu je zásadné pochopiť a popísať mechanické vlastnosti betónu taktiež počas tuhnutia a tvrdnutia.

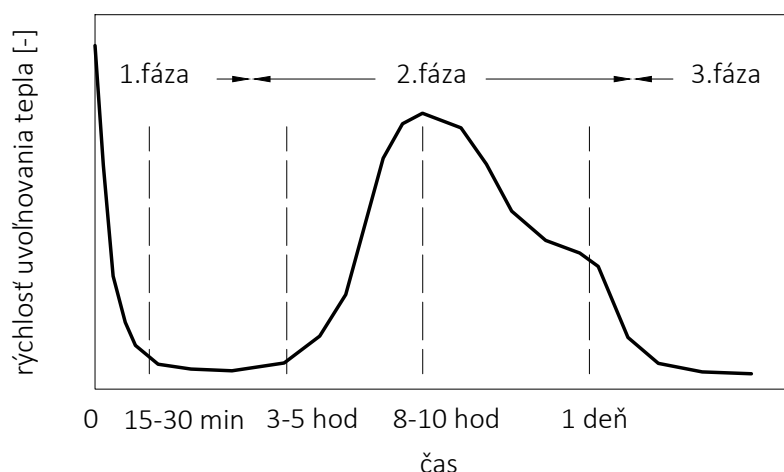
Po zmiešaní cementu, základnej zložky betónu, s vodou dochádza vplyvom chemických reakcií ku vzniku jemných kryštálikov, ktoré vzájomne prerastajú a tým cement postupne tvrdne. Daný proces je označovaný pojmom hydratácia cementu. Počas hydratácie prechádza betón zložitými fyzikálno-chemickými procesmi, ktoré sú sprevádzané neustálym menením mechanických vlastností betónu. Pre popísanie týchto zmien v čase bola vytvorená rada modelov od zložitejších po jednoduchšie. Zložitejšie modely v sebe zahrňujú viacero ovplyvňujúcich faktorov a potrebujú veľké množstvo presných vstupných premenných. Naopak zjednodušené modely niektoré faktory zanedbávajú a preto popisujú skutočné správanie betónu s menšou presnosťou.

Medzi najdôležitejšie faktory, ktoré ovplyvňujú vývoj mikroštruktúry tuhnúceho betónu, sa zaraďujú vlastnosti jednotlivých zložiek betónu, ich pomerné množstvo, objem a geometria betónového prvku a tiež vonkajšie podmienky ako napr. teplota alebo vlhkosť okolného vzduchu. Vývoj mechanických vlastností má tiež priamu súvislosť s priebehom hydratácie. Už v jej počiatočných štádiách, keď sa betón z viskózneho tmelu začína meniť na pevnú látku, je možné s konkrétnymi úpravami merať mechanické vlastnosti podobne ako u stvrdnutého betónu (Kalafutová, 2007). Preto bol zavedený pojem stupeň hydratácie, ktorý vyjadruje závislosť postupu hydratácie na čase. Veľkosť parametra stupňa hydratácie môže dosiahnuť hodnoty v obore $(0; 1)$ a znamená buď pomer doposiaľ zhydratovaného cementu k celkovému množstvu cementu, alebo tiež pomer zatiaľ uvoľneného tepla k celkovému potenciálnemu teplu, ktoré je schopný cement počas hydratácie uvoľniť.

Okrem doteraz spomenutých faktorov ovplyvňujúcich priebeh hydratácie, predstavuje jemnosť mletia cementu ďalší dôležitý činiteľ, na ktorom závisí rýchlosť hydratácie. Čím je cement jemnejšie namletý, tým je merný povrch jeho zŕn väčší. Voda má tak jednoduchší prístup k cementu a to urýchľuje prebiehajúce chemické reakcie. Na rýchlosť prebiehajúcich reakcií pôsobí tiež vodný súčiniteľ w , ktorý vyjadruje pomer hmotnosti účinného obsahu vody k hmotnosti cementu v čerstvom betóne. Tiež na ňom závisí stupeň hydratácie, ktorý môže betón dosiahnuť. Pri malých hodnotách vodného súčiniteľa sa voda obsiahnutá v betóne spotrebuje už v počiatku hydratácie a stupeň hydratácie tak nikdy nedosiahne teoretickú hodnotu 1. Betóny s nižším

stupňom hydratácie ale vykazujú lepšie mechanické vlastnosti a vyššiu trvanlivosť v porovnaní s betónmi s vyšším vodným súčiniteľom.

Na obrázku č. 5.3 je zobrazený priebeh hydratácie. Sledovaným parametrom je množstvo uvoľneného tepla v čase. Z obrázka možno vyčítať, že v rámci hydratácie sa rozlišujú tri fázy (Kubík, 2007).



Obr. 5.3: Produkcia tepla v rôznych fázach hydratácie (Kubík, 2007)

Prvá fáza je označovaná ako indukčná a rýchlosť uvoľňovania tepla je okrem prudkého počiatočného nárastu kludná. Po skončení prvej fázy, ktorá trvá približne 2 hodiny, nastáva fáza akceleračná. Je pre ňu typické zrýchlenie prebiehajúcich reakcií a v štruktúre betónu sa objavujú prvé väzby, vďaka ktorým sa betón postupne začína pretvárať na pevné skupenstvo a tuhne. V tejto časti hydratácie sa uvoľní väčšina vyprodukovaného tepla a približne po 1-2 dňoch nastáva tretia, deceleračná fáza. V nej je možné pozorovať opätovné znižovanie sa rýchlosti hydratácie.

5.3 Materiálové modely pre betón raného veku

Aby boli materiálové modely schopné správne vystihnúť krivku hydratácie, mali by byť závislé na všetkých ovplyvňujúcich parametroch, ktoré boli popísané v časti kapitoly 5.2. Dnes najpoužívanejšie modely možno na základe prístupu k popisu hydratácie rozdeliť na makroskopické a mikroskopické (Šmilauer V., 2003).

Makroskopické modely zaobchádzajú s betónom pri popísaní hydrátacie ako s celkom a snažia sa jej priebeh vyjadriť konkrétnou krivkou závislou na čase t . K popísaniu je tu zavedené väčšie množstvo ovplyvňujúcich konštánt a_i . Zavedené konštanty možno vyjadriť na základe vzorcov získaných z experimentálnych pozorovaní. Vypozorované vzťahy sú závislé na vodnom súčiniteli.

Mikroskopické modely sú založené na numerickej simulácii, ktorá umožňuje vyjadrenie správania jednotlivých minerálov cementu pri prebiehajúcich reakciách. Tým je dosiahnutý čo najlepší popis všetkých fáz hydrátacie s vysokou presnosťou.

5.4 Experimentálne skúmanie vlastností betónu raného veku

K vytvoreniu vhodného modelu umelej neurónovej siete, ktorý bude schopný vyjadriť vlastnosti betónu raného veku, bolo najprv potreba získať reálne namerané dáta na základe prevedených experimentov. Z časových dôvodov boli ako základ pre následnú tvorbu neurónovej siete použité už dostupné dáta. Jednalo sa o výsledky zo skúšok, ktoré boli uskutočnené na katedre betonových a zděných konstrukcí fakulty stavebnej ČVUT. Získané dáta slúžili jednak ako prostriedok pre porovnanie a kontrolu dosiahnutých výstupov z vytvoreného modelu umelej neurónovej siete, a tiež bola na ich základe zhotovená diplomová práca, ktorá experimentálne vyšetřovala vlastnosti betónu raného veku (Kalafutová, 2007).

Skúšky boli zrealizované na betónových vzorkách veľmi skorého veku. Konkrétne sa jednalo o čas 3, 4, 5, 6, 7 a 8 hodín. Uskutočnili sa celkom 4 zaťažovacie série s použitím štandardných oceľových foriem na výrobu skúšobných telies s rozmermi 150 x 150 x 150 mm. Po potrebných prípravách, ktoré boli nutné k začatiu experimentu boli v prvej časti skúšky vzorky zaťažované jednoosím krátkodobým zaťažením. V tejto časti došlo k úplnej deštrukcii jednotlivých betónových kociek a na základe toho bolo možné určiť pevnosť a tuhosť konkrétnych vzoriek. Zistené hodnoty boli potrebné pri druhej časti experimentu, v ktorej sa analyzovalo dotvarovanie skúšobných telies po dobu 20 minút. Na nasledujúcom obrázku je zobrazený stav skúšanej vzorky po realizácii experimentu.



Obr. 5.4: Deštruktívny stav skúšaného betónu (Kalafutová, 2007)

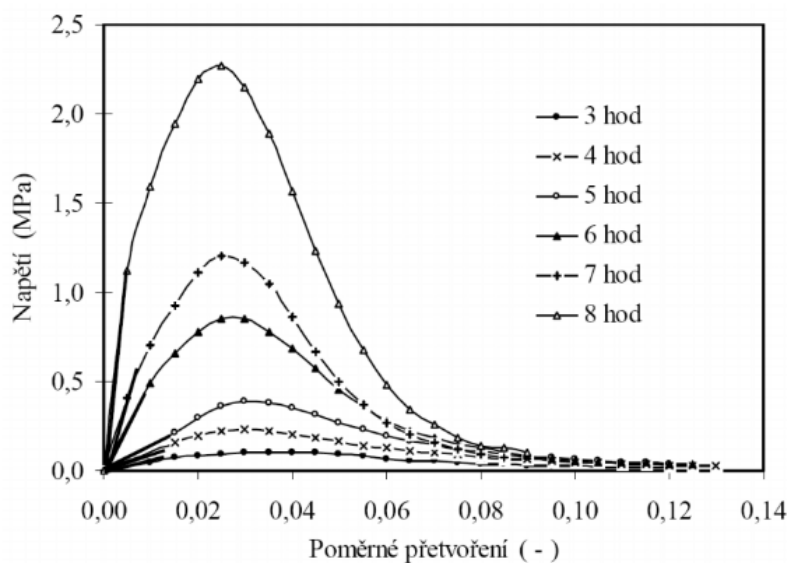
Na výrobu skúšaného betónu bol použitý cement *CEMI/42,5R*¹. V tabuľke 5.2 sú uvedené ďalšie základné vlastnosti betónu pri realizácii experimentov.

Tabuľka 5.1: Vlastnosti skúšaného betónu (Kalafutová, 2007)

Vlastnosť	Hodnota
Množstvo cementu	430 kg/m^3
Vodný súčiniteľ	0,42
Objemová hmotnosť	2442 kg/m^3
Pevnosť po 28 dňoch	$f_{c,28} = 66,0 MPa$

V nasledujúcich častiach bakalárskej práce sa pracovalo len s výsledkami z prvej časti experimentu, tzn. s nameranými dátami, na základe ktorých boli zhotovené pracovné diagramy betónu veľmi skorého veku pre časy postupne 3,4,5,6,7 a 8 hodín. Uvažovali sa spriemerované získané hodnoty z jednotlivých skúšobných sérií. Výsledky sú zobrazené na obrázku č. 5.5. S nameranými hodnotami, ktoré vyjadrovali dotvarovanie skúšobných telies sa ďalej nepracovalo.

¹Na základe označenia cementov používaného v norme ČSN EN 197-1 sa jedná o portlandský cement. Pevnosť cementu po 28 dňoch má hodnotu 42,5 MPa a cement sa vyznačuje rýchlym vývojom počiatočnej pevnosti (rýchloväzný).



Obr. 5.5: Spriemerované pracovné diagramy z experimentu (Kalafutová, 2007)

Cielom pokračujúcej časti práce bude vyjadriť pomocou umelých neurónových sietí priebeh pracovných diagramov z popísaných zaťažovacích skúšok nie len pre skúmané časy, ale zároveň tiež pre všetky medzilahlé časové hodnoty. Bude teda hľadaná spojitá plocha znázorňujúca vývoj pevnosti betónu raného veku v ľubovoľnom čase medzi 3 a 8 hodinami. Jedná sa o úlohu, ktorá v sebe zahŕňa dva vstupy (pretvorenie, čas) a jeden výstup (napätie). Pretvorenie bolo pre potreby tejto bakalárskej práce obmedzené na interval $\langle 0; 0,08 \rangle$ ².

V nasledujúcich podkapitolách bude popísaný postup pri tvorbe modelu umelej neurónovej siete, ktorý by bol schopný spĺňať predchádzajúce požiadavky. Podobne ako v predošlých častiach práce, bude aj tu kladený dôraz na vytvorenie vhodného modelu siete, ktorý zamedzí overfittingu a výstupom z neho budú výsledky, ktoré čo najviac odpovedajú reálnym očakávaniam.

²V niektorých časoch veku betónu končili experimentálne dáta krátko za pretvorením 0,08 a práve preto bol zvolený udelený interval.

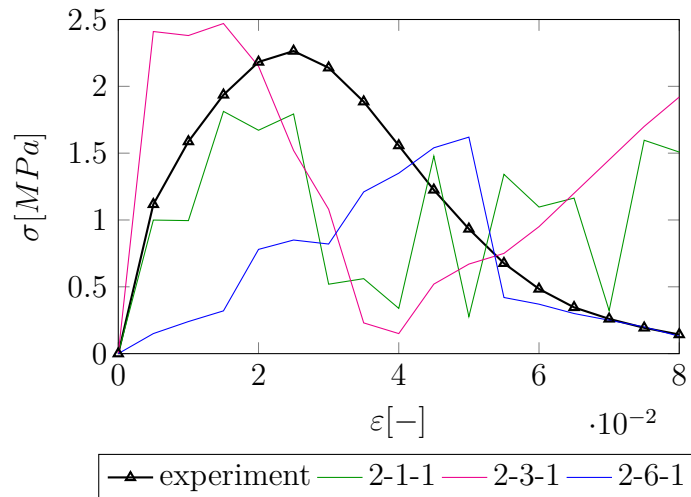
5.5 Tvorba modelu umelej neurónovej siete

Aby bola vytvorená neurónová sieť schopná správne vyjadriť pracovné modely aj v časoch, v ktorých neboli uskutočnené skúšky, je na začiatku potreba správne aproximovať výsledky dosiahnuté v experimente. Preto sa najprv vyskúšali výstupy zo sietí s rôznym počtom neurónov v skrytých vrstvách. Následne bola tiež overená a zamedzená prítomnosť overfittingu. Pre lepšiu názornosť budú bližšie popísané a vykreslené výstupy pre čas $t = 8 \text{ hod}$. V nasledujúcej tabuľke sú uvedené namerané výsledky poskytnuté z uskutočneného experimentu pre vybraný čas.

Tabuľka 5.2: Namerané výsledky pre čas $t = 8 \text{ hod}$ (Kalafutová, 2007)

t= 8 hod	
pretvorenie $\varepsilon[-]$	napätie $\sigma[MPa]$
0,005	1,117
0,010	1,587
0,015	1,936
0,020	2,182
0,025	2,264
0,030	2,139
0,035	1,885
0,040	1,557
0,045	1,225
0,050	0,933
0,055	0,676
0,060	0,484
0,065	0,346
0,070	0,260
0,075	0,193
0,080	0,143

Na začiatku boli vykreslené výstupy z nenatrenovaných sietí, ktoré mali každá iný počet neurónov v skrytej vrstve. Každá sieť má 2 vstupné neuróny, keďže výstupy zo siete závisia na čase a pretvorení a takisto má každá sieť 1 výstupný neurón, pretože hľadanou veličinou je napätie. Počet neurónov v skrytej vrstve siete s označením 2-1-1 je 1, analogicky má 3 neuróny sieť 2-3-1 a 6 neurónov v skrytej vrstve sa nachádza v sieti 2-6-1. Výstupy boli vykreslené pre príkaz `np.random.seed(1)`.



Obr. 5.6: Porovnanie výstupov z nenatrenovaných sietí s rôznymi počtami neurónov s experimentom (Kalafutová, 2007)

Následne boli všetky siete natrenované. V tomto prípade boli výstupy zo siete vykreslené aj pre medzilahlé body, ktoré neslúžili ako vstupné dáta z dôvodu overenia overfittingu. Zatiaľ, čo na základe predchádzajúcich výsledkov nebolo možné porovnaním zistiť, ktorý model siete je lepší, v danom prípade to už možné bolo. Vo všetkých prípadoch sa síce vyskytoval overfitting, tzn. výstupy zo siete odpovedali približne hodnotám získaných v experimente len v poskytnutých vstupných bodoch, ale najlepší celkový odhad dosiahla sieť 2-3-1. V nasledujúcej tabuľke sú zhromaždené výstupy zo všetkých troch sietí. Pre názornosť sú tu vypísané len výstupy pre poskytnuté vstupné dáta. V poslednom riadku tabuľky sú vypočítané hodnoty penalizačnej funkcie J^3 , ktorá slúžila ako základ pre porovnanie výsledkov. Jej princíp bol vysvetlený v časti 2.3.3. Pre výpočet hodnoty J už ale boli použité všetky dáta, pre ktoré boli vykreslené odhady zo siete, teda nie len dáta zobrazené

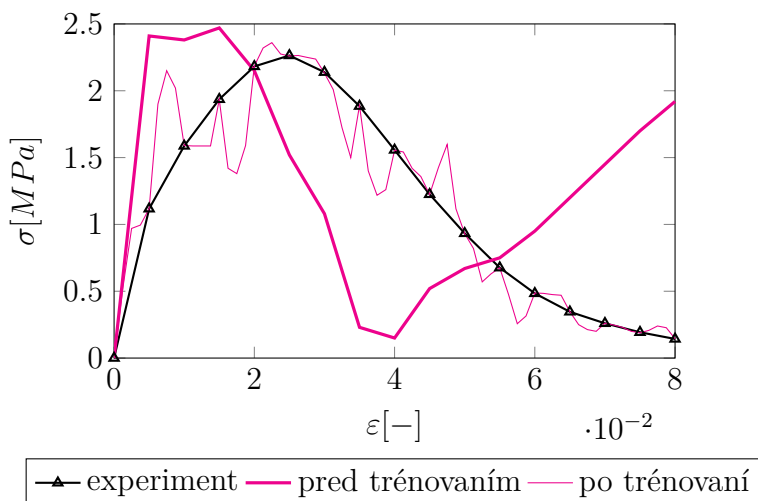
$$^3J = \sum \frac{1}{2}(y_i - \hat{y}_i)^2$$

v tabulke 5.3.

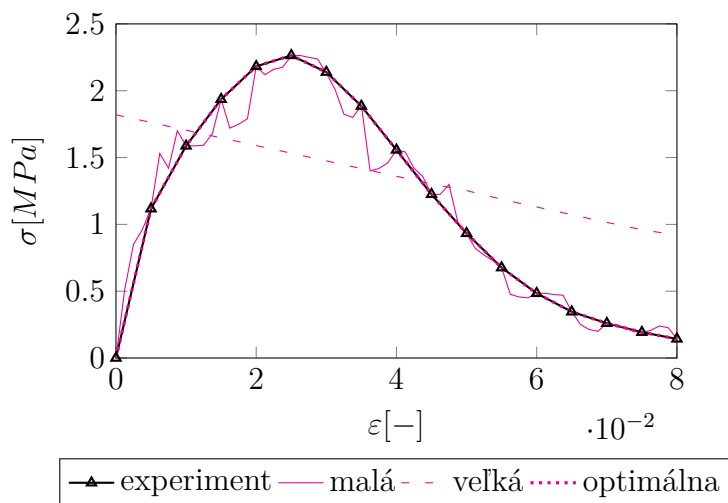
Tabuľka 5.3: Porovnanie výstupov z natréňovaných sietí s experimentom pre čas $t = 8 \text{ hod}$ (Kalafutová, 2007)

t= 8 hod				
pretvorenie $\varepsilon[-]$	napätie $\sigma[MPa]$			
–	experiment	sieť 2-1-1	sieť 2-3-1	sieť 2-6-1
0,005	1,117	1,115	1,118	1,125
0,010	1,587	1,570	1,590	1,593
0,015	1,936	1,930	1,935	1,934
0,020	2,182	1,184	2,185	1,175
0,025	2,264	1,263	2,260	1,259
0,030	2,139	1,143	2,132	1,140
0,035	1,885	1,892	1,889	1,893
0,040	1,557	1,549	1,555	1,552
0,045	1,225	1,220	1,224	1,227
0,050	0,933	1,930	0,931	1,930
0,055	0,676	1,671	0,675	1,680
0,060	0,484	1,485	0,487	1,480
0,065	0,346	1,350	0,347	1,342
0,070	0,260	1,258	0,261	1,255
0,075	0,193	1,190	0,189	1,195
0,080	0,143	1,150	0,145	1,140
J [-]				
–	0	3.925	2.909	5.827

Výstupy boli na obrázku č. 5.7 vykreslené len pre sieť, ktorá dosiahla najlepšiu aproximáciu a aj ostatné výstupy boli opäť zobrazené len pre danú sieť (tj. sieť 2-3-1). Po natrénovaní siete bol znovu zavedený regularizačný parameter λ a boli porovnané odlišné miery regularizácie (obrázok č. 5.8).

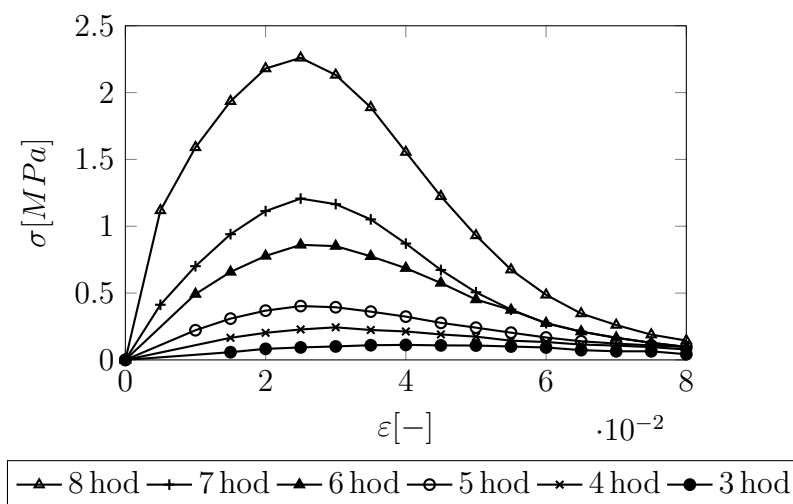


Obr. 5.7: Porovnanie výstupov z natrénovanej siete 2-3-1 s experimentom (Kalafutová, 2007)

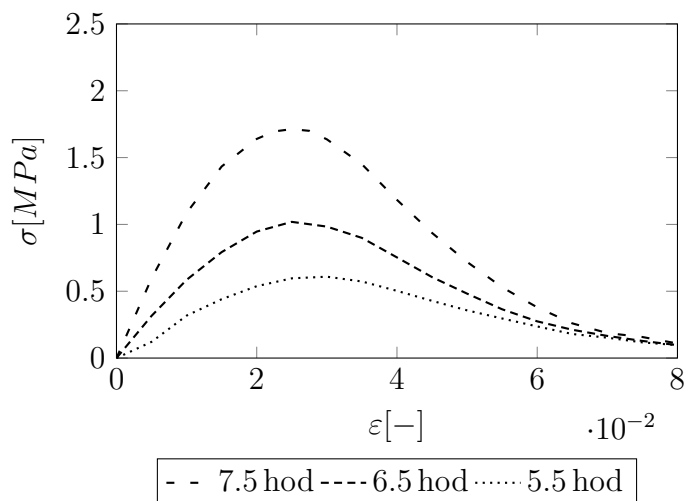


Obr. 5.8: Odlišné miery regularizácie na výstupy z natrénovanej siete

Rovakým spôsobom boli zo siete 2-3-1 vygenerované aj výstupy pre všetky ostatné časy, v ktorých sa uskutočnil experiment. Na obrázku č. 5.9 sú zobrazené výstupy z natrénovaného modelu siete s optimálnou mierou regularizácie (výstupy z nenatrénovaného resp. natrénovaného modelu siete s odlišnými mierami regularizácie nebudú vykreslené). Na ďalšom obrázku č. 5.10 sú v 2D vykreslené výstupy pre niektoré medzilahlé časy.

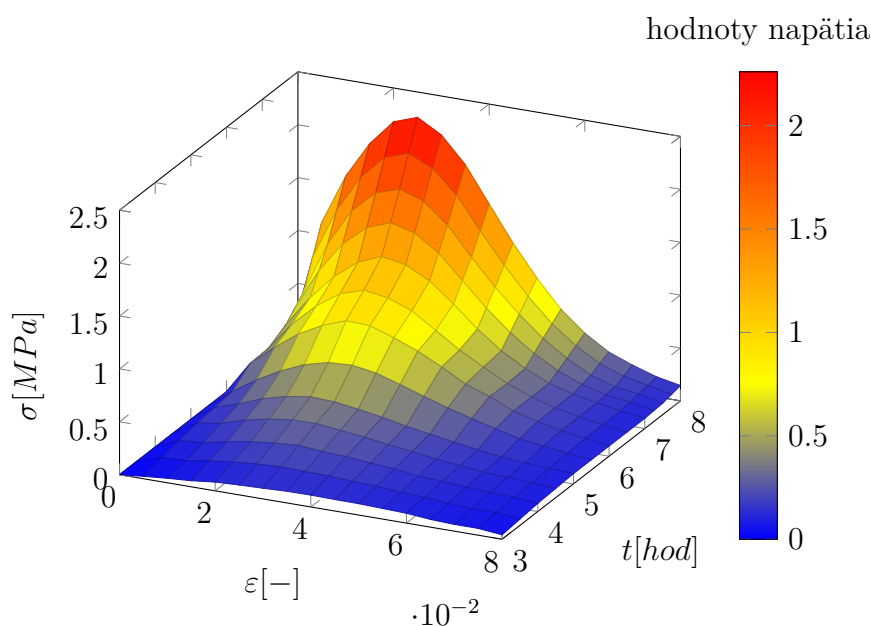


Obr. 5.9: Výsledné optimálne výstupy zo siete pre všetky časy



Obr. 5.10: Výsledné optimálne výstupy zo siete pre medzilahlé časy

Na základe získania pracovných diagramov aj pre medzilahlé časy bola následne tiež vykreslená výsledná plocha v 3D zobrazujúca závislosť napätia na pretvorení v čase.



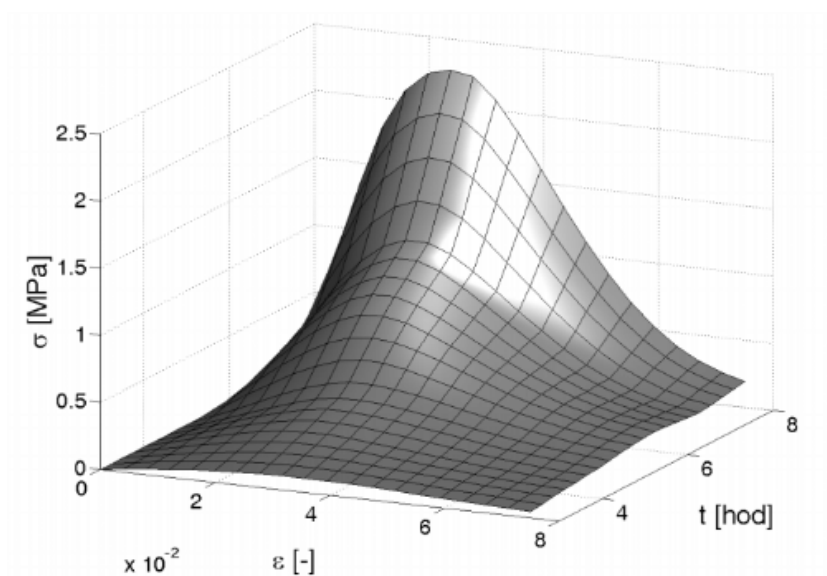
Obr. 5.11: Výsledná plocha pre všetky časy

5.6 Analýza výsledku

V závere tejto kapitoly bola výsledná plocha zobrazená na obrázku č. 5.11 porovnaná s výstupom z bakalárskej práce, ktorej cieľom bolo vyjadrenie pracovného diagramu betónu počas tuhnutia využitím fuzzy riadenia (Tomek, 2008). Zatiaľ, čo pri neurónových sieťach je prvotný výstup získaný na základe náhodného počítateľného vygenerovania váh a následne prebieha zlepšovanie výstupu pomocou tréningu siete, v ktorom sa prepočítavajú dané hodnoty, tak fuzzy riadenie sa usmerňuje úplne odlišnými pravidlami. Jedná sa o jednu z hlavných aplikácií fuzzy množín a nástrojom riadenia sú regulátory, ktoré možno označiť za automaty schopné v rozhodovacom procese zastúpiť človeka. Úspech týchto regulátorov vždy závisí na tom, ako dobre sa podarilo popísať regulovaný jav.

Základom pre obe práce boli rovnaké dáta poskytnuté zo skúšok uskutočnených na katedre betonových a zděných konštrukcií fakulty stavebnej, a preto sa mohli

výstupy z prác porovnať. Na obrázku č. 5.12 je zobrazená výsledná plocha fuzzy riadenia.



Obr. 5.12: Výsledná plocha fuzzy riadenia (Tomek, 2008)

Použitím dvoch odlišných prístupov k vyjadreniu pracovného diagramu betónu v ranom veku sa teda dosiahlo podobných výsledkov.

Kapitola 6

Záver

Bakalárska práca sa zaoberala využitím neurónových sietí pre tvorbu materiálových modelov. V úvode práce boli vysvetlené hlavné charakteristiky a pojmy typické pre umelé neurónové siete. Po ich vysvetlení bola vytvorená základná architektúra siete, ktorá umožňovala na základe vstupných dát predikovať prvé výstupy. Ako vstupné dáta v prvom prípade slúžili hodnoty rozponu prostého nosníku l , a rovnomerného zaťaženia q . Po ich zadaní do vytvoreného výpočtového modelu mala neurónová sieť za úlohu správne stanoviť veľkosť maximálneho ohybového momentu M_{max} . Na získaných výstupoch sa mohla overiť správnosť fungovania neurónovej siete. Požadovaná vhodnosť modelu bola tiež analyzovaná v nasledujúcej kapitole, ktorej cieľom bola aproximácia jednake funkcie s nespojitou funkciou a jednake nespojitej funkcie.

Cieľom tejto práce bolo porovnanie rôznych modelov umelých neurónových sietí pre tvorbu pracovného diagramu betónu v jednoosom tlakovom namáhaní v čase počas tuhnutia. Základom pre modelovanie vhodnej neurónovej siete boli výsledky zo skúšok tuhúcej betónovej zmesi vo veku od 3 do 8 hodín. V úvode kapitoly 5, ktorá sa venovala samotnému riešeniu daného problému, bol vysvetlený pojem pracovný diagram betónu a tiež boli zmienené základné deje prebiehajúce pri tuhnutí betónovej zmesi. Po uvedení do danej problematiky sa práca zaoberala tvorbou vhodnej umelej neurónovej siete.

Postupne boli vyskúšané siete s odlišným počtom neurónov v skrytej vrstve. Najprv boli vykreslené výstupy z nenatréovaných, a potom z natréovaných sietí.

Práve druhé spomenuté zobrazení sloužilo jako základný nástroj pro porovnání vhodností jednotlivých sítí. Po porovnání se dále pracovalo jen s sítí, která měla po uskotočnění trénování celkově nejlepší odhad. Následně byla overená přítomnost overfittingu v výstupu a boli graficky znázornené odlišné miery regularizácie, ktorých úlohou bolo zamedzenie overfittingu. Po získaní pracovných diagramov betónu v časoch, v ktorých sa experimenty uskotočnili, bolo dôležité tiež vygenerovať materiálový model aj pre zvyšné medzilahlé časy. Práve vďaka tomu bolo následne možné získať, a graficky znázorniť súvislú plochu vyjadrujúcu závislosť napätia na pretvorení v čase.

V práci bolo ukázané, že oblasť použitia neurónových sítí je takmer bez obmedzení. Do širokého intervalu ich využitia rozhodne možno zaradiť aj stavbársky priemysel a s tým spojené modelovanie vlastností materiálov, ktoré je pre následný návrh konštrukcií neoddeliteľnou súčasťou. Velkú výhodu predstavuje schopnosť predikcie chovania, a v neposlednom rade tiež redukcia nutného počtu vstupných dát k získaniu potrebných údajov.

Literatúra

Yaser S Abu-Mostafa, Malik Magdon-Ismail, and Hsuan-Tien Lin. *Learning from data*, volume 4. AMLBook New York, NY, USA:, 2012.

Mario Collepardi. *The new concrete*. Grafiche Tintoretto, 2006.

EN ČSN. 12350-2. zkoušení čerstvého betonu: Část 2: Zkouška sednutím, 2009.

Dario Floreano and Claudio Mattiussi. *Bio-inspired artificial intelligence: theories, methods, and technologies*. MIT press, 2008.

Joel Grus. *Data science from scratch: first principles with python*. Ö'Reilly Media, Inc.", 2015.

Jiří Šmejkal Hana Hanzlová. *Betonové a zděné konstrukce 1 Základy navrhování betonových konstrukcí*. České vysoké učení technické v Praze, 2015.

John Hopfield. Reflections on the aps and the evolution of physics, sep 2007. URL <https://www.aps.org/publications/apsnews/200708/backpage.cfm>.

Ivan Idris. *NumPy: Beginner's Guide*. Packt Publishing Ltd, 2015.

Petra Kalafutová. *Experimentální vyšetřování betonů raného stáří*. Diplomová práce. Praha: ČVUT FSV, 2007.

Martin Kubík. *Modelování vývinu hydratačního tepla na betonových konstrukcích*. Diplomová práce. Praha: ČVUT FSV, 2007.

Vladimír Kvasnička. Neurónové siete - kapitola 01, may 2014. URL <http://www2.fiit.stuba.sk/~kvasnicka/CognitiveScience/2.prednaska/Neuronove%20siete-kapitola01.pdf>.

- Bernardo Llanas, Sagrario Lantarón, and Francisco J Sáinz. Constructive approximation of discontinuous functions by neural networks. *Neural Processing Letters*, 2008.
- Warren S McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4):115–133, 1943.
- Tariq Rashid. *Make your own neural network*. CreateSpace Independent Publishing Platform, 2016.
- Eric Roberts. Neural networks, apr 2009. URL <http://cse.stanford.edu/class/sophomore-college/projects-00/neuralnetworks/index.html>.
- Rastko R Selmic and Frank L Lewis. Neural-network approximation of piecewise continuous functions: application to friction compensation. *IEEE transactions on neural networks*, 2002.
- Mark Summerfield. *Programming in Python 3: a complete introduction to the Python language*. Addison-Wesley Professional, 2010.
- Pavel Tomek. *Vyjádření vývoje pracovního diagramu betonu během tuhnutí pomocí fuzzy množin*. Bakalářská práce. Praha: ČVUT FSV, 2008.
- Bittnar Z. Šmilauer V. *Hydratace cementové pasty a model CEMHYD3D*. Beton TKS, 2003.
- Stephen Welch. Neural-networks-demystified, nov 2014. URL <http://https://github.com/stephencwelch/Neural-Networks-Demystified>.

Dodatok A

Tabuľka aktivačných funkcií

Tabuľka A.1: Tabuľka aktivačných funkcií

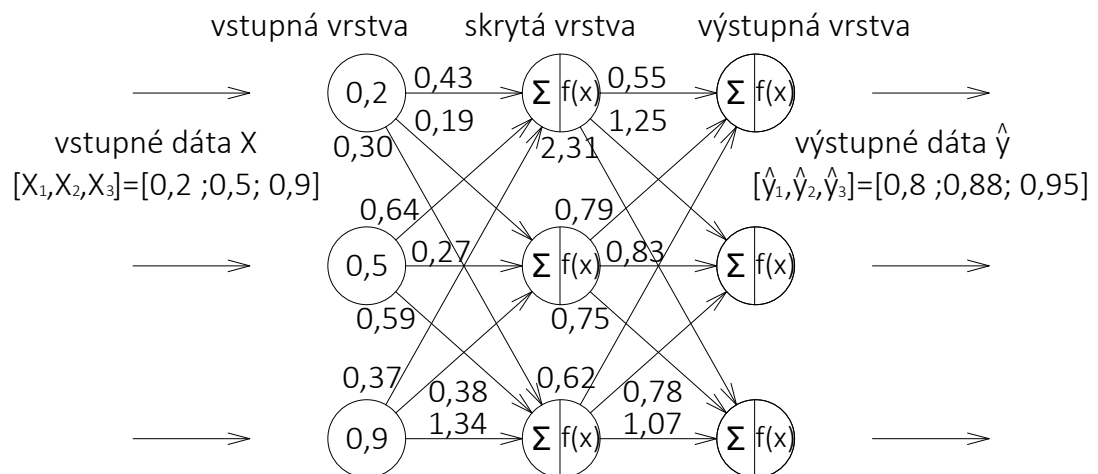
Názov funkcie (z angličtiny)	Predpis	Derivácia
ArcTan	$f(x) = \tan^{-1}(x)$	$f'(x) = \frac{1}{x^2+1}$
Rectified Linear Unit (ReLU)	$f(x) = \begin{cases} 0 & x < 0 \\ x & x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$
Parametric Rectified Linear Unit (PReLU)	$f(x) = \begin{cases} \alpha x & x < 0 \\ x & x \geq 0 \end{cases}$	$f'(x) = \begin{cases} \alpha & x < 0 \\ 1 & x \geq 0 \end{cases}$
Exponential Linear Unit (ELU)	$f(x) = \begin{cases} \alpha(e^x - 1) & x < 0 \\ x & x \geq 0 \end{cases}$	$f'(x) = \begin{cases} \alpha e^x & x < 0 \\ 1 & x \geq 0 \end{cases}$
Soft Plus	$f(x) = \log_e(1 + e^x)$	$f'(x) = \frac{1}{1+e^{-x}}$

Dodatok B

Vzor forward propagation

Pre vygenerovanie počiatkových náhodných hodnôt váh bol použitý príkaz `numpy.random.randn(r,s)` s argumentami r a s , ktoré vyjadrujú počet riadkov resp. stĺpcov matíc $\mathbf{W}^{(1)}$ a $\mathbf{W}^{(2)}$.

$$\mathbf{W}^{(1)} = \begin{bmatrix} 0,43 & 0,19 & 0,30 \\ 0,64 & 0,27 & 0,59 \\ 0,37 & 0,38 & 1,34 \end{bmatrix}, \mathbf{W}^{(2)} = \begin{bmatrix} 0,55 & 1,25 & 2,31 \\ 0,79 & 0,83 & 0,75 \\ 0,62 & 0,78 & 1,07 \end{bmatrix}$$



Obr. B.1: Vzorový príklad

Použitie rovnice 2.4:

$$\begin{aligned} \mathbf{S}^{(2)} &= \mathbf{X}\mathbf{W}^{(1)} = \begin{bmatrix} X_1 & X_2 & X_3 \end{bmatrix} * \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{bmatrix} = \\ &= \begin{bmatrix} 0,2 & 0,5 & 0,9 \end{bmatrix} * \begin{bmatrix} 0,43 & 0,19 & 0,30 \\ 0,64 & 0,27 & 0,59 \\ 0,37 & 0,38 & 1,34 \end{bmatrix} = \begin{bmatrix} 0,739 & 0,515 & 1,561 \end{bmatrix} \end{aligned} \quad (\text{B.1})$$

Použitie rovnice 2.5:

$$\mathbf{a}^{(2)} = f(\mathbf{S}^{(2)}) \quad (\text{B.2})$$

Aktivačná funkcia sigmoidu:

$$\begin{aligned} f(x) &= \frac{1}{1 + e^{-x}}, \\ f(0,739) &= \frac{1}{1 + e^{-0,739}} \doteq 0,67, \\ f(0,515) &= \frac{1}{1 + e^{-0,515}} \doteq 0,63, \\ f(1,561) &= \frac{1}{1 + e^{-1,561}} \doteq 0,83. \end{aligned} \quad (\text{B.3})$$

Použitie rovnice 2.8:

$$\begin{aligned} \mathbf{S}^{(3)} &= \mathbf{a}^{(2)}\mathbf{W}^{(2)} = \begin{bmatrix} a_1^{(2)} & a_2^{(2)} & a_2^{(2)} \end{bmatrix} * \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{bmatrix} = \\ &= \begin{bmatrix} 0,67 & 0,63 & 0,83 \end{bmatrix} * \begin{bmatrix} 0,55 & 1,25 & 2,31 \\ 0,79 & 0,83 & 0,75 \\ 0,62 & 0,78 & 1,07 \end{bmatrix} = \begin{bmatrix} 1,381 & 2,008 & 2,908 \end{bmatrix} \end{aligned} \quad (\text{B.4})$$

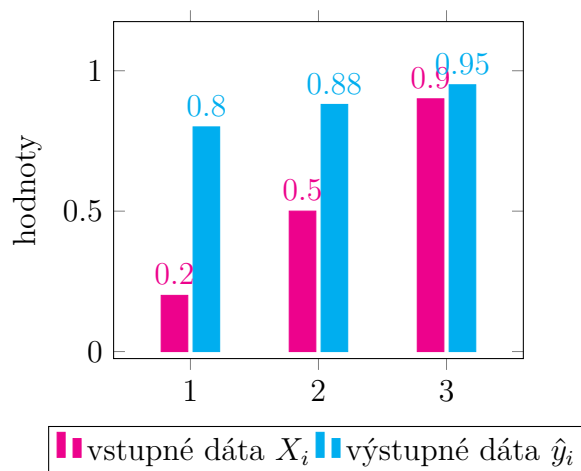
Použitie rovnice 2.9 a aplikácia aktivačnej funkcie sigmoidu:

$$\begin{aligned} \hat{y} &= f(\mathbf{S}^{(3)}), \\ f(x) &= \frac{1}{1 + e^{-x}}, \\ f(1,381) &= \frac{1}{1 + e^{-1,381}} \doteq 0,80, \\ f(2,008) &= \frac{1}{1 + e^{-2,008}} \doteq 0,88, \\ f(2,908) &= \frac{1}{1 + e^{-2,908}} \doteq 0,95. \end{aligned} \quad (\text{B.5})$$

Tabuľka B.1: Porovnanie vstupných a výstupných dát

	Vstupné dáta X_i	Výstupné dáta \hat{y}_i
1	0,2	0,80
2	0,5	0,88
3	0,9	0,95

Tým boli vypočítané výstupy zo siete, ktoré sa líšia od vstupných dát, pretože sieť nebola natrénovaná. Grafické porovnanie:



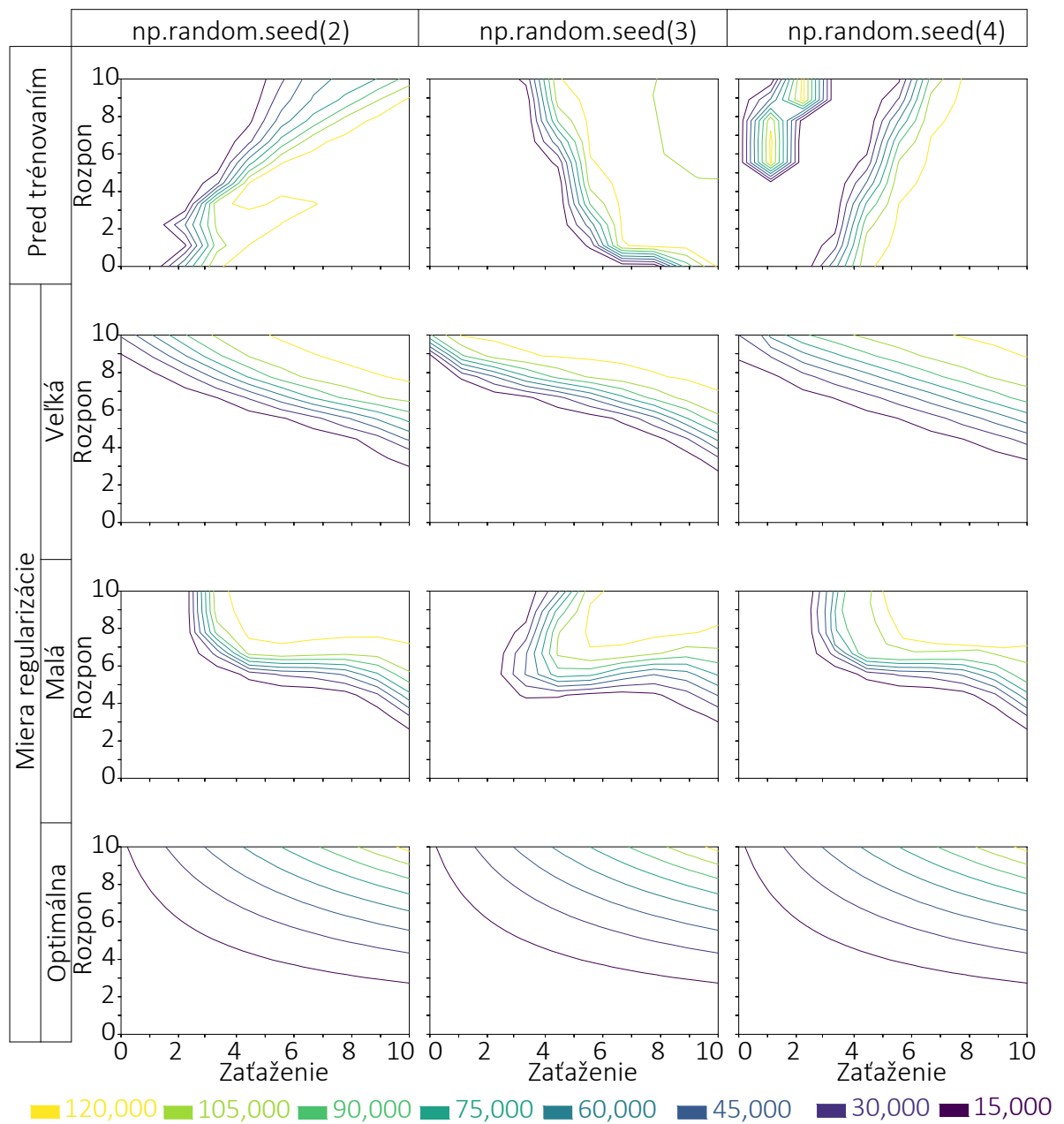
Obr. B.2: Grafické porovnanie

Dodatok C

Náhodnosť počiatočných hodnôt váh

Použitím príkazu `np.random.seed(x)` sa v Pythone dosiahne rovnaké nastavenie počiatočných náhodných hodnôt váh po každom opakovanom spustení programu. Na porovnanie boli použité tri odlišné argumenty príkazu `np.random.seed(x)` pre $x = 2$, $x = 3$, a $x = 4$. Tiež bola použitá postupne veľká, malá a optimálna miera regularizácie. Kvôli tomu bol pre každé x v rámci jednotlivých stupňov regularizácie vykreslený iný graf. V každom prípade bolo použitých 10 vstupných dát rozdelených na 6 tréningových a 4 testovacích a všetky grafy boli tiež porovnané s výstupmi z netrénovanej siete. Na odlišnosti grafov bolo dokázané správne vytvorenie modelu pre umelú neurónovú sieť. Pred tréňovaním siete sú medzi dosiahnutými výstupmi a výsledkami získanými použitím vzorca pre výpočet M_{max} veľké rozdiely.

Pri použití veľkej miery regularizácie mali všetky tri grafy tendenciu „narovnať sa“, tzn. výstup zo siete možno označiť za približne lineárnu závislosť momentu M_{max} na hodnotách zataženia q a rozponu nosníku l . Pri použití malej miery regularizácie sa výstupy len minimálne líšili od prípadu, keď bolo použitých 10 vstupných dát ale bez regularizácie, viď obrázok č. 3.8 na strane 25. Použitím vhodnej miery regularizácie sa už výstupy nelíšili a boli zhodné s prípadom, keď bolo použitých 90 vstupných dát bez regularizácie (obrázok č. 3.11d na strane 28). Výstupy odpovedajú tiež hodnotám dosiahnutých použitím vzorca $M_{max} = \frac{1}{8}ql^2$ s minimálnymi odchýlkami.



Obr. C.1: Vplyv náhodnosti počiatkových váh na výsledný výstup