

*IBM SPSS Neural Networks 29*



**Poznámka**

Před použitím těchto informací a produktu, který podporují, si přečtěte informace v tématu [“Upozornění” na stránce 17](#).

**Informace o produktu**

Toto vydání se vztahuje k verzi 29, vydání 0, modifikaci 1 produktu IBM® SPSS Statistics a ke všem následujícím vydáním a modifikacím, dokud nebude v nových vydáních uvedeno jinak.

© Copyright International Business Machines Corporation .

---

# Obsah

|  |           |
|--|-----------|
| <b>Kapitola 1. neuronové sítě.....</b> | <b>1</b>  |
| Úvod do neuronové sítě.....            | 1         |
| Co je neurální síť?.....               | 1         |
| Struktura neuronové sítě.....          | 2         |
| Multivrstva Perceptron.....            | 3         |
| Oblasti .....                          | 5         |
| Architektura .....                     | 5         |
| Školení .....                          | 7         |
| Výstup .....                           | 8         |
| Uložit .....                           | 9         |
| Export .....                           | 9         |
| Možnosti .....                         | 10        |
| Radiální základní funkce.....          | 10        |
| Oblasti .....                          | 12        |
| Architektura .....                     | 13        |
| Výstup .....                           | 13        |
| Uložit .....                           | 14        |
| Export .....                           | 15        |
| Volby .....                            | 15        |
| <b>Upozornění.....</b>                 | <b>17</b> |
| Ochranné známky.....                   | 18        |
| <b>Rejstřík.....</b>                   | <b>21</b> |



---

# Kapitola 1. neuronové sítě

Do produktu SPSS Statistics Premium Edition nebo Neural Networks Option jsou zahrnuty následující funkce neuronové sítě.

---

## Úvod do neuronové sítě

Neuronové sítě jsou upřednostňovaným nástrojem pro mnoho prediktivních aplikací pro dolování dat z důvodu jejich výkonu, flexibility a snadného použití. Prediktivní nervové sítě jsou užitečné zejména v aplikacích, kde je základní proces komplexní, jako např.:

- Prognózování spotřebitelské poptávky po zefektivnění výrobních a dodacích nákladů.
- Předpovědíte pravděpodobnost odpovědi na marketingový marketing, abyste určili, které domácnosti na seznamu adresátů by měly být zaslány nabídky.
- Ohodnocení žadatele za účelem stanovení rizika rozšíření úvěru na žadatele.
- Detekovat podvodné transakce v databázi pojistných nároků.

Nedirální sítě používané v prediktivních aplikacích, jako jsou multilayer perceptron (MLP) a sítě s radiálním základem (RBF) sítě, jsou kontrolovány v tom smyslu, že model-předpovězené výsledky lze porovnat se známými hodnotami cílových proměnných. Volba Neural Networks vám umožňuje vybavit sítě MLP a RBF a uložit výsledné modely bodování.

## Co je neurální síť?

Termín **neurální síť** se vztahuje na volně související řadu modelů charakterizovaných velkou prostorovou mezerou a pružnou strukturou, sestupné ze studií fungujících na mozku. Vzhledem k tomu, že rodina rostla, většina nových modelů byla navržena pro nebiologické použití, i když značná část související terminologie odráží její původ.

Specifické definice neurálních sítí se lišily jako oblasti, v nichž se používají. Zatímco žádná jednotlivá definice řádně nepokrývá celou řadu modelů, pro tuto chvíli zvažte následující popis<sup>1</sup>:

Neurální síť je masivně paralelní distribuovaný procesor, který má přirozený sklon k ukládání zkušenostního poznání a jeho zpřístupnění pro použití. Připomíná to mozek ve dvou ohledech:

- Znalost získává síť prostřednictvím procesu učení.
- Interní připojení silné stránky známé jako synaptické váhy se používají k ukládání znalostí.

Pro diskusi o tom, proč je tato definice možná příliš restriktivní, viz<sup>2</sup>.

Aby bylo možné odlišit neuronové sítě od tradičních statistických metod používajících tuto definici, *co není* uvádí, je stejně významné jako skutečný text definice. Například tradiční lineární regresní model může získat znalosti prostřednictvím metody nejmenších čtverců a uložit tyto znalosti v regresní koeficienty. V tomto smyslu je to neurální síť. Ve skutečnosti můžete tvrdit, že lineární regrese je zvláštním případem některých nervových sítí. Lineární regrese má však pevnou strukturu modelu a sadu předpokladů, které jsou uloženy před tím, než se data učíme.

Oproti tomu výše uvedená definice klade minimální požadavky na strukturu modelu a předpoklady. Neuronová síť tak může přibližovat širokou škálu statistických modelů, aniž by vyžadovala, abyste předem hypotetizovali určité vztahy mezi závislými a nezávislými proměnnými. Místo toho se forma vztahů určuje během procesu učení. Je-li lineární vztah mezi závislou a nezávislou proměnnou vhodný, výsledky neuronové sítě by měly úzce sblížit výsledky modelu lineární regrese. Je-li nelineární vztah vhodnější, neurální síť se automaticky přibližuje "správné" modelové struktuře.

---

<sup>1</sup> Haykin, S. 1998. *Neural Networks: A Comprehensive Foundation*, 2nd ed. New York: Macmillan College Publishing.

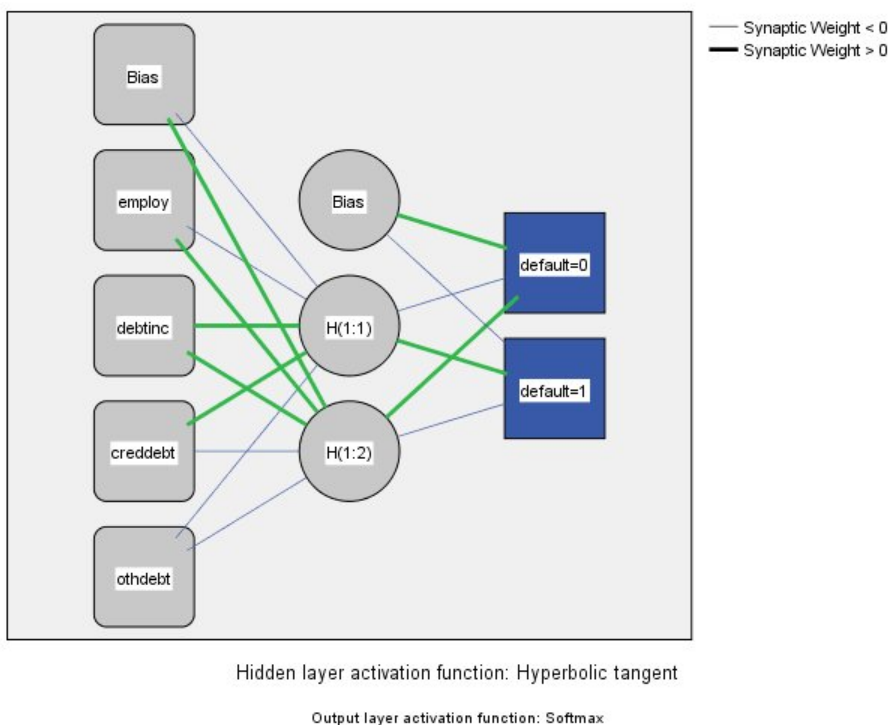
<sup>2</sup> Ripley, B. D. 1996. *Rozpoznávání vzorů a neuronové sítě*. Cambridge: Cambridge University Press.

Tento kompromis pro tuto flexibilitu spočívá v tom, že synaptické váhy neuronové sítě nejsou snadno interpretovatelné. Pokud se tedy pokoušíte vysvětlit základní proces, který vytváří vztahy mezi závislými a nezávislými proměnnými, bylo by lepší použít tradičnější statistický model. Nicméně, pokud modelový interpretatelnost není důležitý, můžete často získat dobrý model výsledků rychleji pomocí neuronové sítě.

## Struktura neuronové sítě

Přestože neuronové sítě kladou minimální požadavky na strukturu a předpoklady modelu, je užitečné pochopit obecnou **síťovou architekturu**. Síť multilayer perceptron (MLP) nebo radiální (radiální) funkce (RBF) je funkce prediktorů (také nazývaných vstupy nebo nezávislé proměnné), které minimalizují predikci chyb cílových proměnných (také nazývaných výstupy).

Uvažte datovou sadu *bankloan.sav*, která se dodává spolu s produktem, v němž chcete být schopni identifikovat možné defaulty mezi fondy žadatelů o úvěr. Síť MLP nebo RBF použitá na tento problém je funkce měření, která minimalizují chybu při předvídání výchozí hodnoty. Následující obrázek je užitečný pro související formát této funkce.



Obrázek 1. Feedforward architektura s jednou skrytou vrstvou

Tato struktura je známá jako **architektura přenosu kanálů**, protože připojení v toku sítě od vstupní vrstvy do výstupní vrstvy bez zpětných smyček se zpětnou vazbou. Na tomto obrázku:

- **Vstupní vrstva** obsahuje prediktory.
- **Skrytá vrstva** obsahuje nepozorovatelné uzly nebo jednotky. Hodnota každé skryté jednotky je některá z funkcí prediktorů; přesný tvar funkce závisí zčásti na typu sítě a v části na uživatelem ovladatelných specifikacích.
- **Výstupní vrstva** obsahuje odpovědi. Vzhledem k tomu, že historie výchozí je kategorická proměnná se dvěma kategoriemi, je přesměrováno jako dvě proměnné indikátorů. Každá výstupní jednotka je některá z funkcí skrytých jednotek. Opět platí, že přesný tvar funkce závisí zčásti na typu sítě a v části na uživatelem ovladatelných specifikacích.

Síť MLP umožňuje druhou skrytou vrstvou; v takovém případě je každá jednotka druhé skryté vrstvy funkcí jednotek v první skryté vrstvě a každá odezva je funkcí jednotek ve druhé skryté vrstvě.

## Multivrstva Perceptron

Procedura Multilayer Perceptron (MLP) produkuje prediktivní model pro jednu nebo více závislých (cílových) proměnných na základě hodnot proměnných prediktoru.

**Příklady.** Níže jsou uvedeny dva scénáře použití procedury MLP:

Úvěrový pracovník v bance musí být schopen identifikovat ukazatele, které svědčí o lidech, kteří pravděpodobně nevyužívají půjčky, a používají tyto charakteristiky k identifikaci dobrých a špatných úvěrových rizik. Pomocí vzorku minulých zákazníků může vyškolit vícevrstvý perceptron, potvrdit analýzu pomocí holdout vzorku minulých zákazníků a poté pomocí sítě klasifikovat potenciální zákazníky jako dobré nebo špatné úvěrové riziko.

Nemocniční systém má zájem na sledování nákladů a délky pobytu u pacientů přijatých k léčbě infarktu myokardu (MI, nebo "infarkt"). Získání přesných odhadů těchto opatření umožňuje aplikaci řádně spravovat dostupný prostor v lůžku tak, jak je zacházeno s pacienty. Při použití záznamů o léčbě u vzorku pacientů, kteří dostávali léčbu pro indikátor MI, může administrátor vycvičit síť, aby předpověděl jak náklady, tak délku pobytu.












Aspekty dat

**Závislé proměnné.** Závislé proměnné mohou být:

- **Nominální.** Proměnná může být považována za nominální, když její hodnoty představují kategorie bez vlastního hodnocení (například oddělení společnosti, v níž zaměstnanec pracuje). Příklady nominálních proměnných zahrnují region, poštovní směrovací číslo a náboženskou příslušnost.
- **Pořadové.** Proměnnou lze považovat za ordinální, když její hodnoty představují kategorie s nějakým vlastním hodnocením (například úroveň spokojenosti služeb od vysoce nespokojené s vysoce spokojené). Příklady ordinálních proměnných zahrnují skóre přístupu představující stupeň spokojenosti nebo důvěryhodnost a skóre hodnocení preferencí.
- **Měřítko.** Proměnná může být považována za měřítko (souvislá), když její hodnoty reprezentují uspořádané kategorie se smysluplnou metrikou, takže porovnání vzdálenosti mezi hodnotami je vhodné. Příklady proměnných měřítka zahrnují věk v rocích a příjem v tisících dolarů.

Procedura předpokládá, že byla přiřazena vhodná úroveň měření pro všechny závislé proměnné; avšak můžete dočasně změnit úroveň měření pro proměnnou tak, že klepnete pravým tlačítkem myši na proměnnou ze seznamu zdrojových proměnných a vyberete úroveň měření z rozevírací nabídky.

Ikona vedle každé proměnné v seznamu proměnných identifikuje úroveň měření a datový typ:

| Tabulka 1. Ikony úrovně měření |   |   |  |   |
|--------------------------------|---|---|--|---|
|                                | Číselné   | Řetězec   | Datum  | Čas   |
| Měřítko (spojité)              |  | není k dispozici  |  |  |
| Pořadové                       |  |  |  |  |
| Nominální                      |  |  |  |  |

**Proměnné prediktoru.** Prediktory lze zadat jako faktory (kategoriální) nebo kovariáty (měřítko).

**Kategorické kódování proměnných.** Procedura dočasně přeprogramuje kategoriální prediktory a závislé proměnné pomocí kódu one-of-c kódu po dobu trvání procedury. Pokud existuje c kategorií proměnné, pak je proměnná uložena jako c vektory, s první kategorií označenou (1,0, ..., 0), další kategorie (0,1,0, ..., 0), ..., a konečné kategorie (0,0, ..., 0, 1).

Tento kódovací schéma zvyšuje počet synaptické váhy a může vést k pomalejšímu tréninku; nicméně, více "kompaktní" kódovací metody obvykle vedou k špatně přizpůsobenými nervových sítí. Pokud vás

síťový trénink probíhá velmi pomalu, pokuste se snížit počet kategorií ve vašich kategoriálních prediktorů kombinací podobných kategorií nebo vypuštění případů, které mají extrémně vzácné kategorie.

Všechna jedna-of-c kódování je založena na údajích o školení, a to i v případě, že je definován test nebo soubor chladout (viz "Oblasti" na stránce 5). Pokud tedy vzorky zkoušek nebo pouzdro obsahují případy s prediktory, které nejsou obsaženy v údajích o výcviku, pak tyto případy nejsou používány postupem nebo bodovým hodnocením. Pokud testování nebo pouzdro vzorků obsahuje případy se závislými kategoriemi proměnných, které nejsou obsaženy v údajích o školení, pak tyto případy nejsou použity postupem, ale mohou být bodohodnoceny.

**Rescaling.** Proměnné závislé na měřítku a kovariáty jsou při výchozím nastavení proškolení pro zlepšení školení v síti. Veškeré opakované skenování se provádí na základě údajů o školení, a to i v případě, že je definován test nebo soubor chladout (viz "Oblasti" na stránce 5). To znamená, že v závislosti na typu rescalování se průměrná směrodatná odchylka, minimální hodnota nebo maximální hodnota proměnné kovariátu nebo závislé proměnné vypočítává pouze pomocí údajů o školení. Určíte-li proměnnou pro definování oblastí, je důležité, aby tyto proměnné nebo závislé proměnné měly podobné distribuce v rámci různých výukových programů, testování a odebrání vzorků.

**Frekvence frekvencí.** Tato procedura se ignoruje váhami frekvence.

**Replikace výsledků.** Chcete-li přesně replikovat výsledky, použijte stejnou inicializační hodnotu pro generátor náhodných čísel, stejné pořadí dat a stejné pořadí proměnných, jako dodatek k použití stejných nastavení procedury. Další podrobnosti o tomto problému jsou následující:

- **Generování náhodných čísel.** Procedura používá náhodné generování náhodných čísel při náhodném přiřazení logických oblastí, náhodný subvzorkování pro inicializaci synaptických vah, náhodný podvzorkování pro automatickou volbu architektury a simulovaný algoritmus annealing použité při inicializaci váhy a automatické výběru architektury. Chcete-li v budoucnu reprodukovat stejné náhodné výsledky, použijte pro generátor náhodných čísel stejnou inicializační hodnotu před každým spuštěním procedury Multilayer Perceptron.
- **Pořadí případů.** Metody školení online a Mini-batch (viz "Školení" na stránce 7) jsou explicitně závislé na pořadí případů; nicméně i dávkový trénink je závislý na pořadí případů, protože inicializace synaptických vah zahrnuje dílčí vzorkování z datové sady.

Chcete-li minimalizovat efekty objednávky, náhodně objednejte případy. Chcete-li ověřit stabilitu daného řešení, možná budete chtít získat několik různých řešení s případy setříděnými v různých náhodných objednávkách. V situacích s extrémně velkými velikostmi souborů lze provést více spuštění se vzorkem případů seřazených v různých náhodných příkazech.

- **Pořadí proměnných.** Výsledky mohou být ovlivněny pořadím proměnných v faktoru a kovariátních seznamů kvůli odlišnému vzorku počátečních hodnot přiřazených při změně pořadí proměnných. Stejně jako u případů s objednávkou případu můžete zkusit různé proměnné objednávky (jednoduše přetahovat pomocí myši v seznámech faktoru a kovariate), abyste zhodnotili stabilitu daného řešení.

Vytvoření sítě Perceptron Multilayer

Z nabídky vyberte:

#### **Analýza > Neutrální síť > Multilayer Perceptron ...**

1. Vyberte alespoň jednu závislou proměnnou.
2. Vyberte alespoň jeden faktor nebo kovarite.

Volitelně na kartě Proměnné můžete změnit metodu pro opakované změny proměnných. Volby jsou:

- **Standardizováno.** Odečtení střední hodnoty a dělení se směrodatnou odchylkou  $(x - \text{mean})/s$ .
- **Normalizované.** Odečtení minima a dělení podle rozsahu,  $(x - \text{min})/(\text{max} - \text{min})$ . Normalizované hodnoty spadají mezi 0 a 1.
- **Upraveno normalizováno.** Upravená verze odečtením minimálního a děleného rozsahu rozsahem,  $[2 * (x - \text{min})/(\text{max} - \text{min})] - 1$ . Upravené normalizované hodnoty spadají mezi -1 až 1.
- **Není nutno vykonat žádnou akci.** Žádné opakované okasování kovariátů.

Pole s neznámou úrovní měření



Výstraha na úrovni měření se zobrazí, když je úroveň měření pro jednu nebo více proměnných (polí) v datové sadě neznámá. Jelikož úroveň měření ovlivňuje výpočet výsledků pro tuto proceduru, všechny proměnné musí mít definovanou úroveň měření.

**Data skenování.** Přečte data v aktivní datové sadě a přiřadí výchozí úroveň měření k jakýmkoli polím s momentálně neznámou úrovní měření. Je-li datová sada velká, může to nějakou dobu trvat.

**Přiřadit ručně.** Otevře dialogové okno se seznamem všech polí s neznámou úrovní měření. Toto dialogové okno můžete použít k přiřazení úrovně měření k těmto polím. Úroveň měření můžete také přiřadit v pohledu Proměnné v editoru dat.

Vzhledem k tomu, že úroveň měření je pro tuto proceduru důležitá, nemůžete přistupovat k dialogovému oknu pro spuštění této procedury, dokud nebude mít všechna pole definovanou úroveň měření.

## Oblasti

**Datová sada oblasti.** Tato skupina uvádí metodu rozdělení aktivní datové sady na školení, testování a pouzdro vzorků. **Školící vzorek** se skládá z datových záznamů používaných pro výcvik neuronové sítě; některé procento případů v datové sadě musí být přiděleno vzorku výcviku, aby bylo možné získat model. **Testovací ukázka** je nezávislá sada datových záznamů používaných ke sledování chyb během školení, aby se zabránilo nadměrnému školení. Doporučuje se, abyste vytvořili testovací vzorek, a školení v síti bude obecně neefektivnější, pokud je testovací vzorek menší než vzorek odborné přípravy. **Výpočetní ukázka** je další nezávislá sada datových záznamů používaných k odhadu konečné neuronové sítě; chyba v souboru Holdout poskytuje "čestný" odhad prediktivní schopnosti modelu, protože pouzdro pouzdro nebylo použito k sestavení modelu.

- **Náhodně přiřazujte případy na základě relativního počtu případů.** Uvedte relativní číslo (poměr) případů náhodně vybraných do každého vzorku (trénink, testování a pouzdro). Sloupec **%** uvádí procentní část případů, které budou přiřazeny ke každému vzorku na základě zadaných relativních čísel.

Například, uvedte 7, 3, 0 jako relativní čísla pro výcvik, testování a výtěrové vzorky odpovídají 70%, 30% a 0%. Zadání 2, 1, 1 jako relativní čísla odpovídá 50%, 25% a 25%; 1, 1, 1 odpovídá rozdělení datové sady na stejné třetiny mezi tréninkem, testováním a holdout.

- **Použití rozdělovací proměnnou k přiřazení případů.** Uvedte numerickou proměnnou, která přiřadí každý případ v aktivní datové sadě k výcviku, testování nebo vzorkování dlouh. Případy s kladnou hodnotou proměnné jsou přiřazeny ke vzorku školení, případy s hodnotou 0, do vzorku zkoušek a případy se zápornou hodnotou do vzorku odchvovacího vzorku. Případy se systémem-chybějící hodnota jsou z analýzy vyloučeny. Jakýkoli uživatel-chybějící hodnoty proměnné logické oblasti je vždy považován za platný.

*Poznámka:* Použití proměnné rozdělení nezajistí stejné výsledky v následných spuštěních procedury. Viz téma "Replikace výsledků" v hlavním tématu [Multilayer Perceptron](#).

## Architektura

Karta Architecture se používá k určení struktury sítě. Procedura může automaticky vybrat "nejlepší" architekturu nebo můžete zadat vlastní architekturu.

Automatický výběr architektury sestaví síť s jednou skrytou vrstvou. Zadejte minimální a maximální počet jednotek povolených ve skryté vrstvě a výběr automatické architektury vypočítá "nejlepší" počet jednotek ve skryté vrstvě. Automatický výběr architektury používá výchozí aktivační funkce pro skryté a výstupní vrstvy.

Výběr vlastní architektury vám dává expertní kontrolu nad skrytými a výstupními vrstvami a může být nejužitečnější, když víte předem, jakou architekturu chcete, nebo když potřebujete naladit výsledky výběru automatické architektury.

Skryté vrstvy

Skrytá vrstva obsahuje nepozorovatelné síťové uzly (jednotky). Každá skrytá jednotka je funkce váženého součtu vstupů. Funkce je aktivační funkce a hodnoty vah jsou určeny algoritmem odhadování. Pokud

síť obsahuje druhou skrytou vrstvu, každá skrytá jednotka ve druhé vrstvě je funkcí váženého součtu jednotek v první skryté vrstvě. Stejná aktivační funkce se používá v obou vrstvách.

*Počet skrytých vrstev.* Vícevrstvý perceptron může mít jednu nebo dvě skryté vrstvy.

*Aktivační funkce.* Funkce aktivace "spojuje" vážené součty jednotek ve vrstvě na hodnoty jednotek v následné vrstvě.

- **Hyperbolický tangens.** Tato funkce má tento tvar:  $\gamma(c) = \tanh(c) = (e^c - e^{-c}) / (e^c + e^{-c})$ . Vezme argumenty s reálnými hodnotami a transformuje je na rozsah (-1, 1). Je-li použit výběr automatické architektury, jedná se o aktivační funkci pro všechny jednotky ve skrytých vrstvách.
- **Sigmoid.** Tato funkce má tvar:  $\gamma(c) = 1 / (1 + e^{-c})$ . Vezme argumenty s reálnými hodnotami a transformuje je na rozsah (0, 1).

*Počet jednotek.* Počet jednotek v každé skryté vrstvě lze zadat explicitně nebo automaticky určit pomocí algoritmu odhadu.

Výstupní vrstva

Výstupní vrstva obsahuje cílové (závislé) proměnné.

*Aktivační funkce.* Funkce aktivace "spojuje" vážené součty jednotek ve vrstvě na hodnoty jednotek v následné vrstvě.

- **Identita.** Tato funkce má tvar:  $\gamma(c) = c$ . Vezme argumenty s reálnými hodnotami a vrátí je nezměněné. Je-li použit výběr automatické architektury, jde o aktivační funkci pro jednotky ve výstupní vrstvě, pokud existují nějaké proměnné závislé na měřítku.
- **Softmax.** Tato funkce má tento tvar:  $\gamma(c_k) = \exp(c_k) / \sum_j \exp(c_j)$ . Vezme vektor argumentů s reálnými hodnotami a transformuje je na vektor, jehož prvky spadají do rozsahu (0, 1) a součtu na 1. Softmax je k dispozici pouze v případě, že všechny závislé proměnné jsou kategorické. Je-li použit výběr automatické architektury, jde o aktivační funkci pro jednotky ve výstupní vrstvě, pokud jsou všechny závislé proměnné kategorické.
- **Hyperbolický tangens.** Tato funkce má tento tvar:  $\gamma(c) = \tanh(c) = (e^c - e^{-c}) / (e^c + e^{-c})$ . Vezme argumenty s reálnými hodnotami a transformuje je na rozsah (-1, 1).
- **Sigmoid.** Tato funkce má tvar:  $\gamma(c) = 1 / (1 + e^{-c})$ . Vezme argumenty s reálnými hodnotami a transformuje je na rozsah (0, 1).

**Změna měřítka závislých proměnných stupnice závislé na velikosti.** Tyto ovládací prvky jsou dostupné pouze v případě, že byla vybrána alespoň jedna proměnná závislá na měřítku.

- **Standardizováno.** Odečtení střední hodnoty a dělení se směrodatnou odchylkou  $(x - \text{mean}) / s$ .
- **Normalizované.** Odečtení minima a dělení podle rozsahu,  $(x - \text{min}) / (\text{max} - \text{min})$ . Normalizované hodnoty spadají mezi 0 a 1. Jedná se o vyžadovanou metodu opětovného skenování proměnných závislých na měřítku, pokud výstupní vrstva používá aktivační funkci sigmoid. Volba pro korekci určuje malý počet  $\epsilon$ , který se použije jako korekce na vzorec pro změny měřítka. Tato oprava zaručuje, že všechny hodnoty proměnné závislé na rescaled budou ležet v rozsahu aktivační funkce. Zejména hodnoty 0 a 1, které se vyskytují v neopraveném vzorci, když  $x$  vezme své minimální a maximální hodnoty, definují limity rozsahu funkce sigmoid, ale nejsou v tomto rozsahu. Opravený vzorec je  $[x - (\text{min} - \epsilon)] / [(\text{max} + \epsilon) - (\text{min} - \epsilon)]$ . Zadejte číslo větší než nebo rovno 0.
- **Upraveno normalizováno.** Upravená verze odečtením minimálního a dělením rozsahu rozsahem,  $[2 * (x - \text{min}) / (\text{max} - \text{min})] - 1$ . Upravené normalizované hodnoty spadají mezi -1 až 1. Toto je požadovaná metoda opětovného skenování pro proměnné závislé na měřítku, pokud výstupní vrstva používá aktivační funkci hyperbolické tangenty. Volba pro korekci určuje malý počet  $\epsilon$ , který se použije jako korekce na vzorec pro změny měřítka. Tato oprava zaručuje, že všechny hodnoty proměnné závislé na rescaled budou ležet v rozsahu aktivační funkce. Zejména hodnoty -1 a 1, které se vyskytují v nekorigované vzorci, když  $x$  vezme své minimální a maximální hodnoty, definují limity rozsahu hyperbolické funkce tangens, ale nejsou v tomto rozsahu. Opravený vzorec je  $\{2 * [(x - (\text{min} - \epsilon))] / ((\text{max} + \epsilon) - (\text{min} - \epsilon))\} - 1$ . Zadejte číslo větší než nebo rovno 0.
- **Není nutno vykonat žádnou akci.** Proměnné závislé na měřítku nejsou k dispozici žádné opakované volání.

# Školení

Karta Školení se používá k určení, jak by měla být síť školená. Typ školení a optimalizační algoritmus určují, které možnosti školení jsou k dispozici.

**Typ školení.** Typ školení určuje, jak síť zpracovává záznamy. Vyberte jeden z následujících typů školení:

- **Dávka.** Aktualizuje synaptické váhy pouze po předání všech záznamů o školeních; to znamená, že dávkový trénink používá informace ze všech záznamů v datové sadě školení. Dávkové školení je často upřednostňováno, protože přímo minimalizuje celkovou chybu; dávkové školení však může vyžadovat aktualizaci vah tolikrát, dokud není splněna jedna z pravidel zastavení, a proto může potřebovat mnoho datových průchodů. Nejužitečnější je pro "menší" datové sady.
- **Online.** Aktualizuje synaptické váhy po každém záznamu dat odborné přípravy; to znamená, že online školení používá informace z jednoho záznamu po druhém. Školení online průběžně získává záznam a aktualizuje váhy do doby, než je splněna jedna z pravidel zastavení. Pokud jsou všechny záznamy použity jednou a žádná z pravidel pro zastavení není splněna, pak proces pokračuje v recyklaci datových záznamů. Online školení je lepší než dávková data pro "větší" datové sady s přidruženými prediktory; to znamená, že pokud existuje mnoho záznamů a mnoho vstupů a jejich hodnoty nejsou nezávislé na sobě navzájem, pak on-line školení může rychleji získat rozumnou odpověď, než dávkový trénink.
- **Mini-dávka.** Dělí záznamy dat o školení do skupin o přibližně stejné velikosti, pak aktualizuje synaptické váhy po předání jedné skupiny; to znamená minidávkové školení používá informace ze skupiny záznamů. Pak proces recykluje skupinu dat, je-li to nutné. Minidávková odborná příprava nabízí kompromis mezi dávkovým a online školením a může být nejlepší pro datové soubory "medium size". Procedura může automaticky určit počet záznamů školení pro minidávku, nebo můžete zadat celé číslo větší než 1 a menší nebo rovno maximálnímu počtu případů, které se mají uložit do paměti. Maximální počet případů pro uložení v paměti můžete nastavit na kartě [Volby](#).

**Optimalizační algoritmus.** To je metoda použitá k odhadu synaptické váhy.

- **Skalovaný přechod konjugentů.** Předpoklady, které odůvodňují použití metod gradientu konjugace, se vztahují pouze na dávkové tréninkové typy, takže tato metoda není k dispozici pro online nebo pro minidávkové školení.
- **Sestup gradientu.** Tato metoda musí být použita při online nebo minidávkovém školení; lze ji také použít s dávkovým výcvikem.

**Volby školení.** Volby školení vám umožňují dokončit optimalizační algoritmus. Obvykle nebudete muset tato nastavení měnit, pokud síť nevstoupí do problémů s odhadem.

Volby školení pro algoritmus škálovaného přechodového gradientu zahrnují:

- **Počáteční lambda.** Počáteční hodnota parametru lambda pro algoritmus škálovaného hašovacího přechodu. Zadejte číslo větší než 0 a menší než 0,000001.
- **Počáteční sigma.** Počáteční hodnota parametru sigma pro algoritmus škálovaného přechodu konjugentu. Zadejte číslo větší než 0 a menší než 0,0001.
- **Časové centrum a Posunutí intervalu.** Středový střed ( $a_0$ ) a posun intervalu ( $a$ ) definují interval  $[a_0 - a, a_0 + a]$ , v němž jsou vektory váhy náhodně generovány, když se používá simulovaná žihání. Simulované annealing se používá k vytržení z místního minima, s cílem najít globální minimum, během aplikace optimalizačního algoritmu. Tento přístup se používá při inicializaci váhy a při výběru automatické architektury. Zadejte číslo pro střed intervalu a číslo větší než 0 pro posun intervalu.

Volby školení pro přechodový algoritmus přechodu zahrnují:

- **Počáteční míra výuky.** Počáteční hodnota rychlosti učení pro algoritmus přechodu gradientu. Vyšší míra učení znamená, že síť bude trénovat rychleji, možná za cenu, že se stane nestabilní. Zadejte číslo větší než 0.
- **Spodní hranice Learning Rate.** Spodní hranice rychlosti učení pro algoritmus gradientu gradientu. Toto nastavení platí pouze pro online a minidávkové školení. Zadejte číslo větší než 0 a menší než počáteční míra učení.
- **Momentum.** Počáteční parametr hybnosti pro algoritmus přechodu gradientu. Termín hybnosti pomáhá předcházet nestabilitě způsobenému too-vysokým studijním očením. Zadejte číslo větší než 0.

- **Snižování míry výuky v Epochs.** Počet epoch ( $p$ ) nebo předání údajů o odborné přípravě, nezbytné ke snížení počáteční míry učení na nižší hranici míry učení, kdy se barevný přechod mění s on-line nebo minidávkovým výcvikem. To vám dává kontrolu nad faktorem rozpadu učení  $\beta = (1/p K) * \ln(\eta_0/\eta_{nizká})$ , kde  $\eta_0$  je počáteční míra učení, je hodnota  $\eta_{nizká}$  dolní hranice úrovně učení a  $K$  je celkový počet minidávek (nebo počet školících záznamů pro online školení) v souboru údajů o odborné přípravě. Zadejte celé číslo větší než 0.

## Výstup

**Struktura sítě.** Zobrazí souhrnné informace o neuronové síti.

- **Popis.** Zobrazí informace o neuronové síti, včetně závislých proměnných, počtu vstupních a výstupních jednotek, počtu skrytých vrstev a jednotek a aktivačních funkcí.
- **Diagram.** Zobrazí diagram sítě jako neupravitelný graf. Všimněte si, že jak se zvyšuje počet proměnných a úrovní faktoru, bude se diagram obtížněji interpretovat.
- **Sypité váhy.** Zobrazuje odhady koeficientu, které ukazují vztah mezi jednotkami v dané vrstvě a jednotkami v následující vrstvě. Synchronizace vah je založena na vzorku školení, i když je aktivní datová sada rozdělena na oblasti vzdělávání, testování a výzdro. Všimněte si, že počet synaptických vah může být spíše velký a že tyto váhy se obvykle nepoužívají pro interpretaci výsledků sítě.

**Výkon sítě.** Zobrazí výsledky použité k určení, zda je model "dobrý". *Poznámka:* Grafy v této skupině jsou založené na kombinovaných vzorcích školení a testování nebo pouze na ukázkách školení, pokud není k dispozici žádný testovací vzorek.

- **Souhrn modelu.** Zobrazuje souhrn výsledků neuronové sítě podle oblasti a celkové, včetně chyby, relativní chyby nebo procentní části chybných předpovědí, pravidlo zastavení použité k zastavení školení a času školení.

Chyba je chyba součtu hodnot druhých mocnin při použití aktivační funkce identity, sigmoid nebo hyperbolické tangenty na výstupní vrstvě. Je to chyba cross-entropie, když se na výstupní vrstvě použije aktivační funkce softmax.

Relativní chyby nebo procenta chybných předpovědí se zobrazí v závislosti na úrovních měření závislých proměnných. Má-li nějaká závislá proměnná úroveň měření, zobrazí se průměrná celková relativní chyba (relativní vzhledem k průměrnému modelu). Jsou-li všechny závislé proměnné kategorické, zobrazí se průměrný procentní podíl chybných předpovědí. Relativní chyby nebo procenta chybných předpovědí se také zobrazí pro jednotlivé závislé proměnné.

- **Výsledky klasifikace.** Zobrazí tabulku klasifikace pro každou kategorickou závislou proměnnou podle oblastí a celkově. Každá tabulka uvádí počet případů, které jsou správně a nesprávně klasifikovány pro každou kategorii závislých proměnných. Procento z celkového počtu případů, které byly správně klasifikovány, je také hlášeno.
- **Křivka ROC.** Zobrazí křivku ROC (Receiver Operating Characteristic) pro každou kategorickou závislou proměnnou. Také zobrazuje tabulku poskytující plochu pod každou křivkou. Pro danou závislou proměnnou zobrazí graf ROC jednu křivku pro každou kategorii. Má-li závislá proměnná dvě kategorie, pak každá křivka zachází s danou kategorií jako s pozitivním stavem a s druhou kategorií. Má-li závislá proměnná více než dvě kategorie, pak každá křivka zachází s danou kategorií jako s pozitivním stavem a s agregací všech ostatních kategorií.
- **Graf kumulativních zisků.** Zobrazí kumulativní přírůstkový graf pro každou kategorickou závislou proměnnou. Zobrazení jedné křivky pro každou kategorii závislých proměnných je stejné jako u křivek ROC.
- **Zvedací graf.** Zobrazí zvedací graf pro každou kategorickou závislou proměnnou. Zobrazení jedné křivky pro každou kategorii závislých proměnných je stejné jako u křivek ROC.
- **Předpovědi podle pozorovaného grafu.** Zobrazí graf předpovědané hodnoty podle pozorovaného grafu pro každou závislou proměnnou. Pro kategorické závislé proměnné jsou pro každou kategorii odezvy zobrazeny klastrované krabičkové grafy predikované pseudo-pravděpodobnosti, přičemž kategorie odezvy je pozorovaná jako proměnná klastru. U proměnných závislých na měřítku se zobrazí bodový graf.

- **Zbývající podle predikovaného grafu.** Zobrazí graf se zbytkovou hodnotou podle předpovědi pro každou proměnnou závislou na měřítku. Mezi reziduálními hodnotami a předpovězenými hodnotami by neměly být žádné viditelné vzory. Tento graf se vytváří pouze pro proměnné závislé na měřítku.

**Souhrn zpracování případu.** Zobrazí souhrnnou tabulku zpracování případu, která shrnuje počet zahrnutých případů a vyloučených v analýze, celkem a podle školení, testování a odebrání vzorků.

**Analýza důležitosti nezávislé proměnné.** Provede analýzu citlivosti, která vypočítá důležitost každého prediktoru při určování neuronové sítě. Analýza je založena na kombinované cvičné odborné přípravě a zkouškách vzorků nebo pouze na zkušební vzorek, pokud není k dispozici žádný zkušební vzorek. Tím se vytvoří tabulka a graf zobrazující důležitost a normalizovaný význam pro každý prediktor. Všimněte si, že analýza citlivosti je výpočetně nákladná a časově náročná, pokud existuje velký počet prediktorů nebo případů.

## Uložit

Karta Uložit se používá k ukládání předpovědi jako proměnných v datové sadě.

- **Ukládat předpovězenou hodnotu nebo kategorii pro každou závislou proměnnou.** Tím uložíte předpovězenou hodnotu pro proměnné závislé na měřítku a předpovídanou kategorii pro kategoriální závislé proměnné.
- **Ukládat předpokládanou pseudo-pravděpodobnost nebo kategorii pro každou závislou proměnnou.** Tento příkaz uloží predikované pseudo-pravděpodobnosti pro kategoriální závislé proměnné. Samostatná proměnná se uloží pro každou z prvních  $n$  kategorií, kde  $n$  je uveden ve sloupci **Kategorie určené k uložení**.

**Názvy uložených proměnných.** Automatické generování názvu zajišťuje, že zachováte veškerou svou práci. Vlastní názvy vám umožňují vyřadit/nahradit výsledky z předchozích spuštění bez předchozího odstranění uložených proměnných v Editoru dat.

Pravděpodobnosti a pseudopravdnosti

Categorical dependent variables with softmax activation and cross-entropy error will have a predicted value for each category, where each predicted value is the stability that the case belongs to the category.

Kategorické závislé proměnné s chybou sum-of-squares budou mít předpovězenou hodnotu pro každou kategorii, ale predikované hodnoty nelze interpretovat jako pravděpodobnosti. Procedura ukládá tyto predikované pseudo-pravděpodobnosti i v případě, že jakákoli hodnota je menší než 0 nebo větší než 1, nebo součet pro danou závislou proměnnou není 1.

ROC, kumulativní zisky a zvedací grafy (viz "Výstup" na stránce 8) jsou založeny na pseudo-pravděpodobnosti. V případě, že kterýkoli z pseudopravděpodobnosti je menší než 0 nebo větší než 1, nebo součet pro danou proměnnou není 1, jsou nejprve překalibrovány tak, aby byly mezi 0 a 1 a součet 1. Pseudo-pravděpodobnosti se rescaled tím, že se dělí podle jejich součtu. Například, pokud případ předpovídal pseudo-pravděpodobnosti 0,50, 0,60 a 0,40 pro tří-kategorii závislou proměnnou, pak každé pseudo-pravděpodobnost je vydělí součtem 1,50, aby získal 0,33, 0,40, a 0,27.

Je-li některý z pseudopravděpodobnosti negativní, pak se absolutní hodnota nejnižší hodnoty přidá ke všem pseudopravděpodobnostem před výše uvedeným rescaling. Je-li například hodnota pseudo-pravděpodobnosti 0,30, 0,50 a 1,30, pak nejprve přidejte 0,30 do každé hodnoty pro získání 0,00, 0,80 a 1,60. Dále rozdělte každou novou hodnotu součtem 2,40 pro získání 0,00, 0,33 a 0,67.

## Export

Karta Export se používá k ukládání odhadů synaptické váhy pro každou závislou proměnnou na soubor XML (PMML). Tento modelový soubor můžete použít k použití informací modelu na jiné datové soubory pro účely hodnocení. Tato volba není k dispozici, pokud byly definovány rozdělené soubory.

## Možnosti

**Uživatel-Chybějící hodnoty.** Faktory musí mít platné hodnoty pro případ, který má být zahrnut do analýzy. Tyto ovládací prvky umožňují rozhodnout o tom, zda uživatelské hodnoty jsou považovány za platné mezi faktory a kategoriálními závislými proměnnými.

**Zastavení pravidel.** To jsou pravidla, která určují, kdy přestat trénovat neurální síť. Školení probíhá prostřednictvím alespoň jednoho datového toku. Školení lze poté zastavit podle následujících kritérií, která jsou kontrolována v uvedeném pořadí. V definicích zastavovacího pravidla, které následují, odpovídá jeden krok přenosu dat pro online a mini-dávkové metody a iteraci pro dávkovou metodu.

- **Maximum kroků bez snížení chyby.** Počet kroků, které je třeba provést před kontrolou snížení chyby. Pokud po uplynutí zadaného počtu kroků nedojde k žádnému snížení chyby, pak se školení zastaví. Zadejte celé číslo větší než 0. Můžete také určit, který vzorek dat se použije k výpočtu chyby. Volba **Zvolit automaticky** použije testovací vzorek, pokud existuje, a použije výběrový soubor s jiným způsobem. Vezměte na vědomí, že dávková odborná příprava zaručuje po každém průchodu dat pokles v ukázkové chybě školení, a proto se tato volba vztahuje pouze na dávkové školení, pokud existuje testovací ukázka. **Výcvik a testovací data** zkontroluje chybu pro každý z těchto ukázek. Tato volba se použije pouze v případě, že se testovací ukázka ukončí.

*Poznámka:* Po každém úplném předání dat vyžaduje online a mini-dávkový trénink další předání dat, aby bylo možné vypočítat chybu školení. Tento přebytečný přenos dat může výrazně zpomalit školení, proto se obecně doporučuje, abyste dodali testovací vzorek a vybrali **Zvolit automaticky** v každém případě.

- **Maximální doba školení.** Vyberte, zda chcete určit maximální počet minut, kdy se má algoritmus spustit. Zadejte číslo větší než 0.
- **Maximální odborná příprava Epochs.** Je povolen maximální povolený počet epoch (předání dat). Dojde-li k překročení maximálního počtu epoch, pak se odborná příprava zastaví. Zadejte celé číslo větší než 0.
- **Minimální relativní změna v chybě školení.** Školení se zastaví, pokud je relativní změna v chybě školení v porovnání s předchozím krokem menší než hodnota kritéria. Zadejte číslo větší než 0. V případě online a minidávkového školení je toto kritérium ignorováno, pokud jsou k výpočtu chyby použita pouze testovací data.
- **Minimální relativní změna poměru chyb školení.** Školení se zastaví, je-li poměr chyb učení k chybě modelu null menší než hodnota kritéria. Model s hodnotou null předpovídá průměrnou hodnotu pro všechny závislé proměnné. Zadejte číslo větší než 0. V případě online a minidávkového školení je toto kritérium ignorováno, pokud jsou k výpočtu chyby použita pouze testovací data.

**Maximální počet případů pro uložení v paměti.** Tento ovládací prvek řídí následující nastavení v rámci algoritmů vícevrstvého perceptron. Zadejte celé číslo větší než 1.

- Ve výběru automatické architektury je velikost vzorku použitého k určení architektury sítě min  $(1000, memsize)$ , kde  $memsize$  je maximální počet případů, které se mají uložit v paměti.
- V minidávkovém školení s automatickým výpočtem počtu minidávek, je počet minidávek je min  $(M/10, 2), memsize)$ , kde  $M$  je počet případů v rámci školení Ukázka.

## Radiální základní funkce

Procedura RBF (Radial Basis Function) produkuje prediktivní model pro jednu nebo více závislých (cílových) proměnných na základě hodnot proměnných prediktoru.

**Příklad.** Poskytovatel telekomunikačních služeb segmentoval své zákaznické základny podle vzorce užívání služeb a kategorizuje zákazníky do čtyř skupin. Síť RBF používající demografická data k předpovídání členství ve skupině umožňuje společnosti přizpůsobit nabídky pro jednotlivé potenciální zákazníky.

Aspekty dat












**Závislé proměnné.** Závislé proměnné mohou být:

- **Nominální.** Proměnná může být považována za nominální, když její hodnoty představují kategorie bez vlastního hodnocení (například oddělení společnosti, v níž zaměstnanec pracuje). Příklady nominálních proměnných zahrnují region, poštovní směrovací číslo a náboženskou příslušnost.
- **Pořadové.** Proměnnou lze považovat za ordinální, když její hodnoty představují kategorie s nějakým vlastním hodnocením (například úroveň spokojenosti služeb od vysoce nespokojené s vysoce spokojené). Příklady ordinálních proměnných zahrnují skóre přístupu představující stupeň spokojenosti nebo důvěryhodnost a skóre hodnocení preferencí.
- **Měřítko.** Proměnná může být považována za měřítko (souvislá), když její hodnoty reprezentují uspořádané kategorie se smysluplnou metrikou, takže porovnání vzdálenosti mezi hodnotami je vhodné. Příklady proměnných měřítka zahrnují věk v rocích a příjem v tisících dolarů.

Procedura předpokládá, že byla přiřazena vhodná úroveň měření pro všechny závislé proměnné, ačkoli můžete dočasně změnit úroveň měření proměnné tak, že pravým tlačítkem myši klepnete na proměnnou ze seznamu zdrojových proměnných a vyberete úroveň měření z rozevírací nabídky.

Ikona vedle každé proměnné v seznamu proměnných identifikuje úroveň měření a datový typ:

*Tabulka 2. Ikony úrovně měření*

|                   | Číselné  | Řetězec  | Datum   | Čas  |
|-------------------|--|--|---|--|
| Měřítko (spojité) |   | není k dispozici   |   |   |
| Pořadové          |   |   |   |   |
| Nominální         |  |  |  |  |

**Proměnné prediktoru.** Prediktory lze zadat jako faktory (kategoriální) nebo kovariáty (měřítko).

**Kategorické kódování proměnných.** Procedura dočasně přeprogramuje kategoriální prediktory a závislé proměnné pomocí kódu one-of-c kódu po dobu trvání procedury. Pokud existuje c kategorií proměnné, pak je proměnná uložena jako c vektory, s první kategorií označenou (1,0, ..., 0), další kategorie (0,1,0, ..., 0), ..., a konečné kategorie (0,0, ..., 0, 1).

Toto kódovací schéma zvyšuje počet synaptické váhy a může vést k pomalejšímu tréninku, ale více "kompaktní" kódovací metody obvykle vedou k špatně přizpůsobenými nervových sítí. Pokud váš síťový trénink probíhá velmi pomalu, pokuste se snížit počet kategorií ve vašich kategoriálních prediktorů kombinací podobných kategorií nebo vypuštění případů, které mají extrémně vzácné kategorie.

Všechna jedna-of-c kódování je založena na údajích o školení, a to i v případě, že je definován test nebo soubor chluďout (viz "Oblasti" na stránce 12). Pokud tedy vzorky zkoušek nebo pouzdro obsahují případy s prediktory, které nejsou obsaženy v údajích o výcviku, pak tyto případy nejsou používány postupem nebo bodovým hodnocením. Pokud testování nebo pouzdro vzorků obsahuje případy se závislými kategoriemi proměnných, které nejsou obsaženy v údajích o školení, pak tyto případy nejsou použity postupem, ale mohou se skórovali.

**Rescaling.** Proměnné závislé na měřítku a kovariáty jsou při výchozím nastavení proškolení pro zlepšení školení v síti. Veškeré opakované skenování se provádí na základě údajů o školení, a to i v případě, že je definován test nebo soubor chluďout (viz "Oblasti" na stránce 12). To znamená, že v závislosti na typu rescalování se průměrná směrodatná odchylka, minimální hodnota nebo maximální hodnota proměnné kovariátu nebo závislé proměnné vypočítává pouze pomocí údajů o školení. Určíte-li proměnnou pro definování oblastí, je důležité, aby tyto proměnné nebo závislé proměnné měly podobné distribuce v rámci různých výukových programů, testování a odebrání vzorků.

**Frekvence frekvencí.** Tato procedura se ignoruje váhami frekvence.



**Replikace výsledků.** Chcete-li přesně replikovat výsledky, použijte stejnou inicializační hodnotu pro generátor náhodných čísel a stejný datový příkaz, kromě použití stejných nastavení procedury. Další podrobnosti o tomto problému jsou následující:

- **Generování náhodných čísel.** Procedura používá náhodné generování čísel během náhodného přiřazení oblastí. Chcete-li v budoucnu reprodukovat stejné náhodné výsledky, použijte pro generátor náhodných čísel stejnou inicializační hodnotu před každým spuštěním procedury Radial Basis Function.
- **Pořadí případů.** Výsledky jsou také závislé na pořadí dat, protože se používá dvoukrokový klastrový algoritmus k určení radiálních základních funkcí.

Chcete-li minimalizovat efekty objednávky, náhodně objednejte případy. Chcete-li ověřit stabilitu daného řešení, možná budete chtít získat několik různých řešení s případy seřazenými v různých náhodných objednávkách. V situacích s extrémně velkými velikostmi souborů lze provést více spuštění se vzorkem případů seřazených v různých náhodných příkazech.

Vytváření radiální funkce základu pro práci

Z nabídky vyberte:

#### **Analýza > Nezvukové sítě > Radiální základní funkce ...**

1. Vyberte alespoň jednu závislou proměnnou.
2. Vyberte alespoň jeden faktor nebo kovariate.

Volitelně na kartě Proměnné můžete změnit metodu pro opakované změny proměnných. Volby jsou následující:

- **Standardizováno.** Odečtení střední hodnoty a dělení se směrodatnou odchylkou  $(x - \text{mean})/s$ .
- **Normalizované.** Odečtení minima a dělení podle rozsahu,  $(x - \text{min})/(\text{max} - \text{min})$ . Normalizované hodnoty spadají mezi 0 a 1.
- **Upraveno normalizováno.** Upravená verze odečtením minimálního a dělením rozsahu rozsahem,  $[(x - \text{min})/(\text{max} - \text{min})] - 1$ . Upravené normalizované hodnoty spadají mezi -1 až 1.
- **Není.** Žádné opakované okasování kovariátů.

Pole s neznámou úrovní měření

Výstraha na úrovni měření se zobrazí, když je úroveň měření pro jednu nebo více proměnných (polí) v datové sadě neznámá. Jelikož úroveň měření ovlivňuje výpočet výsledků pro tuto proceduru, všechny proměnné musí mít definovanou úroveň měření.

**Data skenování.** Přečte data v aktivní datové sadě a přiřadí výchozí úroveň měření k jakýmkoli polím s momentálně neznámou úrovní měření. Je-li datová sada velká, může to nějakou dobu trvat.

**Přiřadit ručně.** Otevře dialogové okno se seznamem všech polí s neznámou úrovní měření. Toto dialogové okno můžete použít k přiřazení úrovně měření k těmto polím. Úroveň měření můžete také přiřadit v pohledu Proměnné v editoru dat.

Vzhledem k tomu, že úroveň měření je pro tuto proceduru důležitá, nemůžete přistupovat k dialogovému oknu pro spuštění této procedury, dokud nebude mít všechna pole definovanou úroveň měření.

## **Oblasti**

**Datová sada oblastí.** Tato skupina uvádí metodu rozdělení aktivní datové sady na školení, testování a pouzdro vzorků. **Školící vzorek** se skládá z datových záznamů používaných pro výcvik neuronové sítě; některé procento případů v datové sadě musí být přiděleno vzorku výcviku, aby bylo možné získat model. **Testovací ukázka** je nezávislá sada datových záznamů používaných ke sledování chyb během školení, aby se zabránilo nadměrnému školení. Doporučuje se, abyste vytvořili testovací vzorek, a školení v síti bude obecně nejefektivnější, pokud je testovací vzorek menší než vzorek odborné přípravy. **Výpočetní ukázka** je další nezávislá sada datových záznamů používaných k odhadu konečné neuronové sítě; chyba v souboru Holdout poskytuje "čestný" odhad prediktivní schopnosti modelu, protože pouzdro pouzdro nebylo použito k sestavení modelu.



- **Náhodně přiřazujte případy na základě relativního počtu případů.** Uvedte relativní číslo (poměr) případů náhodně vybraných do každého vzorku (trénink, testování a pouzdro). Sloupec % uvádí procentní část případů, které budou přiřazeny ke každému vzorku na základě zadaných relativních čísel. Například, uvedte 7, 3, 0 jako relativní čísla pro výcvik, testování a výtěrové vzorky odpovídají 70%, 30% a 0%. Zadání 2, 1, 1 jako relativní čísla odpovídá 50%, 25% a 25%; 1, 1, 1 odpovídá rozdělení datové sady na stejné třetiny mezi tréninkem, testováním a holdout.
- **Použití rozdělovací proměnnou k přiřazení případů.** Uvedte numerickou proměnnou, která přiřadí každý případ v aktivní datové sadě k výcviku, testování nebo vzorkování dlouh. Případy s kladnou hodnotou proměnné jsou přiřazeny ke vzorku školení, případy s hodnotou 0, do vzorku zkoušek a případy se zápornou hodnotou do vzorku odchvovacího vzorku. Případy se systémem-chybějící hodnota jsou z analýzy vyloučeny. Jakýkoli uživatel-chybějící hodnoty proměnné logické oblasti je vždy považován za platný.

## Architektura

Karta Architecture se používá k určení struktury sítě. Postup vytváří neurální síť s jednou skrytou "radiální základní funkcí" vrstvy; obecně to nebude nutné pro změnu těchto nastavení.

**Počet jednotek ve skryté vrstvě.** Existují tři způsoby, jak vybrat počet skrytých jednotek.

1. **Najděte nejlepší počet jednotek v rámci automaticky vypočteného rozsahu.** Procedura automaticky vypočítá minimální a maximální hodnoty rozsahu a vyhledá nejlepší počet skrytých jednotek v rozsahu. Je-li definován testovací vzorek, pak procedura používá kritérium testovacích dat: Nejlepším počtem skrytých jednotek je ten, který dává nejmenší chybu v testovacích datech. Není-li zkušební vzorek definován, použijte tento postup Bayesovo informační kritérium (BIC): Nejlepší počet skrytých jednotek je ten, který nese nejmenší kód BIC na základě údajů o odborné přípravě.
2. **Najděte nejlepší počet jednotek v uvedeném rozsahu.** Můžete zadat svůj vlastní rozsah a procedura najde "nejlepší" počet skrytých jednotek v daném rozsahu. Stejně jako před tím je nejlepší počet skrytých jednotek z rozsahu určen pomocí kritéria testovacích dat nebo BIC.
3. **Použít uvedený počet jednotek.** Můžete potlačit použití rozsahu a zadat konkrétní počet jednotek přímo.

**Aktivační funkce pro skrytou vrstvu.** Funkce aktivace pro skrytou vrstvu je radiální základní funkcí, která "spojuje" jednotky ve vrstvě k hodnotám jednotek v následné vrstvě. Pro výstupní vrstvu je aktivační funkce funkce identity; výstupní jednotky jsou jednoduše váženými součty skrytých jednotek.

- **Normalizovaná radiální základní funkce.** Používá aktivační funkci Softmax, takže aktivace všech skrytých jednotek je normalizována tak, aby byla čtveta 1.
- **Normální radiální základní funkce.** Používá exponenciální aktivační funkci, takže aktivace skryté jednotky je Gaussovým "busem" jako funkce vstupů.

**Přesah mezi skryté jednotky.** Překrývající se faktor je násobitel použitý na šířku radiálních základních funkcí. Automaticky vypočtená hodnota překrývajícího se faktoru je  $1+0.1d$ , kde  $d$  je počet vstupních jednotek (součet počtu kategorií přes všechny faktory a počet kovariany).

## Výstup

**Struktura sítě.** Zobrazí souhrnné informace o neuronové síti.

- **Popis.** Zobrazí informace o neuronové síti, včetně závislých proměnných, počtu vstupních a výstupních jednotek, počtu skrytých vrstev a jednotek a aktivačních funkcí.
- **Diagram.** Zobrazí diagram sítě jako neupravitelný graf. Všimněte si, že jak se zvyšuje počet proměnných a úrovní faktorů, bude se diagram obtížněji interpretovat.
- **Sypité váhy.** Zobrazuje odhady koeficientu, které ukazují vztah mezi jednotkami v dané vrstvě a jednotkami v následující vrstvě. Synchronizace vah je založena na vzorku školení, i když je aktivní datová sada rozdělena na oblasti vzdělávání, testování a výzdro. Všimněte si, že počet synaptických vah může být poměrně velký a tyto váhy se obvykle nepoužívají pro interpretaci výsledků sítě.

**Výkon sítě.** Zobrazí výsledky použité k určení, zda je model "dobrý." *Poznámka:* Grafy v této skupině jsou založené na kombinovaných vzorcích školení a testování nebo pouze v ukázce školení, pokud není k dispozici žádný testovací vzorek.

- **Souhrn modelu.** Zobrazí souhrn výsledků neuronové sítě podle oblasti a celkové, včetně chyby, relativní chyby nebo procentní části chybných předpovědí a doby školení.

Chyba je chyba součtu čtverců. Kromě toho se zobrazí relativní chyby nebo procenta chybných předpovědí, v závislosti na úrovních měření závislých proměnných. Má-li nějaká závislá proměnná úroveň měření, zobrazí se průměrná celková relativní chyba (relativní vzhledem k průměrnému modelu). Jsou-li všechny závislé proměnné kategorické, zobrazí se průměrný procentní podíl chybných předpovědí. Relativní chyby nebo procenta chybných předpovědí se také zobrazí pro jednotlivé závislé proměnné.

- **Výsledky klasifikace.** Zobrazí klasifikační tabulku pro každou kategorickou závislou proměnnou. Každá tabulka uvádí počet případů, které jsou správně a nesprávně klasifikovány pro každou kategorii závislých proměnných. Procento z celkového počtu případů, které byly správně klasifikovány, je také hlášeno.
- **Křivka ROC.** Zobrazí křivku ROC (Receiver Operating Characteristic) pro každou kategorickou závislou proměnnou. Také zobrazuje tabulku poskytující plochu pod každou křivkou. Pro danou závislou proměnnou zobrazí graf ROC jednu křivku pro každou kategorii. Má-li závislá proměnná dvě kategorie, pak každá křivka zachází s danou kategorií jako s pozitivním stavem a s druhou kategorií. Má-li závislá proměnná více než dvě kategorie, pak každá křivka zachází s danou kategorií jako s pozitivním stavem a s agregací všech ostatních kategorií.
- **Graf kumulativních zisků.** Zobrazí kumulativní přírůstkový graf pro každou kategorickou závislou proměnnou. Zobrazení jedné křivky pro každou kategorii závislých proměnných je stejné jako u křivek ROC.
- **Zvedací graf.** Zobrazí zvedací graf pro každou kategorickou závislou proměnnou. Zobrazení jedné křivky pro každou kategorii závislých proměnných je stejné jako u křivek ROC.
- **Predicted podle pozorovaného grafu.** Zobrazí graf předpovídané hodnoty podle pozorovaného grafu pro každou závislou proměnnou. Pro kategorické závislé proměnné jsou pro každou kategorii odezvy zobrazeny klastrované krabičkové grafy predikované pseudo-pravděpodobnosti, přičemž kategorie odezvy je pozorovaná jako proměnná klastru. U proměnných závislých na měřítku se zobrazí bodový graf.
- **Zbývající podle predikovaného grafu.** Zobrazí graf se zbytkovou hodnotou podle předpovídaných hodnot pro každou proměnnou závislou na měřítku. Mezi reziduálními hodnotami a předpovězenými hodnotami by neměly být žádné viditelné vzory. Tento graf se vytváří pouze pro měřítko závislých proměnných.

**Souhrn zpracování případu.** Zobrazí souhrnnou tabulku zpracování případu, která shrnuje počet zahrnutých případů a vyloučených v analýze, celkem a podle školení, testování a odebrání vzorků.

**Analýza důležitosti nezávislé proměnné.** Provede analýzu citlivosti, která vypočítá důležitost každého prediktoru při určování neuronové sítě. Analýza je založena na kombinované cvičné odborné přípravě a zkušebních vzorcích nebo pouze odběru vzorků, pokud neexistuje žádný zkušební vzorek. Tím se vytvoří tabulka a graf zobrazující důležitost a normalizovaný význam pro každý prediktor. Mějte na zřeteli, že analýza citlivosti je výpočetně nákladná a časově náročná, pokud existuje velký počet prediktorů nebo případů.

## Uložit

Karta Uložit se používá k ukládání předpovědi jako proměnných v datové sadě.

- **Ukládat předpovězenou hodnotu nebo kategorii pro každou závislou proměnnou.** Tím uložíte předpovězenou hodnotu pro proměnné závislé na měřítku a předpovídanou kategorii pro kategoriální závislé proměnné.
- **Ukládat předpovídanou pseudo-pravděpodobnost pro každou závislou proměnnou.** Tento příkaz uloží predikované pseudopravděpodobnosti pro kategoriální závislé proměnné. Samostatná proměnná se uloží pro každou z prvních  $n$  kategorií, kde  $n$  je uveden ve sloupci *Kategorie určené k uložení*.

**Názvy uložených proměnných.** Automatické generování názvu zajišťuje, že zachováte veškerou svou práci. Vlastní názvy vám umožňují vyřadit nebo nahradit výsledky z předchozích spuštění bez předchozího odstranění uložených proměnných v Editoru dat.

Pravděpodobnosti a pseudopravděpodobnosti

Předpovídat pseudo-pravděpodobnosti nelze interpretovat jako pravděpodobnosti, protože procedura Radial Basis Function používá pro výstupní vrstvu funkci součtu chyb a identity identity. Procedura ukládá tyto predikované pseudo-pravděpodobnosti i v případě, že jakákoli hodnota je menší než 0 nebo větší než 1, nebo součet pro danou závislou proměnnou není 1.

ROC, kumulativní zisky a zvedací grafy (viz "[Výstup](#)" na stránce 13) jsou založeny na pseudo-pravděpodobnosti. V případě, že kterýkoli z pseudopravděpodobnosti je menší než 0 nebo větší než 1 nebo součet pro danou proměnnou není 1, je první znovu nastaven tak, aby byl mezi 0 a 1 a aby byl součet 1. Pseudo-pravděpodobnosti se rescaled tím, že se dělí podle jejich součtu. Například, pokud případ předpovídal pseudo-pravděpodobnosti 0,50, 0,60 a 0,40 pro tří-kategorii závislou proměnnou, pak každé pseudo-pravděpodobnost je vydělí součtem 1.50, aby získal 0,33, 0,40, a 0,27.

Je-li některý z pseudopravděpodobnosti negativní, pak se absolutní hodnota nejnižší hodnoty přidá ke všem pseudopravděpodobnostem před výše uvedeným rescaling. Je-li například pseudo-pravděpodobnosti 0,30, .50 a 1,30, pak nejprve přidejte 0,30 do každé hodnoty pro získání 0,00, 0,80 a 1,60. Dále rozdělte každou novou hodnotu součtem 2,40 pro získání 0,00, 0,33 a 0,67.

## Export

Karta Export se používá k ukládání odhadů synaptické váhy pro každou závislou proměnnou na soubor XML (PMML). Tento modelový soubor můžete použít k použití informací modelu na jiné datové soubory pro účely hodnocení. Tato volba není k dispozici, pokud byly definovány rozdělené soubory.

## Volby

**Uživatel-Chybějící hodnoty.** Faktory musí mít platné hodnoty pro případ, který má být zahrnut do analýzy. Tyto ovládací prvky umožňují rozhodnout o tom, zda uživatelské hodnoty jsou považovány za platné mezi faktory a kategoriálními závislými proměnnými.



## Upozornění

---

Tyto informace byly vytvořeny pro produkty a služby poskytované v USA. Tento materiál může být dostupný od IBM v jiných jazycích. K povolení přístupu však může být vyžadováno vlastnictví kopie produktu nebo verze produktu v tomto jazyce.

Společnost IBM nemusí nabízet produkty, služby nebo funkce uvedené v tomto dokumentu v jiných zemích. Informace o produktech a službách, které jsou aktuálně k dispozici ve vaší oblasti, získáte od lokálního zástupce společnosti IBM. Odkazy na produkty, programy nebo služby společnosti IBM neuvádí ani neimplikují, že lze použít pouze daný produkt, program nebo službu společnosti IBM. Lze použít libovolný funkčně ekvivalentní produkt, program nebo službu neporušující práva duševního vlastnictví společnosti IBM. Vyhodnocení a ověření funkčnosti produktů, programů nebo služeb, které nepatří společnosti IBM, je však zodpovědností uživatele.

Společnost IBM může vlastnit patenty nebo nevyřízené žádosti o patenty zahrnující předměty popsané v tomto dokumentu. Vlastnictví tohoto dokumentu neposkytuje licenci k těmto patentům. Dotazy na licence můžete písemně odeslat na následující adresu:

*IBM Director of Licensing  
IBM Corporation  
North Castle Drive, MD-NC119  
Armonk, NY 10504-1785  
USA*

S dotazy na licence týkající se dvoubajtových informací (DBCS) se obraťte na oddělení intelektuálního vlastnictví společnosti IBM v dané zemi, nebo je odešlete písemně na následující adresu:

*Intellectual Property Licensing  
Legal and Intellectual Property Law  
IBM Japan Ltd.  
19-21, Nihonbashi-Hakozakicho, Chuo-ku  
Tokyo 103-8510, Japan*

SPOLEČNOST INTERNATIONAL BUSINESS MACHINES CORPORATION POSKYTUJE TUTO PUBLIKACI "TAKOVOU, JAKÁ JE", BEZ JAKÝCHKOLIV ZÁRUK, VYJÁDŘENÝCH NEBO ODVOZENÝCH VČETNĚ, MIMO JINÉ, ODVOZENÝCH ZÁRUK NEPORUŠENÍ PRÁV TŘETÍCH STRAN, ZÁRUKY PRODEJNOSTI NEBO VHODNOSTI PRO URČITÝ ÚČEL. Některé právní řády u určitých transakcí nepřipouštějí vyloučení záruk výslovně vyjádřených nebo vyplývajících z okolností, a proto se na vás výše uvedené omezení nemusí vztahovat, a proto se vás toto prohlášení nemusí týkat.

Uvedené údaje mohou obsahovat technické nepřesnosti nebo typografické chyby. Údaje zde uvedené jsou pravidelně upravovány a tyto změny budou zahrnuty v nových vydáních této publikace. Společnost IBM může kdykoli bez upozornění provádět vylepšení nebo změny v produktech či programech popsaných v této publikaci.

Jakékoliv odkazy v této publikaci na webové stránky jiných společností než IBM jsou poskytovány pouze pro pohodlí uživatele a nemohou být žádným způsobem vykládány jako doporučení těchto webových stránek. Materiály uvedené na těchto webových stránkách nejsou součástí materiálů pro tento produkt IBM a použití uvedených stránek je pouze na vlastní nebezpečí.

IBM může použít nebo distribuovat jakékoli informace, které jí poskytnete, libovolným způsobem, který společnost považuje za odpovídající, bez vzniku jakýchkoliv závazků vůči vám.

Vlastníci licence k tomuto programu, kteří chtějí získat informace o možnostech (i) výměny informací s nezávisle vytvořenými programy a jinými programy (včetně tohoto) a (ii) oboustranného využití vyměňovaných informací, mohou kontaktovat informační středisko na adrese:

*IBM Director of Licensing  
IBM Corporation*

North Castle Drive, MD-NC119  
Armonk, NY 10504-1785  
USA

Poskytnutí takových informací může být podmíněno dodržením určitých podmínek a požadavků zahrnujících v některých případech uhrazení stanoveného poplatku.

Licencovaný program popsáný v tomto dokumentu a veškerý licencovaný materiál k němu dostupný jsou společností IBM poskytovány na základě podmínek uvedených ve smlouvách IBM Customer Agreement, IBM International Program License Agreement nebo v jiné ekvivalentní smlouvě.

Citovaná data o výkonu a příklady klienta jsou uvedeny pouze pro názornost. Skutečné výsledky výkonu se mohou lišit v závislosti na specifických konfiguracích a provozních podmínkách.

Informace týkající se produktů jiných společností než IBM byly získány od dodavatelů těchto produktů, z jejich publikovaných sdělení, nebo z jiných veřejně dostupných zdrojů. IBM tyto produkty netestovala a nemůže potvrdit přesnost údajů o výkonu, kompatibilitě nebo jiná tvrzení týkající se produktů jiných společností než IBM. Otázky týkající se možností produktů jiných společností než IBM by měly být adresovány dodavatelům těchto produktů.

Prohlášení týkající se budoucího směru vývoje nebo záměrů společnosti IBM se mohou změnit nebo mohou být zrušena bez předchozího upozornění a představují pouze cíle a záměry.

Tyto údaje obsahují příklady dat a sestav používaných v běžných obchodních operacích. Aby byla představa úplná, používají se v příkladech jména osob, společností, značek a produktů. Všechna tato jména jsou fiktivní a jakákoliv podobnost se skutečnými lidmi nebo obchodními podniky je čistě náhodná.

#### COPYRIGHT - LICENCE:

Tyto informace obsahují ukázkové aplikační programy ve zdrojovém jazyku a ilustrují různé programovací techniky na různých operačních platformách. Tyto ukázkové programy můžete bez závazků vůči společnosti IBM jakýmkoli způsobem kopírovat, měnit a distribuovat za účelem vývoje, používání, odbytu či distribuce aplikačních programů odpovídajících rozhraní API pro operační platformu, pro kterou byly ukázkové programy napsány. Tyto příklady nebyly důkladně testovány ve všech podmínkách. Společnost IBM proto nemůže zaručit spolehlivost, upotřebitelnost nebo funkčnost těchto programů. Ukázkové programy jsou poskytovány "JAK JSOU", bez záruky jakéhokoli druhu. IBM nenes odpovědnost za žádné škody vzniklé ve spojení s Vaším užíváním ukázkových programů.

Jakákoli kopie nebo část těchto ukázkových programů nebo jakékoli odvozené dílo musí obsahovat následující poznámku o autorských právech:

© Copyright IBM Corp. 2021. Části tohoto kódu jsou odvozeny ze vzorových programů společnosti IBM Corp. Vzorové programy.

© Copyright IBM Corp. 1989-2021. Všechna práva vyhrazena.

## Ochranné známky

---

IBM, logo IBM a [ibm.com](http://ibm.com) jsou ochranné známky nebo registrované ochranné známky společnosti International Business Machines Corp., registrované v mnoha jurisdikcích po celém světě. Ostatní názvy produktů a služeb mohou být ochrannými známkami společnosti IBM nebo jiných společností. Aktuální seznam ochranných známek společnosti IBM je k dispozici na webu na stránce "Copyright and trademark information" na adrese [www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml).

Adobe, logo Adobe, PostScript a logo PostScript jsou buď registrované ochranné známky, nebo ochranné známky společnosti Adobe Systems Incorporated ve Spojených státech anebo v dalších zemích.

Intel, logo Intel, Intel Inside, logo Intel Inside, Intel Centrino, logo Intel Centrino, Celeron, Intel Xeon, Intel SpeedStep, Itanium a Pentium jsou ochranné známky nebo registrované ochranné známky společnosti Intel Corporation nebo jejich dceřiných společností ve Spojených státech a případně v dalších jiných zemích.

Linux je registrovaná ochranná známka Linuse Torvaldse ve Spojených státech a případně v dalších jiných zemích.

Microsoft, Windows, Windows NT a logo Windows jsou ochranné známky společnosti Microsoft Corporation ve Spojených státech a případně v dalších jiných zemích.

UNIX je registrovaná ochranná známka společnosti The Open Group ve Spojených státech a případně v dalších jiných zemích.

Java a všechny ochranné známky a loga založené na jazyce Java jsou ochranné známky nebo registrované ochranné známky společnosti Oracle anebo příbuzných společností.





# Rejstřík

## A

- aktivační funkce
  - ve funkci Radial Basis [13](#)
  - ve vícevrstvé Perceptron [5](#)
- architektura
  - neuronové sítě [2](#)
- architektura sítě
  - ve funkci Radial Basis [13](#)
  - ve vícevrstvé Perceptron [5](#)

## D

- dávkové školení
  - ve vícevrstvé Perceptron [7](#)

## CH

- chybějící hodnoty
  - ve vícevrstvé Perceptron [10](#)

## K

- Křivka ROC
  - ve funkci Radial Basis [13](#)
  - ve vícevrstvé Perceptron [8](#)

## M

- Multivrstva Perceptron
  - architektura sítě [5](#)
  - export modelu [9](#)
  - Oblasti [5](#)
  - Trénování [7](#)
  - ukládání proměnných do aktivní datové sady [9](#)
  - volby [10](#)
  - výstup [8](#)

## N

- neuronové sítě
  - architektura [2](#)

## O

- online školení
  - ve vícevrstvé Perceptron [7](#)

## P

- Pravidla zastavení
  - ve vícevrstvé Perceptron [10](#)
- přírůstkový graf
  - ve funkci Radial Basis [13](#)
  - ve vícevrstvé Perceptron [8](#)

## R

- Radiální základní funkce
  - architektura sítě [13](#)
  - export modelu [15](#)
  - Oblasti [12](#)
  - ukládání proměnných do aktivní datové sady [14](#)
  - volby [15](#)
  - výstup [13](#)

## S

- Síťový diagram
  - ve funkci Radial Basis [13](#)
  - ve vícevrstvé Perceptron [8](#)
- skrytá vrstva
  - ve funkci Radial Basis [13](#)
  - ve vícevrstvé Perceptron [5](#)

## Š

- školení mini-batch
  - ve vícevrstvé Perceptron [7](#)
- školení v síti
  - ve vícevrstvé Perceptron [7](#)

## U

- ukázka školení
  - ve funkci Radial Basis [12](#)
  - ve vícevrstvé Perceptron [5](#)
- ukázka testování
  - ve funkci Radial Basis [12](#)
  - ve vícevrstvé Perceptron [5](#)

## V

- výstupní vrstva
  - ve funkci Radial Basis [13](#)
  - ve vícevrstvé Perceptron [5](#)
- výtahový graf
  - ve funkci Radial Basis [13](#)
  - ve vícevrstvé Perceptron [8](#)
- Vzorek pro odchvatu
  - ve funkci Radial Basis [12](#)
  - ve vícevrstvé Perceptron [5](#)





