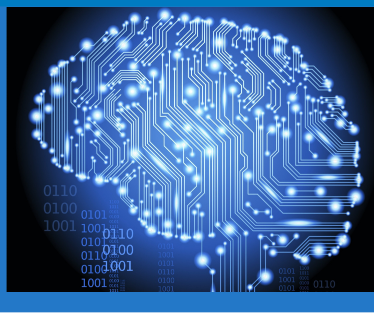
Oddělení strojového učení

Ústav informatiky AV ČR, v.v.i.



Analýza možností data-miningu pomocí univalentních neuronových sítí a zobrazení.

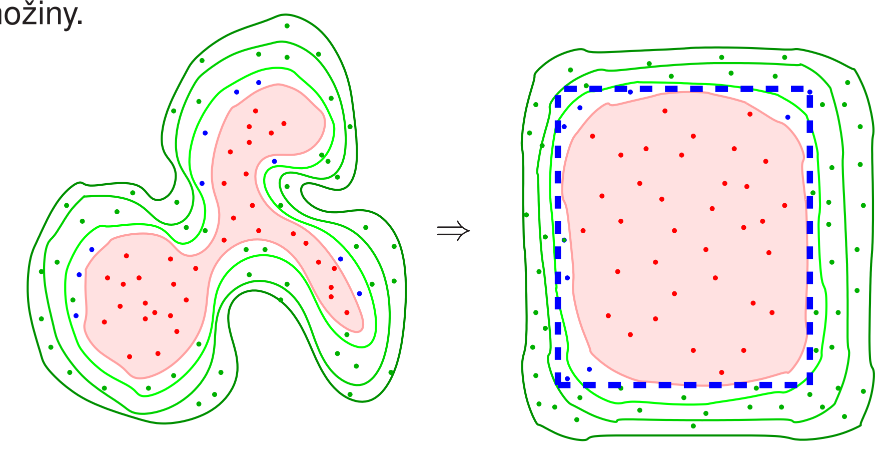
Ing. František Hakl, CSc.

Popis tématu

Univalentnost (vzájemná jednoznačnost) zobrazení mezi dvěma Eukleidovskými prostory, které je použito pro řešení úloh separace množin, umožňuje analyzovat separované (resp. aproximované) množiny z jejich geometrického hlediska, kdy je možno definovat "vnitřek množiny" v běžném slova smyslu a vybírat ze separované množiny vzory, které jsou "uvnitř". Toto je v případě vzájemně jednoznačného zobrazení umožněno skutečností, že každé dvě disjunktní množiny musí být zobrazeny opět na disjunktní množiny.

Základní princip využití univalentnosti zobrazení

Body v příhraničních oblastech zkoumané množiny (značeny zeleně) jsou zobrazovány do "příhraničních" oblastí univalentního obrazu této množiny. Body, které budeme považovat za "vnitřní" – vyznačeny červeně –, jsou zobrazovány do "vnitřní" oblasti obrazu množiny. Díky námi předem definovanému cílovému tvaru zobrazované množiny, lze definovat vnitřní oblast zobrazované množiny v obrazovém prostoru (vyznačena modře) a body zobrazené do této vnitřní oblasti pak považovat za vnitřní body původní množiny. Modře značené body jsou díky předpokládané spojitosti použitého univalentního zobrazení stále v blízkosti vnitřku množiny.



Je zjevné, že odebráním bodů původní množiny, které se zobrazí do vnitřní (modré) oblasti, zredukujeme velikost původní množiny při zachování informace o jejím geometrickém tvaru. Tohoto postupu lze využít pro redukci velikosti datových množin. Současně si můžeme definovat více do sebe vnořených "vnitřních" množin a ohodnocovat body původní datové množiny podle příslušnosti do těchto námi vytvořených "vnitřních" množin. Hodnotu tohoto ohodnocení pak můžeme chápat jako míru toho, nakolik je daný bod vnitřním bodem. To může mít praktický dopad v aplikacích, kde je stav systému spojitě závislý na parametrech a je v jistých oblastech na těchto parametrech závislý proporcionálně – např. čím vyšší hodnota parametru, tím větší přiblížení se systému k nežádoucímu stavu systému. Lze-li systém navíc externě řídit, je využití univalentnosti možné i v opačném směru, kdy můžeme udržovat parametry systému mimo modrou oblast, která garantuje existenci systému v akceptovatelné oblasti stavů.

Univalentní neuronová síť typu MLP

Jedním z užívaných nástrojů pro separaci množin jsou umělé neuronové sítě, které mohou být uzpůsobeny tak, aby realizovaly univalentní zobrazení. Univalentní neuronové sítě jsou odvozeny ze standardních modelů vícevrstvých neuronových sítí s tím, že jejich topologie a parametrický prostor je modifikován takovým způsobem, že je garantována vzájemná jednoznačnost mezi vstupním a výstupním prostorem univalentní sítě (zobrazení). Základní snahou při syntéze těchto sítí je navrhnout vhodou restrikci architektury a parametrů tak, aby bylo dosaženo univalentnosti zobrazení, při pokud možno nejširším stupni volnosti používaných parametrů. Dále je potřeba navrhnout vhodný algoritmus pro učení sítě, kdy nelze použít standardní gradientní iterační metody, které nereflektují zvolená omezení parametrického prostoru.

Věta o univalentnosti vrstevnatých neuronových sítí

Předpokládejme, že $\mathbf{A}_i, \mathbf{H}_i, \mathbf{B} \in R^{n\times n}$, $\mathbf{f}_i \in R^n$, $F : R^n \Rightarrow R^n$, $\frac{\partial F_i}{\mathbf{x}_i} > 0$, $\frac{\partial F_i}{\mathbf{x}_j} = 0$, $i, j \in \{1, \dots, L\}$. Pak matice $\sum_{i=1}^{L} \mathbf{A}_i\mathbf{D}_i\mathbf{H}_i + \mathbf{B}$ je ve třídě P pro libovolné kladné diagonální matice $\mathbf{D}_i$, $i \in \{1, \dots, L\}$ právě když rovnice

$$\sum_{i=1}^{L} \mathbf{A}_i F\left(\mathbf{H}_i\mathbf{x} - \mathbf{f}_i\right) + \mathbf{B}\mathbf{x} = \mathbf{c} \tag{1}$$

má pro libovolný vektor $\mathbf{c}$ právě jedno řešení.
Navíc toto řešení spojitě závisí na vektoru $\mathbf{c}$.

Praktická aplikovatelnost univalence

Praktické použití univalentních sítí se dá očekávat v oblastech, kde nutnost reakce na data v separovaných množinách je spojité povahy a data která jsou uvnitř množin jsou více kritická než data na povrchu (např. medicínské aplikce, úlohy vyhodnocování dat bez kategoriálních proměnných, systémy sledující technologické procesy spojitě závisející na vnějších parametrech, atd.).

Čím by jste se mohli zabývat

- Obecné studium univalentnosti zobrazení. Stanovení podmínek na parametrický prostor zobrazení, které zachovávají univalentnost.
- Odvození učících algoritmů pro univalentní separační nástroje, např. gradientní metody s omezujícími podmínkami či jiné metody globální optimalizace umožňující podchycení omezujících podmínek v definičním oboru.
- Aplikace univalentních separátorů na vybraná reálná data.

http://ml.cs.cas.cz hakl@cs.cas.cz