

Intelligence psů vzhledem k velikosti plemene

Využití shlukové analýzy

Psi jsou jedním z nejdéle domestikovaných druhů živočichů, a žijí ve společnosti lidí už tisíce let. Pro mnoho lidí jsou dnes právoplatnými členy rodiny. A podobně jako si většina z nás přeje mít chytré děti, chceme mít i chytré psy.

Jedním z nejznámějších vědců, který se zabýval oblastí intelligence psů, je emeritní profesor University of British Columbia, psycholog Stanley Coren.



V průběhu své kariéry publikoval stovky odborných článků v oblasti neuropsychologie (zabýval se např. senzoryckými procesy), celosvětový věhlas mu ale přinesla až jeho kniha *The Intelligence of Dogs* z roku 1994 (Frame, 2017). Coren ve spolupráci s rozhodčími soutěží v poslušnosti psů seřadil 133 plemen psů podle jejich pracovní intelligence – hodnotil jak pravděpodobnost splnění povelu, tak i schopnost psa učit se nové povely (Coren, 2009).

Data

Pro účely této analýzy byly použity dva datasety. První z nich pochází z již zmíněného výzkumu Stanleyho Corena. Obsahuje seznam plemen psů s dvěma proměnnými – **poslušnost** (vyjadřuje procentuální pravděpodobnost splnění povelu na první pokus) a **počet opakování povelu** (vyjadřuje, kolikrát bylo nutné zopakovat nový povel k jeho naučení). Druhý dataset obsahuje informace o jednotlivých plemenech psů (výška, váha), zdrojem údajů byl American Kennel Club. Spojením těchto dvou datasetů je možné zkoumat, zda existují nějaké souvislosti mezi velikostí plemene a jeho pracovní inteligencí.

Před samotnou analýzou jsme provedli úpravu dat. Byly vyřazena ta plemena, u kterých nebyly dispozici všechny údaje. V případě proměnných, kde byla data uvedena v rozmezí (výška od-do, váha od-do, min. počet opakování povelu-max. počet opakování povelu) byly hodnoty pro zjednodušení a názornost zprůměrovány. Výsledná datová matice obsahuje 112 pozorování (plemen psů).

Plemeno	průměrná výška v cm	průměrná váha v kg	poslušnost	počet opakování povelu
Affenpinscher	26,67	4,54	70%	20,50
Airedale Terrier	58,42	20,41	70%	20,50
Akita	68,58	45,36	50%	33,00
American Cocker Spaniel	39,37	11,34	85%	10,00
American Foxhound	59,69	30,62	50%	33,00
American Staffordshire Terrier	45,72	20,41	70%	20,50
American Water Spaniel	41,91	15,88	50%	33,00
Australian Cattle Dog	46,99	18,14	95%	2,50
Australian Shepherd	52,07	22,68	50%	33,00
Australian Silky Terrier	24,13	4,31	70%	20,50
Australian Terrier	25,40	5,44	70%	20,50
Bearded Collie	53,34	22,68	70%	20,50

Základní představu by nám na základě těchto dat mohly poskytnout už popisné statistiky. Co kdybychom ale chtěli plemena psů podle zjištěných hodnot rozdělit do určitých skupin? V tom nám pomůže **shluková analýza**.

Shluková analýza

Shluková analýza je jednou z vícerozměrných statistických metod, která umožňuje klasifikaci objektů na základě jejich podobnosti nebo příbuznosti – jejím cílem je tedy nalézt shluky objektů, které jsou si podobné, a zároveň nejsou příliš podobné objektům mimo tento shluk. Shlukovou analýzu můžeme dále dělit na **hierarchické** a **nehierarchické** metody.

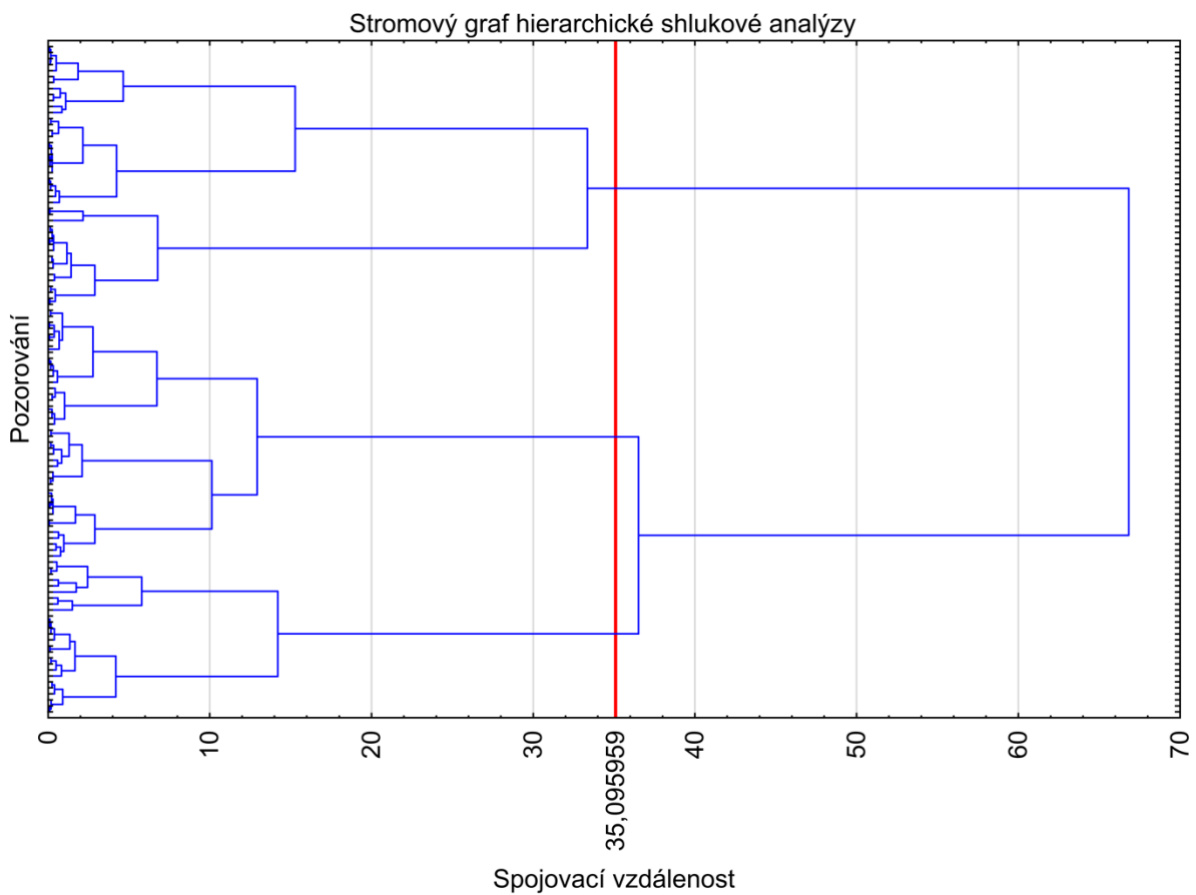
Hierarchická shluková analýza

Principem hierarchické shlukové analýzy je postupné vytváření hierarchické struktury shluků. Existují dvě hlavní strategie – při **aglomerativním přístupu** začínáme s jednotlivými objekty, které na základě podobnosti spojujeme, dokud nakonec nedojdeme k jednomu velkému shluku. Při **divizním přístupu** je to přesně naopak –

začínáme s jedním velkým shlukem, který na základě podobností dělíme na stále menší shluky, dokud nedosáhneme jednotlivých objektů.

Existuje více přístupů, jak objekty na základě podobnosti shlukovat. V našem případě jsme použili **Wardovu metodu**, která vychází analýzy rozptylu a zaměřuje se na minimalizaci variančního rozptylu uvnitř shluků. Tímto způsobem umožňuje vytvořit shluky, které jsou kompaktní a homogenní.

Výsledek hierarchické shlukové analýzy můžeme vizualizovat pomocí stromového grafu neboli **dendrogramu**.



Na dendrogramu z našich dat můžeme vidět, jak byly postupně tvořeny jednotlivé shluky. Na ose Y jsou vyneseny jednotlivá pozorování (objekty), které jsou na základě podobností zařazovány do stále větších shluků. Na ose X pak vidíme spojovací vzdálenosti mezi jednotlivými shluky. Rozhodnutí, do kolika shluků budeme jednotlivé objekty slučovat, vyžaduje určitou znalost dat, intuici, případně podporu v nějaké teorii. V našem případě

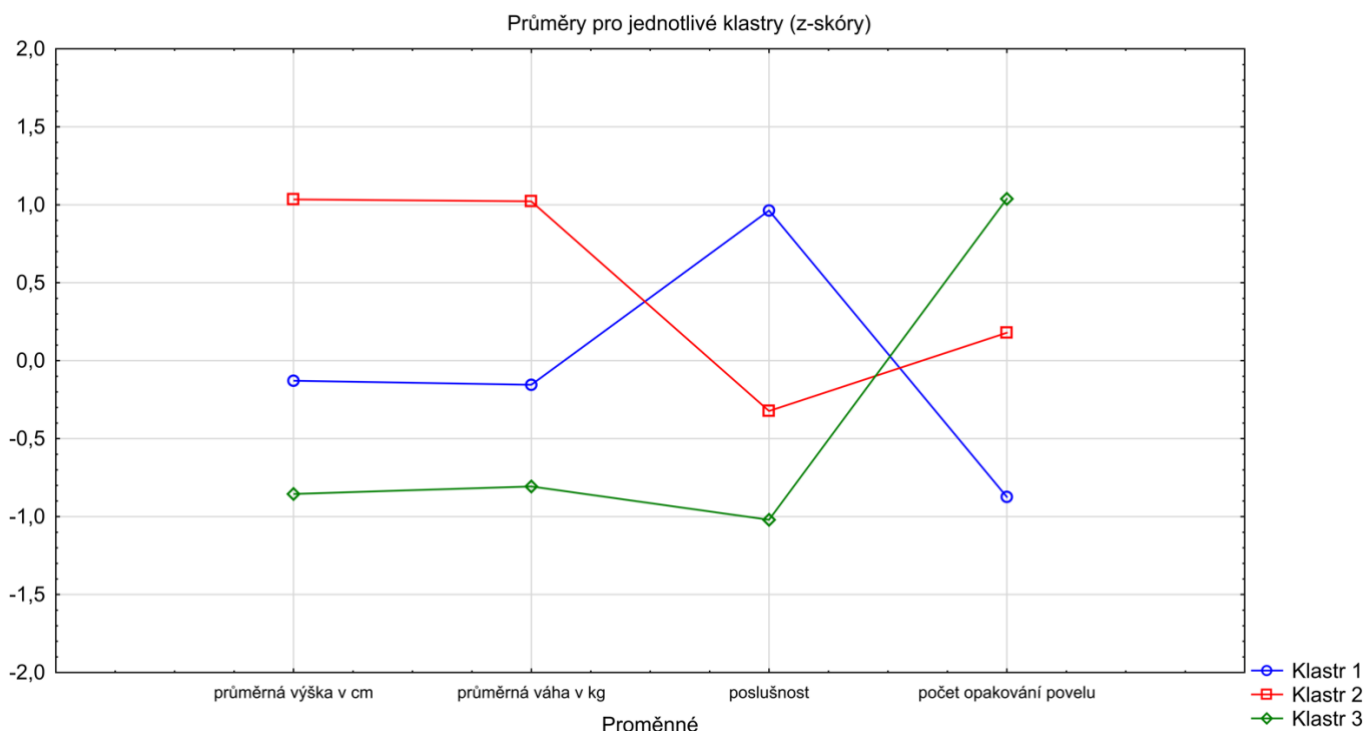
bychom mohli například dělit plemena psů na malá, střední, a velká – rozdělení do 3 shluků (klastřů) je na dendrogramu znázorněn červenou vertikální čarou.

Nehierarchická shluková analýza

Narozdíl od předchozí metody, při použití nehierarchické shlukové analýzy nedochází k vytvoření hierarchické struktury shluků.

Jednou z nejčastěji používaných metod nehierarchické shlukové analýzy je metoda k-means. Metoda k-means je algoritmus pro shlukování dat do předem určeného počtu shluků. Jednoduše řečeno, algoritmus se snaží najít střední hodnoty (**centroidy**) shluků tak, aby se minimalizovala vzdálenost mezi každým bodem datového souboru a středem příslušného shluku.

V našem případě jsme se rozhodli, že budeme objekty slučovat do 3 klastřů. Výsledek takové nehierarchické shlukové analýzy pak můžeme vizualizovat např. pomocí grafu:



Na ose Y vidíme vyneseny jednotlivé proměnné, na ose X hodnoty standardizované do z-skóru. Tento graf nám již přináší poměrně názorný přehled o tom, jak to s našimi 3 skupinami plemen psů dopadlo. Podle *průměrné výšky a váhy* vidíme, že v Klastru 1 jsou

zařazena střední plemena psů. Hodnoty proměnné *poslušnost* jsou u tohoto klastru vysoké, hodnoty *počtu opakování povelu* naopak nízké. Na základě našich dat bychom tedy mohli říct, že střední plemena psů budou dosahovat nejvyšší pracovní inteligence.

Klastr 2 obsahuje velká plemena. Hodnoty proměnných *poslušnost* i *počet opakování povelu* jsou u tohoto klastru na střední úrovni.

Klastr 3 pak obsahuje malá plemena psů. Vidíme, že hodnoty proměnné *poslušnost* jsou nízké, naopak *potřebný počet opakování povelu* k naučení je vysoký. Zde bychom tedy pravděpodobně usuzovali nižší pracovní inteligenci.

Zdroje

Coren, S. (2009). *Canine Intelligence – Breed Does Matter*. Psychology Today. <https://www.psychologytoday.com/intl/blog/canine-corner/200907/canine-intelligence-breed-does-matter>

Frame, S. (2017). *Adventures in dog research with Stanley Coren*. American Psychological Association. <https://www.apa.org/members/content/stanley-coren>