

**Název práce:** Dobývání znalostí z dat pomocí kopulí

**Autor:** Martin Ščavnický

**Katedra:** Katedra teoretické informatiky a matematické logiky

**Vedoucí diplomové práce:** RNDr. Ing. Martin Holeňa CSc., Katedra teoretické informatiky a matematické logiky

**Abstrakt:** Kopule jsou funkce, které popisují vztah mezi sdruženým rozdělením náhodného vektoru a jeho marginály. Kopule umožňují modelovat vícerozměrná rozdělení a jsou hojně využívány ve finančnictví a zkoumány v dobývání znalostí. V praxi existuje mnoho různých rodin kopulí, ale žádný standardní způsob, jak vybrat tu správnou. V naší práci se zabýváme porovnáním vhodnosti různých rodin kopulí k dobývání znalostí. Prokládáme klasifikační data pomocí osmi rodin kopulí a porovnáváme je pomocí tří metrik kvality proložení. Také používáme klasifikační algoritmus založený na kopulích a porovnáváme jeho přesnost pro různé rodiny kopulí. Výsledky naznačují, že eliptické kopule popisují naše data nejlépe, ale v přesnosti klasifikace jsou s nimi porovnatelné hierarchické Archimedovské kopule. V práci také navrhuje a testujeme modifikovanou metodu pro modelování dat pomocí hierarchických Archimedovských kopulí, která je schopna lépe popsat některá data s negativní závislostí mezi atributy. Na bázi této metody pak navrhuje vizualizaci závislostí v datech a pozorujeme, jestli zachycuje rozdíly mezi třídami.

**Klíčová slova:** data mining, vztahy mezi atributy, pravděpodobnostní vztahy, kopule, typy kopulí