

# Rozhodovací stromy pro Poker Hands

Martin Lukeš

TABLE I  
TAULKA ATRIBUTŮ

Atribut	Význam	Rozsah
C1	Barva první karty	1-4
S1	Hodnota první karty	1-13
C2	Barva druhé karty	1-4
S2	Hodnota druhé karty	1-13
C3	Barva třetí karty	1-4
S3	Hodnota třetí karty	1-13
C4	Barva čtvrté karty	1-4
S4	Hodnota čtvrté karty	1-13
C5	Barva páté karty	1-4
S5	Hodnota páté karty	1-13
Hand Power	Síla ruky	0-11

**Abstrakt— Rozpoznání síly ruky v pokeru spolu s pozorováním tellů (manýr, výraz, emoce či zvyk vyjadřující informaci o stavu ruky hráče) oddělují špatné hráče od dobrých. Tento dokument se zaměřil na zkoumání síly jednotlivých rukou.**

## I. ASSIGNMENT

Vyberte si data, která jsou použitelná pro klasifikaci nebo regresi. Data předzpracujte, a nainportujte do DM aplikace. Zvolte si klasifikátor (kNN, lineární a polynomiální separace, rozhodovací strom) nebo regresní algoritmus (kNN, lineární a polynomiální regrese, regresní strom). Najděte nastavení parametrů zvoleného algoritmu tak, aby produkoval co nejlepší modely.

## II. INTRODUCTION

Vytvořil jsem model pro rozhodovací stromy. Nicméně pro mnou zvolená data se rozhodovací stromy ukázaly jako nešťastné (moc rozvětvené). Ač byla data rozsáhlá, tak byla již v číselné podobě a bylo nutné je jen normalizovat. Atributy a jejich významy jsou zachyceny v tabulce 1. Poslední atribut Hand Power je tzv. učitel, který vyjadřuje sílu dané ruky. Položek v datech je přes jeden milion, proto jsem použil náhodný vzorek asi 3000 instancí, který pro mé účely stačil.

## III. METHODOLOGY

Rozhodovací stromy obsahují v každém listu jednu klasifikační třídu. Každý nelistový uzel stromu zastupuje jednu podmínku na nějaký příznak. Atribut, podle kterého tvoříme rozhodovací pravidlo je vybrán podle hodnoty entropie. Na Figure 1 vidíme vývoj chyby v závislosti na hodnotě parametru splitmin. Z hlediska nejnižší chyby je nejvýhodnější hodnota splitmin = 60. Pokus jsem opakovat, abych si ověřil domněnku o parametru splitmin. Graf druhého pokusu je na Figure 3 a nevyvrací první měření .

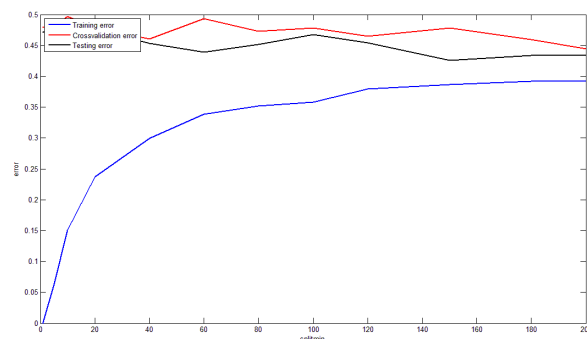


Fig. 1. Chyba v závislosti na ohebnosti modelu

vidíme, že nejlepší je vytvořit trénovací množinu z asi 75% dat.

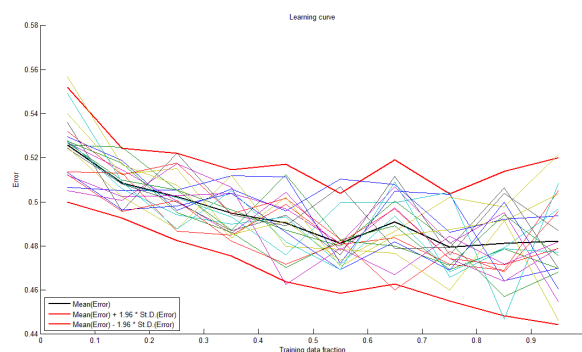


Fig. 2. Učící křivka.

## IV. CONCLUSION

Zjistil, že je nejvýhodnější vybrat do trénovací množiny 75 % dat. Parametr splitmin je nejvýhodnější pro hodnotu 60.

## REFERENCES

- [1] Materiály z přednášek.

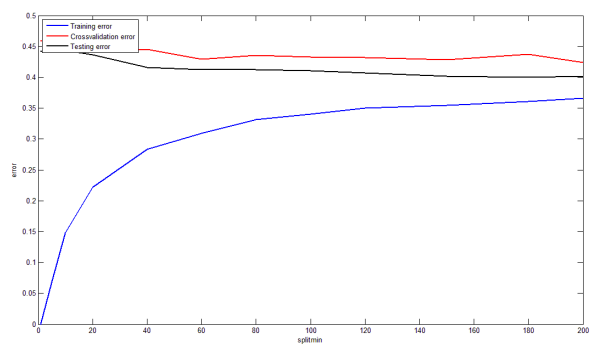


Fig. 3. Druhý pokus chyby