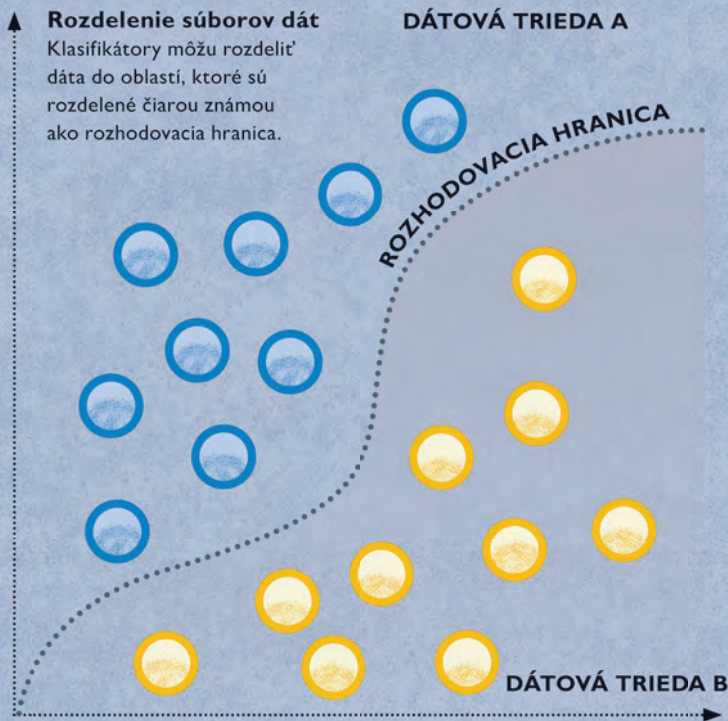
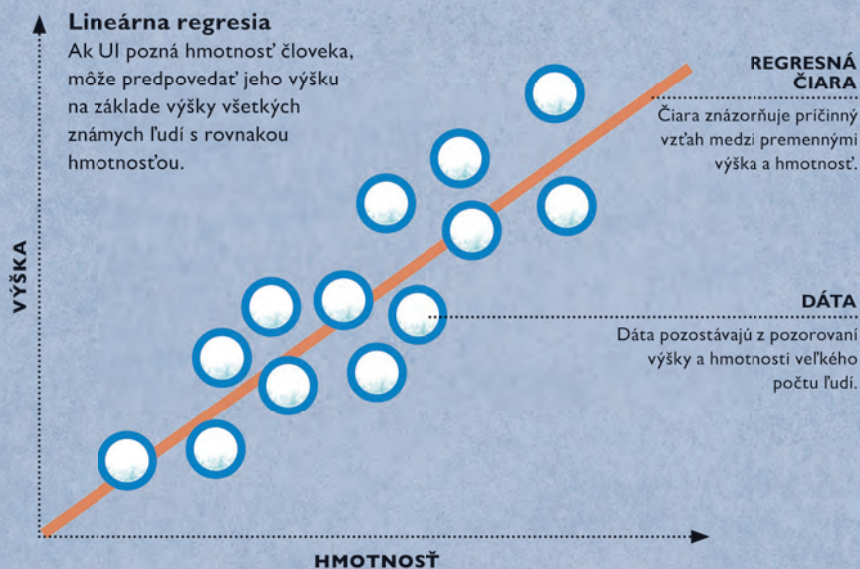


TYPY DÁT

Algoritmus, ktorý priraduje prvkom označenia (pozri s. 62 – 63) a potom ich triedi do kategórií alebo „tried“, sa nazýva „klasifikátor“.

Prostredníctvom procesu učenia s učiteľom (pozri s. 72) sa UI učí klasifikovať položky pomocou označeného súboru tréningových dát (pozri s. 61), z ktorého sa učí rozpoznávať vzory spojené s rôznymi označeniami. Napríklad spamový filter sa pomocou tréningových dát učí rozpoznávať znaky žiadaných a nevyžiadanych e-mailov zo súboru označených e-mailov. Na základe nich môže UI následne automaticky priradiť novým e-mailom označenie „spam“ alebo „nie je spam“.





PRIAMKA NAJLEPŠEJ ZHODY

Regresia je proces, ktorý sa používa v mnohých oblastiach vrátane strojového učenia (pozri s. 58 – 59), kde algoritmus predpovedá správanie jednej alebo viacerých premenných v závislosti od hodnoty inej premennej. Využíva sa v mnohých aplikáciách učenia s učiteľom (pozri s. 72), najmä v tých, ktorých cieľom je nájsť kauzálne vzťahy medzi viacerými premennými. Napríklad sa môže použiť na predpovedanie teploty nasledujúceho dňa vzhľadom na dnešnú vlhkosť, rýchlosť vetra a atmosférický tlak a dáta o tom, ako sa všetky štyri premenné správali v minulosti. „Lineárna regresia“ (pozri vyššie) je najbežnejšou formou regresnej analýzy a používa sa najmä v oblasti financií a ekonomiky.