

kapitola 12

Porazil bys hloupého soupeře?

Předpověď, které se smáli

1861: Robert FitzRoy vydal v *The Times* první veřejnou předpověď počasí na světě. Razil pro ni i slovo — **forecast**.

Tisk se mu smál za každý netrefený den, věda jím pohrdala: *bádání není věda*. Zemřel v posměchu a ruině.

Zbyla otázka: byla jeho předpověď k něčemu — nebo jen opisovala včerejšek?

Nejdřív najdi hloupého soupeře

- „Model má 75 % úspěšnost!“ V Praze neprší tři dny ze čtyř — tupé „zítra nebude pršet“ má taky 75 %.
- Hloupý soupeř je laťka na zemi: průměr, poslední hodnota, nejčastější třída. Kdo ho neporazí, jen dráž dělá totéž.
- FitzRoyův soupeř má jméno: persistence — „zítra bude jako dnes“.

Prevalence je krutější než přesnost

Test na vzácnou nemoc, 90 % přesný v obou směrech. Nemoc má jeden ze sta. Přejde ti pozitivní výsledek — jsi nemocný?

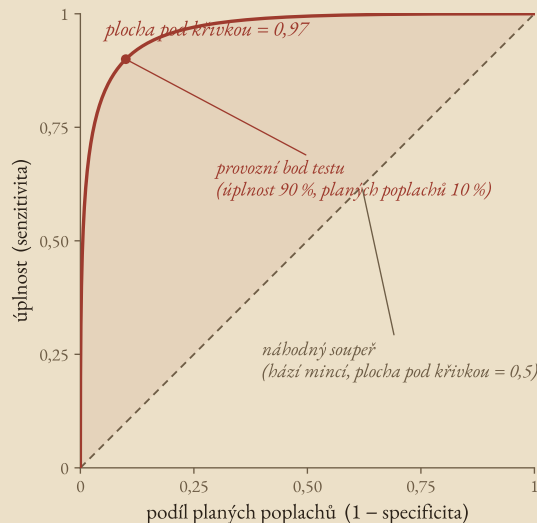
Z tisíce lidí: 9 skutečných nálezů, ale 99 planých poplachů ze zdravé většiny. Pozitivních 108, nemocných 9.

Šance, že jsi nemocný, je osm procent — ne devadesát.

Tři čísla, která se pletou

- **Správnost** — kolik ze všech odpovědí je dobře. Zdravá většina ji drží vysoko.
- **Přesnost** — z označených kolik doopravdy platí. Spadla na 8 %.
- **Úplnost** — z nemocných kolik jsi našel. A cena chyby se úlohu od úlohy obrací: u spamu bolí planý poplach, u rakoviny prošvihnutý nádor.

ROC: kdo hází mincí, leží na úhlopříčce



Náhodný soupeř má AUC 0,5. Ale i vysoká AUC u vzácné nemoci klame.

„Jak bych poznal, že se pletu?„

Než uvěříš modelu nebo AI demu, polož dvě otázky:

Co je hloupý soupeř — a porazil ho vůbec? A: jak vzácné je to, co hledá?